EVALUATION OF EXOME SEQUENCING DATA & MUTATION SCREENING IN GDAP1, GJB1, MPZ, NDRG1 AND PMP22 GENES IN TURKISH CMT PATIENTS

by

Merve SIVACI

B.S., Molecular Biology and Genetics, Istanbul Technical University, 2011

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Molecular Biology and Genetics
Boğaziçi University

2013

To my family...

ACKNOWLEDGEMENTS

I would like to express my gratitude to my supervisor Prof. Esra Battaloğlu for the opportunities which she provided me and her guidance, valuable critics during this study. I would like to thank the other members of the thesis committee, Prof. Aslı Tolun and Prof. Albena Jordanova for evaluating my thesis. I also would like to express my appreciation to Prof. James R. Lupski, Davut Pehlivan and Claudia Gonzaga-Jauregui who performed exome sequencing analysis and guided me during the evaluation of the data.

I appreciate Prof. Yeşim Parman and all other clinicians for providing blood samples and clinical data of the patients. I am also grateful to patients for participating in this study.

I would deeply thank to Burçak Özeş and Kaya Akyüz for their friendship and endless support whenever I needed. I would like to express my special thanks to my lab members Cansu Kuey, Elif Begüm Gökerküçük, Kerem Yıldırım, Merve Kılınç, Alperen Erdoğan and Duygu Dağlıkoca. They not only supported me for my thesis, but also enabled me to experience one of the greatest experiences of my life which encouraged me to hope to achieve my future goals. I feel really lucky to have such great friends.

I would like to thank Aslı Uğurlu, Neslihan Zöhrap and Özden Akay for their great friendship and support. I would also like to thank other members of Lab 204; Burcu Ersoy, Emirhan Taşöz and Mahmut Can Hız for their friendships.

Last but not least, I very much thank to my dear family for their endless love, patience, support and for being there for me during this study. I always feel their confidence in me in all my way till now and I hope it will last forever.

Finally, I am grateful to TUBITAK for awarding me via the project, 112T275, during my master.

ABSTRACT

EVALUATION OF EXOME SEQUENCING DATA & MUTATION SCREENING IN GDAP1, GJB1, MPZ, NDRG1 AND PMP22 GENES IN TURKISH CMT PATIENTS

Charcot-Marie-Tooth (CMT) disease which is an inherited peripheral neuropathy is an excellent candidate to be analyzed by exome sequencing because of its locus heterogeneity and low prevalence. Protein-coding regions, 'the exome', constitute only 1% of the human genome, but 85% of the functional variations are found within these regions. Recent developments in sequencing techniques have made exome sequencing an alternative approach for disease-causing variation identification with its cost-effective feature when compared with whole genome sequencing. In the first part of this study, seventeen Turkish CMT patients were exome sequenced and the data obtained from sequencing was evaluated. Two novel variations in MFN2 gene as well as a novel splicing variant in *GJB1* were confirmed in three families. In addition, in the other patients, novel variations in DNM2, LRSAM, FGD4 and KIF1B genes were identified, but they should be validated in other family members and also in controls. Previously identified mutations in AARS and DNM2 genes were found to be causative in two other families. Besides, possible novel genes were identified but needs confirmation in other patients. The data also revealed seven variations in CMT genes observed in all seventeen patients. Six of them have been previously reported as SNPs and one was novel probably representing a common variation in our population. In the second part of the study, eight other Turkish CMT patients were screened for mutations in the genes, GDAP1, GJB1, MPZ, NDRG1 and PMP22. A novel variation in *PMP22* was identified for one patient and three previously reported SNPs were detected in *GDAP1* for another patient. The presence of the founder p.R148X mutation in *NDRG1* was also confirmed in two affected brothers. These data show the importance of clinical use of whole exome sequencing technology in CMT patients and other genetically heterogenous diseases as it allows the identification of disease-causing variants in a shorter period of time, compared to mutation screening in known CMT genes.

ÖZET

TÜRK CMT AİLELERİNDE EKZOM DİZİLEME VERİLERİNİN DEĞERLENDİRİLMESİ VE GDAP1, GJB1, MPZ, NDRG1 VE PMP22 GENLERİNDE MUTASYON TARAMASI

Kalıtımsal periferal sinir hastalığı olan Charcot-Marie-Tooth (CMT), gerek lokus çeşitliliği gerekse görülme sıklığının az olmasından dolayı ekzom dizileme analizi için oldukça uygun bir adaydır. Protein kodlayan bölgeler (ekzom), insan genomunun sadece %1'ini oluştursa da, fonksiyonel varyasyonların %85'i bu bölgelerde bulunur. DNA dizileme tekniklerindeki yeni gelişmeler, tüm genom dizilemesiyle karşılaştırıldığında ekzom dizilemeyi, hastalığa sebep olan varyasyonları bulmada tercih edilen bir yöntem kılmıştır. Bu çalışmanın ilk kısmında, on yedi Türk CMT hastasına ekzom dizileme uygulandı ve elde edilen veriler değerlendirildi. Üç ailede iki yeni *MFN2* varyasyonu ile bir yeni *GJB1* kırpılma varyantının kalıtıldığı doğrulandı. Ayrıca, DNM2, LRSAM, FGD4 ve KIF1B genlerinde yeni varyasyonlar saptandı. Ancak bu varyasyonların diğer aile bireylerinde ve sağlıklı kontrollerde doğrulanması gerekmektedir. Farklı iki ailede AARS ve DNM2 genlerinde daha önceden tanımlanmış mutasyonların hastalığa neden olduğu bulundu. Geniş CMT hasta gruplarında çalışılarak doğrulanması gerekli olan ve hastalığa sebep olabilecek muhtemel yeni genler belirlendi. Ekzom verileri arasında tüm incelenen hastalarda CMT genlerinde yedi ortak varyasyon ortaya çıkarıldı. Bunların altısının önceden rapor edilen polimorfizler olduğu belirlenirken birinin toplumumuzda sıklıkla görülen yeni bir varyasyon olabileceği öne sürüldü. Çalışmanın ikinci kısmında, sekiz CMT hastasında GDAP1, GJB1, MPZ, NDRG1 ve PMP22 genlerinde mutasyon taraması gerçekleştirildi. Bir hastada PMP22 geninde yeni bir varyasyon görüldü. Diğer bir hastada GDAP1 geninde daha önce rapor edilen üç polimorfizm bulundu. Romen kökenli iki kardeşte ise atasal p.R148X *NDRG1* mutasyonu saptandı. Bu bulgular CMT gibi heterojen hastalıklarda ekzom dizilemesinin klinik olarak kullanılmasının önemini göstermektedir. Bu yöntemle hastalığa sebep olan varyasyonların saptanması, CMT genlerinin klinik bulgulara dayanılarak sırayla taramasına oranla daha kısa bir zamanda mümkün olabilmektedir.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS iv

ABSTRACT v

ÖZET vi

LIST OF FIGURES x

LIST OF TABLES xv

LIST OF SYMBOLS xviii

LIST OF ACRONYMS / ABBREVIATIONS xx

1. INTRODUCTION 1

1.1. Charcot-Marie-Tooth (CMT) Disease 1

1.1.1. Charcot-Marie-Tooth Disease Type 1 (CMT1) 5

1.1.1.1. PMP22 5

1.1.1.2. MPZ 7

1.1.1.3. GJB1 8

1.1.2. Charcot-Marie-Tooth Disease Type 4 (CMT4) 9

1.1.2.1. GDAP1 10

1.1.2.2. NDRG1 10

1.1.2.3. FGD4 11

1.1.3. Charcot-Marie-Tooth Disease Type 2 (CMT2) 12

1.1.3.1. MFN2 13

1.1.3.2. AARS 13

1.1.3.3. DNM2 14

1.2. Exome Sequencing 15

1.2.1. Analysis of Exome Sequencing Data 16

1.2.2. Exome Sequencing in CMT 17

1.3. Genetic Testing in CMT 18

1.3.1. Phenotype/Genotype Correlation 18

2. AIM OF THE STUDY 22

3. MATERIALS 23

3.1. Patients 23

3.2. Chemicals 23
3.3. Buffers and Solutions 24
3.3.1. DNA Extraction from Peripheral Blood 24
3.3.2. Polymerase Chain Reaction (PCR) 24
3.3.3. Agarose Gel Electrophoresis 25
3.3.4. Polyacrylamide Gel Electrophoresis (PAGE) 25
3.3.5. Silver Staining 26
3.4. Fine Chemicals 26
3.4.1. Enzymes 26
3.4.2. Oligonucleotide Primers 26
3.4.3. DNA Ladders 33
3.5. Equipments 33
4. METHODS 35
4.1. DNA Extraction from Peripheral Blood 35
4.2. Quantitative Analysis of Extracted DNA 35
4.3. Mutation Screening 35
4.3.1. Mutation Analysis Using PCR-SSCP 36
4.3.1.1. Polymerase Chain Reaction (PCR) 36
4.3.1.2.Single-Strand Conformation Polymorphism (SSCP) 37
4.3.1.3. Silver Staining 38
4.3.2. Mutation Analysis Using PCR-Sanger Sequencing 38
4.3.2.1. Polymerase Chain Reaction (PCR) 39
4.3.2.2. Sanger Sequencing 39
4.3.3. Mutation Analysis by Restriction Enzyme Digestion 39
4.3.3.1. Polymerase Chain Reaction (PCR) 39
4.3.3.2. Restriction Enzyme Digestion 40
4.4. Exome Sequencing 40
5. RESULTS 43
5.1. Exome Sequencing 43
5.1.1. P114 50
5.1.2. P138 53
5.1.3. P163 56
5.1.4. P209 58

5.1.5. P212 59
5.1.6. P245 62
5.1.7. P313 64
5.1.8. P378 66
5.1.9. P851 68
5.2. Mutation Screening 74
5.2.1. Mutation Analysis of *MPZ* 74
5.2.2. Mutation Analysis of *GDAP1* 77
5.2.3. Mutation Analysis of *GJB1* 78
5.2.4. Mutation Analysis of *PMP22* 79
5.2.5. Mutation Analysis of *NDRG1* 79
6. DISCUSSION 81
6.1. Exome Sequencing 81
6.2.Mutation Screening 99
7. CONCLUSION 103
REFERENCES 104

LIST OF FIGURES

Figure 1.1. Main characteristics of CMT; pes cavus and clawed toes, claw hand, proximal and distal leg wasting. 1

Figure 1.2. Genetic classification of CMT and causative genes. 3

Figure 1.3. Proteins and intracellular pathways involved in CMT. 4

Figure 1.4. Protein structure of PMP22 with point mutations associated with CMT. 6

Figure 1.5. Protein structure of MPZ with the mutations known to cause neuropathy. 7

Figure 1.6. Proteins in the myelin sheath that cause CMT. 8

Figure 1.7. Hemichannels and protein structure of connexins. 9

Figure 1.8. Protein structure and genomic organization of GDAP1. 10

Figure 1.9. Protein structure of NDRG1 protein. 11

Figure 1.10. Protein structure of FGD4 and mutations associated with CMT. 12

Figure 1.11. Protein structure of MFN2. 13

Figure 1.12. Protein structure of AARS located in cytoplasm and its mutations associated with CMT. 14

Figure 1.13 Protein structure of DNM2 and its mutations associated with CMT. . . 15

Figure 1.14. General filtering steps of exome sequencing data. 17

Figure 5.1. Pedigree of the family P114. 50

Figure 5.2. Chromatograms showing the *SBF2* variant in patient P114 and family members. 51

Figure 5.2. Chromatograms showing the AARS variant in patient P114 and family members. 52

Figure 5.3. Pedigree of the family P138. 53

Figure 5.4. Chromatograms showing the hemizygous *GJB1* variant in patient P138. 54

Figure 5.5. Chromatograms showing the variantions in LDB3, OTOF, CNNM4, and ARHGAP4 genes in patient 138. 55

Figure 5.6. Pedigree of the family P163. 56

Figure 5.7. Chromatograms showing the variants in *DNM2* and *LRSAM1* in patient P163. 57

Figure 5.8. Pedigree of the family P209. 58

Figure 5.9. Chromatograms showing the *MFN2* variant in patient P209 and his affected son. 58

Figure 5.10. Pedigree of the family P212. 59

Figure 5.11. Chromatograms showing the absence of *AARS* variants in patient P212 and his unaffected brother. 60

Figure 5.12. Chromatograms showing the variant in GMFG, FLNA, SIGMAR1, CXorf23 genes in patient P212 and his unaffected brother. 61

Figure 5.13. Pedigree of the family P245. 62

Figure 5.14. Chromatograms showing the variants in *FGD4*, *KIF1B*, *THBS4*, *ZNF880* in patient P245. 63

Figure 5.15. Pedigree of the family P313. 64

Figure 5.16. Chromatograms showing the *MFN2* variant in patient P313 and family members. 65

Figure 5.17. Pedigree of the family P378. 66

Figure 5.18. Chromatograms showing the variants in *SBF2*, *SLC6A11*, *DLGAP3* in patient P378 and his affected and unaffected brothers. 67

Figure 5.19. Pedigree of the family P851. 68

Figure 5.20. Chromatograms showing the *DNM2* variant in patient P851 and her family members. 69

Figure 5.21. SSCP gels (8%) for all six exons of *MPZ*. 75

Figure 5.22. SSCP gels (8%) for all six exons of *GDAP1*. 77

Figure 5.23. Chromatograms showing that variants in *GDAP1* in patient P1070. . . 78

Figure 5.24. Chromatogram showing the splicing variant in *PMP22* in patient P868. 79

Figure 5.25. *TaqI* restriction enzyme digestion results for patients P983.3 and P983.4. 80

Figure 6.1. Functional groups (R) of arginine to histidine. 83

Figure 6.2. The effect of the splicing variation in *GJB1*. 84

Figure 6.3. Protein structure of CNNM4 and the location of p.E511K variation. . . 85

Figure 6.4. Protein structure of ARHGAP4 and the location of p.S76I variation. . 86

Figure 6.5. Functional groups (R) of serine to isoleucine. 86

Figure 6.6. Conservation of the amino acid at position 341 among mammalian species. 87

Figure 6.7. Protein structure of LRSAM1. 88

Figure 6.8. Conservation of the amino acid at position 112 among mammalian species. 88

Figure 6.9. Locations of MFN2 mutations linked to CMT. 89

Figure 6.10. Functional groups (R) of valine to glysine. 90

Figure 6.11. Functional groups (R) of glysine to valine. 91

Figure 6.12. Functional groups (R) of arginine to cysteine. 91

Figure 6.13. Protein structure of KIF1B. 92

Figure 6.14. Functional groups (R) of tryptophan to arginine. 93

Figure 6.15. Conservation of the amino acid at position 703 among mammalian species. 93

Figure 6.16. Functional groups (R) of asparagine to serine. 94

Figure 6.17. Functional groups (R) of glycine to serine. 94

Figure 6.18. Conservation of the amino acid at position 176 among mammalian species. 95

Figure 6.19. Conservation of the amino acid at position 358 among mammalian species.. 96

Figure 6.20. The effect of the splicing variation in *PMP2*. 101

LIST OF TABLES

Table 1.1.	Genotype/phenotype association in CMT.	19
Table 3.1.	Materials used for DNA extraction from peripheral blood.	25
Table 3.2.	Materials used for PCR.	25
Table 3.3.	Materials used for agarose gel electrophoresis.	26
Table 3.4.	Materials used for PAGE.	26
Table 3.5.	Materials used for Silver Staining.	27
Table 3.6.	Primers used for exon amplification of *GDAP1*.	28
Table 3.7.	Primers used to amplify coding region of *GJB1*.	28
Table 3.8.	Primers used for exon amplification of *MPZ*.	29
Table 3.9.	Primers used to amplify exon 7 of NDRG1 gene.	29
Table 3.10.	Primers used to amplify coding regions of *PMP22*.	30
Table 3.11.	Primers used to confirm variations identified by exome sequencing for patient P114.	30
Table 3.12.	Primers used to confirm variations selected from the exome sequencing data for patient P138.	31

Table 3.13. Primers used to confirm variations selected from the exome sequencing data for patient P163. 31

Table 3.14. Primers used to confirm the variations selected from the exome sequencing data for patients P209 and P313. 32

Table 3.15. Primers used to confirm variations selected from the exome sequencing data for patient P212. 32

Table 3.16. Primers used to confirm variations identified by exome sequencing for patient P245. 33

Table 3.17. Primers used to confirm variations selected from the exome sequencing data for patient P378. 33

Table 3.18. Primers used to confirm the variation revealed by exome sequencing for patient P851. 34

Table 3.19. Equipments used in this study. 35

Table 4.1. PCR reagents for exon amplification of GDAP1 and MPZ genes. 37

Table 4.2. PCR conditions for exon amplification of GDAP1 and MPZ genes. . . 38

Table 4.3. Reagents used to prepare 8% (v/v) polyacrylamide gels. 38

Table 4.4. PCR reagents for exon amplification of GDAP1, GJB1 and PMP22 genes. 40

Table 4.5. Restriction enzyme digestion mix. 41

Table 4.6. Reagents for PCR amplification of targeted regions in SBF2 gene of patients P114 and P378 and in DNM2 gene of patient P851. 41

Table 4.7. Reagents for PCR amplification of targeted regions in DNM2 gene of patient P163 and the GMFG and the FLNA genes of patient P212. . . . 42

Table 5.1. Clinical characteristics of seventeen CMT patients that were exome sequenced. 44

Table 5.2. Number of variations determined by filtering criteria for exome sequencing data analysis. 47

Table 5.3. Candidate variants for nine CMT patients. 48

Table 5.4. Candidate variants selected for the other eight CMT patients. 71

Table 5.5. Variations found in all seventeen CMT patients. 73

Table 5.6. Clinical characteristics of eight CMT patients screened for mutations in targeted genes. 76

Table 6.1. Variations validated in CMT patients. 98

LIST OF SYMBOLS

A	Adenine
C	Cytosine
D	Aspartic acid
E	Glutamic acid
F	Phenylalanine
G	Guanine
Gb	Gigabase
H	Histidine
I	Isoleucine
K	Lysine
kDa	Kilodalton
L	Leucine
m	Meter
M	Methionine
Mb	Megabase
mg	Milligram
min	Minute
ml	Milliliter
mM	Millimolar
N	Asparagine
ng	Nanogram
P	Proline
pmol	Picomole

Q	Glutamine
R	Arginine
s	Second
S	Serine
T	Thymine
V	Valine
v	Volume
W	Tryptophan
w	Weight
X	Stop codon
Y	Tyrosine
µg	Microgram
µl	Microliter

LIST OF ACRONYMS / ABBREVIATIONS

Aa	Amino acid
AARS	Alanyl-tRNA synthetase
AD	Autosomal dominant
$AgNO_3$	Silver nitrate
AIMP1	Aminoacyl tRNA synthetase complex-interacting multifunctional protein 1
APS	Ammonium peroxodisulphate
AR	Autosomal recessive
ARHGAP4	Rho GTPase-activating protein 4
ARS	Aminoacyl-tRNA synthetase
ATP7A	ATPase, Cu++ transporting, alpha polypeptide
BAM	Binary alignment/map
BICD2	Bicaudal D homolog 2
bp	Basepair
BWA	Burrows-Wheeler Aligner
Ca	Calcium
CBS	Cystathionine β-synthase
CC	Coil-coiled domains
CD	Cytoplasmic domain
Chr	Chromosome
CL	Cytoplasmic loop
CLIC2	Chloride intracellular channel 2
CMAP	Compound muscle action potentials
CMT	Charcot-Marie-Tooth

CNNM4	Cyclin M4
cNMP	Cyclic nucleotide monophospate
CNS	Central nervous system
Cx32	Connexin 32
CXorf23	Chromosome X open reading frame 23
dbSNP	SNP database
del	Deletion
DH	Dbl homology
DHTKD1	Dehydrogenase E1 and transketolase domain containing 1
DI	Dominant Intermediate
DLGAP3	Discs, large (Drosophila) homolog-associated protein 3
DNA	Deoxyribonucleic acid
DNM2	Dynamin 2
dNTP	Deoxyribonucleotide Triphosphate
DOB	Date of birth
DOCK9	Dedicator of cytokinesis 9
DYNC1H1	Dynein, Cytoplasmic 1, Heavy Chain 1
ECD	Extracellular domain
EGR	Epidermal growth factor
EMG	Electromyography
ER	Endoplasmic reticulum
ESE	Exonic splicing enhancer xix
FAB	F-actin binding
FGD4	FYVE, RhoGEF and PH domain-containing 4
FLNA	Filamin A, alpha
FNBP1L	Formin binding protein 1-like

GALC	Galactosylceramidase
GAP	GTPase activating
GAS3	Growth arrest specific 3
GDAP1	Ganglioside-induced differentiation-associated-protein 1
GED	GTPase effector domain
GEF	Guanine nucleotide exchange factor
GJB1	Gap junction protein, beta 1
GMFG	Glia maturation factor, gamma
GST	Glutathione-S-transferases
GTP	Guanosine triphosphate
GWAS	Genome-wide association studies
HapMap	Haplotype mapping
HCl	Hydrochloric acid
HD	Hydrophobic domain
Het	Heterozygous
HGMD	Human gene mutation database
HK1	Hexokinase 1
HMSN	Hereditary Motor and Sensory Neuropathy
HMSNL	Hereditary Motor and Sensory Neuropathy - Lom
HNPP	Hereditary Neuropathy with Liability to Pressure Palsy
Hom	Homozygous
HR	Heptad repeat
IP_3	Inositol trisphosphate
KCl	Potassium chloride
KHCO 3	Potassium bicarbonate
KIAA0415	Adaptor-related protein complex 5, zeta 1 subunit

KIF	Kinesin superfamily
KIF1B	Kinesin motor protein 1B
LDB3	LIM domain binding 3
LR	Leucine-rich repeats
LRSAM1	Leucine rich repeat and sterile alpha motif containing 1
LRT	Likelihood ratio test
MAM	Mitochondrial associated membrane
MFN2	Mitofusion 2
$MgCl_2$	Magnesium chloride
MICAL1	Microtubule associated monoxygenase, calponin and LIM domain containing 1
MNCV	Motor nerve conduction velocity
MPZ	Myelin protein zero
mRNA	Messenger RNA
MYO10	Myosin X
MYO18A	Myosin XVIIIA
Na_2EDTA	Disodium ethylenediamine tetraacetate
$NaBH_4$	Sodium borohydride
NaCl	Sodium chloride
NaOH	Sodium hydroxide
NCBI	National center for biotechnology information
NCV	Nerve conduction velocity
NDRG1	N-myc downstream-regulated gene-1
NEFL	Neurofilament light polypeptide
NGS	Next generation sequencing
NH_4Cl	Ammonium chloride

Nt	Nucleotide
OMIM	Online Mendelian Inheritance in Man
OMM	Outer mitochondrial membrane
OTOF	Otoferlin
P2RX5	Purinergic receptor P2X, ligand-gated ion channel, 5
PAGE	Polyacrylamide Gel Electrophoresis
PCR	Polymerase chain reaction
PH	Pleckstrin homology
PHF16	PHD finger protein 16
PI	Phosphoinositide
PMP22	Peripheral myelin protein-22
PNS	Peripheral nervous system
POLG	Polymerase (DNA directed), gamma
PolyPhen	Polymorphism phenotyping
PR	Proline-rich
PRD	Proline-arginine-rich domain
PRX	Periaxin
PSAT1	Phosphoserine aminotransferase 1
RBC	Red blood cell
ref	Reference
RNA	Ribonucleic acid
ROBO3	Roundabout, axon guidance receptor, homolog 3
RUFY4	RUN and FYVE domain containing 4
SAM	Sequence alignment/map
SBF2	SET binding factor 2
SDK1	Sidekick cell adhesion molecule 1

SDS	Sodium dodecyl sulfate
SH3TC2	Src homology 3 domain and tetratricopeptide repeats-containing protein 2
SIFT	Sorting tolerant from intolerant
SIGMAR1	Sigma non-opioid intracellular receptor
SLC6A11	Solute carrier family 6 (neurotransmitter transporter, GABA), member 11
SNP	Single nucleotide polymorphism
SNV	Single nucleotide variations
SSCP	Single strand conformation polymorphism
STARD8	StAR-related lipid transfer (START) domain containing 8
SYNGR1	Synaptogyrin 1
TBE	Tris-Borate-EDTA
TE	Tris-EDTA
TEMED	Tetramethylethylenediamine
Temp	Temperature
THBS4	Thrombospondin-4
TMD	Transmembrane domain
tRNA	Transfer RNA
TRIM2	Tripartite motif containing 2
TRPV4	Transient Receptor Potential Cation Channel, Subfamily V, Member 4
UTR	Untranslated region
Var	Variant
ZFP62	ZFP62 zinc finger protein
ZnF	Zinc-finger motif of the RING type

ZNF880	Zinc finger protein 880
Zyg	Zygosity

1. INTRODUCTION

1.1. Charcot-Marie-Tooth (CMT) Disease

Charcot-Marie-Tooth (CMT) disease is the most common inherited peripheral neuropathy with a prevalence of 1 in 2500 (Skre 1974). The disease was first described in 1886 by Jean Martin Charcot, Pierre Marie (Charcot and Marie 1886) and Henry Howard Tooth (Tooth 1886). It is also known as Hereditary Motor and Sensory Neuropathy (HMSN) since it affects both motor and sensory components of the peripheral nervous system (Reilly *et al*., 2011).

CMT is a progressive disorder associated with distal muscle atrophy and weakness, decreased or absent deep tendon reflexes and foot deformities (e.g. pes cavus). Hand deformities manifest as claw hands and generally appear as the disease progresses (Figure 1.1). Distal sensory loss, decreased pain and temperature sensation are also common (Szigeti and Lupski, 2009). However, clinical features vary among affected individuals in terms of disease progression, severity and age of onset, indicating the heterogeneity of the disease (Azzedine *et al*., 2012). The disease usually does not alter lifespan, but affects the quality of life.

Figure 1.1. Main characteristics of CMT; pes cavus and clawed toes, claw hand, proximal and distal leg wasting (Reilly *et al*., 2011).

CMT is classified as demyelinating (CMT1) and axonal (CMT2) according to electrophysiological and nerve biopsy findings. Upper limb motor nerve conduction velocities (MNCV) (median and ulnar nerves) are below 38 m/s in demyelinating form. For this type, there are also histologic findings of demyelination and remyelination with onion bulb formation. Conversely, upper limb MNCV is nearly normal or higher than 38 m/s with axonal degeneration in CMT2. Additionally, in axonal CMT, there is a decrease in compound muscle action potential (CMAP) amplitudes (Jordanova *et al*., 2003). Intermediate CMT, on the other hand, is associated with MNCV values between 25 and 45 m/s. Patients who cannot be classified as either CMT1 or CMT2 are diagnosed as intermediate CMT (Davis *et al*., 1978).

The disease can also be classified according to the inheritance pattern. It can be inherited as an autosomal dominant (AD), autosomal recessive (AR) or X-linked recessive/dominant trait (Siskind *et al*., 2013).

Up to now, at least 40 causative genes have been identified and the number of the loci and genes associated with the disease is being increasing rapidly. Additionally, some of these genes are known to cause both demyelinating and axonal forms of the disease and may present AD or AR traits (Tazir *et al*., 2013) (Figure 1.2).

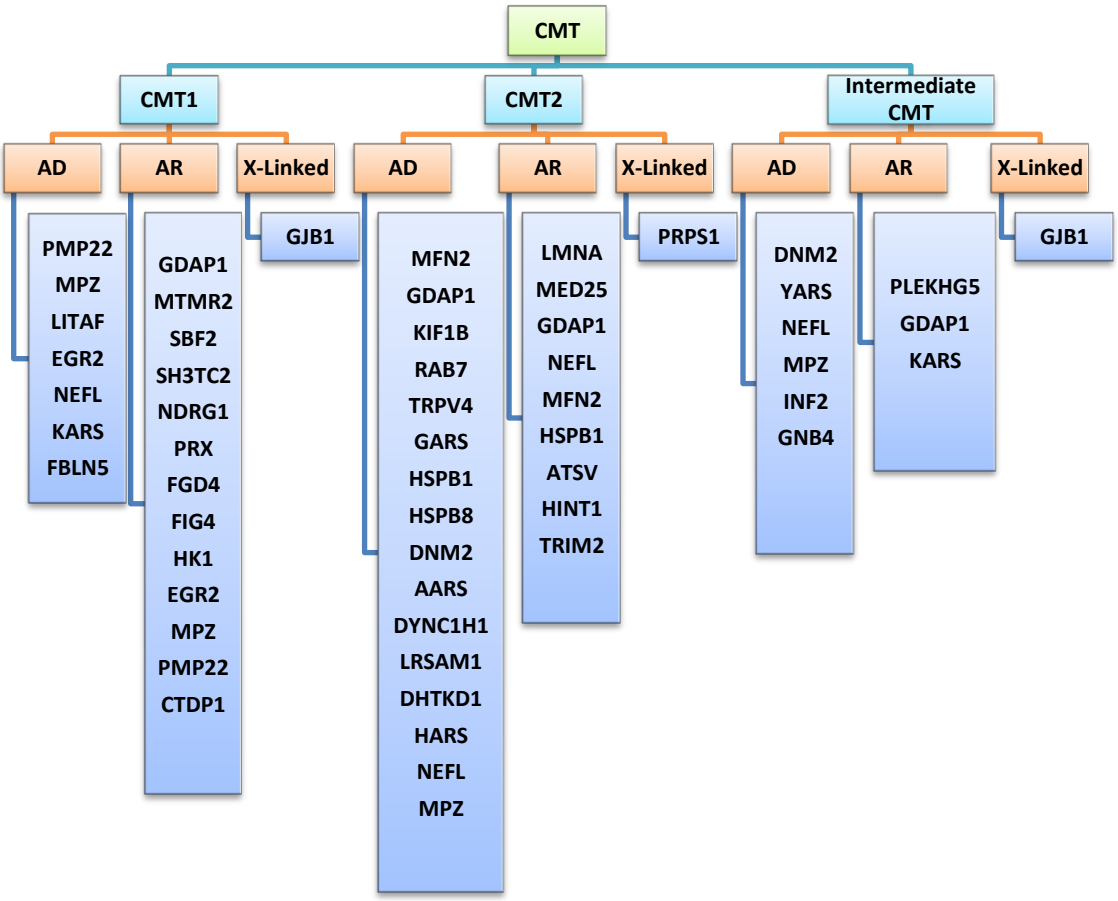


Figure 1.2. Genetic classification of CMT and causative genes.

The variety of clinical symptoms and genetic characteristics of CMT indicate the heterogeneity of the disease. Thus, a pure classification is not yet possible. Currently, the categorization combines electrophysiological studies, clinical characteristics, family history (inheritance pattern) and the causative gene.

CMT is caused by mutations in genes that encode proteins located in axons, Schwann cells and myelin structures (Figure 1.3) (Patzkó and Shy *et al*., 2011).

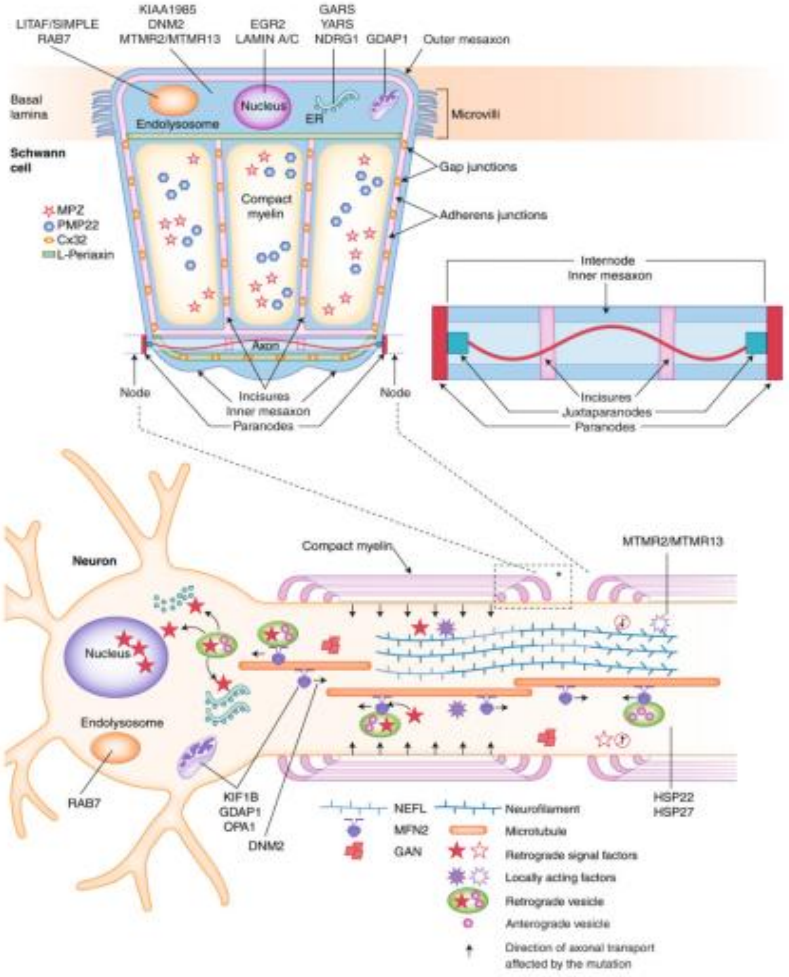


Figure 1.3. Proteins and intracellular pathways involved in CMT (Patzkó and Shy, 2011).

