ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL OF SCIENCE ENGINEERING AND TECHNOLOGY

INVESTIGATION OF MOLECULAR PATHWAYS AND BIOMARKERS IN MULTIPLE SCLEROSIS CLINICAL SUBTYPES

Ph.D. THESIS

Timuçin AVŞAR

Molecular Biology – Genetics and Biotechnology Department

Molecular Biology – Genetics and Biotechnology Programme

JUNE 2015

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL OF SCIENCE ENGINEERING AND TECHNOLOGY

INVESTIGATION OF MOLECULAR PATHWAYS AND BIOMARKERS IN MULTIPLE SCLEROSIS CLINICAL SUBTYPES

Ph. D. THESIS

Timuçin AVŞAR
(521092073)

Molecular Biology – Genetics and Biotechnology Department

Molecular Biology – Genetics and Biotechnology Programme

Thesis Advisor: Assoc. Prof. Dr. Eda TAHİR TURANLI

JUNE 2015

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ FEN BİLİMLERİ ENSTİTÜSÜ

MULTİPL SKLEROZ KLİNİK ALT TİPLERİNDE MOLEKÜLER YOLAKLARIN VE BİYOBELİRTEÇLERİN ARAŞTIRILMASI

DOKTORA TEZİ

Timuçin AVŞAR
(521092073)

Moleküler Biyoloji, Genetik ve Biyoteknoloji Anabilim Dalı

Moleküler Biyoloji – Genetik ve Biyoteknoloji Programı

Tez Danışmanı: Doç. Dr. Eda TAHİR TURANLI

HAZİRAN 2015

Timuçin Avşar, **a Ph.D.** student of ITU **Graduate School of Science Engineering and Technology** student ID **521092073,** successfully defended the **thesis** entitled "**INVESTIGATION OF MOLECULAR PATHWAYS AND BIOMARKERS IN MULTIPLE SCLEROSIS CLINICAL SUBTYPES**", which he prepared after fulfilling the requirements specified in the associated legislations, before the jury whose signatures are below.

Thesis Advisor : Assoc. Prof. Dr. Eda TAHİR TURANLI .....
İstanbul Technical University

Jury Members : Prof. Dr. Aksel SİVA .....
İstanbul University

Jury Members : Assoc. Prof. Dr. Gizem DİNLER DOĞANAY .....
İstanbul Technical University

Jury Members : Prof. Dr. Ayten Yazgan KARATAŞ .....
İstanbul Technical University

Jury Members : Assist. Prof. Dr. Can ERZİK .....
Marmara University

Date of Submission : 19 June 2015
Date of Defense : 30 June 2015

To my wife and son,

FOREWORD

First, I would like to express my deep appreciation and thanks for my advisor Assoc. Prof. Dr. Eda TAHİR TURANLI for the inspiring guidance, support and the opportunities she provided me. I was especially fortunate regarding her unique way of approach to scientific phenomena and her enthusiasm, which makes possible to ask the right questions with courage and devotion. It was a great pleasure for me to carry on this thesis under her supervision. I want to thank her for encouraging me to establish this thesis, which could not have been accomplished without her.

I would also like to thank to Prof. Dr. Aksel SİVA for implementing and supporting this project; Prof. Dr. Ayten Yazgan KARATAŞ, Assoc. Prof. Dr. Gizem DİNLER DOĞANAY and Assist. Prof. Dr. Can ERZİK for evaluating my thesis and being a member of thesis jury. A special thanks to Prof. Dr. O. Uğur SEZERMAN for helping and guiding us in bioinformatics studies and to Prof. Dr. Arzu KARABAY KORKMAZ for following my thesis during all progresses and critical evaluations in all steps.

I would like to thank to Prof. Dr. Sabahattin SAİP, Dr. Uğur UYGUNOĞLU, Dr. Melih TÜTÜNCÜ and Onat DEMİRCİ for his meticulous work in clinical part of my study, obtaining patients' data and the collaborating nurses and assistants from Neurology Department, Nuray OLPAK and Mrs. NURSEL for their kind help in collecting the samples. Another special thanks for İ. Melis DURASI for her precious help and support in bioinformatics analysis. It was a great pleasure and luck for me to meet and study with all them.

I would like to thank my lab partners, Aslı KİREÇTEPE AYDIN, Gökçe ÇELİKYAPI ERDEM, Elif EVEREST, Ayşe ERÖZENCİ, İlker KARACAN, Merve ÖZKILINÇ, Pelinsu KÖPRÜLÜ, Begüm FİDAN, Ayşe BALAMİR and Neslihan SEVİNÇ for their collaboration and help in setting up our lab from ground. It was an invaluable experience for me to work with them.

I have been extraordinarily fortunate to have İrem AVŞAR in my life. She is the source of my happiness, that I am not in need to pursuit anymore and moreover the mother of my son Yusuf Mert. And I thank God for enabling me to achieve all that I have achieved, and I thank to my parents Ayşe and Mehmet Ali AVŞAR and my brother Hüseyin AVŞAR and sister Emine AVŞAR ŞİŞMAN for their endless love, patience and having every confidence in my ability to succeed. They all have a pure effect in everything I accomplish.

I would like to thank to The Scientific and Technological Research Council of Turkey (TUBITAK) (Project No: 109S070), Scientific Research Projects of Istanbul Technical University and Bayer & Schering Ploug Company (Turkey) for the financial support they provided for this study.

Last but the most important thanks are for multiple sclerosis patients and their families. The support they provided for us is not measurable with words and compliments. They did not only shared their blood, cerebrospinal fluid, and personal information etc. with us, they also shared their hopes, expectations and joy of livings. I hope we will fulfill their expectations in near future...

June 2015 Timuçin AVŞAR
(Ph.D.)

TABLE OF CONTENTS

FOREWORD ..... **ix**
ABBREVIATIONS ..... **xv**
LIST OF TABLES ..... **xix**
LIST OF FIGURES ..... **xxi**
SUMMARY ..... **xxiii**
ÖZET ..... **xxv**
1. INTRODUCTION **1**
1.1 Purpose Of Thesis 1
1.2 Literature Review 1
1.2.1 Multiple sclerosis 1
1.2.1.1 Multiple sclerosis pathogenesis 2
1.2.1.2 Diagnosis of multiple sclerosis 5
1.2.1.3 Subtypes of multiple sclerosis 6
1.2.1.4 Epidemiology of multiple sclerosis 10
1.2.1.5 Genetics of multiple sclerosis 13
1.2.1.6 Gene expression studies in multiple sclerosis 15
1.2.2 System level approaches in complex diseases 16
1.2.2.1 Proteomic studies in multiple sclerosis 16
1.2.2.2 Pathway and network oriented GWAS analysis 19
1.2.2.3 Molecular Pathway studies in multiple sclerosis 20
1.3 Aim Of The Study: 21
2. MATERIALS AND METHODS **22**
2.1 Sample Selection 22
2.2 Sample Collection And Storage 24
2.3 Proteome And Targeted Protein Analysis 25
2.3.1 Protein isolation from cerebrospinal fluid samples 25
2.3.2 Protein purification and enrichment 25
2.3.3 Bradford assay 26
2.3.4 Isoelectric focusing 26
2.3.5 Sodium dodecly sulphate polyacrylamide gel electrophoresis (SDS-PAGE) 27
2.3.6 Protein spot identification and removal 28
2.3.7 MALDI-TOF mass spectrometry 28

2.3.8 ELISA analysis of candidate protein biomarkers 29
2.3.9 Western blot analysis of selected protein biomarkers 31
2.3.10 Statistical analysis of Proteome Studies 32
2.4 Pathway Analysis Using Bioinformatics Approaches 32
2.4.1 PANOGA tool and KEGG pathway analysis 33
2.4.3 Statistical comparison of molecular pathway analysis 34
2.5 Genomic Level Studies 34
2.5.1 DNA isolation from blood samples 34
2.5.2 Genome-wide association study using familial MS cases 35
2.5.3 Linkage analysis 36
2.5.4 Comparison of proteomics, linkage and GWAS results 37
3. RESULTS 38
3.1 Sample Characteristics Of Study Groups 38
3.2 Optimization Of Proteome Studies 39
3.3 Differentially Expressed Protein Biomarkers In Multiple Sclerosis Subtypes 42
3.3 Differentially Expressed Protein Biomarkers In MS Converting CIS Patients 48
3.4 Differentially Expressed Common And Shared Proteins 49
3.5 Confirmation Of Individual Proteins 50
3.5 Molecular Pathways In MS Subtypes 60
3.6 Shared And Subtype Specific Molecular Pathways 61
3.7 Molecular Pathways In MS Converting CIS Subtype 63
3.8 GWAS Study Revealed Disease Associated Significant SNPs 64
3.9 Linkage Analysis Indicated Two Chromosomal Loci In Chr 13 And Chr 21 68
3.10 Correlation Of Proteomic And Genomic Results 71
4. DISCUSSION: 73
4.1 Potential Protein Biomarkers In Multiple Sclerosis 73
4.2 Subtype Specific And Common Protein Biomarkers In Multiple Sclerosis 76
4.3 Subtype Specific And Common Molecular Pathways In Multiple Sclerosis 77
4.4 The Importance Of Renin-Angiotensin System Pathway And Complement And Coagulation System Pathway In Multiple Sclerosis 83
4.5 The Importance Of Vitamin Metabolism In Multiple Sclerosis 84
4.6 The Importance Of Salt Metabolism And Its Molecular Basis 84
4.7 Important Molecular Pathways In Clinical Conversion From CIS To Multiple Sclerosis 86

4.8 Identification Of Genome Wide Potential Biomarkers And Comparison Of Results With Proteome Studies.....88
5. CONCLUSION.....91
REFERENCES.....93
CURRICULUM VITAE.....107

ABBREVIATIONS

2D – PAGE : 2 Dimensional Polyacrylamide Gel Electrophoresis
A1AC : Alpha-1 Antichymotrypsin
A2MG : Alpha-1 Macroglobulin
ACN : Acetonitrile
ADEM : Acute Disseminated Encephalomyelitis
ANGT : Angiotensinojen
APC : Antigen Presenting Cell
APS : Ammonium Per Sulphate
BBB : Blood Brain Barrier
BCS : Balo's concentric lesions
BSA : Bovine Serum Albumin
CCC : Complement and Coagulation Cascade
CD : Cluster of Differentiation
CDMS : Clinically Definite Multiple Sclerosis
cDNA : Complementary DNA
CHCA : α-Cyano-4-hydroxycinnamic acid
CI : Confidence Interval
CIS : Clinically Isolated Syndrome
CNS : Central Nervous System
CSF : Cerebrospinal Fluid
CV : Coefficient of Variation
DIS : Dissemination in Space
DIT : Dissemination in Time
DMSO : Dimethyl Sulfoxide
DMT : Disease Modifying therapies
EAE : Experimental Autoimmune Encephalomyelitis
EBNA : Epstein Barr Virus Nuclear Antigen
EBV : Epstein-Barr Virus
EDSS : Expanded Disability Status Scale
ELISA : Enzyme Linked Immuno Sorbent Assay
GFAP : Glial Fibrillary Acidic Protein
GWAS : Genome Wide Association Studies
H_2O_2 : Hydrogen Peroxide
HHV : Herpes Human Virus
HLA : Human Leukocyte Antigen
HRP : Horse Radish Peroxidase
HSV : Herpes Simplex Virus
HWE : Hardy Weinberg Equilibrium
ICAM1 : Intracellular Adhesion Molecule-1
IEF : Isoelectric Focusing
IFN : Interferon
IgG : Immunoglobulin G
IL : Interleukin
IMSGC : International Multiple Sclerosis Genetics Consortium

IMSGC : International Multiple Sclerosis Genetics Consortium
INS : Insulin
IR : Insulin Receptor
KEGG : Kyoto Encyclopedia of Genes and Genomes
LHON : Lebers Hereditary Optic Neuropathy
LP : Lumbar Punction
MAB21L1 : Mab-21-Like 1
MAF : Minor Allele Frequency
MALDI-TOF: Matrix Assisted Laser Desorption Ionization – Time Of Flight
MBP : Myelin Basic Protein
MeOH : Methanol
MHC : Major Histocompatibility Complex
miRNA : MicroRNA
MMP : Matrix Metallo Proteinase
MOG : Myelin Oligodendrocyte Glycoprotein
MRI : Magnetic Resonance Imaging
mRNA : Messenger Ribonucleic Acid
MS : Multiple Sclerosis
MSIF : Multiple Sclerosis International Federation
MSs-CIS : Multiple Sclerosis Suggestive Clinically isolated Syndrome
NaOH : Sodium Hydroxide
NEDD4L : E3 Ubiquitin-Protein Ligase
NFL : Neurofilament Light Chain
NH_4HCO_3 : Ammonium bicarbonate
NK : Natural Killer
NMO : Neuromyelitis Optica
NPL : Non-Parametric Linkage
OCB : Oligoclonal Band
OND : Other Neurological Diseases
OR : Odd Ratio
PANOGA : Pathway and Network Oriented GWAS
PBMC : Peripheral Blood Mononuclear Cell
PBS : Phosphate Buffered Saline
PBST : Phosphate Buffered Saline with Tween20
PCR : Polymerase Chain Reaction
PCR : Polymerase Chain Reaction
PDG2 : Prostaglandin D2 Synthase
PI3P : Phosphatidyl Linositol 3-Phopsphate
PML : Progressive Multifocal Leucoencephalopathy
PPI : Protein Protein Interaction
PPMS : Primary Progressive Multiple Sclerosis
PRMS : Progressive Relapsing Multiple Sclerosis
PVC : Poly Vinil Chlorur
RAS : Renin Angiotensin System
RIS : Radiologically Isolated Syndrome
RRMS : Relapsing-Remitting Multiple Sclerosis
SAMS : Single Attack Multiple Sclerosis
SAPMS : Single Attack Progressive Multiple Sclerosis
SDS – PAGE : Sodium Dodedcyl Sulphate Polyacrylamide Gel Electrophoresis
SGK : Serine/Threonine-protein kinase

SNP : Single Nucleotide Polymorphisms
SPMS : Secondary Progressive Multiple Sclerosis
TCA : Trichloro Acetic Acid
TE : Tris EDTA
TEMED : Tetramethylethylenediamine
TFA : Trifluoroacetic acid
TGFbeta-1 : Transforming Growth Factor Beta - 1
Th : T Helper
TIM : T cell Ig- and mucin-domain-containing molecules
TNF : Tumor Necrosis Factor
Treg : Regulatory T Cell
TTR : Transthyretin
US : United States
VCAM1 : Vascular-Cell Adhesion Molecule-1
VDBP : Vitamin D Binding Protein
WHO : World Health Organisation
WTCCC2 : Welcome Trust Case Control Consortium
ZAG : Zinc-alpha-2-glycoprotein
α1-AGP : Alpha-1 acid glycoprotein

LIST OF TABLES

Page

Table 2.1 : Selected candidate proteins for confirmation studies 31
Table 3.1 : Clinic information of case and control samples 39
Table 3.2 : Spot optimization results 42
Table 3.3 : Differentially expressed proteins in RRMS subtype 44
Table 3.4 : Differentially expressed proteins in CIS subtype 46
Table 3.5 : Differentially expressed proteins in PPMS subtype 47
Table 3.6 : Differentially expressed proteins in CIS to MS converters 49
Table 3.7 : Frequency of common and shared proteins in different MS subtypes 51
Table 3.8 : Confirmation study results of all selected proteins 61
Table 3.9 : Shared molecular pathways among all MS subtypes 62
Table 3.10 : Molecular pathways indicated by proteome studies in CIS subtype 63
Table 3.11 : Molecular pathways indicated by proteome studies in RRMS subtype 63
Table 3.12 : Molecular pathways indicated by proteome studies in PPMS subtype. 64
Table 3.13 : Molecular pathways and associated proteins revealed by CIS subgroup analysis 65
Table 3.14 : Significant SNPs with $p<10^{-4}$ 66
Table 3.15 : SNPs with strong associations with $p<10^{-3}$ 67
Table 3.16 : Associated SNPs at the chr 13q13.3 70
Table 3.17 : Associated SNPs at the chr 21q22.2 71
Table 3.18 : The list of correlated genes and molecular pathway from proteome and GWAS results 72

LIST OF FIGURES

Page

Figure 1.1 : Immunopathogenesis of MS 3
Figure 1.2 : MS Subtypes 7
Figure 1.3 : Disease progression and phase switches 9
Figure 1.4 : Epidemiology of MS 11
Figure 1.5 : Recurrence risk for MS 13
Figure 2.1 : Overview and design of the proteomic study 25
Figure 2.2 : Family trees used for the linkage and GWAS studies 36
Figure 3.1 : Proteins spot optimization with different IPG strips 40
Figure 3.2 : Protein spot counts and matching rate between different IPG strips 41
Figure 3.3 : 2D PAGE images of same sample precipitated with different methods 42
Figure 3.4 : Representative 2D gel images for all MS subtypes 44
Figure 3.5 : Confirmation of APOA4 protein in different MS subtypes 52
Figure 3.6 : Confirmation of APOE protein in different MS subtypes 53
Figure 3.7 : Confirmation of TTR protein in different MS subtypes 54
Figure 3.8 : Confirmation of PDG2 protein in different MS subtypes 55
Figure 3.9 : Confirmation of α1-AGP protein in different MS subtypes 56
Figure 3.10 : ELISA confirmation studies for previously selected proteins 57
Figure 3.11 : Confirmation of vitamin D binding protein with ELISA 57
Figure 3.12 : Confirmation of α2 - macroglobulin protein with ELISA 58
Figure 3.13 : Confirmation of trypsin protein with ELISA 59
Figure 3.14 : Confirmation of Sodium reabsorption pathway proteins with ELISA in CSF samples 60
Figure 3.15 : Confirmation of Sodium reabsorption pathway proteins with ELISA in serum samples 60
Figure 3.16 : SimWalk Multipoint NPL Analysis showing the highest scores 69
Figure 3.17 : Fine mapping of chromosome 13 from 28 cM to 46 cM region 70
Figure 3.18 : Fine mapping of chromosome 21 from 35 cM to 58 cM region 71
Figure 4.1 : Common and subtype specific proteins 77
Figure 4.2 : Renin angiotensin system pathway 78
Figure 4.3 : Complement and coagulation cascade 79
Figure 4.4 : Proteasome pathway 80
Figure 4.5 : Vitamin digestion and absorption pathway 81
Figure 4.6 : Notch signaling pathway 82
Figure 4.7 : Pathogenic E. coli infection pathway 83
Figure 4.8 : Aldosterone regulated sodium reabsorption pathway 86

INVESTIGATION OF MOLECULAR BIOMARKERS IN MULTIPLE SCLEROSIS CLINICAL SUBTYPES

SUMMARY

Multiple sclerosis (MS) is a neuro-inflammatory and neuro-degenerative disease of the central nervous system (CNS), in which myelin sheaths and axons are damaged. Inflammatory demyelination and other pathological changes disrupt the ability of the nervous system to communicate resulting the wide variety of signs and symptoms. Heterogeneous clinical presentation and course ıf the disease, without an effective diagnostic laboratory test, make it difficult to draw firm conclusions about the disease's states, drug response and in general prognosis which may lead to more severe forms in some cases.

The purpose of this study is to find out molecular biomarkers and pathways associated with the etiopathogenesis of the disease and its major subclinical forms. To this aim, proteomic, bioinformatics and genomic approaches were used to reveal differentially expressed proteins and their related pathways in cerebrospinal fluids of patients with clinically isolated syndrome (CIS), relapsing remitting MS (RRMS) and progressive MS (PPMS).

Cerebrospinal fluid (CSF) samples of 65 CIS, 72 RRMS, 42 PPMS and 42 controls including healthy subjects and other neurological disease samples were analyzed. For all samples, 2D-PAGE analyses were performed. At least 2 times differently expressed protein spots were identified by PDQuest® software and removed from the gel for MALDI-TOF-MS protein identification analysis. The results of proteomic studies were further confirmed by both ELISA and western blot methods for selected candidate proteins. KEGG pathways analysis was performed with genes corresponding the proteins to identify disease relevant specific molecular pathways.

Further linkage and genome-wide association study (GWAS) were performed by using single nucleotide polymorphism (SNP) genotyping on the Illumina CytoSNP 300K array in 36 individuals from 10 informative MS families having more than one affected member. Nonparametric linkage (NPL) scores were calculated for each of 3118 SNP markers spaced at an average of 1 cM intervals using SimWalk multipoint NPL analysis. Subsequent fine mapping of regions showing higher NPL scores was performed. GWAS analysis was performed by using 18 MS and 60 healthy subjects data by comparing the 298114 different genome wide SNPs.

Comparison of disease groups with controls identified a total of 151 proteins that are differentially expressed in clinically different MS subtypes. KEGG analysis using PANOGA tool revealed disease related pathways including aldosterone-regulated sodium reabsorption ($p=8.02x10^{-5}$) which is important in the immune cell migration, renin-angiotensin ($p=6.88x10^{-5}$) system that induces Th17 dependent immunity,

notch signaling ($p=1.83x10^{-10}$) pathway indicating the activated remyelination and vitamin digestion and absorption pathways ($p=1.73x10^{-5}$).

Linkage study including 10 MS families showed suggestive evidence of linkage (NPL scores above 1.7) to chromosomal regions of 13q12.2-14.11 and 21q22.12-22.3. Fine mapping of these regions revealed that the most promising loci for linkage were mapped to 13q13.3 and 21q22.2, with NPL scores of 1.82 and 1.85, respectively. Detailed search of these regions revealed inflammation and neurodegeneration associated genes some of which previously implicated in MS and also promising novel genes yet to be identified like N4BP2L2 and TRPC4 were found within the suggestive regions. On the other hand, GWAS analysis revealed disease associated SNPs with significant associations on disease pathogenesis associated genes like INS-IGF2 ($p=4.39x10^{-7}$ MAF = Case:Control = 0.36/0.03), MGMT ($p=9.66x10^{-6}$ MAF = Case:Control = 0.32/0.04).

Individual comparison of genomic and proteomic data revealed 9 different genes correlated with 14 different proteins, in 4 different molecular pathways that are associated with proteomic data. Those genes and proteins are the most prominent molecular parameters to understand disease mechanisms and needed to be further verified with various studies.

To conclude, proteomic investigation of CSF samples belonging to different MS phenotypes indicated that all MS clinical forms share common biological pathways such as renin-angiotensin system (RAS) and complement and coagulation cascade (CCC) pathways. There are also clinical subtypes specific and pathophysiology related pathways, which may have further therapeutic implications. Furthermore, our study implicated importance of salt metabolism in MS pathogenesis for the first time. On the other hand, genomic results also showed disease pathogenesis associated genetic variations and chromosomal regions. Our suggestive results provide a framework for deep sequencing to identify new susceptibility genes and novel variants associated with risk of MS.

MULTİPL SKLEROZ KLİNİK ALT TİPLERİNDE MOLEKÜLER YOLAKLARIN VE BİYOBELİRTEÇLERİN ARAŞTIRILMASI

ÖZET

Multipl skleroz (MS) merkezi sinir sisteminde (CNS) nöroenflamasyon ve nörodejenerasyon ile karakterize, otoimmün kökenli olduğu kabul edilen, enflamatuvar, demiyelinizan bir hastalıktır. MS patojenezinde dolaşım sisteminde oto reaktif hale gelen bağışıklık sistemi hücrelerinin kan beyin duvarını (BBB) aşarak ya da CNS'de otoimmun hale gelen çeşitli hücrelerin nöronları çevreleyen myelin kılıfa saldırması görülmektedir. Nöronların ve daha ileri seviyelerde oligodendrosit hücrelerin yıkılması CNS'de farklı bölgelerde ve farklı zamanlarda lezyonların oluşmasına neden olmaktadır. Bu lezyonlarla ilişkili olarak gelişen demiyelinizasyon ve aksonal dejenerasyon MS'deki nörolojik engellilikten sorumlu olarak kabul edilmektedir.

MS'in etiyolojisi tam olarak bilinmemekle birlikte genetik ve çevresel risk faktörlerinin etkileşimiyle kompleks kalıtım paterninde ortaya çıktığı en yaygın görüştür. Genç erişkinlerde nontravmatik nörolojik hastalıklar arasında özürlülüğe yol açan en önemli hastalıktır. Kronik, nöroenflamatuvar, nörodejeneratif bir hastalık olarak tanımlanan Multipl skleroz'un (MS), kompleks patojenezi ve farklı prognozları ile birlikte çeşitli klinik alt tipleri bulunmaktadır. Bu alt tipler hastalığın ilerleyişine ve semptomların ya da radyolojik bulguların ortaya çıkma sıklığına bağlı olarak farklı isimlerle sınıflandırılmaktadır. Klinik izole sendrom (CIS), yineleyici – düzelen MS (RRMS), birincil ilerleyen (PPMS) ve ikincil ilerleyen (SPMS) MS'in ana klinik alt tipleri olmakla birlikte MS'in farklı formlarına benzeyen fakat ortaya çıkışı ve diğer klinik parametrelere bağlı olarak farklı MS varyantları da bulunmaktadır. MS'in klinik alt tipleri arasında geçişler de görülebilmektedir.

Farklı klinik alt tipleri, patojenezinde farklılıklar ve hastalığın kompleks yapısı göz önünde bulundurulduğunda, MS'in tanı ve prognozunun da ne kadar karmaşık olduğu görülmektedir. Bu nedenle hastalığın tanı ve prognozunda kullanılabilecek biyobelirteçlere ihtiyaç duyulmaktadır. Biyobelirteçler, normal biyolojik süreçleri, hastalık süreçlerini, veya ilaç yanıtlarının göstergeleri olarak, nesnel, objektif ölçülebilen ve değerlendirilebilen özellikleri taşıyan, bağımsız biyolojik moleküller ve görüntüleme simgeleri olarak tanımlanırlar. Biyobelirteçler ayrıca hastalıkların biyolojik mekanizmalarının çözümlenmesinde de önemli rol oynarlar. Özellikle MS gibi klinik ve genetik heterojenite gösteren kompleks hastalıklarda hastalığın klinik alt tiplerinin birbirinden ayırt edilmesinde, farklı ilaçlara verilebilecek cevapların önceden anlaşılabilmesinde ayırıcı biyobelirteçlere ihtiyaç vardır. Farklı alt tiplerle ilişkilifarklı biyobelirteçlerin yer aldığı moleküler yolakların ya da protein – protein ilişkilerinin ortaya çıkarılması da hastalığın moleküler mekanizmalarının aydınlatılmasına yardımcı olacağı gibi, tanı ve prognozun yanı sıra tedavi etkinliğinin de artırılabilmesine olanak sağlamaktadır.

Bu çalışmada amaç MS alt tiplerine ait BOS örneklerinde, proteomik ve biyoinformatik yaklaşımlarla elde edilen protein biyobelirteçleri kullanılarak, farklı MS alt tiplerinin etiopatojenezinde rol alan moleküler yolakların ortaya konulmasıdır. Bu amaçla ilk olarak farklı klinik alt tiplerdeki aday protein biyobelirteçler, kontrol örnekleri ile yapılan kıyaslamalar ile ortaya konularak, proteilerin ifadelenme seviyeleri belirlenmiştir. Daha sonra, bu proteinler ve ifadelenme oranları kullanılarak PANOGA (Pathway and Network-Oriented GWAS Analysis) web ara yüzü ile KEGG (Kyoto Encyclopedia of Genes and Genomes) moleküler yolakları, biyoinformatik analizler ile belirlenmiştir. Bu yolaklarda yer alan diğer ilişkili proteinler belirlenerek hedef genler ve proteinler ortaya konulmuştur.

Genom seviyesinde biyobelirteçlerin ortaya konulabilmesi için de ailesel MS olgularında genom boyu bağlantı (linkage) ve ilişkilendirme (association) çalışmalarında da hastalık ile ilişkili genetik varyantlar ve kromozom bölgelerinin belirlenmesi sağlanmıştır. Proteomik ve genomik verilerin karşılaştırılması ile elde korelasyon gösteren protein ve genlerin MS patojenezindeki rolleri göz önüne alınarak açığa çıkarılması hedeflenmiştir.

Proteomik çalışmalarda 65 CIS, 72 RRMS, 42 ilerleyici MS (PMS) (PPMS ve SPMS birlikte) ve 42 kontrol örneğine (sağlıklı kontroller ve diğer nörolojik hastalıklara sahip hastalardan elde edilen örnekler) ait beyin omurilik sıvısı (BOS) ve serum kullanılarak her bir örnekten protein izolasyonu ve saflaştırılması yapıldı. 2D-PAGE (iki boyutlu jel elektroforezi) analizi ile en az 2 kat farklı miktarda ifadelendiği tespit edilen protein spotlarını belirlemek üzere farklı klinik alt tipler ve kontrollerle karşılaştırılmıştır. Tüm protein spotlarının kontrollerle karılaştırıldığında kat değişimleri belirlenerek listelenmiştir. Tüm protein spotları jellerden çıkarılarak kütle spektrometresi için hazır hale getirilmiştir. Matriks assited laser desorption ionization – time of flight kütle spektrometresi (MALDI-TOF-MS) analizi yapılarak spotların ait olduğu peptit dizileri ve proteinlerin isimleri tanımlanmıştır.

Genom boyu bağlantı ve asosiasyon çalışmaları için 28 farklı ailesel MS örnekleri toplanmıştır. Bu ailelerden bağlantı analizi için en uygun 10 aile seçilmiştir. Bu aileler MS tanısı almış 19 hasta ve 17 sağlıklı bireyden oluşmaktadır. Toplam olarak bu ailelerden 36 örneğe mikrodizin çalışması yapılmıştır. Aileler, birinci derece yakınlarında MS tanısı almış hastalar ve en az 1 sağlıklı birey içeren ailelerden oluşmaktadır. Mikrodizin çalışması için Illumina CytoSNP 300K SNP genotiplendirme çipleri kullanılmıştır. Elde edilen ham veriler hem bağlantı analizleri için hem de ilişkilendirme çalışmaları için kullanılmıştır. Aileler içerisinde de bağlantı analizi yapmak için, parametrik olmayan bağlantı (NPL) skoru, SimWalk multipoint NPL analizi ile 3118 belirteç üzerinde ve ortalama 1 cM aralıklarla incelenmiştir. Sonrasında ise, detaylı haritalama ile daha yüksek çözünürlükte analiz yapıldı ve NPL değeri 1.8 üzerinde olan kromozomal bölgeler elde edildi.

GWAS analizi için ailele MS örneklerinde birbiri ile akraba olmayan 11 MS hastası belirlenerek ve yine birbiri ile akrabalık ilişkisi bulunmayan 60 sağlıklı bireyle karşılaştırma yapılmıştır. Sağlıklı bireylerin mikrodizin analizlerinde de aynı Illumina çipleri (300K) kullanılmıştır. Biyoinformatik analiz için Golden Helix® SNP & Variation Suite (SVS) programı kullanılmıştır. Yapılan analizlerde ilk olarak 298114 genom boyu SNP verisi farklı özelliklerine göre filtreleme işlemine tabi tutulmuştur. Kalite ve kantitatif filtrelemeler sonucunda 129547 SNP belirteçi hasta ve sağlıklılarda karşılaştırılmıştır.

Proteomik çalışmalarda, MS klinik alt tipleri ve kontroller arasında yapılan karşılaştırmalarda kontrollere göre istatiksel olarak anlamlı miktarda farklılık gösteren 151 protein tespit edilmiştir. Bu proteinlerden literatürde daha önce MS ile ilişkilendirilmiş olanlar ve ilk kez tespit edilenler listelenmiştir. KEGG yolak analizi sonucunda da, farklı miktarlarda ifadelenen protein biyobelirteçleri kullanılarak, hastalık grupları arasında ortak ve alt gruplara özgü moleküler yolaklar ortaya çıkarılmıştır. Bunlardan *renin-angiotensin system* (RAS) ve *complement and coagulation cascade* (CCC) yolakları tüm MS alt gruplarında ortak olarak kontrol grubuna göre artmış bir aktivite göstermektedir ($p<0.0001$). İmmun hücre göçünde önemli bir yolak olan *Aldosterone-regulated sodium reabsorption* ($p=8.02x10^{-5}$) yolağı, Th17-bağımlı immunite gelişiminde önemli *renin-angiotensin* ($p=6.88x10^{-5}$), aktif remyelinizasyonda etkili *notch-signaling* (p=1.83x10-10) yolağı ve vitamin emiliminde görevli *vitamin digestion and absorption* ($p=1.73x10^{-5}$) yolağı da klinik alt tipler arasında farklılıklar göstermiştir.

Genomik çalışmalarda ise, 13q12.2-14.11 ve 21q22.12-22.3 kromozom bölgelerinde NPL skoru 1.8 ve üzeri bağlantı bölgeleri tespit edilmiştir. Detaylı haritalama ile de en yüksek bağlantı skorlarının 13q13.3 ve 21q22.2 bölgelerinde olduğu gösterilmiştir. Bu bölgelerde yer alan genlerin incelenmesi ile de hem MS ile daha önce ilişkilendirilmiş hem de ilişkilendirilmemiş inflamasyon ve nörodejenerasyon yolaklarında yer alan genler listelenmiştir. Ayrıca N4BP2L2 ve TRPC4 görevleri gereği MS ile daha detaylı incelenmesi gereken yeni aday potansiyeli olan genler tespit edilmiştir. GWAS çalışmalarında ise hastalık ile ilişkili SNP'ler listelenerek bu varyantların bulunduğu genler gösterilmiştir. Bu genlerin bir kısmının (*INS - IGF2*, $p = 4.39x10^{-7}$ MAF = Case : Control = 0.36/0.03, *MGMT,* $p = 9.66x10^{-6}$ MAF = Case : Control = 0.32/0.04) MS patojenezinde yer alan genler olduğu tespit edilmiştir.

Proteomik ve genomik sonuçların, bireysel olarak tüm ailesel MS örneklerinde karşılaştırılması sonucunda, 4 farklı moleküler yolakta, 9 gen ve 14 proteinin korelasyonu gösterilmiştir. Bu genler ve proteinler hastalığın moleküler mekanizmasının anlamamız için en uygun moleküler parametreler olup, farklı diğer çalışmalarla doğrulanması gerekmektedir.

Sonuç olarak bu çalışma ile farklı MS klinik alt tiplerinde etkili moleküler yolaklar belirlenmiştir. Hem klinik alt tiplerde ortak yolaklar hem de farklı yolaklar ortaya çıkarılmıştır. Ayrıca MS'de tuz metabolizmasının, hastalığın gelişiminde ve prognozundaki önemi ilk kez *in vivo* olarak ortaya konulmuştur. Elde ettiğimiz veriler, literatürdeki klinik ve hayvan modeli çalışmaları ile uyumluluk gösterirken, hedef moleküler yolakların hasta örneklerine ait BOS sıvılarındaki değişimleri ilk kez gösterilmiştir. Ayrıca faklı MS klinik alt tiplerine ait potansiyel protein biyobelirteçleri listelenerek hastalığın tanı ve prognozunda kullanılabilme potansiyelleri tartışılmıştır. Genomik sonuçlarla proteomik sonuçların karşılaştırılması ile hastalık patojenezinde önemli ve korelasyon gösteren aday genler ve proteinler ortaya konmuştur. Gelecek çalışmalarda ise aday genlerin ve proteinlerin hastalığın in vivo modellerinde ve klinik çalışmalarda doğrulama çalışmaları yapılması hedeflenmektedir.

1. INTRODUCTION

1.1 Purpose Of Thesis

The aims of this study were to evaluate potential diagnostic and prognostic protein biomarkers via top to down proteomic approaches and identification of disease related, common and shared molecular pathways via bioinformatics approaches. Further, determination of multiple sclerosis (MS) related genes, gene variations and chromosomal regions on familial MS cases by linkage and genome wide association studies were aimed. Finally, we aimed to reveal key molecular players in MS pathogenesis via comparison of proteomic and genomic data.

1.2 Literature Review

1.2.1 Multiple sclerosis

Multiple sclerosis (MS) is also known as disseminated sclerosis and it is first described by Jean Martin-Charcot in 1868 by differentiation of clinical symptoms and grey plaques throughout the brain and spinal cord (Charcot, 1868). MS is a chronic autoimmune neuro-inflammatory and degenerative disease of the central nervous system (CNS), in which the insulating covers of the nerve cells, so called myelin sheaths, and axons are damaged (Goldenberg, 2012). Inflammatory demyelination and other pathological changes disrupt the ability of nervous system to communicate resulting wide variety of signs and symptoms. Highly variable and unpredictable course of the disease is initially characterized by reversible neurological symptoms, which are often followed by progressive neurological deterioration over time.

Several forms of MS can be seen with new symptoms either occurring in independent attacks (relapsing forms) or building up over time (progressive forms). Symptoms of MS may disappear completely after remission period. However, permanent neurological problems often occur, especially within the progressive forms of the disease. The exact cause of MS still is not clear and among many

proposed etiological mechanisms, destruction of nerve cells by immune cells and/or failure of the myelin producing cells appear to gain most supporting evidence. Nevertheless, all factors are linked to both genetic and environmental factors and their interactions.

1.2.1.1 Multiple sclerosis pathogenesis

For many years, the central topic of MS research has been still to find out the disease's pathogensis mechanism. Central mechanisms underlying the disease are inflammation, demyelination and neurodegeneration (Figure 1.1). Several *in vivo* and in *vitro* studies either with tissue samples and animal models, point out the importance of autoimmune processes.

The CNS is always considered as an immunologically preserved region even it includes rather few immunological cells in the presence of any active or ongoing infection. Yet, many studies have proved that a few number of T cells can pass through the CNS by penetrating into the blood-brain barrier (BBB), searching for infection or injury (Hafler and Weiner, 1987). Auto-reactive T and B cells are regular elements of the immune system. Studies indicated that myelin components or other structural proteins can stimulate some of these auto-reactive cells in a healthy body, but this does not mean an autoimmune process unless, the tolerance of immune system gets broken and cells auto-activated (Sakaguchi, 2000). The cause of autoimmune response against myelin components in the CNS develops with one of the two mechanisms that are molecular mimicry or activation and epitope spreading (McFarland and Martin, 2007). Activation of myelin specific T cells lead to induction of microglia, macrophages, astrocytes and B cells, by proliferation in CNS and secretin inflammatory cytokines (Peterson and Fujinami, 2007).

