IDENTIFICATION OF NEW LOCI, GENES, AND MUTATIONS RESPONSIBLE FOR HEREDITARY SPASTIC PARAPLEGIA

by

Burçak Özeş Ak

B.S., Molecular Biology and Genetics, Boğaziçi University, 2008

M.S., Molecular Biology and Genetics, Boğaziçi University, 2011

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

Graduate Program in Molecular Biology and Genetics

Boğaziçi University

May, 2018

IDENTIFICATION OF NEW LOCI, GENES, AND MUTATIONS RESPONSIBLE FOR HEREDITARY SPASTIC PARAPLEGIA

APPROVED BY:

Prof. Esra Battaloğlu (Thesis Supervisor)	……
Prof. Hande Çağlayan	……
Assoc. Prof. Sibel Uğur İşeri	……
Prof. Yeşim Parman	……
Prof. Zehra Oya Uyguner	……

DATE OF APPROVAL: 29.05.2018

To my father...

ACKNOWLEDGEMENTS

First of all, I would like to express my sincere appreciation to my thesis supervisor Prof. Esra Battaloğlu for her invaluable guidance, and endless support throughout the study.

I would like show my gratitude to my thesis progress committee members Prof. Hande Çağlayan and Prof. Yeşim Parman for their valuable criticism. I also would like to extend my thanks to other thesis committee members Assoc. Prof. Sibel Uğur İşeri and Prof. Zehra Oya Uyguner for devoting their time to evaluate this thesis.

I am grateful to Prof. Stephan Züchner for accepting six of the HSP families for WES project performed in University of Miami, USA. I am thankful to Prof. Peter De Jonghe for accepting me as a visitor scientist to his laboratory in University of Antwerp, Belgium and to MBG-BRIDGE project for supporting this visit, financially. I am thankful to Prof. Yesim Parman, and all other clinicians for providing blood samples and clinical data of patients. I would like to thank the families for participating in this study.

I am grateful to all of academic staff of the department for their contribution to my improvement and to Ümit Bayraktar for her endless kindness. I would thank to former and current CMT lab members Alperen Erdoğan, Kaya Akyüz, Kerem Yıldırım, Merve Kılınç, Merve Sıvacı, Begüm Gökerküçük, Cansu Küey, Ayşe Candayan, Nazmiye Özkan and to my friends Neslihan Zohrap, Mahmut Hız, Harun Niron, and Ceren Saygı. I would like to share my warmest thanks with my lifelong friends Yeliz Yılmaz, Aslı Uğurlu and Elif Eren. I am thankful to all interns, especially Fatma Tiryaki, for assisting me with the experiments.

I would like to share my deepest appreciation to my father, my mother and my brother for providing me a wonderful family love. Last but not least, thanks from deep of my heart goes to my husband, Gürbüz Ak, for being in my life and making it fabulous.

This study was supported by The Scientific and Technological Research Council of Turkey (TÜBİTAK-114S725) and Boğaziçi University Research Fund (6047, 11940).

ABSTRACT

IDENTIFICATION OF NEW LOCI, GENES, AND MUTATIONS RESPONSIBLE FOR HEREDITARY SPASTIC PARAPLEGIA

Hereditary Spastic Paraplegia (HSP), characterized by lower limb spasticity and progressive weakness, is a group of inherited neurodegenerative disorders. 47 of the 63 genes responsible for HSP, have been associated with autosomal recessive (AR) HSP. In this study, whole exome sequencing analysis was performed for one patient from each 27 ARHSP families to identify causative genes. When WES data was not informative, homozygosity mapping was performed by using WES and/or whole genome SNP genotyping data. After segregation analyses of candidate variants, the causative variants were identified in fifteen families. *SPG11* gene was causative in four families. Single families had mutations in *CYP7B1* (SPG5A), *SPG7, SPG15,* and *ALS2* genes. *SACS* and *CYP27A1* genes that are associated with Charlevoix-Saguenay Spastic Ataxia and CTX, respectively were determined as the causative genes in two families providing differential diagnosis to these families. *KIF1C* (SPG58) (Caballero-Oteyza *et al.,* 2014) and *PLA2G6* genes (Ozes *et al.,* 2017) identified in two families were reported as novel HSP genes. *SAMHD1, ATAD1,* and *SEMA3C* were identified as strong HSP gene candidates. Immortalized B-lymphocytes derived from family members were analyzed to unravel the involvement of candidate HSP genes on disease pathogenesis. ATAD1, SAMHD1 and CYP27A1 protein levels were shown to be reduced in immortalized cell lines of patients. This study contributed to understanding of the genetic heterogeneity of HSP by identifying five novel HSP genes, one of which locates to SPG27. We also underlined the importance of genetic analysis for differential diagnosis, and the necessity of primary genetic screening of *SPG11* in AR-HSP-TCC and *CYP7B1* in pure AR-HSP cases highlighting the effect of correct genotype-phenotype correlation.

ÖZET

HEREDİTER SPASTİK PARAPAREZİDEN SORUMLU YENİ LOKUS, GEN VE MUTASYONLARIN TANIMLANMASI

Herediter Spastik Paraparezi (HSP) alt bacaklarda spastisite ve ilerleyen zaaf ile karakterize kalıtsal nörodejeneratif bir hastalıklar grubudur. HSP'den sorumlu 63 genin 47'si otozomal çekinik HSP (AR-HSP) ile ilişkilendirilmiştir. Bu çalışmada, 27 AR-HSP ailesinden bir hasta bireyde tüm ekzom dizileme (WES) uygulanarak hastalığa neden olan genlerin belirlenmesi amaçlanmıştır. WES verisinde kuvvetli adayların belirlenemediği ailelerde WES ve/veya tüm genom SNP genotipleme verileri kullanılarak homozigotluk haritalaması çalışması yapılmıştır. Aday varyantlar, ailelerde segregasyon analizleri ile incelenerek on beş ailede hastalığa neden olan gen belirlenmiştir. Dört ailede *SPG11* geninde, birer ailede de *CYP7B1* (SPG5A), *SPG7, SPG15, ve ALS2* genlerinde mutasyon belirlenmiştir. Sırasıyla Charlevoix-Saguenay Spastik Ataksi ve CTX hastalıklarına neden olan *SACS* ve *CYP27A1* genlerinde birer ailede mutasyon görülmüş ve bu iki aile için ayırt edici tanıya ulaşılmıştır. Birer ailede görülen KIF1C (SPG58) ve PLA2G6 genleri yeni HSP genleri olarak belirlenmiş ve rapor edilmişlerdir (Caballero-Oteyza *et al.,* 2014; Ozes *et al.,* 2017). Bunların yanı sıra birer ailede hastalığa neden olduğu görülen *SAMHD1, ATAD1,* ve *SEMA3C* genleri kuvvetli birer yeni HSP geni olarak ilk kez bu çalışmada tanımlanmışlardır. Yeni HSP geni adaylarının hastalığa etkisini göstermek üzere aile bireylerinden elde edilerek ölümsüzleştirilen B-lenfosit hücre hatları analiz edilmiştir ve bu hücre hatlarında *ATAD1, SAMHD1* ve *CYP27A1* genlerinde protein seviyesinde azalma görülmüştür. Bu çalışma ile biri SPG27'ye lokalize beş yeni HSP geni tanımlayarak HSP'nin genetik heterojenliğine katkıda bulunulmuştur. Ayrıca, TCC-AR-HSP vakalarında *SPG11* ve saf AR-HSP vakalarında *CYP7B1* genlerinin öncelikli olarak incelenmesi gerekliliği ve ayırt edici tanının önemi bir kez daha gösterilmiştir.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS iv
ABSTRACT v
ÖZET vi
TABLE OF CONTENTS vii
LIST OF FIGURES xi
LIST OF TABLES xxvi
LIST OF ACRONYMS/ABBREVIATIONS xxxi
1. INTRODUCTION 1
1.1. Autosomal Dominant HSP 2
1.2. Autosomal Recessive HSP (AR-HSP) 3
1.3. X-Linked HSP (XL-HSP) 3
1.4. Other HSPs 3
1.5. Pathophysiology of HSP 3
1.5.1. ER Morphogenesis and membrane trafficking 4
1.5.2. Mitochondrial Function 7
1.5.3. Lipid Synthesis and Metabolism 8
2. PURPOSE 26
3. MATERIALS 28
3.1. Subjects 28
3.2. Primers 29
3.3. Primary and Secondary Antibodies 42
3.4. Cell Line Mediums 44
3.5. Buffers and Solutions 44
3.6. Chemicals and Disposables 48
3.7. Equipment 50
4. METHODS 52
4.1. Molecular Genetic Analysis 52
4.1.1. Pedigree Analysis 52
4.1.2. DNA Extraction from Peripheral Blood 52

4.1.3. Quantitative Analysis of Extracted DNA 53
4.1.4. Polymerase Chain Reaction (PCR) 53
4.1.5. Mutation Screening in Spastin Gene 53
4.1.5.1. HRM Analyses 54
4.1.5.2. Single Strand Conformation Polymorphism (SSCP) Analysis 54
4.1.5.3. Silver Staining 55
4.1.6. Sanger Sequencing 55
4.1.7. Whole Exome Sequencing (WES) 55
4.1.7.1. WES and Initial Analyses 55
4.1.7.2. WES Data Analysis 56
4.1.7.3. Verification of variant 56
4.1.8. SNP Genotyping and Homozygosity Mapping 57
4.1.9. Identification of ROHs by PLINK 57
4.1.10. Comparing WES & SNP Genotyping Data & Searching for Candidates 59
4.1.11. 3D Protein Model Prediction 59
4.2. In vitro Analyses 59
4.2.1. Immortalization of Human B-Lymphocytes 59
4.2.1.1. B-Lymphocyte Isolation from Blood 59
4.2.1.2. Epstein-Barr Virus (EBV) Production 60
4.2.1.3. Transduction of Human B-Lymphocytes 60
4.2.2. Monitoring mRNA and Protein Levels 60
4.2.2.1. Quantitative real-time polymerase chain reaction (qPCR) 60
4.2.2.2. Western Blot Analysis 61
4.2.3. Immunocytochemical Analysis 62
5. RESULTS 64
5.1. Mutation Screening in Spastin Gene 64
5.2. Molecular Genetic Analysis 66
5.2.1. Confirmation of PLINK and Its Comparison to SNP Genotyping data 67
5.2.2. Family P392 70
5.2.3. Family P463 73
5.2.4. Family P627 75
5.2.5. Family H6 76
5.2.6. Family H16 79

5.2.7. Family H28 80
5.2.8. Family H29 85
5.2.9. Family H38 87
5.2.10. Family H44 91
5.2.11. Family H45 93
5.2.12. Family H52 96
5.2.13. Family H53 98
5.2.14. Family H55 98
5.2.15. Family H57 104
5.2.16. Family H59 107
5.2.17. Family H61 107
5.2.18. Family H65 108
5.2.19. Family H72 110
5.2.20. Family H77 111
5.2.21. Family H82 112
5.2.22. Family H98 114
5.2.23. Family H99 114
5.2.24. Family H108 117
5.2.25. Family H110 119
5.2.26. Family H133 121
5.2.27. Family H142 123
5.2.28. Family H143 128
5.2.29. Population screening 129
5.2.30. Summary of genetic findings 130
5.3. Effect of Rare Variants on Disease Manifestation 131
5.4. *In vitro* Analyses 139
5.4.1. Analyses of Immortalized B-Lymphocytes 139
5.4.1.1. Family H28 and ATAD1 Gene 140
5.4.1.2. Family H55 and CYP27A1 and ZNF142 Genes 142
5.4.1.3. Family H57 and SAMHD1 Gene 144
5.4.1.4. Family H142 and SEMA3C Gene 147
6. DISCUSSION 149
7. CONCLUSION 159

REFERENCES 161
APPENDIX A: CHROMATOGRAMS.....178
APPENDIX B: 3D PROTEIN MODELS.....228
APPENDIX C: RESULTS OF TECHNICAL REPLICATES 235

LIST OF FIGURES

Figure 1.1. Common pathogenic cellular processes in the HSPs 4

Figure 5.1. Normalized and temperature-shifted melting curves. 65

Figure 5.2. Eight per cent SSCP gel photographs.....65

Figure 5.3. Calculated effect of c.1222-38_1222-37del variant 66

Figure 5.4. ROHs identified in the WES data of H98.3. 68

Figure 5.5. Distribution of ROHs determined on WES data of H98.3. 68

Figure 5.6. Update of ROH limiting positions.....69

Figure 5.7. Comparison of ROH determined by PLINK with SNP genotyping data.....69

Figure 5.8. ROHs identified in the WES data of P392. 70

Figure 5.9. Distribution of ROHs identified for the patient P392. 71

Figure 5.10. Evolutionary conservation of SPG11 protein.....74

Figure 5.11. Conservation of Glycine amino acid in 213th position in SACS gene. 76

Figure 5.12. ROHs identified for the patients in family H6. 77

Figure 5.13. Distribution of ROHs identified for the patients in family H6.....78

Figure 5.14. ROHs identified for the patients in family H28. 81

Figure 5.15. Distribution of ROHs identified for the patients in family H28.....83

Figure 5.16. Evolutionary conservation of p.Met234.....83

Figure 5.17. ROHs identified for the patients in family H29. 85

Figure 5.18. Distribution of ROHs identified for the patients in family H29.....86

Figure 5.19. ROHs identified for the patients in family H38. 89

Figure 5.20. Distribution of ROHs identified for the patients in family H38.....89

Figure 5.21. ROHs identified for the patients in family H44. 92

Figure 5.22. Distribution of ROHs identified for the patients in family H44.....92

Figure 5.23. ROHs identified for the patients in family H45. 94

Figure 5.24. Distribution of ROHs identified for the patients in family H45.....95

Figure 5.25. The KIF1C mutations.....97

Figure 5.26. ROHs identified for the patients in family H55. 99

Figure 5.27. Alteration in acceptor site when c.256-1G>C variant.....101

Figure 5.28. Conservation of arginine amino acid on ZNF142 protein.....102

Figure 5.29. Distribution of ROHs identified for the patients in family H55.....102

Figure 5.30. Conservation of leucine amino acid on SAMHD1 gene 106

Figure 5.31. ROHs identified for the patients in family H55. 109

Figure 5.32. Scores indicating the possible effect of c.2104-2A>G variant on splicing. ..110

Figure 5.33. Distribution of ROHs identified for the patients in family H72.....111

Figure 5.34. ROHs identified for the patients in family H82. 112

Figure 5.35. Distribution of ROHs identified for the patients in family H82.....112

Figure 5.36. ROHs identified for the patients in family H99. 115

Figure 5.37. Distribution of ROHs identified for the patients in family H99.....116

Figure 5.38. Distribution of ROHs identified for the patients in family H108.....117

Figure 5.39. ROHs identified for the patients in family H108. 118

Figure 5.40. ROHs identified for the patient H110. 119

Figure 5.41. Distribution of ROHs identified for the patient H110.....120

Figure 5.42. ROHs identified for the patient H133. 121

Figure 5.43. Distribution of ROHs identified for the patient H133.....122

Figure 5.44. ROHs identified for the patients in family H142. 124

Figure 5.45. Distribution of ROHs identified for the patients in family H142.....124

Figure 5.46. Position of SEMA3C on distribution of ROHs on chromosomes.....126

Figure 5.47. Evolutionary conservation of tyrosine amino acid on SEMA3C protein.....127

Figure 5.48. Graph showing the number of rare variants in ARHSP patietns.....139

Figure 5.49. RT-PCR analysis of mRNA levels of ATAD1 gene for H28 family 140

Figure 5.50. Western blot analyses of ATAD1 protein. 140

Figure 5.51. Immunostaining for ATAD1 protein.....141

Figure 5.52. RT-PCR analysis of mRNA levels of CYP27A1 gene for H55 family 142

Figure 5.53. RT-PCR analysis of mRNA levels of ZNF142 gene for H55 family.....142

Figure 5.54. Western blot analyses of CYP27A1 protein performed for H55 family.....143

Figure 5.55. Immunostaining for CYP27A1 protein 144

Figure 5.56. RT-PCR analysis of mRNA levels of SAMHD1 gene for H57 family.....144

Figure 5.57. Western blot analyses of SAMHD1 protein performed for H57 family 145

Figure 5.58. Immunostaining for SAMHD1 protein for family H57 146

Figure 5.59. Immunostaining for SAMHD1 protein for H57.3 and control individual.....146

Figure 5.60. RT-PCR analysis of mRNA levels of SEMA3C gene for H142 family 147

Figure 5.61. Immunostaining for SEMA3C protein 148

Figure A.1. H82 pedigree and chromatograms showing c.1222-38_1222-37del variant intron 10 of *SPAST* gene. 178

Figure A.2. P392 pedigree and chromatograms showing the c.157G>A variant in KCNMB3 gene and the c.47G>A variant in SLC7A2 gene.....178

Figure A.3. P392 pedigree and chromatograms showing the c.980T>A variant in PDGFRL gene and the c.590G>A variant in NAT2 gene 179

Figure A.4. P392 pedigree and chromatograms showing the c.1070C>T variant in SCYL2 gene and the c.1505C>G variant in UTP20 gene 179

Figure A.5. P392 pedigree and chromatograms showing the c.1443C>G variant in MYBPC1 gene and c.1906-6C>A variant in CAPRIN2 gene.....180

Figure A.6. P392 pedigree and chromatograms showing the c.2783T>C variant in OVOS2, c.136C>T variant in CCDC114 and the c.734C>T variant in ETFB genes.....180

Figure A.7. P463 pedigree and chromatograms showing c.1235C>G, p.S412X variant in SPG11 gene in affected individuals.....181

Figure A.8. P463 pedigree and chromatograms showing c.3794T>A, p.L1265Q variant in ZFYVE26 gene in affected individuals 181

Figure A.9. P463 pedigree and chromatograms showing c.985A>G, p.S329G variant in ZFYVE26 gene in affected individuals 181

Figure A.10. P627 pedigree and chromatogram showing c.1087-3delT and c.1087-3delTT variants in CES1 gene, in patient P627.....182

Figure A.11. P627 pedigree and chromatogram showing c.637G>A variant in SACS gene, in patient P627.....182

Figure A.12. H6 pedigree and the chromatograms showing n.387A>G variant in SDHAP2 and absence of c.1317-3C>T variant in ANKRD20A4 gene 182

Figure A.13. H6 pedigree and chromatograms showing the c.719T>G variant in OR11H12 gene and c.4708G>A variant in RGPD8 gene.....183

Figure A.14. H6 pedigree and chromatograms showing the c.20C>T variant in GUCA2A gene and the c.1852G>T variant in SLC5A9 gene.....183

Figure A.15. H6 pedigree and chromatograms showing the c.1018C>T variant in PRKAG3 gene and the c.443-6T>C variant in SLC23A3 gene 184

Figure A.16. H6 pedigree and chromatograms showing the c.90G>T variant in RESP18 gene and the c.5+6G>T variant in IFIT5 gene 184

Figure A.17. H6 pedigree and chromatograms showing the c.3576-5C>T variant in MYOM1 gene and the c.161T>C variant in SPINT3 gene 185

Figure A.18. H16 pedigree and chromatograms showing c.6215_6219dupAGAT, p.Phe2074ArgfsTer15 insertion in SPG11 gene.....185

Figure A.19. H16 pedigree and chromatograms showing the sequences including c.*3745G>A and c.424A>G (right panel) in C19orf12 gene 186

Figure A.20. H28 pedigree and chromatogram showing c.-1G>C and c.2287G>A variants in AP5Z1 gene in patient H28.1 and H28.2 186

Figure A.21. H28 pedigree and chromatograms showing the c.411-4C>A and c.411-5G>A variants in CERS2 and the c.9206C>T variant in COL6A3 genes.....187

Figure A.22. H28 pedigree and chromatograms showing the c.3751G>A variant in COL6A3 gene and the c.7235C>T variant in TNXB gene.....187