The peripheral nervous system (PNS) contains myelinated and non-myelinated nerves that can be very long in length and diameter. Rapid and efficient transmission of nerve impulses through axons is enhanced by Schwann cells. With the help of nodes of Ranvier, myelin sheaths increase the speed and efficiency of signal propagation by reducing the dispersion of electrical current (Juarez and Palau, 2012; Langworthy and Appel, 2012). Schwann cells are also responsible for sequential assembly of multiprotein complexes of cell adhesion molecules, ion channels, and scaffolding proteins into distinct domains that are localized around nodes of Ranvier during myelination and remyelination. In addition, they regulate axon cytoskeleton, organelle content and axonal transport. At early developmental stages axonal signals ensure that the number of Schwann cells and axons are matched. Signals from axons also determine whether or not a Schwann cell will

make myelin. So, axons and Schwann cells are anatomically and functionally connected in a highly dynamic community. They regulate each other and any problem in this system usually causes peripheral neuropathies (Niemann *et al*., 2006).

The common molecular pathways that have been identified to play role in CMT are interactions between Schwann cells and axons, compaction and maintenance of myelin, transport through myelin, transcriptional regulation, cell signaling, cytoskeletal formation, protein turnover, axonal transport, mitochondrial fusion and fission, vesicle and endosomal trafficking and chaperons (Patzkó and Shy, 2011; Ju´arez and Palau, 2012).

Currently, there is no specific treatment for CMT. Only supportive therapies such as proper bracing, physical and occupational therapies and orthopedic surgeries are used to increase the quality of life. Orthopedic devices are usually beneficial (Siskind *et al*., 2013).

1.1.1. Charcot-Marie-Tooth Disease Type 1 (CMT1)

CMT1 represents demyelinating form of the disease with a reduction in NCV (<38 m/s). Although it is not usually performed anymore, nerve biopsy of the affected individuals indicates loss of myelinated fibers and "onion bulb" formation due to repeated demyelination and remyelination in peripheral nerves (Banchs *et al*., 2009).

CMT1 can be subdivided according to the causative genes. CMT1A is the most common subtype that accounts for approximately 55% of CMT1 and 36% of all CMT cases (Saporta *et al*., 2011). It is caused by a 1.4 Mb duplication of chromosome 17p11.2 containing the peripheral myelin protein 22 (PMP22) gene (Lupski *et al*., 1991; Raeymaekers *et al*., 1991; Timmerman *et al*., 1992). The deletion of the same region causes another neurological disorder known as Hereditary Neuropathy with Liability to Pressure Palsies (HNPP). Point mutations in PMP22 gene are also causative for CMT1A and HNPP (Barisic *et al*., 2008).

1.1.1.1. PMP22. PMP22 is a glycosylated transmembrane protein which is a member of the Growth Arrest Specific 3 (GAS3) family. It has two alternative promoters; P1 and P2, both transcripts contain different noncoding exons, but produce the same protein. It is supposed that the presence of two promoters induces tissue specific expression. P1

produces transcripts that are found in PNS and CNS. In the PNS, the gene is highly expressed in Schwann cells and its protein product is localized to compact myelin structures. It is known to be required for myelin synthesis and assembly. On the other hand, transcripts produced from P2 are present in neuronal or non-neuronal tissues within the whole body (Rao *et al*., 2011).

PMP22 represents 2-5% of PNS myelin proteins. As its name implies, the protein has a molecular mass of 22 kDa. PMP22 protein consists of four transmembrane and two extracellular domains including a carbohydrate chain which is implicated in intercellular recognition, interaction and adhesion (Figure 1.4). In addition to its role in proper myelination and maintenance, PMP22 also regulates cell spreading, cell migration and apoptosis (Snipes and Suter, 1995; D'Urso *et al*., 1999).

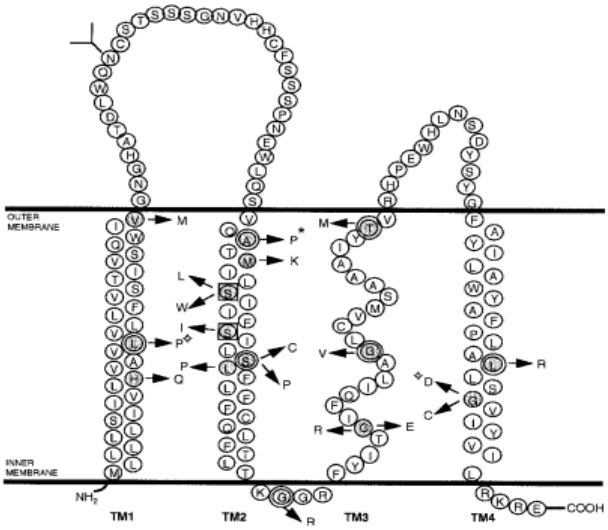


Figure 1.4. Protein structure of PMP22 with point mutations associated with CMT (Kovach *et al*., 1999).

In order to enforce adhesion, PMP22 interacts with another protein found in compact myelin structure; myelin protein zero (MPZ). Mutations in MPZ gene on chromosome 1q22-23 are also causative for CMT and in this case the phenotype is referred as CMT1B. This CMT subtype accounts for 10% of CMT1 cases (Patzkó and Shy, 2011). Mutations in *MPZ* may also cause AD CMT2 and intermediate CMT (Niemann *et al*., 2006).

1.1.1.2. MPZ. MPZ is one of the most important proteins found in compact myelin structure. It represents 50% of peripheral myelin proteins. It is a member of immunoglobulin gene superfamily. The protein contains an extracellular immunoglobulin-like domain, a transmembrane domain and a cytoplasmic domain. Up to now, at least 95 mutations in *MPZ* have been identified to cause peripheral neuropathies and most of them are found within the extracellular domain (Figure 1.5). As a member of immunoglobulin gene superfamily, MPZ protein plays a role in adhesion to connect adjacent myelin lamallae. Because it holds myelin structures together, it is suggested that MPZ is essential in myelination as well as in regulation of this process. Moreover, both extracellular and cytoplasmic domains of the protein are required for proper adhesion (Shy *et al*., 2004).

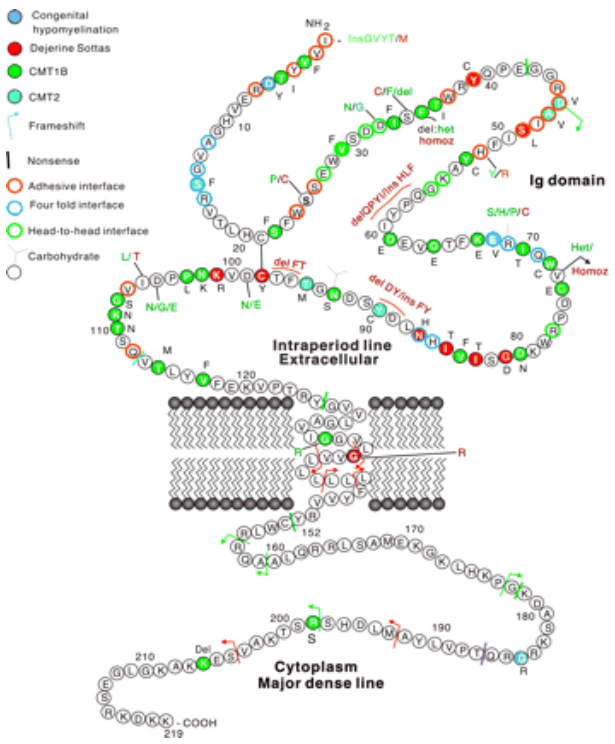


Figure 1.5. Protein structure of MPZ with the mutations known to cause neuropathy (Shy *et al*., 2004).

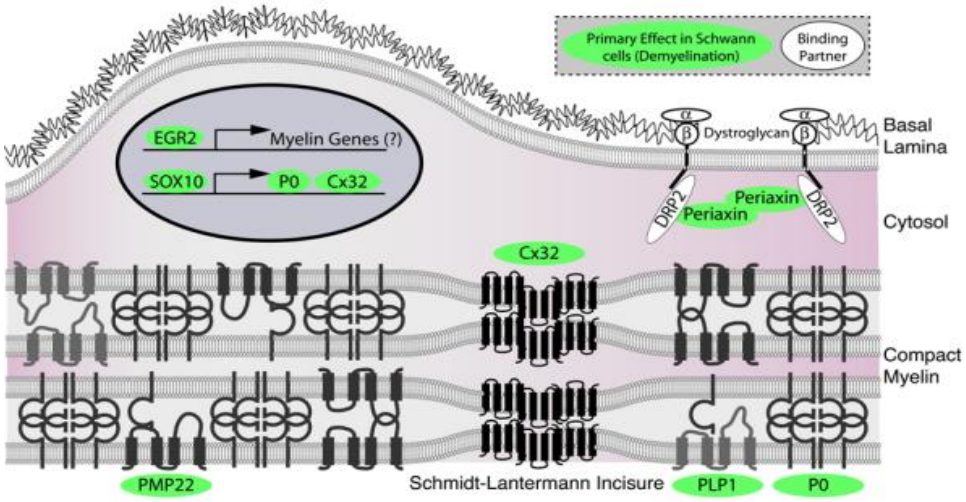


Figure 1.6. Proteins in the myelin sheath that cause CMT (Niemann *et al*., 2006).

1.1.1.3. GJB1. Gap junction beta 1 (GJB1) gene is found on chromosome Xq13.1 and encodes a connexin protein whose molecular mass is 32 kDa (Cx32) (Oyamada *et al*., 2013). Mutations in this gene are responsible for X-linked demyelinating CMT (CMT1X) and this subtype is the second most frequent form of the disease after CMT1A. More than 300 causative mutations in GJB1 gene have been identified up to date (Patzkó and Shy, 2011).

Connexins are a family of tetraspan transmembrane proteins. There are at least 20 isoforms of connexins. They generally contain four membrane-spanning domains (M1-M4), two extracellular loops (E1 and E2), one cytoplasmic loop (CL) and intracellular amino and carboxyl termini. The main divergence occurs in the CL and carboxyl terminal. Six connexin proteins oligomerize and generate a structure called connexon. Connexons, on the other hand, form hemichannels in each cell resulting in gap junction channels in order to transport ions and small metabolites between adjacent cells (Figure 1.7) (Saez *et al*., 2010; Nicholson and Myers, 2006).

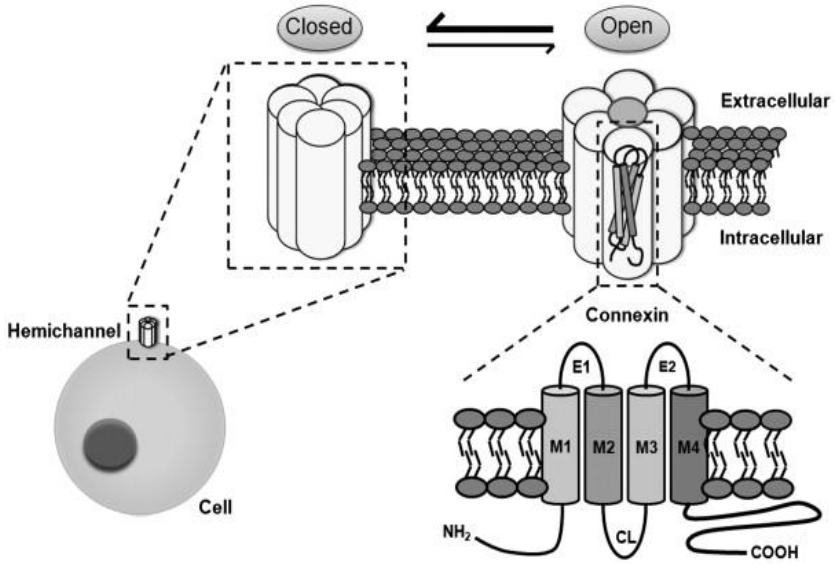


Figure 1.7. Hemichannels and protein structure of connexins (Saez *et al*., 2010).

Cx32 is located mostly in myelinating Schwann cells at Schmidt-Lanterman incisures and paranodal region (Figure 1.6) and also found in oligodendrocytes, but not in neurons. It is also expressed in several other tissues such as liver and lung. In Schwann cells, gap junctions generate intracellular Cx32 channels through the layers of myelin. Aggregates of these channels form a pathway for diffusion of molecules from Schwann cell to the axon. It is involved in proper myelin formation by allowing communication between Schwann cells and axons (Stauch *et al*., 2012).

1.1.2. Charcot-Marie-Tooth Disease Type 4 (CMT4)

Autosomal recessive forms of demyelinating CMT are designated as CMT4. As for other CMT types it is divided into subtypes based on the causative genes. Ganglioside-induced differentiation-associated protein 1 (GDAP1) gene located on chromosome 8q13.3 is the most frequently mutated one among CMT4 and it is classified as CMT4A (Siskind *et al*., 2013). Except CMT4A, mutations in this gene are also responsible for AR/AD CMT2 and intermediate recessive CMT (Tazir *et al*., 2013).

1.1.2.1. GDAP1. GDAP1 is a member of glutathione-S-transferases (GSTs) subfamily. It is located in the outer mitochondrial membrane (OMM) and expressed mainly in myelinating Schwann cells and motor and sensory neurons of the PNS. Because of its location, it has an essential role in regulation of mitochondrial dynamics, especially in fission (Wagner *et al*., 2009).

The protein contains two GST domains; GST-N located at the N terminus and GST-C located at the C terminus. There are also a transmembrane domain (TMD), a hydrophobic domain (HD) and two helices, α4 and α5, resulting in a α4-α5 loop structure (Figure 1.8). It is predicted that C-terminal hydrophobic TMD is localized in the membrane, while N-terminal domain is placed to the cytoplasm. GST domains are supposed to have a role in mitochondrial localization of the protein, but their function in regulation of mitochondrial morphology is still unknown. TMD is responsible for mitochondrial targeting and HD plays a role in mitochondrial fission (Cassereau *et al*., 2011; Otera *et al*., 2013).

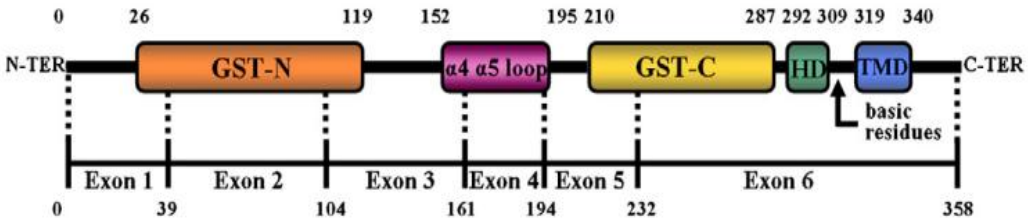


Figure 1.8. Protein structure and genomic organization of GDAP1 (Cassereau *et al*., 2011).

1.1.2.2. NDRG1. CMT4D is another subtype of autosomal recessive demyelinating CMT. It is also known as Hereditary Motor and Sensory Neuropathy - Lom (HMSNL). Because it was first described in a city in Bulgaria called Lom, the disease is named as HMSNL. Mutations in the N-myc downstream-regulated gene-1 (NDRG1) gene located on chromosome 8q24 are responsible for this subtype, and they are generally described in Gypsy populations (Kalaydjieva *et al*., 2000).

NDRG1 is an intracellular protein belonging to the NDRG family. There are three more proteins in this family in addition to NDRG1 and they have an NDR region and an

α/β hydrolase region in common. However, hydrolase region does not contain a hydrolytic catalytic site. Distinctly, NDRG1 has three tandem repeats which are not present in other family members. Additionally, it contains a phosphopantetheine sequence that has a role in attachment of activated fatty acids and amino acids (Figure 1.9) (Melotte *et al*., 2010).

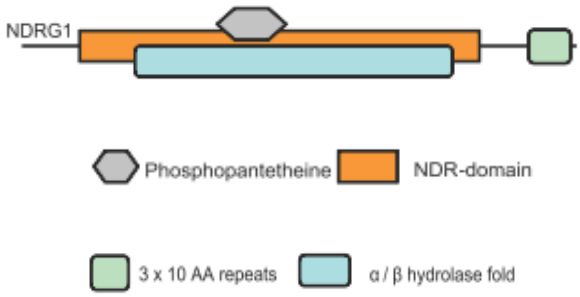


Figure 1.9. Protein structure of NDRG1 protein (Melotte *et al*., 2010).

NDRG1 is expressed in the PNS. It is mainly localized in myelinating Schwann cells and oligodendrocytes, while sensory and motor neurons do not contain NDRG1. Because it is localized to cytoplasm and nucleus, it is supposed to be a signaling protein between these structures. Hence, it is required for Schwann cell signaling resulting in axonal survival. NDRG1 is also connected to CMT pathogenesis by interacting with E-cadherin/catenin complex. Recently, it has been shown that E-cadherin/catenin complex is involved in the initiation of myelination in Schwann cell-axon interface. In cancer, it acts as a metastasis suppressor gene and it is regulated by p53. Moreover, it plays a role in ER stress response (Bucci *et al*., 2012; Kalaydjieva *et al*., 2000).

1.1.2.3. FGD4. Another subtype of autosomal recessive demyelinating CMT is CMT4H. It is caused by mutations in FYVE, RhoGEF and PH domain-containing 4 (FGD4) gene on chromosome 12p11.21.

FGD4 encodes a protein called Frabin that is a guanine nucleotide exchange factor (GEF). It interacts especially with a member of Rho family of GTP binding proteins (Rho GTPases), Cdc42. Cdc42 has a role in regulation of actin cytoskeleton dynamics and

cellular morphogenesis including myelination. Frabin, on the other hand, maintains the correct activation of Cdc42 (Tazir *et al*., 2013).

FGD4 is composed of an F-actin binding (FAB) domain, a Dbl homology (DH) domain, a FYVE domain and two pleckstrin homology (PH) domains (Figure 1.10). FAB domain is responsible for association with F-actin. DH domain is important for the catalysis of GDP-GTP exchange. The protein interacts with phosphoinositides (PIs) through its PH and FYVE domains. Therefore, defects in PI metabolism and cytoskeletal dynamics affect membrane trafficking and myelin deposition resulting in CMT (Baudot *et al*., 2012; Bucci *et al*., 2012).

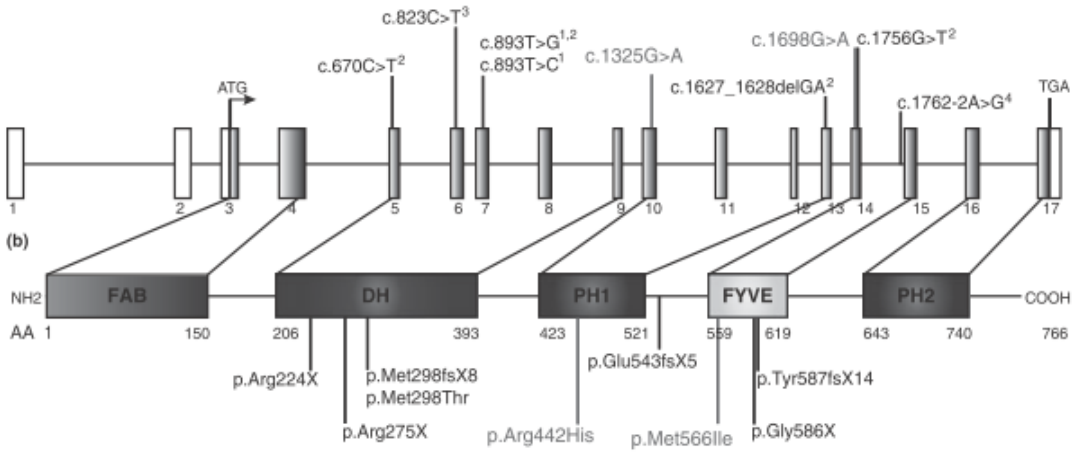


Figure 1.10. Protein structure of FGD4 and mutations associated with CMT (Baudot *et al*., 2012).

1.1.3. Charcot-Marie-Tooth Disease Type 2 (CMT2)

CMT2 represents the axonal form of the disease. It accounts for approximately 20% of all CMT cases. It is distinguished from demyelinating CMT by normal or higher NCVs (>38 m/s) with a decrease in CMAP amplitudes. It is classified based on causative genes like in CMT1, but in the case of CMT2, there are numerous genes each contributing only to a small proportion of CMT2 phenotypes. The most frequent subtype of CMT2 is

CMT2A. It accounts approximately 10-30% of all CMT2 cases. It is caused by mutations in mitofusion 2 (MFN2) gene on chromosome 1p36.22 (Zuchner *et al*., 2004).

1.1.3.1. MFN2. MFN2 is one of the proteins found in OMM. It has an essential role in mitochondrial fusion. It is expressed in peripheral nerves, in addition to spinal cord, muscle and heart. Except OMM, it is also localized in endoplasmic reticulum (ER) in order to control its morphology and facilitate mitochondrial calcium influx from ER stores (Banchs *et al*., 2009; Bucci *et al*., 2012).

MFN2 is a nuclear-encoded dynamin-like GTPase. The GTPase domain is responsible for the oligomerization of mitofusions by GTP hydrolysis. The protein also has a proline-rich (PR) domain known to be required for protein-protein interactions. Additionally, there are two heptad repeat (HR) domains (HR1 and HR2) that are responsible for coiled-coil interactions (Figure 1.11). HR2 domain is especially important for fusion process. When two mitofusin molecules interact, the connection between them is generated by their HR2 domains. Both C- and N- terminus of MFN2 are localized in the cytosol. Up to now, at least 40 mutations in the gene are found to be associated with CMT andmore than 50% of them are located in the GTPase domain (Ranieri *et al*., 2013).

Figure 1.11. Protein structure of MFN2 (Ranieri *et al*., 2013).

1.1.3.2. AARS. Interestingly, several aminoacyl-tRNA synthetases (ARS), namely, ARS, GARS, YARS, KARS and AARS, are involved in CMT pathogenesis representing different types of the disease. Mutations in the alanyl-tRNA synthetase (AARS) gene, located on chromosome 16q22, are classified as CMT2N (Latour *et al*., 2010).

ARSs are essential enzymes that provide transfer of appropriate amino acids to tRNA molecules in the cytoplasm and mitochondria during translation. All reported mutations in ARS genes coding for cytoplasmic proteins are causative for CMT, while the

ones in mitochondria have been implicated in several other diseases. The mechanism of how the mutations in these genes cause CMT is not clearly known, yet (Yao and Fox, 2013).

AARS gene is responsible for the attachment of amino acid alanyl to tRNA molecules. It encodes a protein consisting of a catalytic domain, an editing domain, and a C-terminus domain (Figure 1.12). Aminoacylation process takes place at the catalytic domain. They have also a proofreading mechanism and the hydrolization of incorrect product occurs in the editing domain (Naganuma *et al*., 2009).

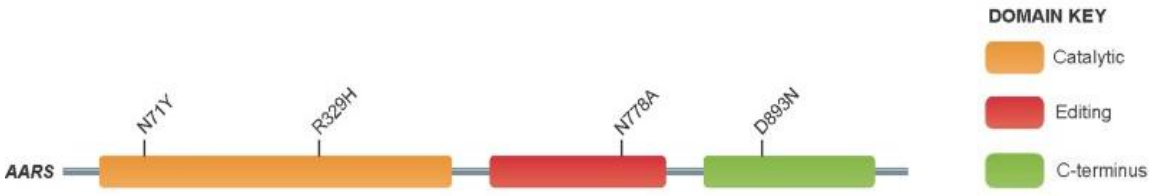


Figure 1.12. Protein structure of AARS located in cytoplasm and its mutations associated with CMT (Yao and Fox, 2013).

1.1.3.3. DNM2. Dynamins are a large family of GTPases. They consist of DNM1, DNM2 and DNM3 in humans. Mutations in DNM2 gene have been associated with CMT2M, a subtype of autosomal dominant CMT2 (Fabrizi *et al*., 2007; Gallardo *et al*., 2008). However, there are cases in which mutations in this gene are involved in the intermediate form of the disease (Zuchner *et al*., 2005).

DNM2 contains an N-terminal GTPase domain, a middle domain (MID), a pleckstrin homology (PH) domain, a GTPase effector domain (GED), and a C-terminal proline-arginine-rich domain (PRD) (Figure 1.13). Mutations associated with CMT are usually located at the PH domain which plays a role in binding membrane phospholipids. On the other hand, GED activates the GTPase and PRD is important for protein-protein interactions (Bohm *et al*., 2012).

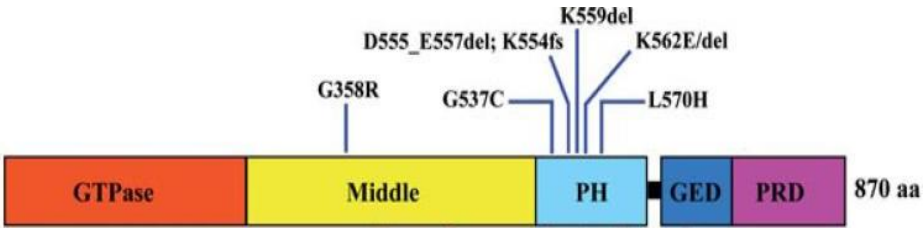


Figure 1.13. Protein structure of DNM2 and its mutations associated with CMT (Bohm *et al*., 2012).

DNM2 is a key player of formation and fission of budding vesicles produced from plasma membrane and Golgi. It regulates membrane trafficking. It is involved in endocytosis, cytokinesis, phagocytosis, and cell migration. It is also responsible for centrosome cohesion and actin assembly through direct and indirect interactions. Most mutations causative for CMT reduce binding efficiency of DNM2 protein to phosphoinositides resulting loss of function (Cowling *et al*., 2012; Bohm *et al*., 2012).

1.2. Exome Sequencing

Genes responsible for Mendelian diseases as well as other complex traits have been investigated since many years by using conventional approaches such as candidate gene sequencing or linkage analysis. However, only one-half to one-third of all cases have been identified and still the knowledge about genes and mutations that are causative for a lot of rare phenotypes remains unknown. Many loci and genes related with complex traits have been identified by genome-wide association studies (GWASs), but these genes account for only a small fraction of all unresolved cases (Bamshad *et al*., 2011).

In conjuction with the availability of next generation sequencing (NGS) methods since 2005, it has become easier to search genes for complex traits. Especially the development of a technique for sequencing of only the coding regions is a milestone for molecular diagnosis and identifying novel disease causing genes. This technique termed as ‘exome sequencing’ is a powerful method that accelerated the identification of associated genes (Bamshad *et al*., 2011).

‘Exome’ refers to the complete set of protein coding regions and it accounts for only 1% of the whole genome. In monogenic diseases, more than 85% of mutations are found within these regions (Choi *et al*., 2012). Whole genome sequencing targets about three Gigabases (Gb), whereas the typical exome target is about 33 Megabases (Mb). While approximately three million variants are detected in the whole genome context, the number decreases to about 15,000-20,000 for the exome. Hence, exome sequencing is technically feasible and extremely cost effective. Since it was discovered, many novel causative mutations and genes for numerous rare disorders have been identified (Ku *et al*., 2012).

1.2.1. Analysis of Exome Sequencing Data

In order to narrow down the search and identify the causal variant from about 20,000 variations identified by exome sequencing, several filtering steps are used (Figure 1.14). Generally, variants are classified as synonymous and nonsynonymous alterations. It is assumed that the causative variant will change the amino acid sequence. Thus, focusing on nonsynonymous ones eliminates approximately 50-75% of the variations. Frameshifts, stop codons and splice site mutations can be considered as important. Additionally, predicted effects of protein alterations assumed by some tools such as PolyPhen or SIFT can also be used to reduce the number of possible variants. Moreover, variants found in common databases such as dbSNP, HapMap or 1000 Genome Project are usually excluded by assuming that the causative variant will not be an alteration that has been previously reported. Because novel single nucleotide variations (SNVs) identified by exome sequencing usually accounts for only about 2%, this is a very effective step to reduce the number of possible causal variants. On the other hand, a small fraction of these previously reported variants can be pathogenic and they can be eliminated due to their low frequencies in population. So, if the filtering approach fails, it has to be taken into consideration. The dominant/recessive mode of inheritance is also used to restrict the search. If the disease is inherited as a recessive trait, homozygous variations are chosen. However, compound heterozygosity has to be also considered. In the case of dominant mode of inheritance,

heterozygous variants are selected, while homozygous ones are excluded (Stitziel *et al*., 2011; Bamshad *et al*., 2011).

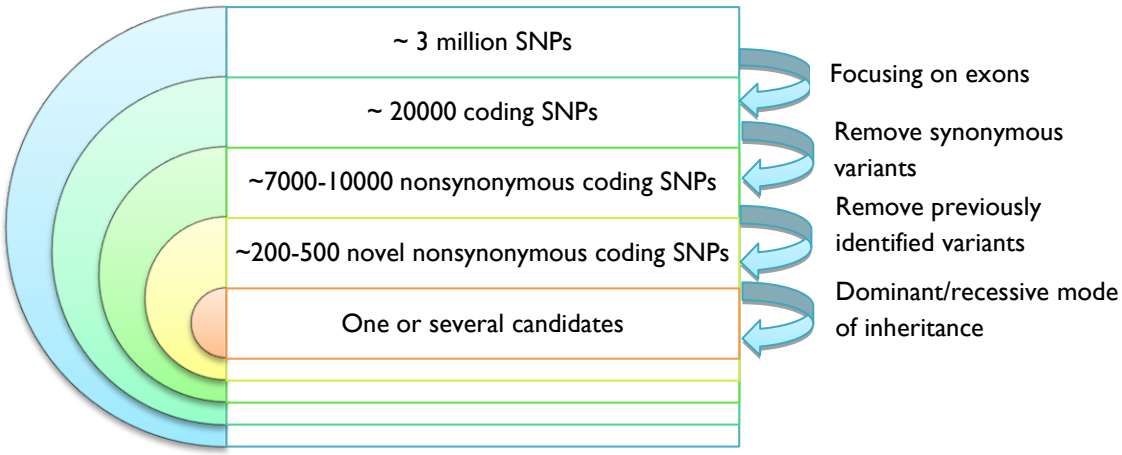


Figure 1.14. General filtering steps of exome sequencing data (Modified from Stitziel *et al*., 2011).

1.2.2. Exome Sequencing in CMT

Exome sequencing method well suits the genetic analysis in CMT patients since it is clinically and genetically heterogeneous with still unidentified loci. Phenotype rarely directs geneticists to certain CMT genes among more than 30 causative genes. The disease shows all Mendelian inheritance types and the genes overlap between these different modes of inheritance. On the other hand, it is one of the most common inherited neurological diseases with a prevalence of 1 in 2500. Exome sequencing can be used as a diagnostic screening tool for CMT (Azzedine *et al*., 2012). Additionally, considering the fact that only 50-70% of CMT cases can be genetically diagnosed by screening the known CMT genes (Saporta *et al*., 2011), the technique has a potential to discover further CMT genes.