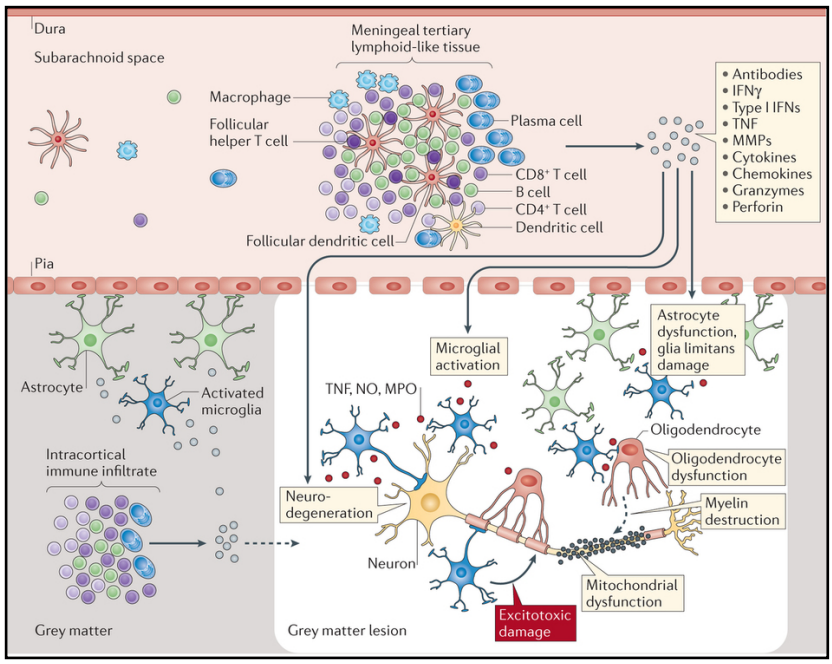

Figure 1.1: Immunopathogenesis of MS (Calabrese et al., 2015).

BBB functions as a barrier to protect CNS from immune cells and unwanted molecules by restricting the entry. It also prevents the transfer of CNS antigens to the systemic circulation. Permeability of BBB can change depending on the intrinsic (genetic) and extrinsic (environmental) factors like viral or bacterial infections (Obermeier et al., 2013). The increase of permeability initially starts with increase of the level of adhesion molecules like ICAM-1 ("Intracellular Adhesion Molecule-1"), VCAM-1 ("Vascular-Cell Adhesion Molecule-1") and E-Selectin (Kebir et al., 2007). Matrix metallo - proteases initiates the destruction of BBB (Minagar and Alexander, 2003). Activated T cells in the CNS begin to synthesize of proinflammatory cytokines, interleukin (IL) -1, -6, -8, interferone (IFN) -gamma and Tumor Necrosis Factor (TNF) -alpha (Cannella and Raine, 1995). Proinflammatory cytokines increase the expression of surface molecules of lymphocytes and antigen presenting cells (APCs). Especially, IFN gamma stimulates the microglia cells to activate as APCs (Aloisi et al., 2000). When a class II Major Histocompatibility Complex (MHC) carrying APC associates with an undifferentiated T cell and MS antigen, it can enter to one of following two pathways depending on the co-stimulatory molecule on the surface of T cell. If the co-stimulatory molecule on the surface of the T cell is CD28 and its ligand B7-1 is present, T cells differentiate into CD4+ Th1 cells and synthesize pro-inflammatory IL-12 and IFN-gamma cytokines. This pathway is called as T helper Type 1 (Th1) mediated inflammatory

pathway.(Lincoln et al., 2005). If the co-stimulatory molecule on the surface of the T cell is CTLA-4 and its ligand B7-2 is present, T cells differentiate into CD4+ Th2 cells and initiate the anti-inflammatory IL-1, IL-4 and IL-10 cytokines synthesis. This pathway is called as T helper Type 2 (Th2) mediated inflammatory pathway (Kidd, 2003). IL-12 stimulates the Th1 pathway whereas IFN-gamma inactivates the Th2 pathway. On the other hand, IL-4 and IL-10 stimulate the Th2 pathway, IL-1 inactivates the Th1 pathway. In Th1 pathway, TNF-α cytokine initiates the death of oligodendrocyte (OLG) cells and damages to the myelin sheath. Damage occurs with the stimulation of TNF-α receptors leading to activation of apoptosis and necrosis. In Th2 pathway, TNF-α receptors are activated, and remyelination process can start. Macrophages which are activated by IFN-gamma cytokines cause myelin sheath damage either by stimulating complement response or directly. In a healthy body, normally Th1 and Th2 cell and their products are in a balanced state (Kidd, 2003).

For more than 40 years immunologists have divided T helper cells into Th1 and Th2 subsets. Identification of Th17 cells led to breaking the monopoly of Th1/Th2 mediated inflammation understanding in immunopathogenesis of MS, and its experimental models (Steinman, 2007). Autoreactive Th17 cells can penetrate into CNS and by the production of IL-17 and IL-22 cytokines, which disrupt tight junction proteins in the CNS. As a result, Th17-mediated inflammation is characterized by neutrophil recruitment into the CNS and oligodendrocyte destruction leading to death of neurons (Jadidi-Niaragh and Mirshafiey, 2011).

Inflammation is thought to be the primary sign of MS for many years. The inflammatory process caused by T cells had always been taken as a proof of concept that the inflammation is the main detrimental process in MS. However emerging genetic, pathologic and magnetic resonance imaging (MRI) data indicate that inflammation has both beneficial and unfavorable features in the pathogenesis of MS (Martino et al., 2002). The advantage of inflammation is that it is fundamentally important for destruction of pathogens, activating immune system, which is strictly regulated, and it is able to trigger counter reactive mechanisms in order to suppress the inflammation. However, the loss of control in inflammatory process would lead to chronic and unpreventable inflammation in many systems including CNS.

Demyelination is the loss or destruction of myelin sheath of nerves and it is the hallmark of many autoimmune neurodegenerative diseases. It commonly results in

various symptoms determined by the functions of the affected nerve cells. The characteristic feature of MS plaques is demyelination (Lucchinetti et al., 2000). Many possible biochemical and cellular mechanisms may lead to destruction of myelin and oligodendrocytes via immune cells and their elements. Direct binding of T or B cells to myelin epitopes can lead to activation of macrophages. Similarly, release of cytokines from T cells or microglia / macrophages can lead to demyelination (Brück, 2005).

As described above, the architecture of MS pathogenesis is highly complex and include many players like immune cells, nerve cells and their supporting elements. A typical MS lesion contains these elements. However, depending on the disease stage, interactions all of these may change. In the early stages, inflammatory demyelination causes to the formation of so-called 'MS plaques'. In later stages or progressive forms of MS, demyelination in the cerebral cortex becomes pervasive and degenerative changes diffuse throughout the entire white and grey matter. Therefore, in patients with PPMS and SPMS, brain and spinal cord atrophy with extensive nerve loss and dilatation of cerebral ventricles can be seen. Thus, while the disease process starts with inflammation driven focal demyelinating lesions, later diffuse neurodegeneration appears which affects the entire CNS (Lassmann and van Horssen, 2011).

1.2.1.2 Diagnosis of multiple sclerosis

Diagnosis of MS is mainly based on clinical, laboratory and imaging assessments. The aim is to emphasize disseminations of lesions in time (DIT) and in space (DIS). Furthermore, excluding the alternative diagnoses is also necessary as most of the disease symptoms can be seen in other neurological diseases (OND) (Polman et al., 2011). There is no single blood test or other radiological technic that is able to diagnose definite MS. In some cases, history and symptoms of the patients are enough to diagnose the disease. But some patients may need to have multiple tests and prolonged observation for diagnosis. To standardize the diagnosis procedure, there are operational criteria that combine clinical parameters and radiologic findings with supportive laboratory analyses.

Several schemes for diagnosis of MS have been established up to date. Previously, diagnosis was solely based on clinical and imaging findings (Schumacher, 1950).

But later, these criteria have been revised depending on the new advancements in our understanding of MS. First, Poser et. al. emphasized the importance of supportive laboratory findings for MS diagnosis (Poser et al., 1983). Later, Poser et. al, reviewed the criteria based on the magnetic resonance imaging (MRI) variabilities among MS patients (Poser and Brinar, 2001). In 2001, McDonald Criteria have been evolved in international Panel on MS diagnosis, which focused on objective demonstration of dissemination of lesions in both time and space (McDonald et al., 2001). Later these criteria were revised by Polman et al, at 2005 (Polman et al., 2005) and 2010 (Polman et al., 2011). In 2005 revision, together with the new studies, international committee of MS have pointed to the need for simplification of parameters to improve their comprehension and utility and for evaluating their appropriateness in different populations. In 2010, international committee met for a third time to examine requirements for demonstrating DIS and DIT and to focus on application of the McDonald Criteria in pediatric, Asian, and Latin American populations. Finally, besides the clinical and radiological parameters, increase of IgG in the cerebrospinal fluid (CSF), IgG positivity and presence of plasma cells in demyelinating areas is also necessary for MS diagnosis.

1.2.1.3 Subtypes of multiple sclerosis

The standard terminology used to describe the patterns and courses of MS is important in order to correctly classify the disease. Since MS is not a single disease, rather it is regarded as a spectrum of disorders and it is classified into different subtypes and its variants, depending on the progression patterns and the frequency of disease symptoms. Clinically isolated syndromes (CIS), relapsing-remitting MS (RRMS), primary progressive (PPMS) and secondary progressive MS (SPMS) are the different MS clinical phenotypes (Figure 1.2). There are also MS variants that the symptoms and radiological findings resemble MS, but they are different from definite MS (Simon and Kleinschmidt-DeMasters, 2008). Opticospinal MS, Balo's and Marburg's diseases, acute disseminated encephalomyelitis (ADEM) are considered as different forms of MS due to their heterogeneous clinical, imaging and pathological features (Siva, 2006).

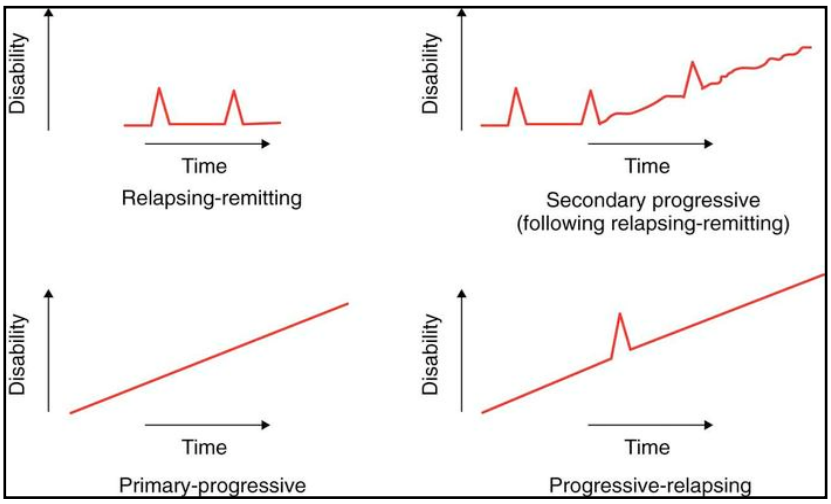

Figure 1.2: MS Subtypes (Hannoun et al., 2012).

CIS is an early clinical variant of MS that is described as first clinical episode with MS suggestive features (Miller et al., 2012). It usually occurs in young adults and in 85% of MS cases onset is a subacute clinically isolated syndrome (Miller et al., 2005b). In CIS, optic nerves, brainstem, and the spinal cord are primarily affected (Miller et al., 2012). In many patients who later covert to clinically definite (CDMS), the clinical onset starts with an acute or subacute episode of monoregional neurological involvement of the CNS. In some cases of CIS, MRI may reveal polyregional involvement of the CNS. Although normal brain MRI and CSF IgG index indicate the low risk of developing MS in following 5 years, on the other hand, positive findings increases this risk 80-90% of the cases. Currently, conversion to CDMS requires an additional attach or dissemination of lesion in time or space. However, not all the cases definitely convert to CDMS. Estimating the risk of conversion is important regarding to early treatment decisions. Because, some CIS cases are self-limited and do not require long-term treatment (Dalton et al., 2002). Therefore, an early and accurate identification of patients who are going to convert to CDMS has become of great importance, since it has been reported that early treatment of these patients improves their clinical outcome (Comi et al., 2001). The presence of oligoclonal IgG bands is highly specific and sensitive for early prediction of conversion to multiple sclerosis. MRI criteria have a high specificity but less sensitivity. The simultaneous use of both tests shows high sensitivity and specificity in predicting clinically isolated demyelinating syndrome conversion to CDMS (Masjuan et al., 2006).

Relapsing-remitting MS (RRMS) is the most common form of the disease and characterized by the acute attacks (relapses) and following total or partial recovery (remissions). The initial presentation in most patients (frequency in all newly diagnosed MS is 85%) is RRMS with relapses and remissions due to self-limiting plaques of inflammatory demyelination disseminated in time and place in the CNS. Following relapses arise in an unpredictable fashion and after each attack level of disability of the patient can increase. RRMS shows the tendency to turn into SPMS form in the following years (%50) (Siva, 2006).

PPMS is a continuous form of the disease. The course of disease continues without clear relapses from the beginning. There can be plateaus (periods of stabilization) time intervals. It is characterized by steady worsening of neurological functions and 10–15% of all MS patients are in this group (Miller and Leary, 2007). In a rare form of progressive MS, - progressive relapsing MS (PRMS) - progression occurs from the beginning and occasional exacerbations appears during the disease and people with PRMS may not experience recovery. The clinical course of PPMS is variable, although most patients develop significant motor disability in 15 to 30 years after onset. Compared with RRMS, people with PPMS are older at onset and a higher proportion of affected are men (Miller and Leary, 2007). Another progressive form of the disease is secondary progressive MS (SPMS) which starts as RRMS and later (approximately 5 to 6 years) steady progression appears with or without clinical relapses. The level of disability increases with each relapses (Figure 1.3). The frequency of conversion from RRMS to SPMS depends on the duration of the follow - up study. In the longer follow ups, the higher (up to 90%) conversion is observed (Vukusic and Confavreux, 2003).

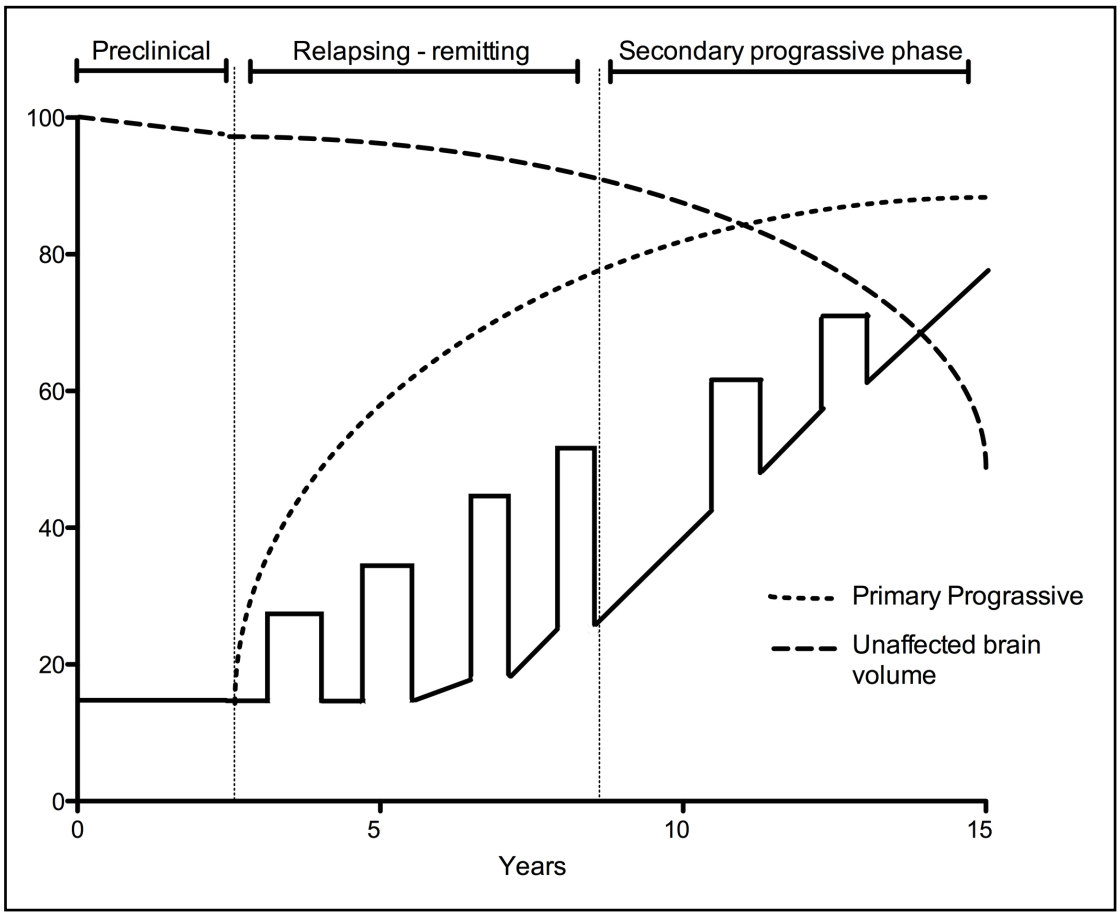

Figure 1.3: Disease progression and phase switches (Turanli et al., 2015).

Radiologically isolated syndrome (RIS) is characterized by the presence of MRI findings resembling to MS, but with normal neurological functions and without any other MS suggestive symptoms (De Stefano and Siva, 2013). It is previously indicated that approximately two-thirds of RIS cases show radiological progression and the rest develop neurological symptoms during mean follow-up times of up to five years (Granberg et al., 2012). Treatment options for RIS are still in debate and it is still unclear whether RIS individuals should be treated before the symptoms appear.

There are severe acute and subacute forms of MS which are uncommon and considered as MS variants due to demyelinating aspects of the neuropathological lesions (Capello and Mancardi, 2004). Marburg type of MS is rare and malign form of disease with rapid progression. Cognitive impairments, advanced level of vision loss, dysarthria, dysphagia and respiratory problems are observed (Capello and Mancardi, 2004). Balo's concentric lesions (BCS) are also considered as a variant of MS due to demyelinating resemblance. The course of BCS continues with destruction of brain white matter in concentric layers in a manner that leaves the axis

cylinders intact. Clinical course is similar to Marburg Type, which has a monophasic rapidly progressive course with a fatal outcome. Most of the reported cases were young adults (Karaarslan et al., 2001). Neuromyelitis optica (also known as Devic's syndrome) is an idiopathic, severe, demyelinating disease of CNS that affects the optic nerve and spinal cord. In many patients, optic neuritis (ON) continues with the MS like relapses and remittings. Spinal cord lesions can be permanent. Its prognosis is as serious as MS and can lead to vision loss. A patient with isolated ON may recover completely, or have recurrences, or develop a progressive optic neuropathy, or may remain symptom-free, or eventually convert to RRMS and SPMS (Lennon et al., 2004). Acute Disseminated Encephalomyelitis (ADEM) is a nonvasculitic, inflammatory, demyelinated, immune response originated form of MS. It is usually monophasic and frequently seen in 6-8 years old children and adolescents. It usually occurs following a viral infection, but may appear after one week or month of vaccination, bacterial or parasitic infection (Wender, 2011).

1.2.1.4 Epidemiology of multiple sclerosis

The prevalence of MS varies among populations and geographical distributions depending on environmental and genetic factors. The Multiple Sclerosis Foundation reported that more than 400,000 people diagnosed with MS in the United States (US) and currently there are more than 2.5 million MS people around the world. Each week average of 200 new MS cases are diagnosed in the US (Higuera, 2015). Population prevalence of MS increases with distance from the equator (Simpson et al., 2011). It is estimated that in southern part of the world, the rate of MS frequency is between 57 and 78 cases per 100,000 people. The rate is 110 to 140 cases per 100,000 in Northern part (Figure 1.4). (Tullman, 2013). Climate also appears to be related with the MS incidence, because people of Northern European decents have the highest risk of developing MS. These population based studies also reveal that the descents of people have more vulnerability, no matter where they live (Simpson et al., 2011). Native Americans, Africans, and Asians have the lowest risk of MS development based on MS prevalence studies (Tullman, 2013). MS occurs two to three times more in women than men. Although the disease has a broad range of age at onset (85% of cases occur between the ages of 14 and 55), initial symptoms typically present in early adulthood (between ages 20 and 40). Incidence of MS has

been found to be increasing over time, which may be related to more advanced diagnostic schemes or other unknown etiologies (Kenealy et al., 2003).

Unfortunately, there is no up to date prevalence and incidence report for MS in Turkey. The latest study by Çelik Y et al, at 2011 indicated that prevalence of MS is approximate 34/100000 in Turkey. The mean annual incidence was 3.48/100000 (Çelik et al., 2011). The other European countries showed variable MS prevalence: $290/10^5$ in Ireland (Lonergan et al., 2011), $173/10^5$ in Sweden (Ahlgren et al., 2011), $8/10^5$ in France (Fromont et al., 2012), $95/10^5$ in Italy (Millefiorini et al., 2010), $182/10^5$ in Denmark (Eaton et al., 2007), $148/10^5$ in Norway (Smestad et al., 2008) and $53/10^5$ in Spain (Fernandez et al., 2008). These data and numbers may change in different studies depending on the regional differences in populations and regarding pediatric cases; but, the Atlas of MS 2013 reported a global median prevalence of $33/10^5$ in 2013 (Federation, 2013).

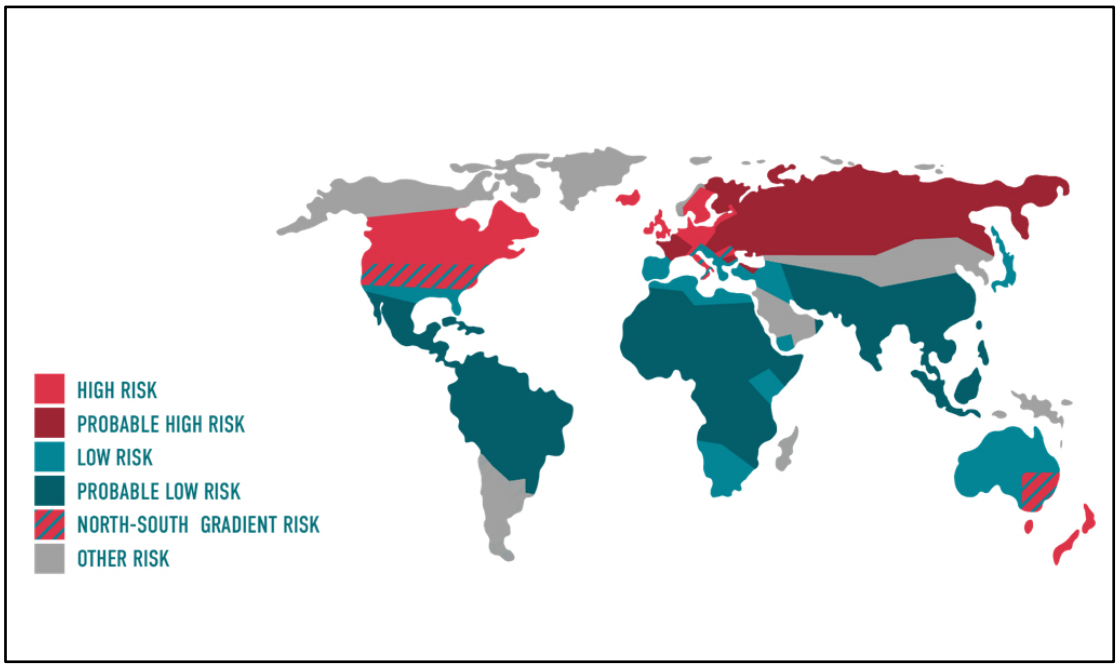

Figure 1.4: Epidemiology of MS. Global Distribution of disease (Adapted from MultipleSclerosis.net).

Besides genetic factors, it is known that variable distribution of MS prevalence can be explained in part by both environmental factors and population-specific genetics. Other environmental factors besides the latitude include dietary vitamin D levels, infectious agents and cigarette smoking. Studies on vitamin D serum levels indicated that higher vitamin D levels are associated with decreased exacerbation risk in RRMS. However, the possibility of reverse causality cannot be ruled out completely

(Runia et al., 2012). In another study, it is shown that low level of serum vitamin D is a strong risk factor for long-term MS activity and progression (Ascherio et al., 2014). The second is the *prevalence hypothesis* which postulates that MS is caused by a pathogen (viral or microbe) that is common in high MS prevalence regions (Kurtzke, 1993). Studies on infectious causes of MS are mostly based on the clinical and pathological evidences. There are two common hypotheses about the infectious causes of MS. First one is the *poliomyelitis hypothesis,* which indicates that the presence of a virus acquired in childhood or adulthood increases the risk of MS, whereas it may also serve as a potential protective factor, if it is acquired in infancy (Poskanzer et al., 1976). The first hypothesis appears to have gained more favor than the second and it is later evolved to *hygiene hypothesis*. Hygiene hypothesis indicates the exposure of infectious agents early in life is protective against MS. This hypothesis was later supported by many studies indicating the high MS prevalence in higher education and income populations (Ascherio and Munger, 2007a). In other studies, Epstein Barr Virus (EBV) is associated with higher MS risk. It is shown that people with EBV infection have more tendency to develop MS (Levin et al., 2010), but the pathologic evidences of EBV infection is still missing (Willis et al., 2009). However, it is previously shown that being carrier for EBV antigens (EBNA-1, EBNA-2 or EA-2), increases the risk of MS development 1.6 to 2.5 times (Ascherio and Munger, 2007a). There are also other studies indicating the potential association of different viruses like Herpes Human Virus 6 (HHV6) (Álvarez-Lafuente et al., 2006), Herpes Simplex Virus (HSV) (Hawkes et al., 2006), measles (Krakowka et al., 1983), varicella zoster (Sotelo et al., 2008) and rubella (Brettschneider et al., 2009). However, up to date none of these studies implicated a strong correlation that best fits to disease mechanisms.

Cigarette smoking is a risk factor for MS development. (Ascherio and Munger, 2007b). Several biologic models could explain the increased risk of MS among including the effects of smoking on the immune system, BBB compactness, and toxic effects of on the CNS (Hernan et al., 2001). Several studies indicate that conversion from RRMS to SPMS and CIS to CDMS is also associated with cigarette smoking (Hernán et al., 2005)(Di Pauli et al., 2008). Even parental smoking was found to increase the risk with the early onset MS development (Mikaeloff et al., 2007). Regarding the metabolic effects of smoking, it can be concluded that smoking has been associated with both risk of MS and progressive disease.

1.2.1.5 Genetics of multiple sclerosis

The genetic contribution to susceptibility for MS development is supported by several lines of studies involving classical family and twin studies as well as genome wide association (GWAS) and genetic screening studies.

Family studies indicating genetic contribution to the disease based on the first, second and third degree of familial relations indicate 2 to 10 times increased risk for MS development (Nielsen et al., 2005). Numerous familial aggregation studies have shown that recurrence risk ratio for MS decreases with the degree of relationship between individuals meaning that more close individuals of family have more risk of MS development (Nielsen et al., 2005). In a family based study, Sadovnick AD. et. al. showed that monozygotic siblings (30.3%) have more risk to develop MS than dizygotes (%4.6) and siblings (5.1%) (Sadovnick and Ebers, 1995). In another family based population study, adopted children were analyzed to understand the environmental effect to the disease, Ebers GC. et. al.'s results indicated that prevalence of MS in adopted children (%0.05) was not significantly higher than general population. This result defined that genetic sharing is more important than shared family environment (Figure 1.5) (Ebers et al., 1995). Twin studies also indicate the relative importance of genetic sharing. Twin MS patients' a sibling with MS were analyzed and results indicated that twin MS patients have more recurrence risk (RR) (19%) than general population (Ebers et al., 2004). Migration studies indicated that migration from a high-risk area to the low risk area decreases the possibility of developing MS. Compared to the general population, these elevated risks suggest a strong complex inheritance of MS susceptibility (Sadovnick, 2012).

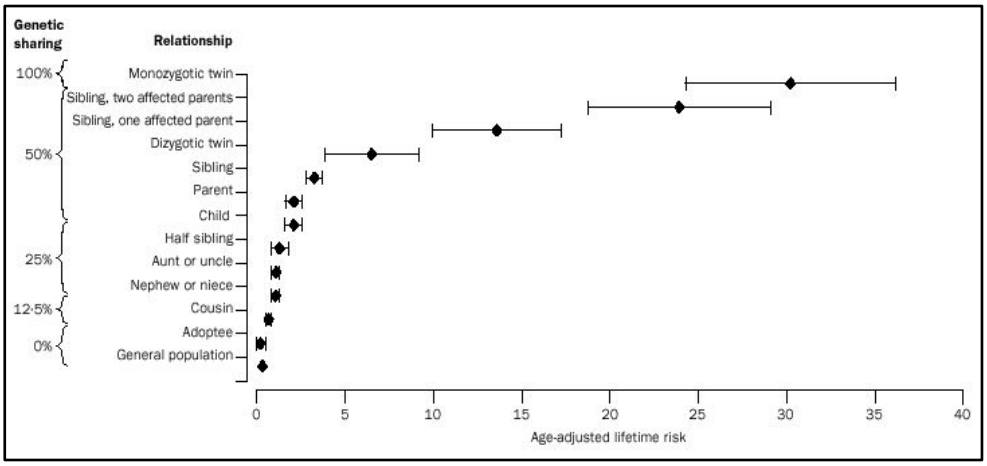

Figure 1.5: Recurrence risk for MS (Compston and Coles, 2002).

In the early studies of MS molecular genetics, a possible association was found in Human Leukocyte Antigen (HLA) genes. For more than 50 years, contribution of HLA locus to disease development has still been not certain. Studies showed association with certain HLA genes on chromosome 6p21, including HLA-A, HLA-DRB1, HLA-DQB1, and HLA-DRA on chromosome 6p21.3 (Sadovnick et al., 1996). Initial associations were observed in the HLA class I region of the MHC and later with polymorphisms in the HLA class II region. In Northern Europeans, susceptibility has been well mapped to an extended HLA class II haplotypes DQA1*0102-DQB1*0602-DRB1*1501-DRB5*0101 (Gourraud et al., 2012). Carrier frequencies of these haplotypes ranges in Sardinian families ranges between 38.5% - 58.6% in patients and 25% - 30% in healthy members (Marrosu et al., 1998).

Developments in high throughput genomic screening technology facilitated the identification of new MS related genetic gene and chromosomal loci. Analysis of whole genome SNPs and exom sequencing technologies led to discovery of new candidate MS genetic susceptibility genes. The first comprehensive GWAS study was conducted by De Jager et. al in 2009 with 2624 subject and 2116 control samples. They showed new susceptibility loci at TNFRSF1A ($p=1.59x10^{-11}$), IRF8 ($p= 3.73x10^{-9}$) and CD6 ($p=3.79x10^{-9}$) genes with strong odd ratios (De Jager et al., 2009). International Multiple Sclerosis Genetics Consortium (IMSGC) genome wide analysis study (GWAS) in 2013 analyzed 14498 MS cases and 24091 healthy controls by microarray technology and identified 110 risk variants in 103 discrete loci outside of the HLA complex genes. Among them, variants in VCAM1, CD58, CD86, IL12A, IL7R, IL2RA, CD69, CD40, MAPK1 and CXCR5 have the higher odd ratios ranging between 1.1 - 2.4. They also revealed 48 new susceptibility variants including STAT4, CCR4, FOXP1, CARD11 and MAPK3. So far, this study is the most comprehensive genetic study on MS with the highest sample number and largest analyzed genomic coverage (Consortium, 2013). Whole exome sequencing of 43 probands from MS families revealed variants in CYP27B1 gene which is associated in vitamin D metabolism (Ramagopalan et al., 2011). Currently IL7R and IL2RA are the known associated genetic variants in MS. Despite the conflicting study results, Traboulsee et al genotyped 1978 MS and 830 control samples. They indicated that IL7R is not a strong genetic variant however, there is a significant association in PPMS, therefore it may not be a causing determinant but it is a strong

determinant of disease course in MS. The effect of IL2RA genetic variation in patients without a family history of MS indicated that the genetic components for familial and sporadic disease are perhaps distinct (Traboulsee et al., 2014). In another study, together with the IL2RA, IL7R and IL12A, a new risk variant is detected in SOCS1 gene. Study screened 55 genes in 3919 MS and 4003 control samples (Vandenbroeck et al., 2012). Meta analysis of all disease related genes, SNPs and loci enabled investigation of the total effect of genetic variations in MS pathogenesis. Regarding all genomic studies including GWAS and linkage studies, it can be concluded that immunologically relevant genes, vitamin metabolism genes and particularly T-helper-cell differentiation pathway genes are affected in the pathogenesis of MS.

1.2.1.6 Gene expression studies in multiple sclerosis

Microarray based comparative gene profiling technique provides an overview of genes underlying a particular condition. Several large scale microarray studies using various samples types (brain tissue, blood, T cells and CSF) of MS patients have been reviewed (Dutta and Trapp, 2012). In a microarray study conducted with 24 MS and 17 control sample's PBMCs were analyzed. 7500 cDNA clones were investigated and results indicated increased levels of platelet activating factor (PAF), TNF receptor (CD27), zeta-chain associated protein kinase (ZAP70) and CD3e genes compared to healthy control samples. Regarding the interactions of all increased and decreased mRNA products, authors concluded that large number of genetic factors have to be kept in place to maintain immunological tolerance, however a slight dysregulation may cause its unbalance and development of MS pathology. Their findings also showed that the activation of autoreactive T cells is of primary importance to disease development (Bomprezzi et al., 2003). Expression study in active MS plaques with and without demyelination showed increased expression of transforming growth factor beta - 1 (TGF-beta1), Jagged 1 and Notch genes. These data demonstrated the restriction of Notch pathway during demyelination which indicated the potential therapeutic target for progressive MS forms (John et al., 2002). In a cDNA array study on 15 RRMS and 15 age-matched healthy controls, more than 400 genes were analyzed and 34 were significantly differed in RRMS group. 12 of these genes were immune related genes indicating the role of active inflammatory genes in RRMS patients (Ramanathan et al., 2001). In another study,

in order to understand the pathogenesis of MS lesions, the comparison of active and inactive plaques of MS patients were analyzed. It showed that different sets of genes playing roles in inflammatory, apoptosis related and stress-induced pathways were increased (Mycko et al., 2004). Gene expression studies in a subset of T helper cells obtained from CSF is challenging and difficult to perform because of low level of T helper cells. T cells obtained from active lesion carrying MS patients indicated elevated level of T cell Ig- and mucin-domain-containing molecules (TIM-1) and TIM-3 as well as IFN-γ and TNF-α (Khademi et al., 2004). The level of chemokines (CXCR3) and chemokine receptors (CCR5), (Sørensen et al., 1999), IL-15 (Kivisäkk et al., 1998), interleukin 7 receptor alpha (IL7-Rα) also increased in CSF of MS patients having acute demyelination (Lundmark et al., 2007). As a result gene expression studies mainly focused on targeted genes in various sample types in order to show the expression level of MS pathogenesis genes. Currently, TGF-beta1, TNF-α, CCR5, IL7-Rα are the most prominent ones.

1.2.2 System level approaches in complex diseases

1.2.2.1 Proteomic studies in multiple sclerosis

The proteome is the total sum of proteins in particular cells, tissue or a biological sample from a healthy or diseased state. Therefore, profiling of the proteome for MS like complex diseases is fundamental to understand the disease pathophysiology. Therefore, as explained above, conflicting and complex results obtained from genomic and transcriptomics studies needed to be correlated and elucidated with proteomic studies. Proteomics technologies have evolved and got easier to be performed at last decade. The use of proteomic techniques and strategies in complex diseases increased our knowledge of protein structure, function and variations in certain conditions or during the disease course. This is especially crucial in MS because of the increasing incidence of the disease and the challenges in the treatment and prognosis and the need for a biomarker.

A biomarker is a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes or pharmacological responses to a therapeutic intervention. There are various pathophysiological processes contributing to the development of MS or to the evolution of MS subtypes among each other. On the basis of analysis of the published studies examining

etiological mechanisms in MS, candidate biomarkers maybe classified into one of these five categories depending on the processes that those proteins are involved in: Biomarkers of immune system, biomarkers of demyelination, biomarkers of axonal/neuronal damage, biomarkers of remyelination and repair, biomarkers of BBB disruption.

Proteomic studies reveal many targets and biomarker candidates are still in the infancy stages to be able to be used in either diagnosis or the prognosis of the disease stages (Turanli et al., 2015). There are various methodological approaches, each has advantages and disadvantages depending on the aim of the study. 2D-PAGE (Two Dimensional Polyacrylamide Gel Electrophoresis) is the central methodology to separate the proteins. In proteomic techniques, following a separation approach (SDS-PAGE, chromatography, centrifugation etc.), an identification technique (mass spectrometry, protein chips, protein microarrays, western blot, ELISA etc.) is applied. Finally, bioinformatics analysis of obtained data is performed to understand the function and characteristic features of proteins.

Proteomic studies are mainly focused on the biomarker discovery for diagnosis, prognosis and understanding pathophysiology of MS. Studies from different disease subtypes on serum/CSF or CNS tissue samples, disease stages or response to therapeutic interventions point many candidate protein biomarkers for MS. One of the most important issues in biomarker identification is the selection of proper biological sample that should be suitable for clinical use in terms of its ease of use and repeatability of the test as well as sensitivity and specificity. The first option is usually considered to be the blood, which can easily be used for diagnosis and follow-up in clinical practice due to its accessibility and minimally invasive collection procedure. However, there are several drawbacks that are required to be considered. Albumin and IgG constitute approximately 75% of the total protein weight in blood plasma/serum, and 20 additional proteins make up most of the remaining weight. The other hundreds of low-abundance proteins account for only approximately 1% of the protein weight in plasma/serum (Jaros et al., 2013). The depletion of the most abundant proteins is the gold standard employed in proteomic studies, but it is always important to note the possibility of the depletion of these proteins might also deplete other proteins that may have importance in pathogenesis (Koutroukides et al., 2011). Furthermore, depletion of high-abundance proteins, such

as albumin, IgG, transferrin, and transthyretin should also be performed prior to the proteomic analysis of CSF. Nevertheless, CSF is an ideal sample type for identifying modifications in the CNS and is therefore an interesting source of biomarker candidates. The major drawback of employing CSF as the biological fluid for clinical applications is the invasive nature of its collection and requirement of strict ethical rules. In an ideal world, CNS tissue would be the most correct tissue for discovering biomarkers. However, the difficulty of collecting such tissue from living patients prevents using CNS tissue markers routinely. Despite this limitation, CNS tissue is still a rich source of information for better understanding and characterizing the pathophysiology of MS. With this aim, CNS tissues from post-mortem human samples and animal models of MS (cuprizone and Experimental Autoimmune Encephalomyelitis – EAE - model) have been investigated by proteomic techniques and cross-compared with the CSF and serum results in order to establish the best candidate biomarker for MS.