Figure A.23. H28 pedigree and chromatograms showing the c.8111G>A variant in TNXB gene and c.3239G>A variant in LAMA4 gene.....188

Figure A.24. H28 pedigree and chromatograms showing the c.431T>G variant in GPRC6A gene and the c.730A>G variant in FAM35A gene.....188

Figure A.25. H28 pedigree and chromatograms showing the the c.103C>T variant in NUTM2D gene and the c.701T>C variant in ATAD1 gene 189

Figure A.26. H28 pedigree and chromatograms showing the c.511A>G variant in LIPF gene and the c.980C>G variant in HSDL1 189

Figure A.27. H28 pedigree and chromatograms showing the c.524G>A variant in ARSD gene and the c.959G>A variant in ARSD gene. 190

Figure A.28. H28 pedigree and chromatograms showing the c.2299G>C variant in ATP7A gene and the c.3060dupC variant in DGKK gene 190

Figure A.29. H28 pedigree and chromatograms showing the c.160A>C variant in FAM104B gene and the c.217C>T variant in H2BFM gene.....191

Figure A.30. H28 pedigree and chromatograms showing the c.82-7_82-6dupCT variant in IDH3G gene and the c.544G>A variant in SSX7 gene 191

Figure A. 31. H29 pedigree and chromatograms showing c.4804C>T variant 192

Figure A.32. H29 pedigree and chromatograms showing c.1168G>A, p.Gly390Ser variant in RTN2 gene.....192

Figure A.33. H29 pedigree and chromatograms showing the c.40C>G variant in TTC22 gene and the c.15446-7G>A variant in SSPO gene 193

Figure A.34. H29 pedigree and chromatograms showing the c.590G>A variant in NAT2 gene and the c.1547C>T variant in SLC22A10 gene.....193

Figure A.35. H29 pedigree and chromatograms showing the variant c.401G>A in KRR1 gene and the c.1871G>T variant in AKAP13 gene.....194

Figure A.36. H29 pedigree and chromatograms showing the the c.1144C>G variant in PDIA2 gene and the c.2594C>T variant in WDR90 gene.....194

Figure A.37. H29 pedigree and chromatograms showing the c.1798-8G>A variant 195

Figure A.38. H38 pedigree and chromatograms showing the c.39962C>T variant 195

Figure A.39. H38 pedigree and chromatograms showing the c.63917G>A variant 195

Figure A.40. H38 pedigree and chromatograms showing the c.-29delG variant 196

Figure A.41. H38 pedigree and chromatograms showing the c.2128G>A variant in EMC1 gene and the c.67G>A variant in RNF186 gene 196

Figure A.42. H38 pedigree and chromatograms showing the c.3972-5T>C variant in MYOM3 gene and c.901A>G variant in TRAPPC12 gene.....196

Figure A.43. H38 pedigree and chromatograms showing the c.293C>T variant in APOB gene and the c.847G>A variant in CAPN14 gene 197

Figure A.44. H38 pedigree and chromatograms showing the c.1294C>G variant in WFS1 gene the c.2128G>A variant in LOXL4 gene.....197

Figure A.45. H38 pedigree and chromatograms showing the c.67G>A variant in PKD2L1 gene and the c.2128G>A variant in PREX1 gene 197

Figure A.46. H44 pedigree and chromatograms showing c.7417-5G>C variant in 198

Figure A.47. H44 pedigree and chromatograms showing c.7607G>A variant 198

Figure A.48. H44 pedigree and chromatograms showing c.503T>C variant 198

Figure A.49. H45 pedigree and chromatograms showing c.2245-1G>A variant 199

Figure A.50. H45 pedigree and chromatograms showing c.2057T>C variant 199

Figure A.51. H45 pedigree and chromatograms showing c.1499_1500delAG, p.E500Vfs57X variant in SPG11 gene 200

Figure A.52. H45 pedigree and chromatograms showing the c.1999G>A variant in PLCL1 gene and the c.2487A>C variant in ALPK2 gene.....200

Figure A.53. H45 pedigree and chromatograms showing c.2748T>A variant and the c.4345C>T variant in ALPK2 gene. 200

Figure A.54. H52 pedigree and chromatograms showing c.1214G>A variant 201

Figure A.55. H53 pedigree and chromatograms showing c.2239G>A variant and domain structure of the PLA2G6 gene indicating the position of the mutations identified in this study. 201

Figure A.56. Chromatogram showing c.2239G>A variant in patient HSS-2. 201

Figure A.57. H55 pedigree and the chromatograms showing c.571C>T variant 202

Figure A.58. H55 pedigree and chromatograms showing c.3373C>T variant in ZNF142 gene and the c.256-1G>C variant in CYP27A1 gene 202

Figure A.59. H55 pedigree and chromatograms showing c.443-6T>C variant in SLC23A3 gene and the c.90G>Tvariant in RESP18 gene 203

Figure A.60. H55 pedigree and chromatograms showing c.1615G>A variant in FAM198A gene and the c.1348-7T>A variant in NT5DC2 gene 203

Figure A.61. H55 pedigree and chromatograms showing c.1754A>T variant in ITIH1 gene and the c.1379G>A variant in CCDC66 gene.....204

Figure A.62. H55 pedigree and chromatograms showing c.3923A>G variant in DNAH12 gene and the c.77T>C variant in CDC27 gene.....204

Figure A.63. H55 pedigree and chromatograms showing c.806-1G>A variant in EFCAB13 gene.....205

Figure A.64. H55 pedigree and chromatograms showing c.1627G>A variant in HEXB gene and the c.208A>G variant in DST gene. 205

Figure A.65. H55 pedigree and chromatograms showing c.589G>A variant in FKBP14 gene and the c.175T>A variant in SYT12 gene.....206

Figure A.66. H55 pedigree and chromatograms showing c.452C>G variant in MYH4 gene and the c.2395C>T variant in CTDP1 gene.....206

Figure A.67. H57 pedigree and chromatograms showing c.674C>A variant in ERLIN2 gene and c.6062G>A variant in ZFYVE26 207

Figure A.68. H57 pedigree and chromatograms showing c.8611T>C variant in SACS gene and variant in position chr9:92715016 in BICD2 gene.....207

Figure A.69. Chromatograms showing variants in genes CEP350 and SH3TC1.....208

Figure A.70. H57 pedigree and chromatograms showing c.754G>A variant in PIP5K1B (FAM122A) gene and c.79C>G variant in COQ9 gene.....208

Figure A.71. H57 pedigree and chromatograms showing c.1360C>G (left panel) and c.1373C>T variants.....209

Figure A.72. H57 pedigree and chromatograms showing c.2591G>A variant in RBP3 gene and c.533T>C (p.L178P) variant in SAMHD1 gene.....209

Figure A.73. H59 pedigree and chromatograms showing c.325_326insTGTC variant 210

Figure A.74. H61 pedigree and chromatograms showing c.4321C>T variant 210

Figure A.75. H65 pedigree and chromatograms showing c.11264T>C variant 211

Figure A.76. H65 pedigree and chromatograms showing c.14378C>A variant.....211

Figure A.77. H65 pedigree and chromatograms showing the c.4211_4213delGTA variant in VPS13B gene.....212

Figure A.78. H65 pedigree and chromatograms showing the c.6034C>T variant 212

Figure A.79. H65 pedigree and chromatograms showing the c.334T>G variant 213

Figure A.80. H72 pedigree and chromatograms showing c.2104-2A>G variant 213

Figure A.81. H77 pedigree and chromatograms showing c.3036C>A variant.....214

Figure A.82. H82 pedigree and chromatograms showing the c.7612C>G variant in POLQ gene and the c.4515_4516delTC variant in 13th exon of ABCA10 gene.....214

Figure A.83. H82 pedigree and chromatograms showing the c.3964C>T variant and the c.1331_1334delCTGT variant in ABCA10 gene.....215

Figure A.84. H82 pedigree and chromatograms showing the c.290C>T variant in ZNF781 gene and the c.313delA variant in 4th exon of ZNF880 gene 215

Figure A.85. H82 pedigree and chromatograms showing the c.1412A>G variant in 4th exon of ZNF880 gene and the c.155C>T variant in VN1R4 gene.....216

Figure A.86. H82 pedigree and chromatograms showing the c.1010G>T variant in 7th exon the c.149T>A variant in 4th exon of ZSCAN5A gene.....216

Figure A.87. H98 pegidree and chromatograms showing c.825T>A (p.Tyr275Ter) variant in CYP7B1 gene 217

Figure A.88. H99 pedigrees and chromatograms showing c.7417-5G>C, c.5806T>A and c.1844C>T variants in ZFYVE26 gene 217

Figure A.89. H99 pedigree and chromatograms showing the c.4C>T variant FUCA1 gene, the c.9211_9212insC variant in SSPO and c.247C>T variant in GIMAP7 genes.....218

Figure A.90. H108 pedigree and chromatograms showing c.*1020C>A variant.....218

Figure A.91. H108 pedigree and chromatograms showing c.1343+6T>G variant.....219

Figure A.92. H108 pedigree and chromatograms showing c.820G>A variant 219

Figure A.93. H108 pedigree and chromatograms showing the c.1122G>C variant in SLC45A2 gene and the c.524G>A variant in AMACR gene.....220

Figure A.94. H108 pedigree and chromatograms showing the c.1132G>C variant in NEBL gene and the c.353C>T variant in C10orf53 gene 220

Figure A.95. H110 pedigree and chromatogram showing c.12166A>C variant 221

Figure A.96. Chromatograms showing the variants listed in Table 34. 221

Figure A.97. H133 pedigree and chromatogram showing c.437A>T variant 221

Figure A.98. Chromatograms showing the variants listed in Table 35. 222

Figure A.99. H142 pedigree and chromatograms showing n.450_451insAAC insertion in C14orf23 gene.....222

Figure A.100. H142 pedigree and chromatograms showing c.103T>C variant in SDHAF2 gene and n.1482G>A variant in SDHAP2 gene 223

Figure A.101. H142 pedigree and chromatograms showing c.1420A>G variant in the GPATCH2 gene and c.4897-9_4897-8delTT variant in DNAH7 gene.....223

Figure A.102. H142 pedigree and chromatograms showing the c.1999G>A variant in PLCL1 gene and the c.169-5T>G variant in DGKB gene.....224

Figure A.103. H142 pedigree and chromatograms showing the c.1286A>G variant in SEMA3C gene and the c.742-7T>C variant in KIAA1324L gene.....224

Figure A.104. H142 pedigree and chromatograms showing the c.812-8T>C variant in PRPF4 gene and the c.820T>C variant in UBQLNL gene.....225

Figure A.105. H142 pedigree and chromatograms showing the c.887dupC variant in UNC93B1 gene and the c.290G>T variant in C11orf24 gene.....225

Figure A.106. H142 pedigree and chromatograms showing the c.665C>T variant in the c.706G>A variant.....226

Figure A.107. H142 pedigree and chromatograms showing the c.*15+6T>A variant in TCL1B gene.....226

Figure A.108. H143 pedigree and chromatograms showing the c.767A>C, c.883-3C>T and c.1261A>C variants in KIF1A gene 226

Figure A.109. H143 pedigree and chromatograms showing the c.3904_3907delTCTA variant in SPG11 gene 227

Figure B.1. 3D models for SPG11 protein predicted by SWISS-MODEL algorithm.....228

Figure B.2. 3D models for SPG11 protein predicted by RaptorX algorithm 228

Figure B.3. 3D models for SACS protein predicted by SWISS-MODEL algorithm 229

Figure B.4. 3D models for ATAD1 protein predicted by iTASSER algorithm 229

Figure B.5. Conservation of hydrophobic amino acids in the very same position of p.234Met in ATAD1 gene. 230

Figure B.6. 3D models for KIF1C protein predicted by SWISS-MODEL algorithm 230

Figure B.7. 3D models for PLA2G6 protein predicted by SWISS-MODEL algorithm....231

Figure B.8. 3D models for SAMHD1 protein predicted by iTASSER algorithm 231

Figure B.9. 3D models for ALS2 protein predicted by SWISS-MODEL algorithm.....232

Figure B.10. 3D models for SPG15 protein predicted by Raptor-X algorithm 232

Figure B.11. 3D models for SPG11 protein predicted by SWISS-MODEL algorithm.....232

Figure B.12. 3D models for CYP7B1 protein predicted by SWISS-MODEL algorithm..233

Figure B.13. 3D models for SEMA3C protein predicted by SWISS-MODEL algorithm.233

Figure B.14. 3D models for SPG11 protein predicted by Raptor-X algorithm 234

Figure C.1. Western blot analyses of ATAD1 protein performed for H28.1 and H28.2 (Replicate #2). 235

Figure C.2. Western blot analyses of ATAD1 protein performed for H28.1 and H28.2 (Replicate #3). 235

Figure C.3. Western blot analyses of CYP27A1 protein for family H55 (Replicate #2). .235

Figure C.4. Western blot analyses of SAMHD1 protein for family H57 (Replicate #2). .236

Figure C.5. Western blot analyses of SAMHD1 protein for family H57 (Replicate #3). .236

LIST OF TABLES

Table 1.1. Classification of HSP based on intracellular pathophysiological mechanisms 4

Table 1.2. Autosomal dominant forms of HSP 14

Table 1.3. Autosomal recessive forms of HSP 16

Table 1.4. X-linked forms of HSP 21

Table 1.5. Other types of HSP 22

Table 3.1. Clinical findings for the families included in this study. 28

Table 3.2. Sequences of the primers used for exon amplification of *Spastin.* 30

Table 3.3. Sequences of primers used for candidate variants 31

Table 3.4. Sequences of primers for variants identified after homozygosity mapping 34

Table 3.5. Sequences of primers for H28 for a possible XL inheritance. 41

Table 3.6. Primers used for RT-PCR experiments 42

Table 3.7. Primary antibodies used in this study 43

Table 3.8. Secondary antibodies used in this study 43

Table 3.9. Cell lines and mediums used in this study 44

Table 3.10. Buffers used in this study 44

Table 3.11. Solutions used in this study. 46

Table 3.12. Gels used in this study 47

Table 3.13. All solid and liquid chemicals and disposable materials used in this study. 48

Table 3.14. All equipment used in this study.....50

Table 4.1. PLINK parameters used in WES analysis 58

Table 5.1. Candidate variants identified in the WES data of patient P392.....71

Table 5.2. The eliminated splice site variants identified in the WES data of P392.....72

Table 5.3. Sanger sequencing results for family P392.....72

Table 5.4. Compound heterozygous variants in ZFYVE26 gene for patient P463. 73

Table 5.5. Candidate variants for patient P627.....75

Table 5.6. Candidate variants in neurological disorders for patient H6.1. 76

Table 5.7. Candidate variants for patient H6.1 after stringent filtering.....77

Table 5.8. Candidate variants identified in the WES data of H6.1.....78

Table 5.9. Sanger sequencing results for family H6.....79

Table 5.10. Compound heterozygous variants in C19orf12 gene for patient H16. 79

Table 5.11. Candidate variants identified in the WES data of H28.1.....81

Table 5.12. Sanger sequencing results for family H28.....82

Table 5.13. Candidate variants identified in X-linked genes in WES data of H28.1. 84

Table 5.14. Sanger sequencing results for family H28 in X-linked genes. 84

Table 5.15. Candidate variants identified in the WES data of H29.1. 86

Table 5.16. Sanger sequencing results for family H29. 87

Table 5.17. Candidate variants for patient H38.3. 88

Table 5.18. Candidate variants identified in the WES data of H38.3. 89

Table 5.19. Sanger sequencing results for family H38. 90

Table 5.20. Candidate variants for patient H44.7. 91

Table 5.21. Candidate variant identified in the WES data of H44.7. 93

Table 5.22. Candidate variants identified in SPG11 gene for patient H45.3. 93

Table 5.23. Sanger sequencing results in H45 family. 94

Table 5.24. Candidate variants identified in the WES data of H45.3. 95

Table 5.25. Sanger sequencing results for family H45. 96

Table 5.26. Candidate variant identified in SHMT1 gene. 99

Table 5.27. Candidate variants identified in the WES data of H55.3. 99

Table 5.28. Sanger sequencing results for family H55. 100

Table 5.29. Candidate variants determined for H55 for possible AD inheritance. 103

Table 5.30. Sanger sequencing results for H55 for possible AD inheritance. 103

Table 5.31. Heterozygous variants observed in the AR-HSP genes for patient H57.3. 104

Table 5.32. Homozygous variants in WES data of patient H57.3 105

Table 5.33. Candidate variants analyzed by direct sequencing in H57 family. 106

Table 5.34. Candidate variants analyzed by direct sequencing in family H65. 108

Table 5.35. Candidate variants identified in the WES data of H65.3. 109

Table 5.36. Sanger sequencing results for family H65. 109

Table 5.37. Candidate variants identified in the WES data of H82.3. 113

Table 5.38. Sanger sequencing results for family H82. 113

Table 5.39. Compound heterozygous variants in ZFYVE26 gene for patient H99.3. 115

Table 5.40. Sanger sequencing results in H99 family for variants in ZFYVE26 gene. 115

Table 5.41. Candidate variants identified in the WES data of H99.3. 116

Table 5.42. Sanger sequencing results for family H99. 116

Table 5.43. Candidate variants identified in the WES data of H108.3. 118

Table 5.44. Sanger sequencing results for family H108. 119

Table 5.45. Candidate variants identified in the WES data of H6.1. 120

Table 5.46. Candidate variants identified in the WES data of H133. 122

Table 5.47. Candidate variants in neurological disease-genes for patient H142. 123

Table 5.48. Candidate variants identified in the WES data of H142.3. 125

Table 5.49. Sanger sequencing results for family H142. 125

Table 5.50. Compound heterozygous variants in KIF1A gene for patient H143.3. 128

Table 5.51. Sanger sequencing results for H143 family for the variants in KIF1A gene..128

Table 5.52. Frequencies of variants in NDAL-WES database. 129

Table 5.53. The variants identified as causative in the families. 130

Table 5.54. List of number of variants identified in 'Effect of Rare Variant' Analysis....132

Table 5.55. List of variants identified after LRT, SIFT and MutationTaster predictions..133

Table 5.56. Prediction algorithm scores for HPMV and rare variants in patient H17.1....136

Table 5.57. Prediction algorithm scores for HPMV and rare variants in patient H31.1....136

Table 5.58. Prediction algorithm scores for HPMV and rare variants in patient H32.2....136

Table 5.59. Prediction algorithm scores for HPMV and rare variants in patient H52.5....137

Table 5.60. Prediction algorithm scores for HPMV and rare variants in patient H53.3....137

Table 5.61. Prediction algorithm scores for HPMV and rare variants in patient H59.3....138

Table 5.62. Prediction algorithm scores for HPMV and rare variants in patient H61.3....138

Table 7.1. Five novel HSP genes identified in this study.. 159

LIST OF ACRONYMS/ABBREVIATIONS

AAA ATPase	ATPases associated with diverse cellular activities
AD	Autosomal dominant
AD-HSP	Autosomal dominant Hereditary Spastic Paraplegia
ALS	Amyotrophic Lateral Sclerosis
APS	Ammonium Persulfate
AR	Autosomal recessive
AR-HSP	Autosomal recessive Hereditary Spastic Paraplegia
AR-HSP-TCC	AR-HSP with thin corpus callosum
ARSACS	Autosomal Recessive Spastic Ataxia of Charlevoix-Saguenay
ATP	Adenosine Triphosphate
BCA	Bicinchoninic Acid Assay
bp	Base Pair
BPB	Bromophenol Blue
BSA	Bovine Serum Albumin
CaCl2	Calcium Chloride
cDNA	Complementary DNA
CO_2	Carbon Dioxide
Ct	Cycle Threshold
CTX	Cerebrotendinous Xanthomatosis
DABCO	1,4-diazabicyclo[2.2.2]octane
DAPI	Diaminophenylindolamine
DMEM	Dulbecco's Modified Eagle Medium
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic Acid
EDTA	Ethylenediaminetetraacetic Acid
EGF	Epidermal Growth Factor
ER	Endoplasmic reticulum
ERAD	ER-associated degradation
ESCRT	Endosomal Sorting Complexes Required for Transport

FBS	Fetal Bovine Serum
HCl	Hydrochloric Acid
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic Acid
HSP	Hereditary Spastic Paraplegia
IAHSP	Infantile ascending hereditary spastic paralysis
ICC	Immunocytochemistry
KCl	Potassium chloride
kDa	Kilodalton
LB	Luria Broth
LCLs	Lymphoblastoid cell lines
LD	Lipid droplet
mg	Miligram
MIT domain	Microtubule interacting and trafficking domain
ml	Mililiter
mM	Milimolar
MP	Milk Powder
MVB	Multivesicular body
MW	Molecular Weight
Na_2HPO^4	Sodium Phosphate Dibasic
NaCl	Sodium Chloride
ng	Nanogram
nm	Nanometer
OD	Optical Density
PAGE	Polyacrylamide Gel Electrophoresis
PBS	Phosphate Buffered Saline
PFA	Paraformaldehyde
PVDF	Polyvinylidene Fluoride
qRT-PCR	Quantitative Real-Time Polymerase Chain Reaction
RMSD	root-mean-square deviation
RNA	Ribonucleic Acid
rpm	Rotations Per Minute
SDS	Sodium Dodecyl Sulfate

SDS-PAGE	SDS-Polyacrylamide Gel Electrophoresis
sgRNA	single guide RNA
SPG	Spastic paraplegia gene
TBS-T	Tris Buffered Saline Tween
TCC	Thin corpus callosum
TE	Tris-EDTA
TEMED	Tetramethylethylenediamine
UTR	Untranslated Region
UV	Ultraviolet
WASH	Wiskott-Aldrich syndrome protein and scar homolog
WB	Western Blot
XL-HSP	X-linked Hereditary Spastic Paraplegia

1. INTRODUCTION

Hereditary Spastic Paraplegia (HSP), also called Strümpell-Lorrain syndrome, comprises a group of clinically and genetically heterogeneous disorders (Casari and Rugarli, 2001). The defining features of the disease are lower limb spasticity and progressive weakness (Fink, 2003). HSP is classified as 'pure' or 'complicated' and in pure HSPs in addition to main features, minor vibration sense loss and bladder disturbances are observed in some cases. 'Complicated' HSP can be associated with some additional neurological and non-neurological symptoms including thin corpus callosum (TCC), mental retardation, cognitive decline, dementia, deafness, cerebellar ataxia, retinopathy, epilepsy, dysarthria, ichthyosis, optic atrophy, peripheral neuropathy, retinitis pigmentosa, and cataract (Depienne *et al.,* 2007; Battini *et al.,* 2011; Reid, 1999).