Recent studies have demonstrated that exome sequencing is a very powerful and effective tool in order to diagnose CMT and/or to discover novel causative mutations and

genes. The first exome sequencing effort for a CMT1 case resulted in the identification of compound heterozygous mutations located in SH3TC2 gene. One of these mutations was an unreported missense mutation, p.Y169H, while the other was a known disease causing nonsense mutation, p.R954X (Lupski *et al*., 2010). In another study, three affected individuals of a four generation family were screened by exome sequencing. All three patients represented CMT2 characteristics and the mode of inheritance was autosomal dominant. Accordingly, after the evaluation of exome sequencing data, only one mutation was detected to be shared by all three patients. It was a missense mutation, p.H306R, found in the novel DYNC1H1 gene (CMT2O) (Weedon *et al*., 2011). Similarly, a novel alteration, p.R186Q, in TRPV4 gene was identified in an affected individual representing CMT2C phenotype. Although TRPV4 gene has been previously associated with CMT2C subtype, exome sequencing succeeded to discover a novel mutation in this gene (Landoure *et al*., 2012).

1.3. Genetic Testing in CMT

Since CMT is a heterogeneous disorder, it is generally challenging for clinicians to determine the underlying cause of the disease. A detailed knowledge of clinical and genetic features of the subtypes and their frequencies in different populations can be of help to clinicians for an accurate diagnosis. Even though there is no treatment for CMT currently, knowing the genetic cause of the disease for an affected individual is important for family planning, prognosis and treatment trials (Miller *et al*., 2011).

1.3.1. Phenotype/Genotype Correlation

A correlation between genotype and phenotype is useful to unravel the pathomechanisms responsible for the disease and can lead clinicians in diagnosis. The association of certain clinical characteristics with the most common causative genes is shown in Table 1.1.

Table 1.1. Genotype/phenotype association in CMT (Barisic *et al*., 2008; Tazir *et al*., 2013; Siskind *et al*., 2013).

Type	Gene	Age of onset	Specific phenotype	MNCV
AD CMT1				
CMT1A	PMP22 *Duplication* PMP22 *Point mutations*	1-10 years Early onset	Slowly progressive More severe, deafness	15-30 m/s 15-20 m/s
HNPP	PMP22 *Deletion*	2-64 years	Transient episodes of weakness or sensory loss	Normal or mild decreasing
CMT1B	MPZ	Late onset Early onset	Typical CMT phenotype Severe, hearing loss	>35 m/s <20 m/s
AR CMT1				
CMT4A	GDAP1	Early onset	Severe, progressive, vocal cord paralysis, scoliosis	25-35 m/s
CMT4D	NDRG1	1–10 years	Deafness	9-20 m/s
CMT4H	FGD4	<2 years	Severe, scoliosis, delayed milestones	<15 m/s
X-linked CMT1				
CMT1X	GJB1	Males: <20 years Females: later onset	Males: severe Females: milder	10-37 m/s 30-50 m/s

Table 1.1. Genotype/phenotype association in CMT (Barisic *et al*., 2008; Tazir *et al*., 2013; Siskind *et al*., 2013) (cont.).

AD CMT2				
CMT2A	MFN2	<10 years	Severe, optic atrophy, hearing loss	Normal or slightly reduced NCV
CMT2B	KIF1B	Late or early onset	Sensory loss, feet ulcerations, hyperkeratosis	Normal NCV
CMT2N	AARS	6-54 years	Slowly progressive	32-50 m/s
CMT2K	GDAP1	2nd or 3rd decade	Slowly progressive, mild, vocal cord paralysis	Slightly decreased NCV
CMT2I/J	MPZ	Late onset	Deafness	< 38 m/s, Progress to reduced NCV
AR CMT2				
CMT2K	GDAP1	Early onset	hypotonia, vocal cord paralysis, respiratory insufficiency	>40 m/s
HMSN VI	MFN2	1 to 3 years	Optic atrophy, scoliosis	Slightly decreased NCV

Table 1.1. Genotype/phenotype association in CMT (Barisic *et al*., 2008; Tazir *et al*., 2013; Siskind *et al*., 2013) (cont.).

Intermediate CMT				
DI-CMTD	MPZ	Late onset	Variable and moderate severity	Intermediate NCV
DI-CMT2B	DNM2	1st or 2nd decade	Normal or increased tendon reflexes	24-54 m/s
CMT-RIA	GDAP1	Early onset	Typical CMT phenotype	25-35 m/s

General symptoms of CMT1A caused by the duplication of a region on chromosome 17 including PMP22 gene are foot drop, tight heel cords, pes cavus, hammertoes, sensory loss and hand weakness. The majority of patients begin to show these symptoms in childhood. The disease progresses slowly and only 5-7% need a wheelchair. NCVs are slower than 35 m/s (usually <25 m/s). On the other hand, deletion of the same region causes palsy of nerves resulting transient episodes of weakness or sensory loss, while point mutations in PMP22 often give rise to an early onset, severe phenotype. Affected individuals usually need wheelchair before adulthood (Siskind *et al*., 2013).

Furthermore, when there is no male to male transmission in the case of CMT1, GJB1 is thought to be the possible disease causing gene. Mutations cause a loss of function in males due to hemizygosity. They generally present symptoms early in childhood. Pronounced atrophy of intrinsic hand muscles and distal calves are usually observed. The progression of the disease is aggressive. On the other hand, females with GJB1 mutations usually have mild phenotype and slower progress. NCVs are 25-45 m/s for males, while females' NCVs are above 35 m/s (Siskind *et al*., 2013; Saporta *et al*., 2011).

2. AIM OF THE STUDY

In this study, we aimed to identify unknown causative genes and novel mutations for Charcot-Marie Tooth disease by exome sequencing approach. In addition, by screening known causative CMT genes for mutations in the light of clinical characteristics of the patients, we aim to perform molecular diagnosis of the disease to direct clinicians and patients with.

3. MATERIALS

3.1. Patients

Seventeen unrelated Turkish CMT patients were analyzed by exome sequencing in the first part of the study. Based upon clinical and electrophysiological studies, they were all diagnosed as CMT. The majority of them had an axonal type of the disease. Among fifteen CMT2 patients, nine presented autosomal recessive inheritance, while six had an autosomal dominant trait. In two patients, autosomal dominant demyelinating CMT (CMT1) was suspected due to pedigree and clinical investigations.

In the second part of the study mutation screening was perfomed in candidate CMT genes determined based on clinical findings by conventional Sanger sequencing for four unrelated families. These cases were diagnosed as demyelinating CMT by clinical and electrophysiological studies. Seven of them had autosomal recessive inheritance, and one had autosomal dominant trait.

Peripheral blood samples from all patients included in this study and their family members were provided by various neurology departments around Turkey. Detailed clinical features of patients and pedigrees were obtained from clinicians. Informed consent was obtained from all patients and family members.

3.2. Chemicals

Solid and liquid chemicals used in this study were purchased from Sigma (USA), Merck (Germany) and Riedel de-Häen (Germany) unless stated otherwise in the text.

3.3. Buffers and Solutions

3.3.1. DNA Extraction from Peripheral Blood

Table 3.1. Materials used for DNA extraction from peripheral blood.

Red Blood Cell (RBC) Lysis Buffer	155 mM NH_4Cl 10 mM $KHCO_3$ 1 mM Na_2EDTA (pH 7.4)
Nuclei Lysis Buffer	10 mM Tris-HCl (pH 8.0) 400 mM NaCl 2 mM Na_2EDTA (pH 7.4)
Sodiumdodecylsulphate (SDS)	10% SDS (w/v) (pH 7.2)
Proteinase K	20 mg/ml in dH_2O
5 M NaCl Solution	292.2 g NaCl in 1 L dH_2O
Ethanol (EtOH)	Absolute EtOH
Tris-EDTA (TE) Buffer	20 mM Tris-HCl (pH 8.0) 1 mM Na_2EDTA (pH 8.0)

3.3.2. Polymerase Chain Reaction (PCR)

Table 3.2. Materials used for PCR.

10 X $MgCl_2$ Free *Taq* Buffer	750 mM Tris-HCl (pH 8.8 at 25^0C) 200 mM $(NH_4)_2SO_4$ 0.1% (v/v) Tween 20 (Thermo Scientific, USA)
Magnesium Chloride ($MgCl_2$)	25 mM $MgCl_2$ (Thermo Scientific, USA)
Deoxyribonucleotide Triphosphates (dNTPs)	100 mM of each dNTP (Fermentas, Lithuania)

3.3.3. Agarose Gel Electrophoresis

Table 3.3. Materials used for agarose gel electrophoresis.

5 X Tris-Borate-EDTA (TBE) Buffer	0.89 M Tris-Base 0.89 M Boric Acid 20 mM Na_2EDTA (pH 8.3)
1 or 4% Agarose Gel	1or 4% Agarose (w/v) (Peqlab, Germany) in 0.5 X TBE Buffer
Ethidium Bromide	10 mg/ml
6 X Loading Dye	2.5 mg/ml Bromophenol Blue 1% SDS (w/v) in 2 ml glycerol

3.3.4. Polyacrylamide Gel Electrophoresis (PAGE)

Table 3.4. Materials used for PAGE.

5 X TBE Buffer	0.89 M Tris-Base 0.89 M Boric Acid 20 mM Na_2EDTA (pH 8.3)
30% Acrylamide Stock (29:1)	29% Acrylamide (w/w) 1% N, N'-methylenebisacrylamide (w/w)
Ammoniumpersulfate (APS)	1% APS (w/v)
6 X Denaturing Loading Dye	95% Formamide (w/v) 20 mM EDTA 0.05% Xylene Cyanol (w/v) 0.05%
Glycerol	99% Glycerol
Tetramethylethylenediamine (TEMED)	

3.3.5. Silver Staining

Table 3.5. Materials used for Silver Staining.

Solution 1	0.1% $AgNO_3$ in dH_2O (w/v)
Solution 2	1.5% NaOH (w/v) 0.01% $NaBH_4$ (w/v) 0.015% Formaldehyde (v/v)

3.4. Fine Chemicals

3.4.1. Enzymes

Taq DNA Polymerase was purchased from Thermo Scientific (USA). *Taq*I restriction enzyme was obtained from Fermentas (Lithuania).

3.4.2. Oligonucleotide Primers

All the primers used in this study were designed using the program ApE (A Plasmid Editor, Wayne Davis, University of Utah) and OligoAnalyzer 3.1 tool (*Integrated DNA Technologies, Inc)*. The primers used in PCR to confirm variations selected from exome sequencing data and the primers used for mutation screening in GDAP1, GJB1, MPZ, NDRG1 and PMP22 genes were synthesized by Macrogen (Seoul, Korea) using the same primers by Macrogen (Seoul, Korea).

The primers used for mutation screening are given in Tables 3.6 to Table 3.10.

Table 3.6. Primers used for exon amplification of *GDAP1*.

Region	Primer (F/R)*	Primer Sequence (5' → 3')	Annealing Temp. (°C)	Product Size (bp)
Exon 1	G1 F1	GCCCTTCATAACCAGGGTCTC	62	456
	G1 R1	GCCCGATCCGCGGGACCCTCCG		
	G1 F2	AGCAACCCTCAGTATCTTGG	55	618
	G1 F2	CACTGGAGGCGGATTTCTAG		
Exon 2	G2 F	GGCTGCTTAGCGGTGTCCAGG	55	330
	G2 R	GGGAACACATAGTTGTGTTG		
Exon 3	G3 F	GCTTTTGAGTGTAACAACTCATG	64.4	317
	G3 R	GACCATGAGACATGCTAGGTC		
Exon 4	G4 F	CAGGGTAAGCCCAAGGCAGAG	66	288
	G4 R	GTAGAACATTTACTCCGTGCA		
Exon 5	G5 F	GGCTGAACTCTGTAAGAGTTT	54.9	281
	G5 R	GACCTAAGAATGTTCCCATG		
Exon 6	G6 F	CCACTGATACCAGCTGG	54	530
	G6 R	CAGAGAGCCACGGGCAATCAC		

***** F: Forward primer, R: Reverse primer

Table 3.7. Primers used to amplify coding region of *GJB1*.

Region	Primer (F/R)*	Primer Sequence (5' → 3')	Annealing Temp. (°C)	Product Size (bp)
Intron 1-Exon 2	GJB1 F	AGGAGAAGCTGGCAAGGG	64.5	323
	GJB1 R	TCATCACCCCACACACTCTC		
Exon 2	Cx32 F	TGAGGCAGGATGAACTGGACA	65	885
	Cx32 R	TGGCAGGTTGCCTGGTATGT		

***** F: Forward primer, R: Reverse primer

Table 3.8. Primers used for exon amplification of *MPZ*.

Region	Primer (F/R)*	Primer Sequence (5' → 3')	Annealing Temp. (°C)	Product Size (bp)
Exon 1	MPZ1 F	CCCGTTCAGTTCCTGGT	55	173
	MPZ1 R	GTCCCAAGACTCCCAGAGTA		
Exon 2	MPZ2 F	TTCCAAATTGGCCTCTGACC	64.5	440
	MPZ2 R	AGGACAATGTAGTCAGGGTG		
Exon 3	MPZ3 F	AGCTGTGTTCTCATTAGGGTCC	61.4	432
	MPZ3 R	GCATTGAGGATGTAGGACTCCCAG		
Exon 4	MPZ4 F	GGAGTCCTACATCCTCAATGC	64.5	383
	MPZ4 R	AGAGACTGCTGTACGTTTGG		
Exon 5	MPZ5 F	GGAGACTACTCCAGTGGGTG	65	255
	MPZ5 R	CTTGACTGTTCCCATCCCAC		
Exon 6	MPZ6 F1	ACAAGATGGGAAGGAGAACC	65	652
	MPZ6 R1	GACAGGCACCAAACACACAG		
	MPZ6 F2	GCCCAGGAGCTAATCCTACC	57	774
	MPZ6 R2	CTAGACAAGGCAGCAGCTTC		

***** F: Forward primer, R: Reverse primer

Table 3.9. Primers used to amplify exon 7 of NDRG1 gene.

Region	Primer (F/R)*	Primer Sequence (5' → 3')	Annealing Temp. (°C)	Product Size (bp)
Exon 7	NDRG1 F	AGGCTCCCGTCACTCTG	57	176
	NDRG1 R	GTCTCCTTCATCTTAAAATG		

***** F: Forward primer, R: Reverse primer

Table 3.10. Primers used to amplify coding regions of *PMP22*.

Region	Primer (F/R)*	Primer Sequence (5' → 3')	Annealing Temp. (°C)	Product Size (bp)
Exon 2	PMP22-2 F	CTCCTCGCAGGCAGAAACTC	64.5	202
	PMP22-2 R	CTGAACCAGCAGGAGCACGGG		
Exon 3	PMP22-3 F	TCAGGATATCTATCTGATTCTC	55	156
	PMP22-3 R	AAGCTCATGGAGCACAAAACC		
Exon 4	PMP22-4 F	TGGCCAGCTCTCCTAAC	55.8	217
	PMP22-4 R	CACCCCGCTTCCACATG		
Exon 5	PMP22-5 F	GCCATGGACTCTCCGTC	55	249
	PMP22-5 R	CCTATGTACGCTCAGAG		

***** F: Forward primer, R: Reverse primer

The primers that were used for confirmation of possible disease causing variations chosen by evaluating exome sequencing data for each patient are shown in Tables 3.11 to 3.18.

Table 3.11. Primers used to confirm variations identified by exome sequencing for patient P114.

Gene	Region	Primer (F/R)*	Primer Sequence (5' → 3')	Annealing Temp. (°C)	Product Size (bp)
AARS	Exon 8	AARS-8 F	AAGTTCCTGAAAGAGGTGTTGC	65	292
		AARS-8 R	GCTGCTTAACAGTATGCCCTC		
SBF2	Exon 36	SBF2-36 F	CCATGTCCACCACTTTCCTCAAC	63.3	387
		SBF2-36 R	TGGACCTCCAGGACATGACC		

***** F: Forward primer, R: Reverse primer

Table 3.12. Primers used to confirm variations selected from the exome sequencing data for patient P138.

Gene	Region	Primer (F/R)*	Primer Sequence (5' → 3')	Annealing Temp. (°C)	Product Size (bp)
GJB1	Intron 1 Exon 2	GJB1 F	TCATCACCCCACACACTCTC	61.4	323
		GJB1 R	AGGAGAAGCTGGCAAGGG		
LDB3	Exon 10	LDB3 F	TCTGTGAACACCCTGCTGAG	58.9	800
		LDB3 R	GGCTCAGAGAGGTAAGACCAG		
OTOF	Exon 9	OTOF F	AGGCAGTGGTCAGAATGGAG	57.1	420
		OTOF R	GGGTTGTAACAGGGTGGATC		
CNNM4	Exon 2	CNNM4 F	GCAGTAGAAACGTCAGGAAAG ATGG	64.5	416
		CNNM4 R	ACTGGGCCTCTGGATGCTAG		
ARHGAP4	Exon 2	ARHGAP4 F	TAGCTCTTGCCCTCAGCCTAG	61.4	568
		ARHGAP4 R	CAATGTGACTCAGGCGCTGG		

***** F: Forward primer, R: Reverse primer

Table 3.13. Primers used to confirm variations selected from the exome sequencing data for patient P163.

Gene	Region	Primer (F/R)*	Primer Sequence (5' → 3')	Annealing Temp. (°C)	Product Size (bp)
DNM2	Exon 8	DNM2-8 F	GAACAGTAAACCCTGGCTTG	55	322
		DNM2-8 R	AGCTTGAGACCTTATTGCCTG		
LRSAM1	Exon 7	LRSAM1-7 F	AACTCTTGACCTTGTG ATCCACCTG	60	235
		LRSAM1-7 R	AGGATGGCCCAGAGCTGATG		

***** F: Forward primer, R: Reverse primer

Table 3.14. Primers used to confirm the variations selected from the exome sequencing data for patients P209 and P313.

Gene	Region	Primer (F/R)*	Primer Sequence (5' → 3')	Annealing Temp. (°C)	Product Size (bp)
MFN2	Exon 5	MFN2-5 F	GTAGCAGATGGGCAGCATTC	65	401
		MFN2-5 R	TCCAGAGGAAAGCCAAGAGG		

***** F: Forward primer, R: Reverse primer

Table 3.15. Primers used to confirm variations selected from the exome sequencing data for patient P212.

Gene	Region	Primer (F/R)*	Primer Sequence (5' → 3')	Annealing Temp. (°C)	Product Size (bp)
AARS	Exon 21	AARS-21 F	GTAAAGTCTGTCTCGGCCTC	57.1	620
		AARS-21 R	AAACGATTCCTACTTGCTGACG		
CXorf23	Exon 3	CXorf23 F	TCCTCTTTAACTGCCTTGTGC	65	938
		CXorf23 R	GCAATCCCTCCCACTTGATG		
SIGMAR1	Exon 2	SIGMAR1 F	GCGAAGAGATAGCGCAGTTG	61.4	500
		SIGMAR1 R	TAGCGCCCTGAGTGACAATC		
FLNA	Exon 6	FLNA F	GTGCCACAGCTCACACTGAATG	63.3	404
		FLNA R	GACGTACCAGACGGAGAAGGTG		
GMFG	Exon 6	GMFG F	CCAAGCCAACCTAGAAACTGAGTAC	64.5	378
		GMFG R	CCTGGATGACAGAGCGAGAC		

***** F: Forward primer, R: Reverse primer

Table 3.16. Primers used to confirm variations identified by exome sequencing for patient P245.

Gene	Region	Primer (F/R)*	Primer Sequence (5' → 3')	Annealing Temp. (°C)	Product Size (bp)
FGD4	Exon 14	FGD4-14 F	AGAGAGAGGGTTTCACCATG	65	467
		FGD4-14 R	AGTCCTCTGCACGTAATTGG		
KIF1B	Exon 22	KIF1B-22 F	AGCTGAATGGAGAGAAACTGG	55	344
		KIF1B-22 R	CAAGGTGAGAACCCAGATGC		
THBS4	Exon 21	THBS4 F	AGTGGCAACAGCTACAGTAC	64.5	450
		THBS4 R	CTTCTGGTACACAGGGACTATC		
ZNF880	Exon 4	ZNF880 F	TTCTGCACTGCCAGCATGATAC	55.8	454
		ZNF880 R	AAGGCTTTGCCATGCTCATTAC		

***** F: Forward primer, R: Reverse primer

Table 3.17. Primers used to confirm variations selected from the exome sequencing data for patient P378.

Gene	Region	Primer (F/R)*	Primer Sequence (5' → 3')	Annealing Temp. (°C)	Product Size (bp)
SBF2	Exon 37	SBF2-37 F	GGACGTGTAGGTTGTTTCCAG	55	457
		SBF2-37 R	TAGCTTATTGAGGGACCTGTGG		
SLC6A11	Exon12	SLC6A11	CATCTCGGTCGTTGTTTAAGC	55	393
		SLC6A11	CTTTGTAAGGTGGCCTGAAC		
DLGAP3	Exon 1	DLGAP3 F	ACCTAAAGGGATACAAGGATGC	55.8	1322
		DLGAP3 R	CTTTCCACTGCTCCCTAGTTG		

***** F: Forward primer, R: Reverse primer

Table 3.18. Primers used to confirm the variation revealed by exome sequencing for patient P851.

Gene	Region	Primer (F/R)*	Primer Sequence (5' → 3')	Annealing Temp. (°C)	Product Size (bp)
DNM2	Exon 8	DNM2-8 F	GAACAGTAAACCCTGGCTTG	55	322
		DNM2-8 R	AGCTTGAGACCTTATTGCCTG		

***** F: Forward primer, R: Reverse primer

3.4.3. DNA Ladders

Size markers used in this study were 100 bp DNA ladder with a range of 100-1000 bp and 1kb DNA ladder with a range of 250-10000 bp (Fermentas, Lithuania).

3.5. Equipments

Exome sequencing was performed by the group of Prof. James Lupski in the Department of Molecular and Human Genetics of Baylor College of Medicine (Houston, Texas, USA). Sanger sequencing was conducted by Macrogen (Seoul, Korea). The equipments used for experiments performed in our department are listed in Table 3.19.

Table 3.19. Equipments used in this study.

Autoclave	Model MAC-601 (Eyela, Japan)
Centrifuge	Allegra X-22R Centrifuge (Beckman Coulter, USA)
Deep Freezers	-20 ^{0}C (Bosch, Germany) -20 ^{0}C (Arçelik, Turkey)
Documentation System	GelDoc Documentation System (Bio-Rad, USA)
Electrophoresis	Mini-Sub Cell (Bio-Rad, USA) PROTEAN Vertical Electrophoresis System (Bio-Rad, USA)
Incubators	DRI Block DB-2A (Techne, UK) Shake'n'Stack (Hybaid, UK) Oven EN400 (Nuve, Turkey)
Magnetic Stirrer	Speed Safe (Hanna Instruments, USA) MK 418 (Nüve, Turkey)
Oven	Microwave Oven (Vestel, Turkey)
Power Supplies	Power Pac Model 3000 (Bio-Rad, USA) PS 304 (Apelex, France)
Refrigerator	+4 ^{0}C (Arçelik, Turkey)
Shaker	SL 350 (Nüve, Turkey)
SpeedVac	SPD111V (Thermo Scientific, USA)
Spectrophotometer	NanoDrop ND-1000 (NanoDrop, USA)
Thermal Cyclers	Icycler (Bio-Rad, USA) Mycycler (Bio-Rad, USA)
Vortex	Nuvemix (Nuve, Turkey)

4. METHODS

4.1. DNA Extraction from Peripheral Blood

Ten ml blood sample that was stored in K3EDTA tube at 4 ^{0}C was transferred into 50 ml centrifuge tube. After addition of 30 ml RBC lysis buffer in order to lyse erythrocyte membranes, the sample was mixed by vortex and left at 4 ^{0}C for 25 minutes. Then it was centrifuged at 5000 rpm (revolutions per minute) at 4 ^{0}C for 10 minutes. After discarding the supernatant, 10 ml RBC lysis buffer was added and the sample was mixed well to resuspend the pellet. It was centrifuged at 5000 rpm, at 4 ^{0}C for 10 minutes again. The supernatant was discarded and 5 ml nuclei lysis buffer, 40 µl proteinase K and 50 µl SDS (10%) were added to the pellet for overnight incubation at 37 ^{0}C. Thus, cellular proteins were degraded. After the incubation, 10 ml of 2.5 M NaCl was added to the sample. It was mixed well and centrifuged at 5000 rpm, at room temperature for 30 minutes. The supernatant was transferred into a new 50 ml centrifuge tube and approximately 30 ml absolute ethanol (-20 ^{0}C) was added to the sample. Addition of EtOH induced the precipitation of DNA. Clustered DNA was transferred into a 1.5 ml microcentrifuge tube. After air drying, the DNA was dissolved in 100-300 µl of Tris-EDTA (TE) buffer.

4.2. Quantitative Analysis of Extracted DNA

Genomic DNA concentration was measured at 260 nm by using Nanodrop ND-1000 spectrophotometer. In general, it is expected to have an absorbance of 1.0 at 260 nm for 50 µg of pure double stranded DNA. The ratio of A_{260}/A_{280} should be around 1.8 in order to ensure the purity of DNA.

4.3. Mutation Screening

In this study, mutation screening was performed by SSCP, direct Sanger sequencing or restriction enzyme digestion after PCR amplification of the relevant exon.

4.3.1. Mutation Analysis Using PCR-SSCP

PCR-SSCP analysis is one of the methods used to identify sequence differences between individuals. After the target sequence is amplified by PCR, the products are denatured by exposing them to high temperature and then cooling at low temperature. Then, they are run on polyacrylamide gel. Depending on the altered mobility of single strands, even a single nucleotide change is detected (Hayashi, 1991).

4.3.1.1. Polymerase Chain Reaction (PCR). Amplification of six exons of GDAP1 and MPZ genes was performed by using the primers shown in Tables 3.6 and 3.8. PCR reagents and conditions are given in Tables 4.1 and 4.2.

Table 4.1. PCR reagents for exon amplification of GDAP1 and MPZ genes.

Reagents	**Volume (µl)**
DNA (100 ng)	2
10 X $MgCl_2$ Free *Taq* Buffer	2.5
$MgCl_2$ (25mM)	2
dNTPs (10 Mm)	0.6
Forward/Reverse Primers (10pmol/µl)	1 from each
Taq DNA Polymerase (1U/ µl)	0.2
dH_2O	15.7
TOTAL	**25**

Table 4.2. PCR conditions for exon amplification of GDAP1 and MPZ genes.

PCR Conditions	
Inıtial denaturation	4 min at 94 ^{0}C
Denaturation Annealing Extension } 35 cycles	45 s at 94 ^{0}C 30 s at annealing temperature 60 s at 72 ^{0}C
Final Extension	10 min at 72 ^{0}C
Cooling	10 min at 4 ^{0}C

In order to visualize PCR products, three µl PCR amplicon and three µl 6X agarose loading dye were mixed and run on 1% agarose gel in the presence of five µl DNA ladder.

4.3.1.2. Single-Strand Conformation Polymorphism (SSCP). After detection and size control of PCR amplicons, SSCP analysis was performed. For this purpose, two 8% (v/v) acrylamide gels were prepared by using the reagents shown in Table 4.3.

Table 4.3. Reagents used to prepare 8% (v/v) polyacrylamide gels.

Reagents	**Volume**	
	With glycerol	**Without glycerol**
30% Acrylamide Stock (29:1)	16 ml	16 ml
5 X TBE buffer	7.2 ml	7.2 ml
10% APS	600 µl	600 µl
TEMED	60 µl	60 µl
Glycerol	2 ml	-
dH_2O	35 ml	37 ml

TEMED and APS were added lastly. Then, the solution was immediately poured between two glasses separated by spacers. The combs were placed and the gels were left to polymerize for at least half an hour. Meanwhile, seven µl of amplicons were mixed with seven µl of 6X denaturing loading dye. The mixtures were kept at 95 ^{0}C for five minutes

and then on ice for another five minutes. Seven µl of denatured PCR amplicon was loaded to the gel without glycerol, while the remaining seven µl was loaded to the gel with glycerol. Both gels were run in 0.5 X TBE buffer at 110 V for 16 hours.

4.3.1.3. Silver Staining. DNA fragments that were run on polyacrylamide gels were visualized by silver staining. Solutions used to stain the gels are given in Table 3.5. The gels were incubated with $AgNO_3$ solution (Solution 1) for 10 minutes. Then, they were washed and incubated with Solution 2. Both incubation steps took place on shaker. When the bands were visible, the gels were placed in transparent folders and the bands obtained for patients and controls were compared on white light box.

4.3.2. Mutation Analysis Using PCR-Sanger Sequencing

A combination of PCR and Sanger sequencing methods were used to detect variations in the GDAP1, GJB1 and PMP22 genes . After amplification of targeted DNA regions, the amplicons were sequenced by Macrogen (Seoul, Korea). The data for Sanger sequencing was compared with the sequences in general databases (NCBI, Ensemble) by using ApE program (A Plasmid Editor, Wayne Davis, University of Utah) and deviations from the normal nucleotide sequence were determined.

4.3.2.1. Polymerase Chain Reaction (PCR). Amplification of coding sequences of the GDAP1, GJB1 and PMP22 genes was performed in a total volume of 50 µl. The reagents used in the reactions are given in Table 4.4. Primers that were used are given in Tables 3.6, 3.7 and 3.10. PCR conditions were the same as for the conditions used for PCR-SSCP analysis (Table 4.2).

PCR products were visualized and confirmed by agarose gel electrophoresis. Three µl PCR amplicon and three µl 6X agarose loading dye were mixed and run on 1% agarose gel. Five µl DNA ladder was also used to confirm the size of amplicons.

Table 4.4. PCR reagents for exon amplification of GDAP1, GJB1 and PMP22 genes.

Reagents	**Volume (µl)**
DNA (100 ng)	2
10 X $MgCl_2$ Free *Taq* Buffer	5
$MgCl_2$ (25mM)	4
dNTPs (10 Mm)	1.2
Forward/Reverse Primers (10pmol/µl)	2 from each
Taq DNA Polymerase (1U/ µl)	0.4
dH_2O	33.4
TOTAL	**50**

4.3.2.2. Sanger Sequencing. After confirmation by agarose gel electrophoresis,10 ng/100 bases of the PCR product with 10 pmole/µl of each primer (forward/reverse) for each reaction were sent to Macrogen (Seoul, Korea) for Sanger sequencing.

4.3.3. Mutation Analysis by Restriction Enzyme Digestion

The founder R148X mutation in NDRG1 gene was tested in two affected brothers by restriction enzyme analysis.

4.3.3.1. Polymerase Chain Reaction (PCR). The targeted region of NDRG1 was amplified by using the reagents shown in Table 4.1. PCR conditions were demonstrated in Table 4.2. The primers used in the reaction are given in Table 3.9.

After amplification, the length of the PCR product was confirmed by running on 1% agarose gel. The mixtures of three µl PCR amplicons and three µl 6X agarose loading dye were run on the gel with five µl DNA ladder solution.

4.3.3.2. Restriction Enzyme Digestion. After confirmation of the presence and size of PCR products by agarose gel electrophoresis, they were digested by *TaqI*. It was performed in restriction enzyme digestion mix shown in Table 4.5 at 65 ^{0}C for four hours.

Table 4.5. Restriction enzyme digestion mix.

Reagents	**Volume (µl)**
PCR Product	10
10 X Buffer (TaqI)	2
TaqI (10U/ µl)	2
dH_2O	16
TOTAL	**30**

After restriction enzyme digestion, five µl of PCR amplicons were mixed with five µl loading dye and loaded on 4% agarose gel. The restriction patterns of samples were analyzed under UV light.

4.4. Exome Sequencing

In this study, seventeen patients who were previously found to be negative for CMT1A duplication and HNPP deletion were exome sequenced.

After DNA extraction from peripheral blood samples, the concentration of them was determined. About 200 ug of DNA samples were sent to the Department of Molecular and Human Genetics of Baylor College of Medicine (Houston, Texas, USA) for exome sequencing. The reaction was performed on Illumina platform. General protocol of exome sequencing process is described below.