Serum proteomic analysis of 103 patients with CIS compared with healthy controls indicated the antibodies against myelin oligodendrocyte glycoprotein (MOG) and myelin basic protein (MBP) as predictors of CDMS after a first demyelinating event (Berger et al., 2003). CSF 2D-PAGE analysis of 12 CIS, 12 RRMS and 20 control samples revealed the one up-regulated, ten down-regulated proteins in CIS and two up-regulated, eleven down-regulated proteins in RRMS group including Apolipoprotein A (ApoA), Apolipoportein E (ApoE), Transferrin, Protein Kinase A and Albumin family proteins those that are later confirmed by many other studies (Lehmensiek et al., 2007). Ottervald J's study, analyzed 112 MS 209 coMMPntrol samples including RRMS in relapse and remission state, CIS, SPMS, other neurological diseases (OND) and healthy controls. They used 2D-PAGE and ELISA validation approach to reveal candidate biomarkers. In RRMS patients, there were significant increases in the CSF levels of alpha-1 antichymotrypsin (A1AC), alpha-1 macroglobulin (A2MG) and fibulin 1 as compared to control subjects. In SPMS contactin 1, fetuin A, vitamin D binding protein (VDBP) and angiotensinogen (ANGT) were increased compared to control subjects. In particular, ANGT was increased 3-fold in SPMS, indicating a potential for its role as a biomarker of disease progression in MS. In PPMS, A1AC and A2MG proteins showed significantly higher CSF levels than controls, with a trend of increase for ANGT. Further

classification models using decision trees could identified 70% of the RRMS and 80% of the SPMS patients correctly based on the protein biomarkers (Ottervald et al., 2010).

Ideally, studies indicating potential biomarkers in CSF of MS should be correlated and confirmed in serum samples in order to be used in less invasively. To this aim, serum proteomic analysis studies focused on to reveal reliable protein biomarkers. Avasarala JR et al's study (2005), tested the sera from 25 healthy controls and 25 RRMS patients using MALDI-TOF MS approach. ApoE, MOG, CXCL10, CXCL12, CXCL13, MMP-9 and MMP-2 proteins have been shown to be correlated with CSF proteome compared to healthy controls (Avasarala et al., 2005). Estimation of the rates of conversion between clinical subtypes as well as underlying molecular signatures have been very important for MS research. Tumani H. et al's study (2009) analyzed the CIS patients remained as CIS (n=8) and CIS converting to RRMS (n=8) who were followed for at least 2 years. They identified one protein that was up-regulated in CIS-RRMS (serin peptidase inhibitor) and eight proteins (alpha-1-B-glycoprotein, Fetuin-A, apolipoprotein A4, haptoglobin, human Zinc-alpha-2-glycoprotein (ZAG), Retinol-binding protein, superoxid dismutase 1, transferrin) that were down-regulated in CIS-RRMS vs. CIS-CIS (Tumani et al., 2009).

Despite the rapidly rising number of studies aiming to identify potential biomarkers for MS, no reliable protein candidate has been approved up to date.

1.2.2.2 Pathway and network oriented GWAS analysis

The advancements in high throughput laboratory technologies enabled us to analyze -omic features of complex diseases. As it becomes easier and cheaper, researchers are able to perform whole genome, proteome and transcriptome analysis to complex diseases in order to understand the mechanisms of the diseases or discover diagnostic and prognostic biomarkers. However, the data obtained from these studies are complicated and requires bioinformatics knowledge. PANOGA is a web based bioinformatics tool to meet these requirements.

PANOGA (pathway and network oriented GWAS analysis) method is used for identification of molecular pathways by using pathway markers which have significant evidence for genetic association with protein-protein interaction (PPI) networks, functional information of selected SNPs, and current knowledge of biochemical pathways (Bakir-Gungor and Sezerman, 2012). Since complex diseases are developed or aroused by the complex interactions of multigenic, environmental and epigenetic factors, discovering their interactions and total effects are challenging. PANOGA maps the genes to a PPI network and then determine the connected subnetworks targeted by the SNPs or proteins. Then, KEGG pathways in these subnetworks are found and significance of these modifications (variation, increase or decrease in level) are determined. The pathways are ranked according to the significance scores and are referred as the SNP targeted pathways (Bakir-Gungor et al., 2014). PANOGA is tested on several complex diseases including rheumatoid arthritis, intracranial aneurysm, epilepsy, and Behcet's disease; and proved to be useful. In these tests, PANOGA found several SNP targeted pathways, which were also identified in previous studies, and additional pathways, where the relations of the genes in these pathways to the trait were confirmed (Bakir-Gungor et al., 2013).

1.2.2.3 Molecular Pathway studies in multiple sclerosis

Molecular pathway is an ordered molecular activity between molecules (genes, proteins, mRNA, miRNA, metabolic factors etc.) in a cell, which leads to a particular product or a change in a cell. In complex diseases, interaction of several genes and other factors may affect the disease development.

Since MS is a complex and multifactorial disease, uncovering of affected molecular pathways is important. However hundreds of thousands markers in genes are tested in these GWAS studies, making them difficult to compare and differentiate between disease subtypes (Oksenberg et al., 2008). Baranzini SE. et al, (2009) performed a meta analysis from 11 GWAS study, including 2 for MS and 9 for other neurological diseases and healthy controls in order to reveal disease associated molecular pathways. In the MS data sets, they identified sub-networks of genes from several immunological pathways including cell adhesion, communication and signaling, hematopoietic cell lineage, type I diabetes mellitus pathways. Remarkably, neural pathways, including axon-guidance and synaptic potentiation pathways, were also over-represented in MS (Baranzini et al., 2009). Lundmark F. et. al's study (2007)

investigated the role of interleukin 7 receptor alpha and interleukin 7 signaling pathway (IL7R-alpha-IL7) in MS by designing mRNA and genotyping assays. They analyzed 75 MS and 48 control samples for expression analysis. For genotyping of selected 10 SNPs, 1820 Nordic MS and 1234 control samples obtained. They showed the increased expression of IL7, but not IL7R-alpha in CSF samples. However, they showed that three SNPs (rs987107, rs987106, rs3194051) were significantly associated with MS group. Furthermore variations of all these SNPs were significantly correlated with the increased level of IL7. These results revealed the association of IL7R-alpha-IL7 pathway for the first time (Lundmark et al., 2007).

1.3 Aim Of The Study:

The aim of this study is to reveal potential biological biomarkers and disease related molecular pathways to understand the biology of MS and attempt to simplify the diagnosis and prognosis of the disease. To this aim, we first analyzed the CSF proteomic profile from a prospective cohort that has been followed for seven years, containing discrete clinical subtypes with a substantial sample size (n=221). We performed proteomic approaches to CSF and serum samples of clinically isolated syndrome (CIS), Relapsing remitting MS (RRMS), primary progressive MS (PPMS) and secondary progressive MS (SPMS) patients using 2D and MS. Further, the etiopathogenesis related molecular pathways in subclinical types of MS were investigated through bioinformatics approaches in order to find common and specific pathways. ELISA and western blot studies were performed verifications potential candidate protein biomarkers. Second, to uncover the disease related genetic markers, we performed whole genome genomic hybridization assay in familial MS families including disease and healthy individuals. Linkage and genome wide association analyses were performed to reveal disease associated genetic loci and variants. Finally, we looked for a significant correlation between proteomic and genomic results.

2. MATERIALS AND METHODS

2.1 Sample Selection

CSF and serum samples were collected from all patients at the time of entry to the study after reading and approving the informed consent of the study. The ethics approval (Protocol number: 930211) was obtained from Istanbul University, Faculty of Medicine at 2009. The study subjects were enrolled according to their clinical diagnosis, which were CIS, RRMS, PPMS, SPMS and control groups. After the enrollment, each patient group was clinically followed prospectively. The CIS patients were treatment-naïve and endpoints for the CIS patients were having a second clinical attack or appearance or new MRI lesions on follow-up MRI studies fulfilling McDonald 2005 (Polman et al., 2005) criteria for the diagnosis of MS. In order to evaluate the difference between CIS patients remaining CIS and converting to clinically definite MS subtype, CIS group was separated into three different subgroups:

1. MS suggestive CIS (MSs-CIS) indicates the CIS patients remaining CIS with a single clinical episode and MRI findings suggestive of MS but not fulfilling the diagnostic MRI criteria (either dissemination in time or space) to be diagnosed as MS during the follow up period.

2.Single attack MS (SA-MS) indicates the patients who fulfill 2010 McDonald Criteria either at onset of clinical episode or during the follow up period.

3. Clinically definite MS (CDMS) patients indicates the patients who fulfill the McDonald 2010 by having a second clinical attack which is previously known as Poser Criteria (Polman et al., 2011).

RRMS patient were followed for any change in their disability levels by Expanded Disability Status Scale (EDSS) and composite index and whether they would convert to SPMS or not. Single attack progressive MS (SAP-MS) patients were not included in the study. On the other hand, for the CIS group, serum and CSF samples were collected within the first 10 days of neurological episode of CIS and prior to any

treatment. The patients then were clinically followed prospectively until the termination of the study. Additionally, CIS patients were followed for the first 36 months both by clinical and imaging examination every 6 months and then on yearly basis or they were seen whenever they had a new clinical symptom and a sign suggestive of an MS episode. EDSS, MRI activity and clinical symptoms were followed to follow up disease progression. EDSS is a method of quantifying disability in MS. It ranges 0.0 to 10.0. EDSS steps 1.0 to 4.5 refer to people with MS who are fully ambulatory. EDSS steps 5.0 to 9.5 are defined by the impairment to ambulation (Bushnik, 2011).

The samples comprised cerebrospinal fluid (CSF) and serum samples from 65 patients initially diagnosed with clinically isolated syndrome (CIS), 72 patients with relapsing remitting MS (RRMS), and 42 progressive MS (PMS) group which was compromised of 8 secondary progressive MS (SPMS) and 34 primary progressive (PPMS). Control samples included 22 patients with non-inflammatory neurologic disorders such as acute headache and pseudotumor cerebri and 20 other neurological diseases (OND) with neuro-inflammatory properties such as neuro-Behçet disease, neuro-sarcoidosis and polyneuropathies. CIS patients were also subclassified depending on their clinical conversion to clinically definite MS (CDMS). Converters (SA-MS) and nonconverters (MSs-CIS) were sampled at the time of their enrollment to the study. Therefore, comparison of proteomic data between SA-MS and MSs-CIs subgroups were performed with their initial samples. Whole methodological approaches applied in proteomic studies given in Figure 2.1.

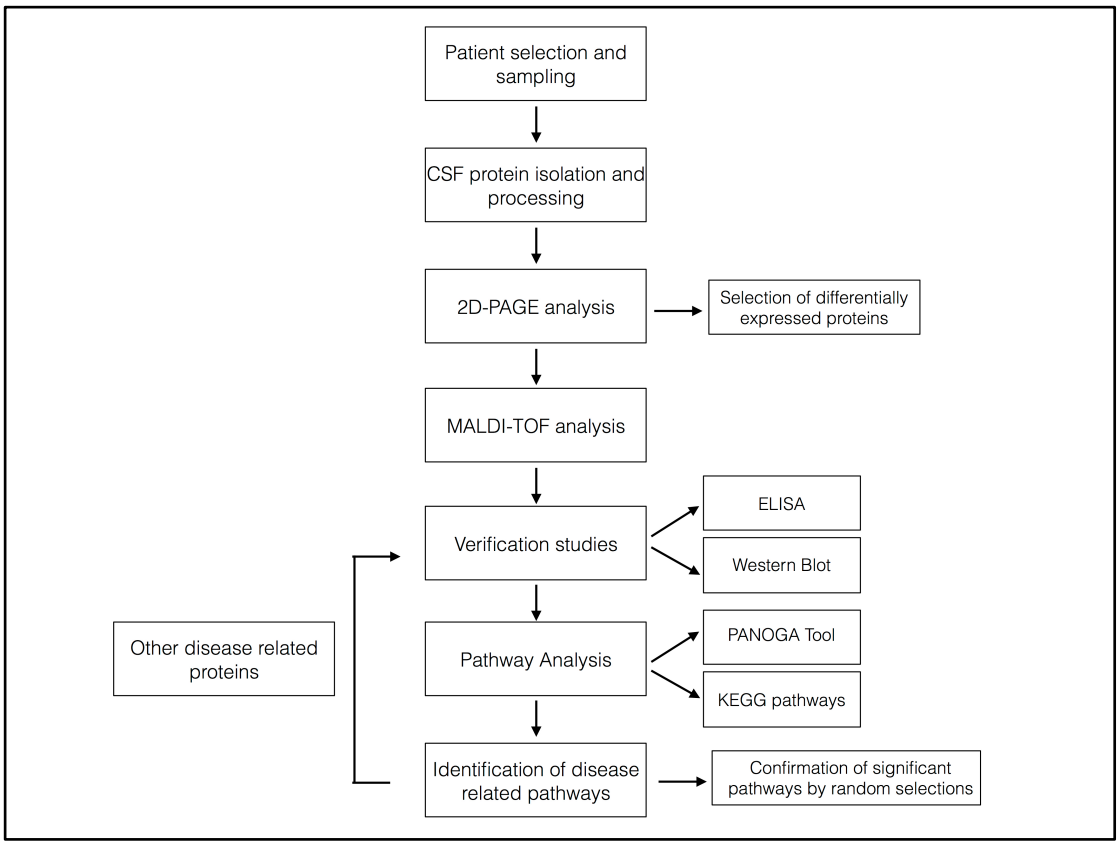

Figure 2.1: Overview and design of the proteomic study.

2.2 Sample Collection And Storage

Sample collection of patients and control groups is done at the Neuroimmunology and Demyelination Service clinics of Cerrahpaşa Faculty of Medicine. Collection of CSF is performed by lumbar punction (LP) (Freedman et al., 2005). LP is a diagnostic procedure in order to collect CSF for further analyses. To obtain the CSF, first the patient is placed in right or left lateral position and his/her neck-chest distance is reduced as possible as approximating a fetal position. LP is performed in the interspaces between the lumbar vertebrae, usually at the L4 – L5 level. Before the procedure, local anesthetic is applied to that region. With a spinal needle, 3cc of CSF is taken in to sterile, dry, empty tubes. Then, 3cc CSF is separated into 500 µl sterile eppendorf tubes to aliquote them. Eppendorf tubes were kept at -80^{o}C until analyses.

This study includes only CSF and serum samples, however other body fluids were collected from same cohort and stored for biobanking and further study purposes. For expression profile studies, blood samples were taken into PAXGENE RNA (PreAnalytix - Qiagen-BDTM Company) tubes in order to study RNA and stored first 24 hour on -20^{o}C then on -80^{o}C freezer. To study on plasma proteins, blood

samples were collected in P100 Blood Collection System (BDTM) for plasma protein preservation. Then tubes were centrifuged at 4000 rpm for 15 minutes, thus plasma was separated and tubes were stored at -20^{o}C. Serum samples were also taken by using mechanical serum separator tubes and after serum was separated, samples were aliquoted into sterile eppendorf tubes and stored at -80^{o}C. For further genomic analyses blood samples were taken into EDTA tubes and stored at -80^{o}C.

2.3 Proteome And Targeted Protein Analysis

2.3.1 Protein isolation from cerebrospinal fluid samples

Total proteome was isolated from CSF samples by following Trichloroacetic acid (TCA) protocol. TCA is widely used in biochemistry for the precipitation of macromolecules such as proteins, DNA and RNA. Week chlorine bonds in TCA bind to positively charged amino acids in proteins and helps to precipitation (Jiang et al., 2004).

Frozen CSF samples were thawed on ice, 1 volume of TCA (100%) stock was added to 4 volumes of protein sample in 1,5 ml sterile eppendorf tube and incubated on ice for 15 minutes. Then, it was centrifuged at 14.000 rpm for 10 minutes and supernatant is removed to leave protein pellet intact. Pellet was washed with 200 µl ice-cold acetone and centrifuged again at 14.000 rpm for 5 minutes; this washing step was repeated again with 200 µl acetone. Than pellet was dried by placing tube in 65^{o}C for 5 minutes to drive off acetone. Finally, pellet was suspended in 120 or 150 µl sterile dH_2O and mixed by vortexing for 5 minutes.

2.3.2 Protein purification and enrichment

Protein samples were purified and concentrated by using commercial kits. ProteoPrep Immunoaffinity Albumin and IgG Depletion Kit (Sigma-Aldrich, St. Louis, Missouri, USA) were used for albumin and IgG depletion. Manufacturer's protocol was followed for the purification which is a column based purification including silica beads for binding of albumin and IgG proteins.

2.3.3 Bradford assay

Bradford Assay was performed before all proteomic applications. Bradford assay was firstly described by Dr. Marion Bradford in 1976 (Bradford, 1976), and it defines the use of Coomassie Brilliant Blue Dye in a colorimetric reagent for detection and quantification of total proteins in a solution. In the acidic environment of the reagent, protein binds to the coomassie dye, which results in a spectral shift from the reddish/brown form of the dye (absorbance max. at 465 nm) to the blue form of the dye (absorbance max. at 610 nm). The difference between the two forms of the dye is greatest at 595 nm, therefore it is the optimal wavelength to measure the blue color from the coomassie dye-protein complex. Besides, the blue color can be measured at any wavelength between 575 nm and 615 nm. Bradford protein assay has been associated with the presence of certain basic amino acids (primarily arginine, lysine and histidine) in the protein. Van der Waals forces and hydrophobic interactions also participate in the binding of the dye to protein. The number of coomassie dye ligands bound to each protein molecule is approximately proportional to the number of positive charges found on the protein. Free amino acids, peptides and low molecular weight proteins do not produce color with coomassie dye reagents. The assay is performed at room temperature. The measurement was accomplished at 595 nm by using a 96 well plate. In order to measure protein concentrations ranging between 100-1500 µg/ml, 10 µl protein samples were prepared from protein stocks in 1:10 ratio, using distilled water as diluent. Bovine serum albumin (BSA) dilutions in a range of 0.125 mg/ml to 2 mg/ml diluted with distilled water were used as protein standards. Five µl of BSA standards and protein samples were loaded in a 96 well plate. 200 µl of Bradford Reagent (Bio-Rad Protein Assay, Bio-Rad Laboratories GmbH, Munich, Germany) was added onto each well and homogenized via pipetting. After incubation at room temperature for 2 minutes, the absorbance values were measured at 595 nm by Microplate Reader 3550-UV (Biorad).

2.3.4 Isoelectric focusing

Validation of isoelectric focusing (IEF) studies was accomplished on 7 cm strips. Later IEF studies were performed using 17 cm IPG strips pH 4-7, 5-8 and 6-9 (Bio-Rad) in order to increase resolution and separation of proteins. For 7 cm IPG strips 100 µg total protein, for the 17 cm IPG strips, 300 µg total protein was loaded in a

100 to 200 µl resuspension buffer. The first step is the rehydration of IPG strips, for consistently even distribution of the sample along the length of the strip. IPG strip was placed gel side down onto sample then 2-3 ml mineral oil was overlaid on the gel to prevent evaporation. Rehydration/equilibration tray was covered with the plastic lid and left on a level bench overnight (11-16 hour) to rehydrate the IPG strips and load the protein sample. Protocol consists of 7 steps by gradually increasing the voltage after rehydration for 4 hours: 30 V for 12 hours, from 30 V to 500 V in 30 minutes, 500 V for 1 hour, from 500 V to 1000V in 30 minutes, 1000 V for 1 hour, gradient from 1000 V to 8000 V in 30 minutes, 8000 V for 2 hours, making a total of approximately 30000 Vh.

2.3.5 Sodium dodecly sulphate polyacrylamide gel electrophoresis (SDS-PAGE)

Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) allows separation of proteins according to their sizes. Following the IEF, to separate the proteins based on their molecular size, SDS-PAGE was performed as second dimension. Different gel concentrations in different times were tried in order to observe best protein spot separation. During optimization studies, 2D-PAGE images were compared and analyzed regarding their maximum obtained protein spots. %10, 12 and 15 concentrations and 2 hours, 3 hours, 4 hours times were tried. As a result, 12% concentration with 3h electrophoresis time was selected as optimal conditions.

%12 Separating gel was prepared with 4ml of %30 acrylamide, 2.5 ml of1, 5M Tris pH 8.8, 100 µl of %10 SDS, 100 µl of %10 APS, 8 µl of TEMED, and 3.3 ml of dH_2O. Separating gel was transferred into SDS gel cassette via pipetting. A thin layer of isopropanol was added onto the gel in the cassette. The gel was let at room temperature for 30 minutes for polymerization. Next, the isopropanol in the cassette was discarded with a tissue paper and upper part of the gel is completely cleaned from the isopropanol by dH_2O. %5 stacking gel was prepared via mixing; 670 µl of %30 acrylamide, 500 µl of 0.5M Tris pH 6.8, 40 µl of %10 SDS, 40 µl of %10 APS, 4 µl of TEMED, and 2.7 ml of dH_2O. The stacking gel was added onto the polymerized separating gel in the cassette and left 45 minutes at room temperature for polymerization. 5 µl of ~10 mg/ml protein samples was mixed with 4X loading buffer and denatured at 95°C for 5 min. ~20 µl of each sample was loaded onto SDS-PAGE gel. As protein ladder, Precision Plus Protein™ Prestained Standards (BIORAD) was used. Measurement of loaded total protein is formulated depending

on concentration of sample, used sample buffer and volume of buffer. 7 µl of protein ladder was loaded onto the SDS-PAGE gel. Electrophoresis was accomplished in Tris-glycine running buffer at 90 V, for first 30 minutes and then voltage is increased to 150 V for 3 hours.

When the electrophoresis was finished, the gel was removed and stained with SYPRO Ruby Gel Stain. SYPRO Ruby is a permanent stain composed of ruthenium as part of an organic complex that interacts noncovalently with proteins. Fluorescence emission is approximately 618 nm. Detection sensitivity is about 0.5-1.0 ng protein/mm^2. It is fully compatible with mass spectrometry. Gels were fixed in 10% MeOH and 7% acetic acid for 30 – 60 minutes. Fixation solution was discarded and gel was embedded into 200 ml 1 × SYRPO Ruby stain for 1 hour at room temperature. Then stain was removed and washed with 10% MeOH/7% acetic acid solution for 45 minutes. Finally, gel was illuminated on transilluminating UV box and images of gel was taken.

2.3.6 Protein spot identification and removal

Image files were processed using the Image analysis and quantification. Detected protein spots were then matched between gels and a synthetic master image, which was prepared to represent a majority of the protein spots present in all gels belonging to control samples. Matching rate of the spot analysis was evaluated by comparison of each sample gel with the master gel. Spots corresponding to the same spots in both gels indicate the matching rate. Determination of matching rate maintains the efficacy of the comparison in terms of the variation. Spot identification was followed by image analysis and quantification (relative integrated optical density of protein spots) using PDQuest software. All patient gels were individually compared with control master gel. At least two times differentially expressed protein spots were selected and removed from the gels. Protein spots were embedded into distilled water.

2.3.7 MALDI-TOF mass spectrometry

Matrix assisted laser desorption ionization - time of flight (MALDI-TOF) mass spectrometry (MS) is a technique to identify biomolecules, in this case proteins, depending on their charge and molecular weights (Caprioli et al., 1997). It is basically based on the ionization of amino acids in a suitable matrix and radiation in

a certain wavelengths. Common matrixes used in MALDI-TOF experiments are 2,5-dihydroxy benzoic acid, 3,5-dimethoxy-4-hydroxycinnamic acid, α-Cyano-4-hydroxycinnamic acid (CHCA) and 3-hydroxy picolinic acid that their radiation wavelengths varies as 337, 355 and 266 nm.

In this study, MALDI-TOF analysis was performed in Greifswald University, Microbiology Department (Berlin, Germany). Excised protein spots from SYPRO Ruby (BioRad) stained gels were destained with 200 mL 50% acetonitrile (CAN) in 10 mM NH_4HCO_3 and dried in a SpeedVac concentrator. Protein was digested overnight at 37^{o}C with trypsin (5 mg/mL; Pro- mega Madison, WI, USA) in 50 mM NH_4HCO_3. The resulting peptides were extracted twice with 25 mL 50% ACN/0.1% TFA. The collected extracts were lyophilized and resuspended in 10 mL 0.1% TFA and desalted on ZipTip C18-microcolumns (Millipore, Bedford, MA, USA). Elution was performed with CHCA (5 mg/mL) directly onto the MALDI matrix (2 mL of the solution was applied to a plated sample holder and introduced into the mass spectrometer after drying). Solarix MI (Bruker, England) was used as instrument and the data was acquired using the 4700 Explorer software (Applied Biosystems, Foster City, CA, USA). Ions were accelerated at 20 kV and reflected at 21.3 kV. Spectra were acquired in the delayed extraction, reflectron R mode. Scans (100–300) were averaged to produce final spectra. Spectra were externally calibrated using the monoisotopic MH+ ion from peptide standards. Results obtained in .xml format and top five possible protein candidates were listed. Among them first candidate proteins were selected for identification.

2.3.8 ELISA analysis of candidate protein biomarkers

Enzyme linked immunosorbent assays is a technique for the identification and quantification of target proteins (Engvall and Perlmann, 1971). In this study, ELISA performed for selected 21 proteins. The list of selected proteins were given in table 2.1, which were found to be differentially expressed among MS subtypes and have important implication in pathogenesis or previously associated with MS. We performed validation experiments via indirect ELISA protocol and same antigens were also verified with commercially available ELISA kits for four of them. Original samples from proteome study were used for individual verifications. As commercial ELISA kits, Blue Gene (Shangai Biotech) kits were used according to manufacturers specifications; but for the in house ELISA, Abcam's commercial protocol was

followed (Abcam, 2012). First precipitated CSF proteins and serum samples were diluted to a final concentration of 20 µg/ml in carbonate buffer (100 mM), and wells of PVC 96 well microtiter plate were coated with this antigen by pipetting into 50 µl of the antigen dilution. Plate was covered with adhesive plastic and incubated at 4^{o}C for overnight. Coating buffer was removed by aspirating and plate was washed with 200 µl PBS solution for three times. Then, antigen coated wells were blocked with blocking 200 µl buffer (5% non fat dry milk in PBS) for 2 hours at room temperature. Later, blocking buffer was removed and plate was washed for three times. 100 µl of diluted primary antibody solution was added to wells and incubated at 4^{o}C for overnight. Solution was removed by aspirating and wells were washed again for three times. 100 µl of conjugated secondary antibody solution was applied into wells and incubated at 37^{o}C for 1 hour. The last wash step was performed before detection with horse radish peroxidase enzyme and hydrogen peroxide (H_2O_2). 100 µl of hydrogen peroxide solution was dispensed and for the color development, plate was incubated for 15 minutes at room temperature and dark room. After sufficient colour development 100 µl of stop solution (0,1M NaOH) was applied to the wells and then absorbance was read in microplate reader (BioRad).

Table 2.1: Selected candidate proteins for confirmation studies.

Protein Symbol	Protein Name	Antibody
APOE	Apolipoprotein E	Abcam (ab20261)
APOA4	Apolipoprotein A-IV	Santa Cruz Biotechnology Inc. apoA-IV (H-240): sc-50375
PDG2S	Prostaglandin D2 Synthase	Santa Cruz Biotechnology Inc., PDG2S (F-24) sc-133904
TTR	Transthyretin	Santa Cruz Biotechnology Inc. TTR (FL-147) sc-13098
A2MG	Alpha-2 Macroglobulin	Abcam [2D9] (ab36995)
VDBP	Vitamin D Binding Protein	Abcam (65636)
A1AT	Alpha - 1 anti-Trypsin	Abcam (ab17263)
AGP	Alpha 1-acid glycoprotein	Abcam (ab58291)
GFAP	Glial fibrillary acidic protein	Santa Cruz Biotechnology Inc. sc-33673
MBP	Myelin basic protein	Santa Cruz Biotechnology Inc. sc-25665
NFL	Neurofilament light chain	Santa Cruz Biotechnology Inc. sc-25652
Tau	Tau protein	Santa Cruz Biotechnology Inc. sc-32274

Table 2.1 (Continued): Selected candidate proteins for confirmation studies.

MOG	Myelin oligodendrocyte glycoprotein	Santa Cruz Biotechnology Inc. sc-66968
PI3P	Phosphatidylinositol 3-phosphate	Santa Cruz Biotechnology Inc. sc-221817
SGK	Serine/threonine-protein kinase	Santa Cruz Biotechnology Inc. (c-20) sc-15885
IR	Insulin receptor protein	Santa Cruz Biotechnology Inc. (L-20) sc-19851
NEDD4	E3 ubiquitin-protein ligase	Santa Cruz Biotechnology Inc.
Notch	Notch protein	Santa Cruz Biotechnology Inc. sc-6014
NCT	Nicastrin	Santa Cruz Biotechnology Inc. sc-14369
NUMBL	Numb like protein	Santa Cruz Biotechnology Inc. sc-54525
PEN2	Presenilin enhancer 2	Santa Cruz Biotechnology Inc. sc-30232

2.3.9 Western blot analysis of selected protein biomarkers

In western blotting, the protein samples separated on SDS-PAGE according to their molecular weight and they were transferred to a nitrocellulose membrane and analyzed for presence of target proteins using antibodies (Blake et al., 1984). In this study same antibodies used in ELISA studies, were also used for western blotting.

Frozen samples dissolved on ice. 16 µL of each sample was mixed with 4 µL of 4X loading dye and mixture were vortexed gently. The tubes were heated in 95^{o}C heat block for 5 minutes, and transferred onto ice immediately to prevent protein renaturation. 20 µL of each sample and 5 µL of marker were loaded into SDS-PAGE gel (12% separating gel, 5% stacking gel). Protein separation performed by electrophoresis for 2 hours at 110V. Proteins were transferred onto nitrocellulose membranes from the gels by a semi-wet electro-blotting system. Transfer was performed for 1 hour at 1A. Membranes were transferred into falcon tubes and washed with 1x Phosphate-Buffered Saline with Tween20 (PBST) for 5 minutes at room temperature twice and they were incubated with blocking solution (5% non-fat dry milk) for 3 hours at room temperature. Membranes were washed with PBST for 5 minutes at room temperature twice. To detect, antibody solutions were prepared from primary antibodies in a 1:500 ratio using 5% non-fat dry milk as diluent. Then, membranes were washed with PBST for 5 minutes at room temperature twice. As secondary antibody (Santa Cruz Biotechnology Inc. goat anti-rabbit, anti-mouse and

anti-goat IgG-HRP: sc-2030) was diluted in 5% non-fat dry milk to 1:6000 ratio. Membranes incubated with secondary antibody solutions for 3 hours at room temperature and then washed with PBST for 5 minutes at room temperature twice. Membranes incubated with Bio-Rad Amplification Reagent Solution (Bio-Rad Western Blot Amplification Module) for 10 minutes at room temperature. Membranes were washed with 20% Dimethyl Sulfoxide (DMSO) solution for 5 minutes at room temperature and washing step repeated three more times for a total of four washes. Later, membranes were washed with PBST for 5 minutes at room temperature twice and incubated with 1:10000 diluted Streptavidin-HRP (Bio-Rad Opti-4CN Substrate Kit) for 30 minutes at room temperature. Streptavidin-HRP was diluted in 5% non-fat dry milk. Falcon tubes were covered with aluminum folio to avoid light. Membranes were washed with PBST for 5 minutes at room temperature twice. Finally, 18 mL distilled water, 2 mL Opti-4CN diluent, and 400 µL Opti-4CN substrate (Bio-Rad Opti-4CN Substrate Kit) were mixed and transferred into tubes. Tubes were covered with aluminum folio and incubated for 15 minutes at room temperature. After the incubation was over, membrane was washed with dH_2O and let to dry. Quantitative assessment of the protein bands on membranes after photoluminescence was performed using PDQuest Analysis Software.

2.3.10 Statistical analysis of Proteome Studies

Statistical analyses of ELISA and western blot studies have been performed using One-Way Analysis of Variance (One-way ANOVA) and nonparametric Student's t tests. The purpose of One-way ANOVA is to find out whether data from several groups have a common mean. Student t test is a statistical hypothesis test in which the test statistic has probability distribution, which was applied to find out whether protein levels were different between groups. Graphpad prism v2 program was used for the statistical analyses.

2.4 Pathway Analysis Using Bioinformatics Approaches

Biological data production has been increasing at an unprecedented pace with the advancements of high throughput biological data production in genomics, transcriptomics and proteomics. Bioinformatics approaches were used in this study for two main purposes; one was to handle many data points, elucidate uninformative

protein results and reveal most prominent ones to draw statistically meaningful results, second was to reveal novel interacting pathways of the identified proteins important in MS using a systematic approach.

2.4.1 PANOGA tool and KEGG pathway analysis

For each patient in each MS group (CIS, RRMS, PMS) mean protein fold change values were calculated. T-test statistics is applied for each group to identify significantly changing proteins (p value obtained) in that group compared to two different control groups. The lists of significant genes along with their p values are given as input to PANOGA (Bakir-Gungor et al., 2014) to identify pathways with proteins that are significantly altered for each MS group.

PANOGA first searches out active sub-networks containing most of the disease affected proteins in the human PPI network. We used Goh et al.'s human PPI network in this study (Goh et al., 2007). jActive Modules algorithm (Ideker et al., 2002) is employed to identify the sub-networks taking into account the P-values of each gene with the network topology to extract potentially meaningful active sub-networks that overlaps at most 50% with each other.

The next step following the identification of sub-networks is to evaluate whether these sub-networks are biologically meaningful. For each sub-network, PANOGA computes the number of the genes in an identified sub-network that are also found in a specific human biochemical pathway, compared to the overall number of genes described for that pathway. In this functional enrichment step, PANOGA uses a two-sided increase or decrease test based on the hypergeometric distribution to examine the association between MS-related genes and the genes in each KEGG pathway. To correct the p-values for multiple testing, the Bonferroni correction procedure was applied on the p-values of each identified pathway. If a KEGG pathway is determined statistically significant for at least one of the active sub-networks, PANOGA adds this pathway into our final list of significant pathways associated with disease. If a pathway appears in more than one subnetwork analysis only the most significant one is reported (Bakir-Gungor and Sezerman, 2012).

2.4.3 Statistical comparison of molecular pathway analysis

Statistical significance of the pathways in relation to MS groups is tested by shuffling the MS groups randomly 10 times and selecting CIS, PPMS and RRMS patients from each shuffled patient data. Then PANOGA is employed for each random MS group to identify significant pathway. The significance was considered as $p<0.0001$ for pathway analysis.

2.5 Genomic Level Studies

Genome analysis for MS is performed in familial cases. 28 multiply affected MS families were collected for the study. Among them 10 informative families including 36 samples (18 patient and 17 healthy subjects) were used. Clinical subtypes of patients were 6 CIS, 9 RRMS and 3 PPMS. Healthy subject were used as controls for linkage analysis. Family trees indicating the diseased and healthy individuals were shown in Figure 2.2. For the GWAS analysis, a different (non-family member cohort) control samples were used. Control samples were obtained from Done Genetics Co., which were enrolled as control samples to another study investigating the cardiovascular diseases. All permissions and confirmation of study inclusion documents were obtained from healthy subjects. One patient sample was discarded from linkage analysis due to uncertain familial information. For the GWAS analysis, 1 of MS patient belonging to same family discarded due to homogenization of GWAS analysis. As a result, unrelated 11 MS patients and 60 unrelated healthy subjects were included in the GWAS analysis.

2.5.1 DNA isolation from blood samples

DNA from 18 MS and 17 healthy subjects were isolated by using Roche Blood DNA isolation kit. One ml of blood was mixed with 3 ml of red blood cell lysis buffer in a sterile 50 ml centrifuge tube and vortexed. Tubes were inverted for 10 minutes, and centrifuged at 875 x g for 10 minutes. Supernatants were discarded, and white cell containing pellets were thoroughly resuspended in residual supernatant at the bottom of the tube by vortexing. Last three steps were repeated to obtain completely red blood cell-free pellets. 0.5 ml of white cell lysis buffer was added onto each pellet and tubes were vortexed thoroughly. Tubes were incubated at $37^{o}C$ and vortexed in every 10 minutes until the solutions contained no undissolved particles. 260 µl of

Protein Precipitation Solution was added into each tube and tubes were vortexed for 25 seconds. Tubes were centrifuged at 12.000 x g for 10 minutes. DNA containing supernatants were poured into 50 ml centrifuge tubes containing 5 ml of absolute ethanol. DNA samples were collected and transferred into 1.5 ml ependorf tubes containing 200 ml of 70% ethanol. Tubes were centrifuged at 875 x g for 5 minutes, and supernatants were discarded. DNA pellets were allowed to air dry, and resuspended in 100 µl of TE Buffer by incubating at 65^{o}C for 10 minutes and vortexing in every 2-3 minutes. Concentration and purity of isolated DNA samples were measured by Nanodrop at 260/280 nanometers (nm), and samples were stored in 4^{o}C until use.

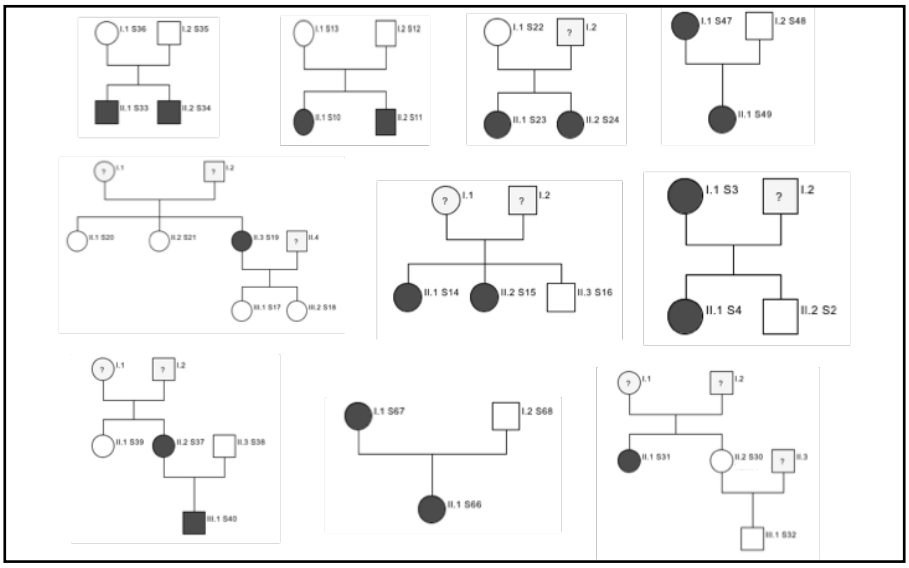

Figure 2.2: Family trees used for the linkage and GWAS studies.