Pathologically, HSP is characterized by retrograde distal axonopathy of the upper motor neurons with longest axons in the corticospinal tract and posterior columns (Crosby and Proukakis, 2002; Salinas *et al*., 2008). The answer to the main question of why the longest neurons in the spinal cord are affected, lies in the functions of the proteins that are shown to be associated with the disease. These proteins have roles in several important cellular processes such as membrane trafficking, axonal transport, cytoskeletal organization, and synaptic transmission (Stevanin *et al.,* 2008; Ozes and Battaloglu, 2011). When these proteins do not function properly, axonal traffic of macromolecules, organelles, and other cargoes are disrupted., which mainly influence the distal locations of the neurons. Other HSP proteins important in mitochondrial function point out the reliance of neurons on the energy utilization for cellular transport. The other main theme highlighted in recent years in HSP is the importance of lipid metabolism due to studies reporting genes involved in lipid metabolism and signaling (Gonzalez *et al.,* 2013a).

A review on HSP genetics in the year 2015 reported a total of 72 spastic paraplegia gene (*SPG*) loci and 59 genes (Klebe *et al.,* 2015). In the recent three years, these numbers increased to 78 loci and 63 genes implicating the fast improvement in HSP genetics research.

According to mode of inheritance, HSP can be classified as autosomal dominant (AD-HSP), autosomal recessive (AR-HSP), and X-linked (XL-HSP). Maternal inheritance might also be observed in some HSP cases (Verny *et al.,* 2011). A recently identified type of HSP, SPG72, was shown to be inherited in AD or AR manner complicating the interpretation of the pedigrees prior to genetic analysis (Esteves *et al*., 2014). Highly variable progression rate and age of onset, which range from infancy to eight decade, further supports the heterogeneity of the disease (Roşulescu *et al.,* 2009). Symptoms are often observed in the early adulthood (20-30 years old) in AD-HSP whereas in other HSP types they emerge in childhood and early adolescence.

AD-HSPs are generally associated with pure HSP and their mutations are the predominant reason in the European populations with a prevalence ranging from 1.27/100,000 in Ireland to 5.5/100,000 in Norway (Finsterer *et al*., 2012). On the other hand, AR-HSPs generally have a complicated phenotype and more frequently observed in consanguineous populations (Blackstone*,* 2012). A study by Ruano *et al*., examined 22 studies with 14,539 patients from 16 countries and calculated the prevalence of both AD-HSP and AR-HSP as 1.8/100,000 (2014). Epidemiological studies vary widely depending on the regions of the study.

Prevalence data show that approximately 20% of the HSP cases do not have a mutation in known HSP genes/loci. The missing heritability may arise from unidentified rare forms of the disease, undiscovered non-genetic causes, or potential misinterpretation of the data from candidate gene approach-studies based on pedigree analysis (Varga *et al*., 2013). The main reason for this misinterpretation is incomplete penetrance observed in some HSP families.

1.1. Autosomal Dominant HSP

Twenty loci have been mapped and thirteen genes have been identified responsible for autosomal dominant HSP (AD-HSP). SPG3A and SPG4 are the most common two types of AD-HSP, responsible for 50% of the mutations identified. Phenotype is pure in majority of these cases (Finsterer *et al*., 2012) (Table 1.2).

1.2. Autosomal Recessive HSP (AR-HSP)

Autosomal recessive HSP (AR-HSP) is generally observed in inbred populations. Patients have a complex phenotype with variable age of onset (Blackstone *et al.,* 2012). Causative genes for 47 loci out of the 53 have been identified for this mode of inheritance.

In approximately 20% of the AR-HSP patients, the disease is associated with *SPG11*. When HSP is accompanied with TCC and mental impairment the frequency of *SPG11* reaches up to 59% (Stevanin *et al.,* 2008). Although SPG15 constitutes only 3% of AR-HSP cases, ARHSP-TCC patients that are negative for *SPG11* mutations have *SPG15* mutations in 28% rate. In pure AR-HSP, 16% of cases can be explained by *SPG5* mutations (Schüle and Schöls, 2011). The frequencies of other AR-HSPs are very low; nineteen of them were reported in single families and another nineteen of them were found in less than five families. Genes and loci associated with AR-HSP are listed in Table 1.3.

1.3. X-Linked HSP (XL-HSP)

X-linked HSP is associated with five loci and disease-causing genes are identified in three of them. The clinical features and the protein products of X-linked HSP loci are summarized in Table 1.4.

1.4. Other HSPs

Beside the genes given in the previous parts, there are genes that are not listed as an SPG gene but the phenotypes were shown to be associated with HSP. These genes and their clinical effects are given in Table 1.4. Mitochondrial inheritance is observed very rarely (de Souza *et al*., 2016).

1.5. Pathophysiology of HSP

As for the functions of the HSP proteins, it is possible to group them into a number of key cellular processes such as membrane trafficking, myelin formation, mitochondrial function and lipid metabolism (Figure 1.1, Table 1.1).

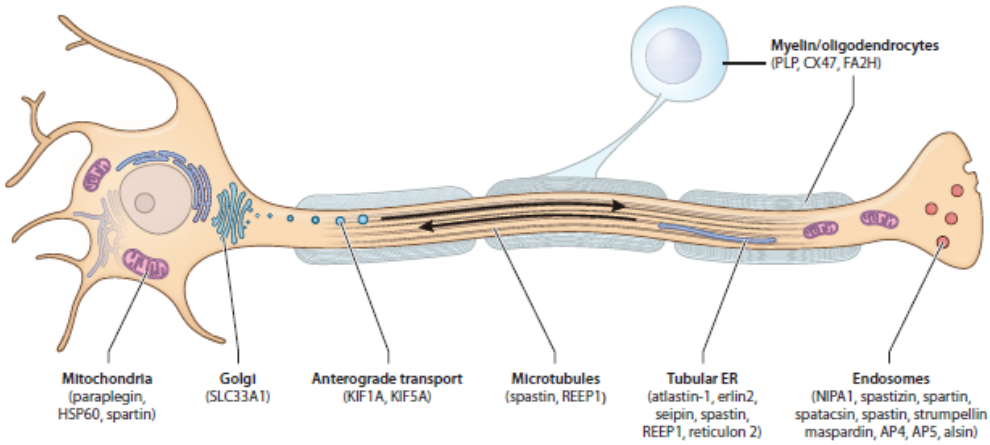


Figure 1.1. Common pathogenic cellular processes in the HSPs (Blackstone, 2012)

Table 1.1. Classification of hereditary spastic paraplegia based on intracellular pathophysiological mechanisms (de Souza *et al*., 2016).

Intracellular pathophysiological mechanisms	SPG Type
Membrane trafficking and organelle shaping	SPG3A, SPG4, SPG6, SPG11, SPG15, SPG18, SPG20, SPG31, SPG59, SPG60, SPG61, SPG62, SPG69, SPG72
Axonal transport	SPG4, SPG10, SPG30, SPG58
Mitochondrial dysfunction	SPG7, SPG20, SPG31
Lipid metabolism disturbances	SPG5, SPG26, SPG28, SPG35, SPG39, SPG46, SPG54, SPG56
Myelination abnormalities	SPG1, SPG2, SPG39, SPG42, SPG67

1.5.1. ER Morphogenesis and membrane trafficking

Endoplasmic reticulum (ER) is a highly crucial organelle for its role in membrane expansion during formation of axons and dendrites (Lynes & Simmen 2011). SPG3A- SPG4 and SPG31, which are the three most frequently mutated genes in AD-HSP, and SPG12, function in the ER network formation (Park *et al*., 2010). *REEP1* (SPG31) which is localized to ER not only interacts with atlastin-1 (SPG3A) but also binds to microtubules via its

hydrophobic and C-terminal domains, respectively (Blackstone *et al*., 2012). *Reticulon 2* (SPG12) is implicated in ER shaping (Montenegro *et al*., 2012).

Interaction with microtubule cytoskeleton is essential for ER motility. *SPG4* encodes for a microtubule-interacting and severing AAA ATPase which binds *Atlastin-1, REEP1* and *Reticulon1* (Connell *et al*., 2009, Park *et al.,* 2010) and this partnership shows a strong link between ER network and microtubule cytoskeleton. *BSCL2 (*Berardinelli-Seip congenital lipodystrophy protein 2) gene, codes for an integral membrane protein of the ER that leads to both SPG17 (Silver syndrome) and Berardinelli-Seip congenital lipodystrophy (Windpassinger *et al.,* 2004). Seipin protein functions in the size regulation of lipid droplets (LDs) (Fei *et al.,* 2011). Misfolded seipin proteins in the SPG17 form aggregates triggering ER stress (Yagi *et al.,* 2011). Another HSP protein, ERLIN2 (SPG18) also localizes to ER and is suggested to be part of ER-associated degradation (ERAD) pathway, which uses ubiquitin-proteosome system to degrade ER proteins (Pearce *et al.,* 2009). Moreover, ERLIN2 binds gp78, a key enzyme in cholesterol metabolism (Jo *et al.,* 2011). Various cellular stress conditions including unfolded protein response pathways can cause LD accumulation (Hapala *et al.,* 2011) that in turn, can affect ER network and function. However, ER has roles in many cellular pathways and it is not known which of them underlies the pathogenesis of HSP.

Spartin (SPG20) functions in LD turnover and EGF (Epidermal Growth Factor) receptor degradation (Bakowska *et al.,* 2007). It interacts with E3 ubiquitin ligases so it uses ubiquitin modification to regulate a number of signaling pathways. Like Spastin, Spartin has a MIT (microtubule interacting and trafficking) domain and interacts with IST1- component of ESCRT-III complex. Cytosolic protein complexes form ESCRT (Endosomal Sorting Complexes Required for Transport) that serves as membrane remodeling system for various cellular processes such as multivesicular body (MVB) biogenesis, cellular abscission, and viral budding (Hurley and Hanson 2010). Spartin participates in cytokinesis (Renvois´e *et al.,* 2010) and interestingly, interaction of Spastin with ESCRT-III subunits CHMP1B and IST1 is also required for cytokinesis (Yang *et al.,* 2008, Connell *et al.,* 2009, Guizetti *et al.,* 2011). These data show us that many of the HSP proteins have common functions; they interact with common complexes indicating a common pathogenic mechanism underlying neurodegeneration. At this point it should be mentioned that *Spartin* (SPG20) mutations are

mainly observed in Amish families and rarely in Omani families (Manzini *et al.,* 2010). This shows us that, a protein with very critical functions like *Spartin* can be identified by analyzing even a single family, regardless of the previous knowledge about the frequency and population of HSP subtypes.

Strumpellin (SPG8) is identified as a subunit of WASH (Wiskott-Aldrich syndrome protein and scar homolog) complex which resides at the surface of endosomes that recruits and activates the Arp2/3 complex to generate branched actin networks (Derivery and Gautreau, 2010). This shows that defects in membrane modeling and cytoskeletal organization may be one of the pathogenic mechanisms leading to HSP. WASH is associated with autosomal recessive intellectual disability indicating a general role for this complex in neurological diseases (Ropers *et al.,* 2011). KIAA0415 (SPG48) interacts with Spatacsin (SPG11) and Spastizin (SPG15), which co-localize in cytoplasmic structures (Slabicki *et al.,* 2010, Murmu *et al.,* 2011). Spastizin is implicated in cytokinesis with ESCRT proteins by binding to the lipid PI(3)P (Sagona *et al.,* 2010). KIAA0415 belongs to an adaptor protein complex, AP-5, which acts in endosomal dynamics (Hirst *et al.,* 2011). Moreover, there are other HSP proteins (SPG47, SPG50-52), which are subunits of another adaptor protein complex AP-4. Amyloid precursor proteins are transported from the *trans*- Golgi to endosomes with the function of this complex (Burgos *et al.,* 2010). These findings point out to a general role for adaptor complexes in neurological diseases as well as HSP (Abou Jamra *et al.,* 2011, Moreno-De-Luca *et al.,* 2011). Alsin protein *(ALS2* gene) is a guanine nucleotide exchange factor (GEF) for the small GTPases and its mutations are associated with Infantile ascending hereditary spastic paralysis (IAHSP) and Juvenile amytrophic lateral sclerosis (JALS). The protein is localized to endosomal compartments and functions as a modulator of endosomal dynamics (Hadano *et al.,* 2007). The genes that are mutated in both HSP and in other neurodegenerative diseases like ALS or Parkinson's disease, indicate a broad range of underlying mechanism leading to neurodegeneration.

KIF5A (kinesin heavy chain 5A) is an ATP-dependent motor protein that perform anterograde transport of cargo molecules along axons and functions in membrane traffic pathways. Its association to SPG10 shows involvement of motor-based transport in HSP pathogenesis (Reid *et al.,* 2002; Goizet *et al.,* 2009).

L1CAM (SPG1) encoding for NCAM (neural cell adhesion molecule) is an axonal glycoprotein that is implicated in neuronal migration and differentiation of and identification of its mutant forms in HSP patients indicates the axon pathfinding as a possible pathogenic mechanism (Bechara *et al.,* 2008). Myelination is another cellular process related to HSP. PLP1 (tetraspan integral membrane proteolipid protein), one of the major components of myelin in CNS, is associated with SPG2. However, in *Plp1* null mice although myelin is not defective, anterograde transport in the underlying axon is affected probably due to a signaling cascade starting from the oligodendrocyte (Edgar *et al.,* 2004; Gruenenfelder *et al.,* 2011). Connexin47 proteins that form gap junction channels have a key role in central myelination and are associated with SPG44. It is suggested that channel dysfunction leads to HSP in the patients (Orthmann-Murphy *et al.,* 2009). Another protein to be included in this context is FA2H (SPG35). It encodes for a nicotinamide adenine dinucleotide phosphate (NADPH)-dependent monooxygenase which has role in the synthesis of 2-hydroxysphingolipids, a subset of sphingolipids. These molecules are part of the myelin galactolipids. Demyelination is observed in *Fa2h* null mice (Potter *et al.,* 2011). Overall, oligodendrocyte-mediated myelination can affect the underlying axon.

1.5.2. Mitochondrial Function

Proper functioning of mitochondria is crucial for neuronal health. When impaired, it can cause various developmental and degenerative neurological disorders (Di-Mauro and Schon, 2008). Paraplegin (SPG7), an *m*-AAA metalloprotease that locates in the inner mitochondrial membrane and HSP60/SPGD1 (SPG13), which is implicated in mitochondrial protein import and macromolecular assembly, are two examples linking mitochondrial function with HSP (Ferreirinha et al. 2004). SPG55 associated-gene *C12ORF65* encodes for a protein that functions in releasing newly synthesized proteins from mitochondrial ribosome (Antonicka *et al.,* 2010). Despite identification of mitochondria-related genes as HSP genes, the exact patho-mechanism leading to HSP is not known. Mitochondrial disorganization and dysfunctions leading to reduction in energy production might be considered as one possible underlying mechanism (Goizet *et al.,* 2011). Another possible patho-mechanism suggested is impairment in ER-mitochondrial membrane contact sites (Krols *et al.,* 2016). The interaction between ER and mitochondria is crucial for various processes including exchange of phospholipids and calcium between two organelles,

mitochondrial dynamics, sterol and energy metabolism. Disruption of the ER-mitochondria membrane contact might affect neuronal health causing neurodegeneration as implied in various neurological diseases including Amyotrophic Lateral Sclerosis (ALS), and Charcot–Marie–Tooth disease (CMT) (Krols *et al.,* 2016; Bernard-Marissal *et al;* 2015; de Brito *et al.,* 2008).

1.5.3. Lipid Synthesis and Metabolism

Lipid distribution and its metabolism is another key function of ER and there are several proteins associated with HSP implicating involvement of lipid metabolism. Role of Spartin (SPG20) and Seipin (SPG17) in lipid droplet (LD) turnover is investigated, however, effects of LDs in axons are not known in detail. Essential cellular processes such as organelle shaping, lipid distribution, or signaling pathways are suggested to be affected in axons (Blacktone, 2012). *SLC33A1* (SPG42) gene encodes for the acetyl-CoA transporter which functions in monitoring carbohydrate and fat levels. Since acetyl-CoA enters the citric acid cycle, it is important for the energy metabolism of the cell.

Defects in metabolism of phospholipids and fatty acids, that functions as precursor of various lipid messengers, can have an impact on the mitochondrial function and can alter the mitochondrial-membrane lipid composition. These two possible mechanisms can finally affect the mitochondrial respiration and energy production in cell (Gonzalez *et al.,* 2013a; Tesson *et al.,* 2012). Another hypothesis associating HSP to lipid metabolism could be the accumulation of one or more products in the cell. One of the genes which belong to this category is SPG5 associated-protein CYP7B1 (cytochrome P-450 oxysterol 7 α-hydroxylase). This protein is mainly expressed in extra-hepatic tissues, including the brain, and affects dehydroepiandrosterone (DHEA) neurosteroids and oxysterols metabolism (Wu *et al.,* 1999; Rose *et al.,* 1997). Increased 27-hydroxycholesterol levels in plasma and cerebrospinal fluid is observed in *CYP7B1*-mutated patients indicating that aberrant oxysterol levels might be the underlying pathogenic mechanism in SPG5-associated HSP (Schüle *et al.,* 2009; Tsaousidou *et al.,* 2008). NTE/ SPG39 (neuropathy target esterase) is an ER integral membrane protein and it deacylates phosphatidylcholine, which is an important membrane phospholipid. When the protein is mutated, membrane formation is affected leading to distal axonal degeneration (Rainier *et al.*, 2008, Read *et al.,* 2009).