Firstly, a genomic DNA library is prepared. Several micrograms of purified genomic DNA are fragmented by sonication. The fragmented DNA molecules that are approximately 200-400 bp are repaired enzymatically and adaptors are ligated to the ends. Exome capture is performed by using solution based capture method. DNA fragments are hybridized to biotinylated probes bound to streptavidin coated magnetic beads. Several washing steps are performed in order to remove non-specific bindings. Hence, an eluted library pool is obtained. Enrichment is performed by amplification of adaptor ligated DNA fragments. Finally, sequencing reaction is carried out (Bamshad *et al*., 2011).

After sequencing, sequence alignment is performed by Burrows-Wheeler Aligner (BWA) software. Sequence alignment/map (SAM) files that contain the raw data are converted into binary alignment/map (BAM) files. Then, a pileup format of sequence reads that consist of every base observed at each aligned position is generated. Because the majority of bases are the same with the reference, a process called variant calling is carried out to obtain a list of variations. Once variants are obtained, variant annotation is performed prior to the evaluation of sequencing data to identify causative variants. For this purpose, a variety of databases providing information about polymorphisms, human diseases, causative mutations, etc. are used (Dolled-Filhart *et al*., 2013).

After evaluation of exome sequencing data, possible disease causing variants were selected and their segregation was analyzed in patients and their family members (if available) by PCR-Sanger sequencing method. PCR was performed in a total volume of 50 µl by using conditions shown in Table 4.2. In order to amplify targeted regions in SBF2 genes for patients P114 and P378 as well as in DNM2 gene for patient P851, the reagents shown in Table 4.6 were used. Amplification of targeted regions in the DNM2 gene of patient P163 and the GMFG and FLNA genes of patient P212 was carried out by using the reagents given in Table 4.7. Amplification reaction of all the remaining variants contained the reagents shown in Table 4.4. Primers used to amplify each targeted region are given in Tables 3.11-3.18.

Table 4.6. Reagents for PCR amplification of targeted regions in SBF2 gene of patients P114 and P378 and in DNM2 gene of patient P851.

Reagents	**Volume (µl)**
DNA (100 ng)	2
10 X $MgCl_2$ Free *Taq* Buffer	5
$MgCl_2$ (25mM)	3
dNTPs (10 Mm)	1.2
Forward/Reverse Primers (10pmol/µl)	2 from each
Taq DNA Polymerase (1U/ µl)	0.4
dH_2O	34.4
TOTAL	**50**

Table 4.7. Reagents for PCR amplification of targeted regions in DNM2 gene of patient P163 and the GMFG and the FLNA genes of patient P212.

Reagents	**Volume (µl)**
DNA (100 ng)	2
10 X $MgCl_2$ Free *Taq* Buffer	5
$MgCl_2$ (25mM)	5
dNTPs (10 Mm)	1.2
Forward/Reverse Primers (10pmol/µl)	2 from each
Taq DNA Polymerase (1U/ µl)	0.4
dH_2O	32.4
TOTAL	**50**

The size of the PCR products was confirmed by electrophoresis on 1% agarose gel and Sanger sequencing was performed in Macrogen (Seoul, Korea).

5. RESULTS

In order to understand the genetic basis of CMT disease in our population, exome sequencing and mutation screening in selected target genes were performed in this study. The results obtained from these approaches are demonstrated in detail below.

5.1. Exome Sequencing

Seventeen unrelated CMT patients whose clinical features were given in Table 5.1 underwent exome sequencing and the resulting data for each individual was analyzed. The first filtering of the data for all seventeen patients was performed to identify the variations in the known CMT genes. heterozygous/homozygous state of variants, according to the most probable inheritance pattern in the affected individuals, was the first filtering criteria. The alterations which have been previously reported as polymorphisms were filtered out. Synonymous variations were excluded and nonsynonymous variants were selected. The read depth, the ratio of total reads and variant reads, was used as another criteria and variants with a depth ratio below 30 were eliminated. In addition, novel genes, MICAL1, TRIM2, BICD2, ATP7A, that have been recently found to be associated with CMT disease were checked in the exome data of every patient. However, no variations in these genes were identified. The numbers of variations for each patient obtained after each filtering step are given in Table 5.2. After selection of candidate variants, their presence in the patients was confirmed by conventional Sanger sequencing and their segregation was investigated among family members. If no candidate causative variation can be identified for an individual, the data was evaluated again by focusing on other genes which have not been previously associated with CMT. The variations found in the genes whose function seemed to be related with CMT pathogenesis were selected. The expression levels in PNS of these genes were also considered using miRBase and RNAdb databases. The segregation analysis of the selected variants was completed for nine of seventeen patients, and the results are summarized in Table 5.3.

Table 5.1. Clinical characteristics of seventeen CMT patients that were exome sequenced.

	DOB	Age of onset	Severity	Median MNCV	Type of the disease	Mode of inheritance	Additional features	Consanguinity
P38	1982	11-20 y	-	-	CMT2	AR	-	Consanguineous
P114	1981	11-20 y	Normal mobility Upper limbs: normal	50 m/s	CMT2	AD	Tremor Hammer toe	-
P138	1956	21-40 y	Abnormal mobility- no aids required Upper limbs: distal weakness (hand wrist)	30 m/s	CMT1	AD	Sensory ataxia	-
P157	1959	21-40 y	Abnormal mobility- no aids required Upper limbs: normal	60 m/s	CMT2	AD	-	-
P163	1926	>40 y	Normal mobility Upper limbs: normal	-	CMT2	AD	-	-
P171	1950	>40 y	Normal mobility Upper limbs: normal	-	CMT2	AR	Sensory ataxia: mild	Consanguineous
P209	1949	21-40 y	Normal mobility Upper limbs: distal weakness (hand wrist)	51.6 m/s	CMT2	AD	-	-
P212	1972	11-20 y	Abnormal mobility- no aids required Upper limbs: distal weakness (hand wrist)	40 m/s	CMT2	AR	Scoliosis Bilateral temporal muscles: atrophic	Consanguineous

Table 5.1. Clinical characteristics of seventeen CMT patients that were exome sequenced (cont.).

P245	1983	2-10 y	Normal mobility Upper limbs: normal	45.4 m/s	CMT2	AR	-	Consanguineous
P254	1970	21-40 y	Abnormal mobility- walking aids used Upper limbs: normal	59.8 m/s	CMT2	AR	-	Consanguineous
P261	1971	2-10 y	Abnormal mobility- walking aids used Upper limbs: distal weakness (hand wrist)	57.8 m/s	CMT2	AR	Tremor Sensory ataxia Cerebellar dysfunction Scleral telangiectasia	Consanguineous
P266	1976	11-20 y	abnormal mobility- no aids required Upper limbs: distal weakness (hand wrist)	60 m/s	CMT2	AR	-	-
P313	2000	Delayed motor milestones walking >18 months	Abnormal mobility-no aids required Upper limbs: distal weakness (hand-wrist)	55.6 m/s	CMT2	AR	-	Consanguineous
P343	1973	11-20 y	Abnormal mobility- no aids required Upper limbs: distal weakness (hand wrist)	44 m/s	CMT2	AR	Tremor	-

Table 5.1. Clinical characteristics of seventeen CMT patients that were exome sequenced (cont.).

P378	1965	11-20 y	Abnormal mobility- no aids required Upper limbs: distal weakness (hand wrist)	48 m/s	CMT2	AD	Tremor	-
P738	-	2-10 y	Abnormal mobility- no aids required Upper limbs: distal weakness (hand wrist)	15 m/s	CMT1	AD	Tremor Sensory ataxia	-
P851	1984	11-20 y	Normal mobility Upper limbs: distal weakness (hand wrist)	39 m/s	CMT2	AD	-	-

Table 5.2. Number of variations determined by filtering criteria for exome sequencing data analysis.

Patient	Total number of variants	Synonymous SNVs filtered out	Not annotated SNP	Zygosity	Read depth (>30)	Clinic and function prediciton algorithms (SIFT, polyphen, LRT, OMIM)	Analysis of CMT genes	Sanger sequencing validation
P38	22789	11765	1278	97	38	2		Not analyzed
P114	24366	12979	1947	1768	348	2		AARS
P138	23745	12397	1612	1449	234	1		GJB1
P157	23688	12348	1569	1384	191	3		Not analyzed
P163	24172	12730	1832	1660	314	232	2	DNM2, LRSAM1
P171	24344	12986	2226	204	97	2		Not analyzed
P209	23873	12595	1677	1522	295	1		MFN2
P212	31640	16466	2216	203	84	5		SIGMAR1, CXorf23
P245	23934	12590	1729	176	73	4		FGD4, KIF1B,THBS4
P254	24262	12904	2131	199	85	2		Not analyzed
P261	23529	12389	1815	192	95	2		Not analyzed
P266	24116	12578	1659	187	94	3		Not analyzed
P313	23952	12734	2087	203	93	1		MFN2
P343	24027	12659	1828	173	61	3		Not analyzed
P378	24067	12671	1823	1663	286	3		Not confirmed
P738	24339	12878	1958	1775	317	1		Not analyzed
P851	24582	13041	2144	1953	343	1 – polymorphism ?	1	DNM2

Table 5.3. Candidate variants for nine CMT patients.

	Cytoband	Location	Gene	Ref	Var	Zyg	Annotation	Nt Change	Aa Change	SNP Annotation
P114	16q22.1	70302259	AARS	C	T	Het	nonsynonymous SNV	c.G986A	p.R329H	rs267606621
	11p15.4	9809202	SBF2	TTT	-	Het	nonframeshift deletion	c.5014_5016 del	p.1672_1672 del	-
P138	Xq13.1	70443540	GJB1	A	G	Hem	splicing mutation	c.-18A>G	-	-
	10q23.2	88476170	LDB3	T	C	Het	nonsynonymous SNV	c.T1333C	p.S445P	-
	2p23.3	26717905	OTOF	C	T	Het	nonsynonymous SNV	c.G802A	p.G268S	-
	2q11.2	97462877	CNNM4	G	A	Het	nonsynonymous SNV	c.G1531A	p.E511K	-
	Xq28	153187103	ARHGAP4	C	A	Hem	nonsynonymous SNV	c.G227T	p.S76I	-
P163	19p13.2	10904424	DNM2	G	A	Het	nonsynonymous SNV	c.G1021A	p.E341K	-
	9q33.3	130223464	LRSAM1	G	A	Het	nonsynonymous SNV	c.G334A	p.E112K	-
P209	1p36.22	12057358	MFN2	T	G	Het	nonsynonymous SNV	c.T479G	p.V160G	-

Table 5.3. Candidate variants selected for nine CMT patients (cont.).

P212	16q22.1	70286631	AARS	T	A	Het	nonsynonymous SNV	c.A2900T	p.K967M	rs35744709
	19q13.2	39819681	GMFG	C	T	Het	nonsynonymous SNV	c.G316A	p.G106R	-
	Xq28	153595129	FLNA	C	T	Hem	nonsynonymous SNV	c.G958A	p.V320M	-
	9p13.3	34637357	SIGMAR1	C	A	Hom	nonsynonymous SNV	c.G212T	p.G71V	-
	Xp22.12	19983811	CXorf23	G	A	Hem	nonsynonymous SNV	c.C625T	p.R209C	-
P245	12p11.21	32778699	FGD4	C	T	Het	stopgain SNV	c.C1747T	p.R583X	-
	1p36.22	10381802	KIF1B	T	C	Het	nonsynonymous SNV	c.T2107C	p.W703R	-
	19q13.41	52887145	ZNF880	TA	-	Hom	Frameshift deletion	c.312_313 del	p.104_105 del	-
	5q14.1	79378366	THBS4	A	G	Hom	nonsynonymous SNV	c.A2822G	p.N941S	-
P313	1p36.22	12057405	MFN2	G	A	Hom	nonsynonymous SNV	c.G526A	p.G176S	-
P378	11p15.4	9806690	SBF2	G	T	Het	nonsynonymous SNV	c.C5213A	p.S1738Y	-
	1p34.3	35370888	DLGAP3	A	G	Het	nonsynonymous SNV	c.T97C	p.Y33H	-
	3p25.3	10975821	SLC6A11	G	T	Het	nonsynonymous SNV	c.G1544T	p.C515F	-
P851	19p13.2	10904475	DNM2	G	A	Het	nonsynonymous SNV	c.G1072A	p.G358R	rs267606772

5.1.1. P114

Nerve conduction velocity studies (50 m/s) and the segregation pattern of the disease in the family (Figure 5.1) indicated that patient P114 was presenting autosomal dominant CMT2.

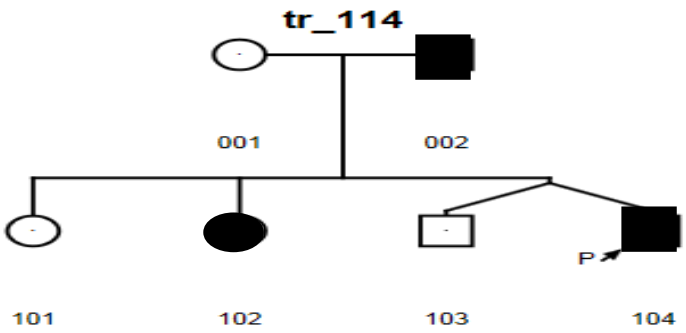


Figure 5.1. Pedigree of the family P114. “P” indicates the proband.

Two variations were candidates for patient P114, a novel *SBF2* alteration and an alteration in the AARS gene which has been previously reported to cause CMT (Latour *et al*., 2010). The patient was heterozygous for the SBF2 variant, but mutations in this gene generally cause autosomal recessive CMT (CMT4B). As shown in Figure 5.2, the unreported heterozygous *SBF2* variant, c.5014_5016del, was present in the proband, in the unaffected mother (001) and unaffected dizygotic twin (103). On the other hand, it was absent in the affected father (002) and in two sisters -one was affected (102), the other was unaffected (101). Consequently, because the variant was found in the heterozygous state in the patient and is not segregating with the disease in the family, it was excluded.

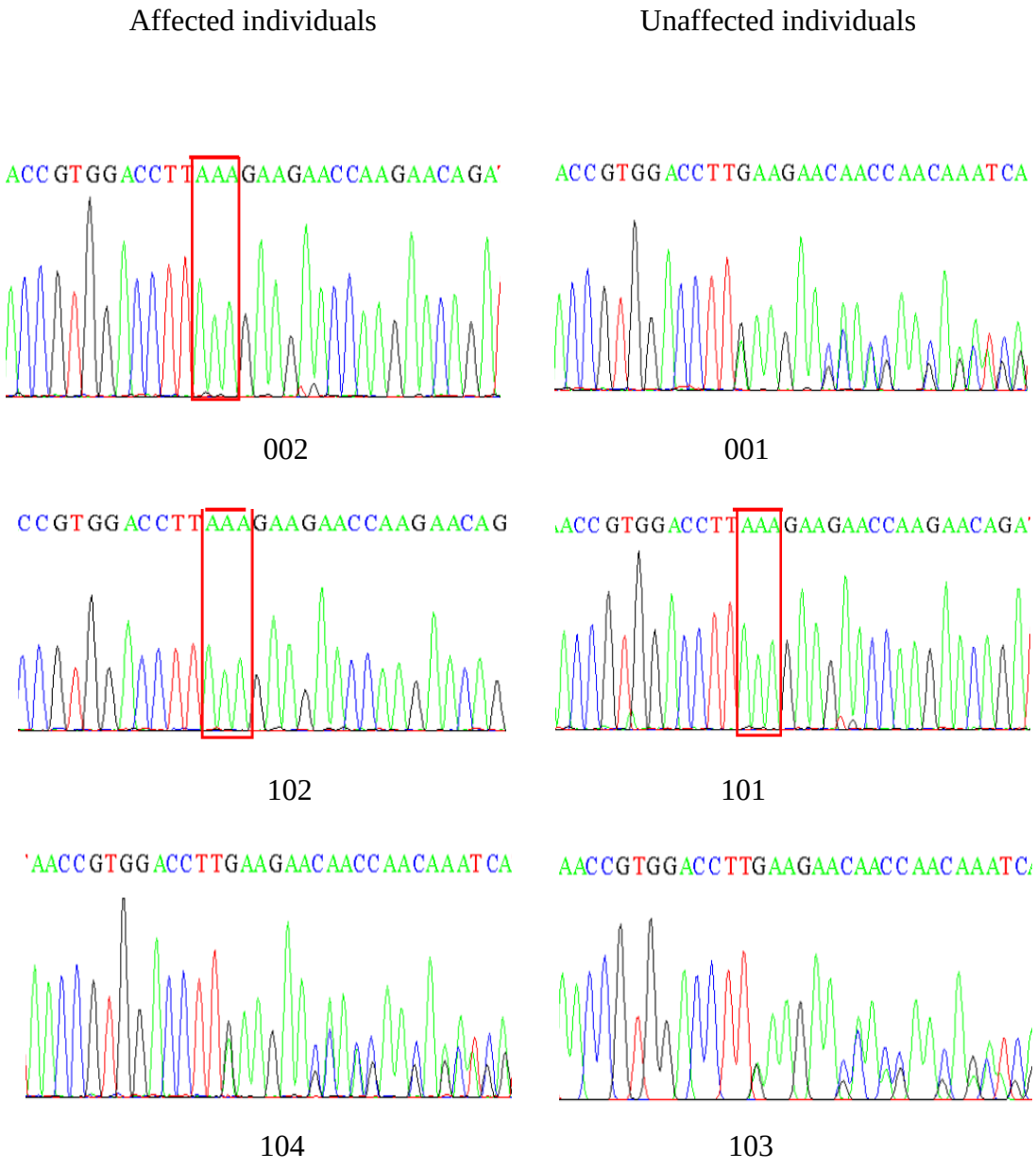


Figure 5.2. Chromatograms showing the *SBF2* variant, c.5014_5016del, in patient P114 and family members. The region in red rectangel is the deletion site that was absent in the individuals 001, 103 and 104. Overlapping peaks after deletion shows heterozygosity for the alteration in individuals 001, 103, and 104.

As shown in Figure 5.3, the heterozygous c.G986A variant in AARS gene (rs267606621) was present in all of the affected members of family P114, while the unaffected individuals were not presenting the variation. The findings indicated that this variation was probably the cause of the disease in family P114.

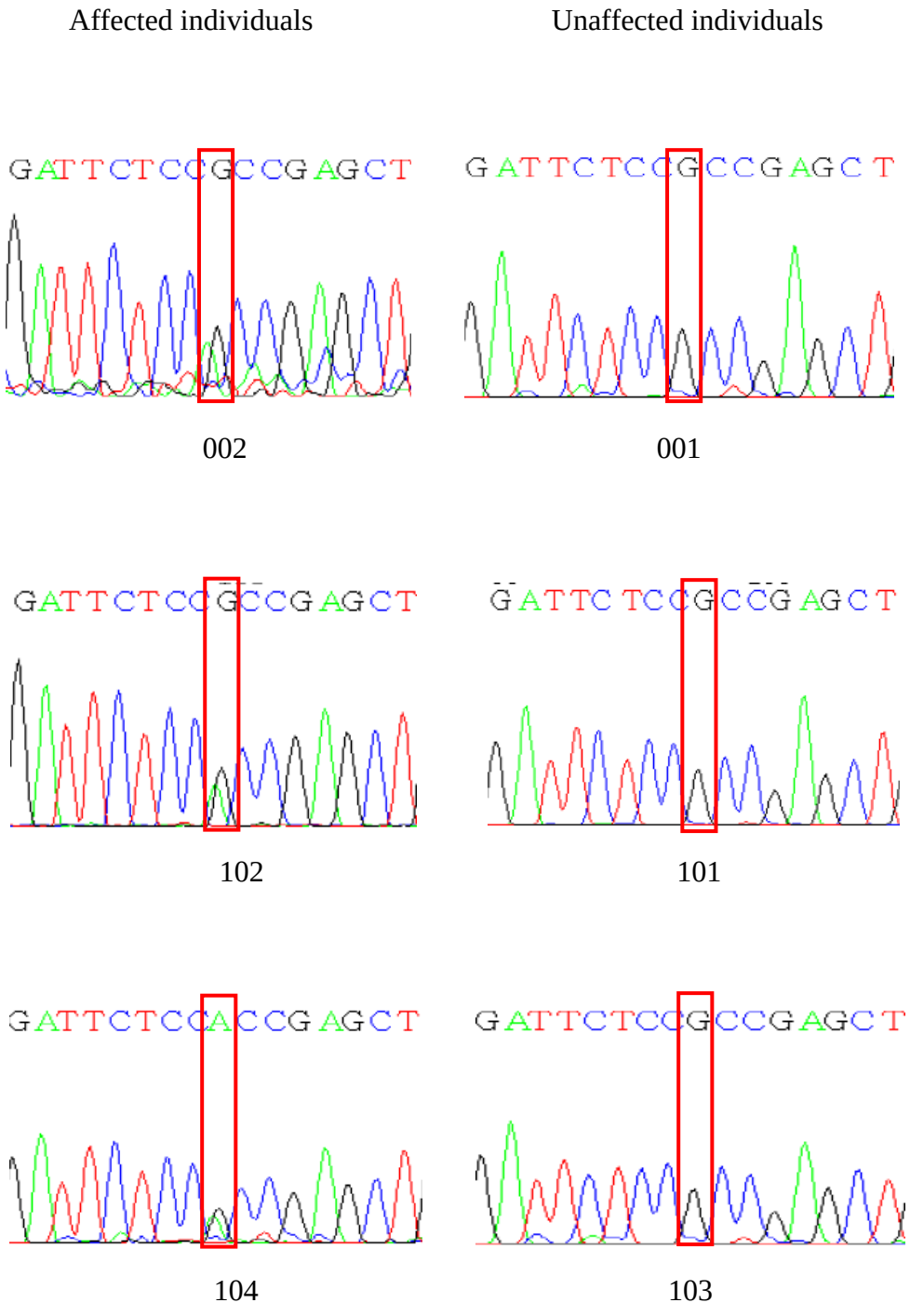


Figure 5.2. Chromatograms showing the AARS variant, c.G986A, in patient P114 and family members. The region in red rectangel is the alteration site. The individuals 002, 102 and 104 carried the alteration, while the individuals 001, 101 and 103 were homozygous for the wild type allele.

5.1.2. P138

Patient P138 had a median NCV of 30 m/s indicating presence of the demyelinating or intermediate type of the disease (CMT1). The pedigree of the family (Figure 5.3), was not helpful to determine the inheritance pattern.

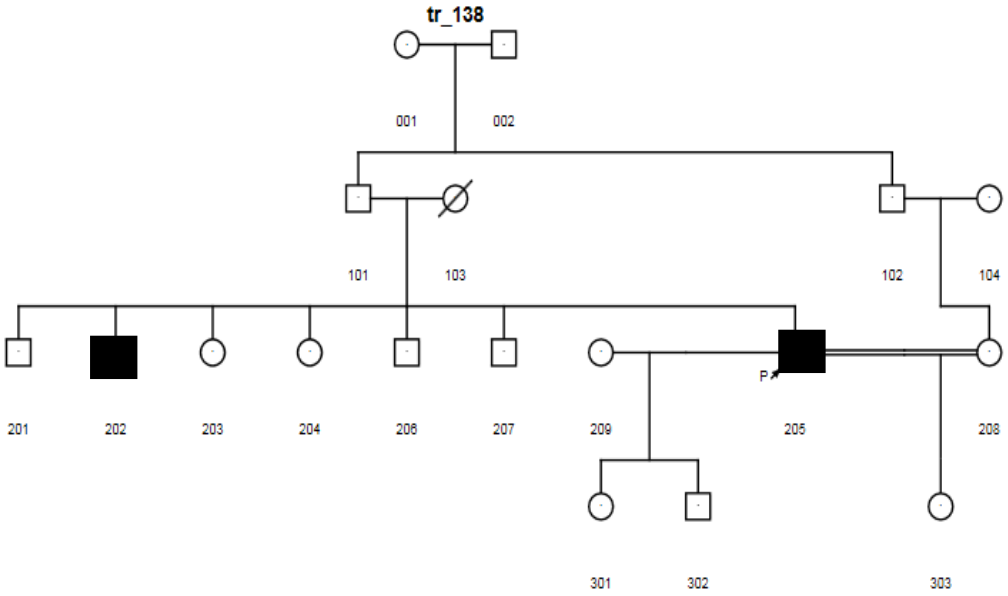


Figure 5.3. Pedigree of the family P138. One of the proband's probable daughters (301) was reported to have abnormal EMG.

The variations that have been chosen as possible disease causing alterations for patient P138 are shown in Table 5.1. The strongest candidate was the novel variant, the c.-18A>G alteration, located at the splice site of *GJB1*. Since the data obtained did not include this variation initially, four other novel variants located in the genes, LDB3, OTOF, CNNM4, and ARHGAP4 had been selected as candidates. Although these genes have not been associated with CMT up to date, they were chosen due to their function and expression profile as possible causative genes.

Unfortunately, except from the patient's DNA, there was only one other DNA sample from a female relative. The degree of relativeness of this female in the family was

not known. It was supposed to belong to the daughter of the patient whose EMG was reported as abnormal.

As demonstrated in Figure 5.4, the splicing variation, c.-18A>G, found in *GJB1* was present in the patient (205) and also in his daughter who is supposed to be affected (301). As expected, the proband was hemizygous, while the female relative was heterozygous for the variation. The alteration was located at 3' end of intron 1 of *GJB1*, as the second last base. It was predicted to affect splicing by abolishing the binding site of SRp40 protein by Human Splicing Finder Tool (Desmet el al., 2009).

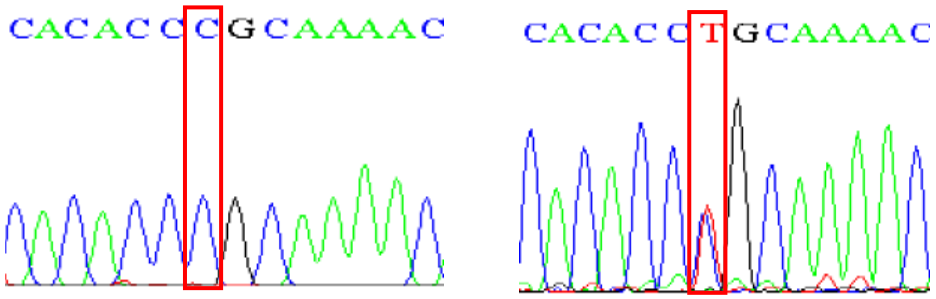


Figure 5.4. Chromatograms showing the hemizygous *GJB1* variant, c.-18A>G, in the proband (205) and heterozygosity for his probable daughter (301).

Sanger sequencing results for other four variations located in possible candidate genes are shown in Figure 5.5. Accordingly, the heterozygous c.T1333C variation in LDB3 gene was not present in both individuals indicating that it was a false positive. The heterozygous c.G802A variant in OTOF gene was present in the patient, but not present in his daughter. Because it was not segregating with the disease in the family, it was excluded. However, both the heterozygous c.G1531A alteration in CNNM4 gene and the hemizygous c.G227T variant in ARHGAP4 gene were present in the patient and also in the affected daughter.

These results showed that the alterations in CNNM4 and ARHGAP4 could be causative in family P138 like the splice site variation in the GJB1 gene.

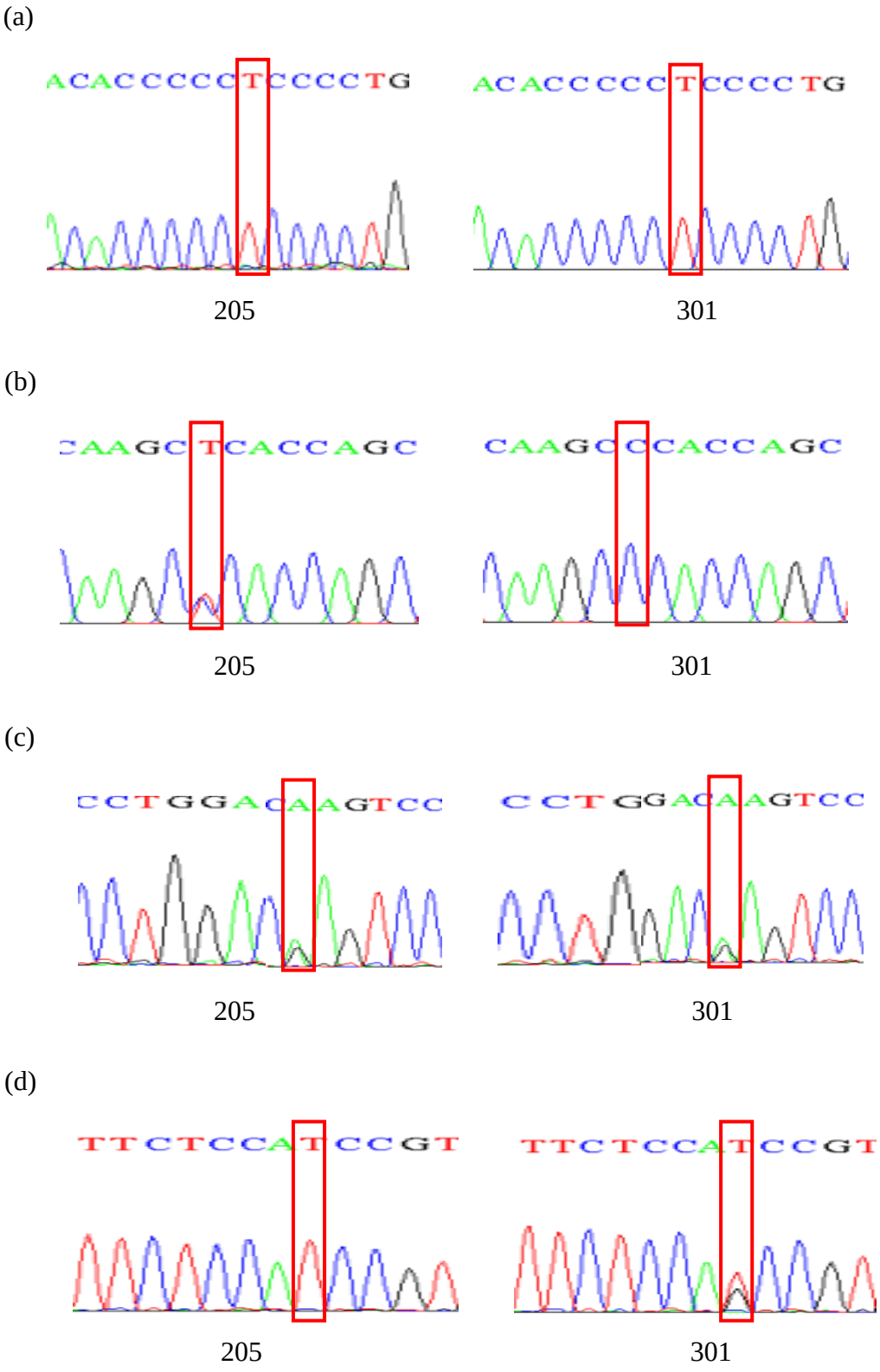


Figure 5.5. Chromatograms showing the (a) absence of c.T1333C variant in LDB3 and the (b) presence of the c.G802A variant in OTOF, (c) the c.G1531A variant in CNNM4, (d) the hemizygous c.G227T variant in ARHGAP4 in the proband (205) and his daughter (301).

5.1.3. P163

Patient P163 had autosomal dominant CMT2. The pedigree of the family is shown in Figure 5.6.

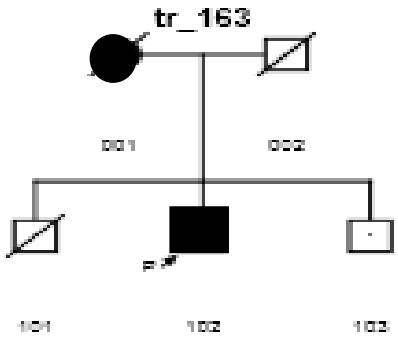
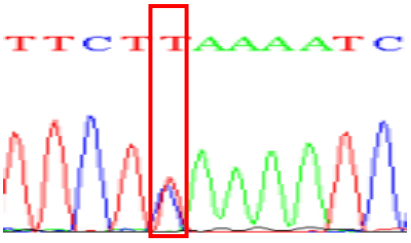


Figure 5.6. Pedigree of the family P163.