2.5.2 Genome-wide association study using familial MS cases

Genome wide association studies (GWAS) compare the differences between case and control groups with a high coverage that is representative of the whole genome in order to identify genetic variations by searching for related single nucleotide polymorphsims (SNPs) (Wang et al., 2005). Main goal is to associate the variation profiles with a disease or a phenotype along with finding the key genetic variations. To this aim, we performed a microarray study comprising the samples from multiply affected MS families. Illumuni Human CytoSNP-12 v2.1 bead chip was used. CytoSNP chip, can also be used for linkage analysis, therefore,

Sample preparation protocol includes many steps like PCR free Whole genome amplification, DNA fragmentation and selective hybridization. All protocol

continues for three days and there are more than 200 steps (Silberstein et al., 2013). Finally chips were scanned on Illumina iScan array scanner. Bioinformatics analyses of the results were performed manually. First filtering and classifying of the genotype data was accomplished by following the chip protocol based on call rate and minimum hybridized probe of 90%. Raw genotyping data was processed in GenomeStudio v.11.1 software and project file was created. Quality controls of the samples were assessed by the 'call rate' parameter, which indicates the efficiency of binding of each probe to the target DNA fragments. The threshold of call rate was 95%, but none of the samples discarded due to quality control issues. Project file was transferred to Golden Helix® SNP & Variation Suite (SVS) program for GWAS analysis. Then, at first step relativity degrees of samples were determined. "Identity by Descent - IBD- Estimation" was applied and estimated PI values above 0.2 were discarded from the analysis. SNP markers on Y chromosome were discarded from the analysis, 296106 SNPs remained. Genotyping rate below 95% (11638 SNPs) and minor allele frequency (MAF) rates below 0.01 (36136 SNPs) were also filtered from the data. Fischer's exact test for the Hardy Weinberg Equilibrium (HWE) is applied. Furthermore, the effect of linkage disequilibrium was calculated and SNPs with LD value above 0.5 were filtered. As a results, among 129547 SNP markers, comparison of genotypes between healthy and patients was accomplished via Additive model, "(dd)>(Dd)>(DD)" and "Correlation/Trend" test was applied. For the multiple correction analysis, Bonferroni adjustment and False Discovery Rate (FDR) tests were applied. Finally p values below $1*10^{-4}$ were listed as significantly differed SNPs.

2.5.3 Linkage analysis

Genetic linkage studies have high success rates in identification of single gene disorders. However conducting linkage studies for complex diseases like MS, can only point candidate disease susceptibility genes.

SNP genotyping on the Illumina CytoSNP 300K array was performed for genomes of 36 individuals from the included families according to the manufacturer's protocol. Approximately, 300.000 SNP markers were genotyped in each individual. For data mining and filtering, SNPs on chromosome Y, SNPs with call rates lower than 95% and minor allele frequency (MAF) lower than 0.01, and in strong linkage disequilibrium ($r^2>0.5$) were excluded from the data. For the linkage analysis

multipoint non-parametric linkage (NPL) analysis was performed by using SimWalk tool (Sobel and Lange, 1996). Filtering and elucidating of uninformative SNPs resulted with 245.008 SNPs. NPL results were calculated for each of 3118 informative SNP markers spaced at an average of 1 cM intervals, and expressed as NPL Z scores. For the regions with NPL score of 1.7 and higher fine mapping analysis was performed for each of 639 and 831 informative SNP markers spaced at every 0.2 cM for chromosome 13 and 21, respectively. NPL scores for each region were analyzed, and linkage suggestive chromosomal positions were determined. Analyses were performed using easyLINKAGE software v5.08 (Hoffmann and Lindner, 2005). From the resulting loci, protein coding genes were listed and further filtering was done based on MS and other inflammatory/neurological conditions literature using MalaCards Human Disease Database and PubMed, unrelated genes were excluded. Among the selected genes, determination of possible functional relevance with MS was assumed by considering biological processes and pathways in which each corresponding protein encoded by each gene involves (GeneCards The Human Gene Compendium, KEGG Pathway Database).

2.5.4 Comparison of proteomics, linkage and GWAS results

The correlation of linkage, GWAS and proteomics results would facilitate our understanding of MS pathogenesis and reveals the candidate potential biomarkers. To this aim, protein coding genes located on the suggestive linkage regions, genes carrying the SNPs with significant and suggestive associations in the GWAS, and genes encode for the proteins that found to be differentially expressed in MS subtypes were listed and compared within each other. Each protein having altered expression level in each patient was individually compared with the linkage and the GWAS results, and overlapping genes were listed. Additionally, pathways in which proteins encoded by each gene revealed in the genetic analyses involve were noted (KEGG Pathway Database, GeneCards The Human Gene Compendium), and genes encoding for proteins that involve in the affected pathways implicated in the proteomic analyses were determined. The best-correlated genes and proteins were determined for further analyses.

3. RESULTS

3.1 Sample Characteristics Of Study Groups

A total of 260 CSF and serum samples of differentially classified MS patients were collected. Among them 179 MS patients including, 65 patients with CIS (30 MSc-CIS and 30 SA-MS), 72 with RRMS, 42 with PMS and 42 control samples (20 ONDs and 22 nonMS control) were recruited to study. Demographic data of patients and control subjects are summarized in table 3.1. The female to male ratio was 1.8:1 (114:65) in subject group and 2.5:1 (30:12) in control groups. No significant association was observed among age of onset of the first neurological event, duration of the disease, oligoclonal band formation. However, as expected, expanded disability status score (EDSS) and CSF IgG index scores differ significantly among different disease subtypes (ANOVA, $p<0.001$).

Table 3.1: Clinical information of case and control samples.

	Age (Year)	Age at Onset	Gender (F:M)	Disease Duration (Months)	EDSS	Follow up period (Months)	CSF IgG Index
CIS (65)	33.4±6.2	30.2±6.0	43:22	32.3±9.2	1.1±0.3	31.7±5.8	1.1±0.4
MSs-CIS (35)	31.8±6.4	28.6±5.8	20:12	30.81 ± 9.4	1.2±0.4	30.1±5.2	1.1±0.3
SA-MS (30)	35.3±7.1	32.2±6.3	23:10	33.60 ± 8.1	1.1±0.1	33.5±6.2	1.2±0.6
RRMS (72)	34.5±6.3	28.9±5.2	46:26:00	52.4 ± 9.2	3.3±1.2	46.4±8.4	1.2±0.7
PMS (42)	44.2±8.1	33.1±7.5	25:17:00	70.6 ± 11.4	6.2±2.7	60.1±9.5	1.5±0.9
NonMS Control (22)	32.7±3.3	-	16:06	-	-	-	-
OND (20)	38.3±4.7	-	14:06	-	-	-	-

Age at analysis, age of disease onset, gender, expanded disability status scale (EDSS), disease duration which indicates the time between first clinical attack and last examination, follow-up period which indicates the time between first and last examination and finally CSF IgG index, which is calculated as IgG to CSF albumin ratio compared to the serum IgG to serum albumin ratio, were summarized. NonMS control samples consisted of acute headache and pseudotumour cerebri patients where as, other neurological disease (OND) consisted of neuro-Behçet disease, neuro-sarcoidosis and polyneuropathies.

3.2 Optimization Of Proteome Studies

Proteomic methods are complex and multi-parametric studies including many steps and parameters that affect the final results. Therefore, we needed to optimize many parameters before studying patient samples. CSF is a promising source of biomarkers however CSF proteins are generally less abundant than their corresponding serum counterparts. Therefore, optimization results of this study revealed a comprehensive know-how about further CSF proteome studies. Here, we present the results of different parameter change consequences.

Firstly, use of narrow pH range IPG strips resulted increased yield of protein spots on 2D-PAGE gels. Comparison of IPG strips with pH 3-6, 4-7, 5,8 and 3-10 resulted with varying number of spots (Figure 3.1 and 3.2). The mean spot counts of IPG strips were evaluated for five 2D-PAGE runs with same sample: 323 spots for pH 3-10, 186 spots for pH 3-6, 251 spots for pH 4-7 and 177 spots for pH 5-8 (Table 3.2). The differences among different IPG strips were significant ($p<0.001$). Matched spot counts and matching rate values were also maximum for pH 3-10 strip among all IPG strips.

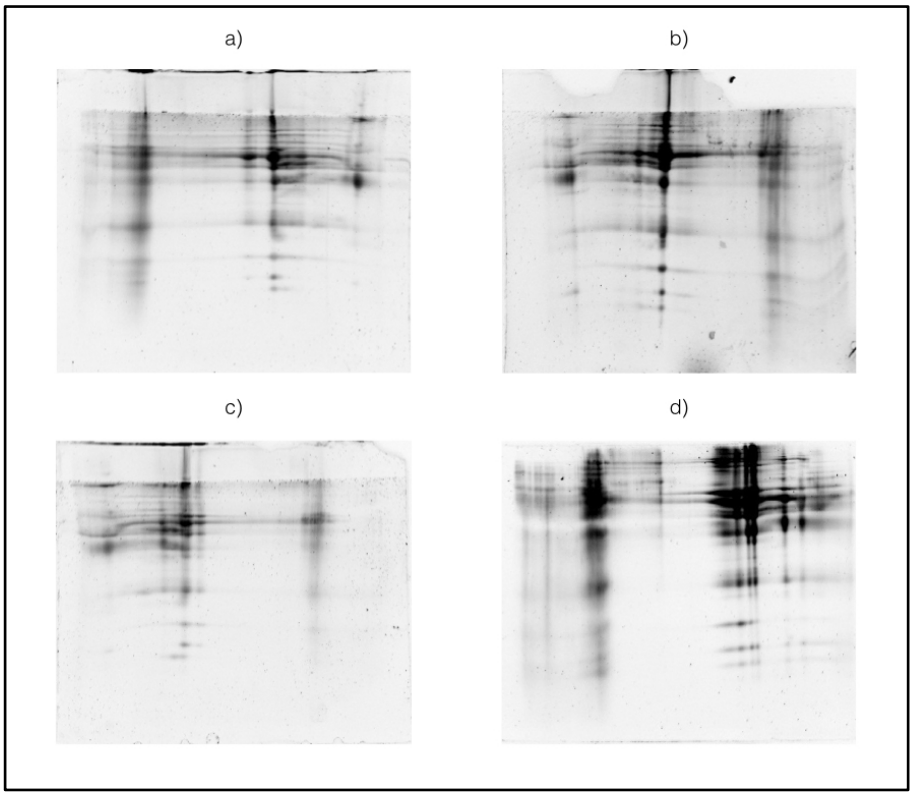

Figure 3.1: Proteins spot optimization with different IPG strips. a) pH: 3-6, b) pH: 4-7, c) pH:5-8, d) pH:3-10.

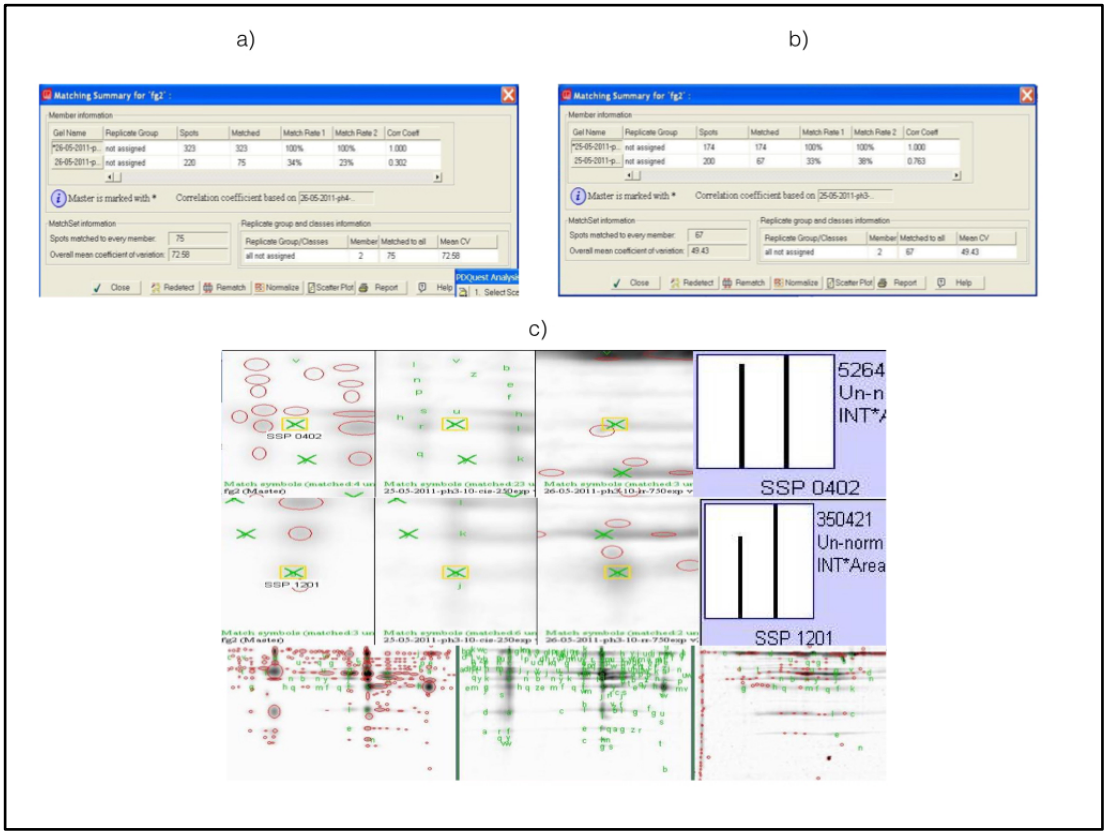

Figure 3.2: Protein spot counts and matching rate between different IPG strips a) pH:3-10 and pH:3-6, b) pH: 4-7 and pH:5-8, c) Example of protein spot matching between different gels.

Concentration of SDS-PAGE in the second dimension of 2D PAGE study is critical factor to get better protein spot separation. 10%, 12% and 15% gel concentrations were compared. Among different electrophoresis durations (2h, 3h and 4h), it was observed that 3h in 150 V is the best electrophoresis time and 15% gel is the best concentration for better separation.

Table 3.2: Spot optimization results.

IPG Strips	Counted Spot Number	Matched Spot Number	Matching Rate (%)
pH 3-6	251 ± 41	163 ± 12	56
pH 4-7	186 ± 37	112 ± 18	52
pH 5-8	177 ± 32	120 ± 11	58
pH 3-10*	323 ± 44	220 ± 24	62

- pH 3 – 10 was the optimum IPG strip due to its counted spot number. Matching spot number refers to matching of same sample between different IPG strips and matching rate is the average of all IPG strip results.

Salt ingredient of CSF is another important issue to that presence of high salt concentration causes improper protein separation in 2D-PAGE. Therefore, we compared to protein precipitation method in order to remove salt from the CSF. TCA/Acetone precipitation resulted with better protein separation compared to acetone precipitation alone. Protein spots were clearer and spot masking protein bulk in the middle of the gels were less in TCA/Acetone precipitation method (Figure 3.3).

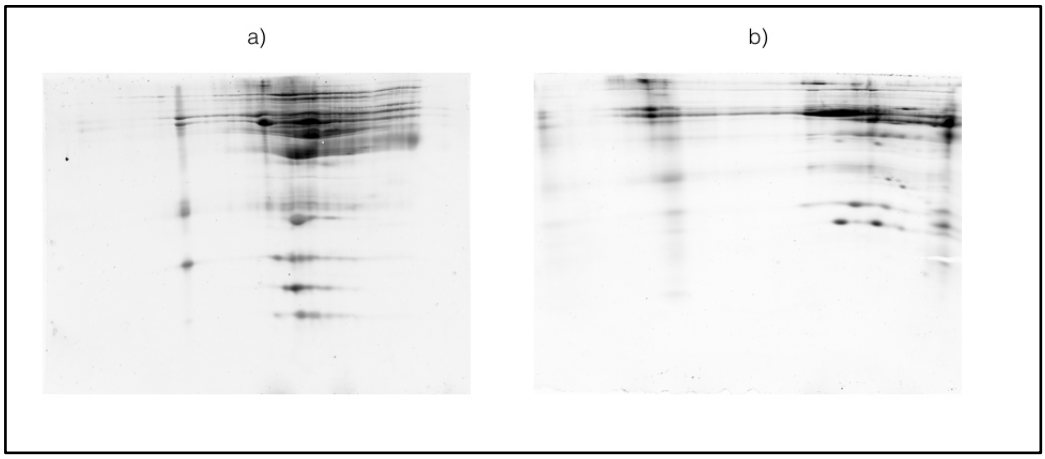

Figure 3.3: 2D PAGE images of same sample precipitated with different methods. a) Acetone precipitation b) TCA/Acetone precipitation.

As a result, 2D-PAGE optimization studies before sample analysis is important and needed to be standardized. Our optimization studies indicated that IPG strip with pH3-10, 15% SDS PAGE gel concentration for 3h at 150 V and TCA/Acetone protein precipitation methods were most probable conditions for better CSF

proteome results. All patient and control samples were studies performed based on these optimizations.

3.3 Differentially Expressed Protein Biomarkers In Multiple Sclerosis Subtypes

Differentially expressed proteins spots were selected and MALDI-TOF analysis performed for all individually obtained protein spots. Results of mass spectrometry analysis revealed an average of 982 + 34 protein spots in the CSF samples from MS and control patients. Analyses were performed based on the 2-4-6 and 8 fold differences between individual patient samples and control master gel. 382 protein spots with more than 2-fold difference in expression levels were listed for all of subtypes compared to controls. A two-fold difference for the comparison was selected based on the optimum spot difference and previous literature findings (Alban et al., 2003). Spot abundance variability is calculated by coefficient of variation (CV) value. Average CV of matching between all MS and control samples was about 24.21%, 23% CV between all control samples, 26% for CIS samples, 22% for PMS samples and about 22% for RRMS samples. The difference between matching spots was assessed by the independent samples t-test ($P<0.05$). Differentially expressed selected spots were searched in Swiss-Prot database and showed 151 distinct proteins. Subtype specific biomarkers and common biomarkers differentially expressed in all subtypes compared to controls were listed in Table 3.3, 3.4 and 3.5. Representative gel images were given in Figure 3.4.

Beside the subtype specific abundant proteins, regarding all individual CSF samples of patients, comparison between total control group and disease subtypes revealed 61 different proteins in RRMS, 58 different protein in CIS and 60 proteins in PMS subtypes that are at least 2 fold differentially expressed.

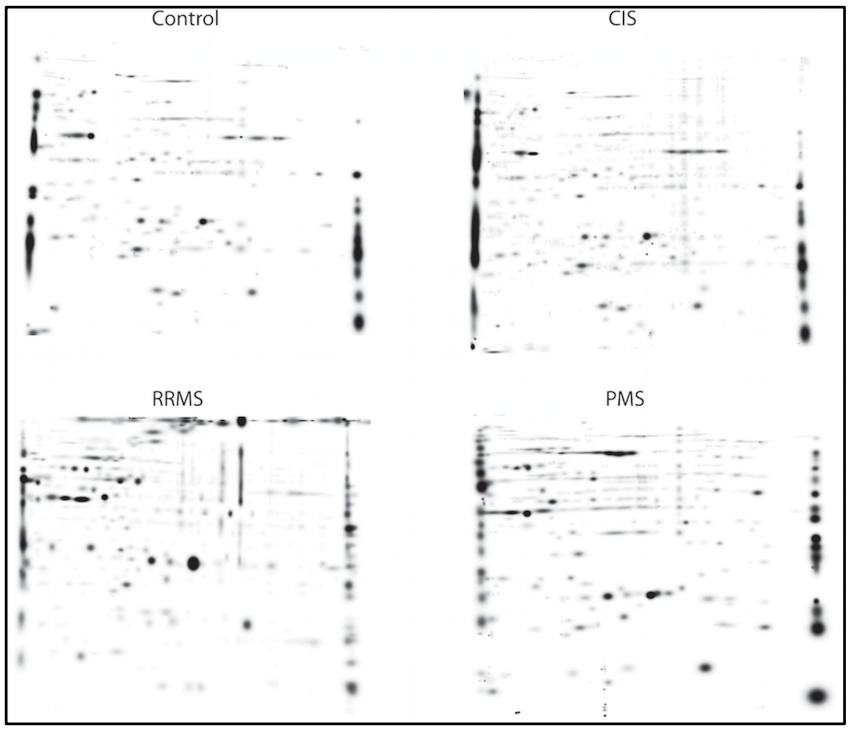

Figure 3.4: Representative 2D gel images for all MS subtypes.

Table 3.3: Differentially expressed proteins in RRMS subtype.

RRMS vs Controls		
Protein Name	**Accession number**	**Change**
Splicing factor, arginine/serine-rich 14, isoform CRA_f	gi|119605178	(+) 2.10
Chain A, Tertiary Structures Of Three Amyloidogenic Transthyretin Variants	gi|2098257	(+) 3.10
Hypothetical protein	gi|30268247	(+) 1.81
Unnamed protein product	gi|63089882	(+) 2.14
Unnamed protein product	gi|215415636	(+) 1.92
Chain B, Molecular Basis For Complement Recognition And Inhibition Determined By Crystallographic S	gi|305677899	(+) 2.63
Chain F, Crystal Structure Of Complement C3b In Complex With Factor B	gi|317455070	(+) 2.20
apolipoprotein A-IV precursor	gi|178779	(+) 2.32
cyclin G associated kinase, isoform CRA_b	gi|119603051	(-) 2.44
unnamed protein product	gi|194375550	(+) 2.13
Chain A, Dna Polymerase Beta With A Gapped Dnd Substrate And Dtmp (Cf2)	gi|229597882	(+) 1.91
MAP9 protein	gi|116283343	(+) 2.24
Chain C, Complex Of Ttr And Rbp4 And Oleic Acid	gi|305677617	(+) 2.13
hCG2038441	gi|119584862	(+) 2.62
Chain A, Solution Structure Of Double Super Helix Model	gi|293651900	(+) 2.55
immunoglobulin light chain	gi|149673891	(+) 2.76
Chain M, Crystal Structure Of A Broadly Neutralizing Anti-Hiv-1 Antibody In Complex With A Peptide	gi|48425065	(+) 1.27
Chain A, Germline V-Genes Sculpt The Binding Site Of A	gi|258588258	(+) 2.21

Table 3.3 (Continued): Differentially expressed proteins in RRMS subtype.

Chain A, Crystal Structure Of Human Alpha 1 Acid Glycoprotein	gi|215794600	(+) 2.33
Albumin, isoform CRA_s	gi|119626082	(+) 2.41
Chain A, Structure Of Human Serum Albumin With S-Naproxen And The Ga Module	gi|168988718	(+) 2.57
Unnamed protein product	gi|194391112	(+) 2.29
albumin [Peptide Partial , 25 aa]	gi|247527	(+) 2.33
methionine sulfoxide reductase A2	gi|62868459	(+) 3.44
Chain A, Crystal Structure Of Human Serum Albumin Complexed With Myristic Acid And Lysophosphatidyl	gi|228311905	(+) 2.25
Albumin, isoform CRA_h	gi|119626071	(+) 4.20
Alpha 2 macroglobulin	gi|46812315	(+) 2.77
DHX57 protein	gi|40850932	(+) 1.78
Unnamed protein product	gi|300621695	(+) 1.39
Albumin-like	gi|763431	(+) 3.30
PRO2619	gi|11493459	(+) 1.81
Titin-cap (telethonin)	gi|15426456	(+) 2.34
Unnamed protein product	gi|1340170	(+) 2.33
Parafibromin	gi|40018640	(+) 1.98
TTLL5 protein	gi|118142839	(+) 2.55
hCG1776873	gi|119620109	(+) 2.26
RecName: Full=Serotransferrin; Short=Transferrin	gi|313104271	(+) 2.77
Inter-alpha-trypsin inhibitor heavy chain H1	tr|A6NBZ8|	(+) 2.18
Cyclic AMP-dependent transcription factor ATF-7	sp|Q9UI66	(+) 2.59
Protein POF1B	gi|158258947	(+) 2.50
Signal recognition particle 14 kDa protein	sp|P16840|	(+) 1.77
Cytotoxic T-lymphocyte protein 4	tr|Q569I7	(+) 2.42
Fibroblast growth factor receptor substrate 3	tr|B7WNR0|	(+) 2.30
Human androgen-induced prostate proliferative shutoff associated protein (AS3)	gi|21758324	(+) 2.14
Nucleolin	gi55956788	(-) 2.15
Vit.D binding	gi32483410	(-) 3.16
Kallikrein 6	gi21466131	(-) 2.47
Contactin	gi414791	(+) 1.98
putative myosin 15B	gi150384690	(-) 1.79
amyloid beta A4 protein	gi41406057	(+) 1.80
Immunoglobulin kappa variable (IGKV1)	gi21669539	(+) 2.51
Immunoglobulin kappa constant (IGKC)	gi125145	(+) 2.82
Clusterin	gi4502905	(+) 2.32
Angiotensinogen	gi4557287	(+) 2.54
Apolipoprotein AI	gi4557321	(+) 2.75
ceruloplasmin	gi4557485	(+) 1.56
Fibulin 1 isoform D	gi5922009	(+) 2.37
Fibulin 1 isoform C	gi4503663	(+) 2.28
Complement C4d	gi223961	(+) 2.29
Fetuin A	gi156523970	(+) 2.50
Alpha-1 antichymotrypsin	gi112874	(+) 2.71

Table 3.4: Differentially expressed proteins in CIS subtype.

CIS vs Control		
Protein Name	**Accession number**	**Change**
Plectin 2	gi|40849932	(-) 1.92
Chain A, Alpha1-Antitrypsin	gi|157831596	(-) 2.21
unnamed protein product	gi|194391112	(+) 1.66
anti-pneumococcal antibody 1.2 light chain variable region	gi|7673381	(+) 2.25
Chain L, Crystal Structure Of Hiv-Gp120 Core In Complex With Cd4- Binding Site Antibody B13	gi|268612336	(+) 2.51
Titin-cap (telethonin)	gi|15426456	(+) 2.73
unnamed protein product	gi|194374381	(+) 2.20
TCRDV2	gi|2181880	(+) 2.42
HP protein	gi|47124562	(+) 1.97
Steerin2 protein	gi|27526809	(+) 2.33
Chain D, Structure Of Il-13 Antibody H2l6, A Humanized Variant Of C836	gi|311772086	(+) 2.91
Truncated telethonin	gi|161137716	(+) 2.24
Chain B, A Covalent Dimer Of Transthyretin That Affects The Amyloid Pathway	gi|55669576	(+) 2.23
Chain A, Human Serum Albumin Complexed With Myristate And Aspirin	gi|122920512	(+) 2.57
Chain A, Crystal Structure Of Human Serum Albumin Complexed With Myristic Acid And Lysophosphatidyl	gi|228311905	(+) 2.29
unnamed protein product	gi|218650345	(+) 2.33
Ig gamma 2 H chain, BUR [Peptide, 348 aa]	gi|6137432	(-) 3.44
alloalbumin Venezia	gi|178345	(+) 2.21
Chain A, 2.0 Angstrom Structure Of Intact Alpha-1-Antitrypsin: A Canonical Template For Active Serp	gi|6137432	(-) 2.33
unnamed protein product	gi|40982242	(+) 4.20
hypothetical protein	gi|60219224	(+) 2.77
methylmalonyl Coenzyme A mutase precursor variant	gi|62087270	(+) 1.78
hCG2024404, isoform CRA_c	gi|119570052	(+) 2.33
Chain A, Human Serum Albumin Complexed With Myristate And Aspirin	gi|122920512	(+) 3.44
albumin, isoform CRA_h	gi|119626071	(+) 2.25
acylphosphatase 2, muscle type, isoform CRA_a	gi|119620560	(+) 4.20
prostaglandin D2 synthase 21kDa (brain)	gi|55962673	(+) 2.51
Chain A, Structure Of Human Serum Albumin With S-Naproxen And The Ga Module	gi|168988718	(+) 1.96
Albumin, isoform CRA_m	gi|119626076	(+) 1.99
Orosomucoid 1	gi|219519923	(+) 3.35
Apolipoprotein A-IV precursor	gi|178779	(+) 2.81
RecName: Full=Serotransferrin; Short=Transferrin; AltName: Full=Beta-1 metal-binding globulin; AltN	gi|313104271	(+) 2.34
Homo sapiens WD repeat domain 59	gi|193787129	(+) 2.33

Table 3.4 (Continued): Differentially expressed proteins in CIS subtype.

Integrin alpha-IIb	tr|B7WPB5|	(+) 1.98
Cytotoxic T-lymphocyte protein 4	tr|Q569I7|	(+) 2.55
Calcium/calmodulin-dependent protein kinase type II subunit beta	gi|194391112	(+) 2.26
Dyslexia-associated protein KIAA0319-like protein	gi|158259857	(+) 2.77
Coiled-coil domain-containing protein 171	sp|Q6TFL3|CI093	(+) 2.18
Insulin-like growth factor-binding protein 3	sp|P20570|	(+) 2.59
Rho-related GTP-binding protein Rho6	tr|Q8TCD0|	(+) 2.50
Vesicle-associated membrane protein 2	tr|Q7L876|	(+) 1.77
Cdc42-interacting protein 4	gi|28590	(+) 2.42
Cyclic AMP-dependent transcription factor ATF-7	sp|Q9UI66|CI038	(+) 2.37
C-C motif chemokine 24	tr|A6NH40	(+) 2.28
Platelet glycoprotein 4	tr|A6NKU1	(+) 2.15
CUGBP Elav-like family member 2	tr|Q6N095	(+) 1.16
P3 protein	tr|Q6N093	(+) 2.17
Component C9	gi4502511	(-) 1.94
Serotransferrin	gi4557871	(+) 1.79
Contactin	gi414791	(+) 1.80
Ceruloplasmin	gi4557485	(+) 2.51
Angiotensinogen	gi4557287	(+) 2.82
Leucine rich alpha 2 glycoprotein	gi16418467	(-) 2.23
Ig kappa chain VLJ	gi21669449	(+) 2.54
Igl chain VLJ	gi21669631	(+) 2.55
Phosphatidylethanolamine binding protein	gi406290	(+) 1.56
Plasma retinol binding protein	gi62298174	(+) 2.30
Haptoglobulin	gi3337390	(+) 2.14

Table 3.5: Differentially expressed proteins in PPMS subtype.

PPMS vs Control		
Protein Name	**Accession number**	**Change**
Human androgen-induced prostate proliferative shutoff associated protein (AS3)	gi|21758324	(+) 2.73
Chain A, Solution Structure Of Double Super Helix Model	gi|293651900	(+) 1.62
unnamed protein product	gi|215415638	(+) 1.76
apolipoprotein A-IV precursor	gi|178779	(+) 2.85
TMEM95 protein	gi|76825355	(+) 2.21
HP protein	gi|47124562	(+) 1.73
Chain A, Zn-Alpha-2-Glycoprotein; Baculo-Zag No Peg, No Glycerol	gi|58176767	(+) 2.81
hCG2011852	gi|119629468	(+) 2.23
immunoglobulin light chain	gi|149673891	(+) 2.57
Chain D, Structure Of Il-13 Antibody H2l6, A Humanized Variant Of C836	gi|311772086	(+) 2.29
Chain A, 2.0 Angstrom Structure Of Intact Alpha-1-Antitrypsin: A Canonical Template For Active Serpin	gi|6137432	(+) 2.33
Chain A, Alpha1-Antitrypsin	gi|157831596	(+) 3.44

Table 3.5 (Continued): Differentially expressed proteins in PPMS subtype.

Chain A, Crystal Structure Of Human Serum Albumin Complexed With Myristic Acid And Lysophosphatidyl	gi|228311905	(+) 2.03
unnamed protein product	gi|218650345	(+) 1.97
unnamed protein product	gi|40981074	(+) 2.09
Chain A, Apolipoprotein E2 (Apoe2, D154a Mutation)	gi|157832106	(+) 2.63
interferon-induced protein with tetratricopeptide repeats	gi|72534658	(+) 2.44
Ig gamma 2 H chain, BUR [Peptide , 348 aa]	gi|243169	(+) 3.21
Ig G1 H Nie	gi|229601	(+) 2.33
unnamed protein product	gi|284388777	(-) 4.20
RecName: Full=Serotransferrin; Short=Transferrin; AltName: Full=Beta-1 metal-binding 47nhibitö; AltN	gi|313104271	(+) 2.77
47nhibit, isoform CRA_h	gi|119626071	(-) 1.78
Alloalbumin Venezia	gi|178345	(+) 2.33
Armadillo repeat containing 3, isoform CRA_b	gi|119606544	(+) 2.44
Chain A, Human Transferrin N-Lobe H249q	gi|16975242	(+) 2.25
alloalbumin Venezia	gi|178345	(+) 2.20
Unnamed protein product	gi|1335218	(+) 2.51
Chain A, Human Serum Albumin In A Complex With Myristic Acid And Tri- Iodobenzoic Acid	gi|157830361	(+) 1.96
Albumin [human, Peptide Partial , 25 aa]	gi|247527	(+) 1.99
Albumin-like protein	gi|763431	(+) 3.35
47nhibit, isoform CRA_s	gi|119626082	(+) 2.81
47nhibit, isoform CRA_h	gi|119626071	(+) 2.54
unnamed protein product	gi|221042622	(+) 2.33
Chain A, Crystal Structure Of The Ga Module Complexed With Human Serum Albumin	gi|55669910	(+) 1.98
Chain A, Crystal Structure Of Human Serum Albumin Complexed With Myristic Acid And Lysophosphatidyl	gi|228311905	(+) 2.55
Hypothetical protein	gi|51476390	(+) 2.26
unnamed protein product	gi|213499853	(+) 2.77
unnamed protein product	gi|284388777	(+) 2.18
Ig G3m(g) pFc HER	gi|222995	(+) 2.59
Ig gamma 2 H chain, BUR [Peptide , 348 aa]	gi|243169	(+) 2.20
immunoglobulin kappa light chain variable region	gi|48475432	(+) 2.77
Chain A, Crystal Structure Of The Fab Fragment Of A Human Monoclonal Igm Cold Agglutinin	gi|10835792	(+) 2.42
Chain A, Solution Structure Of Double Super Helix Model	gi|293651900	(+) 2.37
C-C motif chemokine 24	tr|A6NH40|	(+) 2.28
Platelet glycoprotein 4	tr|A6NKU1|	(+) 2.15
Vesicle-associated membrane protein 2	tr|Q7L876|	(+) 1.16
Inter-alpha-trypsin inhibitör heavy chain H1	tr|A6NBZ8|	(+) 2.17
Fibroblast growth factor receptor substrate 3	tr|B7WNR0|	(+) 1.94
Integrin alpha-Iıb	tr|B7WPB5|	(+) 1.79
SH3 domain-binding glutamic acid-rich-likeprotein	tr|B5MD33|	(+) 1.80
Cytotoxic T-lymphocyte protein 4	tr|Q569I7|	(+) 2.51
amyloid beta A4 protein	gi41406057	(+) 2.82
Kallikrein 6	gi21466131	(-) 2.53

Table 3.5 (Continued): Differentially expressed proteins in PPMS subtype.

Superoxide dismutase	gi928825	(-) 2.54
Agrin	gi1947080	(-) 2.55
Chitinase 3 like protein	gi61741895	(+)2.56
Neuronal cell adhesion molecule	gi51095143	(-) 2.30
Alpha-1 antichymotrypsin	gi112874	(-) 2.14
14-3-3 protein	gi531590	(+) 2.51
Annexin-1	gi113944	(+) 2.82

3.3 Differentially Expressed Protein Biomarkers In MS Converting CIS Patients

CIS patients were clinically and radiologically followed up to average of 3 years for evaluation of MS conversion. At the end of follow-up period, converters and nonconverter patients' proteome were compared. Comparison revealed differentially expressed 28 proteins (Table 3.6). Among them, alpha-1-B-glycoprotein, Annexin A4 and C1 inhibitor proteins were also observed in RRMS patients.

Table 3.6: Differentially expressed proteins in CIS compared to RRMS.

MS Converters vs. Nonconverters		
Protein Name	**Accession number**	**Change**
immunoglobulin light chain	gi|149673887	(+) 2.21
unnamed protein product	gi|21758324	(+) 1.73
Chain A, Human Transferrin N-Lobe Mutant H249e	gi|6980467	(+) 2.81
RecName: Full=Serotransferrin; Short=Transferrin; AltName: Full=Beta-1 metal-binding globulin; AltN	gi|313104271	(+) 2.23
albumin, isoform CRA_o	gi|119626078	(+) 2.57
anti-Rh(D) antibody immunoglobulin heavy chain variable region	gi|93007610	(+) 2.29
Ig gamma 2 H chain, BUR [human, Peptide Mutant, 348 aa]	gi|243169	(+) 2.33
Chain A, The 1.6 A Crystal Structure Of Human Bruton's Tyrosine Kinase Bound To A Pyrrolopyrimidine	gi|284055568	(+) 3.44
albumin, isoform CRA_h	gi|119626071	(+) 2.03
alpha-1-B-glycoprotein - human	gi|69990	(+) 1.27
alloalbumin Venezia	gi|178345	(+) 2.21
hypothetical protein	gi|51476390	(+) 2.33
RecName: Full=Serotransferrin; Short=Transferrin; AltName: Full=Beta-1 metal-binding globulin; AltN	gi|313104271	(+) 2.41
C1-inhibitor	gi|179617	(+) 2.57
albumin, isoform CRA_h	gi|119626071	(+) 2.29
Chain M, Crystal Structure Of A Broadly Neutralizing Antibody In Complex With A Peptide	gi|48425065	(+) 2.34
immunoglobulin light chain	gi|149673887	(+) 2.33
immunoglobulin variable region	gi|323433034	(+) 1.98

Table 3.6 (Continued): Differentially expressed proteins in CIS compared to RRMS.

Chain L, Hiv-1 Hxbc2 Gp120 Envelope Glycoprotein Complexed With Cd4 And Induced Neutralizing Antibo	gi|12084266	(+) 2.55
unnamed protein product	gi|215415638	(+) 2.26
CUGBP Elav-like family member 2	tr|Q6N095|	(-) 2.77
P3 protein	tr|Q6N093|	(+) 2.18
Human androgen-induced prostate proliferative shutoff associated protein (AS3)	gi|21758324	(+) 2.59
Human skeletal mRNA for myosin heavy chain light meromyosin region	gi|1335218	(+) 2.73
Annexin A4	gi|194375550	(+) 1.98
C1 inhibitor	gi|1340170	(-) 1.79
Leukotriene A-4 hydrolase	gi|194374381	(+) 1.80
Homo sapiens WD repeat domain 59	gi|193787129	(+) 2.51

3.4 Differentially Expressed Common And Shared Proteins

Differentially expressed proteins were also classified depending on their presence in all subtypes individually. Among them transferrin, Chain A Alpha 1 Anti - Trypsin, Apolipoprotein A4 and Apolipoprotein E with several albumin isoforms, were abundantly observed in CSF of disease clinical subtypes.

The differentially expression presence of alloalbumin was 100% in CIS, 72% in RRMS and 16% in PMS. For alpha – 2 macroglubulin presence were, 36% in CIS, 100% in RRMS and %22 in PMS, Immunoglobulin proteins G (IgG2, IgG 3, immunoglobulin light chain) proteins have been found in %48 of CIS, 54% of RRMS and 100% of PMS patients.

Some proteins showed abundance in more than one subtype. Human alpha – 1 Acid Glycoprotein have been differentially found in 100% of CIS and RRMS patients but 42% of PMS patients. Likewise prostaglandin D2 Synthase (21 kDa) found in 100% of CIS and RRMS patients but 34% of PMS patients. Superoxide dismutase was found in 82% of CIS, 26% of RRMS and CIS patients. 14-3-3 protein was found in 92% of CIS, 96% of RRMS patients and 14% of PMS patients. SPARC like protein was 18% of CIS, 92% of RRMS patients and 18% of PMS patients. Cytotoxic T Lymphocyte protein 4 was found in 100% of CIS, %76 of RRMS patients and %72 of PMS patients. The percentages of presence of those proteins in different clinical subtypes were given in supplementary table 3.7.