As a summary, HSP proteins act not only in axonal and membrane trafficking but also in a wide spectrum of cellular processes including mitochondrial function and lipid metabolism. Regarding the importance and complexity of these processes, there should be more genes waiting to be discovered to complete the puzzle of HSP pathogenicity.

Table 1.2. Autosomal dominant forms of HSP (Finsterer *et al.,* 2012).

Locus	**Chromosome region**	**Gene (Protein)**	**Function of the protein**	**Frequency of mutations**	**Clinical features**
SPG3A	14q12-q21	SPG3A (atlastin-1)	GTPase, ER to Golgi transfer, spastin partner	10% of ADHSP	Early-onset pure, slow progression HSP
SPG4	2p22	SPG4 (spastin)	Microtubule-severing activity, early secretory pathway	40% of pure AD HSP	Variable-onset mainly pure HSP
SPG6	15q11.1	NIPA1	Mg2+ transporter, endosomal trafficking	<1% (9 families)	Adult-onset pure HSP
SPG8	8q24	KIAA0196 (strumpellin)	Spectrin domain protein	8 families	Adult-onset pure HSP, marked spasticity
SPG9	10q23.3-q24.2	Unknown	-	1 family	Cataracts, motor neuropathy, skeletal abnormalities, gastrooesophageal reflux
SPG10	12q13	KIF5A	Kinesin heavy chain motor protein	4-10 % of SPG4 neg. AD-HSP	Early-onset pure HSP, can be complicated with distal amyotrophy
SPG12	19q13	RTN2 (Reticulon2)	Participates in the network of hairpin loop-containing ER morphogens, including REEP1, ATL1 and spastin.	3 families	Pure HSP
SPG13	2q24 –q34	SPGD1 (Mt. HSP60)	Implicated in mitochondrial protein import and macromolecular assembly.	2 families	Pure and complicated HSP – dystonia

Table 1.2. Autosomal dominant forms of HSP (Finsterer *et al.,* 2012) (cont.).

Locus	**Chromosome region**	**Gene (Protein)**	**Function of the protein**	**Frequency of mutations**	**Clinical features**
SPG17	11q12–q14	BSCL2	regulator of lipid catabolism	12 families	Silver-syndrome, allelic to CMT2
SPG19	9q33–q34	unknown	-	1 family	Late onset pure HSP
SPG29	1p31.1–21.1	unknown	-	1 family	Hearing impairment, persistent vomiting from hiatal hernia
SPG31	9p21.2-q21.12	REEP1	Endosomal trafficking, mitochondrial chaperone	8% of SPG4 neg. AD-HSP	pure or complicated HSP with Axonal neuropathy, Silver-like-syndrome, cerebellar ataxia, tremor, dementia
SPG33	10q24.2	ZFYVE27 (Protrudin)	Intracellular trafficking	1 family	Complicated HSP with foot deformities
SPG36	12q23–24	Unknown	-	1 family	Complicated HSP with polyneuropathy
SPG37	8p21.1–q13.3	Unknown	-	1 family	Pure HSP
SPG38	4p16–p15	Unknown	-	2 families	Silver-syndrome
SPG40	Unidentified	-	-	1 family	Pure or complicated HSP with hyperrreflexia of upper limbs, memory deficits
SPG41	11p14.1 – p11.2	Unknown	-	1 family	Mild weakness of small hand muscles
SPG42	3q24–q26	SLC33A1 (AcoA-carrier)	Acetyl-CoA transporter	1 family	Pure HSP
SPG73	19q13.33	CPT1C	Beta-oxidation and transport of long-chain fatty acids into mitochondria	1 family	Mild muscle atrophy, hyperreflexia with extensor plantar responses

Table 1.3. Autosomal recessive forms of HSP (Schüle and Schlös, 2011; Finsterer *et al.,* 2012).

Locus	**Chromosome region**	**Gene (Protein)**	**Function of the protein**	**Frequency**	**Clinical features**
SPG5A	8q12 – q13	*CYP7B1* (OAH1)	Cholesterol and neurosteroid metabolism	16% pure AR-HSP	Pure or complicated HSP – WMLs
SPG7	16q24.3	*SPG7* (Paraplegin)	Mitochondrial ATPase	1,5-7% of AR-HSP (20 families)	Pure or complicated HSP – Optic or cerebellar atrophy, abnormal mitochondria, dysarthria, dysphagia, PNP, TCC
SPG11	15q13–q15	*KIAA1840* (Spatacsin)	Accessory protein of the AP5 complex of sorting of vesicles	~20% of AR-HSP	Pure or complicated HSP – Mental retardation, sPNP, TCC, dysarthria, nystagmus, Ulweakness, ataxia, parkinsonism, maculopathy, WMLs, allelic to ALS-5 and Kjellin syndrome
SPG14	3q27–q28	-	-	1 family	Pure or complicated HSP – Mental retardation, mPNP
SPG15	14q22–q24	*ZFYVE26* (Spastizin)	Accessory protein of the AP5 complex of sorting of vesicles, putatively involved in endosomal trafficking	3% AR-HSP, 28% SPG11 neg ARHSP-TCC	Complicated – Maculopathy, PNP, dysarthria, mental retardation, pigmental retinopathy, TCC, WMLs, allelic to Kjellin syndrome
SPG18	8p12–8q11.22	*ERLIN2* (SPFH2)	Lipid 16ulphate16l (IP3 dgeradation)	2 families	Complicated HSP – Intellectual decline, epilepsy, congential hip dislocation
SPG20	13q12.3	*SPG20* (Spartin)	Involved in degradation of epidermal growth factor receptor, associated with mitochondria	>25 families	Complicated HSP – Troyer-syndrom (dysarthria, distal wasting)

Table 1.3. Autosomal recessive forms of HSP (Schüle and Schlös, 2011; Finsterer *et al.,* 2012) (cont.).

Locus	**Chromosome region**	**Gene (Protein)**	**Function of the protein**	**Frequency**	**Clinical features**
SPG21	15q22.31	*ACP33* (Maspardin)	Protein transport and sorting in trans-Golgi transportation vesicles	2 families	Pure and complicated HSP – Dementia, cerebellar involvement, extrapyramidal signs, TSS,WMLs
SPG23	1q24–q32	-	-	1 family	Complicated HSP – Pigmentary skin lesions
SPG24	13q14	-	-	1 family	Complicated HSP – Variably associated with spastic dysarthria and pseudobulbar signs
SPG25	6q23–q24	-	-	1 family	Complicated HSP – Disc herniation
SPG26	12p11.1 – q14	B4GALNT1	Involved in the biosynthesis of gangliosides GM2, GD2 and GA2	2 families	Complicated HSP – Dysarthria, distal amyotrophy in upper and lower limbs, intellectual impairment
SPG27	10q22.1 – q24	-	-	2 families	Complicated HSP –Dysarthria, PNP, mental retardation
SPG28	14q21.3 – q22.3	*DDHD1* (DDHD1)	Lipid metabolism	2 families	Pure HSP
SPG30	2q37	*KIF1A* (Kinesin3)	Involved in axonal transport and trafficking	2 families	Pure and complicated HSP – Mild cerebellar signs, sensory neuropathy, allelic to HSAN
SPG32	14q12–q21	-		1 family	Complicated HSP – Mental retardation, brainstem 17ulphate17l, cerebellar atrophy
SPG35	16q21–q23	*FA2H* (FA2H)	Involved in lipid hydroxylation	6 families	Pure and complicated HSP – Cognitive decline, seizures, WMLs, allelic to NBIA
SPG39	19p13.2	*NTE* (Esterase)	phospholipase/ lysophospholipase, Lipid metabolism	2 families	Complicated HSP – Axonal neuropathy, spinal cord atrophy

Table 1.3. Autosomal recessive forms of HSP (Schüle and Schlös, 2011; Finsterer *et al.,* 2012) (cont.).

Locus	**Chromosome region**	**Gene (Protein)**	**Function of the protein**	**Frequency**	**Clinical features**
SPG43	19p13.11–q12	C19ORF12	Fatty acid metabolism	2 families	Complicated HSP – Silver-syndrome, allelic to NBIA
SPG44	1q42.13	*GJA12/GJC2* (Connexin47)	Cell junction protein, 18ulphate18l18on	1 family	Complicated HSP – WMLs, hypomyelination, allelic to Pelizaeus–Merzbacher disease
*SPG45	10q24.3 – q25.1	NT5C2	Purine/pyrimidine nucleotide metabolism	1 family	Complicated HSP – Mental retardation, ocular abnormalities
SPG46	9p21.2–q21.12	GBA2	Lipid metabolism	5 families	Complicated HSP –Mental retardation, cataract, cerebellar dysfunction, TCC
SPG47	1p13.2–1p12	*AP4B1, E1, M1, S1* (AP4B1, E1, M1, S1)	Involved in forming and sorting of vesicules	6 families	Complicated HSP – Mental retardation, seizures, TCC, periventricular WMLs
SPG48	7p22.2	*KIAA0415* (AP5Z1 protein)	Member of the AP5 complex involved in vesicles sorting and putatively in DNA-repair	1 family	Pure and Complicated HSP – Spinal cord hyperintensities
SPG49	14q32.31	*TECPR2* (TECPR2)	Unknown	3 families	Complicated HSP- delayed psychomotor development, mental retardation, dysmorphic features, TCC
SPG53	8p22	*VPS37A* (HCRP1)	Endosomal trafficking	2 families	Complicated HSP-kyphosis, pectus carinatum, cognitive and speech delay

Table 1.3. Autosomal recessive forms of HSP (Schüle and Schlös, 2011; Finsterer *et al.,* 2012) (cont.).

Locus	Chromosome region	Gene (Protein)	Function of the protein	Frequency	Clinical features
SPG54	8p11.23	*DDHD2* (DDHD2)	Membrane trafficking and lipid metabolism	4 families	Complicated HSP – TCC, mental retardation, periventricular WMLs
SPG55	12q24.31	*C12ORF65* (C12ORF65)	Mitochondrial translation	1 family	Complicated HSP – 19ulph atrophy, hyperreflexia, atrophy
SPG56	4q25	*CYP2U1* (CYP2U1)	Lipid metabolism and mitochondrial function	5 families	Complicated HSP – hyperreflexia, mental retardation
SPG57	3q12.2	TFG	ER microtubular architecture and function	1 family	Earlı onset, complicated HSP-optic atrophy, axonal demyelinating sensorimotor neuropathy
SPG58	17p13.2	KIF1C	retrograde transport of Golgi vesicles to the endoplasmic reticulum	5 families	Complicated HSP – spastic ataxia
SPG59	15q21.2	USP8	Regulates endosomal ubiquitin dynamics, cargo sorting, membrane traffic, and maintenance of ESCRT-0 stability	1 family	Complicated HSP – hypertonicity
SPG60	3p22.2	*WDR48* (AUF1)	Implicated in DNA damage repair and translesional synthesis pathway	1 family	Complicated HSP – hypertonicity, nystagmus, 19ulphate19l neurpath
SPG61	16p12.3	ARL6IP1	Protein transport	1 family	Complicated HSP – polysensory and motor neuropathy
SPG62	10q24.31	ERLIN1	ER-associated degradation	3 families	Pure HSP

Table 1.3. Autosomal recessive forms of HSP (Schüle and Schlös, 2011; Finsterer *et al.,* 2012) (cont.).

Locus	Chromosome region	Gene (Protein)	Function of the protein	Frequency	Clinical features
SPG63	1p13.3	AMPD2	Purine nucleotide metabolism	1 family	Complicated HSP – white matter changes in the corpus callosum
SPG64	10q24.1	ENTPD1	Regulation of purinergic transmission	2 families	Complicated HSP – borderline IQ, aggressiveness, and microcephaly
*SPG65	10q24.32-q24.33	NT5C2	Purine/pyrimidine nucleotide metabolism	5 families	Complicated HSP – abnormal gait, corpus callosum hypoplasia, and learning disability
SPG66	5q32	ARSI	Hydrolyze 20ulphate esters, Hormone biosynthesis	1 family	Complicated HSP – corpus callosum and cerebellar hypoplasia with colpocephaly
SPG67	2q33.1	PGAP1	GPI biosynthesis	1 family	Complicated HSP – corpus callosum, cerebellar vermis hypoplasia, and defective myelination
SPG68	11q12-q13	FLRT1	Cell adhesion and receptor signaling	1 family	Complicated HSP
SPG69	1q41	RAB3GAP2	Exocytosis of neurotransmitters and hormones	1 family	Complicated HSP
SPG70	12q13.3	MARS	Cytosolic methionyl-tRNA synthetase	1 family	Complicated HSP – deep tendon reflexes and bilateral Achilles contractures, borderline IQ
SPG71	5p13.3	ZFR	RNA localization?	1 family	Complicated HSP – TCC
SPG72	5q31.2	REEP2	ER-shaping proteins	2 families	Pure HSP
SPG74	1q42	IBA57	Iron-Sulfur Cluster Assembly Factor For Biotin Synthase- And Aconitase-Like Mitochondrial Proteins	1 family	Complicated HSP – optic atrophy and peripheral neuropathy

Table 1.3. Autosomal recessive forms of HSP (Schüle and Schlös, 2011; Finsterer *et al.,* 2012) (cont.).

Locus	Chromosome region	Gene (Protein)	Function of the protein	Frequency	Clinical features
SPG75	19q13	MAG	Myelination	2 families	Complicated HSP – cerebellar signs, nystagmus, and clonus
SPG76	11q13.1	CAPN1	Neruon migratin, and possible involvement in microtubule network	3 families	Complicated HSP - Hyperreflexia, dysarthria, ataxia
SPG77	6p25.1	FARS2	Phenylalanyl-TRNA Synthetase 2, Mitochondrial	1 family	Pure HSP
SPG78	1p36.13	ATP13A2	Transport inorganic cations and other substrates across cell membranes	4 families	Complicated HSP
SPG79	4p13	UCHL1	Hydrolyze small C-terminal adducts of ubiquitin to ubiquitin monomers	2 families	Complicated HSP
*NT5C2 gene is located in the SPG45 region however the Navorino et al., (2014) named the HSP subtype as SPG65. The gene is shown to be related to SPG45 in OMIM database.					

Table 1.4. X-linked forms of HSP (Finsterer *et al.,* 2012).

Locus	Chr. region	Gene (Protein)	Function of the protein	Frequency	Clinical features
SPG1	Xq28	L1CAM (NCAM)	Cell adhesion, neurite outgrowth, myelination	>100 family	Mental retardation, adducted thumbs, hydrocephalus (MASA-syndrome), TCC

Table 1.4. X-linked forms of HSP (Finsterer *et al.,* 2012) (cont.).

Locus	Chromosome region	Gene (Protein)	Function of the protein	Frequency	Clinical features
SPG2	Xq21	PLP1 (MPLP)	Primary constituent of myelin	<100 family	Quadriplegia, nystagmus, mental retardation, seizures, allelic to Pelizaeus-Merzbacher
SPG16	Xq11.2	Unknown	-	3 families	HSP with onset in infancy, aphasia, sphincter disturbance, mental retardation
SPG22	Xq21	*SLC16A2* (MCT8)	-	10 families	Pure or complex HSP - Allan Herndon Dudley syndrome
SPG34	Xq25–27	Unknown	-	1 family	Pure HSP

Table 1.5. Other types of HSP.

HSP Type / Associated Gene (Protein)	Chromosome region	Inheritance	Function of the protein	Frequency	Clinical features
IAHSP / ALS2	2q33.1	AR	Intracellular endosomal trafficking	14 families	Ascending phenotype, marked pyramidal hypersignal
SPOAN / KLC2	11q13.2	AR	Anterograde axoplasmatic transport of organelles and macromolecular cargoes	>70 families	Brazil, Egypt; nystagmus, optic atrophy, hyperhidrosis, marked acoustic startle reflex, distal amyotrophy, neuropathy; Parkinsonism; mild spine atrophy

Table 1.5. Other types of HSP (cont.).

HSP Type / Associated Gene (Protein)	**Chromosome region**	**Inheritance**	**Function of the protein**	**Frequency**	**Clinical features**
CPSQ-I / GAD1	2q31.1	AR	Conversion of glutamic acid to (GABA)	1 family	Mental retardation, occasional seizures, microcephaly, multiple contractures
Spastic paraplegia with deafness		XL		1 family	Tremor, cataracts, deafness, short stature, hypogonadism
BICD2-associated HSP	9q22.31	AR	Involved in dynein-mediated transport	1 family	Amyotrophy
CCT5 -associated HSP	5p15.2	AR	Involved in proper folding of cytoskeletal proteins	1 family	Cavanagh variant, severe neuropathy, mutilating acropathy, vagal hyperactivity, mild spine atrophy
FAM134B-associated HSP	5p15.1	AR	cis-Golgi transmembrane protein	1 family	Painless neuropathy, mutilating acropathy, hyperhidrosis
EXOSC3-associated HSP	9p13.2	AR	A core component of the human RNA exosome complex	1 family	Short stature, mental retardation, cerebellar ataxia, strabismus, distal amyotrophy, tongue atrophy, adducted thumbs, cerebellar atrophy, enlarged cisterna magna
LYST-associated HSP	1q42.3	AR	Intracellular protein trafficking in endosomes	1 family	Cerebellar ataxia, neuropathy, thoracic spine and cerebellar atrophy

Table 1.5. Other types of HSP (cont.).

HSP Type / Associated Gene (Protein)	**Chromosome region**	**Inheritance**	**Function of the protein**	**Frequency**	**Clinical features**
GRID2-associated HSP	4q22.1-q22.2)	AR	Ionotropic glutamate receptor family of excitatory neurotransmitter receptors	1 family	Cerebellar ataxia, frontotemporal dementia, lower motor neuron syndrome, ALS-like phenotype, cerebellar atrophy
IFIH1-related HSP	2q24.2	AD	Viral RNA receptor that activates type I interferon signaling	2 families	British families, multisystem inflammatory disorders
ADAR1-related HSP	1q21.3	AD	Responsible for RNA editing	1 family	Pure HSP, high interferon-1 levels
RNASEH2B-related HSP	13q14.3	AD	Digests RNAs from RNA:DNA duplexes	2 families	Pure HSP, unspecific white matter changes
KLC4-associated HSP	6p21.1	AR	Organel transport	1 family	Joint contractures, marked pyramidal tract hypersignal, thin corpus callosum, white matter changes, cerebellar atrophy, dentate nucleus hypersignal
PMCA4-associated HSP	1q32.1	AD	Plasma Membrane Calcium ATPase	2 families	Chinese families, pure HSP
TUBB4A-related HSP	19p13.3	AR	Microtubule formation	1 family	Cerebellar ataxia, hypomyelinating leukodystrophy
DNM2-associated HSP	19p13.2	AD	Endocytosis and cell motility	1 family	Pure HSP

Table 1.5. Other types of HSP (cont.).

HSP Type / Associated Gene (Protein)	**Chromosome region**	**Inheritance**	**Function of the protein**	**Frequency**	**Clinical features**
MT-CO3 gene mutations		Mitochondrial	Respiratory Complex IV	1 family	Mental retardation, ophthalmoplegia, severe lactic acidosis, Leigh-like features, COX deficiency
MT-TI gene mutations		Mitochondrial	Mitochondrial tRNA for isoleucine	1 family	Cerebellar ataxia, mental retardation, chronic progressive external ophthalmoplegia; normal muscle biopsy; cardiomyopathy, hearing loss, diabetes
MT-ND4 gene mutations		Mitochondrial	Respiratory complex I	1 family	Visual loss (Leber-like)
MT-ATP6 gene mutations		Mitochondrial	Complex V, mitochondrial ATP synthase	1 family	Neuropathy, normal lactate, normal muscle biopsy

2. PURPOSE

HSP is a highly heterogeneous disease not only clinically but also genetically. Parallel to the technological advancements in genetic research, the number of HSP genes increased remarkably in the last five years. Majority of the recent studies focus on the AR-HSP cases and make use of a combination of whole exome sequencing (WES) and homozygosity mapping approaches. In these studies, 24 novel HSP genes were reported and increased the total number of HSP genes to 63. Despite to this high number of known HSP genes, approximately 50% of the cases cannot be explained. This high ratio suggest possibility of many genes to be discovered.