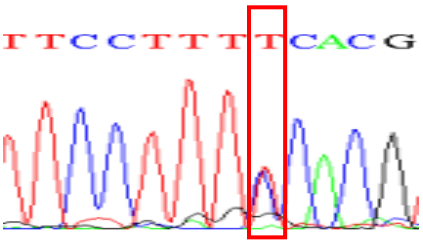
The variations selected for patient P163 were the heterozygous c.G1021A alteration in DNM2 gene and the heterozygous c.G334A alteration in LRSAM1 gene. Since there was no DNA sample from relatives, PCR-Sanger sequencing method was performed only for the proband, As demonstrated in Figure 5.7, the variation in *DNM2* was observed in the patient as well as the variant in *LRSAM1*. *DNM2* is responsible for CMT2M and *LRSAM1* is associated with CMT2P.

(a)



102

(b)



102

Figure 5.7. Chromatograms for the reverse strand showing (a) the heterozygous c.G1021A variant in *DNM2*, (b) the heterozygous c.G334A variant in *LRSAM1* in the proband (102).

5.1.4. P209

Patient P209 presented autosomal dominant CMT2 with a median NCV of 51.6 m/s and the pedigree analysis (Figure 5.8).

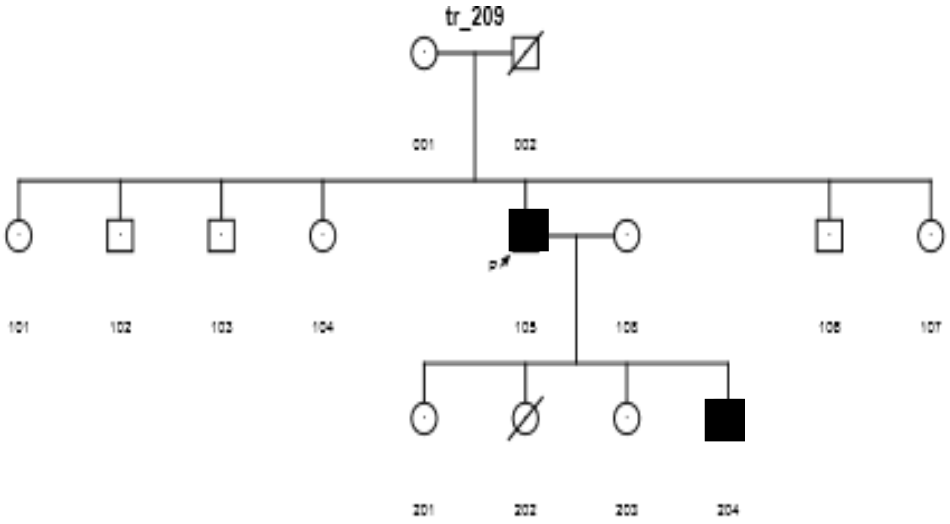


Figure 5.8. Pedigree of the family P209.

The variant that has been chosen for this patient was the novel heterozygous c.T479G alteration in the MFN2 gene. It was present in both the patient (105) and his affected son (204) indicating that it could be the possible disease causing variation in family P209 (Figure 5.9).

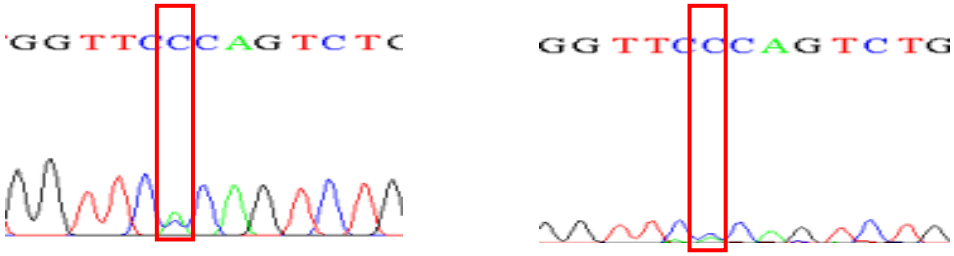


Figure 5.9. Chromatograms showing the *MFN2* variant, c.T479G, in the heterozygous patient (105) and his affected son (204) in the reverse strand.

5.1.5. P212

According to the median NCV (40 m/s) and a report of consanguinity among the parents, patient P212 (203) was thought to present autosomal recessive CMT2. However, because the knowledge for the generation of individual 002 was not enough, it was also considered that there could be an incomplete penetrance (Figure 5.10).

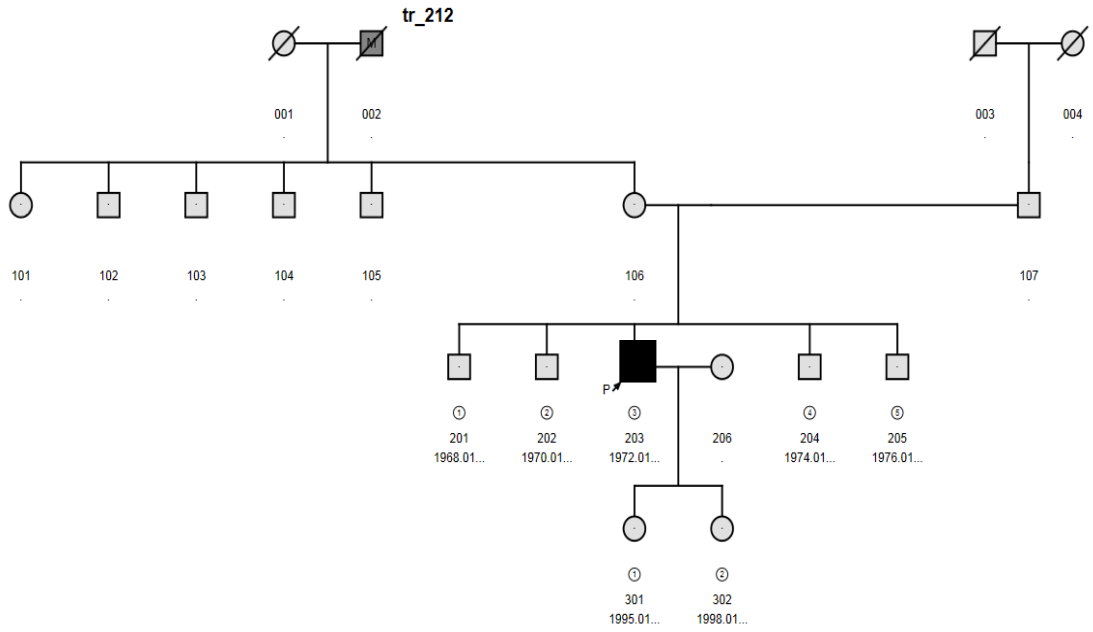


Figure 5.10. Pedigree of the family P212. The consanguinity between individuals 001/002/003/004 was not demonstrated. Individual 002 had pes cavus, "M" indicates his mild phenotype.

For patient P212, a variation which has been previously reported as a SNP in the AARS gene and four novel variations in four novel genes, GMFG, FLNA, SIGMAR1 and CXorf23, were selected as possible disease causing candidates. The SNP in *AARS* was chosen because of its low MAF (0.005) and the predictions as "deleterious" and "probably damaging" in SIFT and PolyPhen2 databases, respectively.

As shown in Figure 5.11, the heterozygous c.A2900T variation (rs35744709) in the AARS gene was absent in both the patient (203) and the unaffected brother (201). Thus, this was a false positive result of exome sequencing.

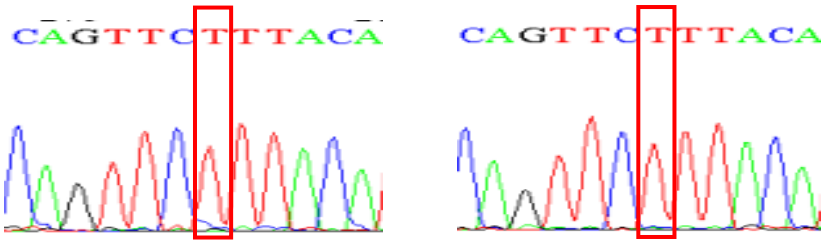


Figure 5.11. Chromatograms showing the absence of *AARS* variant, c.A2900T (rs35744709), in the patient (203) and his unaffected brother (201).

As demonstrated in Figure 5.12, the heterozygous c.G316A alteration in GMFG gene was not present in both the patient and his unaffected brother, indicating another false positive result. The hemizygous c.G958A variation in FLNA gene was present in the patient and also in the unaffected brother. Thus, it was excluded, too. The patient was homozygous for the c.G212T variation in the SIGMAR1 gene , and the unaffected brother was presenting the alteration in heterozygous state. The hemizygous c.C625T variant in CXorf23 gene was present in the patient, and absent in the unaffected brother. In summary, variations in SIGMAR1 and CXorf23 genes were determined to be candidates for patient P212.

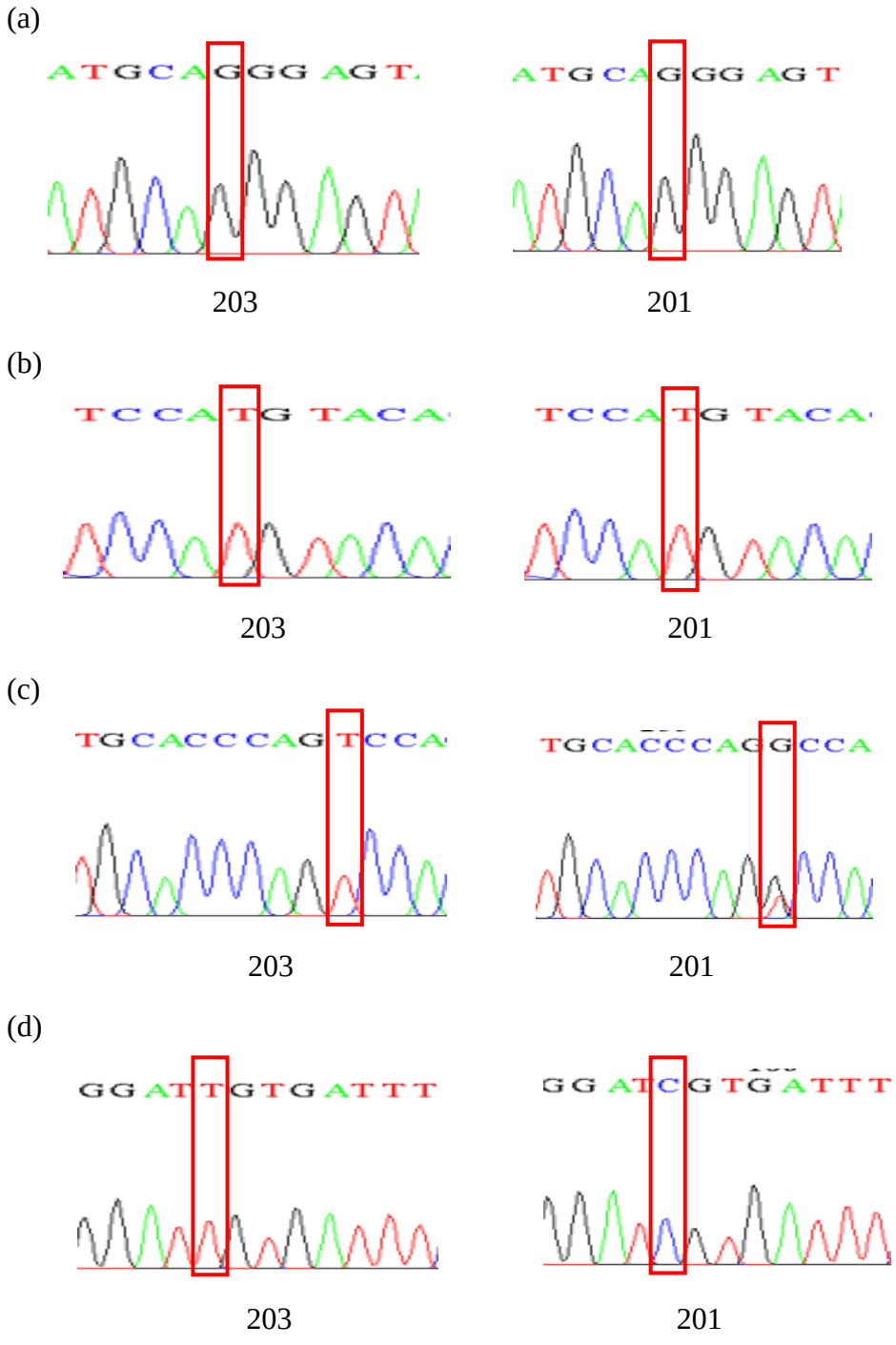


Figure 5.12. Chromatograms showing (a) the absence of c.G316A variant in GMFG, (b) the c.G958A variant in FLNA, (c) the presence of c.G212T variant in SIGMAR1, (d) the c.C625T variant in CXorf23 in the proband (203) and his unaffected brother (201).

5.1.6. P245

Patient P245 had an autosomal recessive CMT2 (AR CMT2) with a median NCV of 45.4 m/s and the following complicated pedigree (Figure 5.13).

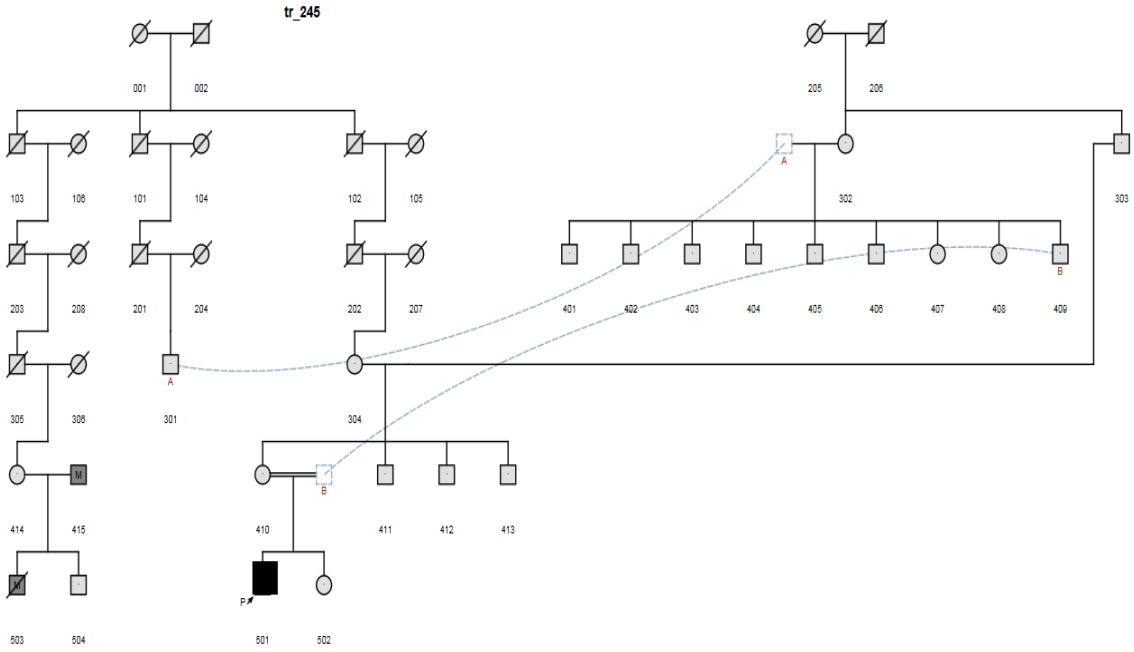


Figure 5.13. Pedigree of the family P245. Individuals 415 and 503 had pes cavus, but they were not reported as affected.

The variations chosen for patient P245 were the novel heterozygous c.C1747T alteration in FGD4 gene (CMT2H), the novel heterozygous c.T2107C variant in KIF1B gene (CMT2A1) and two other novel homozygous alterations in two novel genes, ZNF880 and THBS4. Unfortunately, there was no DNA sample from relatives and the candidate variations could only be tested in the proband (501). As shown in Figure 5.14, the patient was heterozygous for the c.C1747T alteration in *FGD4* as well as the c.T2107C variant in *KIF1B*. He was homozygous for the c.A2822G alteration in *THBS4*. However, the homozygous c.312_313del deletion in ZNF880 gene was not observed indicating that it was a false positive.

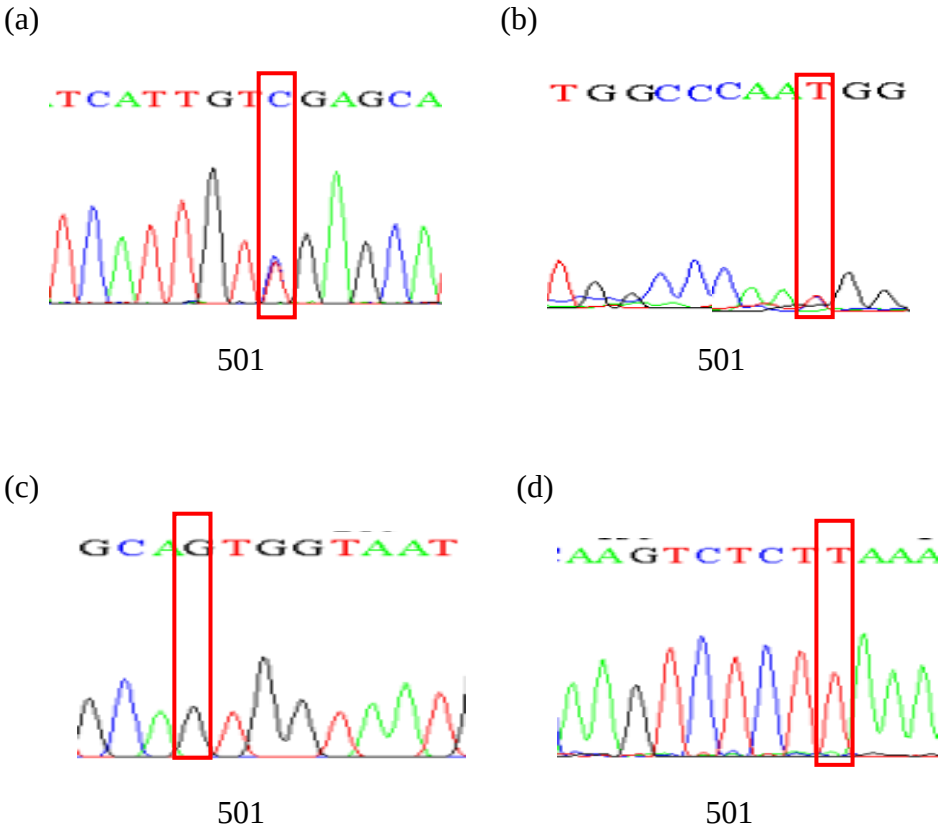


Figure 5.14. Chromatograms showing (a) the c.C1747T variant in *FGD4*, (b) the c.T2107C variant in *KIF1B*, (c) the c.A2822G variant in *THBS4*, (d) the absence of homozygous c.312_313 deletion in *ZNF880* in patient P245 (501).

5.1.7. P313

Patient P313 presented autosomal recessive CMT2 with a severe phenotype. Median NCV was recorded as 55.6 m/s and the pedigree is given in Figure 5.15.

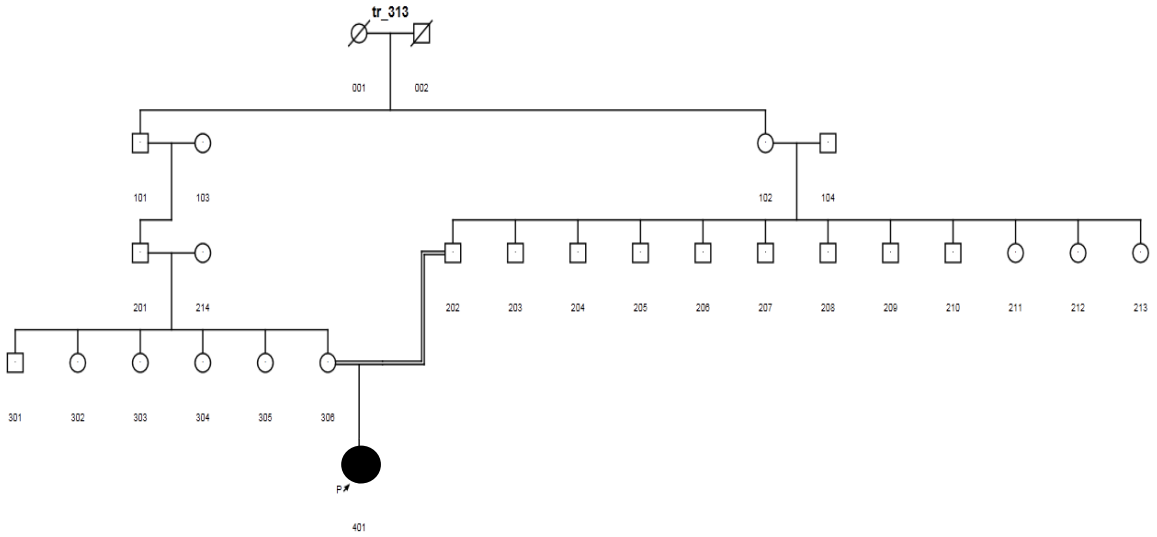


Figure 5.15. Pedigree of the family P313.

The only variation remained after filtration for patient P313 was the homozygous c.G526A alteration in the MFN2 gene. It has not been previously reported. As demonstrated in Figure 5.16, its presence was confirmed in the patient (401). The unaffected father (202), mother (306) and aunt (211) were all heterozygous for this alteration, thus, the variation was probably causative in patient P313 and the other family members tested were carriers.

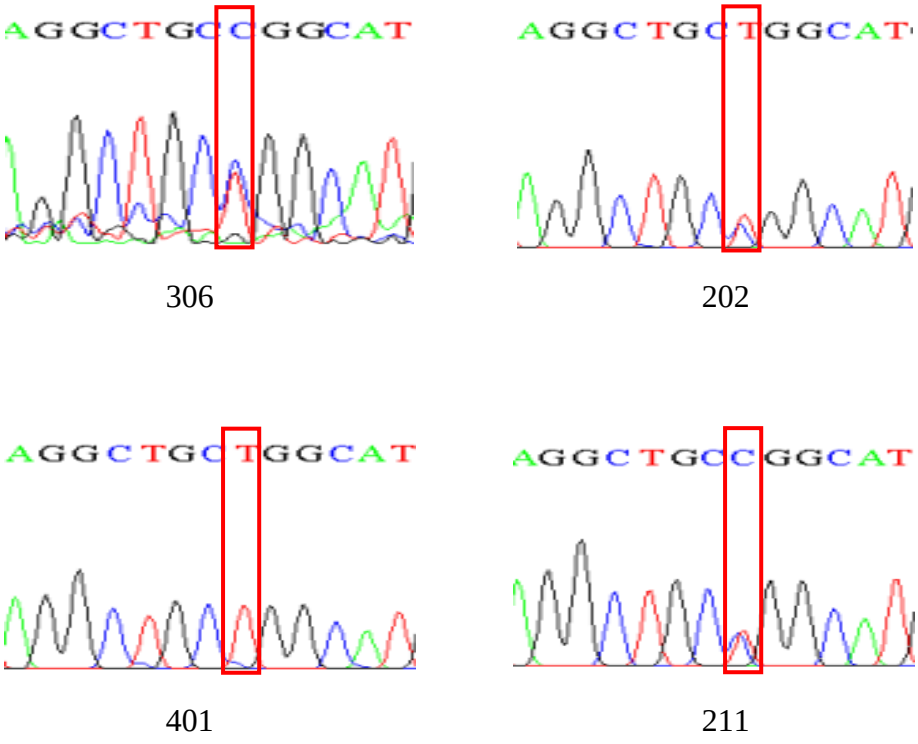


Figure 5.16. Chromatograms showing the *MFN2* variant, c.G526A, in patient P313 (401) in homozygous state and her heterozygous healthy family members.

5.1.8. P378

According to the median NCV (48 m/s) and the inheritance pattern in the family (Figure 5.17), patient P378 presented autosomal dominant CMT2.

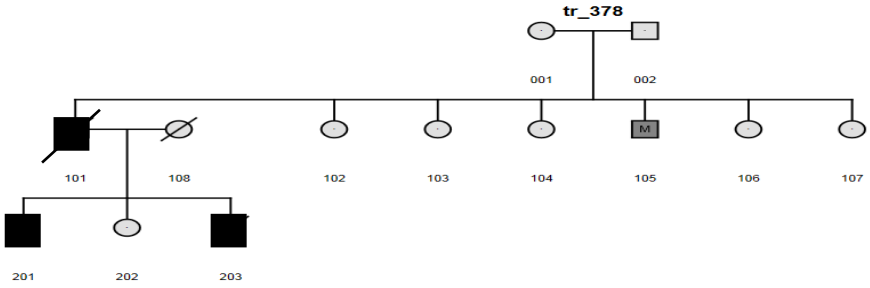


Figure 5.17. Pedigree of the family P378. Individual 105 had a mild phenotype.

For patient P378, the novel heterozygous c.C5213A variation in the SBF2 gene and two novel variations in two novel genes, SLC6A11 and DLGAP3, were selected. The variation in *SBF2* was observed only in the proband (203), while the affected (201) and unaffected (202) brothers did not have the alteration. Similarly, the heterozygous c.G1544T variant in *SLC6A11* was present in the proband and absent in two brothers. Conversely, the proband and his affected and unaffected brothers had the heterozygous c.T97C alteration in *DLGAP3* (Figure 5.18). These results supported that none of the variations selected were the causal alteration in family P378.

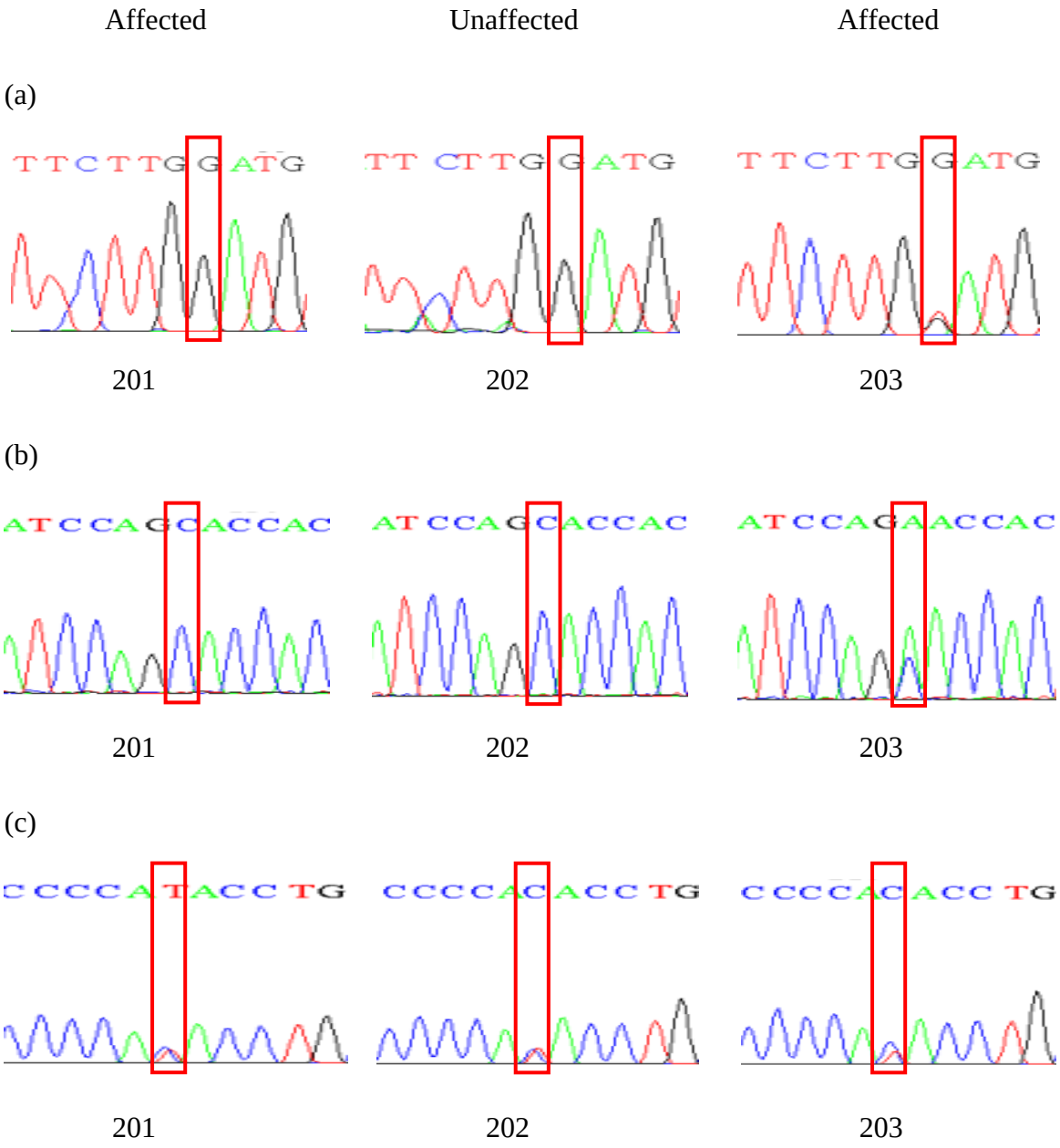


Figure 5.18. Chromatograms showing (a) the c.C5213A variant in *SBF2*, (b) the heterozygous c.G1544T variant in *SLC6A11*, (c) the heterozygous c.T97C variant in *DLGAP3* in the proband, P378 (203) and his affected (201) and unaffected (202) brothers.

5.1.9. P851

For patient P851, median NCV was recorded as 39 m/s and the disease was inherited as an autosomal dominant trait (Figure 5.19).

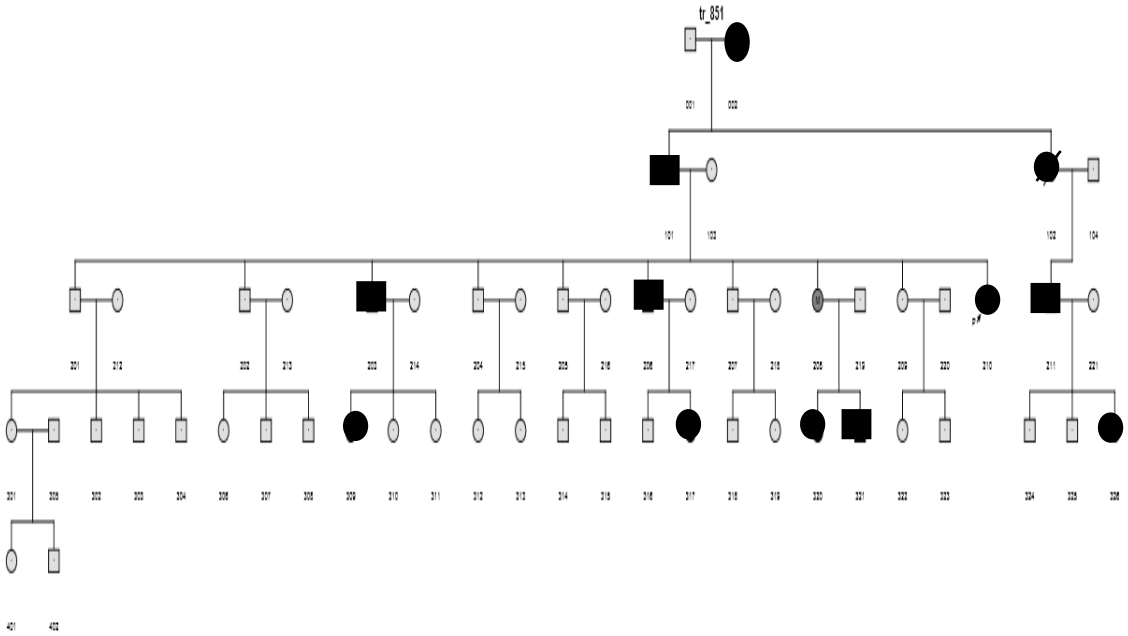


Figure 5.19. Pedigree of the family P851. Individual 208 affected mildly.

The only variant selected for patient P851 was the heterozygous c.G1072A alteration in DNM2 gene. It has been previously identified as a pathogenic SNV (rs267606772) resulting in AD CMT2 (Gallardo *et al*., 2008). As shown in Figure 5.20, it was present in the proband (210) as well as in the affected father (101), in two affected brothers (206, 203) and in the affected daughter of one of these brothers (309). The unaffected mother (103) and unaffected brother (207) did not have the variation, supporting that the c.G1072A alteration in *DNM2* could be the cause of the disease in family P851.