Table 3.7: Frequency of common and shared proteins in different MS subtypes. Values belonging to different proteins refer to percentages of indicated proteins in all related MS subtype individuals.

Common Proteins	CIS (%)	RRMS (%)	PPMS (%)
Human Transferrin	100	100	100
Chain A Alpha 1 Anti-trpsin	100	96	100
Apolipoprotein A4	100	76	100
Albumin, isoform CRA_h	100	88	100
Apolipoprotein E	100	100	92
Subtype Specific Proteins			
Alloalbumin Venezia	100	72	16
Alpha 2 macroglobulin	36	100	22
Ig gamma 2 H chain, [348 aa]	48	54	100
Chain A, Crystal Structure Of Human Alpha 1 Acid Glycoprotein	100	100	42
Prostaglandin D2 synthase 21kDa (brain)	100	100	34
Superoxide dismutase	82	26	26
14-3-3 protein	92	96	14
SPARC like protein	18	92	18
Cytotoxic T-lymphocyte protein 4	100	76	72
Calcium/calmodulin-dependent protein kinase type II subunit beta	64	96	32
Cyclic AMP-dependent transcription factor ATF-7	88	24	32

3.5 Confirmation Of Individual Proteins

Selected protein markers were confirmed by ELISA and western blot studies. Antibodies specific to target proteins were used for above listed proteins (Table 2.1). Selections of candidate proteins for confirmation studies were based on the function of the molecules and their role possible role in multiple sclerosis pathophysiology or in other inflammatory/ neurodegenerative disorders. At this step we also studied with serum samples of same patients group, in order to observe serum level changes of target proteins. Because, biomarker potential of a candidate protein can be evaluated by the correlated changes of protein in serum.

The level of CSF APOA4 protein was significantly differed in CIS (1,82 fold, p=0.004) and RRMS (1,41 fold, p=0.011), whereas PPMS group did not show significant difference compared to control group (Figure 3.5 a). Serum ApoA4 levels did not significantly correlated between any of subtypes, also did not correlated with CSF levels (Figure 3.5 b). Western blot studies also confirmed the results (Figure 3.5 c).

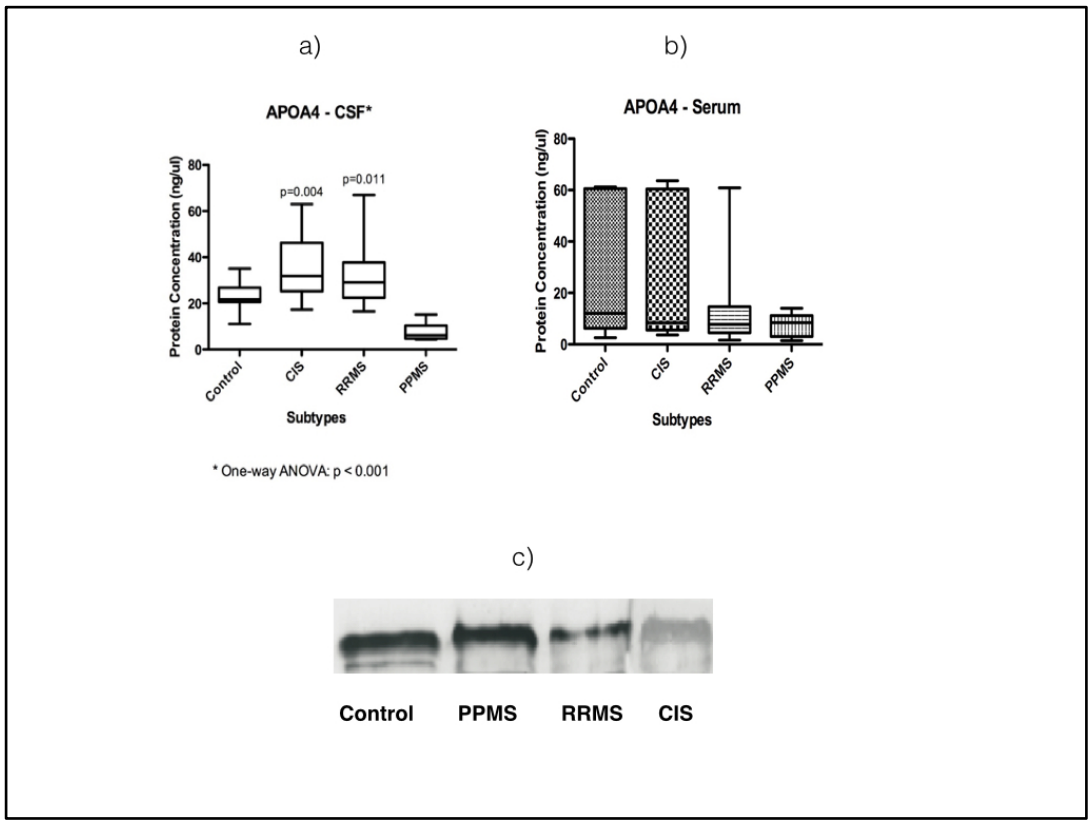

Figure 3.5: Confirmation of APOA4 protein in different MS subtypes. a) ELISA results of APOA4 - CSF protein. b) ELISA results of APOA4 - serum protein. c) Western blot results of APOA4 CSF protein.

Apolipoprotein E showed significant difference in both CSF and serum. CSF ApoE in RRMS decreased compared to controls (1,44 fold, p=0.0481), but did not changed in other subtypes (Figure 3.6 a). Serum level of ApoE showed significant decrease in both RRMS (1.92 fold, p=0.0223) and PPMS groups (2.21 fold, p=0.0459) (Figure 3.6 b). Western blot results also showed consistency with the ELISA results (Figure 3.6 c).

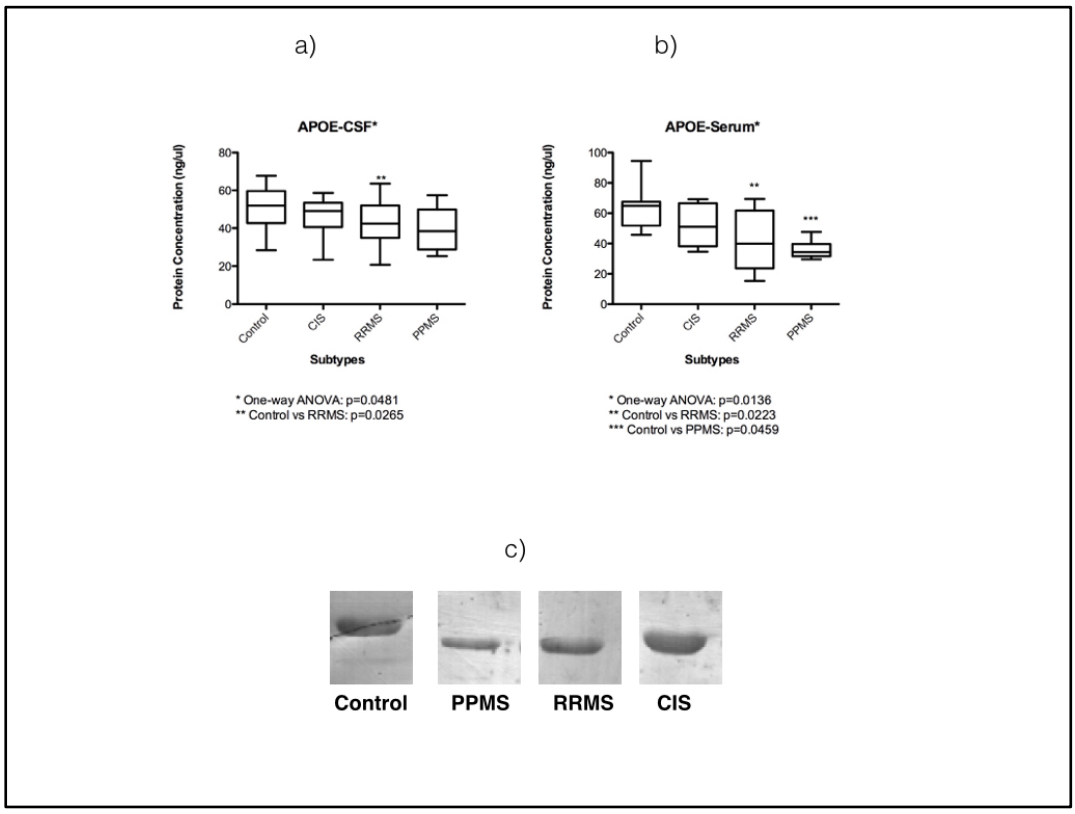

Figure 3.6: Confirmation of APOE protein in different MS subtypes. a) ELISA results of APOE - CSF protein. b) ELISA results of APOE - serum protein. c) Western blot results of APOA4 CSF protein.

CSF Transthyretin (TTR) concentration in CIS and RRMS groups significantly decreased compared to controls (1.66 fold, p=0.0021 and 1.72 fold, p=0.004, respectively) (Figure 3.7 a). Serum TTR levels were also significantly decreased in CIS (2.40 fold, p=0.018) and RRMS groups (2.15 fold, p=0.047) (Figure 3.7 b). PPMS group did not show any significant change either in CSF or serum of MS patients. This may be caused by the clinical variations among PPMS patients. Western blot results also confirmed the ELISA results (Figure 3.7 c).

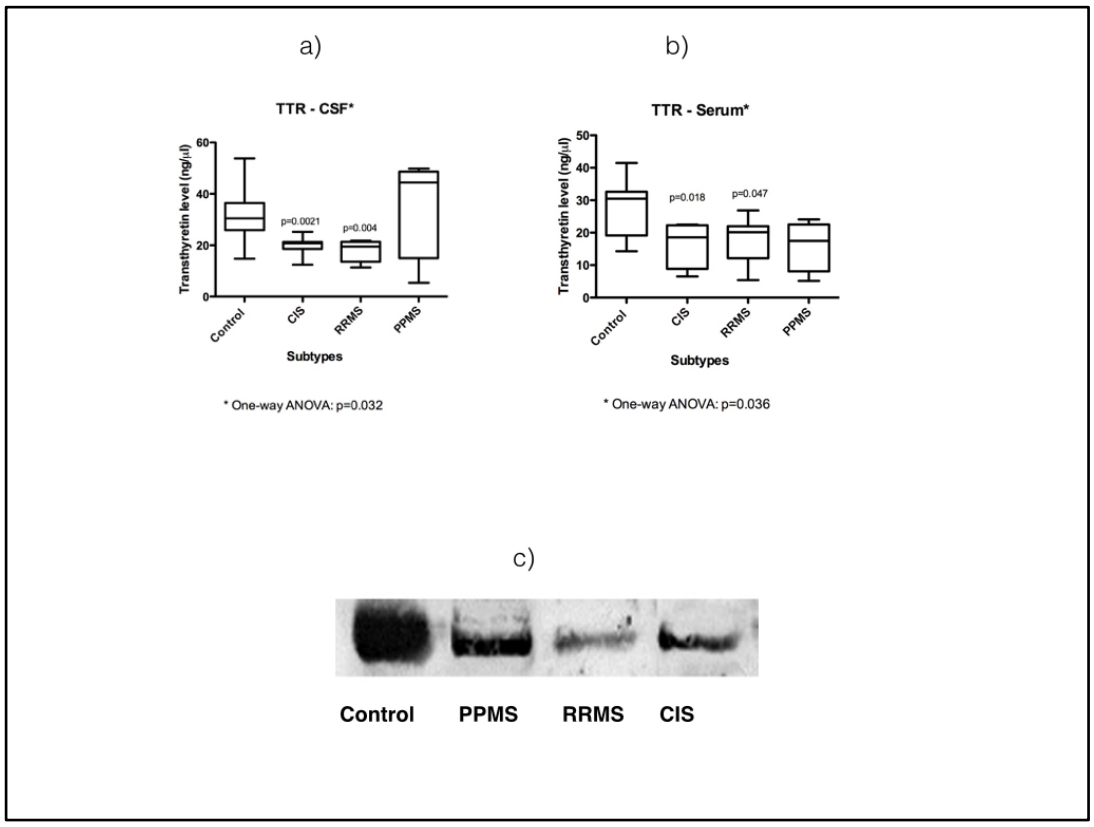

Figure 3.7: Confirmation of TTR protein in different MS subtypes. a) ELISA results of TTR - CSF protein. b) ELISA results of TTR - serum protein. c) Western blot results of TTR CSF protein.

Prostaglandin D2 synthase level in CSF was also significantly differed in CIS and RRMS whereas PPMS did not show significant change. Both groups showed elevated protein level compared to control (RRMS: 1,24 fold, $p<0.0001$, CIS: 1,44 fold, $p<0.0001$) (Figure 3.8 a). However serum level of PGD2 neither showed a significant change, nor correlation with CSF results (Figure 3.8 b). Western blot results also confirmed the CSF ELISA results; CIS and RRMS subtypes showed significance increase in western blot studies (Figure 3.8 c).

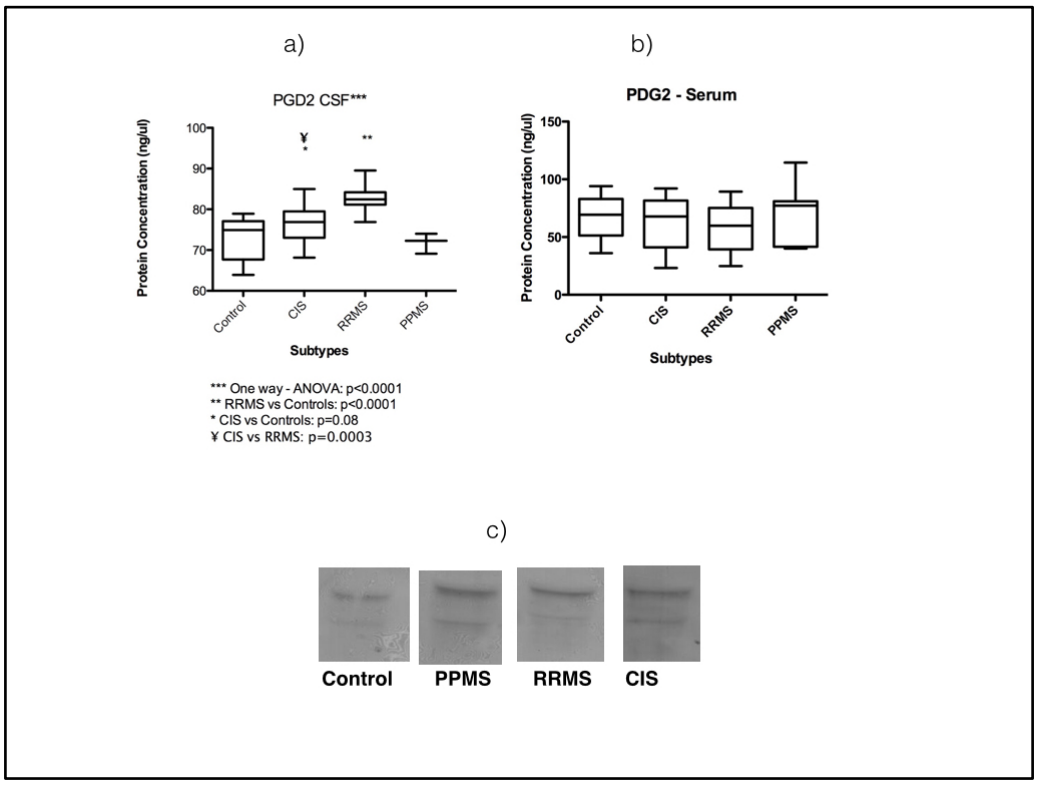

Figure 3.8: Confirmation of PDG2 protein in different MS subtypes. a) ELISA results of PDG2- CSF protein. b) ELISA results of PDG2- serum protein. c) Western blot results of PDG2 CSF protein.

Alpha-1 acid glycoprotein level in CSF and serum significantly differed in RRMS (1.32 fold, p=0.0021) and CIS (1.34 fold, p=0.0021) subtypes compared to control group. The level of α1-AGP in serum and CSF increased in CIS (1.41 fold, p=0.0038) and RRMS (1.52 fold, p=0.0046) subtypes, whereas not in PPMS (Figure 3.9 a, b). Western blot also confirmed the results (Figure 3.9 c).

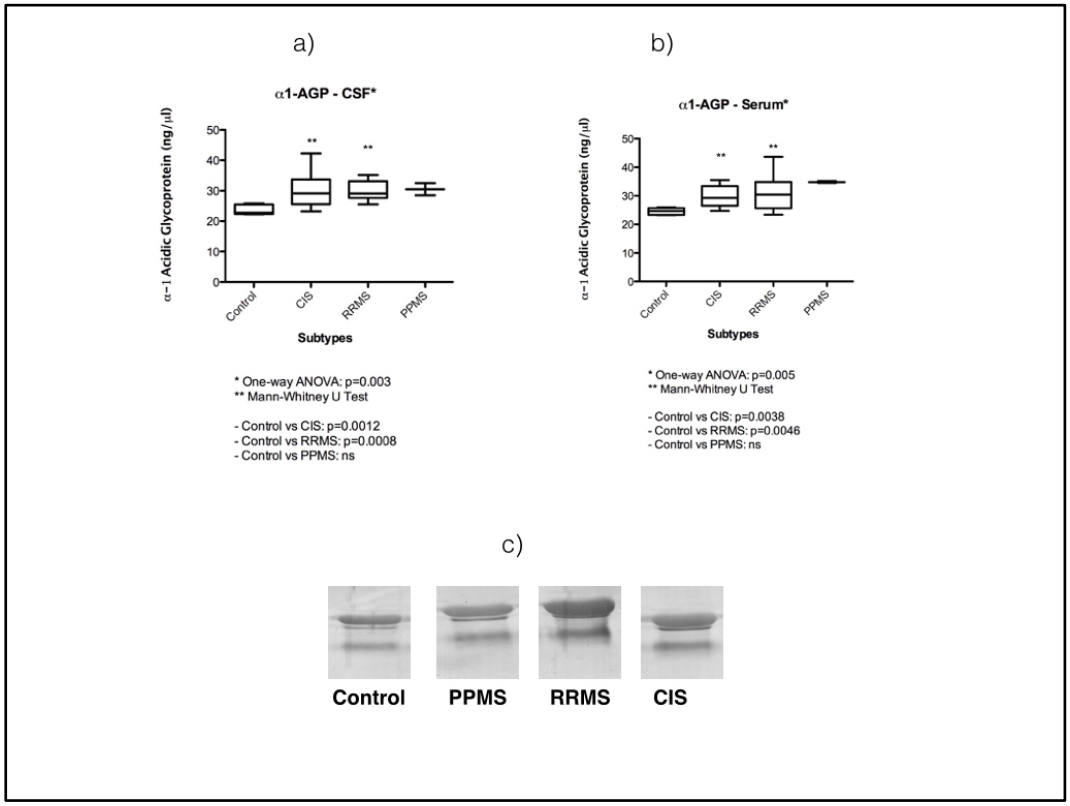

Figure 3.9: Confirmation of α1-AGP protein in different MS subtypes. a) ELISA results of α1-AGP - CSF protein. b) ELISA results of α1-AGP - serum protein. c) Western blot results of α1-AGP CSF protein.

Tau protein, Glial Fibrillary Acidic Protein (GFAP), Neurofilament Light Chain (NFL), Myelin Oligodendrocyte Glycoprotein (MOG) and myelin basic protein (MBP) concentrations were only confirmed by ELISA. In previous study we confirmed the result by western blot in same study group (Avsar et al., 2012). Tau, GFAP, MOG, NFL proteins showed elevated CSF concentrations in all MS subtypes, but MBP did not (Figure 3.10 a, b, c, d, e). ELISA results showed consistency with the proteomic studies.

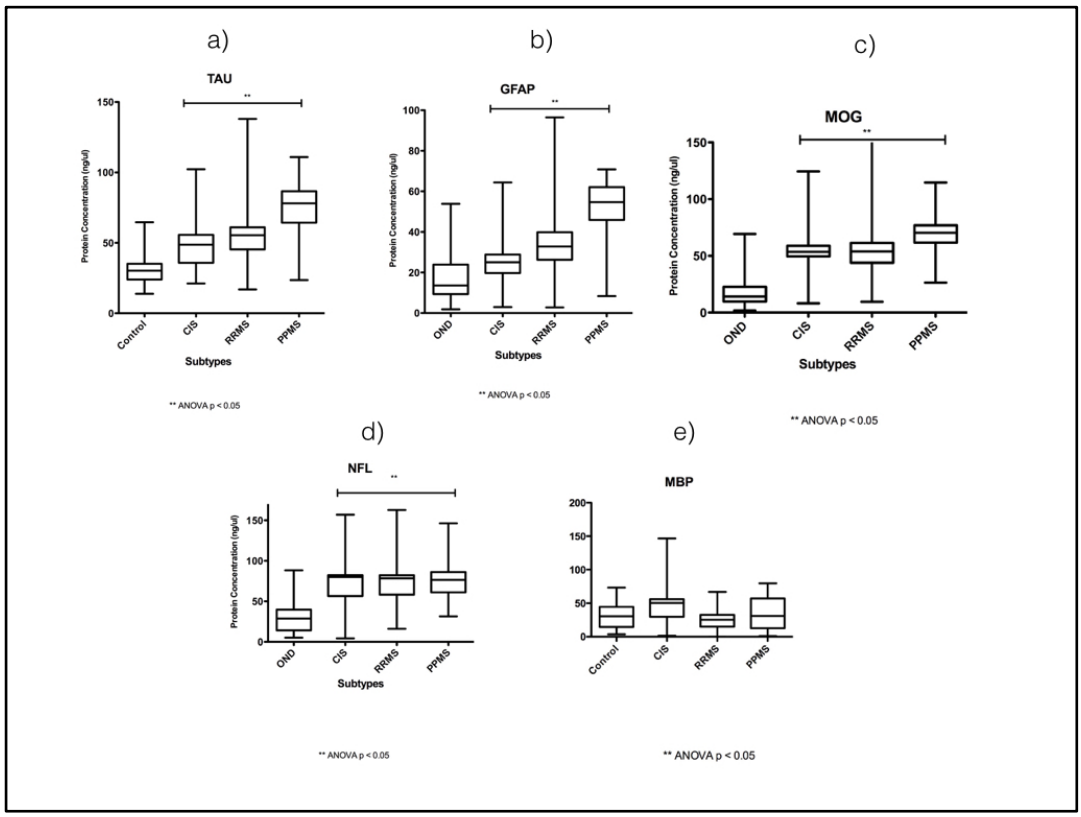

Figure 3.10: ELISA confirmation studies for previously selected proteins. a) Tau protein, b) GFAP, c) MOG protein, d) NFL protein, e) MBP.

Vitamin D binding protein (VDBP) concentration CSF was significantly in CIS (1.22 fold, $p<0.0001$)subtype compared to controls whereas RRMS and PPMS subtypes showed no difference ($p<0.001$) (Figure 3.11 a). However serum level of VDBP did not correlated with CSF results (Figure 3.11 b).

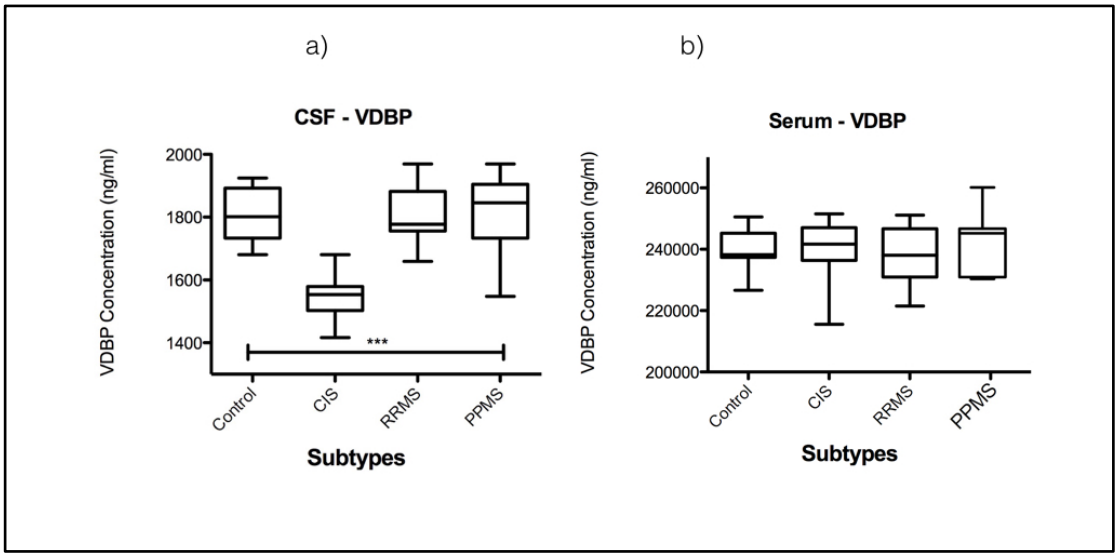

Figure 3.11: Confirmation of vitamin D binding protein with ELISA. a) CSF VDBP results, b) Serum VDBP results.

Alpha 2-macroglobulin level in CSF was also significantly differed in MS subtypes, but not in serum samples. The level of α2-macroglobulin was significantly increased in CIS, RRMS and PPMS subtypes ($p<0.005$) (Figure 3.12 a). Serum level of α2-macroglobulin did not significant difference in any subtypes (Figure 3.12 b).

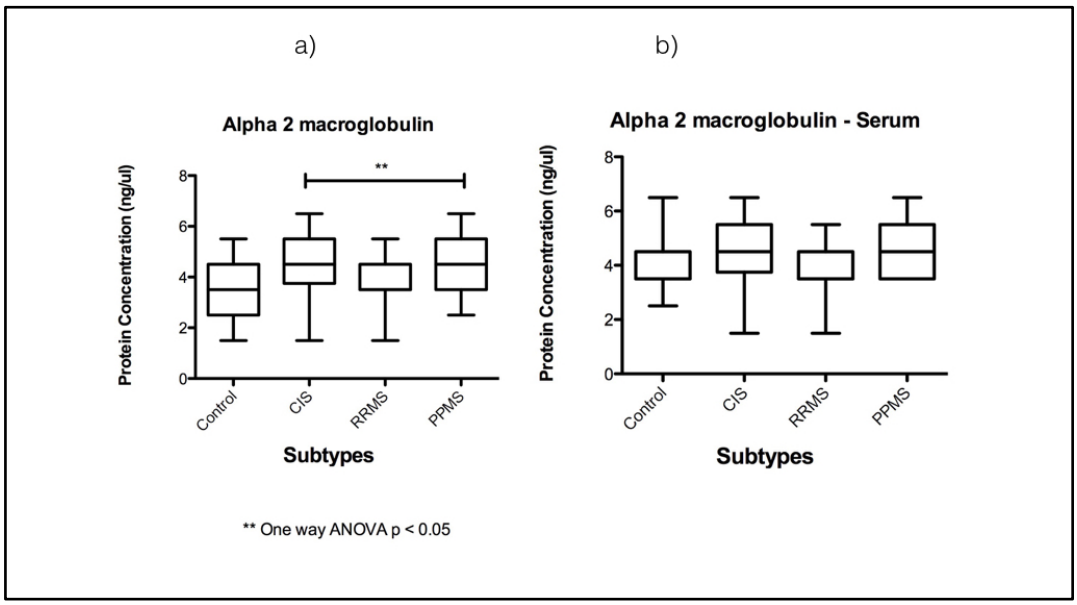

Figure 3.12: Confirmation of α2-macroglobulin protein with ELISA. a) CSF α2-macroglobulin protein results, b) Serum α2-macroglobulin protein results.

The CSF level of trypsin concentration in CIS, RRMS, PPMS, and control groups was analyzed by One-way ANOVA test and there was no significant difference among these groups ($p=0.08$; Figure 3.10 a). However, the level of CSF trypsin was found to be significantly lower in RRMS subtype compared to the CIS subtype ($p=0.032$; Figure 3.13 a). The serum level of trypsin in CIS, RRMS, and control groups showed significant difference between these subtypes. The level of trypsin protein in the serum of RRMS and CIS subtype is significantly lower than in the control group ($p=0.0206$, $p=0.0308$; Figure 3.13 b).

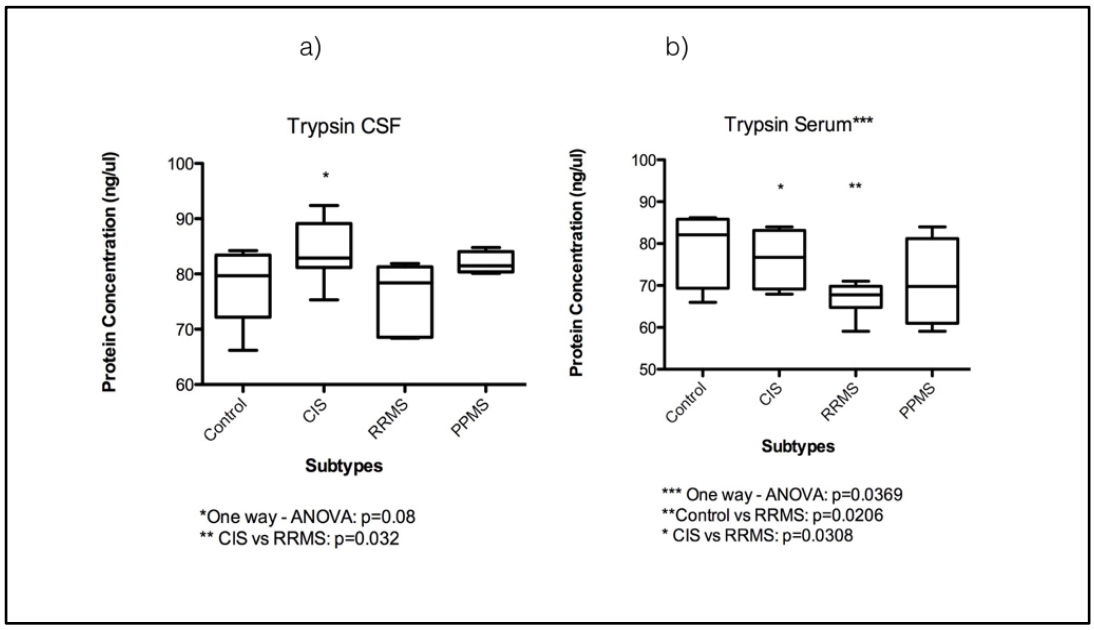

Figure 3.13: Confirmation of trypsin protein with ELISA. a) CSF trypsin protein results, b) Serum protein results.

Pathway analysis of listed proteins revealed MS pathogenesis related molecular pathways. Therefore proteins belonging to these pathways were also confirmed apart from the proteomic studies. Phosphatidyl linositol 3-phopsphate (PI3P), Serine/Threonine-protein kinase (SGK), Insulin receptor (IR) and E3 ubiquitin-protein ligase (NEDD4L) proteins were compared between MS subtypes and controls. Among them only PI3P and NEDD4L proteins were significantly elevated in PPMS subtype (Figure 3.11), The rest did not show significant difference in CSF. Any of serum results did not show significant change, but PPMS subtype could not be tested in serum samples for these proteins due to lack of serum samples (Figure 3.12).

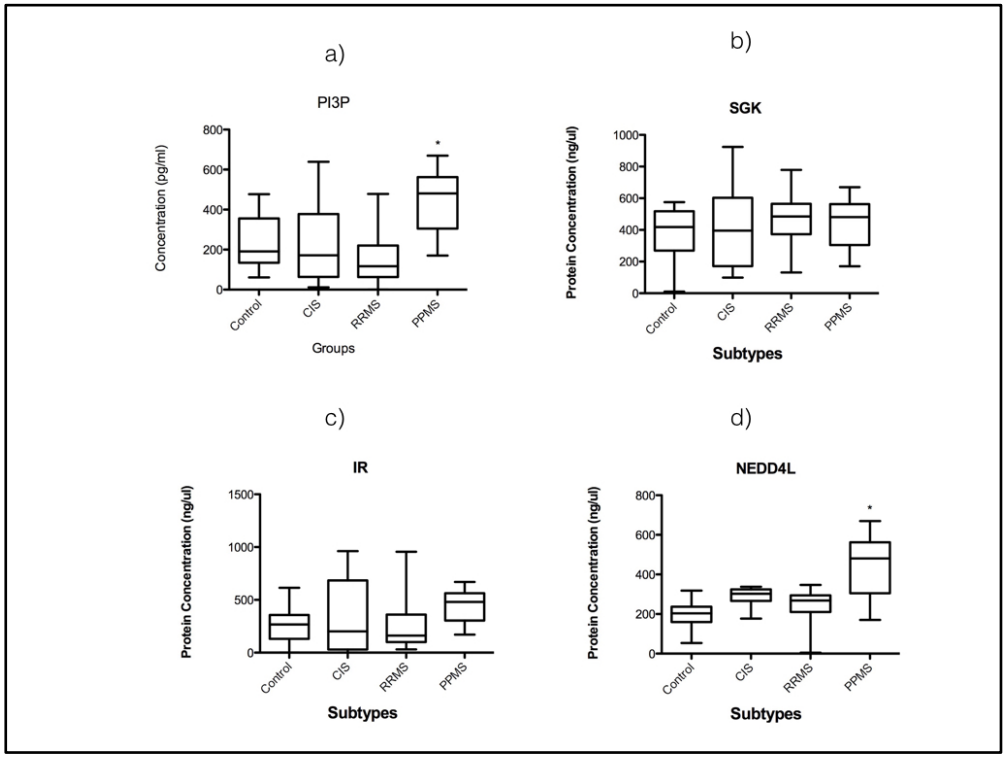

Figure 3.14: Confirmation of Sodium reabsorption pathway proteins with ELISA in CSF samples. a) PI3P, b) SGK, c) IR, d) NEDD4L.

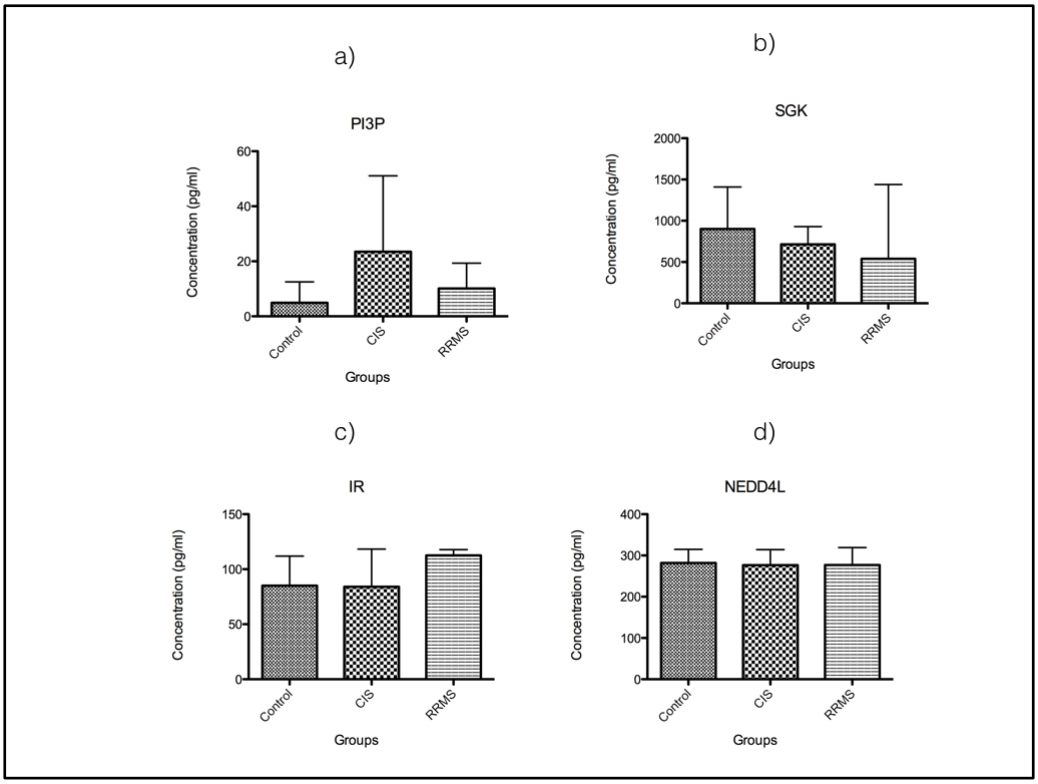

Figure 3.15: Confirmation of Sodium reabsorption pathway proteins with ELISA in serum samples. a) PI3P, b) SGK, c) IR, d) NEDD4L.

ELISA results of selected proteins were listed in table 8. Concentration of proteins were given as μg/ml with 95% CI.

Table 3.8: Confirmation study results of all selected proteins.

Protein	Control		CIS		RRMS		PPMS	
	CSF	Serum	CSF	Serum	CSF	Serum	CSF	Serum
APOE	38,36	62,54	54,63	51,23	40,11**	40,21**	39,74	38,37**
APOA4	21,20	15,24	30,41**	12,46	28,94*	10,12	10,96	9,74
PDG2S	78,98	66,28	79,36	65,33	81,38*	63,74	71,56	70,55
TTR	36,41	30,17	19,87**	19,91**	18,66**	19,33**	42,14	17,54
A2MG	3,67	3,54	4,12**	4,01	4,08**	3,54	4,07*	4,03
VDBP	18,04	228,16	15,11**	220,24	16,24*	217,12	19,03	245,41
A1AT	24,41	23,56	28,32**	25,02**	27,11**	28,03**	30,56	30,14
GFAP	16,01	10,41	22,71**	14,30	34,46**	17,27	58,22**	19,33
MBP	24,13	11,12	50,16	10,98	22,22	13,20	24,32	14,21
NFL	44,31	40,17	68,10**	60,11	67,14**	61,01	74,17*	58,11
Tau	40,34	34,22	48,11**	30,10	51,22**	33,21	70,34**	28,12
MOG	22,21	24,01	51,56**	26,08	51,24**	27,13	56,71**	20,1
Trypsin	79,96	83,32	82,12*	76,11*	76,76	64,33**	50,01	66,12
PI3P	198,14	8,94	193,21	21,17	176,11	12,34	436,71*	-
SGK	412,20	800,76	399,16	770,13	432,33	594,12	411,32	-

* refers to $p<0.01$
** refers to $p<0.001$
*** refers to $p<0.0001$

3.5 Molecular Pathways In MS Subtypes

Interaction of differentially expressed proteins investigated in terms of associated molecular pathways. Gene annotations of corresponding proteins with their degree of change (increase / decrease) were used in KEGG pathway analysis and appointed to significant biological pathways. In this context, identified molecular pathways were compared between disease subtypes and control group. Firstly, results showed that renin-angiotensin system (RAS) and complement and coagulation cascade (CCC) pathways were common in all disease subtypes, with varying significances (given in table 3.9). Pathway associated proteins found in sub-networks were the angiotensin and MAS1 proteins. Mean level of angiotensin protein in the CIS group was increased 4.2 folds, 3.5 fold in RRMS group and 2.8 in PMS group compared to control group ($p<0.05$). For complement and coagulation cascade pathway, associated proteins were Kininogen 1, Complement Component 7, Alpha-2-

Macroglobulin, Serpin E1, Serpin G1, Serpin A1, Plasminogen proteins. Among them, Serpin E1 and Kininogen were decreased, whereas alpha-2-Macroglobulin levels were increased compared to control groups.