Dominant and sporadic HSP cases are observed in a higher frequency in Western populations. However, the frequency of recessive cases exceeds the frequency of the dominant cases in the populations with higher consanguineous marriage rates. In our laboratory cohort of 178 HSP cases, 63,5% of the cases had familial HSP and 43,8% of the cases had an AR inheritance according to pedigree analyses. Power of the combination of WES and homozygosity mapping methods and higher frequency of AR cases in our cohort directed us to genetically analyze these families.

The main aim of this study is to identify novel causative genes in AR-HSP cases by using WES and homozygosity mapping. The identification of novel genes might help us to explain the underlying genetic reason for the unknown cases. The ultimate goal is to unravel further genetic heterogeneity that will eventually lead to understanding of the molecular mechanisms underlying HSP. In a wider perspective, we aim to add new building blocks to the road leading to the cure of the disease.

High clinical heterogeneity and symptoms overlapping with other neurological disorders complicate clinical diagnosis of HSP. Another benefit of identification of more HSP genes might be providing differential diagnosis to a higher number of patients. It is known that use of disease-specific genetic diagnostic panels is on rise and as the number of

genes responsible for HSP increase, better HSP-specific diagnostic panels might be developed in near future.

This project also aims to specify the genes of implicated HSP loci for which the causative gene has not been determined, yet. We also aim to identify novel mutations in novel or known HSP genes to better characterize the mutational hotspots in the genes leading to HSP. The highly mutated domains of the proteins might give clues about the effect of domains on protein functionality and interactions.

Another aim of the study is to understand the effect identified mutations via *in vitro* analyses. To reveal how mutation affected the expression level and cellular localization of the protein, immortalized B-lymphocytes of the patients were analyzed by RT-PCR, Western-blot and immunocytochemical analyses. Observed differences might help us to shed light on the function of the protein and role of the identified protein on neurodegeneration.

As possessing some symptoms and genes that are common to other neurological disorders, it might be suggested that HSP occupies a noteworthy position as a good model to understand the pathomechanism underlying neurodegeneration.

3. MATERIALS

3.1. Subjects

27 autosomal recessive HSP families included in the study (Table 3.1). Peripheral blood samples of HSP patients and their family members were provided by the following neurology departments: Department of Neurology, Istanbul Medical School and Cerrahpasa Medical School, Istanbul University; Department of Neurology, Pamukkale University; Department of Neurology, Marmara University; Department of Neurology, Gazi University; Department of Pediatric Neurology, Çukurova University; Clinics of Neurology, Izmir Tepecik Research and Training Hospital; Department of Neurology, Bakirkoy Training and Research Hospital for Psychiatry and Neurological Diseases; Metin Sabancı Baltalimanı Research and Training Hospital; Medipol Hospital; Acıbadem Hospital; Department of Neurology, Ondokuz Mayıs University; and Clinics of Neurology, Izmir Bozyaka Research and Training Hospital. Informed consent was obtained from all family members and the study was approved by Ethics Review Committees of participating institution.

Table 3.1. Clinical findings for the families included in this study.

Family ID	Number of affected individuals	Age of onset	HSP type	Clinical Findings
P392	2	>40	Complicated	Polyneuropathy
P463	2	19	Complicated	Pes cavus, mental retardation, atrophy in carpus callosum, cortical atrophy
P627	2	21-40	Complicated	Peripheral neuropathy, vibration sense loss
H6	2	11	Pure	No additional finding
H16	3	9	Complicated	TCC, cerebellar findings, hearing loss
H28	2	>20	Pure	No additional finding
H29	2	12	Complicated	TCC, cerebellar findings, hyperreflexia, pigmentation abnormalities
H38	3	>40	Complicated	TCC, WMLs

Table 3.1. Clinical findings for the families included in this study (cont.).

Family ID	Number of affected individuals	Age of onset	HSP type	Clinical Findings
H44	3	2	Pure	Pes cavus, cerebellar findings, extrapyrimidal signs
H45	3	25	Complicated	Peripheral neuropathy, mild demantia, hearing loss
H52	2	17	Pure	No additional finding
H53	2	18	Pure	pes cavus, dysarthria
H55	3	32	Complicated	Pes cavus, claw toe, kyphosis, scoliosis, epileptic seizures
H57	2	2	Pure	No additional finding
H59	3	<5	Complicated	Not mentioned
H61	1	1	Complicated	Peripheral neuropathy, TCC, mental retardation, mild vibration loss at lower extremities
H65	8	40	Complicated	TCC, sphincter dysfunction,distal amyotrophy
H72	2	36	Pure	-
H77	3	20	Complicated	TCC, dysarthria
H82	2	12	Complicated	Demantia and mental retardation
H98	2	3	Pure	-
H99	8	45	Pure	Sphincter disturbance
H108	3	31	Pure	-
H110	2	30	Pure	-
H133	4	<10	Pure	-
H142	3	-	Pure	-
H143	2	20	Complicated	TCC, dementia and mental retardation

3.2. Primers

Primers used during the study were designed by Primer3 software and are listed in the Tables 3.2, 3.3, 3.4, 3.5 and 3.6.

Table 3.2. Sequences of the primers used for exon amplification of *Spastin.*

Exon	Primer	Forward (up) and reverse (down) primers
Exon 1	SPG4 1aF	TACTATTTCTCCTACCCGCTGTT
	SPG4 1aR	CGACCCCACCGCCTTCT
	SPG4 1bF	CGGCGGCGGCAGTGAGAG
	SPG4 1bR	ATGAGGGCGCGGGAGAAG
Exon 2	SPG4 2F	TTTTTATGTATTACCTCTCAAC
	SPG4 2R	TGGGATGGCTATAAACAAAT
Exon 3	SPG4 3F	TTAGTTGGGAAATGTAGAT
	SPG4 3R	TATGTTAAAAAGCCTGGAC
Exon 4	SPG4 4F	TTTTT ACCTTCTCTGTTG
	SPG4 4R	AAGCTTTATTATTTTATGTTAGT
Exon 5	SPG4 5F	GTTCAGCTACAATTTTCTAATCC
	SPG4 5R	TATGATCAACTTAAGCAGGAAT
Exon 6	SPG4 6F	ATGTTAGGTTGTATTTTCA
	SPG4 6R	CAAGGTATTTATTATCTATTTCC
Exon 7	SPG4 7F	TCATAGGGCTTAGGCTTCA
	SPG4 7R	ATGGATTCAGTAACAGATGGTAT
Exon 8	SPG4 8F	CTGTTTGGGAAGATGCT
	SPG4 8R	CTCAAGGACAAGATAAAGTT
Exon 9	SPG4 9F	TGGCCTCATAGCTTACATTTTTAG
	SPG4 9R	TACGACAATATTGGAAACAGAG
Exon 10	SPG4 10F	GTGCTAGATTTTCAACATA
	SPG4 10R	GCCCTTCTTTAAAACTTCTTCC
Exon 11	SPG4 11F	GAATTTAGTAGGACCCACT
	SPG4 11R	GCCACATTAAAAATATCATA
Exon 12	SPG4 12F	ATGGCCAAGGTTAAAAATACAA
	SPG4 12R	CTGGAAGAAAATAGTGAAT
Exon 13	SPG4 13F	CTTTTCCTGTCATTTGCTGTTT
	SPG4 13R	GATGGTAGTTCTTGTTTCTGCTCT
Exon 14	SPG4 14F	ATCATTAATTCTGAAATTAG
	SPG4 14R	ATAAACCAAATCCAAA
Exon 15	SPG4 15F	AAAAAGCGGGAGGGGAAATA
	SPG4 15R	TGGGCAACAGAGTGAGACC
Exon 16	SPG4 16F	TGTATGTATTTTTAAGTGCCTGAC
	SPG4 16R	TACAATATAGAAGACAAAGAAA
Exon 17	SPG4 17F	AACAGCAGCATCATTACTTT
	SPG4 17R	GTTCTGCAGGTTTACAA

Table 3.3. Sequences of primers used for exon amplification of candidate variants identified for HSP families.

Family Number	Gene	Variant	Forward (up) and reverse (down) primers
P463	*SPG11*	c.1235C>G	TACAAGTGTGCAGAGCTGGG
			CCAAGGGAAAAACACTGCAT
	ZFYVE26	c.3794T>A	CTGGCCCAAGAGAATCTCAG
			TTAAGAAAGGCCAAGGCAGA
		c.985A>G	CTTGGTTGTTCAAATGGGCT
			GCCAAATGAGTGGGAGAGAA
P627	*SACS*	c.637G>A	TTTAATGGCTCATGCTTTTTCA
			CAACAAATGGTGCAAACTGG
	CES1	c.1087-3delT	ACTGCATTGGTTTGGTTGGT
			CCCAGTCTTCATTCTGCCAT
H6	*SDHAP2*	n.387A>G	TCACCATGTGAGTAGCTGGC
			AAACAGACCACCAGAGCACC
	ANKRD20A4	c.1317-3C>T	TGCAAAGGAAAAATAATCCAGAA
			TGGTTCAAGTTGTTTTGCTTCA
	OR11H12	c.719T>G	ATTGGATTGTGTTTCTGCCC
			GAGTGGGGTCACCATAGCAT
	LRCH2, RBMXL3	c.128G>A	GTTGTGGTGGAGGTGGAAGT
			AGTTTCCGACCACTGAGGCT
H16	*SPG11*	c.6215_6219dupAGAT	AAGGGTTTCAAGCTCAGCAA
			GGGGAACGGAGGAAATCTTA
	C19orf12	c.*3745G>A	CCACTGTGCCTAACCATGTG
			CCAGCAATTTCACTTTGAGGA
	C19orf12	c.424A>G	TGGACAGTTTAAGCCGGTTC
			AACTCCCAAGCCACCTCTTC
H28	*AP5Z1*	c.-1G>C	GTGACGCAGCCACGTAAG
			TAGGGAGAGATGCAGGAAGG
	AP5Z1	c.2287G>A	GCTGTCAAAGATGAGGACCC
			AGGCTCTTAATCTGCACCGA
	COL11A1	c.652-6delT	CCACATCTCGCTTTGGTATATG
			TGGGTGCTGAAGAGTCACAG
	PCDH11Y	c.3036T>G	AGCCTGAAACTCCCCTGAAT
			ACGGACACAGGTACCTCGAA

Table 3.3. Sequences of primers used for exon amplification of candidate variants identified for HSP families (cont.).

Family #	*Gene*	Variant	Forward (up) and reverse (down) primers
H29	*ZFYVE26*	c.4804C>T	TCTGGTCCTGTCAAAGCCAT
			GAAGTGAGAAGTGGCCAAGC
	RTN2	c.1168G>A	GGAAAGGGAGGTGAGGAGAC
			TTCTCACTGGAAAGGGTTGG
	ATXN1	c.624_626delGCA	TGGAGGCCTATTCCACTCTG
			TGGACGTACTGGTTCTGCTG
H38	*TTN*	c.39962C>T	TTTAAAGAACCTGAGATGCCA
			TGCAACAACAATGAGGACAA
	TTN-AS1, TTN	c.63917G>A	TTACAGACACTCCTGGGCCT
			GCACTGGTTTTGGGACATCT
	ARNT	c.-29delG	CTATATAGGCGGGGTCTCCC
			CTTAGTTGTCAGCCCCTTCG
	WFS1	c.1294C>G	TATGCCCATTTCCTGCTCTC
			CTTCAGGTAGGGCCAATTCA
H44	*ZFYVE26*	c.7417-5G>C and c.7607G>A	TCCTGTTATGAACGTGGCAG
			CAGAGCCAGAGAAGTCCCAC
	CTAGE15	c.1262A>G	TGAGAATCAGAAGCTTCAACAGA
			CGAGCTGCCAACCAATTATC
	POTED	c.337G>A	GAGCAAGATGGGCAAGTGTT
			CCTTTCTGGGGACTTTACCC
H45	*SPG11*	c.2245-1G>A	TTGAAAGAGCAGAAAGCTATGG
			TTCCAAGTTTTTCACCCAGG
	SPG11	c.2057T>C	TCCTTGGAAGCTAACAGATGC
			ATCCCCAAACCGATAAAACC
	SPG11	c.1499_1500delAG	CGCCTATAAAAATGTTGAAAAACC
			CCAGCCATTGAGATGACAAA
H55	*COL7A1*	c.5086C>T	TGAAGGCTACCAACATTCCC
			CCTTGTGTGAAGGCACACTG
	SHMT1	c.571C>T	GTTCAAGTGGGAGGACAGGA
			TCTGTAAGAGGCACCAGGCT
	CDC27	c.77T>C	TTTGATGTAAGGCTGTGTGCT
			CAATCAGAAAAACAACAAGATGAT

Table 3.3. Sequences of primers used for exon amplification of candidate variants identified for HSP families (cont.).

Family Number	Gene	Variant	Forward (up) and reverse (down) primers
H72	RGPD3	c.331G>A	CAGTGGAATTAAACCCAACACA
			GAAACACAAACACCTGGGCT
	CD163L1	c.1567T>A	GAGAGTGGGGGACTGTGTGT
			CTTTCCCCATCCACTGTGTT
	ACAN	c.4207A>G	GAGTAGAGGACATCAGCGGG
			ACTTCTCCAGAAGGAAGCCC
	CDC27	c.823C>G	CCCACTGGGAACAGGAACT
			TTCTGAAAACAGAATGCCCC
	XKR3	c.765T>G	CTACCATTAAGCTACCGCCG
			TTTTCCAAAACTCCAGCCAC
H77	SPG11	c.3036C>A	ACTGAGCCGTATTGGAGGTG
			CCAACTGTTGAGATGGAGAAAA
	PDHA2	c.875G>C	CTGTGTGTTCGTGAGGCAAC
			GGGCAGCATCATCAATTTCT
H99	ZFYVE26	c.5806T>A	CAGGGTCTTTCAGCCTTTTG
			TTGACGAACATCATCTTGGC
	ZFYVE26	c.1844C>T	TCTCTGTGCCTGGAGCTTCT
			AGGCCCTGGGTAACTGTCTT
H108	C19orf12	c.*1020C>A	GCTGACACAGAAACGCAAAG
			TTCCTATCCAAGGCTGGTTG
	ZFYVE27	c.820G>A	TGTGCTCGTGTTTCCATCTC
			GCTGAGTAGAGGATGCTGGG
	CPT1C	c.1343+6T>G	CGCTTTCTTTGTGTCACTGG
			GGGTCCTCAGCTCTATGCAG
H110	SACS	c.12166A>C	TTTGTCAGTTTGGAGCGTTG
			TGACTGCATTACCAAATCGC
H133	CCT5	c.437A>T	TGAACTGATTGTCTGCTGGC
			CGTGGTTTTTGCTGTCTGAA
H142	SDHAP2	n.1482G>A	GGTACCAGTTTGGCACCATT
			TTTCTTGCTCGTGCCTTTCT
	C14orf23	n.450_451insAAC	AAAGGCATTTGGGGAAGAAC
			GGCAAGGAAGATGACCAAGA

Table 3.3. Sequences of primers used for exon amplification of candidate variants identified for HSP families (cont.).

Family Number	Gene	Variant	Forward (up) and reverse (down) primers
H142	SDHAF2	c.103T>C	TTCAGATGCTTGCTCTGTCAA
			TTCCCCTCTTTCTGCTCTCA
H143	SPG11	c.3904_3907delTCTA	TCTTGGGCACTTTAATTGGG
			TCTTGGGCACTTTAATTGGG
	KIF1A	c.767A>C	GCATCTTGTGGCCTTTCCT
			CTATGCAGGCTCACCCTCAG
	KIF1A	c.883-3C>T	AGCCTGAGAATGTGTCCAGC
			ACAGGCAAACCCACAAGTTC
	KIF1A	c.1261A>C	TACCCCTCCTCCCCATAGAC
			TGCCTCTCTAACTCCTCCGA

Table 3.4. Sequences of primers used for exon amplification of candidate variants identified after SNP genotyping and PLINK analyses.

Family #	Gene	Variant	Forward (up) and reverse (down) primers
P392	KCNMB3	c.157G>A	ATCTCCTTCTCCCTTTCCCC
			CCAGCCTCTCCTCCTTTACC
	SLC7A2	c.47G>A	TCCATCAGTCCCAAAACAGA
			TCTTTCCCTTCCTCCCCTAA
	PDGFRL	c.980T>A	TGCTCTCTGAAAGGAACGCT
			CCTGCATCAATCGTCTCAAA
	NAT2	c.590G>A	CCTGGACCAAATCAGGAGAG
			GATGAAGCCCACCAAACAGT
	SCYL2	c.1070C>T	TCCCTTCTTTGATGATGTTGG
			GGCTTCAGTTGGCAGTGTTT
	UTP20	c.1505C>G	AGAGGGCAAGAGCGATTACA
			ATGAGGTAACACCACGAGGG
	MYBPC1	c.1443C>G	TGGAAAAGAAATCTGCCTGA
			TCCCTATGGCTACCCTTCAA
	CAPRIN2	c.1906-6C>A	TTTGGCATCCTGAGACAGTG
			GAGTAGCTGAAGGCGGTTGT

Table 3.4. Sequences of primers used for exon amplification of candidate variants identified after SNP genotyping and PLINK analyses (cont.).

Family Number	Gene	Variant	Forward (up) and reverse (down) primers
P392	OVOS2	c.2783T>C	TTTATCCCCCAAATCTGCTG
			CCACCTCAACCAAGATTTTCA
	CCDC114	c.136C>T	GCTGAAGACTGGACTCCTGG
			AGAACCACCTCTCAAACCCC
	ETFB	c.734C>T	AAAGACCCTCTCCCAGGTGT
			CTCAGGTCAGCTGTCACCAC
H6	GUCA2A	c.20C>T	GACTGACAGAGGCTTCCAGG
			GGGTAGGGACACAGGACTCA
	SLC5A9	c.1852G>T	GCATGACTTCTCTCCAAGGC
			AGAGTCCCCATCCCAAAGAC
	PRKAG3	c.1018C>T	CTAGCAGTCGTGGGGAAGAG
			GATGTCCAGTGCAGTCAGGA
	SLC23A3	c.443-6T>C	TGACACTATAGGGGGCTTGG
			CCTTTCCAGGAGCATCTCAG
	RESP18	c.90G>T	GCCGAGGAGGGTTAAACAG
			AGACCCCCACTTCTCAGGAT
	IFIT5	c.5+6G>T	CCTTTAACAAAGACACGCCG
			GAACTGCGCTGGAAATGAAT
	MYOM1	c.3576-5C>T	GGAGGTGTCCAGATGAAAGG
			TTCGATGTGTGAATGCACTG
	SPINT3	c.161T>C	TGGGTTTGGTCTCCTTCAAC
			AAATGCCTTCTATGGCCCTC
H28	CERS2	c.411-4C>A and c.411-5G>A	GGGAGATGTGAGGAGGTGAA
			CTCCCAGTAATCCCCACTCA
	COL6A3	c.9206C>T	AAGCAGAACCTCACGGTCAC
			CCCAGGCTGAGTTGTTTTGT
		c.3751G>A	CATTTGTTTCGCAGGTGTTG
			GACACGTACTCCAGGGCATT
	TNXB	c.8111G>A	CAGTACAGGAATGGGGATGG
			TCACCAAAGAGCAAGAGGGT
	TNXB	c.7235C>T	CTGCCATCCTCTCTGTGTCC
			GTCCTTGTACTGCACGGTGA

Table 3.4. Sequences of primers used for exon amplification of candidate variants identified after SNP genotyping and PLINK analyses (cont.).