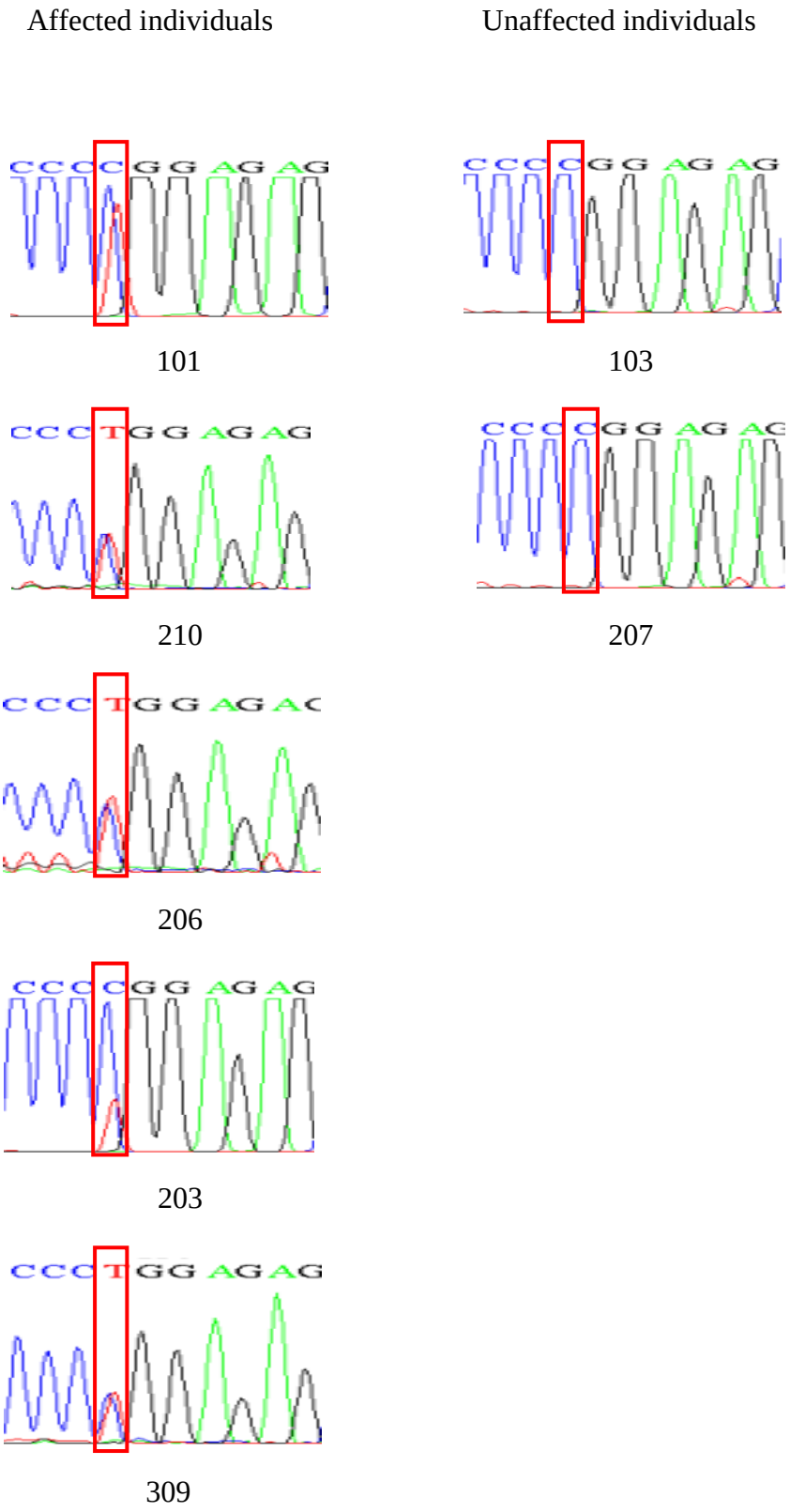


Figure 5.20. Chromatograms showing the *DNM2* variant, c.G1072A, in patient P851 (210) and her family members in the reverse strand.

The filtration of exome sequencing data of the other eight patients was also completed. The possible causal variations selected for them were given in Table 5.4. Accordingly, no possible causative alterations in known CMT genes were observed except the novel, heterozygous *MPZ* variant, c.A403C, detected for patient P738. Unfortunately, none of these variations could be confirmed in the probands and the family members due to the lack of DNA samples or time.

Moreover, variations given in Table 5.5 were found in all seventeen patients supporting that they could be common in our population. They were all located in known CMT genes. Six of them have been identified previously as SNPs, but the *NEFL* alteration, c.1414_1414del, was not present in available databases showing that it is a novel polymorphism that is common in our population or it may have a modifier effect on the CMT phenotype.

Table 5.4. Candidate variants selected for the other eight CMT patients.

	Cytoband	Location	Gene	Ref	Var	Zyg	Annotation	Nt Change	Aa Change	SNP Annotation
P38	Xq13.1	67940149	STARD8	C	T	Hom	Nonsynonymous SNV	c.C1693T	p.R565W	-
	4q24	107249343	AIMP1	C	T	Hom	Stopgain SNV	c.C334T	p.Q112X	-
P157	15q26.1	89876479	POLG	C	A	Het	Nonsynonymous SNV	c.G507T	p.W169C	-
	17q11.2	27434192	MYO18A	A	T	Het	Nonsynonymous SNV	c.T3347A	p.M1116K	-
	1p22.1	94014948	FNBP1L	C	T	Het	Nonsynonymous SNV	c.C1442T	p.A481V	-
P171	7p22.1	4823070	KIAA0415	A	G	Hom	Nonsynonymous SNV	c.A490G	p.S164G	-
	7p22.2	4091417	SDK1	G	C	Hom	Nonsynonymous SNV	c.G2866C	p.D956H	-
P266	9q21.2	80943932	PSAT1	C	T	Hom	Nonsynonymous SNV	c.C1043T	p.A348V	-
	Xp11.23	46844317	PHF16	A	G	Hom	Nonsynonymous SNV	c.A22G	p.S8G	-
	5q35.3	180278385	ZFP62	G	C	Hom	Nonsynonymous SNV	c.C11G	p.S4C	-
P738	1q23.3	161276543	MPZ	T	G	Het	Nonsynonymous SNV	c.A403C	p.I135L	-
P254	Xq28	154508586	CLIC2	C	T	Hom	Nonsynonymous SNV	c.G434A	p.R145H	-
	2q35	218939834	RUFY4	G	A	Hom	Nonsynonymous SNV	c.G619A	p.G207R	-

Table 5.4. Candidate variants selected for the other eight CMT patients (cont.).

P261	22q13.1	39777822	SYNGR1	-	CAA	Hom	Nonframeshift Insertion	c.605_606ins CAA	p.P202delins PN	-
	17p13.2	3594276	P2RX5	TG	-	Hom	Frameshift Deletion	c.333_334del	p.111_112del	-
P343	13q32.3	99532917	DOCK9	T	C	Het	nonsynonymous SNV	c.A2755G	p.K919E	-
	11q24.2	124747426	ROBO3	G	A	Het	nonsynonymous SNV	c.G2858A	p.G953D	-
	5p15.1	16818138	MYO10	A	G	Het	nonsynonymous SNV	c.T259C	p.Y87H	-

Table 5.5. Variations found in all seventeen CMT patients.

Cytoband	Location	CMT Gene	Ref	Var	Zyg	Annotation	Nt Change	Aa Change	Snp Annotation
8p21.2	24811064	NEFL	AG	-	Hom	Frameshift Deletion	c.1414_1414 del	p.472_472del	-
9q33.3	130242166	LRSAM1	A	G	Hom	nonsynonymous SNV	c.A952G	p.N318D	rs1539567
10p14	12111090	DHTKD1	T	C	Hom	nonsynonymous SNV	c.T58C	p.F20L	rs1279138
10p14	12143105	DHTKD1	C	G	Hom	nonsynonymous SNV	c.C1821G	p.I607M	rs2062988
19q13.2	40900865	PRX	C	T	Het	nonsynonymous SNV	c.G3394A	p.G1132R	rs268674
10q22.1	71060610	HK1	A	G	Het	nonsynonymous SNV	c.A20G	p.H7R	rs906220
14q31.3	88401213	GALC	T	C	Hom	nonsynonymous SNV	c.A1852G	p.T618A	rs421262

5.2. Mutation Screening

In this study, mutation screening was performed in eight CMT1 patients that were previously reported to be negative for the 1.4 Mb CMT1A duplication and HNPP deletion. The clinical characteristics of the patients are given in Table 5.5. The genes that were screened for mutations were determined according to the clinical features of the patients.

5.2.1. Mutation Analysis of *MPZ*

Mutation analysis of *MPZ* in two unrelated patients, P686 and P963, was carried out by using PCR-SSCP method. As demonstrated in Figure 5.21, there was no single strand conformation change in any of the six exons of this gene for neither of the patients.

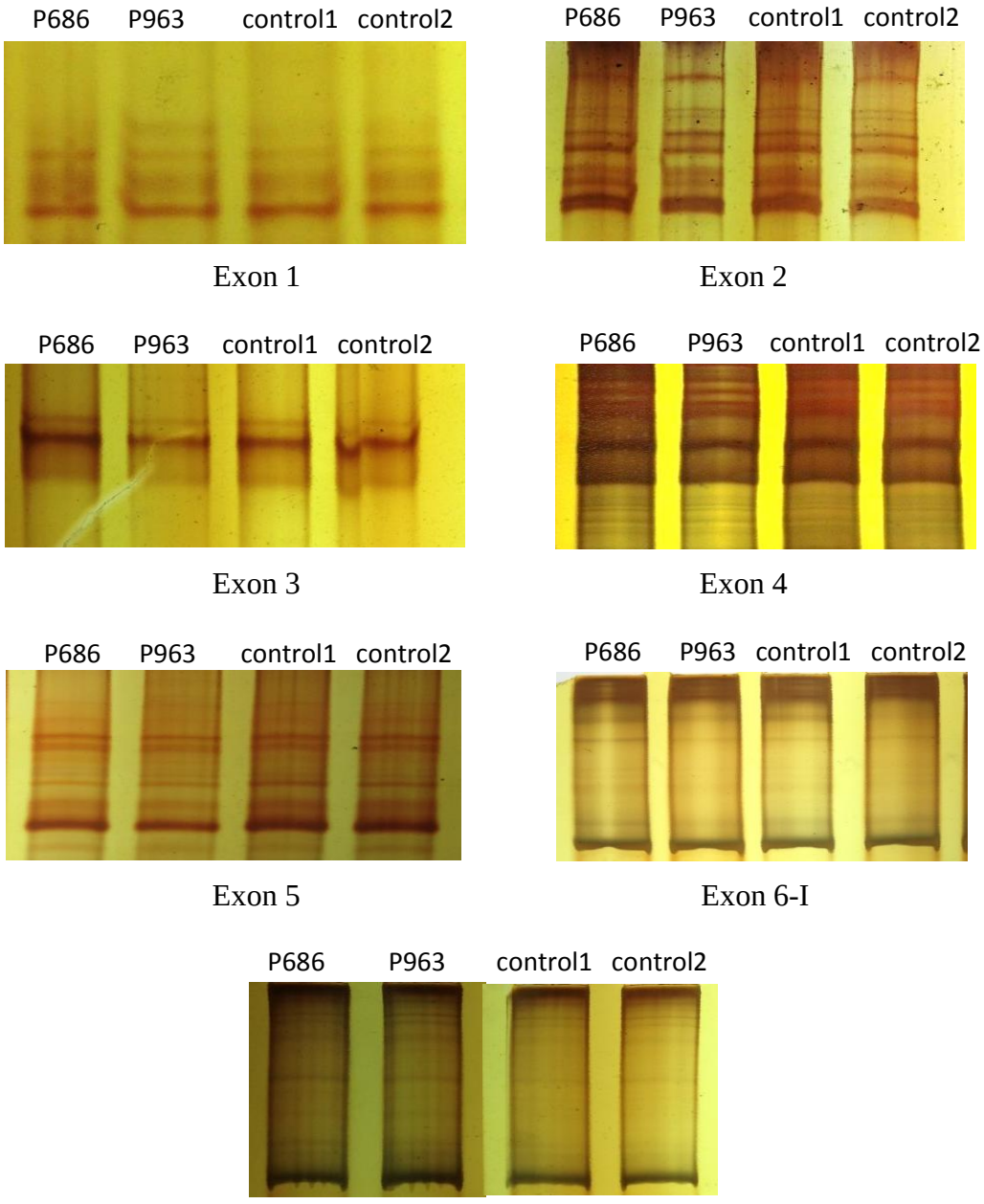


Figure 5.21. SSCP gels (8%) for all six exons of *MPZ*. Sixth exon was amplified as two amplicons due to its length.

Table 5.5. Clinical characteristics of eight CMT patients screened for mutations in targeted genes.

	DOB	Age of onset	Severity	Median MNCV	Type of the disease	Mode of inheritance	Additional features	Consanguinity	Targeted gene
P686	-	Birth/first months	Abnormal mobility- no aids required Upper limbs: distal weakness (hand wrist)	3.9 m/s	CMT1	AR	Hyperlaxity	-	MPZ
P868	2005	2-10 y	Abnormal mobility- no aids required Upper limbs: normal	-	CMT1	AD	-	-	PMP22
P963	1998	2-10 y	Abnormal mobility- no aids required	12 m/s	CMT1	AR	-	Consanguineous	MPZ
P983.3	2000	2-10 y	abnormal mobility- walking aids used upper limbs: distal weakness (hand wrist)	11.2 m/s	CMT1	AR	-	Consanguineous	NDRG1
P983.4	2009	2-10 y	-	-	CMT1	AR	-	Consanguineous	NDRG1
P1070	1926	>40 y	normal mobility upper limbs: normal	35.1 m/s	CMT1	AR	-	-	GJB1, GDAP1
P1075.3	1986	21-40 y	abnormal mobility- walking aids used upper limbs: distal weakness (hand wrist)	20 m/s	CMT1	AR	Scoliosis Bulbar dysfunction	Consanguineous	GDAP1
P1075.4	1986	21-40 y	-	-	CMT1	AR	-	Consanguineous	GDAP1

5.2.2. Mutation Analysis of *GDAP1*

Mutation analysis of *GDAP1* for patients P1075.3 and P1075.4 was performed by PCR-SSCP method. Patients P1075.3 and P1075.4 were dizygotic twins and both were affected. As shown in Figure 5.22, there were additional bands in exon 3 for both patients compared to the controls. In the case of other exons, no single strand conformation change was observed.

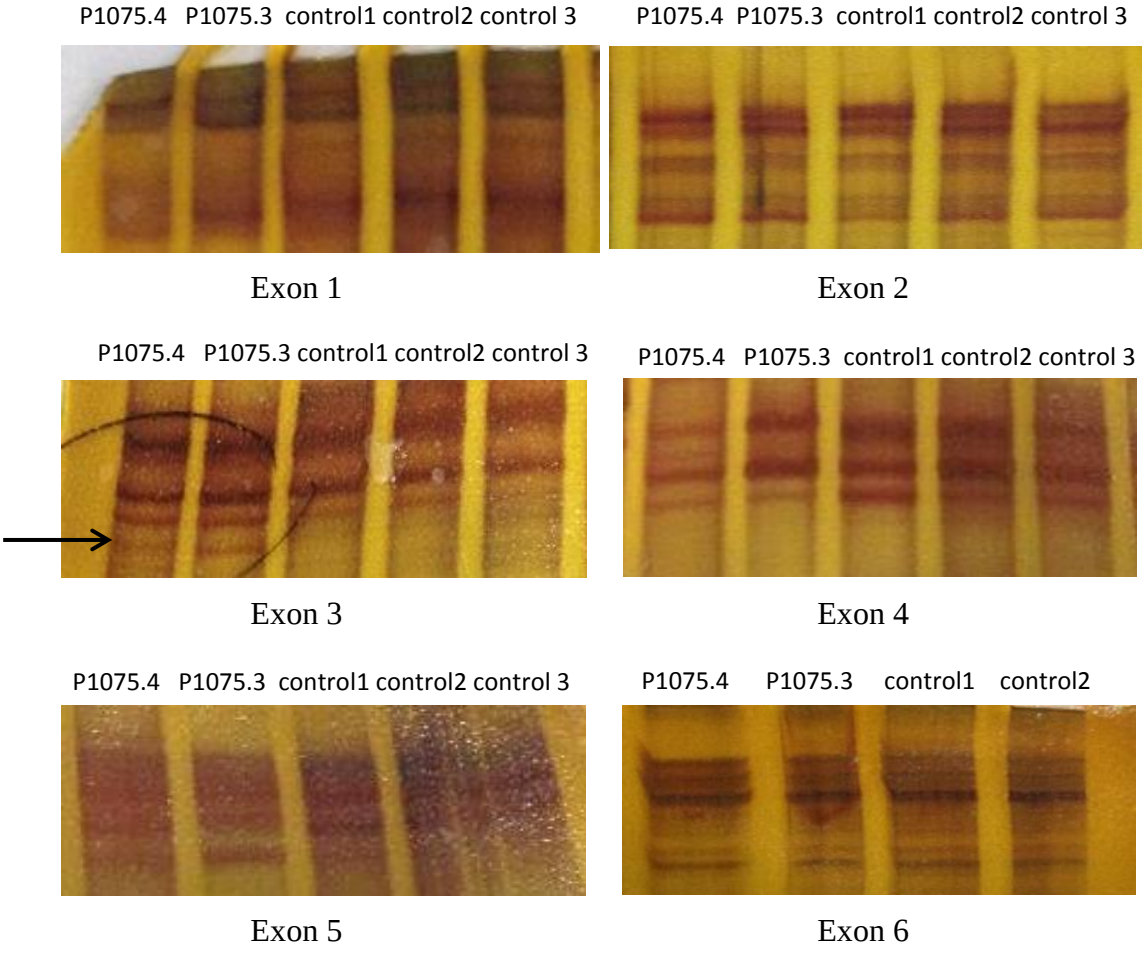


Figure 5.22. SSCP gels (8%) for all six exons of *GDAP1*. The arrow shows extra bands observed for the patients.

Exon 3 was analyzed by Sanger sequencing in order to understand whether an alteration in the nucleotide sequence was indeed present or not. However, DNA sequencing of exon 3 in both patients did not reveal any variation.

For patient P1070, coding regions of *GDAP1* was screened by using PCR-Sanger sequencing method. As shown in Figure 5.23, an upstream gene variant which has been previously identified as a SNP (rs4541908) was detected in exon 1. The patient was homozygous for this c.-161G>T alteration with no known pathogenic effect. In exon 4, the c.T507G variant that has been reported as a synonymous SNP (rs11554166) was observed in the heterozygous state. The third one was an intron variant, c.694+24C>T alteration in exon 5, for which the patient was homozygous. It has been reported as a SNP (rs4463412), but its clinical effect was unknown.

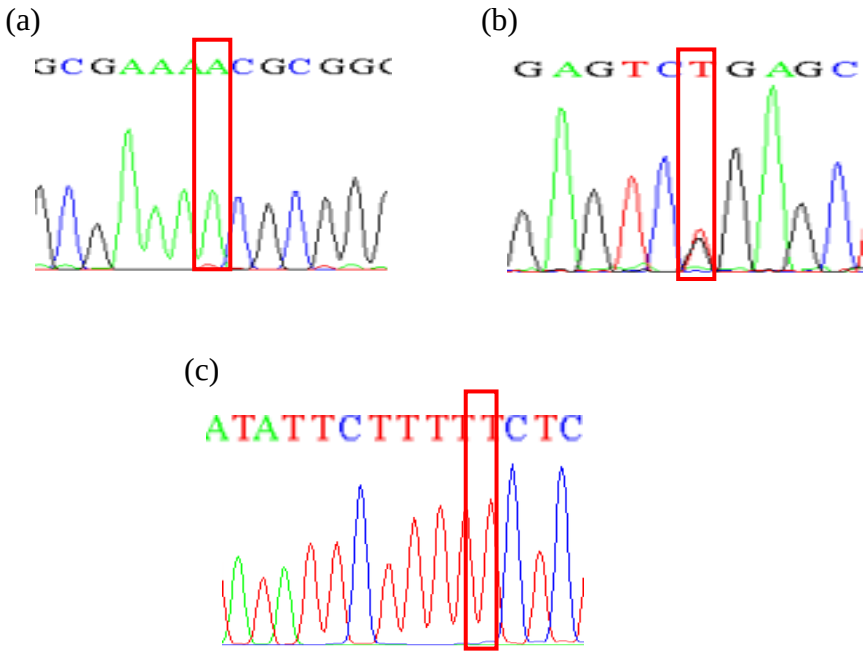


Figure 5.23. Chromatograms showing that (a) the c.-161G>T variant (b) the c.T507G variant and (c) the c.694+24C>T variant in *GDAP1* in patient P1070.

5.2.3. Mutation Analysis of *GJB1*

PCR-Sanger sequencing analysis was carried out in order to screen mutations in *GJB1* for patient P1070. As a result, no alteration in the coding region of the gene was observed.

5.2.4. Mutation Analysis of *PMP22*

PMP22 was screened for mutations by PCR-Sanger sequencing method in patient P868. The heterozygous c.178+2T>C variation (Figure 5.24) which has not been previously reported was detected in intron 3. It was predicted by Human Splicing Finder Tool (Desmet *et al*., 2009) to affect splicing by abolishing the binding site of SRp55 protein. On the other hand, the other three exons did not contain any variation.

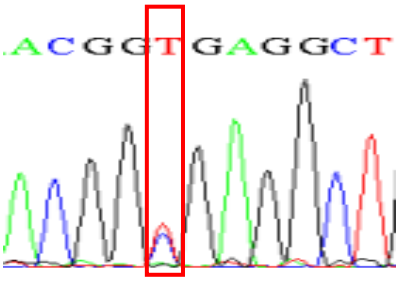


Figure 5.24. Chromatogram showing the c.178+2T>C splicing variant in *PMP22* in patient P868 in heterozygous state.

5.2.5. Mutation Analysis of *NDRG1*

For patients P983.3 and P983.4 who were two affected brothers, the presence of the founder R148X mutation in exon 7 of *NDRG1* was tested by restriction enzyme digestion. In principle, when there is no mutation, PCR product (176 bp) is cut by *TaqI* restriction enzyme resulting in two fragments (100 bp, 76 bp). However, if there is a substitution from C to T at base 564, which is in the restriction site of *TaqI*, digestion cannot occur. As a consequence, on agarose gel, one distinct band is observed for mutated sequences, while wild type fragments generated two bands. As demonstrated in Figure 5.25, a distinct band was observed for each patient. On the other hand, there were two bands for the controls indicating that the mutation was present in patients P983.3 and P983.4.

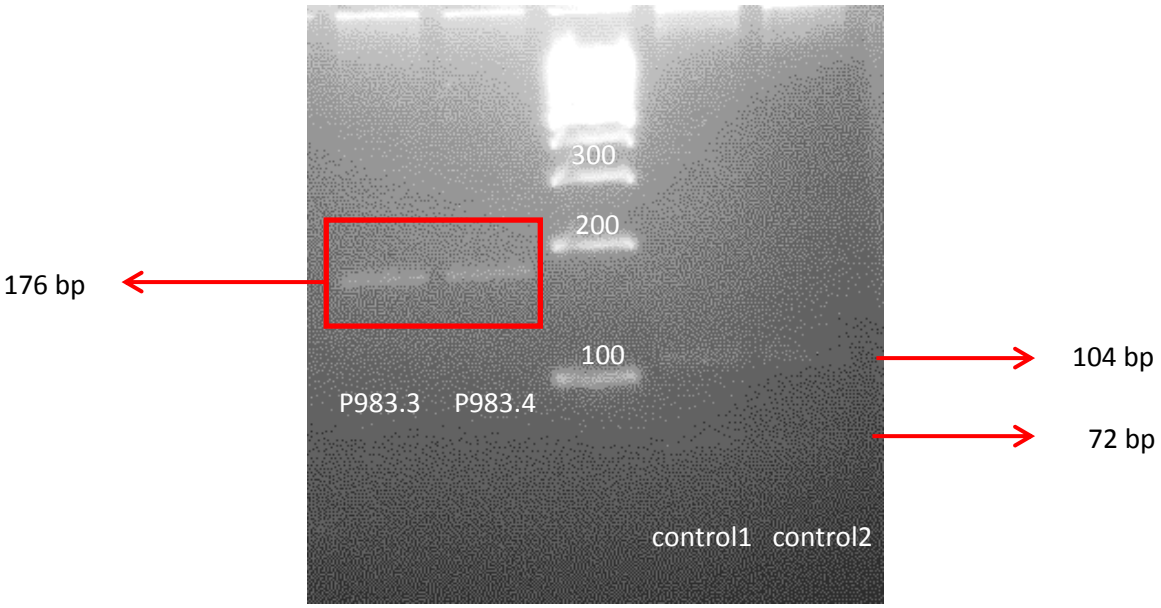


Figure 5.25. *TaqI* restriction enzyme digestion results for patients P983.3 and P983.4.

6. DISCUSSION

In this study, seventeen Turkish CMT patients were exome sequenced and the data obtained from sequencing was evaluated using several filtering approaches. Selected possible disease causing variations in nine of these families were analyzed by using PCR-Sanger sequencing technique in the probands and available family members. At least one alteration was confirmed to segregate with the disease in each family, except for family P378. For the other eight patients, possible candidate variants were also selected using the exome sequencing data. However, variant validation could not be performed because of lack of DNA samples and time. The data revealed seven variations in CMT genes observed in all seventeen patients. Six of these variations have been previously reported as SNPs and one was novel, probably representing a common variation in our population. In the second part of the study, eight other Turkish CMT patients were screened for mutations for one or more of the following genes according to the segregation pattern and clinical symptoms: GDAP1, GJB1, MPZ, NDRG1 and PMP22. A novel variation in *PMP22* was identified for one patient. In addition, three previously reported SNPs were detected in *GDAP1* for another patient. The founder p.R148X mutation in *NDRG1* was also confirmed in two affected brothers.

6.1. Exome Sequencing

All seventeen patients were previously found to be negative for *PMP22* duplication and deletion. The majority of them were also screened for point mutations in the coding regions of common CMT genes such as Cx32, MPZ, PMP22 and GDAP1. However, no alterations in these genes were detected. Because approximately 85% of mutations are found within coding regions and it is more cost effective than whole genome sequencing (Choi *et al.*, 2012; Ku *et al.*, 2012), exome sequencing was carried out for these patients. Consequently, at least one possible causative variation was confirmed for each family except for family P378. The majority of validated alterations were found in the known CMT genes. Although some of them have been reported previously, novel variations were also detected in CMT genes. In addition, several variations in novel possible causative genes were confirmed to segregate with the disease in families.

Exome sequencing data for all seventeen patients was filtered initially by eliminating synonymous variants. Secondly, the SNPs which have been annotated previously were excluded. Then, the variants were chosen according to the heterozygous/homozygous state. Read depth (>30) was also considered. After zygosity selection, functional effects of variations that have been predicted by SIFT, LRT and PolyPhen2 algorithms were considered. Damaging and deleterious ones were selected. They were also checked if they have been reported as associated with CMT in OMIM database. If any candidate could not be chosen after these filtering steps, then all known CMT genes were checked, in the case that databases that were used for variant calling in exome sequencing have missed some of them, especially the rare or new ones. Indeed, the alterations found in *DNM2* and *LRSAM1* which are observed rarely in CMT were detected after the analysis of all CMT genes, individually. On the other hand, if there were no variations in known CMT genes, the possible novel genes were selected due to their functions. The reduction in variation number obtained after each filtering step for each patient are shown in Table 5.2.

Evaluation of exome sequencing data of patient P114 revealed the heterozygous *AARS* variant, c.G986A, p.R329H, as a candidate. The variation was present in all affected family members and absent in the unaffected ones confirming the segregation of the defective allele with the disease in the family. Mutations in AARS gene are known to be associated with autosomal dominant CMT2 that is in accordance with the phenotype of patient P114. The same mutation has been identified previously in two French families (Latour *et al*., 2010), thus the variation is the causal alteration in this family.

Arginine 329 is located in the middle of the catalytic domain of AARS protein which is anα helix structure. In α helices, arginine is significantly more flexible than histidine because of its side chain. In addition, while arginine is a strong proton acceptor, histidine has a protonated imidazole ring. Interestingly, it was reported that the analogous alteration in *E. coli*, p.R314A, caused a reduction in aminoacylation activity of the protein (Ribas de Pouplana *et al*., 1998). Similarly, it was reported that the N71Y mutation in humans located also in the middle domain (Figure 1.12) resulted in a loss of function of AARS enzyme (McLaughlin *et al*., 2012). Consequently, it is supposed that the p.R329H alteration affects the function of AARS, but the mechanism of how it causes CMT is not known, yet.

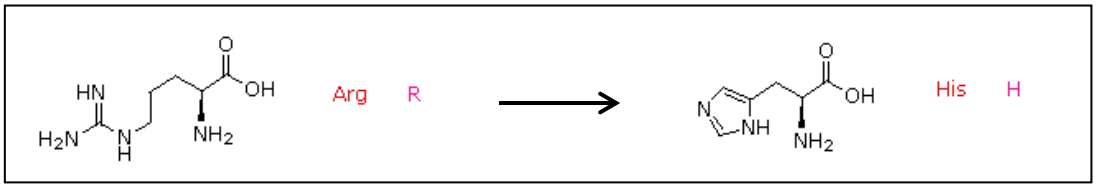


Figure 6.1. Functional groups (R) of arginine to histidine.

One variation in *GJB1* and four other variations in possible candidate genes, LDB3, OTOF, CNNM4 and ARHGAP4, were filtered from the exome sequencing data for patient P138. The alteration in *GJB1* were absent in the first exome sequencing data, so filtering was based on the novel genes. However, it was detected when the raw data was re-analyzed indicating the importance of the evaluation of the raw data. The only DNA sample of relatives belonged to a female individual who was supposed to be the daughter of the patient reported with abnormal EMG records.

The presence of the c.-18A>G variation located at the 3' end of intron 1 in *GJB1* was confirmed in the patient and his probable affected daughter. So far, only five mutations in this region have been reported to cause X-linked CMT which are c.-529T>G, c.-529T>C, c.-527G>C, c.-459C>T and c.-373G>A (Tsai *et al*., 2013). The c.-18A>G variation, on the other hand, has not been reported previously. It was predicted by Human Splicing Finder Tool (Desmet el al., 2009) to affect splicing by abolishing the binding of SRp40 protein. SRp40 activates exonic splicing enhancer (ESE), so it is necessary for efficient splicing. As shown in Figure 6.2, when splicing does not occur properly, intron 1 cannot be cut off from where it is located (in 5'UTR region). As a result, it is predicted by Sixframe tool that a protein containing only 33 amino acids is produced instead of the Connexin 32 protein which consists of 283 amino acids. So, the c.-18A>G alteration can be considered as a novel pathogenic variation resulting in CMT but its effect on splicing can be further tested by RNA analysis in the peripheral blood sample of the patient.