Table 3.9: Shared molecular pathways among all MS subtypes.

KEGG ID	KEGG Term	P Value	Times Found	Associated Genes Found in Subnetworks
KEGG:04610	Complement and coagulation cascades (CCC)	$4.06x10^{-10}$	1	KNG1, A2M, C9, FGB, SERPINE1, SERPING1, SERPINA1
KEGG:04614	Renin-angiotensin system (RAS)	$6.88x10^{-6}$	3	AGT, MAS1

3.6 Shared And Subtype Specific Molecular Pathways

Besides common pathways, bioinformatics analysis also revealed the subtype specific and shared pathways between different subtypes. PPMS and RRMS groups showed an affected vitamin digestion and absorption system ($p = 2.7x10^{-6}$, $p = 1.73x10^{-5}$), and notch signaling pathways ($p=8,9x10^{-7}$, $p = 3.89x10^{-10}$). CIS and RRMS groups shared activated prion disease pathway ($p = 9.05x10-^{10}$, $9.05x10^{-10}$), which is related with neuronal degeneration. Besides shared subtypes, RRMS group showed activated pathogenic *E.coli* absorption pathway ($p = 2.12x10^{-5}$), indicating the importance of inflammatory cell movement.

Comparisons of CIS subtype and control samples indicated that, prion disease ($p=9.06x10^{-10}$) complement and coagulation cascades ($p=4.19x10^{-9}$), renin-angiotensin system ($p=6.88x10^{-6}$), proteasome ($p=4.31x10^{-8}$) pathways were significantly affected by the interaction of differentially expressed proteins (Table 3.10).

Table 3.10: Molecular pathways indicated by proteome studies in CIS subtype.

KEGG ID	KEGG Term	P Value	Times Found	Associated Genes Found in Subnetworks
KEGG:05020	Prion diseases	$9.06x10^{-10}$	1	C8A, C8B, C9, C8G
KEGG:04610	Complement and coagulation cascades (CCC)	$4.19x10^{-09}$	1	KNG1, A2M, C9, FGB, SERPINE1, SERPING1, SERPINA1
KEGG:03050	Proteasome	$4.31x10^{-8}$	1	PSMB10, PSMA2, PSMA1, PSMD13, PSMB2, PSME
KEGG:04614	Renin-angiotensin system (RAS)	$6.88x10^{-6}$	3	AGT, MAS1

Differentially regulated molecular pathways in RRMS subtype are: Notch signaling pathway ($p=3.89x10^{-10}$), prion disease pathway ($p=9.05x10^{-10}$), complement and coagulation cascades pathway ($p=2.10x10^{-9}$), renin-angiotensin system pathway ($p=6.88x10^{-6}$), vitamin digestion and absorption pathway ($p=1.73x10^{-5}$) and pathogenic *Escherichia coli* infection ($p=2.12x10^{-5}$) pathways (Table 3.11).

Table 3.11: Molecular pathways indicated by proteome studies in RRMS subtype.

KEGG ID	KEGG Term	P Value	Times Found	Associated Genes Found in Subnetworks
KEGG:04330	Notch signaling pathway	$3.89x10^{-10}$	2	NOTCH3, NCSTN, NOTCH2, APH1A, RBPJ
KEGG:05020	Prion diseases	$9.05x10^{-10}$	2	C8A, C8B, C9, C8G
KEGG:04610	Complement and coagulation cascades (CCC)	$2.10x10^{-9}$	1	KNG1, A2M, SERPINE1, F2, SERPING1, SERPINA1, PLG
KEGG:04614	Renin - angiotensin system (RAS)	$6.88x10^{-6}$	4	AGT, MAS1
KEGG:05130	Pathogenic *Escherichia coli* infection	$2.12x10^{-5}$	1	ACTB, KRT18, FYN, TUBA4A, NCL, CTNNB1
KEGG:04977	Vitamin digestion and absorption	$1.73x10^{-5}$	1	APOB, CUBN, APOA1

PPMS subtype also showed common and shared molecular pathways. Notch signaling pathway ($p=8.91x10^{-8}$), renin angiotensin system pathway ($p=6.88x10^{-6}$) and vitamin digestion and absorption ($p=1.73x10^{-5}$) pathways were differentially

affected in PPMS subtypes. The list of associated genes found / not found in subnetworks and were given in table 3.12.

Table 3.12: Molecular pathways indicated by proteome studies in PPMS subtype.

KEGG ID	KEGG Term	P Value	Times Found	Associated Genes Found in Subnetworks
KEGG:04330	Notch signaling pathway	$8.91x10^{-8}$	1	NOTCH2, EP300, PSEN1, DTX1, JAG2
KEGG:04614	Renin-angiotensin system (RAS)	$6.88x10^{-6}$	5	AGT, MAS1
KEGG:04977	Vitamin digestion and absorption	$1.73x10^{-5}$	1	APOB, CUBN, APOA1

3.7 Molecular Pathways In MS Converting CIS Subtype

Diagnosis of CIS and MS is a complicated procedure and diagnostic criteria may change depending on the differential criteria that are used (McDonald's and Poser's). Regarding the unpredictable course of CIS patients and the importance of prognosis in MS, a subgroup analysis comprising CIS patients remaining CIS and converting to clinically definite MS subtype, CIS group was separated into three different subgroups. MS suggestive CIS (MSs-CIS) indicates the CIS patients remaining CIS with a single clinical episode and MRI findings suggestive of MS but not fulfilling the diagnostic MRI criteria (either dissemination in time or space) to be diagnosed as MS during the follow up period. Single attack MS (SA-MS) indicates the patients who fulfill 2010 McDonald Criteria either at onset of clinical episode or during the follow up period. Clinically definite MS (CDMS) patients indicate the patients who fulfill the McDonald 2010 by having a second clinical attack, which are previously Poser criteria. Results of subgroup analysis were given in table 3.13.

Table 3.13: Molecular pathways and associated proteins revealed by CIS subgroup analysis.

KEGG ID	KEGG Term	P Value	Times Found	Associated Genes Found in Subnetworks
Multiple Sclerosis suggestive – Clinically Isolated Syndrome				
KEGG:04330	Notch signaling pathway	$3.89x10^{-10}$	1	NOTCH3, NCSTN, NOTCH2, APH1A, RBPJ
KEGG:04610	Complement and coagulation cascades (CCC)	$1.05x10^{-8}$	1	KNG1, C7, A2M, SERPINE1, SERPING1, SERPINA1, PLG
KEGG:04930	Type II diabetes mellitus	$7.73x10^{-7}$	1	IRS4, MAPK1, IRS2, INSR, PIK3R1, PIK3R2
KEGG:04614	Renin-angiotensin system (RAS)	$6.88x10^{-6}$	4	AGT, MAS1
KEGG:04960	Aldosterone-regulated sodium reabsorption	$1.78x10^{-4}$	1	MAPK1, INSR, PIK3R1, PIK3R2
Single Attack – Multiple Sclerosis				
KEGG:04614	Renin-angiotensin system	$6.88x10^{-6}$	7	AGT, MAS1
KEGG:04621	NOD-like receptor signaling pathway	$6.69x10^{-8}$	1	HSP90AA1, RELA, NFKBIA, TRAF6, TNFAIP3, IKBKB
KEGG:04610	Complement and coagulation cascades	$7.89x10^{-2}$	2	A2M, THBD, F5, F2, SERPING1, SERPINA1, F7, PROC
Clinically Definite Multiple Sclerosis				
KEGG:05130	Pathogenic Escherichia coli infection	$1.61x10^{-10}$	1	PRKCA, ARPC3, FYN, ABL1, NCL, ITGB1
KEGG:05020	Prion diseases	$1.45x10^{-10}$	1	NOTCH1, C9, SOD1, PRNP
KEGG:04614	Renin-angiotensin system (RAS)	$6.88x10^{-9}$	6	AGT, MAS1
KEGG:04610	Complement and coagulation cascades (CCC)	$3.80x10^{-7}$	1	KNG1, F12, A2M, F2, SERPING1, SERPINA1, PLG

3.8 GWAS Study Revealed Disease Associated Significant SNPs

Filtering and cleaning of raw data is required for the bioinformatics analysis of GWAS. To this aim Golden Helix SNP and Variation Suite software was used and filtering was performed based on the software's suggestions. Firstly, for the quality

assessments, SNPs with low call rates (lower than 95%) (11638 SNPs), SNPs on Y chromosome (3894 SNPs) and SNPs showing linkage disequilibrium (r^2>0.5) (118727 SNPs) were filtered. Further, 31136 SNPs with minor allele frequency lower than 0.01 were also eliminated in order to obtain disease relation of variants. Additionally, one of the two healthy controls was excluded from the study, since they showed IBD level higher than 0.25, indicating second-degree relativity. As a result GWAS analysis was performed with a total of 129.605 SNP markers in 11 MS patients and 59 healthy controls, revealing 14 SNPs showing significant association ($p<10^{-4}$) (Table 3.14). In addition, 106 SNPs showed higher frequencies in cases compared to controls, even though the differences did not achieve statistical significance ($p<10^{-3}$) (Table 3.15).

Table 3.14: Significant SNPs with $p<10^{-4}$.

Marker	Ref/alt allele	Chr.	Region	MAF (cases)	MAF (controls)	P
rs7873	T|C	11	INS-IGF2	0,36	0,03	$4.39E^{-07}$
rs7092208	G|A	10	MGMT	0,32	0,04	$9.66E^{-06}$
rs16990774	G|A	22	EFCAB6	0,23	0,02	$2.25E^{-05}$
rs844959	G|T	X	SPANXC	0,27	0,01	$3.11E^{-05}$
rs9388582	C|T	6	SOGA3	0,27	0,03	$3.68E^{-05}$
rs4605385	G|A	2	SCN9A	0,41	0,08	$3.75E^{-05}$
rs6737602	T|C	2	STK39	0,68	0,25	$4.46E^{-05}$
rs7377981	C|T	4	WHSC1	0,14	0	$4.69E^{-05}$
rs498973	C|T	17	PLXDC1	0,14	0	$4.69E^{-05}$
rs11696329	C|T	20	ATP9A	0,14	0	$4.69E^{-05}$
rs2656219	G|A	11	CNTN5	0,18	0,01	$4.71E^{-05}$
rs12988176	C|T	2	STON1-GTF2A1L	0,27	0,03	$6.68E^{-05}$
rs474241	C|T	11	CNTN5	0,45	0,11	$7.79E^{-05}$
rs13061453	G|A	3	CHCHD6	0,64	0,2	$9.21E^{-05}$

Table 3.15: SNPs with strong associations with $p<10^{-3}$.

Marker	Ref/alt allele	Chr.	Region	MAF (cases)	MAF (controls)	P
rs2249115	C|A	21	CLDN14	0,5	0,14	$1.05E^{-04}$
rs17058979	C|A	8	PNOC	0,23	0,03	$1.25E^{-04}$
rs11790551	G|T	9	GNA14	0,23	0,03	$1.25E^{-04}$
rs2834646	A|C	21	RUNX1	0,32	0,05	$1.38E^{-04}$
rs7356599	C|T	5	MYO10	0,27	0,03	$2.03E^{-04}$
rs2462915	C|T	14	BCL11B	0,32	0,07	$2.25E^{-04}$
rs4980041	C|T	10	ZMIZ1	0,59	0,19	$2.25E^{-04}$
rs9532539	G|A	13	LINC00598	0,82	0,36	$2.42E^{-04}$
rs2272504	A|C	12	TRHDE	0,5	0,15	$2.46E^{-04}$
rs10160702	C|T	11	KIRREL3	0,41	0,11	$2.54E^{-04}$
rs5910394	C|T	X	DOCK11	0,82	0,31	$2.54E^{-04}$
rs2376481	G|A	15	GABRG3	0,68	0,28	$2.64E^{-04}$
rs7122320	G|A	11	CNTN5	0,59	0,2	$3.03E^{-04}$
rs5994128	A|G	22	GAB4	0,59	0,2	$3.03E^{-04}$
rs599462	A|G	11	BC041900	0,32	0,06	$3.13E^{-04}$
rs6760801	G|A	2	CCDC85A	0,45	0,11	$3.21E^{-04}$
rs4255775	T|G	16	KIFC3	0,22	0,03	$3.27E^{-04}$
rs12123953	C|T	1	GIPC2	0,5	0,14	$3.50E^{-04}$
rs4980047	G|A	10	ZMIZ1	0,5	0,16	$3.65E^{-04}$
rs6869645	C|T	5	SLC6A3	0,18	0,02	$3.70E^{-04}$
rs7850420	C|T	9	MAPKAP1	0,18	0,02	$3.70E^{-04}$
rs12417295	C|T	11	ARHGAP42	0,18	0,02	$3.70E^{-04}$
rs35206614	A|G	13	ATP7B	0,18	0,02	$3.70E^{-04}$
rs34853349	A|G	14	LINC00617	0,18	0,02	$3.70E^{-04}$
rs16984253	C|T	20	MYT1	0,18	0,02	$3.70E^{-04}$
rs12485136	G|A	22	PNPLA5	0,18	0,02	$3.70E^{-04}$
rs2875237	G|A	3	GRM7	0,27	0,05	$3.72E^{-04}$
rs4973986	A|C	3	ULK4	0,27	0,05	$3.72E^{-04}$
rs17311679	A|G	19	VRK3	0,27	0,05	$3.72E^{-04}$
rs1774847	T|C	1	TMEM183A	0,77	0,36	$3.85E^{-04}$
rs2295794	C|T	9	TLN1	0,09	0,48	$3.89E^{-04}$
rs3008639	G|A	1	AIDA	0,18	0,02	$4.21E^{-04}$
rs2414893	C|A	15	MEGF11	0,5	0,18	$4.25E^{-04}$
rs12297776	C|T	12	ERGIC2	0,41	0,12	$4.35E^{-04}$
rs507280	C|T	6	ROS1	0,36	0,1	$4.55E^{-04}$
rs11190831	C|T	10	TLX1NB	0,36	0,1	$4.55E^{-04}$
rs733126	C|T	11	USH1C	0,36	0,1	$4.55E^{-04}$
rs2370473	G|A	2	BC039382	0,23	0,03	$4.78E^{-04}$
rs16899078	C|T	5	CDH12	0,23	0,03	$4.78E^{-04}$
rs502156	T|G	6	AK126334	0,23	0,03	$4.78E^{-04}$
rs4751086	A|G	10	MGMT	0,23	0,03	$4.78E^{-04}$
rs4766819	C|T	12	FBXW8	0,23	0,03	$4.78E^{-04}$
rs7249176	A|G	19	LILRP2	0,23	0,03	$4.78E^{-04}$
rs4611723	G|A	21	LOC339622	0,23	0,03	$4.78E^{-04}$
rs4972701	C|T	2	BC046497	0,77	0,31	$4.86E^{-04}$

Table 3.15 (Continued): SNPs with strong associations with $p<10^{-3}$.

rs8025178	G|A	15	TRPM1	0,64	0,23	$4.91E^{-04}$
rs16971450	A|G	15	MTHFS	0,45	0,12	$4.92E^{-04}$
rs17464525	G|A	1	AP4B1	0,27	0,04	$5.19E^{-04}$
rs2034360	C|T	2	PRKCE	0,45	0,14	$5.37E^{-04}$
rs17080136	T|C	5	MAPK9	0,45	0,14	$5.37E^{-04}$
rs10976051	T|G	9	KDM4C	0,64	0,25	$5.37E^{-04}$
rs626657	C|A	17	CACNB1	0,36	0,09	$5.89E^{-04}$
rs7820617	T|C	8	TRIM55	0,27	0,03	$6.56E^{-04}$
rs6577449	C|T	1	CAMTA1	0,23	0,03	$7.05E^{-04}$
rs12712051	C|T	2	AFF3	0,23	0,03	$7.05E^{-04}$
rs10853465	C|T	18	LINC00669	0,23	0,03	$7.05E^{-04}$
rs6843239	C|T	4	SEC24D	0,18	0,56	$7.36E^{-04}$
rs5918486	A|G	X	SRPX	0,55	0,16	$7.48E^{-04}$
rs2147852	C|T	13	TBC1D4	0,5	0,18	$7.53E^{-04}$
rs12870594	T|C	13	TEX26-AS1	0,32	0,06	$7.83E^{-04}$
rs2133127	G|A	15	SNRPN	0,59	0,22	$8.02E^{-04}$
rs2032701	G|A	3	MECOM	0,36	0,11	$8.25E^{-04}$
rs2072793	G|A	20	RBPJL	0,77	0,38	$8.31E^{-04}$
rs11003872	C|T	10	PCDH15	0,68	0,33	$8.44E^{-04}$
rs6663477	A|G	1	VPS45	0,14	0,01	$8.65E^{-04}$
rs7589234	T|C	2	AFF3	0,14	0,01	$8.65E^{-04}$
rs1515495	A|G	3	TP63	0,14	0,01	$8.65E^{-04}$
rs10045011	C|T	5	FYB	0,14	0,01	$8.65E^{-04}$
rs799157	T|C	7	MLXIPL	0,14	0,01	$8.65E^{-04}$
rs12771692	G|A	10	LOC399715	0,14	0,01	$8.65E^{-04}$
rs4143858	A|G	11	CNTN5	0,14	0,01	$8.65E^{-04}$
rs7957161	T|C	12	PLXNC1	0,14	0,01	$8.65E^{-04}$
rs9565321	C|T	13	MYCBP2	0,14	0,01	$8.65E^{-04}$
rs7233275	G|A	18	TMEM241	0,14	0,01	$8.65E^{-04}$
rs11671231	C|T	19	LILRA5	0,14	0,01	$8.65E^{-04}$
rs7143254	C|A	14	LRFN5	0,5	0,17	$8.82E^{-04}$
rs11791292	T|G	9	ADAMTSL1	0,27	0,06	$9.07E^{-04}$
rs3740829	T|C	11	ARHGAP32	0,27	0,06	$9.07E^{-04}$
rs1330884	T|G	13	RNF219-AS1	0,14	0,51	$9.26E^{-04}$
rs13440231	C|T	9	NR6A1	0,18	0	$9.69E^{-04}$
rs2131905	C|T	1	AKNAD1	0,09	0	$9.69E^{-04}$
rs17258606	T|C	1	NOTCH2	0,09	0	$9.69E^{-04}$
rs12077365	T|C	1	OBSCN	0,09	0	$9.69E^{-04}$
rs7581919	C|T	2	NBAS	0,09	0	$9.69E^{-04}$
rs3917329	G|T	2	IL1R1	0,09	0	$9.69E^{-04}$
rs3813247	G|T	2	TTN	0,09	0	$9.69E^{-04}$
rs9835263	A|G	3	AK124857	0,09	0	$9.69E^{-04}$
rs11098016	T|G	4	COL25A1	0,09	0	$9.69E^{-04}$
rs1521026	G|A	5	CDH18	0,09	0	$9.69E^{-04}$
rs7726659	A|G	5	ANKRD31	0,09	0	$9.69E^{-04}$
rs6964977	A|G	7	WBSCR17	0,09	0	$9.69E^{-04}$
rs10499905	C|T	7	SGCE	0,09	0	$9.69E^{-04}$

Table 3.15 (Continued): SNPs with strong associations with $p<10^{-3}$.

rs10486860	C|T	7	DLX6-AS1	0,09	0	$9.69E^{-04}$
rs6560837	C|T	10	DIP2C	0,09	0	$9.69E^{-04}$
rs3816685	T|C	10	JMJD1C	0,09	0	$9.69E^{-04}$
rs4769191	A|G	13	AK054845	0,09	0	$9.69E^{-04}$
rs17093516	A|G	14	SYNE2	0,09	0	$9.69E^{-04}$
rs2382868	T|C	16	GTF3C1	0,09	0	$9.69E^{-04}$
rs35794595	G|T	17	NTN1	0,09	0	$9.69E^{-04}$
rs414528	T|C	19	USE1	0,09	0	$9.69E^{-04}$
rs4803831	G|T	19	OPA3	0,09	0	$9.69E^{-04}$
rs6051643	G|A	20	DDRGK1	0,09	0	$9.69E^{-04}$
rs12484542	C|T	22	CSNK1E	0,09	0	$9.69E^{-04}$
rs5978925	C|T	X	KAL1	0,09	0	$9.69E^{-04}$
rs1947442	G|A	4	CCSER1	0,41	0,11	$9.91E^{-04}$
rs6753123	C|T	2	LOC284950	0,32	0,08	$9.97E^{-04}$

3.9 Linkage Analysis Indicated Two Chromosomal Loci In Chr 13 And Chr 21

SimWalk tool (version 2.91) was used for the linkage analysis. Since MS pedigrees were small and their sample sizes were not well enough, non-parametric linkage (NPL) analysis was performed. Multipoint estimation regarding the identity by descent probabilities, in 10 MS pedigrees resulted with exclusion of 324 uninformative SNP markers among a total of 3118 informative SNPs. Two chromosomal regions with NPLz score higher than 1.7, which is a threshold for a potential association, were observed on chromosome 13 and 21 (Figure 3.16). The location of linkage region at chr 13 was 37.9 cM with NPLz score of 1.72 (p = 0.019) and at chr 21 it was 41.82 cM with NPLz score of 1.7 (p = 0.019).

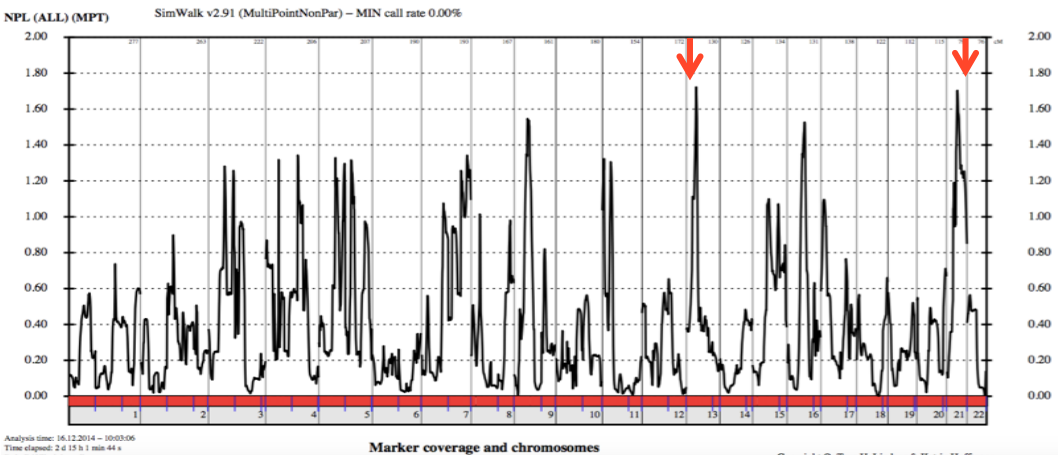

Figure 3.16: SimWalk Multipoint NPL Analysis showing the highest scores. Red arrows indicate chromosome 13 and 21.

For the detailed analysis of regions, fine mapping was applied to those two regions. Results indicated that regions comprising the stated highest NPL scores (33.4 cM - 41.5 cM for chromosome 13 and 36 cM - 52.5 cM for chromosome 21) with 639 and 831 SNPs respectively. Therefore it was revealed that the highest linkage peaks were on 34.11 cM (nearest marker rs1461965, NPL Z = 1.82, p = 0.015) on chromosome 13 and 45.08 cM (nearest marker rs11701543, NPL Z = 1.85, p = 0.014) on chromosome 21 (Figure 3.14 and Figure 3.15), revealing suggestive evidence of linkage for chromosomal regions of 13q13.3 and 21q22.2 to MS.

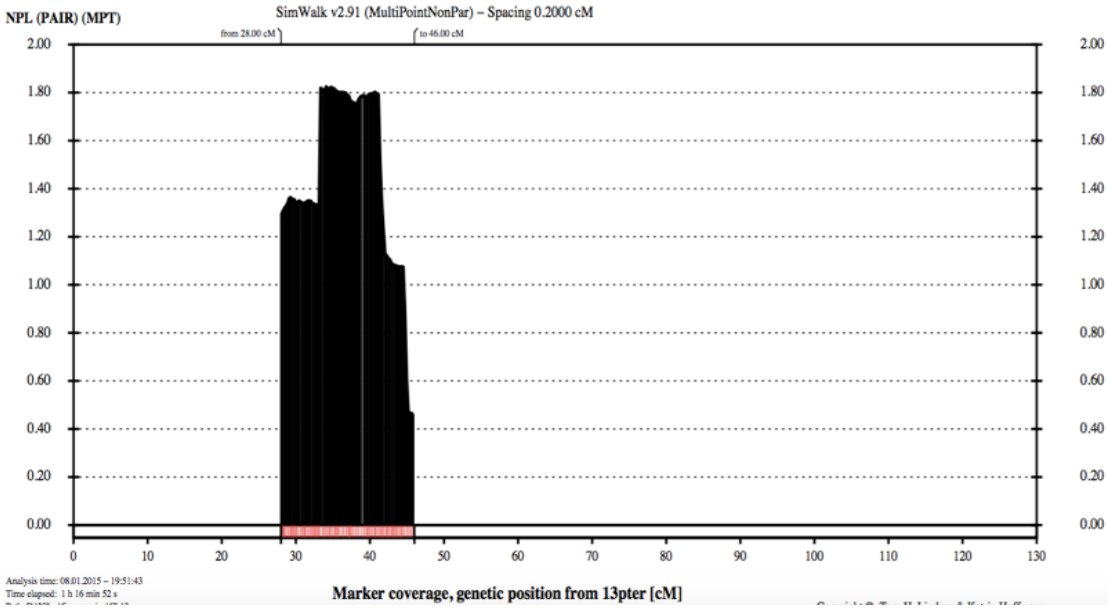

Figure 3.17: Fine mapping of chromosome 13 from 28 cM to 46 cM region

Most significantly associated 5 genotyping markers or SNPs at chr13q13.3 were listed at table 3.16.

Table 3.16: Associated SNPs at the chr 13q13.3.

No	Marker	Chr	cM	pNPLp	pNPLa	NPLp	NLPa	npLODp	npLODa
1	rs1461965	13	34,11	0,015	0,015	1,824	1,824	0,7224	0,7224
2	rs7989245	13	34,73	0,0151	0,0151	1,8222	1,8222	0,721	0,721
3	rs9542104	13	33,37	0,0152	0,0152	1,8182	1,8182	0,7179	0,7179
4	rs17053473	13	34,94	0,0152	0,0152	1,8174	1,8174	0,7172	0,7172
5	rs9315409	13	35,15	0,0153	0,0153	1,8155	1,8155	0,7157	0,7157

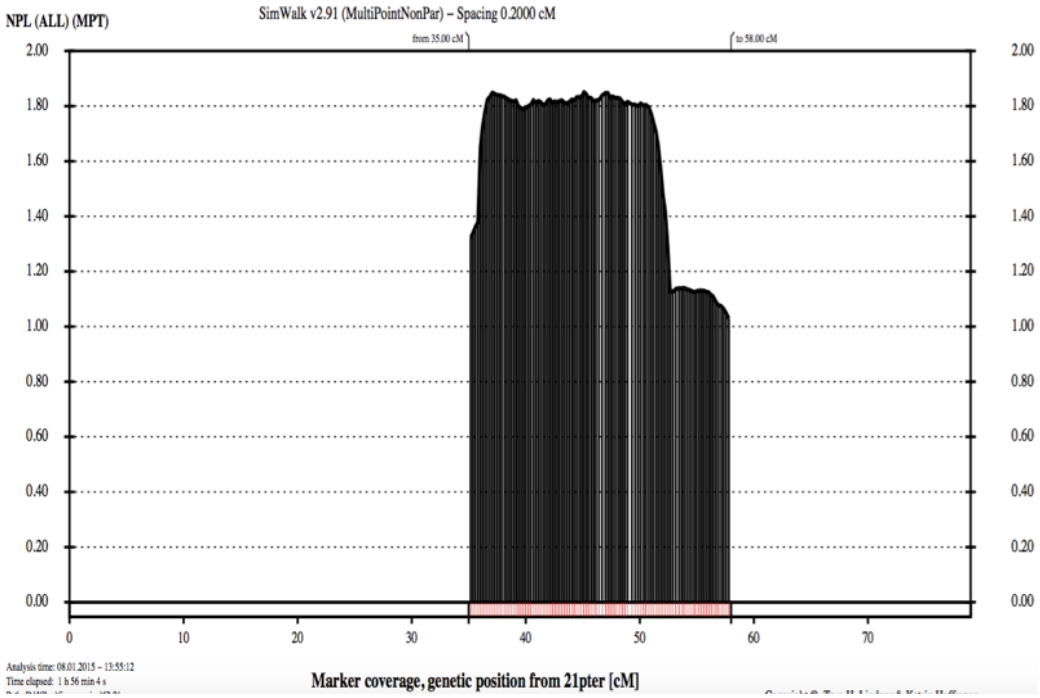

Figure 3.18: Fine mapping of chromosome 21 from 35 cM to 58 cM region

Most significantly associated 5 genotyping markers or SNPs at chr21q22.2 were listed at table 3.17.

Table 3.17: Associated SNPs at the chr 21q22.2.

No	Marker	Chr	cM	pNPLp	pNPLa	NPLp	NLPa	npLODp	npLODa
1	rs11701543	21	45,08	0,0141	0,0141	1,8508	1,8508	0,7438	0,7438
2	rs2834046	21	37,05	0,0142	0,0142	1,8486	1,8486	0,7421	0,7421
3	rs2836096	21	46,98	0,0142	0,0142	1,8473	1,8473	0,741	0,741
4	rs2836400	21	47,21	0,0142	0,0142	1,8473	1,8473	0,741	0,741
5	rs2836271	21	37,26	0,0143	0,0143	1,8452	1,8452	0,7393	0,7393

Finally, genes located on these chromosomal regions of suggestive linkage were analyzed. 24 protein-coding genes in chromosome 13 and 58 protein coding genes in chromosome 21 were listed (Figure 3.16). Furthermore a literature survey was performed in order to observe relation of these genes to MS and other neurological and/or inflammatory conditions. The search was resulted in 22 candidate genes.

3.10 Correlation Of Proteomic And Genomic Results

Correlation between the linkage and the GWAS results were investigated by comparing protein coding genes located on the suggestive linkage regions, genes possessing the SNPs with significant, suggestive associations in the GWAS, and genes encode for the proteins that found to be differentially expressed in MS subtypes CSF proteomic data of MS patients. Furthermore, each patient sample data with protein and genes having altered expression level individually compared with the linkage and the GWAS results, and overlapping genes were listed (Table 3.18). Additionally, pathways in which proteins encoded by each gene revealed in the genetic analyses involved were noted (KEGG Pathway Database, GeneCards The Human Gene Compendium), and genes encoding for proteins that involve in the affected pathways implicated in the proteomic analyses were determined. The best-correlated genes were determined for further analyses.

Table 3.18: The list of correlated genes and molecular pathway from proteome and GWAS results.

Marker	Chr.	Region	Common Pathways	P (GWAS)
rs7873	11	INS-IGF2	Sodium reabsorption, Type II diabetes	$4.39x10^{-07}$
rs2034360	2	PRKCE	Renin-angiotensin, Type II diabetes	$5.37x10^{-04}$
rs17080136	5	MAPK9	NOD-like receptor signaling, Type II diabetes	$5.37x10^{-04}$
rs2072793	20	RBPJL	Notch signaling	$8.31x10^{-04}$
rs11791292	9	ADAMTSL1	Notch signaling	$9.07x10^{-04}$
rs13440231	9	NR6A1	Notch signaling	$9.69x10^{-04}$
rs17258606	1	NOTCH2	Notch signaling	$9.69x10^{-04}$
rs3917329	2	IL1R1	NOD-like receptor signaling	$9.69x10^{-04}$
rs35794595	17	NTN1	Notch signaling	$9.69x10^{-04}$

4. DISCUSSION:

MS is the most common neurological complex disorder causing disability in young adults (Lassmann et al., 2001). Yet its cause is still mainly unknown with an unpredictable course and prognosis. To this aim, we investigated the proteome content of CSF and serum samples of patients representing MS clinical heterogeneity from a prospective MS cohort and used the benefits of bioinformatics approaches that allowed insights into the important pathophysiologic processes in the main subtypes of MS. Furthermore genome wide association and linkage studies performed for the correlation of key proteins in genomic levels. Regarding the data we obtained in this study, result can be discussed in 4 different titles.

1) Identified potential protein biomarkers for MS.

2) MS pathogenesis associated molecular pathways revealed by the differentially expressed proteins.

3) Candidate genomic associations of the disease with regard to chromosomal regions and gene variations.

4) Correlation of proteomics and genomics results to reveal pathogenesis related key molecules.

4.1 Potential Protein Biomarkers In Multiple Sclerosis

We have identified 151 proteins that are differentially expressed and evaluated a panel of 20 proteins among them in CSF and serum as the most potential biomarkers for MS diagnosis and for monitoring of disease stage/progression. The list of differentially expressed proteins with a potential of prognostic or diagnostic values, were listed above in tables 3.2, 3.3 and 3.4. Among them some were selected for the verification of identified molecular pathways.

Apolipoprotein E (APOE) is a plasma protein that has many forms and is functional in the lipid transportation metabolism between astrocytes and neurons. Additionally, when there is damage in neurons or axons, apolipoprotein family proteins facilitate

the regeneration process (Huang et al., 2004). Usually apolipoprotein E is found in astrocytes, however during degeneration it becomes functional on neurons (Reiber, 2001). Studies reporting ApoE concentrations in MS patients' body fluids revealed that the CSF-ApoE level is at least two fold decreased in MS patients (Gerdes, 2003). In this study we also observed that serum and CSF level of APOE protein significantly decreased in RRMS and PPMS subtypes. CIS subtype also showed decreased APOE level, but the difference was not significant. Additionally CSF results also correlate with the serum levels independent of age, sex and disease duration. Regarding the previous findings and our findings in proteomic studies, it may be said that APOE protein has a potential to become a protein biomarker showing the remyelination status in MS subtypes, however further studies are required to correlate with other remyelination markers.

Apolipoprotein A4 (APOA4) is another lipoprotein family member protein and is a major circulating apolipoprotein, its physiological function is not completely clear. Several functions have been ascribed to APOA4 protein. It aggregates in the regenerating sciatic nerve cells for myelin biosynthesis (Olmarker et al., 1996). Previous findings indicated that the level of APOA4 upregulated in CSF of patients with lumbar disc herniation and lebers hereditary optic neuropathy (LHON) (Liu et al., 2006, D'Aguanno et al., 2008). In this study, APOA4 level of CIS and RRMS patients showed significantly elevated CSF level but PPMS patients did not. Serum level of APOA4 protein did not show any consistency and significant difference. APOA4 protein is a relatively small protein that has various forms. Additionally it is difficult to detect all forms via single antibody or 2D-PAGE studies. Furthermore serum level of APOA4 protein is not suitable to be a potential biomarker.

Prostaglandin D2 synthase (PGD2S) is a catalytic enzyme that converts the prostaglandin H2 to prostaglandin D2. It also functions as neuromodulator and trophic factor in CNS. The level of PGD2S increases in neuronal cells (Ragolia et al., 2001) and brain tumors (Saso et al., 1998). It is shown that PGD2 can induce production of Th2-type cytokines (IL-4, IL-5, and IL-13), thus leading to a differentiation of Th2 lymphocytes (Khanapure et al., 2007). Previous studies indicated increased level of PGD2S level in MS CSF but not in controls (Hammack et al., 2004). In this study it is found that, the level of PGD2S upregulated in CSF of CIS and RRMS patients but not in PPMS patients. Serum level did not differed

among subtypes and not correlated with CSF results. Results showed that PGD2S might be a potential biomarker for inflammation and Th2 type cytokine production biomarker in CSF of CIS and RRMS subtypes. It may be used as a prognostic risk factor to follow up of inflammatory status in MS patients.

Alpha 2 macroglobulin (A2MG) is an antiprotease and inactivates variety of proteinases. It functions as an inhibitor of fibrinolysis by inhibiting plasmin and kallikrein. The level of A2MG is associated with many metabolic diseases and cancer types (James et al., 1980, Kanoh et al., 2000). The level of A2MG in CSF was increased in MS patients (Ottervald et al., 2010). In this study we have also found that, the level of A2MG level in CSF of CIS and RRMS patients showed an increased expression, but not in PPMS and serum samples of any subtypes. Regarding the induced activity of kallikrein-kinin system via renin angiotensin and complement and coagulation cascade pathways, which is discussed in more detail below, A2MG concentration is important in CIS and RRMS but not apparently enough alone.

Transthyretin (TTR) is a serum and CSF carrier of the thyroid hormone thyroxine (T4) and retinol-binding protein bound to retinol. Misfolding and aggregation of TTR is previously associated with amyloid diseases (Zeldenrust and Benson, 2010), polyneuropathies (Coelho, 1996) and cardiomyopathies (Jacobson et al., 1997). In MS literature, there are various findings indicating the change of TTR concentration in MS samples. Many proteomic studies indicated contradicting results for TTR concentration. Some studies indicated elevated TTR concentration (Lehmensiek et al., 2007, Dumont et al., 2004, Chiasserini et al., 2008), some indicated no change (Larsen and DeLallo, 1987), and some indicated decreased levels (Hammack et al., 2004). Our results showed decreased expression level of TTR in CSF and serum samples of CIS and RRMS, but not in PPMS patients. Regarding the correlation between serum and CSF samples of CIS and RRMS subtypes with a distinctive value of TTR, it may be concluded that TTR is may be regarded as a potential biomarker for differentiating clinical subtypes of CIS and RRMS from PPMS subtype.

Vitamin D binding protein (VDBP) is an albumin family protein and found in CSF, plasma samples and cell surface of many cell types. It binds to vitamin D and its metabolites to transport them into target tissues (Chun, 2012). The prevalence of MS is highest where environmental supplies of vitamin D are lowest like higher latitudes.

In experimental studies, vitamin D-deficient diet resulted in an increased susceptibility to EAE, which is an animal model of MS, accelerated onset of symptoms and increased clinical symptoms (Cantorna, 2000). Genetic differences in vitamin D binding protein-coding gene influence vitamin D levels in serum and CSF. In several studies VDBP levels in MS patients' CSF correlate with MS course, being lower during relapses and higher in the secondary progressive phase (Disanto et al., 2011). Furthermore in a proteomic study, Ottervald et al, demonstrated almost two fold increased vitamin D binding protein in SPMS patients compared to controls and other disease subtypes (Ottervald et al., 2010). Here we show that VDBP level in CSF significantly differed increased or decreased – more increased in progressive forms in CIS subtype of MS compared to controls and other subtypes. Serum levels of VDBP did not show correlation and significant difference.