Family #	Gene	Variant	Forward (up) and reverse (down) primers
H28	LAMA4	c.3239G>A	GGCTGCCAGTTACTTCTTCG
			TAGACCTGGGAAGGTTCCTG
	GPRC6A	c.431T>G	TACCTGGAGTCAAACTGGGG
			AACCTACCTGTGGCATGAGC
	FAM35A	c.730A>G	GGGCCTGAAGTGGTACAAAG
			ATCTCCCCGTAACTTGCCTT
	NUTM2D	c.103C>T	TCTGGGGTCCTATGTTCCAG
			GTCTTCCCTCCCTCCTCAAC
	ATAD1	c.701T>C	GCCCTAGTAATAACATTCTCAGAGG
			CCCTGTCATAAGCACCCCTA
	LIPF	c.511A>G	CACCTGACATCCACTGATGG
			AAACCACTTGGCCTTGACTG
	HSDL1	c.980C>G	TGTTGTGATTGATTGGTGGG
			CAGAATATCGAGGTTCCCAGG
H29	TTC22	c.40C>G	CTCAGGGACAAGGGTGAGG
			TCCAGCTCCTCCAGGTAGAA
	SSPO	c.15446-7G>A	GCAGATGAAGGGGAAACAGA
			GTGAGAGAGCGTGGAAGGAC
	NAT2	c.590G>A	Same as P392
	SLC22A10	c.1547C>T	AGGGCAAGAGCTTCAGGAAT
			GCTTCTCTCAGGACATGGCT
	KRR1	c.401G>A	TTTGCTTTGTTTGCATCAGTG
			CCAATAAGCCGTTGTCTTCG
	AKAP13	c.1871G>T	GAGGAGAGTGCTGATGCTCC
			ACAAATGGGTGAGGATTTGC
	PDIA2	c.1144C>G	CACAGGACCTCCTGAACCCT
			CTGCCTCCCTCAACACTCAC
	WDR90	c.2594C>T	CCACAGCGGACATGGTATG
			GGCTGATGTGTACTCTGCGA
	CLCN7	c.1798-8G>A	GGTAGGCTCAGGGTTTCTGG
			GCTGTACCTGGCAGTGAGTG
H38	EMC1	c.2128G>A	TGCCCTAGGATCTCACCACT
			TTGAAGCCCTTGTTTTCCTG

Table 3.4. Sequences of primers used for exon amplification of candidate variants identified after SNP genotyping and PLINK analyses (cont.).

Family Number	Gene	Variant	Forward (up) and reverse (down) primers
H38	RNF186	c.67G>A	CAGCTGACACTCTCTGCCAA
			CAGGAGCTTCAGGCAGATG
	NBPF3	c.341A>G	CTGGCAGAGAACAAACAGCA
			ACAAAATGTGAAAATGCCCA
	MYOM3	c.3972-5T>C	CTTTCCTTCATCACCCCTCA
			CCTGCAGCTCTGTAAGACCC
	TRAPPC12	c.901A>G	AGAGCCTTTCGCGCACAT
			TCTCACAGACGCAGAGAGGA
	APOB	c.293C>T	TTCCCTTTGGTGCTCTGATT
			ACTTACCTGGACATGGCTGC
	CAPN14	c.847G>A	TTTCCACTGCAAGTGCTCAG
			CCTGATCCAGCCTAAGACCA
	WFS1	c.1294C>G	TATGCCCATTTCCTGCTCTC
			CTTCAGGTAGGGCCAATTCA
	LOXL4	c.1214A>C	ACAAGTGTGGGGTTACGGAG
			TAGCTCGATGTGGTGTCCTG
	PKD2L1	c.1414C>T	CCCTTTTCCTTCTTCATCCC
			CCCAAGGATACATACACCCG
	PREX1	c.4526+7G>A	GACATCTCCCCTCCCAGTG
			AATCTCACGCAGGCAAGAGT
H44	KLRB1	c.503T>C	GAACCTGATACGTGACAAAGCA
			CAACCAACTTTGAAATCATCCA
H45	PLCL1	c.1999G>A	TTGCAAATGAGTACCCAGAGG
			TCTCGCATTATAGACGGCCT
	ALPK2	c.4345C>T	GAGTACACCAGGGCTGGAAA
			GAATTTGTTCAGCCTCCTGC
		c.2748T>A	AGGGGACATGACTGAGTTGC
			CACTAGGACTGGGGCTCTTG
		c.2487A>C	AGAAGCAGTGTGTGCAATGG
			AGGGGACATGACTGAGTTGC
H55	ZNF142	c.3373C>T	CTGGTTCAAAAGCAGAAGGG
			CATGCAACGGCTCTGTTTAC

Table 3.4. Sequences of primers used for exon amplification of candidate variants identified after SNP genotyping and PLINK analyses (cont.).

Family Number	Gene	Variant	Forward (up) and reverse (down) primers
H55	CYP27A1	c.256-1G>C	TGTAGCCATCAGACCCTTCC
			TTCCGTACTGGGTACTTGCC
	SLC23A3	c.443-6T>C	Same as H6
	RESP18	c.90G>T	
	FAM198A	c.1615G>A	TCGCATCAGGGTGTCTACAG
			GCTTTTCCCATTTGTCCTCA
	NT5DC2	c.1348-7T>A	AGTACATGCACTCGCTGACG
			CACCTGCTCACCTCAGCTC
	ITIH1	c.1754A>T	GACACCATGGGGGCTTAATA
			TGATGCTCATGGAGGTCAGT
	CCDC66	c.1379G>A	TCCCAAACAAAAATGATTTAGC
			TGACACAACTGGACAGGACC
	DNAH12	c.3923A>G	TCACTTTCCTTTTTCTTTGTCCTT
			CCTGCATAGCCAGGATTCAT
	CDC27	c.77T>C	TTTGATGTAAGGCTGTGTGCT
			CAATCAGAAAAACAACAAGATGATG
	EFCAB13	c.806-1G>A	TGTTGGGCGTGTACTGAGTT
			TTGTCATAGCAACAAGAGCTCAC
H82	F7	c.7612C>G	TGCTGATCAGTTTTGCTCCTT
			TTCCAAAATCCTCATGCACA
	TMEM70	c.4515_4516delTC	AAGGTCCATTCCATCTGCAT
			CCAATTTACAACTTGCTTTTTCA
	TMEM70	c.3964C>T	CTGCCCTCAGGAGAACTCAC
			TTCCTTAACGTTTTCATAAATGC
		c.1331_1334delCTGT	TGAATATGTTCATGCATGGCT
			TTCACACAGAGGTCATCTGGA
	CUBN	c.290C>T	TGAATGCAAAGAATGTGGGA
			GCGATTGTTAAATGCCTTCC
	SLIT1	c.313delA	ATGCCATTGTCATGGGTTG
			TCTTGTGGGAGAAATGGAGAA
	RRP12	c.1412A>G	AATGTGCCAAGGTCTTCAGG
			TCGTGCAAGGTGTGAATTGT

Table 3.4. Sequences of primers used for exon amplification of candidate variants identified after SNP genotyping and PLINK analyses (cont.).

Family Number	Gene	Variant	Forward (up) and reverse (down) primers
H82	CHUK	c.155C>T	GAGTCCTGGGGAGCTTCTCT
			ACCGTGATCACCTGGAAGAC
	C10orf120	c.1010G>T	AGCAGTCCCAAAAGAAGCAA
			AGCGCTTCCCACAGAGATTA
	FAM24B	c.149T>A	CAAATTGCACATCCTCATGG
			CACAGATGGCACAGCTCAGT
H99	FUCA1	c.4C>T	GTTAGTCAGAGTGGGCGGAG
			GGATGAACACCCCGAACTT
	SSPO	c.9211_9212insC	GTGGGAGGAGGAGAAGGTG
			CCCTGGGTTGGTAAGAATGA
	GIMAP7	c.247C>T	AACTGTCAAAAAGCATCCCG
			TGTGCTTCATGGCTGACTTC
H108	SLC45A2	c.1122G>C	CCAGAGGTGGAGAAGCAGAG
			TCCAAGTTGTGCTAGACCAGA
	AMACR	c.524G>A	GGCAGAAGTGGTGAGAATCC
			CCTTGGTAACCTGGCTTGAA
	NEBL	c.1132G>C	GGATGACCCTCAGGTGGTAG
			CCAAGATCGTACCTCCCTCA
	C10orf53	c.353C>T	GGCTGACATTTCACCAGCTT
			GAACCCCTTTCCCACAATCT
H110	CNTN4	c.2093-6G>C	AGTGAGTCACCTCTGCCAGG
			GAGGGAGTGTCTGAACCCAA
	LRRN1	c.988C>T	AATCAACAATATGGGCGAGC
			GAAGCGGATGTTGGTTTTGT
	MDN1	c.9010G>A	GTGGCAGAGAAACGGAAAAA
			CAAAACAGACTTGGGTGCCT
	TRIM64C	c.762-6_762-5dupTT	GTTAACCAAAACCGGCAGAA
			AACCAAGGGACCCAATAAGG
	TP53AIP1	c.313C>T	CAAGGACTCCATACGTTTTGC
		c.304A>G	AGTCTCTCTCTGTCGCCCAG
		c.63dupG	GGACTGGCCCTAACAACAAA
			TTCACTGCAGGGACTTACCC

Table 3.4. Sequences of primers used for exon amplification of candidate variants identified after SNP genotyping and PLINK analyses (cont.).

Family Number	Gene	Variant	Forward (up) and reverse (down) primers
H110	PABPC3	c.609T>G	GCTCCTAAATGGTCGCAAAG
			GTGCATCTTCATGCCTTTCA
H133	ALMS1	c.6299C>T	GACCAGAAGACTCCATCCCA
			CAGCTGGTCCAGGAATTGTT
	CCDC66	c.1379G>A	Same as H55
	KALRN	c.3729+6C>T	CGGTGGACAAGCACTACAGA
			CCAGCACTCCCCACTAGAAA
	UGT2A1,UGT2A2	c.949G>A	TGTCAGAAGACAGACAGTTCAACA
			GTACCAGGATTCCAGGCTGT
	C4orf26	c.133C>T	TTCTCCTCTGCCACCACTCT
			ACCTTGAAGGGACGAAAGGT
	SDAD1	c.1724C>G	CCATCAGCACTAGCCGAGTT
			TTTGGCTTTTTATGAAGGCG
	SAA2-SAA4,SAA2	c.92-7_92-3dupTTTCC	GGCACTTGAAGAAGCCAACT
			CCCGAGCATGGAAGTATTTG
	HPS5	c.1249C>A	AGCTATTCCATTGCGCTGTT
			GACAACAAATTTGGGGATCCTA
	IGSF22	c.2014C>T	AAACCACTGCCCAAAGTGAC
			CCTACTCAGCCTGCCTTCTG
	IGSF22	c.696+3A>G	TTGAGAAGGTGTGCATGGAG
			TTGCTGCTTTCACAGTGGAG
	VWA8	c.1978A>G	TTGGCCTTACCCCTTCAGTA
			TGTTTACCTGGAAAGGCAGG
H142	GPATCH2	c.1420A>G	ACGCAACCCCCTCTAGTACC
			TCCTAATCCCTTTGGCCTCT
	DNAH7	c.4897-9_4897-8delTT	TGTATGCTGGGACCTTGTCA
			CCACTGAAGAGGCAAATGCT
	PLCL1	c.1999G>A	Same as H45
	DGKB	c.169-5T>G	TGATTTTGAACACAGCACAGAA
			TTACCGCCTGATAGGAGAGC
	SEMA3C	c.1286A>G	ATATGCGAACCACCAAGGAG
			GTTTCAAGATCATGGCCCTT

Table 3.4. Sequences of primers used for exon amplification of candidate variants identified after SNP genotyping and PLINK analyses (cont.).

H142	KIAA1324L	c.742-7T>C	TTCAGCGGAGGTTAGGTCAG
			TGACCGCCTTAGAACCCATA
	PRPF4	c.812-8T>C	GTGCAGTCACACTCTGGCAT
			TTTTGGGTCCAAGGAGACAG
	UBQLNL	c.820T>C	ATGCAGCAGAACCCAGAAGT
			CTGTGAAAGGGTTTCCTCCA
	UNC93B1	c.887dupC	TGGATTGACTGTGTGCCCTA
			ATCCGATCCAGACCCTTTCT
	C11orf24	c.290G>T	TGAGGATGTAACCATGGCAG
			CACTGGAGGCCACAGTCATA
	AQP11	c.665C>T	TATGCCAGGTGATTCTGCAA
			ACCTAAAGAAGGAGCCAGCC
	USP35	c.706G>A	GGTGGAACCAGAGACTTTGC
			TGCTAATGGCAGTCAGCATC
	TCL1B	c.*15+6T>A	CGTTAAGGCCAACTGGAAGA
			GTACCATCACATCCAAGGGG

Table 3.5. Sequences of primers used for exon amplification of candidate variants identified for H28 for a possible XL inheritance.

Gene	Variant	Forward (up) and reverse (down) primers
ARSD	c.959G>A	TCCTTTCTTTGCTGCATGTG
		CTCCCAAAGTGCTGGGATTA
	c.524G>A	CAGGGTGTGAATTGTGCATC
		GACAGAGAAGAAACCGCAGG
ATP7A	c.2299G>C	CAGTTTTTCGGAGGCTGGTA
		AACACAAACAGCATAGGGGG
DGKK	c.3060dupC	GTTACAGCAGTTGCCCCATT
		CCAAGACAAGATTTTCCCCC
FAM104B	c.160A>C	AAAATGGGGGCTGTTAATCC
		AAGTTTGCATCGGGTTCAGT

Table 3.5. Sequences of primers used for exon amplification of candidate variants identified for H28 for a possible XL inheritance.

Gene	Variant	Forward (up) and reverse (down) primers
H2BFM	c.217C>T	AAGGCCCAGAAGCAGAAGAG
		GTCACGCGCTTGGTGTAAT
IDH3G	c.82-7_82-6dupCT	TTATCTATTGTGCCCCTGGC
		GCCACATGGAGAAAGACACA
SSX7	c.544G>A	ATTCTCCATCACAGGACCCA
		AATGATTTGGGGAGCAAGTG

Table 3.6. Primers used for RT-PCR experiments.

ACTIN-F	CACCAACTGGGACGACAT
ACTIN-R	ACAGCCTGGATAGCAACG
ATAD1-F	AGGTTGCGTTTCCTTCCTGT
ATAD1-R	TGAATGTTAACCCAGGGGCA
CYP27A1-F	GGCCAAGTACGGTCCAATGT
CYP27A1-R	GGTGTCCTTCCGTGGTGAAC
GAPDH-F	GGATTTGGTCGTATTGGG
GAPDH-R	GGAAGATGGTGATGGGATT
SEMA3C-F	TTCAGGAAACTTAAAAGGATGGC
SEMA3C-R	TGGCCAGAAAACACTCAAAGC
ZNF142-F	GTGGCTTCCACATCCCCAATA
ZNF142-R	GGCCTCAGCTTTCCTATCTCC

3.3. Primary and Secondary Antibodies

Primary and secondary antibodies used throughout the study and their purpose of use are given in the tables 3.7 and 3.8.

Table 3.7. Primary antibodies used in this study.

Primary Antibody	Application	Host	Dilution	Supplier	Product Number
Actin	WB	Goat	1:1000	Santa Cruz Biotechnology	sc-1616
ATAD1	WB, ICC	Rabbit	1:500	Atlas Antibodies	HPA037569
CYP27A1	WB, ICC	Rabbit	1:100 (ICC) 1:1000 (WB)	Abcam	ab126785
GM130 (Golgi Marker)	ICC	Rabbit	1:1000	Abcam	ab52649
SAMHD1	WB, ICC	Mouse	1:1000	Abcam	ab67820
SEMA3C	WB, ICC	Rat	1:100 (ICC) 1:500 (WB)	R&D Systems	MAB1728
Vinculin*	WB	Sheep	1:1000	R&D Systems	AF6896
ZNF142	WB, ICC	Rabbit	1:500	Novus Biologicals	NBP2-13554

WB: Western Blot, ICC: Immunocytochemistry.*Vinculin is kindly provided by Assoc. Prof. Umut Sahin.

Table 3.8. Secondary antibodies used in this study.

Target	Brand	Product Code	Host	Dilution	Tag
Mouse	Cell Signaling	7076	Goat	1: 1000	Horse Radish Peroxidase
Rabbit	Santa Cruz Biotechnology	sc-2004	Goat	1: 5000	Horse Radish Peroxidase
Rat	Novus Biologicals	NB7115	Goat	1: 5000	Horse Radish Peroxidase
Sheep	R&D Systems	HAF-016	Donkey	1:5000	Horse Radish Peroxidase
Mouse	Molecular Probes	A21202	Donkey	1:500	Alexa 488
Rabbit	Molecular Probes	A31572	Donkey	1:500	Alexa 488
Rat	NorthernLights	NL013	Goat	1:500	NL-557

3.4. Cell Line Mediums

Cell lines and growth and freezing mediums prepared for these cell lines are given in the Table 3.9.

Table 3.9. Cell lines and mediums used in this study.

Cell Line	Medium	Recipe
B95-8 cell line and Lymphoblastoid cell lines	Growth Medium	20 % FBS
		1 % penicillin/streptomycin
		1.1 % Natrium-Pyruvate (100 mM)
		2 % HEPES
		1 ug/ml Cyclosporin A
		1 % L-Glutamin
		in RPMI
	Freezing Medium	70 % RPMI
		20 % FBS
		10 % DMSO

3.5. Buffers and Solutions

Buffers, solutions, and gels used in this study are given in the Tables 3.10, 3.11 and 3.12, respectively.

Table 3.10. Buffers used in this study.

Buffer	Recipe
Developing buffer	1.5 % NaOH
	0.015 % Formaldehyde
Loading Buffer (10 X)	2.5 mg/ml Bromophenol Blue (BPB)
	1 % SDS in 2 ml glycerol
Nuclei Lysis Buffer	400 mM NaCl
	2 mM EDTA (pH 7.4)
	10 mM Tris-HCl (pH 8.0)
	diluted to 1 L with dH_2O

Table 3.10. Buffers used in this study (cont.).

Buffer	Recipe
Protein Sample Buffer (2X)	4 % SDS
	200 mM 2-mercaptoethanol
	20 % Glycerol
	100 mM Tris
	0.2 % Bromophenol blue
Protein Sample Buffer (6X)	60 % Glycerol
	6 mM 2-mercaptoethanol
	300 mM Tris
	12 mM EDTA
	12 % SDS
	0.04 % Bromophenol blue
Red Blood Cell (RBC) Lysis Buffer	155 mM NH_4Cl
	10 mM $KHCO_3$
	1 mM EDTA (pH 7.4)
	diluted to 1 L with dH_2O
RIPA Buffer	50 mM Tris-HCl (pH 7.4)
	2 mM EDTA
	1 M NaCl
	1 % Triton X-100
	0.1 % SDS
	0.1 % Na-deoxycholate
	Phosphatase inhibitor cocktail
	Mini protease inhibitor cocktail
Running Buffer	250 mM Glycine
	25 mM Tris
	0.2 % SDS
Transfer Buffer	25 mM Tris
	200 mM Glycine
	20 % Methanol
Tris-Boric Acid-EDTA (TBE) Buffer (10 X)	20 mM EDTA (pH 8.3)
	0.89 M Tris-Base
	0.89 M Boric acid
Tris-EDTA (TE) Buffer	20 mM Tris-HCl (pH 8.0)
	1 mM EDTA (pH8.0)

Table 3.11. Solutions used in this study.