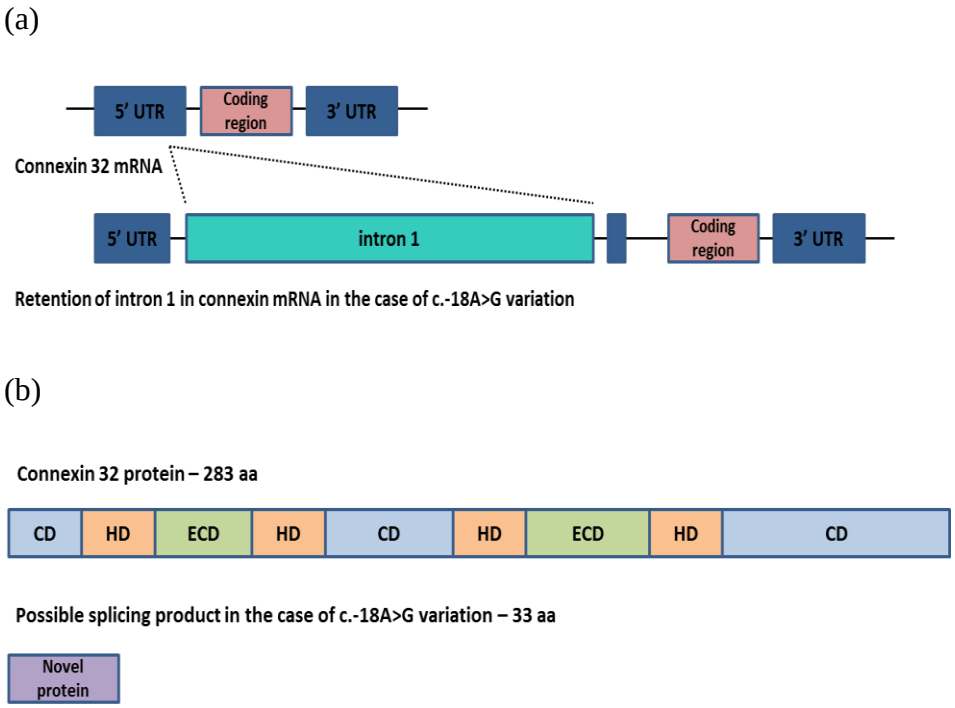


Figure 6.2. The effect of c.-18A>G splicing variation in *GJB1*; (a) retention of intron 1 into mRNA sequence as a result of abolished splicing, (b) the novel protein produced by the mRNA sequence containing intron 1. (CD: cytoplasmic domain, HD: helical domain, ECD: extracellular domain)

Moreover, both the novel heterozygous c.G1531A, p.E511K alteration in *CNNM4* and the novel hemizygous c.G227T, p.S76I variant in *ARHGAP4* were present in the patient and also in the probable affected daughter.

Before referring to the raw data for exome sequencing for this patient, filtering revealed Cyclin M4 (CNNM4) as a candidate because the function of the gene. It is a magnesium transporter. The interaction with the metal ion chaperone COX11 suggests that it may play a role in sensory neuron functions. As shown in Figure 6.3, it contains four domains, DUF21, cystathionine β-synthase (CBS), cyclin box and the cyclic nucleotide monophospate (cNMP) binding domains. The p.E511K alteration found in patient P138 is located in the second CBS domain. CBS domains are supposed to be important as energy-sensing modules that bind adenosin ligands. In addition, they are involved in gating of osmoregulatory proteins, in transport and binding of Mg^{2+}, in mosulation of intracellular

trafficking of chloride channels, in nitrate transport and in inhibition of pyrophosphatase activity (Parry *et al*., 2009 and Garcia *et al*., 2011)

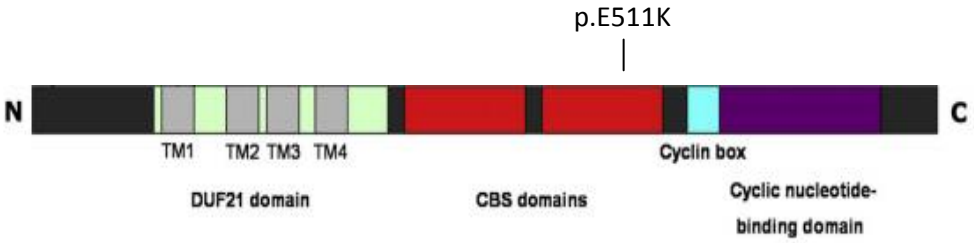


Figure 6.3. Protein structure of CNNM4 and the location of p.E511K variation (Parry *et al*., 2009

On the other hand, in the case of p.E511K, glutamic acid that has a negatively charged side chain changes to lysine which contains a positively charged functional group. So, it can be considered as a transition altering the polarity of the protein and it can affect protein stability.

Rho GTPase-activating protein 4 (ARHGAP4) encodes a member of RhoGAP family. It is responsible for regulation of small GTP-binding proteins. As shown in Figure 6.4, it contains a FCH domain in the N terminus, a GTPase activating (GAP) domain in the centre and an SH3 domain in the C terminus. The p.S76I alteration found in patient P138 is located in the FCH domain. Although its structure and function is not well understood, there are lots of proteins that contain this domain and it is supposed to be involved in the regulation of cytoskeletal rearrangements, vesicular transport, endocytosis and also in MTs binding. Additionally, the SH3 domain located in the C terminus has a role in actin binding. So, it is suggested that the protein is important for regulation of cell migration through actin and MT binding and also for axon motility (Vogt *et al*., 2007).

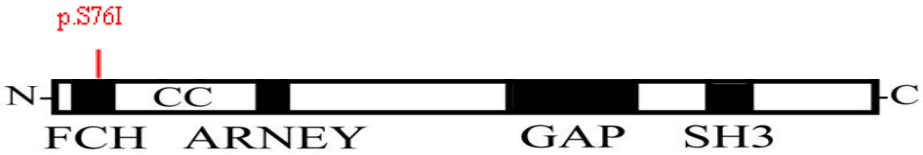


Figure 6.4. Protein structure of ARHGAP4 and the location of p.S76I variation (Vogt *et al.*, 2007).

The amino acid alteration for p.S76I is demonstrated in Figure 6.5. Accordingly, there is a transition from serine that has a polar, uncharged side chain to isoleucine which is an amino acid with hydrophobic side chain. It is suggested that this alteration possibly leads to a change in the structure of the protein.

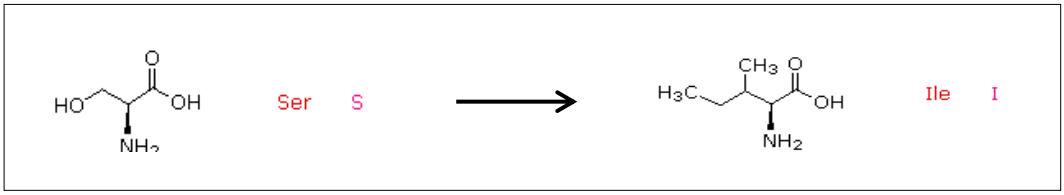


Figure 6.5. Functional groups (R) of serine to isoleucine.

The variation found in *LDB3* for patient P138 was reported as false positive since it was absent in the patient. The alteration in *OTOF* was also excluded, because it was not segregated in the family. All in all, the strongest candidate for patient P138 is the novel splicing variation found in GJB1 gene. On the other hand, given the functions of CNNM4 and ARHGAP4 gene, the variations can be considered as possible disease causing candidates. Further analyses are required to unravel the cause of the the disease in the family.

Analysis of exome sequencing data for patient P163 revealed two variations. Since there was no DNA sample from relatives, they were validated only in the proband by Sanger sequencing. The heterozygous c.G1021A, p.E341K alteration in *DNM2* (CMT2M)

and the heterozygous c.G334A, p.E112K alteration in *LRSAM1* (CMT2P) were confirmed in the patient. None of these variations have been previously reported.

The majority of mutations in *DNM2* causing CMT2M are clustered in the pleckstrin homology (PH) domain which has a role in the interaction with membrane phosphoinositides and regulation of GTPase activity of the protein. The p.E341K variation detected for patient P163 is located in the middle domain which is critical for oligomerization. There is only one reported mutation, G358R, found in this domain (Figure 5.13) (Sidiropoulos *et al*., 2012). p.E341K results in a transition from an amino acid with negatively charged side chain to an amino acid that is positively charged. So, it is suggested that this variation can alter the structure and function of DNM2. As shown in Figure 6.6, the p.E341K alteration occurs in a fully conserved region.

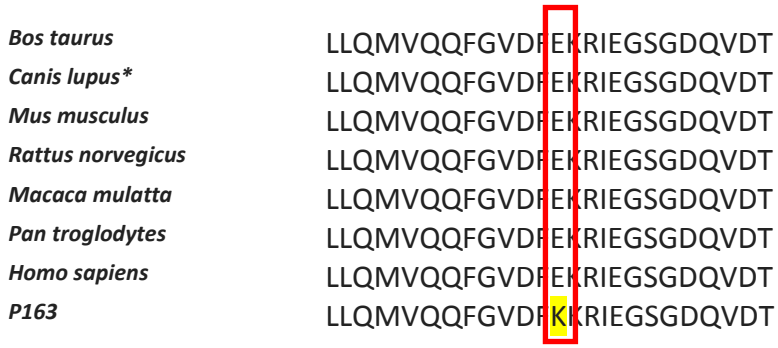


Figure 6.6. Conservation of the amino acid at position 341 among mammalian species, which is shown in red box.

It has been recently identified that mutations in leucine rich repeat and sterile alpha motif containing 1 (LRSAM1) gene causes autosomal dominant axonal neuropathy, CMT2P. It has been also reported to cause recessive CMT2 (AR CMT2) (Guernsey *et al*., 2010). LRSAM1 is an E3 ubiquitin ligase and has a role in membrane fusion. It is implicated in endosomal sorting and trafficking to the lysosome. As shown in Figure 6.7, LRSAM1 protein consists of six leucine-rich repeats (LR), two coil-coiled domains (CC), the sterile-alpha motif (SAM) and a zinc-finger motif of the RING type (ZnF) (Bogdanik *et al*., 2013; Nicolaou *et al*., 2013). The variation found in patient P163, p.E112K, is located in the fourth leucine-rich repeats (LR). The transition from glutamic acid to lysine of

p.E112K, on the other hand, causes a change in the charge which may lead to an alteration in the structure and function of the protein.

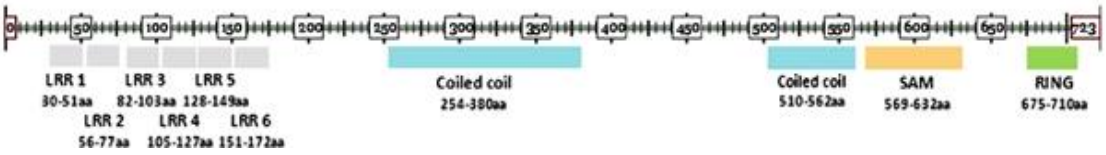


Figure 6.7. Protein structure of LRSAM1 (Nicolaou *et al*., 2013).

In addition, the location where p.E112K alteration occurs is fully conserved among mammalian species as shown in Figure 6.8.

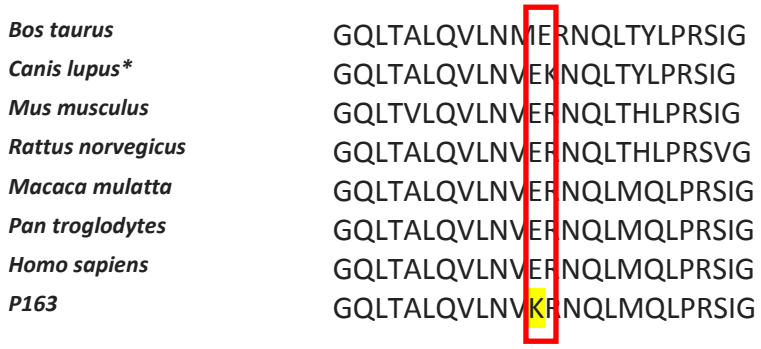


Figure 6.8. Conservation of the amino acid at position 112 among mammalian species, which is shown in red box.

Both *DNM2* and *LRSAM1* cause autosomal dominant CMT2. *DNM2* mutations that result in AD CMT2 have generally a wide range of phenotypic features. It can be mild and slowly progressive or severe with rapid progression (Bitoun *et al*., 2008). On the other hand, *LRSAM1* is a rarely mutated gene and generally it is associated with slowly progressive mild phenotype (Weterman *et al*., 2012). Consequently, because patient P163 had a late onset AD CMT2 with a mild phenotype, the variations found in these genes can be causative for the disease. In order to unravel whether both or one of these are

responsible for the phenotype, the segregation analysis in the family should be performed and if required a population screen should be performed.

The heterozygous c.T479G, p.V160G alteration in the MFN2 gene which has not been reported previously was the only possible candidate selected for patient P209. The alteration was observed in both the proband and his affected son suggesting autosomal dominant segregation.

As demonstrated in Figure 6.9, mutations in *MFN2* resulting in CMT disease are usually found in the GTPase domain which is necessary for mitochondrial tethering and fusion (Eschenbacher *et al*., 2012). The p.V160G alteration found in patient P209 is also located in this domain supporting that this novel variation is causative for the disease in the family. Additionally, the AD CMT2 with a median NCV of 51.6 m/s observed in the patient makes this variation a strong candidate.

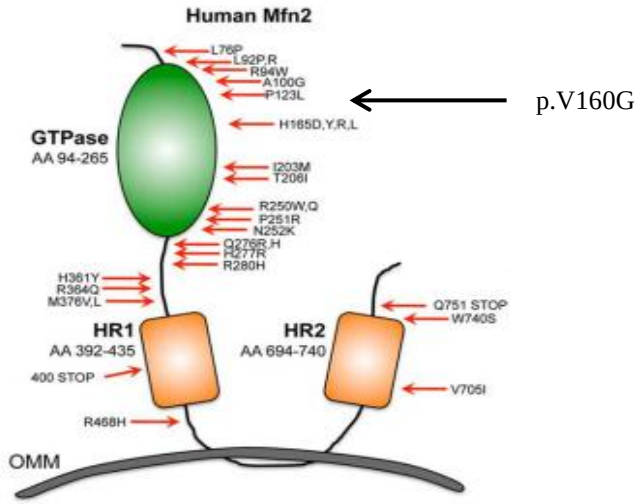


Figure 6.9. Locations of MFN2 mutations linked to CMT (Eschenbacher *et al*., 2012).

Valine is an amino acid that has a hydrophobic group. Glycine, on the other hand, is the amino acid that contains no specific side chain (Figure 6.10). So, the transition from valine to glycine of p.V160G can cause some structural changes in the protein.

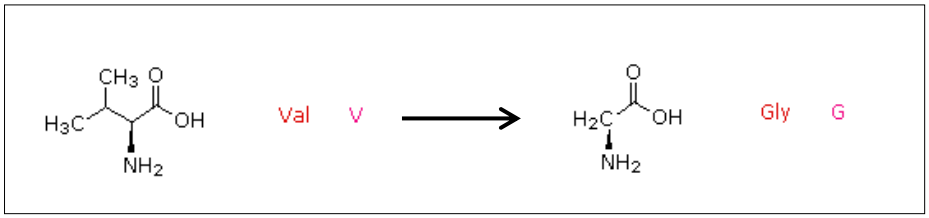


Figure 6.10. Functional groups (R) of valine to glysine.

Consequently, it is suggested that the *MFN2* variation found in patient P209 is the possible cause of the disease in the family.

For patient P212, one reported SNP in *AARS* and four novel variations in novel candidate genes, GMFG, FLNA, SIGMAR1 and CXorf23, were selected. The SNP in *AARS* and the variation in *GMFG* were excluded because they were absent in the patient indicating that they were false positive results. In four of the novel genes, two were confirmed to segregate in the family, the homozygous c.G212T, p.G71V variation in *SIGMAR1* and the hemizygous c.C625T, p.R209C variant in *CXorf23*. They were present in the patient, but absent in the unaffected brother.

Sigma non-opioid intracellular receptor (SIGMAR1) is an integral membrane protein expressed predominantly at ER membranes. It has a role in regulation of a variety of cellular functions in neuronal, glial and peripheral cells. SIGMAR1 protein contains three hydrophobic domains and it is predicted that the first two are transmembrane spanning helices. C terminus, on the other hand, has a molecular chaperon activity. (Tsung-Ping *et al*., 2010 and Teruo *et al*., 2011). Importantly, it is identified that SIGMAR1 proteins are generally found at specialized ER domains associated with mitochondrial outer membranes and may play essential roles in mitochondrial functioning. It has been determined that the chaperone IP3 receptors are localized to MAMs in order to regulate mitochondrial bioenergetics (Teruo *et al*., 2011).

The p.G71V variation is located between the first two hydrophobic domains (extracellular loop). The amino acid transition for the variation is shown in Figure 6.11. Accordingly, there is a change from glycine to valine that contains hydrophobic side chain.

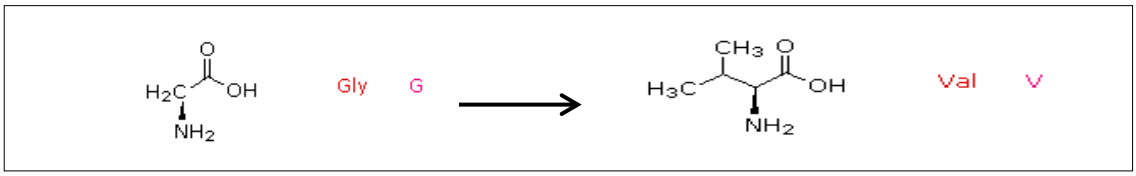


Figure 6.11. Functional groups (R) of glysine to valine.

Chromosome X open reading frame 23 (Cxorf23), for which another variation was identified in the patient, has not been characterized, yet. The transition was from arginine that is an amino acid with positively charged side chain to cysteine (Figure 6.12). Further investigations are required about the function of protein to understand whether it may be causal for CMT or not.

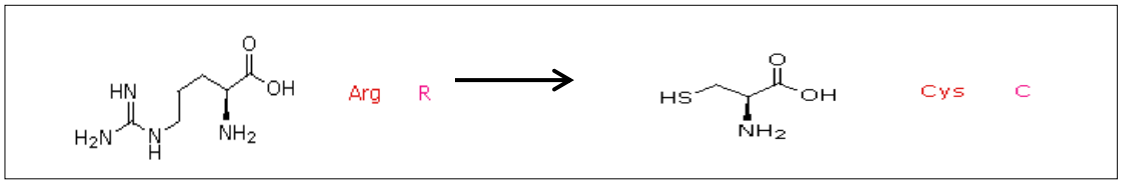


Figure 6.12. Functional groups (R) of arginine to cysteine

GMFG, FLNA, SIGMAR1 and CXorf23 genes were selected as possible candidates for patient P212 in a total number of 31640 variants. By excluding synonymous ones this number was decreased to 16466. Then, it was focused on variants that were predicted as damaging by SIFT, PolyPhen2 and LRT algorithms resulting in 84 alterations in novel genes. Consequently, the variations found in GMFG, FLNA, SIGMAR1 and CXorf23 genes were selected as possible candidates. Given the function of SIGMAR1 gene, it can be considered as a feasible candidate gene for CMT. However, it should be tested in other family members as well as other CMT patients and controls.

Evaluation of exome sequencing data for patient P245 revealed two variations in known CMT genes and two alterations in two possible candidate genes. Unfortunately, there was no DNA sample from relatives, so the selected variations were validated only in

the proband. The variation in ZNF880 gene was a false positive result. On the other hand, the heterozygous c.C1747T, p.R583X alteration in *FGD4* (CMT4H) and the heterozygous c.T2107C, p.W703R variant in *KIF1B* (CMT2A1) were confirmed by conventional sequencing in the patient.

The p.R583X variant is a stop codon change which is located at the FYVE domain of the FGD4 protein (Figure 1.10). FYVE domain has a role in the interaction with phosphoinositides and there are many mutations found within this region causing CMT. So, considering the location of the variation and the stop codon change, p.R582X can be considered as the causal alteration. However, the patient was heterozygous for the mutation in the FGD4 gene that is known to be responsible for the recessive form of the disease.

Kinesin motor protein 1B (KIF1B) belongs to the kinesin superfamily (KIF) of motor proteins and has a role in axonal transport. KIFs contain similar motor domains but have a distinct tail domain allowing association with multiple classes of cargo molecules. *KIF1B* has been reported firstly in a Japanese family to cause CMT2B. In this case, the symptoms began in the proband at early ages. A mild phenotype with moderate muscle weakness and atrophy were observed. However, the patient developed some difficulties in mobility. The causative p.Q98L mutation found in heterozygous state in the motor domain resulted in a reduction in ATPase activity and the motility. So, it was suggested that haploinsufficiency of this protein was responsible for the disease (Zhao *et al*., 2001).

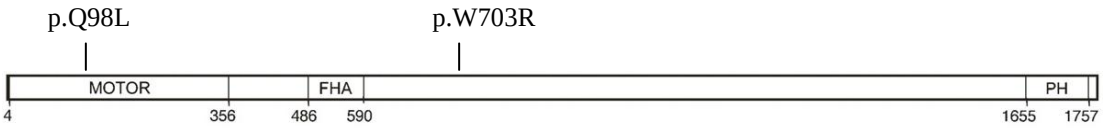


Figure 6.13. Protein structure of KIF1B (Schlisio *et al*., 2008).

Similar to the Japanese patient, patient P245 developed first symptoms at early ages and had a mild phenotype. The p.W703R alteration (Figure 6.13) found in patient P245 leads to a change from tryptophan which is an amino acid with hydrophobic side chain to

arginine amino acid that is positively charged (Figure 6.14). It is located near C terminal region which is less conserved than the motor domain.

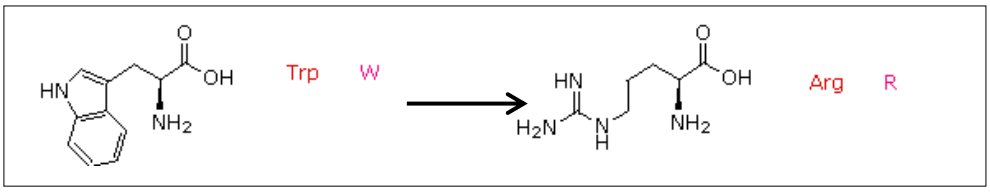


Figure 6.14. Functional groups (R) of tryptophan to arginine.

The p.W703R alteration also occurs in a fully conserved region among mammalian species as given in Figure 6.15.

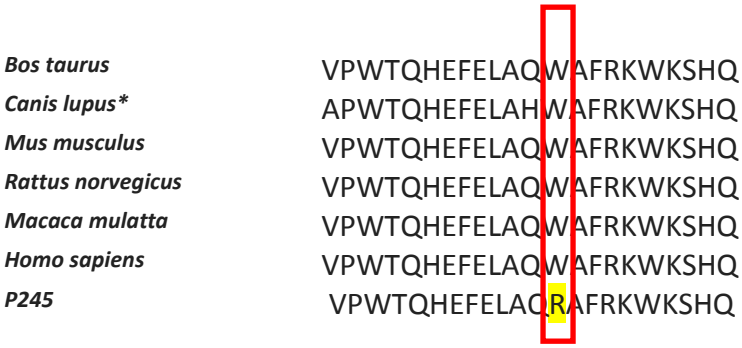


Figure 6.15. Conservation of the amino acid at position 703 among mammalian species, which is shown in red box.

Besides, because of consanguinity among the parents an AR inheritance was suspected in the family. However, if it is the causative gene, the finding will be important confirming the involvement of *KIF1B* in CMT genetics.

On the other hand, the homozygous c.A2822G, p.N941S alteration in *THBS4* was also validated for patient P245. Thrombospondin-4 (THBS4) is a member of thrombospondins that are a family of extracellular matrix glycoproteins. It plays a role for cell-cell or cell-matrix interactions. It is expressed and localized to synapses and neuromuscular junctions at the adulthood (Kim *et al*., 2012). THBS4 protein contains an N-terminal domain, a signature domain harboring four epidermal growth factor (EGR)-like

repeats, calcium-binding wire and a C-terminal domain (Carlson *et al*., 2008; Lawler *et al*., 1995). The p.N941S alteration found in patient P245 is located at the lectin-like C-terminal globe. The transition from asparagine to serine that each has a polar, uncharged side chain (Figure 6.16) does not influence the polarity of the protein, but the structural difference between the functional groups can affect the protein's structure.

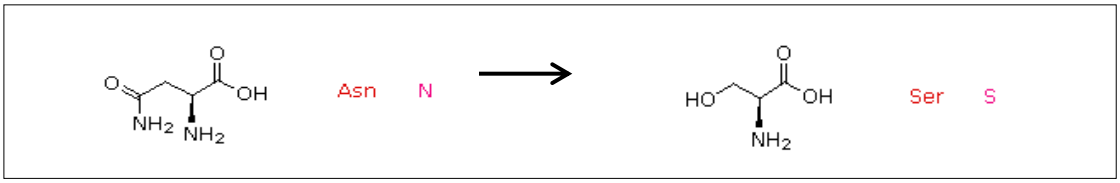


Figure 6.16. Functional groups (R) of asparagine to serine.

The homozygous c.G526A, p.G176S alteration in the MFN2 gene was selected from the exome sequencing data as the possible disease causing candidate for patient P313. It was confirmed to segregate in the family: the patient was homozygous for the alteration and the unaffected father, mother and aunt were heterozygous. As demonstrated in Figure 6.9, the majority of CMT-associated mutations and the p.G176S variant found in the family are located in the GTPase domain. Amino acid transition of the variation is shown in Figure 6.17. Accordingly, there is a change from glycine to serine. Serine has a polar, uncharged side chain, while glycine does not contain any specific functional group.

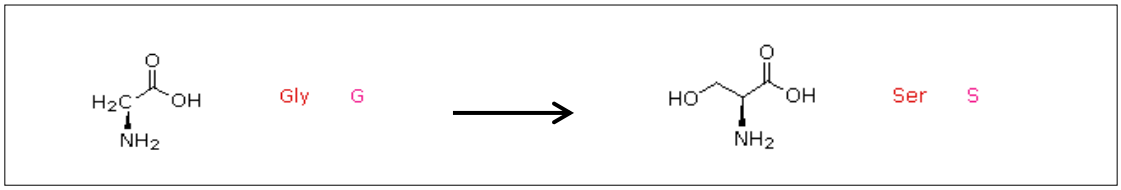


Figure 6.17. Functional groups (R) of glycine to serine.

The alteration is also found in a fully conserved region as shown in Figure 6.18. Importantly, it has not been reported previously indicating that it can be a novel pathogenic mutation that results in CMT.

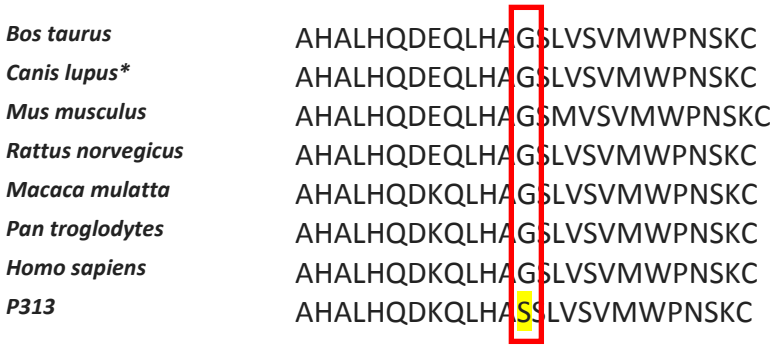


Figure 6.18. Conservation of the amino acid at position 176 among mammalian species, which is shown in red box.

Although the majority of mutations in this gene generally cause autosomal dominant CMT2, there are several cases in which recessive mutations are observed. Both recessive and dominant mutations cause early onset CMT2. Because the recessive mutations in the gene are rare and patient P313 had an early onset CMT2 with almost a mild phenotype, the alteration found in *MFN2* is an important finding.

Analysis of exome sequencing data for patient P378 revealed novel heterozygous c.C5213A variation in SBF2 gene and two novel variations in two novel genes, SLC6A11 and DLGAP3. However, none of them could be confirmed to segregate in the family. Additional filtering approaches are required to choose novel candidates for patient P378.

Evaluation of exome sequencing data of patient P851 revealed only the heterozygous c.G1072A, p.G358R alteration in DNM2 gene. It has been previously identified as a pathogenic SNV (rs267606772) resulting in AD CMT2 with foot deformities and distal muscle weakness and atrophy of the lower limbs. The symptoms developed in the first decade of life (Gallardo *et al*., 2008). Patient P851 presented foot deformity and distal weakness in upper limbs in the first decade of her life supporting that the alteration found in DNM2 gene is possibly causative in this family. The variation was observed in the proband and in the affected father, in two affected brothers and in the affected daughter of one of these brothers. The unaffected mother and unaffected brother were not carrying the variation. These results show that the alteration was segregating perfectly with the disease in the family and is responsible for the disease.

The p.G358R mutation is the only CMT-associated variant located in the middle domain (Figure 5.13). It causes an axonal CMT phenotype. The middle domain is responsible for oligomerization. (Claeys *et al*., 2009). The transition from glycine to arginine changes the polarity. As shown in Figure 6.19, it is located in a fully conserved region among mammalian species.

Figure 6.19. Conservation of the amino acid at position 358 among mammalian species, which is shown in red box.

All in all, considering the patient had AD CMT2 with a mild phenotype, the heterozygous c.G1072A, p.G358R mutation in *DNM2* (CMT2M) gene is suggested to be causative in family P851.

For nine patients, 24 variations in total were analyzed and four of them were recorded as false positives corresponding to approximately %16 of variations analyzed from exome sequencing data. This should be due to both low depth of readings and msreadings during evaluation of the raw data.

Exome sequencing data of the other eight patients was also evaluated. A novel, heterozygous *MPZ* variant, c.A403C, p.I135L was detected for patient P738. Patient P738 was previously found negative for only *PMP22* duplication and deletion. No other CMT genes had been analyzed in the patient. However, because DNA samples of the proband and other family members were absent, the alteration could not be validated. In the remaining patients, there were no possible causative candidates in the known CMT genes.

So, we were focused on the novel genes whose function was possibly related with CMT. However, the selected variants could not be validated due to lack of time.

In addition, seven variations located in different CMT genes were observed in all patients that were exome sequenced (Table 5.5) supporting that they could be common polymorphisms in our population. Six of them were previously reported as SNPs. On the other hand, the homozygous *NEFL* alteration, c.1414_1414del, was not present in dbSNP, Ensemble Variation Database and HGMD. These variations may have an effect on the phenotype and act as modifiers.

All in all, the variations found in seven CMT genes and two possible novel genes were validated (Table 6.1). Exome sequencing revealed two novel variations in MFN2 gene as well as a novel splicing variant in *GJB1*. In addition, novel variations in DNM2, LRSAM, FGD4 and KIF1B genes were detected, but they should be validated in other family members and also in controls. On the other hand, previously identified mutations in AARS and DNM2 genes were found to be causative in two families. Besides, possible novel genes were identified but they are required to be confirmed by studies in other CMT patient cohorts.

Table 6.1. Variations validated in CMT patients.

	Gene	Zyg	Annotation	Nt Change	Aa Change	SNP Annotation
P114	AARS	Het	nonsynonymous SNV	c.G986A	p.R329H	rs267606621
P138	GJB1	Hem	splicing mutation	c.-18A>G	-	Novel
P163	DNM2	Het	nonsynonymous SNV	c.G1021A	p.E341K	Novel
	LRSAM1	Het	nonsynonymous SNV	c.G334A	p.E112K	Novel
P209	MFN2	Het	nonsynonymous SNV	c.T479G	p.V160G	Novel
P212	SIGMAR1	Hom	nonsynonymous SNV	c.G212T	p.G71V	Novel
	CXorf23	Hem	nonsynonymous SNV	c.C625T	p.R209C	Novel
P245	FGD4	Het	stopgain SNV	c.C1747T	p.R583X	Novel
	KIF1B	Het	nonsynonymous SNV	c.T2107C	p.W703R	Novel
P313	MFN2	Hom	nonsynonymous SNV	c.G526A	p.G176S	Novel
P851	DNM2	Het	nonsynonymous SNV	c.G1072A	p.G358R	rs267606772

6.2. Mutation Screening

In this study, eight patients with demyelinating CMT were screened for mutations in candidate genes selected based on their clinical evaluation and EMG data. They were previously found to be negative for CMT1A duplication and HNPP deletion.

In two unrelated affected individuals (patients P686 and P963), *MPZ* was screened using PCR-SSCP method. The disease was inherited in an autosomal recessive trait in these patients. The symptoms were observed at first months in the patient P686. Similarly, there was an early age of onset for the patient P963. They represented difficulties in mobility. The patient P686 had also distal weakness in hand wrist. Median NCVs were recorded as 3.9 m/s for the patient P686 and 12 m/s for the patient P963. Mutations in *MPZ* generally cause autosomal dominant CMT1 (CMT1B) but there are cases in *MPZ* mutations result in AR CMT1, AD CMT2 and intermediate CMT. It has been reported that mutations causing AR CMT1 generate a severe phenotype with early age of onset. NCVs, on the other hand, are very low (Shy *et al*., 2004). Accordingly, MPZ gene was determined as a candidate for these families, however, no variations have been identified in this gene.