We have also validated the results of our previously published study, which also aimed to reveal target biomarkers in MS subtypes using western blot techniques and machine learning algorithms (Avsar et al., 2012). In this study we aimed to validate previous study by 2D-PAGE and ELISA methods to have more quantitative data. Therefore, selected biomarkers Tau, GFAP, MOG, NFL and MBP tested in each of individual samples. Results were consistent with the previous study and ELISA results of those proteins were significantly elevated in all MS subtypes except the MBP.

Besides the proteins detailed in above, we also aimed to validate other potential biomarkers (alpha-1acidic glycoprotein, anti-trypsin, albumin proteins etc.) that were differentially expression levels observed in 2D-PAGE studies but not in ELISA and western blot studies. Therefore regarding the validation of biomarkers studies results, we may conclude that APOE, PGD2S, TTR, Tau, MOG, GFAP, NFL proteins have the potential biomarker capacity. But further validation studies in different cohorts and control samples including different other neurological and inflammatory diseases are required.

4.2 Subtype Specific And Common Protein Biomarkers In Multiple Sclerosis

Differentially expressed proteins were classified based on their relevance in each subtypes. Angiotensinogen protein (AGT), MAS1 protein which is a receptor for angiotensinogen and complement component C9 proteins were widely seen in each

MS subtype. Combination of these 3 proteins together in most of MS samples lead to activation of complement coagulation cascades and renin angiotensin system pathways. These 3 proteins are major players of these pathways (will be discussed below). Besides common proteins there are also subtype specific and shared proteins that were widely seen in specific MS subtypes more than the other subtypes (Figure 4.1). Proteins were also classified based on their functions including, neurodegeneration, cancer and immune system. Results showed that in and CIS subtypes immune system related proteins are abundant, whereas in PPMS neurodegeneration related proteins are more then the other pathways.

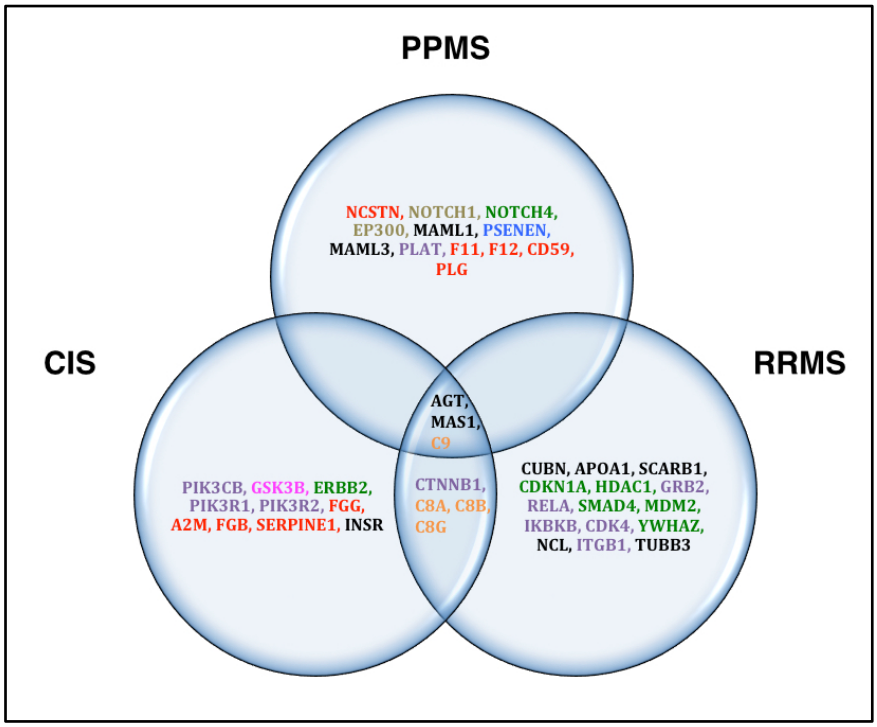

Figure 4.1: Common and subtype specific proteins. Each colour indicates different functional family of proteins. Blue – neurodegeneration, Green – cancer, Brown – neurodegeneration & cancer, Purple – immune system & cancer, Orange – immune system & neurodegeneration, Pink – immune system & neurodegeneration & cancer, Black – others.

4.3 Subtype Specific And Common Molecular Pathways In Multiple Sclerosis

The most prominent finding of this study is the presence of common and subtype specific molecular pathways in MS pathogenesis. Previously interactions of many molecular pathways in MS have been shown via cell culture studies or animal models. However this is the first study revealing the presence of such pathways via proteomic approach and in patient CSF samples. Our findings have revealed renin-

angiotensin system (RAS) and complement and coagulation cascade (CCC) being common pathways shared by all MS clinical subtypes, including early, intermediate and progressive forms of MS. The involvement of CCC and RAS pathways confirm the immune related pathophysiology of MS (Sospedra and Martin, 2005, McFarland and Martin, 2007, Martin et al., 1995). An increase of AGT and MAS1 proteins production of the RAS pathway in combination with the kallikrein-kinin system appear to induce CCC pathway (Figure 4.2).

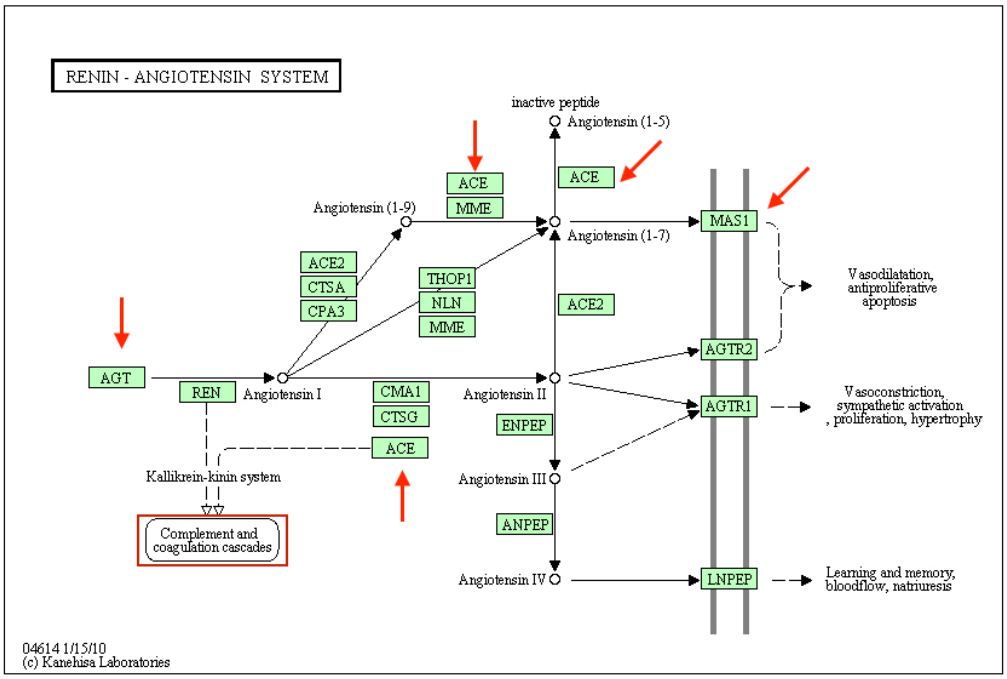

Figure 4.2: Renin angiotensin system pathway. Red arrows indicate the proteins that are found to be elevated in MS subtypes.

Our finding of CCC pathway proteins including upregulation of Kininogen, Plasminogen, Alpha-2-Macroglobulin, SerpinE1, SerpinA1 and SerpinG1 and downregulation of F12 protein all lead to the activation of inflammation with activation of platelets, monocyte and lymphocyte cells (Figure 4.3).

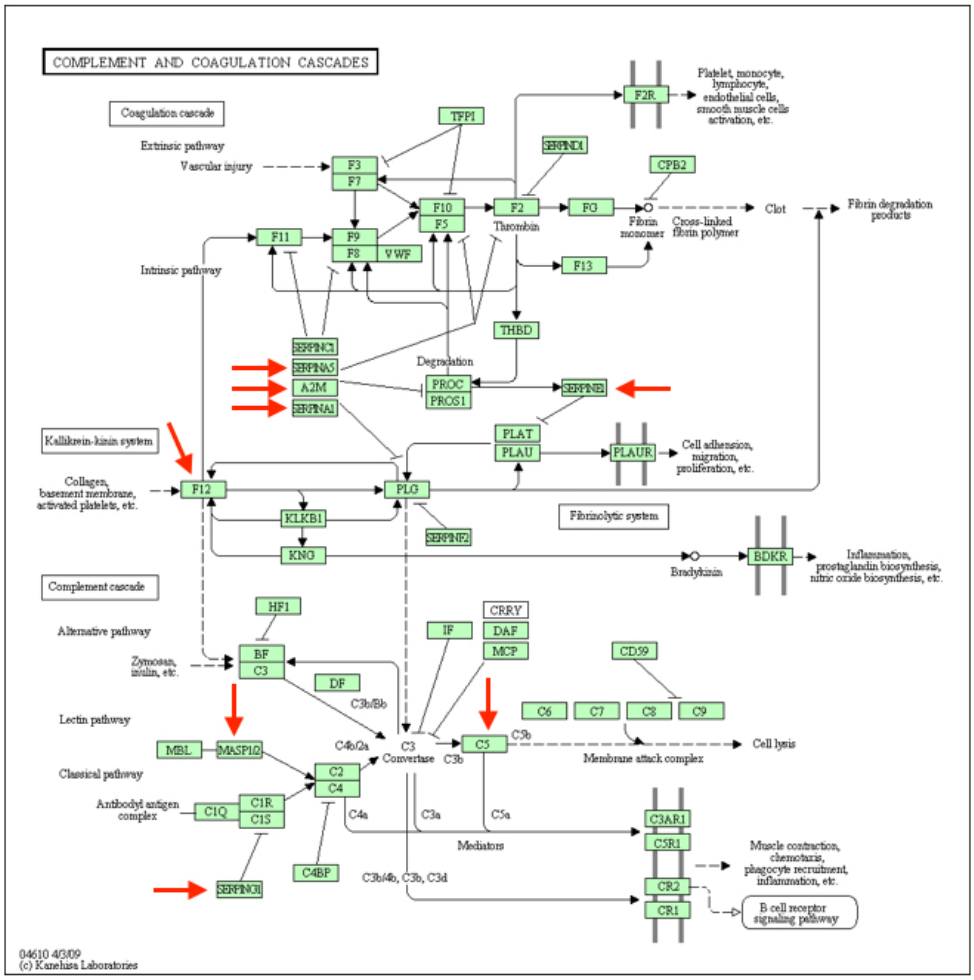

Figure 4.3: Complement and coagulation cascade. Red arrows indicate the proteins that are found to be elevated in MS subtypes.

Alongside with the commonly identified pathways; pathways specific to the clinical subtypes of MS were also revealed. The proteasome pathway in the CIS clinical subtype (including both MSs-CIS and SA-MS) has been over activated with the increased expression of PSMB10, PSMA2, PSMA1, PSMD13, PSMB2, PSME3 proteins (Figure 4.4). This finding may point to the presence of active ubiquitin mediated proteolysis in CIS phenotype via activation of immunoproteasomes.

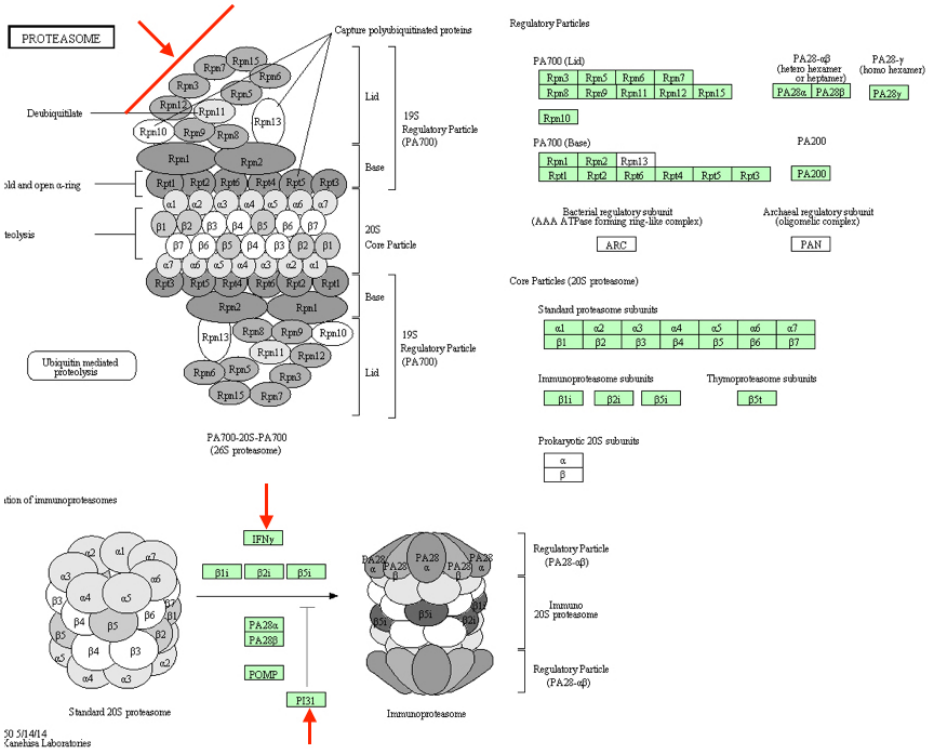

Figure 4.4: Proteasome pathway. Red arrows indicate the proteins that are found to be elevated in MS subtypes.

While the vitamin digestion and absorption pathway in the PMS clinical subtype is activated by the increase in the APOA4, APOA1, SCARB1 proteins (Figure 4.5), it has been observed that the notch signaling pathway is also activated through NOTCH2, EP300, PSEN1, DTX1, JAG2 proteins, which may reflect the fact that the vitamin metabolism of MS is dysregulated and this is supportive with the recently suggested vitamin metabolism involvement in MS pathogenesis (Ascherio et al., 2014).

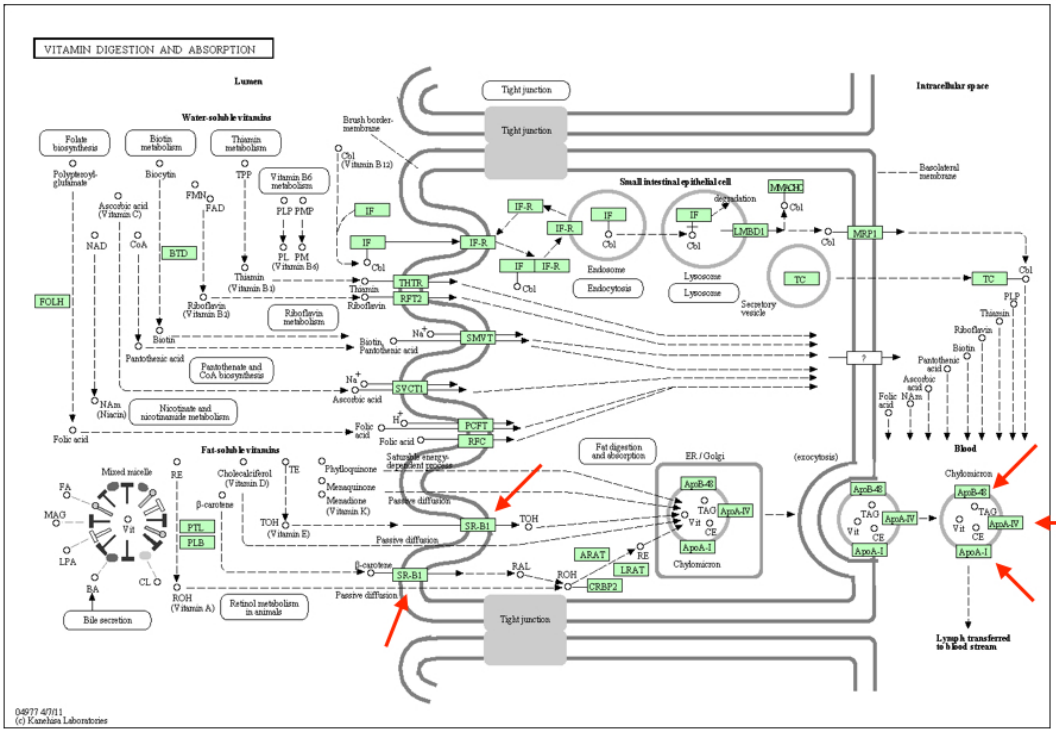

Figure 4.5: Vitamin digestion and absorption pathway. Red arrows indicate the proteins that are found to be elevated in MS subtypes.

In the present study, the increase in the NOTCH2, EP300, PSEN1, DTX1, JAG2 proteins show that NOTCH signaling pathway (Figure 4.6) was also upregulated in the PMS and RRMS clinical subtypes but not in CIS groups. Further detailed analysis of individual protein players of Notch, such as the NOTCH2, JAG2 and PIK proteins, shows that PMS patients appear to have higher expressed levels than RRMS types that is in accordance with the disease severity reflecting neurodegeneration level (2.2 fold, $p<0,05$).

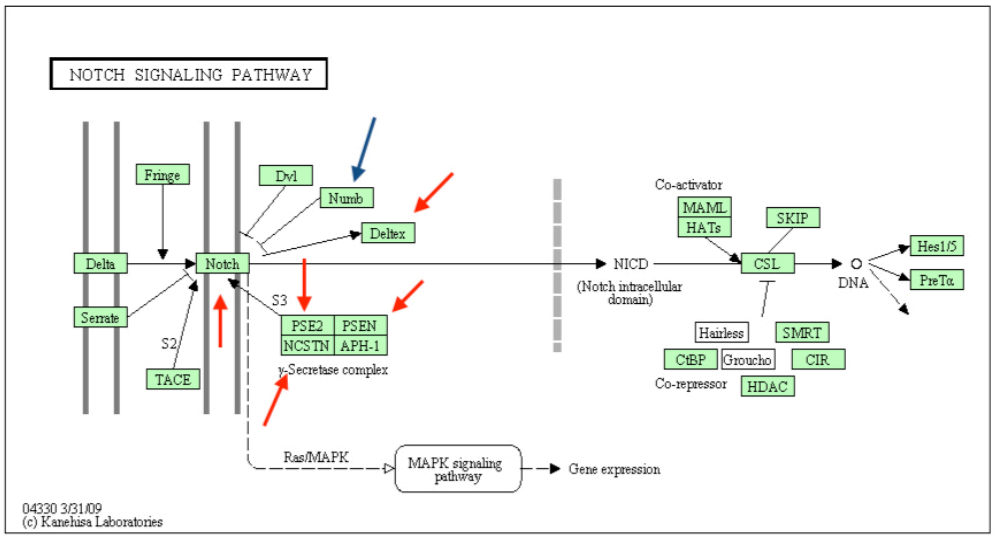

Figure 4.6: Notch signaling pathway. Red arrows indicate the proteins that are found to be elevated, blue arrows decreased proteins, in MS subtypes.

However RRMS clinic subtype patients also show an increase in the ACTB, KRT18, FYN, TUBA4A, NCL, CTNNB1 proteins which are a part of pathogenic E.coli infection pathway, which induce the mobilization of inflammatory cells (Figure 4.7). These findings elucidate for the first time the critical molecular pathways that are closely associated with the inflammatory and neurodegenerative mechanisms in MS. The mentioned pathways are in accordance with the current understanding that inflammatory processes are more activated in the earlier phases of the disease (CIS and RRMS clinical subtypes), more than the progressive forms.

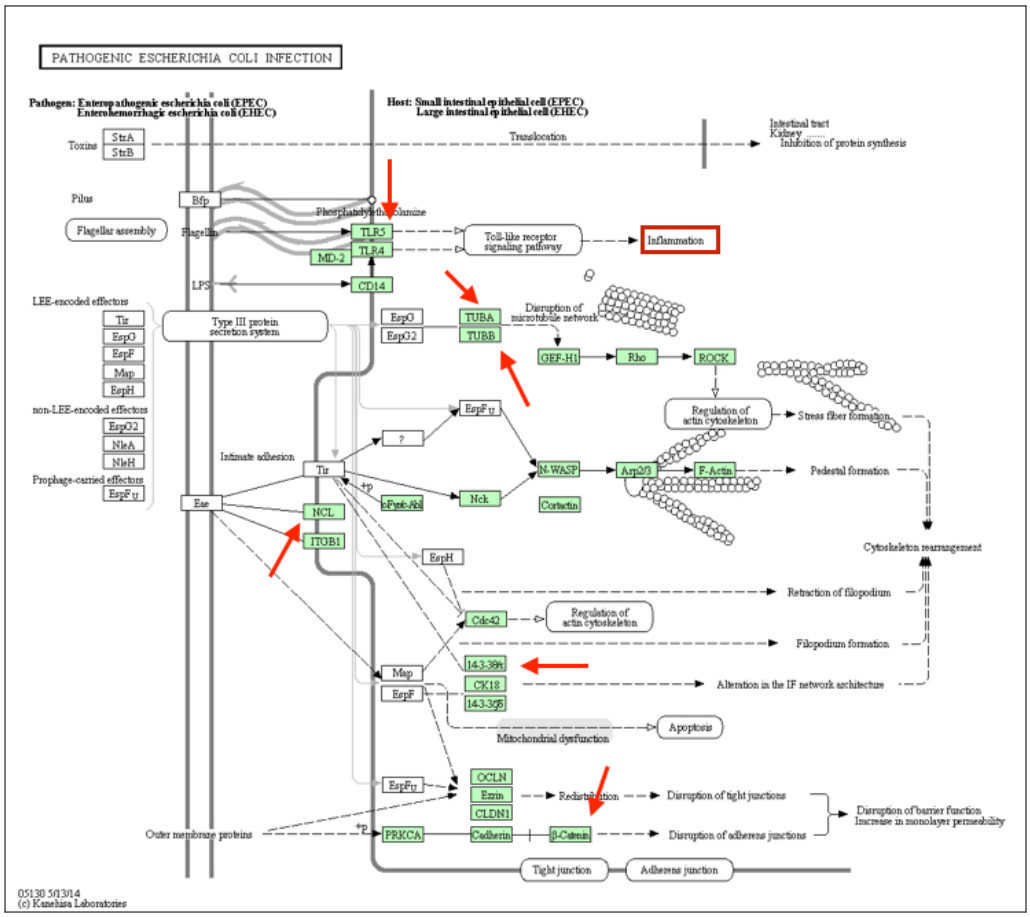

Figure 4.7: Pathogenic E. coli infection pathway. Red arrows indicate the proteins that are found to be elevated, in MS subtypes.

4.4 The Importance Of Renin-Angiotensin System Pathway And Complement And Coagulation System Pathway In Multiple Sclerosis

The role of RAS and CCC involvements in MS derives from several indirect studies of Experimental Autoimmune Encephalomyelitis (EAE) models and directly from post-mortem tissues of MS patients (Stegbauer et al., 2009). Treatments with various renin inhibitors resulted with significantly ameliorated course of EAE in rats and up-regulated RAS proteins were observed in brain lesions of MS patients. Similarly in vivo administration of coagulation cascade inhibitors such as ACE inhibitors reduced the clinical severity EAE model, supporting the view that the blockade of the coagulation cascade would be a beneficial approach for the treatment of MS (Satoh et al., 2009). Also a recent proteomics study of MS lesion-specific proteome profiling showed a pivotal role of coagulation cascade proteins in chronic active demyelination (Han et al., 2008).

Our study indicates the importance of RAS and CCC pathways in CSF samples of clinically different MS patients. Although each clinical subtype shows specific pathways compared to controls, all subtypes commonly showed increased RAS activity. The results of this study for the first time confirm through proteomic and bioinformatics approaches previous findings from EAE models and MS postmortem studies (Stegbauer et al., 2009, Wosik et al., 2007, Platten et al., 2009).

4.5 The Importance Of Vitamin Metabolism In Multiple Sclerosis

The epidemiological evidence suggest that exposure to sunlight may play a protective role in autoimmune diseases including MS, insulin-dependent diabetes mellitus, and rheumatoid arthritis (Ponsonby et al., 2005). Among many, vitamin D is the most prominent and studied one. Genetic factors linked to vitamin D receptor genes and vitamin D – binding proteins have been extensively evaluated and contradictory results have been obtained from different populations (VanAmerongen et al., 2004, Disanto et al., 2011, Cantorna, 2000). However CSF findings of vitamin metabolism is still missing due to the lack of specific proteomic studies. Our results revealed the increased expression of vitamin-metabolism related molecules in CSF of RRMS and PMS patients indicating the significantly increased activity of vitamin digestion and absorption pathway. Apolipoprotein family proteins that are part of the vitamin metabolism (APOA4, APOE, APOB) were significantly increased in the CSF of RRMS and P patients. Further subgroup analysis, comparing both CIS groups (MSs-CIS & SA-MS), with the CDMS in the 31.7 months of follow-up, showed an increased vitamin absorption and digestion related protein expression in CIS patients. ELISA studies with Vitamin D Binding Protein (VDBP) also showed the decreased VDBP level in CIS patients indicating the less amount of vitamin D metabolism in CIS subgroup ($p<0.001$). However, the pathogenetic and theraupetic implications of these findings need to be further explored.

4.6 The Importance Of Salt Metabolism And Its Molecular Basis

Possible effects of salt intake on the development of neuroinflammation has recently been investigated which might link the observed increase in the incidence of multiple sclerosis (MS) and other autoimmune diseases over the past 50 years (Kleinewietfeld et al., 2013, Wu et al., 2013, Yosef et al., 2013, Paling et al., 2013). A current study

by Kleinewietfeld et al. revealed that increased dietary salt intake might represent an environmental risk factor for the development of autoimmune diseases through the induction of pathogenic TH17 cells. In this study, TH17 cells from the EAE mice generated under high-salt conditions display a highly pathogenic and stable phenotype characterized by the upregulation of the pro-inflammatory cytokines GM-CSF, TNF-a and IL-2. Moreover, mice fed with a high-salt diet develop a more severe form of EAE, in line with augmented central nervous system infiltrating and peripherally induced antigen-specific TH17 cells (Kleinewietfeld et al., 2013). In our study, aldosterone regulated sodium reabsorption pathway was significantly affected in CIS patients ($p = 1{,}78x10^{-04}$) compared to the control group. Pathway associated genes found in subnetworks for this pathway were, Phosphatidylinositol-4,5-Bisphosphate 3-Kinase, Catalytic Subunit Beta (PIK3C1), Phosphoinositide-3-Kinase, Regulatory Subunit 1 (Alpha) (PIK3R1), Phosphoinositide-3-Kinase, Regulatory Subunit 2 (Beta) (PIK3R2) and insulin receptor (INSR) proteins (Figure 4.8). Expression of all these proteins were increased in the CIS patients compared to control group. Our results indicate that in the CIS subtype, which represents the first clinical expression of the disease and also the initial phase of other forms, exhibited an upregulation in the reabsorption of salt pathway. Furthermore subgroup analysis of CIS patients also indicated the importance of this pathway in CIS (MSs-CIS and SA-MS) patients who had not shown conversion to RRMS in a mean follow of 30.6 months. These results may signify the importance of increased salt mechanisms in the conversion to the clinically definite MS subtype.

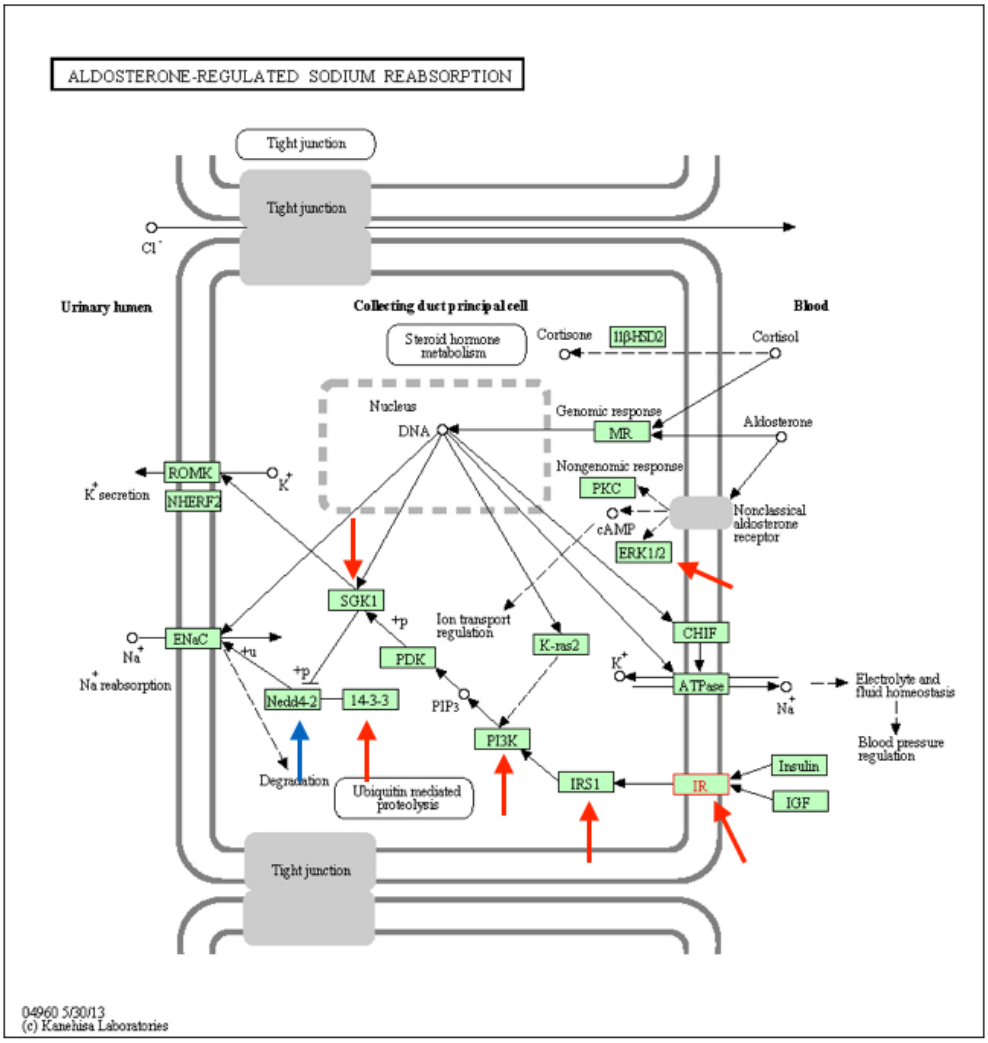

Figure 4.8: Aldosterone regulated sodium reabsorption pathway. Red arrows indicate the proteins that are found to be elevated, in MS subtypes.

4.7 Important Molecular Pathways In Clinical Conversion From CIS To Multiple Sclerosis

The prognosis of patients with CIS are unpredictable, as it is not obvious whether the affected patients will experience further clinical symptoms, and then by definition, go on to be diagnosed with CDMS likely because of the activation or inhibition of different pathological mechanisms in the affected individual with CIS presentation (Miller et al., 2005b, Miller et al., 2005a).

CIS group specific subgroup proteomic analysis indicated that conversion from CIS to CDMS depends on the activation of specific pathways. Non-converters are CIS-remaining-CIS show the increased expression of notch signaling, type II diabetes mellitus and aldosterone regulated sodium reabsorption pathways compared to SA-

MS and RRMS groups. All three pathways are immune related pathways, which are functional in the T cell lineage commitment from common lymphoid precursor and T cell development, immune activation of Treg cells and increase in the expression of inflammatory cytokines. On the other hand, the sodium reabsorption pathway, which is also related with salt metabolism, was also activated in the MSs-CIS group remaining as such during the time course of the study. However, the SA-MS group, which has a higher probability to convert to CDMS and continues to show MRI activity and who by definition are diagnosed as early MS showed an intense inflammation possibly through the activation of the NOD-like receptor-signaling pathway. The CDMS group, showed the involvement of both neurodegenerative and inflammatory pathways, which are prion diseases and pathogenic *E.coli* infection pathways.

All CIS samples were diagnosed as CIS when their CSF samples were taken according to criteria of Poser (Poser and Brinar, 2001). These observations may have therapeutic implications for CIS patients regarding early long-term MS treatment decisions. Our results indicated the different molecular pathways are involved in CIS patients who will show further clinical and imaging disease activity. It may be concluded that SA-MS patients who do not convert to CDMS may not do so because of the upregulation of the sodium reabsorption pathway. This is a thought-provoking hypothesis supportive of Haffler's study (Kleinewietfeld et al., 2013) indicating the increased salt intake might represent an environmental risk factor for the development of MS through increased auto-inflammation. It is difficult to say whether the upregulation of sodium reabsorption pathway is a cause or effect of the change from CIS to MS. Our opinion at this point is that, upregulation of sodium reabsorption pathway, may induce conversion from CIS to RRMS subtype. Previous studies indicated the activation of inflammatory cells in response to higher salt intake in vitro. However this proposition requires further analysis using in vivo studies.

The statistical significance of the pathways is also established by randomization of the patient data. Pathway analysis of randomized data analysis also showed that aldosterone dependent sodium reabsorption pathway is not present in any of the analysis (Table). This pathway is found to be significantly affected only in the analysis of CIS converting to clinically definite MS patients' data. Other commonly

found pathways in randomized MS data analysis are in agreement with the pathways in MS pathophysiology.

4.8 Identification Of Genome Wide Potential Biomarkers And Comparison Of Results With Proteome Studies

In recent years, a considerable effort has been made for the discovery of genetic determinants of MS susceptibility by high-throughput technologies. However, only a small part of the studies revealed a significant heritability regions or genes or variations been elucidated. At this part of this study, genetic basis of MS was explored, conducting a linkage study in MS families comprising of affected and unaffected individuals and a genome-wide association study using unrelated cases and healthy controls, and the results were compared to with the proteome profiling that identified specific and shared pathways in MS clinical subtypes. Linkage studies revealed 2 suggestive regions on chr13 and chr22. These regions include many inflammation and neurodegeneration related regions. Among them MAB2L1 and IFNAR genes were validated as a part of other study in our group. On the other side, there are a number of other candidate genes resulted from our linkage analysis, some of which were previously implicated in MS or other inflammatory/neurological conditions. Among them, SNPs located in CLDN14, RUNX1, and LINC00598 genes showed suggestive association with MS in our GWAS results. Therefore, other candidate genes, particularly having correlation between the genetic analyses, should be carefully taken into consideration in further studies.

Correlation analyses between the GWAS results and the proteome data revealed that among the correlations, only rs7873 located on INS-IGF2 read-through, an important paralog of INS, had significant association, and it was the most significant marker resulted from the GWAS ($P = 4.39E^{-07}$). INS (insulin) gene encodes for pro-insulin, which is then post-translationally cleaved into covalently bound A chain and B chain peptides and C peptide of insulin. The main role of insulin binding to its receptor INSR (insulin receptor) is stimulation of glucose uptake through PI3K (phosphatidylinositol 3-kinase), decreasing glucose concentration in blood. Therefore, many INS variations have been previously associated with type II diabetes mellitus, which is characterized by insulin resistance, thus high blood

glucose levels due to insulin signaling defects. Insulin signaling also stimulates a number of other pathways one of which is aldosterone-regulated sodium reabsorption pathway that found to be activated in MS subtypes together with type II diabetes mellitus in proteomic study. Accordingly, as it is shown in study of Kleinewietfeld et al., increased NaCl concentration dramatically augmented the induction of murine and human Th17 cells (Kleinewietfeld et al., 2013) which is a type of T helper cells that had been found to be critical for development of EAE (Korn et al., 2009). This Th17 promoting effect of high-salt concentration for naïve CD4 cells was through phosphorylation of p38/MAPK, which then activates NFAT5 and SGK1, both of which are components of aldosterone-regulated sodium reabsorption pathway, and downstream components of INSR signaling (Kleinewietfeld et al., 2013). Therefore, variations in INS gene may be responsible for alterations in insulin receptor signaling cascade in CIS patients, causing increased Na+ uptake through SGK1 and conversion of naïve CD4 cells into Th17 cells. In this context, possible effects of sodium reabsorption pathway activation in MS will be explored in our further studies. Along with INS, other 8 genes suggestive of association from the GWAS that found to involve in affected pathways in MS subtypes, as well as correlations observed in individual proteomic data of the patients involved in the GWAS should be investigated in further studies.

5. CONCLUSION

Results of this study revealed common and subtype specific pathways for all MS clinical subtypes. Identified pathways are related with the inflammatory and degenerative characteristics of MS. Besides, the sodium digestion and reabsorption pathway, which is related to salt metabolism, appear to be critical in the conversion of CIS to clinically definite MS. Vitamin absorption and digestion pathway was also important in the disease subclinical forms. Genomic studies were also revealed two candidate locus as a suggestive association in chromosome 13 and chromosome 22. Our results are the first pathway oriented, proteomic based and genomic correlated study indicating disease related molecular pathways, and results were supportive for earlier experimental and clinical molecular studies and demands further confirmative and explorative studies.

REFERENCES

Abcam Inc. 2012. *Sandwich ELISA protocol* [Online]. Available: http://www.abcam.com/protocols/sandwich-elisa-protocol-1 [Retrieved on 22 June 2015].

Ahlgren, C., Odén, A. & Lycke, J. (2011). High nationwide prevalence of multiple sclerosis in Sweden. *Multiple Sclerosis Journal,* 17**,** 901-908.

Alban, A., David, S. O., Bjorkesten, L., Andersson, C., Sloge, E., et al. (2003). A novel experimental design for comparative two−dimensional gel analysis: Two−dimensional difference gel electrophoresis incorporating a pooled internal standard. *Proteomics,* 3**,** 36-44.

Aloisi, F., Ria, F. & Adorini, L. (2000). Regulation of T-cell responses by CNS antigen-presenting cells: different roles for microglia and astrocytes. *Immunology today,* 21**,** 141-147.

Álvarez−Lafuente, R., Heras, V. D., Bartolomé, M., García−Montojo, M. & Arroyo, R. (2006). Human Herpesvirus 6 and Multiple Sclerosis: A One−Year Follow−up Study. *Brain pathology,* 16**,** 20-27.

Ascherio, A. & Munger, K. L. (2007a). Environmental risk factors for multiple sclerosis. Part I: the role of infection. *Annals of neurology,* 61**,** 288-299.

Ascherio, A. & Munger, K. L. (2007b). Environmental risk factors for multiple sclerosis. Part II: Noninfectious factors. *Annals of neurology,* 61**,** 504-513.

Ascherio, A., Munger, K. L., White, R., Köchert, K., Simon, K. C., Polman, C. et. al.(2014). Vitamin D as an early predictor of multiple sclerosis activity and progression. *JAMA neurology,* 71**,** 306-314.

Avasarala, J. R., Wall, M. R. & Wolfe, G. M. (2005). A distinctive molecular signature of multiple sclerosis derived from MALDI-TOF/MS and serum proteomic pattern analysis. *Journal of molecular neuroscience,* 25**,** 119-125.