Solution	Recipe
Acrylamide Solution (30 %)	29 % acrylamide
	1 % N, N'-methylenebisacrylamide
Acrylamide Solution (38.5 %)	37.5 % acrylamide
	1 % N, N'-methylenebisacrylamide
$AgNO_3$ solution	0.1 % $AgNO_3$ in dH_2O
Ammonium Persulfate (APS)	10 per cent APS (w/v) in dH_2O
Blocking Solution (WB)	5 % bovine serum albumin or 5 % skim milk powder in TBST
Blocking Solution #1 (IHC)	5% BSA
	1 % Donkey serum
	0.2% TritonX
Blocking Solution #2 (IHC)	0.25 % TrtX-100
	0.25 % BSA
	1X PBS
Cyclosporin A	1 mg/ml in DMSO
DABCO mounting medium	2.5 % DABCO
	50 mM Tris-HCl (pH 8.0)
	90 % Glycerol
DAPI	0.02 ug/ml DAPI in 1X TBS
Denaturing loading dye	95 % formamide
	20 mM EDTA
	0.05 % Xylene Cyanol
	0.05 % Bromophenol Blue
Deoxyribonucleotide Triphosphates (dNTPs)	100 mM of each dNTP in dH_2O
Ethidium Bromide (EtBr)	10 mg/ml in dH_2O
HEPES-buffered saline (HBS) solution (2X)	50 mM HEPES
	280 mM NaCl
	12 mM dextrose (D-glucose)
	10 mM KCl
	1.5 mM $Na_2HPO_4 \cdot 2H_2O$ (pH 7.05)
	Adjust pH to 7.05 with 10 M NaOH
	filter sterilize and store at -20C as 10 ml aliquots
NaCl Solution (5 M)	292.2 g NaCl in 1 L dH2O

Table 3.11. Solutions used in this study (cont.).

Paraformaldehyde Solution	4 % PFA in 1X PBS
Proteinase K	20 mg/ml in dH_2O
Sodiumdodecylsulfate (SDS) (10 %)	10 % SDS (w/v) (pH 7.2)
Stripping Solution	62.5 mM Tris-HCl (pH 6.8)
	2 % SDS
	0.7 % 2-mercaptoethanol
TBS with Tween-20 (TBS-T)	0.1 % Tween-20 in 1X TBS
Tris Buffered Saline (TBS)	20 mM Tris-HCl (pH 8.0)
	150 mM NaCl

Gels used in this study are given in the table below. For SDS polyacrylamide gel, a sample for 10 % gel was given in the table. For 8 % and 12.5 % gels, Acrylamide:Bisacrylamide solution ratio was changed to 8 % and 12.5 %, respectively.

Table 3.12. Gels used in this study.

Gels	Recipe
Agarose Gel (1 %)	1 % (w/v) Agarose in 0.5 X TBE Buffer
Polyacrylamide Gel (10 %) (SSCP)	30 % Acrylamide:Bisacrylamide (29:1)
	12 % 5X TBE
	1 % TEMED
	0.1 % APS
SDS Polyacrylamide Gel (10 %) (Running Gel)	375 mM Tris-HCl (pH 8.8)
	10 % Acrylamide:Bisacrylamide (37.5:1)
	0.1 % TEMED
	0.1 % SDS
	0.1 % APS
	0.001 M EDTA
SDS Polyacrylamide gel (4.5 %) (Stacking Gel)	4.5 % Acrylamide:Bisacrylamide (37.5:1)
	125 mM Tris-HCl (pH 6.8)
	0.1 % TEMED
	0.1 % SDS
	0.1 % APS
	0.001 M EDTA

3.6. Chemicals and Disposables

All solid and liquid chemicals and disposable materials used in this study are given in the Table 3.13.

Table 3.13. All solid and liquid chemicals and disposable materials used in this study.

Chemical / Disposable	Supplier	Catalog Number
2-Mercaptoethanol	Merck Millipore	805740
ACD-A containing blood collection tubes	BD Vacutainer	366645
Acrylamide	Sigma-Aldrich	A3553
Agarose Gel DNA Extraction Kit	GeneMark	DP03P
$AgNO_3$	Merck Millipore	101512
Ammonium peroxodisulphate (APS)	Fluka	9914
B95-8 cells	Kindly provided by Prof. Rebacca Schüle	
BCA assay kit	ThermoFisher Scientific	23225
Boric acid	Sigma-Aldrich	B6768
Bovine Serum Albumin (BSA)	Sigma-Aldrich	A2153
Bromophenol Blue (BPB)	Sigma-Aldrich	B5525
cDNA Synthesis Kit	BioRad	1708890
Cell scraper	ISOLAB	121.02.020
Centrifuge Tubes (15 ml / 50 ml)	TPP	91015 / 91050
Coverslip		
Culture flasks (75 cm2 / 150 cm2)	TPP	90076 / 90151
Cyclosporin A	Abcam	ab120114
DABCO (1,4-Diazabicyclo[2.2.2]octane)	Sigma-Aldrich	D27802
DAPI	Roche	10 236 276 001
Dextrose (D-glucose)	Sigma-Aldrich	G8270
Dimethylsulfoxide (DMSO)	Sigma-Aldrich	M81802
DNA Ladder (100 bp / 1 kb)	ThermoFisher Scientific	SM0241 / SM0311
dNTP	ThermoFisher Scientific	10297018
Donkey serum		
EDTA	Riedel-de Haen	34549

Table 3.13. All solid and liquid chemicals and disposable materials used in this study (cont.).

Chemical / Disposable	Supplier	Catalog Number
Fetal Bovine Serum (FBS)	ThermoFisher Scientific	10270106
Ficoll	GE Healthcare	17-1440-02
Formaldehyde solution	Sigma-Aldrich	15512
Glycerol	Sigma-Aldrich	G5516
Glycine	Sigma-Aldrich	G8898
HEPES	Sigma-Aldrich	H3375
High Resolution Melting (HRM) Master Mix Kit	Roche	04 909 631 001
K2EDTA blood tube	BD Vacutainer	367844
KCl	Sigma-Aldrich	P9541
$KHCO_3$	Sigma-Aldrich	S5761
L-Glutamin	ThermoFisher Scientific	25030081
Magna lyzer ceramic beads	Roche	03 358 941 001
Methanol	Sigma-Aldrich	32213
Micro-cenrtifuge Tubes (1.5 ml / 0.5 ml/ 0.2 ml)	Axygen Scientific	430909 / 6530 / 6531
Microscope slides / cover glass	Isolab	
N, N'-methylenebisacrylamide	Sigma-Aldrich	T7024
$Na_2HPO_4 \bullet 2H_2O$	Sigma-Aldrich	255793
NaCl	Sigma-Aldrich	S7653
NaOH	Sigma-Aldrich	6203
Natrium-Pyruvate (100 mM)	ThermoFisher Scientific	11360070
NH_4Cl	Santa-Cruz Biotechnology	254134
Paraformaldehyde (PFA)	Sigma-Aldrich	15812-7
Pasteur pipettes	Isolab	
Penicilin/Streptomycin	ThermoFisher Scientific	15140122
Phosphatase inhibitor cocktail tablets	Roche	04 906 845 001
Phosphate-Buffered Saline (PBS)	ThermoFisher Scientific	10010023
Pipette tips (1000 ul/ 200 ul / 10 ul)	Axygen Scientific	4713 / 4845 / 4826
Poly-D-Lysine	Sigma-Aldrich	P7280

Table 3.13. All solid and liquid chemicals and disposable materials used in this study (cont.).

Chemical / Disposable	Supplier	Catalog Number
Polyvinylidene fluoride (PVDF) membrane	Roche	3010040001
Protease inhibitor cocktail tablets	Roche	11 873 580 001
RNA MiniPrep kit	Zymo Research	R1054
RPMI	ThermoFisher Scientific	21875034
Skim milk powder	Sigma-Aldrich	70166
Sodium deoxycholate	Merck	6504
Sodium dodecyl sulphate (SDS)	Sigma-Aldrich	L3771
SYBR Green	BioRad	1725270
Taq Polymerase	ThermoFisher Scientific	EP0402
Trizsma base	Sigma-Aldrich	T1503
Triton X-100	Sigma-Aldrich	T8787
Tween-20	Riedel-de Haen	63158
Western Blotting Luminol Reagent	Santa-Cruz Biotechnology	sc-2048
Xylene Cyanol	Sigma-Aldrich	X4126

3.7. Equipment

All equipment used in this study are given in the Table 3.14.

Table 3.14. All equipment used in this study.

Equipment	Supplier
Autoclave	ASB270NT, Astell Scientific
Balances	Models VA124 and CC081, Gec Avery
Blotting Apparatus	Mini Trans-Blot Cell, Bio-Rad
Centrifuges	Centrifuge 5415R, Eppendorf; Allegra X-22R Centrifuge, Beckman Coulter; Spectrafuge 16 M, Labnet
Deep Freezers	-20C (Bosch, Germany) -20C 2021D, Arçelik; -70C GFL

Table 3.14. All equipment used in this study (cont.).

Equipment	Supplier
Documentation Systems	GelDoc Documentation System, Bio-Rad; Stella, Raytest; G-Box, Syngene
Electrophoretic Equipment	Mini-Protean III Cell, Bio-Rad
Heat block	DRI-Block BD-20, Techne
Homogenizer	MagnaLyser, Roche
Incubators	Shake'n'Stack, Hybaid; Oven EN400, Nüve
Magnetic Stirrer	Speed Safe Hanna Instruments
Micropipettes	Gilson
Microscopy	TCS SP5, Leica Microsystems (Confocal microscopy system) MZ16FA, Leica Microsystems (Fluorescence microscopy)
Microwave oven	Arçelik
Power supply	PowerPac Basic, Bio-Rad
Real time PCR cycler	Exicycler 96, Bioneer
Refrigerator	Arçelik
Shaker	SL350, Nüve
Spectrophotometer	Nanodrop ND-1000 Spectrophotometer, Thermo Fisher Scientific
Thermocycler	iCycler, Bio-Rad
Ultra Low Temperature Freezer	Thermo Forma
Vortex	Nuvemix NM110, Nüve
Water Purification	WA-TECH Ultra Pure Water Purification System, WA-TECH

4. METHODS

4.1. Molecular Genetic Analysis

4.1.1. Pedigree Analysis

This study focuses on molecular genetics of autosomal recessive HSP. Our cohort of 145 HSP patients were directed to our laboratory from different regions of Turkey. The families included in the study were chosen according to pedigree analysis. Since incomplete penetrance is frequently observed in HSP families, pedigree analysis might be susceptible. Therefore, to minimize the error rate in family selection, the families meeting two criteria were sought. According to the first criterion, the parents should be consanguineous or from the same village. The second criterion was presence of more than one affected sibling in the family. Twenty-seven autosomal recessive HSP families meeting these two criteria were chosen for this study (Table 3.1).

4.1.2. DNA Extraction from Peripheral Blood

Ten ml blood sample drawn into tubes containing K_2EDTA was transferred to sterile 50 ml centrifuge tubes and thirty ml red blood cell (RBC) lysis buffer was added. Samples were vortexed and the blood and buffer mix were left at 4°C for 25 minutes. The mix was centrifuged at 5000 rpm, at 4°C for ten min. The supernatant was discarded and the pellet was resuspend in ten ml RBC lysis buffer. The re-suspended solution was centrifuged at 5000 rpm, at 4°C for an additional ten minutes. The supernatant was discarded and five ml nuclei lysis buffer was used to dissolve the pellet. 30 µl Proteinase K and 50 µl of 10% SDS were added and the samples were incubated overnight at 37°C. In the second day of extraction process, ten ml of 2.5 M NaCl was added and the mixture was shaken vigorously. The samples were centrifuged at 5000 rpm at room temperature for 30 min. The supernatant was taken into a new 50 ml centrifuge tube and two volumes of ice-cold absolute ethanol were added for DNA precipitation. The DNA was fished out and transferred to a sterile 1.5

ml micro-centrifuge tube. DNA pellet was air-dried and 200-500 µl of Tris-EDTA (TE) buffer was added. The DNA samples were stored at -20°C.

4.1.3. Quantitative Analysis of Extracted DNA

Nanodrop ND-1000 spectrophotometer was used to determine the concentration of the extracted DNA samples. The measurement was done at 260 nm to determine nucleic acid concentration. Absorbance on 230 and 280 nm were also measured to monitor the purity of the DNA samples. DNA samples were considered as pure when A260/A280 ratio was between 1.8 -2 and A260/A230 ratio was higher than 2. The samples that meet these purity criteria with minimum 100 ng/µl concentration were accepted as appropriate for WES and SNP genotyping. For WES, minimum one µg of sample and for SNP genotyping minimum two µg of samples were used.

4.1.4. Polymerase Chain Reaction (PCR)

The primers for the target regions were designed by using the web-based software; Primer 3 (http://seqtool.sdsc.edu) (Table 3.2, 3.3, 3.4, and 3.5).

PCR reactions included 100 ng DNA, 2.5 µl of 10X polymerase buffer, 2.0 mM MgCl2, 0.2 mM dNTPs, 0.4 mM of each primer and 1U of *Taq* polymerase in a volume of 25 µl. The program was as follows: an initial denaturation step at 95°C for 5 min, followed by 35 cycles of 30 s at 95°C, 30 s at annealing temperature, 30 s at 72°C, a final elongation at 72°C for 5 min, and cooling at 15°C for 1 min. PCR products (5µl) were loaded on 1% agarose gels to control for amplicon size with respect to an appropriate DNA ladder.

4.1.5. Mutation Screening in Spastin Gene

As a preliminary step in our studies, we screened our cohort for SPAST gene mutations to exclude the possibility that they may have causative mutations in this most commonly mutated gene. For this purpose, exons of SPAST gene of 17 index patients were sequenced directly in the University of Antwerp, Belgium. The gene was screened for mutations in

index patients from families H72, H77, H82, H98, and H99 by High Resolution Melting (HRM) and Single Strand Conformation Polymorphism (SSCP) analyses. Variants identified by any of these techniques were then confirmed by sequencing and its nature was identified.

4.1.5.1. HRM Analyses. Roche Light cycler 480 and Light cycler 480 HRM master kit were used for HRM analysis. The general procedure was as follows; PCR reactions were performed in a reaction of 20 µL containing 30 ng DNA, 10 µl 2X HRM master mix, 0,2 µmolar (final concentration) primers and PCR grade water. PCR was performed with an initial denaturation of 5 min at 95°C, followed by 35 cycles of 30 s at 95°C, 30 s at annealing temperature, and 40 s at 68°C, with a final extension of 5 min at 68°C. Primer pairs used for mutation screening in *Spastin* gene were listed in Table 3.2.

After PCR run, the HRM analysis was performed and 'Normalized and temperature-shifted melting curves' were plotted specially to see any variant in the patients. In these curves, the temperature axis of each curve is shifted to superimpose over a certain fluorescence interval to eliminate the temperature differences. When there is a variant in one PCR product, its fluorescence level differs and its curve shifts from the other products. To make a clear discrimination, 'Normalized and temperature-shifted plots' were generated based on the Normalized and temperature-shifted melting curves. In these plots, one melting curve was chosen as a reference and it becomes as horizontal line in zero level. The fluorescence differences of other melting curves are plotted according to this reference curve to generate different paths.

4.1.5.2. Single Strand Conformation Polymorphism (SSCP) Analysis. PCR products were mixed with the denaturing loading dye in 1:1 ratio. Denaturation was performed at 95°C for five minutes followed with a five minute-incubation on ice to prevent renaturation. PCR products were run on an on eight per cent acrylamide gel with glycerol. The amount of PCR product loaded on the gel was determined according to the amount of PCR product observed on 1 % agarose gel. The acrylamide gel was run for approximately 20 hours at 90 Volts, which was prolonged for larger PCR fragments.

Denatured single stranded DNA fragments assume a 3D conformation and run on the gel in single stranded conformation. If the fragment has any variant, this 3D confirmation would be different from the ones that do not bear the variant. Therefore, there would be a visible difference among the position of the DNA bands on the polyacrylamide gels after Silver staining.

4.1.5.3. Silver Staining. In order to visualize the DNA fragments separated on polyacrylamide gels silver staining method was used. The gels were incubated with $AgNO_3$ solution for ten minutes. They were rinsed with dH_2O to prevent non-specific staining. Then, the gels were treated with freshly prepared developing buffer until the bands became visible. Later, they were taken into a transparent folder for documentation.

4.1.6. Sanger Sequencing

Sanger sequencing was out-sourced (Macrogen, South Korea). The DNA region of interest to be sequenced was amplified by polymerase chain reaction (PCR) method. 50 ul PCR reaction was shipped to the company and the results were obtained as ab1 files.

4.1.7. Whole Exome Sequencing (WES)

4.1.7.1. WES and Initial Analyses. Whole exome sequencing analysis of six HSP patients (H52, H53, H57, H59, H61 and H65) was performed in the University of Miami in the scope of a collaboration with Prof. Stephan Züchner. Agilent SureSelect Human All Exon 50 Mb kit was used for exome capture, and samples were sequenced with Hiseq 2000 instruments. Reads of 100 bp length was sequenced with 50X coverage. Sequence alignment and variant calling was performed with BWA and GATK software. WES analysis of these six patients were pursued on the browser interface of the Genomes Management Application (Gonzalez *et al*., 2013b).

For the other 21 families WES analysis was out-sourced to DONE Genetics and Bioinformatics Inc (Turkey). Agilent SureSelect 50 /51 M Capture kit and Illumina NextSeq 500 sequencing system were used. DNA samples were sequenced with 50X coverage and 100 bp paired-end reads. BWA, Samtools, GATK and Variant Studio softwares were used

for alignment, filtering, variant calling and variant annotation, respectively. These initial analyses were performed by service provider and the results were delivered to our laboratory as vcf and excel files.

4.1.7.2. WES Data Analysis. WES data was analyzed for each patient separately. Initially variants in the known HSP genes was analyzed using a flexible filtering criteria; included variant to be missense, nonsense or splice site mutations; minor allele frequency (MAF) to be 1 % at maximum; and read-depth to be 5 at minimum. For a recessive inheritance pattern, variant in the patient was expected to be in homozygous state. Variants in heterozygous state were also analyzed to determine compound heterozygosity and eliminate the possibility of incomplete penetrance in our families. For the single nucleotide variants, prediction algorithms SIFT and PolyPhen were used to determine their possible deleterious and damaging effects.

In the cases where no variants were identified in the known HSP genes, the data was filtered to determine candidate disease-causing variants in the genes that are related to other neurological disorders. In total, 1771 genes were covered known to be associated with neurological disorders including SMA, dHMN, CMT, ALS, Lysosomal Storage, Dilated Cardiomyopathy, Deafness, Ataxia, Muscular Dystrophy, Pharmacogenetics, and MSeqDR-LSDB (Mitochondrial Disease Locus Specific Database). The filtering criteria included a read-depth of 30 at minimum, variant to be at homozygous state, minor allele frequency (MAF) to be 0.5 % at maximum and variant to be missense, nonsense or splice site mutation. SIFT (Kumar *et al.,* 2009) (D: deleterious) and PolyPhen (Adzhubei *et al*., 2010) (PsD: possibly damaging; PrD: probably damaging) prediction algorithms were used to select the strongest candidates when it is not an insertion or deletion. To determine the compound heterozygous mutations same criteria was used.

4.1.7.3. Verification of variant. Forward and reverse primers were designed for the candidate variants and the corresponding region was amplified by PCR in family members. PCR products were then, used for Sanger sequencing.

Analysis of the candidate variants by direct sequencing in the corresponding family members is a crucial step to unravel the segregation of the variant in the family. However,

if the number of candidate variant is high, analyzing all variants by Sanger sequencing would decrease the time- and cost-effectiveness of the study. Therefore, threshold of five was set for the number of candidate variants to keep the study's design time- and cost-effective.

4.1.8. SNP Genotyping and Homozygosity Mapping

Whole genome SNP genotyping and homozygosity mapping was performed for 11 HSP families for which the candidate variants were excluded or the number of candidate variants exceeded five. SNP genotyping is the determination of genetic variants of single nucleotide polymorphisms (SNPs) among individuals. Large-scale whole genome genotyping was performed for two affected individuals from each family using Illumina Infinium OmniExpress-24 array including 710,000 SNPs. The homozygous regions common to both affected individuals were determined for homozygosity mapping. It is a powerful approach to determine candidate loci for AR disease-causing genes in patients from families with consanguinity.