GDAP1 was screened by using PCR-SSCP in two sisters who were dizygotic twins (patients P1075.3 and P1075.4). The disease was inherited as an autosomal recessive trait and there was consanguinity in the family. The sisters, both were affected, began to show symptoms later in their lives. They were presenting abnormal mobility and used walking aids. Scoliosis and bulbar dysfunction were the other symptoms observed in the patients. Median NCVs were recorded as 20 m/s for both patients. On the other hand, GDAP1 causing autosomal recessive CMT1 results in a severe phenotype. NCVs are usually between 25-35 m/s. Scoliosis is one of the additional symptoms that may be observed (Barisic *et al*., 2008). Consequently, *GDAP1* was chosen as the strongest candidate for these patients. However, no variations were observed in the patients.

In one female patient, P1070, both GJB1 and GDAP1 genes were analyzed by PCR and Sanger sequencing methods. The disease was supposed to be inherited as autosomal recessive trait in the family. However, there was no male to male transmission, so it could be X-linked, too. The symptoms were observed after 40s and there was a mild phenotype.

The median NCV was recorded as 35.1 m/s. Since mutations in *GJB1* result in mild phenotype in females and the symptoms are observed late with a NCV range of of 30-50 m/s (Tazir *et al*., 2013), *GJB1* was analyzed initially. However, there were no alterations in the gene. Later, considering the NCV (35.1 m/s), it was suggested that there was an intermediate CMT in the patient rather than CMT1. So, *GDAP1* which cause a typical CMT phenotype (Siskind *et al*., 2013) was screened. The patient was found to have three different SNPs in the gene. She was homozygous for c.-161G>T SNP (rs4541908) with no known pathogenic effect, heterozygous for c.T507G variant (synonymous SNP - rs11554166) and homozygous for c.694+24C>T intron variant (rs4463412). The causative gene and mutation could not be identified for the patient.

In another patient, P868, PMP22 gene was analyzed for point mutations by PCR-Sanger sequencing. In this case, there was an autosomal dominant inheritance pattern. The symptoms were observed at early ages. The patient showed difficulties in mobility, but there was no distal weakness in upper limbs. Point mutations in PMP22 gene result in a severe phenotype with early age of onset (Kovach *et al*., 1999). Hence, PMP22 gene was screened and only the heterozygous c.178+2T>C variation was detected in intron 3. It has not been previously reported. Because it was located in an intron, it was analyzed whether it affected splicing or not. By Human Splicing Finder Tool (Desmet *et al*., 2009), it was predicted that the c.178+2T>C alteration affected splicing by abolishing the binding site of SRp55 protein. In addition, it was predicted that the insertion of intron 3 in the mRNA sequence could cause a truncated protein as shown in Figure 6.20.

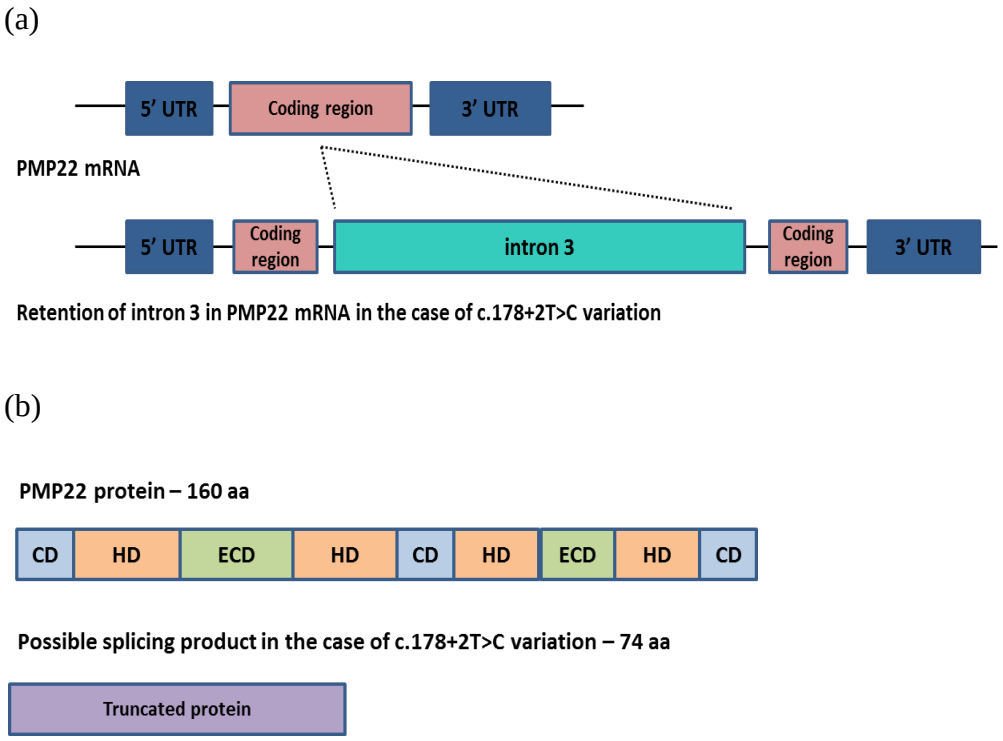


Figure 6.20. The effect of c.178+2T>C splicing variation in *PMP22*; (a) retention of intron 3 into mRNA sequence as a result of abolished splicing, (b) the truncated protein produced by the mRNA sequence containing intron 3. (CD: cytoplasmic domain, HD: helical domain, ECD: extracellular domain)

Consequently, it can be a novel pathogenic splicing mutation in *PMP22* and further analyses are required to validate this possibility.

Lastly, in two Gypsy brothers (patients P983.3 and P983.4), presence of a founder mutation in the NDRG1 gene was tested using PCR and restriction enzyme digestion. The disease was inherited as an autosomal recessive trait in the family and the parents were consanguineous. The onset was at early ages and the median NCV was recorded as 11.2 m/s for patient P983.3. He also used walking aids because of mobility difficulties. Distal weakness with hand wrist was present in upper limbs. Since the known p. R148X founder mutation in *NDRG1* (CMT4D) generally found in Gypsy population, the patients were tested and found to be positive. The founder mutation in *NDRG1* was the causative mutation in the family.

Consequently, these mutation screening results indicate that the determination of causative gene/mutation is very challenging if there is not a specific clinical characteristic or a founder mutation for the patient. It is also time consuming and expensive to screen all known CMT genes. So, this makes exome sequencing a feasible approach to find causative genes and mutations in the case of CMT.

7. CONCLUSION

In this study, we identified some novel and previously reported mutations as causative for the disease by evaluating exome sequencing data and mutation screening in several common CMT genes. Exome sequencing provided to detect three novel variations in CMT genes, MFN2 and GJB1. Further studies should be performed to understand how these novel alterations lead to pathogenicity. In addition, several possible disease causing genes were identified. These genes should be analyzed in other CMT cases in order to identify the exact causative ones. Mutation screening, on the other hand, revealed a novel intron variant in PMP22 gene. Absence of mutations in the known CMT genes in most of the families analyzed in this study provides evidence for further genetic heterogeneity of the CMT disease. Thus, such genetic studies in larger cohorts of patients are required to unravel novel CMT genes. We hope this study can contribute to illumination of genetics of CMT and help development of therapies for patients.

REFERENCES

Azzedine, H., J. Senderek, C. Rivolta, and R. Chrast, 2012, "Molecular Genetics of Charcot-Marie-Tooth Disease: From Genes to Genomes", *Molecular Syndromology*, Vol. 3, No. 5, pp. 204-14.

Bamshad, M. J., S. B. Ng, A. W. Bigham, H. K. Tabor, M. J. Emond, D. A. Nickerson, and J. Shendure, 2011, "Exome Sequencing as a Tool for Mendelian Disease Gene Discovery", *Nature Reviews Genetics*, Vol. 12, No. 11, pp. 745-55.

Banchs, I., C. Casasnovas, A. Alberti, L. De Jorge, M. Povedano, J. Montero, J. A. Martinez-Matos, and V. Volpini, 2009, "Diagnosis of Charcot-Marie-Tooth Disease", *Journal of Biomedicine and Biotechnology*, Vol. 2009, No. 985415.

Barisic, N., K. G. Claeys, M. Sirotkovic-Skerlev, A. Lofgren, E. Nelis, P. De Jonghe, and V. Timmerman, 2008, "Charcot-Marie-Tooth Disease: A Clinico-Genetic Confrontation", *Annals of Human Genetics*, Vol. 72, No. Pt 3, pp. 416-41.

Baudot, C., C. Esteve, C. Castro, Y. Poitelon, C. Mas, T. Hamadouche, M. El-Rajab, N. Levy, A. Megarbane, and V. Delague, 2012, "Two Novel Missense Mutations in Fgd4/Frabin Cause Charcot-Marie-Tooth Type 4h (Cmt4h)", *Journal of the Peripheral Nervous System*, Vol. 17, No. 2, pp. 141-6.

Bitoun, M., T. Stojkovic, B. Prudhon, C. A. Maurage, P. Latour, P. Vermersch, and P. Guicheney, 2008, "A Novel Mutation in the Dynamin 2 Gene in a Charcot-Marie-Tooth Type 2 Patient: Clinical and Pathological Findings", *Neuromuscular Disorders*, Vol. 18, No. 4, pp. 334-8.

Bogdanik, L. P., J. N. Sleigh, C. Tian, M. E. Samuels, K. Bedard, K. L. Seburn, and R. W. Burgess, 2013, "Loss of the E3 Ubiquitin Ligase Lrsam1 Sensitizes Peripheral Axons to Degeneration in a Mouse Model of Charcot-Marie-Tooth Disease", *Disease Models and Mechanisms*, Vol. 6, No. 3, pp. 780-92.

Bohm, J., V. Biancalana, E. T. Dechene, M. Bitoun, C. R. Pierson, E. Schaefer, H. Karasoy, M. A. Dempsey, F. Klein, N. Dondaine, *et al*., 2012, "Mutation Spectrum in the Large Gtpase Dynamin 2, and Genotype-Phenotype Correlation in Autosomal Dominant Centronuclear Myopathy", *Human Mutation*, Vol. 33, No. 6, pp. 949-59.

Bucci, C., O. Bakke, and C. Progida, 2012, "Charcot-Marie-Tooth Disease and Intracellular Traffic", *Progress in Neurobiology*, Vol. 99, No. 3, pp. 191-225.

Carlson, C. B., J. Lawler, and D. F. Mosher, 2008, "Structures of Thrombospondins", *Cellular and Molecular Life Sciences*, Vol. 65, No. 5, pp. 672-86.

Cassereau, J., A. Chevrollier, D. Bonneau, C. Verny, V. Procaccio, P. Reynier, and M. Ferre, 2011, "A Locus-Specific Database for Mutations in Gdap1 Allows Analysis of Genotype-Phenotype Correlations in Charcot-Marie-Tooth Diseases Type 4a and 2k", *Orphanet Journal of Rare Diseases*, Vol. 6, No. 87.

Cassereau, J., A. Chevrollier, N. Gueguen, V. Desquiret, C. Verny, G. Nicolas, F. Dubas, P. Amati-Bonneau, P. Reynier, D. Bonneau, *et al*., 2011, "Mitochondrial Dysfunction and Pathophysiology of Charcot-Marie-Tooth Disease Involving Gdap1 Mutations", *Experimental Neurology*, Vol. 227, No. 1, pp. 31-41.

Charcot, J.M. and P. Marie, 1886, “Sur Une forme Particuliére D`atrophie Musculaire Progressive, Souvent Familiale, Debutant Par Les Pieds Et Les Jambes Et atteignant Plus Tard Les Mains”, *Revue Médicale*, Vol. 6, pp. 97-138.

Choi, B. O., S. K. Koo, M. H. Park, H. Rhee, S. J. Yang, K. G. Choi, S. C. Jung, H. S. Kim, Y. S. Hyun, K. Nakhro, *et al*., 2012, "Exome Sequencing Is an Efficient Tool for Genetic Screening of Charcot-Marie-Tooth Disease", *Human Mutation*, Vol. 33, No. 11, pp. 1610-5.

Claeys, K. G., S. Zuchner, M. Kennerson, J. Berciano, A. Garcia, K. Verhoeven, E. Storey, J. R. Merory, H. M. Bienfait, M. Lammens, *et al*., 2009, "Phenotypic Spectrum of Dynamin 2 Mutations in Charcot-Marie-Tooth Neuropathy", *Brain*, Vol. 132, No. 7, pp. 1741-52.

Cowling, B. S., A. Toussaint, J. Muller, and J. Laporte, 2012, "Defective Membrane Remodeling in Neuromuscular Diseases: Insights from Animal Models", *PLoS Genetics*, Vol. 8, No. 4, pp. e1002595.

Davis, C. J., W. G. Bradley, and R. Madrid, 1978, "The Peroneal Muscular Atrophy Syndrome: Clinical, Genetic, Electrophysiological and Nerve Biopsy Studies. I. Clinical, Genetic and Electrophysiological Findings and Classification", *Journal de Genetique Humaine*, Vol. 26, No. 4, pp. 311-49.

Desmet, F. O., D. Hamroun, M. Lalande, G. Collod-Beroud, M. Claustres, and C. Beroud, 2009, "Human Splicing Finder: An Online Bioinformatics Tool to Predict Splicing Signals", *Nucleic Acids Research*, Vol. 37, No. 9, pp. e67.

Dolled-Filhart, M. P., M. Lee, Jr., C. W. Ou-Yang, R. R. Haraksingh, and J. C. Lin, 2013, "Computational and Bioinformatics Frameworks for Next-Generation Whole Exome and Genome Sequencing", *Scientific World Journal*, Vol. 2013, No. 730210.

Kim, D. S., K. W. Li, A. Boroujerdi, Y. P. Yu, C. Y. Zhou, P. Deng, J. Park, X. Zhang, J. Lee, M. Corpe, *et al*., 2012, "Thrombospondin-4 Contributes to Spinal Sensitization and Neuropathic Pain States", *Journal of Neuroscience*, Vol. 32, No. 26, pp. 8977–8987.

D'Urso, D., P. Ehrhardt, and H. W. Muller, 1999, "Peripheral Myelin Protein 22 and Protein Zero: A Novel Association in Peripheral Nervous System Myelin", *Journal of Neuroscience*, Vol. 19, No. 9, pp. 3396-403.

Eschenbacher, W. H., M. Song, Y. Chen, P. Bhandari, P. Zhao, C. C. Jowdy, J. T. Engelhard, and G. W. Dorn, 2nd, 2012, "Two Rare Human Mitofusin 2 Mutations

Alter Mitochondrial Dynamics and Induce Retinal and Cardiac Pathology in Drosophila", *PLoS One*, Vol. 7, No. 9, pp. e44296.

Fabrizi, G. M., T. Cavallaro, C. Angiari, I. Cabrini, F. Taioli, G. Malerba, L. Bertolasi, and N. Rizzuto, 2007, "Charcot-Marie-Tooth Disease Type 2e, a Disorder of the Cytoskeleton", *Brain*, Vol. 130, No. Pt 2, pp. 394-403.

Gallardo, E., K. G. Claeys, E. Nelis, A. Garcia, A. Canga, O. Combarros, V. Timmerman, P. De Jonghe, and J. Berciano, 2008, "Magnetic Resonance Imaging Findings of Leg Musculature in Charcot-Marie-Tooth Disease Type 2 Due to Dynamin 2 Mutation", *Journal of Neurology*, Vol. 255, No. 7, pp. 986-92.

Guernsey, D. L., H. Jiang, K. Bedard, S. C. Evans, M. Ferguson, M. Matsuoka, C. Macgillivray, M. Nightingale, S. Perry, A. L. Rideout, *et al*., 2010, "Mutation in the Gene Encoding Ubiquitin Ligase Lrsam1 in Patients with Charcot-Marie-Tooth Disease", *PLoS Genetics*, Vol. 6, No. 8.

Jordanova, A., F. P. Thomas, V. Guergueltcheva, I. Tournev, F. A. Gondim, B. Ishpekova, E. De Vriendt, A. Jacobs, I. Litvinenko, N. Ivanova, *et al*., 2003, "Dominant Intermediate Charcot-Marie-Tooth Type C Maps to Chromosome 1p34-P35", *American Journal of Human Genetics*, Vol. 73, No. 6, pp. 1423-30.

Juarez, P., and F. Palau, 2012, "Neural and Molecular Features on Charcot-Marie-Tooth Disease Plasticity and Therapy", *Neural Plasticity*, Vol. 2012, No. 171636.

Kalaydjieva, L., D. Gresham, R. Gooding, L. Heather, F. Baas, R. de Jonge, K. Blechschmidt, D. Angelicheva, D. Chandler, P. Worsley, *et al*., 2000, "N-Myc Downstream-Regulated Gene 1 Is Mutated in Hereditary Motor and Sensory Neuropathy-Lom", *American Journal of Human Genetics*, Vol. 67,No. 1, pp. 47-58.

Kovach, M. J., J. P. Lin, S. Boyadjiev, K. Campbell, L. Mazzeo, K. Herman, L. A. Rimer, W. Frank, B. Llewellyn, E. W. Jabs, *et al*., 1999, "A Unique Point Mutation in the

Pmp22 Gene Is Associated with Charcot-Marie-Tooth Disease and Deafness", *American Journal of Human Genetics*, Vol. 64, No. 6, pp. 1580-93.

Ku, C. S., D. N. Cooper, C. Polychronakos, N. Naidoo, M. Wu, and R. Soong, 2012, "Exome Sequencing: Dual Role as a Discovery and Diagnostic Tool", *Annals of Neurology*, Vol. 71, No. 1, pp. 5-14.

Landoure, G., J. M. Sullivan, J. O. Johnson, C. H. Munns, Y. Shi, O. Diallo, J. R. Gibbs, R. Gaudet, C. L. Ludlow, K. H. Fischbeck, *et al*., 2012, "Exome Sequencing Identifies a Novel Trpv4 Mutation in a Cmt2c Family", *Neurology*, Vol. 79, No. 2, pp. 192-4.

Langworthy, M. M., and B. Appel, 2012, "Schwann Cell Myelination Requires Dynein Function", *Neural Development*, Vol. 7, No. 37.

Latour, P., C. Thauvin-Robinet, C. Baudelet-Mery, P. Soichot, V. Cusin, L. Faivre, M. C. Locatelli, M. Mayencon, A. Sarcey, E. Broussolle, *et al*., 2010, "A Major Determinant for Binding and Aminoacylation of Trna(Ala) in Cytoplasmic Alanyl-Trna Synthetase Is Mutated in Dominant Axonal Charcot-Marie-Tooth Disease", *American Journal of Human Genetics*, Vol. 86, No. 1, pp. 77-82.

Lawler, J., K. McHenry, M. Duquette, and L. Derick, 1995, "Characterization of Human Thrombospondin-4", *Journal of Biological Chemistry*, Vol. 270, No. 6, pp. 2809-14.

Lupski, J. R., R. M. de Oca-Luna, S. Slaugenhaupt, L. Pentao, V. Guzzetta, B. J. Trask, O. Saucedo-Cardenas, D. F. Barker, J. M. Killian, C. A. Garcia, *et al*., 1991, "DNA Duplication Associated with Charcot-Marie-Tooth Disease Type 1a", *Cell*, Vol. 66, No. 2, pp. 219-32.

Lupski, J. R., J. G. Reid, C. Gonzaga-Jauregui, D. Rio Deiros, D. C. Chen, L. Nazareth, M. Bainbridge, H. Dinh, C. Jing, D. A. Wheeler, *et al*., 2010, "Whole-Genome

Sequencing in a Patient with Charcot-Marie-Tooth Neuropathy", *New England Journal of Medicine*, Vol. 362, No. 13, pp. 1181-91.

McLaughlin, H. M., R. Sakaguchi, W. Giblin, Nisc Comparative Sequencing Program, T. E. Wilson, L. Biesecker, J. R. Lupski, K. Talbot, J. M. Vance, S. Zuchner, *et al*., 2012, "A Recurrent Loss-of-Function Alanyl-Trna Synthetase (Aars) Mutation in Patients with Charcot-Marie-Tooth Disease Type 2n (Cmt2n)", *Human Mutation*, Vol. 33, No. 1, pp. 244-53.

Melotte, V., X. Qu, M. Ongenaert, W. van Criekinge, A. P. de Bruine, H. S. Baldwin, and M. van Engeland, 2010, "The N-Myc Downstream Regulated Gene (Ndrg) Family: Diverse Functions, Multiple Applications", *FASEB Journal*, Vol. 24, No. 11, pp. 4153-66.

Miller, L. J., A. S. Saporta, S. L. Sottile, C. E. Siskind, S. M. Feely, and M. E. Shy, 2011, "Strategy for Genetic Testing in Charcot-Marie-Disease", *Acta Myologica*, Vol. 30, No. 2, pp. 109-16.

Naganuma, M., S. Sekine, R. Fukunaga, and S. Yokoyama, 2009, "Unique Protein Architecture of Alanyl-Trna Synthetase for Aminoacylation, Editing, and Dimerization", *Proceeding of the National Academy of Sciences of the United States of America*, Vol. 106, No. 21, pp. 8489-94.

Nicholson, G., and S. Myers, 2006, "Intermediate Forms of Charcot-Marie-Tooth Neuropathy: A Review", *Neuromolecular Medicine*, Vol. 8, No. 1-2, pp. 123-30.

Nicolaou, P., C. Cianchetti, A. Minaidou, G. Marrosu, E. Zamba-Papanicolaou, L. Middleton, and K. Christodoulou, 2013, "A Novel Lrsam1 Mutation Is Associated with Autosomal Dominant Axonal Charcot-Marie-Tooth Disease", *European Journal of Human Genetics*, Vol. 21, No. 2, pp. 190-4.

Niemann, A., P. Berger, and U. Suter, 2006, "Pathomechanisms of Mutant Proteins in Charcot-Marie-Tooth Disease", *Neuromolecular Medicine*, Vol. 8, No. 1-2, pp. 217-42.

Otera, H., N. Ishihara, and K. Mihara, 2013, "New Insights into the Function and Regulation of Mitochondrial Fission", *Biochimica et Biophysica Acta*, Vol. 1833, No. 5, pp. 1256-68.

Oyamada, M., K. Takebe, and Y. Oyamada, 2013, "Regulation of Connexin Expression by Transcription Factors and Epigenetic Mechanisms", *Biochimica et Biophysica Acta*, Vol. 1828, No. 1, pp. 118-33.

Parry, D. A., A. J. Mighell, W. El-Sayed, R. C. Shore, I. K. Jalili, H. Dollfus, A. Bloch-Zupan, R. Carlos, I. M. Carr, L. M. Downey, *et al.*, 2009, "Mutations in Cnnm4 Cause Jalili Syndrome, Consisting of Autosomal-Recessive Cone-Rod Dystrophy and Amelogenesis Imperfecta", *American Journal of Human Genetics*, Vol. 84, No. 2, pp. 266-73.

Patzko, A., and M. E. Shy, 2011, "Update on Charcot-Marie-Tooth Disease", *Current Neurology and Neuroscience Reports*, Vol. 11, No. 1, pp. 78-88.

Raeymaekers, P., V. Timmerman, E. Nelis, P. De Jonghe, J. E. Hoogendijk, F. Baas, D. F. Barker, J. J. Martin, M. De Visser, P. A. Bolhuis, *et al.*, 1991, "Duplication in Chromosome 17p11.2 in Charcot-Marie-Tooth Neuropathy Type 1a (Cmt 1a). The Hmsn Collaborative Research Group", *Neuromuscular Disorders*, Vol. 1, No. 2, pp. 93-7.

Ranieri, M., S. Brajkovic, G. Riboldi, D. Ronchi, F. Rizzo, N. Bresolin, S. Corti, and G. P. Comi, 2013, "Mitochondrial Fusion Proteins and Human Diseases", *Neurology Research International*, Vol. 2013, No. 293893.

Rao, R. G., D. Sudhakar, C. P. Hogue, S. Amici, L. K. Gordon, J. Braun, L. Notterpek, L. Goodglick, and M. Wadehra, 2011, "Peripheral Myelin Protein-22 (Pmp22) Modulates Alpha 6 Integrin Expression in the Human Endometrium", *Reproductive Biology and Endocrinology*, Vol. 9, No. 56.

Reilly, M. M., S. M. Murphy, and M. Laura, 2011, "Charcot-Marie-Tooth Disease", *Journal of the Peripheral Nervous System*, Vol. 16, No. 1, pp. 1-14.

Ribas de Pouplana, L., D. Buechter, N. Y. Sardesai, and P. Schimmel, 1998, "Functional Analysis of Peptide Motif for RNA Microhelix Binding Suggests New Family of RNA-binding Domains", *EMBO Journal*, Vol. 17, pp. 5449–5457.

Saez, J. C., K. A. Schalper, M. A. Retamal, J. A. Orellana, K. F. Shoji, and M. V. Bennett, 2010, "Cell Membrane Permeabilization Via Connexin Hemichannels in Living and Dying Cells", *Experimental Cell Research*, Vol. 316, No. 15, pp. 2377-89.

Saporta, A. S., S. L. Sottile, L. J. Miller, S. M. Feely, C. E. Siskind, and M. E. Shy, 2011, "Charcot-Marie-Tooth Disease Subtypes and Genetic Testing Strategies", *Annals of Neurology*, Vol. 69, No. 1, pp. 22-33.

Schlisio, S., R. S. Kenchappa, L. C. Vredeveld, R. E. George, R. Stewart, H. Greulich, K. Shahriari, N. V. Nguyen, P. Pigny, P. L. Dahia, *et al*., 2008, "The Kinesin Kif1bbeta Acts Downstream from Egln3 to Induce Apoptosis and Is a Potential 1p36 Tumor Suppressor", *Genes and Development*, Vol. 22, No. 7, pp. 884-93.

Shy, M. E., A. Jani, K. Krajewski, M. Grandis, R. A. Lewis, J. Li, R. R. Shy, J. Balsamo, J. Lilien, J. Y. Garbern, and J. Kamholz, 2004, "Phenotypic Clustering in Mpz Mutations", *Brain*, Vol. 127, No. Pt 2, pp. 371-84.

Sidiropoulos, P. N., M. Miehe, T. Bock, E. Tinelli, C. I. Oertli, R. Kuner, D. Meijer, B. Wollscheid, A. Niemann, and U. Suter, 2012, "Dynamin 2 Mutations in Charcot-Marie-Tooth Neuropathy Highlight the Importance of Clathrin-Mediated Endocytosis in Myelination", *Brain*, Vol. 135, No. Pt 5, pp. 1395-411.

Siskind, C. E., S. Panchal, C. O. Smith, S. M. Feely, J. C. Dalton, A. B. Schindler, and K. M. Krajewski, 2013, "A Review of Genetic Counseling for Charcot Marie Tooth Disease (Cmt)", *Journal of Genetic Counseling*, Vol. 22, No. 4, pp. 422-36.

Skre, H., 1974, "Genetic and Clinical Aspects of Charcot-Marie-Tooth's Disease", *Clinical Genetics*, Vol. 6, No. 2, pp. 98-118.

Snipes, G. J., and U. Suter, 1995, "Molecular Anatomy and Genetics of Myelin Proteins in the Peripheral Nervous System", *Journal of Anatomy*, Vol. 186, No. Pt 3, pp. 483-94.

Stauch, K., F. Kieken, and P. Sorgen, 2012, "Characterization of the Structure and Intermolecular Interactions between the Connexin 32 Carboxyl-Terminal Domain and the Protein Partners Synapse-Associated Protein 97 and Calmodulin", *Journal of Biological Chemistry*, Vol. 287, No. 33, pp. 27771-88.

Szigeti, K., and J. R. Lupski, 2009, "Charcot-Marie-Tooth Disease", *European Journal of Human Genetics*, Vol. 17, No. 6, pp. 703-10.

Tazir, M., M. Bellatache, S. Nouioua, and J. M. Vallat, 2013, "Autosomal Recessive Charcot-Marie-Tooth Disease: From Genes to Phenotypes", *Journal of the Peripheral Nervous System*, Vol. 18, No. 2, pp. 113-29.

Teruo, H., T. Shang-Yi, M. Tomohisa, F. Michiko, and S. Tsung-Ping, 2011, "Targeting Ligand-operated Chaperone Sigma-1 Receptors in the Treatment of Neuropsychiatric Disorders", *Expert Opinion on Therapeutic Targets*, Vol. 15, No. 5, pp. 557–577

Timmerman, V., E. Nelis, W. Van Hul, B. W. Nieuwenhuijsen, K. L. Chen, S. Wang, K. Ben Othman, B. Cullen, R. J. Leach, C. O. Hanemann, *et al*., 1992, "The Peripheral Myelin Protein Gene Pmp-22 Is Contained within the Charcot-Marie-Tooth Disease Type 1a Duplication", *Nature Genetics*, Vol. 1, No. 3, pp. 171-5.

Tooth, H. H., 1886, "The Peroneal Type of Progressive Muscular Atrophy", Ph.D Thesis, H. K. Lewis, London.

Tsai, P. C., C. H. Chen, A. B. Liu, Y. C. Chen, B. W. Soong, K. P. Lin, S. F. Yet, and Y. C. Lee, 2013, "Mutational Analysis of the 5' Non-Coding Region of Gjb1 in a Taiwanese Cohort with Charcot-Marie-Tooth Neuropathy", *Journal of the Neurological Sciences*, Vol. 332, No. 1-2, pp. 51-5.

Tsung-Ping, S., T. Hayashi, T. Maurice, S. Buch, and A. E. Ruoho, 2010, "The Sigma-1 Receptor Chaperone as an Inter-organelle Signaling Modulator", *Trends in Pharmacological Sciences*, Vol. 31, No. 12, pp. 557–566

Vogt, D. L., C. D. Gray, W. S. Young, 3rd, S. A. Orellana, and A. T. Malouf, 2007, "Arhgap4 Is a Novel Rhogap That Mediates Inhibition of Cell Motility and Axon Outgrowth", *Molecular and Cellular Neurosciences*, Vol. 36, No. 3, pp. 332-42.

Wagner, K. M., M. Ruegg, A. Niemann, and U. Suter, 2009, "Targeting and Function of the Mitochondrial Fission Factor Gdap1 Are Dependent on Its Tail-Anchor", *PLoS One*, Vol. 4, No. 4, pp. e5160.

Weedon, M. N., R. Hastings, R. Caswell, W. Xie, K. Paszkiewicz, T. Antoniadi, M. Williams, C. King, L. Greenhalgh, R. Newbury-Ecob, *et al*., 2011, "Exome Sequencing Identifies a Dync1h1 Mutation in a Large Pedigree with Dominant Axonal Charcot-Marie-Tooth Disease", *American Journal of Human Genetics*, Vol. 89, No. 2, pp. 308-12.

Weterman, M. A., V. Sorrentino, P. R. Kasher, M. E. Jakobs, B. G. van Engelen, K. Fluiter, M. B. de Wissel, A. Sizarov, G. Nurnberg, P. Nurnberg, *et al*., 2012, "A Frameshift Mutation in Lrsam1 Is Responsible for a Dominant Hereditary Polyneuropathy", *Human Molecular Genetics*, Vol. 21, No. 2, pp. 358-70.

Yao, P., and P. L. Fox, 2013, "Aminoacyl-Trna Synthetases in Medicine and Disease", *EMBO Molecular Medicine*, Vol. 5, No. 3, pp. 332-43.

Zuchner, S., I. V. Mersiyanova, M. Muglia, N. Bissar-Tadmouri, J. Rochelle, E. L. Dadali, M. Zappia, E. Nelis, A. Patitucci, J. Senderek, *et al*., 2004, "Mutations in the

Mitochondrial Gtpase Mitofusin 2 Cause Charcot-Marie-Tooth Neuropathy Type 2a", *Nature Genetics*, Vol. 36, No. 5, pp. 449-51.

Zuchner, S., M. Noureddine, M. Kennerson, K. Verhoeven, K. Claeys, P. De Jonghe, J. Merory, S. A. Oliveira, M. C. Speer, J. E. Stenger, *et al*., 2005, "Mutations in the Pleckstrin Homology Domain of Dynamin 2 Cause Dominant Intermediate Charcot-Marie-Tooth Disease", *Nature Genetics*, Vol. 37, No. 3, pp. 289-94.