Avsar, T., Korkmaz, D., Tütüncü, M., Demirci, N. O., Saip, S., et. al. (2012). Protein biomarkers for multiple sclerosis: semi-quantitative analysis of cerebrospinal fluid candidate protein biomarkers in different forms of multiple sclerosis. *Multiple Sclerosis Journal,* 18**,** 1081-1091.

Bakir-Gungor, B., Baykan, B., İser, S. U., Tuncer, F. & Sezerman, O. (2013). Identifying SNP targeted pathways in partial epilepsies with genome-wide association study data. *Epilepsy research,* 105**,** 92-102.

Bakir-Gungor, B., Egemen, E. & Sezerman, O. U. (2014). PANOGA: a web server for identification of SNP-targeted pathways from genome-wide association study data. *Bioinformatics***,** btt743.

Bakir-Gungor, B. & Sezerman, O. U. (2012). Identification of SNP targeted pathways from genome-wide association study (GWAS) data. *Nature Protocol Exchange,* 10.

Baranzini, S. E., Galwey, N. W., Wang, J., Khankhanian, P., Lindberg, R., Pelletier, D., et. al. (2009). Pathway and network-based analysis of genome-wide association studies in multiple sclerosis. *Human molecular genetics,* 18**,** 2078-2090.

Berger, T., Rubner, P., Schautzer, F., Egg, R., Ulmer, et. al.. (2003). Antimyelin antibodies as a predictor of clinically definite multiple sclerosis after a first demyelinating event. *New England Journal of Medicine,* 349**,** 139-145.

Blake, M., Johnston, K., Russell-Jones, G. & Gotschlich, E. (1984). A rapid, sensitive method for detection of alkaline phosphatase-conjugated anti-antibody on Western blots. *Analytical biochemistry,* 136**,** 175-179.

Bomprezzi, R., Ringner, M., Kim, S., Bittner, M. L., Khan, J., et. al. (2003). Gene expression profile in multiple sclerosis patients and healthy controls: identifying pathways relevant to disease. *Human molecular genetics,* 12**,** 2191-2199.

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical biochemistry,* 72**,** 248-254.

Brettschneider, J., Tumani, H., Kiechle, U., Muche, R., Richards, G., et. al. (2009). IgG antibodies against measles, rubella, and varicella zoster virus predict conversion to multiple sclerosis in clinically isolated syndrome. *PLoS One,* 4**,** e7638.

Brück, W. (2005). The pathology of multiple sclerosis is the result of focal inflammatory demyelination with axonal damage. *Journal of neurology,* 252**,** v3-v9.

Bushnik, T. (2011). Expanded Disability Status Scale. *Encyclopedia of Clinical Neuropsychology.* Springer.

Calabrese, M., Magliozzi, R., Ciccarelli, O., Geurts, J. J., Reynolds, R. et. al. (2015). Exploring the origins of grey matter damage in multiple sclerosis. *Nat Rev Neurosci,* 16**,** 147-58.

Cannella, B. & Raine, C. S. (1995). The adhesion molecule and cytokine profile of multiple sclerosis lesions. *Annals of neurology,* 37**,** 424-435.

Cantorna, M. T. (2000). Vitamin D and autoimmunity: is vitamin D status an environmental factor affecting autoimmune disease prevalence? *Proceedings of the society for Experimental Biology and Medicine,* 223**,** 230-233.

Capello, E. & Mancardi, G. (2004). Marburg type and Balo's concentric sclerosis: rare and acute variants of multiple sclerosis. *Neurological Sciences,* 25**,** s361-s363.

Caprioli, R. M., Farmer, T. B. & Gile, J. (1997). Molecular imaging of biological samples: localization of peptides and proteins using MALDI-TOF MS. *Analytical chemistry,* 69**,** 4751-4760.

Charcot, J.-M. (1868). Histologie de la sclerose en plaques. 1, 221-226.

Chiasserini, D., Di Filippo, M., Candeliere, A., Susta, F., Orvietani, et. al. (2008). CSF proteome analysis in multiple sclerosis patients by two−dimensional electrophoresis. *European Journal of Neurology,* 15**,** 998-1001.

Chun, R. F. (2012). New perspectives on the vitamin D binding protein. *Cell biochemistry and function,* 30**,** 445-456.

Coelho, T. (1996). Familial amyloid polyneuropathy: new developments in genetics and treatment. *Current opinion in neurology,* 9**,** 355-359.

Comi, G., Filippi, M., Barkhof, F., Durelli, L., Edan, G., et. al. (2001). Effect of early interferon treatment on conversion to definite multiple sclerosis: a randomised study. *The Lancet,* 357**,** 1576-1582.

Compston, A. & Coles, A. (2002). Multiple sclerosis. *Lancet,* 359**,** 1221-31.

Consortium, I. M. S. G. (2013). Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. *Nature genetics,* 45**,** 1353-1360.

Çelik, Y., Birgili, Ö., Kiyat, A., Güldiken, B., Özkan, H., et. al. (2011). Prevalence of multiple sclerosis in the metropolitan area of Edirne city, Turkey. *Balkan Medical Journal,* 2011.

D'aguanno, S., Barassi, A., Lupisella, S., Melzi D'eril, G., Del Boccio, P., et. al. 2008). Differential cerebro spinal fluid proteome investigation of Leber hereditary optic neuropathy (LHON) and multiple sclerosis. *Journal of neuroimmunology,* 193**,** 156-160.

Dalton, C. M., Brex, P. A., Miszkiel, K. A., Hickman, S. J., Macmanus, D. G., et. al. (2002). Application of the new McDonald criteria to patients with clinically isolated syndromes suggestive of multiple sclerosis. *Annals of neurology,* 52**,** 47-53.

De Jager, P. L., Jia, X., Wang, J., De Bakker, P. I., Ottoboni, L., et. al. (2009). Meta-analysis of genome scans and replication identify CD6, IRF8 and TNFRSF1A as new multiple sclerosis susceptibility loci. *Nature genetics,* 41**,** 776-782.

De Stefano, N. & Siva, A. (2013). The radiologically isolated syndrome dilemma: just an incidental radiological finding or presymptomatic multiple sclerosis? *Multiple Sclerosis Journal,* 19**,** 257-258.

Di Pauli, F., Reindl, M., Ehling, R., Schautzer, F., Gneiss, et. al. (2008). Smoking is a risk factor for early conversion to clinically definite multiple sclerosis. *Multiple sclerosis*.

Disanto, G., Ramagopalan, S. V., Para, A. E. & Handunnetthi, L. (2011). The emerging role of vitamin D binding protein in multiple sclerosis. *Journal of neurology,* 258**,** 353-358.

Dumont, D., Noben, J. P., Raus, J., Stinissen, P. & Robben, J. (2004). Proteomic analysis of cerebrospinal fluid from multiple sclerosis patients. *Proteomics,* 4**,** 2117-2124.

Dutta, R. & Trapp, B. D. (2012). Gene expression profiling in multiple sclerosis brain. *Neurobiology of disease,* 45**,** 108-114.

Eaton, W. W., Rose, N. R., Kalaydjian, A., Pedersen, M. G. & Mortensen, P. B. (2007). Epidemiology of autoimmune diseases in Denmark. *Journal of autoimmunity,* 29**,** 1-9.

Ebers, G., Sadovnick, A., Dyment, D., Yee, I., Willer, C. et. al. (2004). Parent-of-origin effect in multiple sclerosis: observations in half-siblings. *The Lancet,* 363**,** 1773-1774.

Ebers, G., Sadovnick, A. & Risch, N. (1995). A genetic basis for familial aggregation in multiple sclerosis.

Engvall, E. & Perlmann, P. (1971). Enzyme-linked immunosorbent assay (ELISA) quantitative assay of immunoglobulin G. *Immunochemistry,* 8**,** 871-874.

Federation, M. I. 2013. *Atlas of MS* [Online]. Available: http://www.msif.org/about-us/advocacy/atlas/. (Retrieved on 01.06.2015)

Fernandez, O., Fernandez, V., Martinez‐Cabrera, V., Mayorga, C., Alonso, A., Leon, A., et. al. (2008). Multiple sclerosis in Gypsies from southern Spain: prevalence, mitochondrial DNA haplogroups and HLA class II association. *Tissue Antigens,* 71**,** 426-433.

Freedman, M. S., Thompson, E. J., Deisenhammer, F., Giovannoni, G., Grimsley, et. al. (2005). Recommended standard of cerebrospinal fluid analysis in the diagnosis of multiple sclerosis: a consensus statement. *Archives of neurology,* 62**,** 865-870.

Fromont, A., Binquet, C., Sauleau, E., Fournel, I., Despalins, et. al. (2012). National estimate of multiple sclerosis incidence in France (2001–2007). *Multiple Sclerosis Journal,* 18**,** 1108-1115.

Gerdes, L. U. (2003). The common polymorphism of apolipoprotein E: geographical aspects and new pathophysiological relations. *Clinical chemistry and laboratory medicine,* 41**,** 628-631.

Goh, K.-I., Cusick, M. E., Valle, D., Childs, B., Vidal, M. et. al. (2007). The human disease network. *Proceedings of the National Academy of Sciences,* 104**,** 8685-8690.

Goldenberg, M. M. (2012). Multiple sclerosis review. *Pharmacy and Therapeutics,* 37**,** 175.

Gourraud, P. A., Harbo, H. F., Hauser, S. L. & Baranzini, S. E. (2012). The genetics of multiple sclerosis: an up‐to‐date review. *Immunological reviews,* 248**,** 87-103.

Granberg, T., Martola, J., Kristoffersen-Wiberg, M., Aspelin, P. & Fredrikson, S. (2012). Radiologically isolated syndrome–incidental magnetic resonance imaging findings suggestive of multiple sclerosis, a systematic review. *Multiple Sclerosis Journal***,** 1352458512451943.

Hafler, D. A. & Weiner, H. L. (1987). In vivo labeling of blood T cells: rapid traffic into cerebrospinal fluid in multiple sclerosis. *Annals of neurology,* 22**,** 89-93.

Hammack, B., Fung, K. Y., Hunsucker, S., Duncan, M., Burgoon, M., et. al. (2004). Proteomic analysis of multiple sclerosis cerebrospinal fluid. *Multiple Sclerosis,* 10**,** 245-260.

Han, M. H., Hwang, S.-I., Roy, D. B., Lundgren, D. H., Price, et. al. (2008). Proteomic analysis of active multiple sclerosis lesions reveals therapeutic targets. *Nature,* 451**,** 1076-1081.

Hannoun, S., Bagory, M., Durand-Dubief, F., Ibarrola, D., Comte, J.-C., Confavreux, C., et. al. (2012). Correlation of diffusion and metabolic alterations in different clinical forms of multiple sclerosis. *PloS one,* 7**,** e32525.

Hawkes, C., Giovannoni, G., Keir, G., Cunnington, M. & Thompson, E. (2006). Seroprevalence of herpes simplex virus type 2 in multiple sclerosis. *Acta neurologica scandinavica,* 114**,** 363-367.

Hernán, M. A., Jick, S. S., Logroscino, G., Olek, M. J., Ascherio, A. et. al. (2005). Cigarette smoking and the progression of multiple sclerosis. *Brain,* 128**,** 1461-1465.

Hernan, M. A., Oleky, M. J. & Ascherio, A. (2001). Cigarette smoking and incidence of multiple sclerosis. *American journal of epidemiology,* 154**,** 69-74.

Higuera, A. P. a. V. 2015. *Multiple Sclerosis by the Numbers: Facts, Statistics, and You* [Online]. Available: http://www.healthline.com/health/multiple-sclerosis/facts-statistics-infographic.

Hoffmann, K. & Lindner, T. H. (2005). easyLINKAGE-Plus—automated linkage analyses using large-scale SNP data. *Bioinformatics,* 21**,** 3565-3567.

Huang, Y., Weisgraber, K. H., Mucke, L. & Mahley, R. W. (2004). Apolipoprotein E. *Journal of Molecular Neuroscience,* 23**,** 189-204.

Ideker, T., Ozier, O., Schwikowski, B. & Siegel, A. F. (2002). Discovering regulatory and signalling circuits in molecular interaction networks. *Bioinformatics,* 18**,** S233-S240.

Jacobson, D. R., Pastore, R. D., Yaghoubian, R., Kane, I., Gallo, G., et. al. (1997). Variant-sequence transthyretin (isoleucine 122) in late-onset cardiac amyloidosis in black Americans. *New England Journal of Medicine,* 336**,** 466-473.

Jadidi−Niaragh, F. & Mirshafiey, A. (2011). Th17 cell, the new player of neuroinflammatory process in multiple sclerosis. *Scandinavian journal of immunology,* 74**,** 1-13.

James, K., Merriman, J., Gray, R., Duncan, L. & Herd, R. (1980). Serum alpha 2-macroglobulin levels in diabetes. *Journal of clinical pathology,* 33**,** 163-166.

Jaros, J. A., Guest, P. C., Bahn, S. & Martins-De-Souza, D. (2013). Affinity depletion of plasma and serum for mass spectrometry-based proteome analysis. *Proteomics for Biomarker Discovery.* Springer.

Jiang, L., He, L. & Fountoulakis, M. (2004). Comparison of protein precipitation methods for sample preparation prior to proteomic analysis. *Journal of Chromatography A,* 1023**,** 317-320.

John, G. R., Shankar, S. L., Shafit-Zagardo, B., Massimi, A., Lee, S. C., et. al. (2002). Multiple sclerosis: re-expression of a developmental pathway that restricts oligodendrocyte maturation. *Nature medicine,* 8**,** 1115-1121.

Kanoh, Y., Ohtani, N., Mashiko, T., Ohtani, S., Nishikawa, T., et. al. (2000). Levels of alpha 2 macroglobulin can predict bone metastases in prostate cancer. *Anticancer research,* 21**,** 551-556.

Karaarslan, E., Altintas, A., Senol, U., Yeni, N., Dincer, A., et. al. (2001). Balo's concentric sclerosis: clinical and radiologic features of five cases. *American journal of neuroradiology,* 22**,** 1362-1367.

Kebir, H., Kreymborg, K., Ifergan, I., Dodelet-Devillers, A., Cayrol, R., et. al. (2007). Human TH17 lymphocytes promote blood-brain barrier disruption and central nervous system inflammation. *Nature medicine,* 13**,** 1173-1175.

Kenealy, S. J., Pericak-Vance, M. A. & Haines, J. L. (2003). The genetic epidemiology of multiple sclerosis. *Journal of neuroimmunology,* 143**,** 7-12.

Khademi, M., Illés, Z., Gielen, A. W., Marta, M., Takazawa, N., Baecher-Allan, C., et. al. (2004). T Cell Ig-and mucin-domain-containing molecule-3 (TIM-3) and TIM-1 molecules are differentially expressed on human Th1 and Th2 cells and in cerebrospinal fluid-derived mononuclear cells in multiple sclerosis. *The Journal of Immunology,* 172**,** 7169-7176.

Khanapure, S. P., Garvey, D. S., Janero, D. R. & Gordon Letts, L. (2007). Eicosanoids in inflammation: biosynthesis, pharmacology, and therapeutic frontiers. *Current topics in medicinal chemistry,* 7**,** 311-340.

Kidd, P. (2003). Th1/Th2 balance: the hypothesis, its limitations, and implications for health and disease. *Alternative Medicine Review,* 8**,** 223-246.

Kivisäkk, P., Matusevicius, D., He, B., Söderström, M., Fredrikson, S. et. al. (1998). IL-15 mRNA expression is up-regulated in blood and cerebrospinal fluid mononuclear cells in multiple sclerosis (MS). *Clinical and experimental immunology,* 111**,** 193-197.

Kleinewietfeld, M., Manzel, A., Titze, J., Kvakan, H., Yosef, N., et. al. (2013). Sodium chloride drives autoimmune disease by the induction of pathogenic TH17 cells. *Nature,* 496**,** 518-522.

Korn, T., Bettelli, E., Oukka, M. & Kuchroo, V. K. (2009). IL-17 and Th17 Cells. *Annual review of immunology,* 27**,** 485-517.

Koutroukides, T. A., Guest, P. C., Leweke, F. M., Bailey, D., Rahmoune, H., et. al. (2011). Characterization of the human serum depletome by label−free shotgun proteomics. *Journal of separation science,* 34**,** 1621-1626.

Krakowka, S., Miele, J. A., Mathes, L. E. & Metzler, A. E. (1983). Antibody responses to measles virus and canine distemper virus in multiple sclerosis. *Annals of neurology,* 14**,** 533-538.

Kurtzke, J. F. (1993). Epidemiologic evidence for multiple sclerosis as an infection. *Clinical microbiology reviews,* 6**,** 382.

Larsen, P. D. & Delallo, L. J. (1987). Cerebrospinal fluid transthyretin in multiple sclerosis. *Neurology,* 37**,** 1262-1262.

Lassmann, H., Brück, W. & Lucchinetti, C. (2001). Heterogeneity of multiple sclerosis pathogenesis: implications for diagnosis and therapy. *Trends in molecular medicine,* 7**,** 115-121.

Lassmann, H. & Van Horssen, J. (2011). The molecular basis of neurodegeneration in multiple sclerosis. *FEBS letters,* 585**,** 3715-3723.

Lehmensiek, V., Süssmuth, S., Tauscher, G., Brettschneider, J., Felk, S., et. al. (2007). Cerebrospinal fluid proteome profile in multiple sclerosis. *Multiple sclerosis,* 13**,** 840-849.

Lennon, V. A., Wingerchuk, D. M., Kryzer, T. J., Pittock, S. J., Lucchinetti, C. et. al. (2004). A serum autoantibody marker of neuromyelitis optica: distinction from multiple sclerosis. *The Lancet,* 364**,** 2106-2112.

Levin, L. I., Munger, K. L., O'reilly, E. J., Falk, K. I. & Ascherio, A. (2010). Primary infection with the Epstein−Barr virus and risk of multiple sclerosis. *Annals of neurology,* 67**,** 824-830.

Lincoln, M. R., Montpetit, A., Cader, M. Z., Saarela, J., Dyment, D. A., et. al. (2005). A predominant role for the HLA class II region in the association of the MHC region with multiple sclerosis. *Nature genetics,* 37**,** 1108-1112.

Liu, X. D., Zeng, B. F., Xu, J. G., Zhu, H. B. et. al.. (2006). Proteomic analysis of the cerebrospinal fluid of patients with lumbar disk herniation. *Proteomics,* 6**,** 1019-1028.

Lonergan, R., Kinsella, K., Fitzpatrick, P., Brady, J., Murray, B., et. al. (2011). Multiple sclerosis prevalence in Ireland: relationship to vitamin D status and HLA genotype. *Journal of Neurology, Neurosurgery & Psychiatry***,** jnnp. 2010.220988.

Lucchinetti, C., Bruck, W., Parisi, J., Scheithauer, B., Rodriguez, M. et. al. (2000). Heterogeneity of multiple sclerosis lesions: implications for the pathogenesis of demyelination. *Annals of neurology,* 47**,** 707-717.

Lundmark, F., Duvefelt, K., Iacobaeus, E., Kockum, I., Wallström, E., et. al. (2007). Variation in interleukin 7 receptor α chain (IL7R) influences risk of multiple sclerosis. *Nature genetics,* 39**,** 1108-1113.

Marrosu, M. G., Murru, M. R., Costa, G., Murru, R., et. al. (1998). DRB1-DQA1-DQB1 loci and multiple sclerosis predisposition in the Sardinian population. *Human molecular genetics,* 7**,** 1235-1237.

Martin, R., Mcfarland, H. F. & Boggs, J. (1995). Immunological aspects of experimental allergic encephalomyelitis and multiple sclerosis. *Critical reviews in clinical laboratory sciences,* 32**,** 121-182.

Martino, G., Adorini, L., Rieckmann, P., Hillert, J., Kallmann, B., et. al. (2002). Inflammation in multiple sclerosis: the good, the bad, and the complex. *The Lancet Neurology,* 1**,** 499-509.

Masjuan, J., Alvarez-Cermeno, J., Garcia-Barragan, N., Díaz-Sánchez, M., Espiño, M., et. al. (2006). Clinically isolated syndromes A new oligoclonal band test accurately predicts conversion to MS. *Neurology,* 66**,** 576-578.

Mcdonald, W. I., Compston, A., Edan, G., Goodkin, D., Hartung, et. al. (2001). Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. *Annals of neurology,* 50**,** 121-127.

Mcfarland, H. F. & Martin, R. (2007). Multiple sclerosis: a complicated picture of autoimmunity. *Nature immunology,* 8**,** 913-919.

Mikaeloff, Y., Caridade, G., Tardieu, M. & Suissa, S. (2007). Parental smoking at home and the risk of childhood-onset multiple sclerosis in children. *Brain,* 130**,** 2589-2595.

Millefiorini, E., Cortese, A., Di Rezze, S., Barletta, G., Bellantonio, et. al. (2010). The prevalence of multiple sclerosis in central Italy. *Multiple sclerosis*.

Miller, D., Barkhof, F., Montalban, X., Thompson, A. & Filippi, M. (2005a). Clinically isolated syndromes suggestive of multiple sclerosis, part 2: non-conventional MRI, recovery processes, and management. *The Lancet Neurology,* 4**,** 341-348.

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

Miller, D. H., Chard, D. T. & Ciccarelli, O. (2012). Clinically isolated syndromes. *The Lancet Neurology,* 11**,** 157-169.

Miller, D. H. & Leary, S. M. (2007). Primary-progressive multiple sclerosis. *The Lancet Neurology,* 6**,** 903-912.

Minagar, A. & Alexander, J. S. (2003). Blood-brain barrier disruption in multiple sclerosis. *Multiple sclerosis,* 9**,** 540-549.

Mycko, M. P., Papoian, R., Boschert, U., Raine, C. S. & Selmaj, K. W. (2004). Microarray gene expression profiling of chronic active and inactive lesions in multiple sclerosis. *Clinical neurology and neurosurgery,* 106**,** 223-229.

Nielsen, N. M., Westergaard, T., Rostgaard, K., Frisch, M., Hjalgrim, H., et. al. (2005). Familial risk of multiple sclerosis: a nationwide cohort study. *American journal of epidemiology,* 162**,** 774-778.

Obermeier, B., Daneman, R. & Ransohoff, R. M. (2013). Development, maintenance and disruption of the blood-brain barrier. *Nature medicine,* 19**,** 1584-1596.

Oksenberg, J. R., Baranzini, S. E., Sawcer, S. & Hauser, S. L. (2008). The genetics of multiple sclerosis: SNPs to pathways to pathogenesis. *Nature Reviews Genetics,* 9**,** 516-526.

Olmarker, K., Nordborg, C., Larsson, K. & Rydevik, B. (1996). Ultrastructural changes in spinal nerve roots induced by autologous nucleus pulposus. *Spine,* 21**,** 411-414.

Ottervald, J., Franzén, B., Nilsson, K., Andersson, L. I., Khademi, et. al. (2010). Multiple sclerosis: Identification and clinical evaluation of novel CSF biomarkers. *Journal of proteomics,* 73**,** 1117-1132.

Paling, D., Solanky, B. S., Riemer, F., Tozer, D. J., Wheeler-Kingshott, C. A., et. al. (2013). Sodium accumulation is associated with disability and a progressive course in multiple sclerosis. *Brain,* 136**,** 2305-2317.

Peterson, L. K. & Fujinami, R. S. (2007). Inflammation, demyelination, neurodegeneration and neuroprotection in the pathogenesis of multiple sclerosis. *Journal of neuroimmunology,* 184**,** 37-44.

Platten, M., Youssef, S., Hur, E. M., Ho, P. P., Han, M. H., Lanz, T. et. al. (2009). Blocking angiotensin-converting enzyme induces potent regulatory T cells and modulates TH1-and TH17-mediated autoimmunity. *Proceedings of the National Academy of Sciences,* 106**,** 14948-14953.

Polman, C. H., Reingold, S. C., Banwell, B., Clanet, M., Cohen, J. A., et. al. (2011). Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. *Annals of neurology,* 69**,** 292-302.

Polman, C. H., Reingold, S. C., Edan, G., Filippi, M., Hartung, et. al. (2005). Diagnostic criteria for multiple sclerosis: 2005 revisions to the "McDonald Criteria". *Annals of neurology,* 58**,** 840-846.

Ponsonby, A. L., Lucas, R. & Mei, I. V. D. (2005). UVR, vitamin D and three autoimmune diseases—multiple sclerosis, type 1 diabetes, rheumatoid arthritis. *Photochemistry and photobiology,* 81**,** 1267-1275.

Poser, C. M. & Brinar, V. V. (2001). Diagnostic criteria for multiple sclerosis. *Clinical neurology and neurosurgery,* 103**,** 1-11.

Poser, C. M., Paty, D. W., Scheinberg, L., Mcdonald, W. I., Davis, F. A., Ebers, et. al. (1983). New diagnostic criteria for multiple sclerosis: guidelines for research protocols. *Annals of neurology,* 13**,** 227-231.

Poskanzer, D. C., Walker, A. M., Yonkondy, J. & Sheridan, J. L. (1976). Studies in the epidemiology of multiple sclerosis in the Orkney and Shetland Islands. *Neurology,* 26**,** 14.

Ragolia, L., Palaia, T., Frese, L., Fishbane, S. & Maesaka, J. K. (2001). Prostaglandin D2 synthase induces apoptosis in PC12 neuronal cells. *Neuroreport,* 12**,** 2623-2628.

Ramagopalan, S. V., Dyment, D. A., Cader, M. Z., Morrison, K. M., Disanto, G., et. al.. (2011). Rare variants in the CYP27B1 gene are associated with multiple sclerosis. *Annals of neurology,* 70**,** 881-886.

Ramanathan, M., Weinstock-Guttman, B., Nguyen, L. T., Badgett, D., Miller, et. al. (2001). In vivo gene expression revealed by cDNA arrays: the pattern in relapsing–remitting multiple sclerosis patients compared with normal subjects. *Journal of neuroimmunology,* 116**,** 213-219.

Reiber, H. (2001). Dynamics of brain-derived proteins in cerebrospinal fluid. *Clinica Chimica Acta,* 310**,** 173-186.

Runia, T. F., Hop, W. C., De Rijke, Y. B., Buljevac, D. & Hintzen, R. Q. (2012). Lower serum vitamin D levels are associated with a higher relapse risk in multiple sclerosis. *Neurology,* 79**,** 261-266.

Sadovnick, A., Dyment, D., Ebers, G., Risch, N. & Group, C. C. S. (1996). Evidence for genetic basis of multiple sclerosis. *The Lancet,* 347**,** 1728-1730.

Sadovnick, A. D. (2012). Genetic background of multiple sclerosis. *Autoimmunity reviews,* 11**,** 163-166.

Sadovnick, A. D. & Ebers, G. C. (1995). Genetics of multiple sclerosis. *Neurologic clinics,* 13**,** 99-118.

Sakaguchi, S. (2000). Regulatory T cells: key controllers of immunologic self-tolerance. *Cell,* 101**,** 455-458.

Saso, L., Leone, M. G., Sorrentino, C., Giacomelli, S., Silvestrini, et. al. (1998). Quantification of prostaglandin D synthetase in cerebrospinal fluid: a potential marker for brain tumor. *IUBMB Life,* 46**,** 643-656.

Satoh, J., Tabunoki, H. & Yamamura, T. (2009). Molecular network of the comprehensive multiple sclerosis brain-lesion proteome. *Multiple sclerosis,* 15**,** 531-541.

Schumacher, G. A. (1950). Multiple sclerosis. *Journal of the American Medical Association,* 143**,** 1146-1154.

Silberstein, M., Weissbrod, O., Otten, L., Tzemach, A., Anisenia, A., et. al. (2013). A system for exact and approximate genetic linkage analysis of SNP data in large pedigrees. *Bioinformatics,* 29**,** 197-205.

Simon, J. H. & Kleinschmidt-Demasters, B. K. (2008). Variants of multiple sclerosis. *Neuroimaging Clinics of North America,* 18**,** 703-716.

Simpson, S., Blizzard, L., Otahal, P., Van Der Mei, I. & Taylor, B. (2011). Latitude is significantly associated with the prevalence of multiple sclerosis: a meta-analysis. *Journal of Neurology, Neurosurgery & Psychiatry***,** jnnp. 2011.240432.

Siva, A. (2006). The spectrum of multiple sclerosis and treatment decisions. *Clinical neurology and neurosurgery,* 108**,** 333-338.

Smestad, C., Sandvik, L., Holmoy, T., Harbo, H. F. & Celius, E. G. (2008). Marked differences in prevalence of multiple sclerosis between ethnic groups in Oslo, Norway. *Journal of neurology,* 255**,** 49-55.

Sobel, E. & Lange, K. (1996). Descent graphs in pedigree analysis: applications to haplotyping, location scores, and marker-sharing statistics. *American journal of human genetics,* 58**,** 1323.

Sørensen, T. L., Tani, M., Jensen, J., Pierce, V., Lucchinetti, et. al. (1999). Expression of specific chemokines and chemokine receptors in the central nervous system of multiple sclerosis patients. *Journal of Clinical Investigation,* 103**,** 807.

Sospedra, M. & Martin, R. (2005). Immunology of multiple sclerosis*. *Annu. Rev. Immunol.,* 23**,** 683-747.

Sotelo, J., Martínez‐Palomo, A., Ordoñez, G. & Pineda, B. (2008). Varicella‐zoster virus in cerebrospinal fluid at relapses of multiple sclerosis. *Annals of neurology,* 63**,** 303-311.

Stegbauer, J., Lee, D.-H., Seubert, S., Ellrichmann, G., Manzel, A., et. al. (2009). Role of the renin-angiotensin system in autoimmune inflammation of the central nervous system. *Proceedings of the National Academy of Sciences,* 106**,** 14942-14947.

Steinman, L. (2007). A brief history of TH17, the first major revision in the TH1/TH2 hypothesis of T cell–mediated tissue damage. *Nature medicine,* 13**,** 139-145.

Traboulsee, A. L., Bernales, C. Q., Ross, J. P., Lee, J. D., Sadovnick, et. al. (2014). Genetic variants in IL2RA and IL7R affect multiple sclerosis disease risk and progression. *neurogenetics,* 15**,** 165-169.

Tullman, M. J. (2013). Overview of the epidemiology, diagnosis, and disease progression associated with multiple sclerosis. *Am J Manag Care,* 19**,** S15-20.

Tumani, H., Lehmensiek, V., Rau, D., Guttmann, I., Tauscher, et. al. (2009). CSF proteome analysis in clinically isolated syndrome (CIS): candidate markers for conversion to definite multiple sclerosis. *Neuroscience letters,* 452**,** 214-217.

Turanli, E., Avsar, T., Uygunoğlu, U., Kantarci, O. & Siva, A. (2015). Biomarkers for Phase Switches in Multiple Sclerosis. *In:* PREEDY, V. R. & PATEL, V. B. (eds.) *General Methods in Biomarker Research and their Applications.* Springer Netherlands., 122 - 144

Vanamerongen, B., Dijkstra, C., Lips, P. & Polman, C. (2004). Multiple sclerosis and vitamin D: an update. *European journal of clinical nutrition,* 58**,** 1095-1109.

Vandenbroeck, K., Alvarez, J., Swaminathan, B., Alloza, I., Matesanz, F., et. al. (2012). A cytokine gene screen uncovers SOCS1 as genetic risk factor for multiple sclerosis. *Genes and immunity,* 13**,** 21-28.

Vukusic, S. & Confavreux, C. (2003). Prognostic factors for progression of disability in the secondary progressive phase of multiple sclerosis. *Journal of the neurological sciences,* 206**,** 135-137.

Wang, W. Y., Barratt, B. J., Clayton, D. G. & Todd, J. A. (2005). Genome-wide association studies: theoretical and practical concerns. *Nature Reviews Genetics,* 6**,** 109-118.

Wender, M. (2011). Acute disseminated encephalomyelitis (ADEM). *Journal of neuroimmunology,* 231**,** 92-99.

Willis, S. N., Stadelmann, C., Rodig, S. J., Caron, T., Gattenloehner, S., et. al. (2009). Epstein–Barr virus infection is not a characteristic feature of multiple sclerosis brain. *Brain***,** awp200.

Wosik, K., Cayrol, R., Dodelet-Devillers, A., Berthelet, F., Bernard, M., et. al. (2007). Angiotensin II controls occludin function and is required for blood–brain barrier maintenance: relevance to multiple sclerosis. *The Journal of Neuroscience,* 27**,** 9032-9042.

Wu, C., Yosef, N., Thalhamer, T., Zhu, C., Xiao, et. al. (2013). Induction of pathogenic TH17 cells by inducible salt-sensing kinase SGK1. *Nature,* 496**,** 513-517.

Yosef, N., Shalek, A. K., Gaublomme, J. T., Jin, H., Lee, et. al. (2013). Dynamic regulatory network controlling TH17 cell differentiation. *Nature,* 496**,** 461-468.

Zeldenrust, S. R. & Benson, M. D. (2010). Familial and senile amyloidosis caused by transthyretin. *Protein misfolding diseases: current and emerging principles and therapies***,** 795-815.

CURRICULUM VITAE

Name Surname: Timuçin AVŞAR

Place and Date of Birth: İstanbul 30.05.1986

E-Mail: timucinavsar@gmail.com

EDUCATION:

B.Sc.: (2003 – 2007) İstanbul Technical University - Molecular Biology and Genetics Department

M.Sc.: (2007 – 2009) İstanbul Technical University - Molecular Biology, Genetics and Biotechnology Programme

PROFESSIONAL EXPERIENCE AND REWARDS:

PUBLICATIONS, PRESENTATIONS AND PATENTS ON THE THESIS:

Avsar, T., Durası, İ. M., Uygunoğlu, U., Tütüncü, M., Demirci, N. O., Saip, S., Siva A & Turanlı, E. T. (2015). *CSF Proteomics Identifies Specific and Shared Pathways for Multiple Sclerosis Clinical Subtypes*. **PLoS ONE**. e0122045. **(Research Article)**

Avsar, T., Durasi, İ., Uygunoglu, U., Tutuncu, M., Demirci, N., Saipoglu, S., ... & Turanli, E. (2015). *Proteomic Analysis in CSF Identified Subtype Specific and Shared Molecular Pathways for Multiple Sclerosis Clinical Phenotypes* (P5. 216). **Neurology**, 84(14 Supplement), P5-216 **(Proceedings)**

Eda Tahir Turanli, **Timucin Avsar**, Uğur Uygunoğlu, Orhun H. Kantarci, Aksel Siva. **(2015). General Methods in Biomarker Research and their Applications**, *Biomarkers for Phase Switches in Multiple Sclerosis*, ISBN: 978-94-007-7695-1. **(International Book Chapter)**

Timuçin Avşar, Uğur Uygunoğlu Aksel Siva ve Eda Tahir Turanlı, (2013) **Temel ve Klinik Nöroimmunoloji**, *Multipl Sklerozda Biyobelirteçlerin Tanı ve Prognozdaki Önemi*, ISBN: 3872003, 2013 **(National Book Chapter)**

Avsar T, Uygunoğlu U, Tütüncü M, Demirci ON, Saip S, Siva A, Turanli ET, **(2012)** *Proteom study in clinically different subtypes of multiple sclerosis, by using*

CSF and Serum samples: Subtype specific protein biomarker identification. **National Neurology Congress**, Antalya, Turkey, **(Oral Presentation in national meeting)**

OTHER PUBLICATIONS, PRESENTATIONS AND PATENTS :

Avsar, T., Korkmaz, D., Tütüncü, M., Demirci, N. O., Saip, S., Kamasak, M., Siva A. & Turanli, E. T. **(2012)**. *Protein biomarkers for multiple sclerosis: semi-quantitative analysis of cerebrospinal fluid candidate protein biomarkers in different forms of multiple sclerosis.* **Multiple Sclerosis Journal**, 18(8), 1081-1091.

Yener, U., **Avsar, T**., Akgün, E., Şeker, A., Bayri, Y., & Kılıç, T. **(2012)**. *Assessment of antiangiogenic effect of imatinib mesylate on vestibular schwannoma tumors using in vivo corneal angiogenesis assay: Laboratory investigation*. **Journal of neurosurgery**, 117(4), 697-704.

Özkan, A., Biçer, A., **Avşar, T**., Şeker, A., Toktaş, Z. O., Bozkurt, S. U. & Kılıç, T. **(2012).** *Temporal expression analysis of angiogenesis-related genes in brain development*. **Vascular cell**, 4(1), 1-9.

Karal-Yilmaz, O., Ozkan, A., Akgun, E., Kukut, M., Baysal, K., **Avsar, T.,** & Kilic, T. **(2013)**. *Controlled release of imatinib mesylate from PLGA microspheres inhibit craniopharyngioma mediated angiogenesis*. **Journal of Materials Science: Materials in Medicine**, *24*(1), 147-153.

Clark, V. E., Erson-Omay, E. Z., Serin, A., Yin, J., Cotney, J., Özduman, K., **Avsar T.** et. al. & Günel, M. **(2013)**. *Genomic analysis of non-NF2 meningiomas reveals mutations in TRAF7, KLF4, AKT1, and SMO*. **Science**, 339(6123), 1077-1080.

Yener, U., **Avsar, T**., Akgun, E., Seker, A., Bayri, Y., & Kilic, T. **(2013)**. *Angiogenesis RESPONSE*. **Journal of Neurosurgery**, 119(5), 1353-1353.

Kılıç, K., **Avsar, T**., Akgün, E., Özkan, A., Toktaş, Z. O., Şeker, A., & Kılıç, T. **(2013)**. *Gamma Knife Radiosurgery Inhibits Angiogenesis of Meningiomas: In Vivo Rat Corneal Assay*. **World neurosurgery**, 80(5), 598-604.

Villar, L. M., Casanova, B., Ouamara, N., Comabella, M., Jalili, F., Leppert, D., **Avsar T**. & Bar**-**Or, A. **(2014)**. *Immunoglobulin M oligoclonal bands: Biomarker of targetable inflammation in primary progressive multiple sclerosis*. **Annals of neurology**, 76(2), 231-240.

Ceylan, D., Tatarlı, N., **Avsar, T**., Arslanhan, A., Bozkurt, S. U., Bağcı, P., & Kilic, T. **(2014).** *In vivo effect of pregnancy on angiogenesis potential of arteriovenous malformation tissue samples: an experimental study*. **Journal of neurosurgical sciences**.

Kuhle, J., Disanto, G., Dobson, R., Adiutori, R., Bianchi, L., Topping, J., **Avsar T** et. al. & Martínez, A. D. **(2015)**. Conversion from clinically isolated syndrome to multiple sclerosis: a large multicentre study. **Multiple Sclerosis Journal**, 1352458514568827.

Atabay, K. D., Yildiz, M. T., **Avsar, T**., Karabay, A., & Kiliç, T. **(2015)**. Knockdown of Pin1 leads to reduced angiogenic potential and tumorigenicity in glioblastoma cells. **Oncology Letters**. Online publication.