When analyzing the SNP genotyping data SNPs with MAF values lower than 0.05 were eliminated to remove rare polymorphisms from the data. The haplotypes were determined by coding each SNP allele as AA, AB and BB for each pair of sibs. When both alleles bear the same base they were converted to AA. When one of the affected individuals were heterozygous for that allele then AB code was used. As the last option, when one of the affected individuals was homozygous for the other allele, alleles were coded as BB.

When haplotype construction was completed resulting homozygous regions were examined. Regions that include 50 consecutive SNPs that are homozygous with a minimum rate of 94% and that are longer than 500 kb were accepted as homozygous regions for further studies.

4.1.9. Identification of ROHs by PLINK

SNP genotyping could not be performed for families P392, H110, H133 since DNA samples from the affected family members were not available. Therefore, WES data was used as an alternative to perform homozygosity mapping by using some bioinformatics tools.

PLINK is among the most widely used tool for this purpose (Purcell ve ark., 2007; Pippucci ve ark., 2014). Homozygous regions termed as 'Runs of Homozygosity (ROH)' were determined for each individual by PLINK. The file of ROHs was created according to the parameters listed in the Table 4.1. In cases, where ROH number is too high to analyze or where any ROH was not identified these parameters were altered. These alterations are mentioned in the corresponding parts of the thesis. ROH distribution determined for each individual were transformed into graphical representations by R software.

Table 4.1. PLINK parameters used in WES analysis.

Parameter	Value
Minimum ROH length	500 kb
Minimum number of homozygous SNPs	10
Number of SNPs per window	20
Upper limit for number of missing calls	10
Number of SNPs per window	0.05
SNP density	400 kb
Homozygosity Gap	2000 kb
Upper limit for number of heterozygous SNP per window	1

PLINK was also used to determine the ROHs from SNP genotyping data of other patients. ROHs determined via PLINK were compared with the homozygous regions determined by the SNP genotyping data and they were found to be consistent, as expected. Since SNP genotyping was performed for two affected individuals, it was possible to compare the homozygous regions between two siblings. Some of the ROHs determined by PLINK were eliminated because when SNP data was analyzed one of the affected siblings was identified to be heterozygous for several SNP markers inside the homozygous region.

Effectiveness of PLINK was also tested using the WES data of one patient for which the causative gene has been known. ROHs were determined by PLINK and tested whether any of them implicated the causative gene locus.

4.1.10. Comparing WES and SNP Genotyping Data & Searching for Candidate Genes

To determine novel candidate variants, the WES data was re-examined for variants corresponding to the homozygous regions determined by PLINK and SNP genotyping. Verification of the variants was performed as mentioned in the section 4.1.7.3.

4.1.11. 3D Protein Model Prediction

The effects of variants were analyzed on protein folding by 3D modelling algorithms iTASSER (Zhang, 2008), Phyre-2 (Kelley *et al*., 2015), Raptor-X (Källberg *et al.,* 2012), and SWISS-MODEL (Biasini *et al.,* 2014). Amino acid sequence of native and mutated proteins are uploaded to the algorithms and results were compared to reveal variances among models.

4.2. In vitro Analyses

After novel candidate HSP genes were determined, we tried to elucidate the effect of the variant on disease pathogenesis. With this aim, B-lymphocyte samples were obtained from family members and they were immortalized. Effect of variant on mRNA/protein level and localization were analyzed.

4.2.1. Immortalization of Human B-Lymphocytes

4.2.1.1. B-Lymphocyte Isolation from Blood. Ten ml blood samples were drawn into ACD-A containing blood collection tubes. Blood was then transferred into a 50 ml centrifuge tube and diluted with ten ml PBS. Ten ml ficoll was added into a new 50 ml centrifuge tube and PBS-blood mix was added on top of the ficoll very slowly. At this step, the falcon was held in 45° angle. Blood and ficoll was not mixed and tube was not inverted. The samples were then centrifuged at 1200 g, for 25 minutes. Centrifuge brake setting was adjusted to zero. After centrifugation, a white leukocyte-ring between ficoll and serum became visible. Leukocytes were transferred into a 50 ml centrifuge tube. The supernatant was removed and

the pellet was re-suspended in seven ml RBC-lysis buffer. After incubation for seven minutes, PBS was added up to 50 ml and the sample was centrifuged at 500 g, for ten minutes. The supernatant was removed and the pellet was solved in one ml freezing medium. Cells were stored at -80°C.

4.2.1.2. Epstein-Barr Virus (EBV) Production. B95-8 cells were kindly provided by Prof. Rebecca Schüle from University of Tubingen, Germany. B95-8 cells were seeded with 20 ml of lymphoblast growth media in T75 flask. In the following two days, 20 ml of growth medium was added into the flask. On fourth day, cells were split into two T75 flasks and 30 ml of growth media was added. For the next ten days, cells were left to death which triggers the production of EBV. The media was transferred to a 50 ml falcon and centrifuged for 30 minutes at 600 g. The supernatant was sterilized by 0.45 μm filter and the virus containing media was stored -80°C as two ml-aliquots.

4.2.1.3. Transduction of Human B-Lymphocytes. B-Lymphocyte aliquots were defrosted in 37°C-waterbath and centrifuged at 100 g for 2 minutes. Pellet was suspended in EBV containing supernatant and incubated in 37°C-waterbath for at least 30 minutes. Cells were transferred into a 75 cm cell culture flask with 13 ml of growth media. Fresh media was added into the flask regularly as it got yellow for two-three weeks. The transformed lymphocytes were grown in growth media mentioned in part 3.4. Cells were incubated at 37°C with 5% CO_2. For a healthy growth, cells were supplied with fresh medium as the color of the medium turn yellow.

4.2.2. Monitoring mRNA and Protein Levels

Quantitative real time PCR and Western-blotting methods were used to observe the effect of variant on lymphoblastoid cell lines of the patients.

4.2.2.1. Quantitative real-time polymerase chain reaction (qPCR). mRNA level of the target gene in lymphoblastoid cell lines was monitored by quantitative real time PCR (qRT-PCR). Total RNA was isolated from cells with RNA isolation kit, according to the instructions of the manufacturer. Concentration of the RNA samples was measured by Nanodrop. RNA was

converted to cDNA by cDNA synthesis kit, according to the instructions of the manufacturer. Real time PCR was performed using primers specific to the target gene. PCR was performed with Exicycler 96 Real-time Thermal Cycler (Bioneer, Republic of Korea). PCR included SYBR Green solution, 15 ng cDNA and 0,5 μM primer pairs listed in Table 3.2 in a total of 10 μl final volume. Reaction program included initial denaturation at 95^0C for five minutes, followed by 45 cycles of denaturation at 95^0C for 10 seconds, annealing at the corresponding temperature for 10 seconds and extension at 72°C for 10 seconds. GAPDH and Actin, ubiquitously expressed genes, were used as control and mRNA level of the target gene was compared to that of control gene. The $2^{-\Delta\Delta CT}$ method was used to calculate relative gene expression levels. This method relies on the threshold cycle (*CT*) which indicates the number of cycle where fluorescence signal detected from PCR product reaches a certain level. CT of the gene of interest is compared to reference gene to get the relative expression level.

4.2.2.2. Western Blot Analysis. Western Blotting was the method of choice to observe the protein levels of the target gene in lymphoblastoid cell lines. To extract the proteins, 5-10 ml of cells with approximately $1x10^5$/ml density were centrifuged at 13,000 g for five minutes. Medium was aspirated and cells were washed with ice-cold PBS. 100 - 300 μl of RIPA buffer containing protease and phosphatase inhibitor cocktails was added and distributed over the cells. After incubation on ice for 15 minutes, Magnalyzer seramic beads were added and the cells were sonicated by MagnaLyser homogenizer at 6500 rpm for 60 seconds. Lysate was incubated on ice for an additional 15 minutes and then, centrifuged at 13,000 g for five minutes at 4°C. Supernatant was transferred into a new micro-centrifuge tube. Bicinchoninic acid (BCA) assay was performed to determine the concentration of the protein by BCA assay kit. Total protein lysate were stored at -20°C.

Total protein extracts were loaded on 10% separating gel with 5% stacking gel on top. Concentration of the polyacrylamide gel might vary from 8 % to 12.5 % according to the size of protein of interest. Samples were mixed with in 2X or 6X sample buffer and incubated at 95^0C for five minutes for denaturation followed by an incubation on ice for five minutes. 100 μg of total protein lysates were loaded on the gel and the samples were run on the stacking gel at 80V for 20 minutes and on the polyacrylamide gel at 110V for 2.5 hours.

The samples fractionated on separating gel were electro-transferred to polyvinylidene fluoride (PVDF) membrane at 100 V. Duration of the transfer depends on the size of the protein of interest ranging from 50 – 120 minutes. Transfer buffer was used in this process. Membrane was washed in TBS-T three times for five minutes/each followed with incubation at room temperature for one hour in blocking solution with gentle shaking. Membrane was then incubated overnight at 4^0C with corresponding antibody dissolved in blocking solution. It was washed in TBS-T three times for five minutes/each and then, incubated at room temperature for one hour with HRP-tagged secondary antibody chosen depending on the host of primary antibody. Final wash of the membrane was performed with TBS-T for five minutes/each. Western blotting luminol reagent was applied for one minute and images were captured using fluorescence and chemiluminescence imaging system. The image of blot was analyzed by ImageJ software.

4.2.3. Immunocytochemical Analysis

The co-localization experiment was performed to gain a better understanding for the function of the protein of interest. For this purpose, lymphoblastoid cell lines were used. Since these cells are suspension cells, coverslips were coated with Poly-D-Lysine (PDL), which is a positively charged polymer and its used to enhance cell attachment to coverslip. For this step, 5 mg of were suspenden in 50 ml of sterile water. 100 ul of PDL solution was coated on the coverslip for five minutes. Coverlips were then rinsed with sterile water and air-dried for two hours. Cells were transferred to Poly-D-Lysine-coated coverslips that are on the wells of a 4-well plate. For fixation, cells were incubated with paraformaldehyde (PFA) (4%-pH:7.4) for two min (300 ul for 4-well plate) without removing the cell media. This step is important to enhance fixation of cells which can be easily detached. After PFA and medium mix is removed and cells were incubated with PFA (4%-pH:7.4) for 10 minutes (500 ul for 4-well plate). After fixation, coverslips are removed from 4-well plates to a clean surface. Cells were washed with 1X PBS two times for five minutes. For permeabilization, cells were incubated at -20°C for 20 minutes with -20°C-cold methanol. For blocking, cells were incubated in blocking solution #1 at room temperature for one hour. Then, incubation pursued with primary antibody diluted in blocking solution #2 overnight at 4°C. Cells were washed with PBS, two times for five minutes and were incubated with secondary antibody

in blocking solution #2 for one hour in dark. Cells were then, rinsed two times for five minutes with PBS and incubated with 0.02 ug/ml (20 uM) DAPI (for nuclear staining) in TBS for five minutes. They were rinsed two times for five minutes with PBS and once for five minutes with dH_2O. Cellular images were captured via confocal and fluorescent microscopes.

5. RESULTS

Spastin (SPAST) gene was screened for mutations as a preliminary step before WES analysis since it is the most frequently mutated gene in HSP. Then, genomic DNA of one affected individual from each of the twenty-seven AR-HSP families were analyzed by WES analysis. The results of WES, homozygosity mapping, and segregation analysis were given for each family in separate sections in the text after a description of the general approach to avoid repetitions and keep the integrity of the progressive data.

5.1. Mutation Screening in *Spastin* Gene

DNA samples from seventeen of the patients listed in the Table 3.1 (from P392 to H65) were screened for mutations in the Spastin gene by direct sequencing in University of Antwerp based on a collaboration with Prof. Peter De Jonghe. These individuals were found to be negative for mutations in this gene. Five index patients from families H72, H77, H82, H98, and H99 were screened for mutations in all 17 exons of the *SPAST* gene by High Resolution Melting (HRM) analysis. HRM analysis could not be performed for other exons since the corresponding primers that have been designed previously in our laboratory generate product sizes longer than 250 bp, not purified by HPLC or their TM were not around 60°C. The exons that could not be analyzed by HRM were investigated by SSCP analysis.

Variants were not observed in the exons 2, 4, 7, 13, and 16. For exons 9 and 17, there were non-specific amplifications resulting in two peaks (red arrows) in the melting curve analysis (Figure 5.1), thus, a reliable conclusion was not possible. Under various PCR conditions products could not be generated with the primers of the exon 14 and conditions of HRM could not be optimized. The only variant detected during HRM analysis was observed in exon 11 for patient H82.3 (Figure 5.1-red star).

The exons 1a, 1b, 3, 5, 6, 8, 10, 12, 15 were screened by SSCP analysis since their primers give products longer than 250 bp. The exons 9, 14 and 17, for which the conditions could not be optimized for HRM analysis, were also analyzed by SSCP. Except exon 9 for patient 99.3 no variants could be observed in the SSCP analysis (Figure 5.2, red arrow).

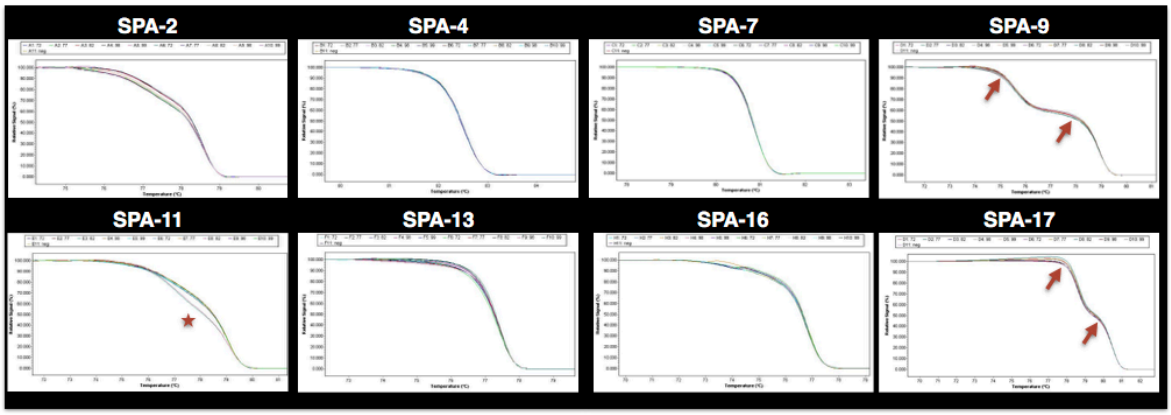


Figure 5.1. Normalized and temperature-shifted melting curves for exons 2, 4, 7, 9, 11, 13, 16, and 17 for patients H72.3, H77.4, H82.3, H98.3, H99.3.

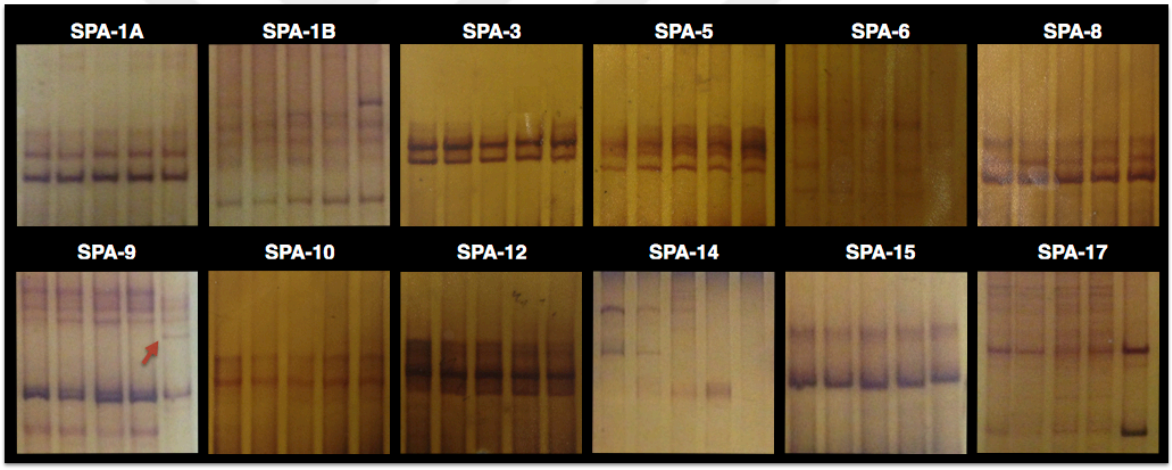


Figure 5.2. Eight per cent SSCP gel photographs showing the SPA mutation screening in exons 1a, 1b, 3, 5, 6, 8, 9, 10, 12, 14, 15 and 17 for patients H72.3, H77.4, H82.3, H98.3, and H99.3 (Red arrow indicates different banding pattern).

In conclusion, H72.3, H77.4, H82.3, H98.3, and H99.3 were found to be negative for mutations for all exons of SPAST except exon 9 for patient H99.3 and exon 11 for patient H82.3. Sanger sequencing was performed to identify the variant in these patients.

Patient H99.3 has a pure HSP phenotype with a late onset implicating a high probability for *Spastin* mutation. Direct sequencing was performed for exon 9 of this patient, however, we could not identify any variants in this exon, implicating a false positive result in SSCP analysis. The different banding pattern might be due to a non-specific PCR product.

Considering the fact that mutations in *SPAST* are mainly associated with late-onset pure HSP and the patient presents a complicated HSP phenotype with early-onset, the probability of having a mutation in this gene was very low for the patient H82.3. A heterozygous deletion c.1222-38_1222-37del variant was observed in intron 10 of *SPAST* gene that was neither listed as polymorphism nor mutation in literature. When segregation of this variant was analyzed family members including the unaffected individuals were found to carry this variant in a heterozygous state except the father (Figure A.1). Besides, Human Splicing Finder bioinformatics tool defined it as a possible enhancer motif for an SR protein SC35, with a variant value of 78 % (Figure 5.3). The indicated strength of the motif was not a strong value. Therefore, it was concluded that the variant cannot be responsible for the disease.

cDNA Position	Linked SR protein	Enhancer motif	Value 0-100 (variation)
c.1322-38	SC35	ggtcttta	78.49 (+13.78 %)

Figure 5.3. Calculated effect of c.1222-38_1222-37del variant on splicing via Human Splicing Finder tool.

Index patients from families H108, H110, H133, H142 and H143 were not screened for mutations in *SPAST* gene because these families arrived to our laboratory after these initial analyses and WES were applied directly to these patients due to time limitations.

5.2. Molecular Genetic Analysis

WES was performed for six of these individuals in the scope of a partnership with Prof. Stefan Züchner's laboratory in the University of Miami, USA. WES data of twenty-seven patients were initially analyzed for the non-synonymous variants in the known HSP genes. In the cases where no variants were identified in the known HSP genes, the data was filtered to determine candidate disease-causing variants in the genes that are related to other neurological disorders. These genes include: SMA, dHMN, CMT, ALS, Lysosomal Storage, Dilated Cardiomyopathy, Deafness, Ataxia, Muscular Dystrophy, Pharmacogenetics, and MSeqDR-LSDB (Mitochondrial Disease Locus Specific Database) genes that makes 1771 genes in total. The filtering criteria included a read-depth of 30 at minimum, variant to be at homozygous state, minor allele frequency (MAF) to be 0,5% at maximum and variant to be

missense, nonsense or splice site mutation. SIFT (D: deleterious) and PolyPhen (PsD: possibly damaging; PrD: probably damaging) prediction algorithms were used to select the strongest candidates when the variant is not an insertion or deletion. Segregation analyses of these candidate variants was performed if the number of candidates was lower than five. Otherwise, linkage analysis and homozygosity mapping was performed to determine the disease-associated region.

5.2.1. Confirmation of PLINK Tool-Set and Its Comparison to SNP Genotyping data

PLINK program was used to determine homozygous regions in WES data of patients, for which SNP genotyping could not be performed. To test the applicability of this program on our WES data set, PLINK was used on the WES data of H98.3 known to have a causative mutation in CYP7B1 gene. Runs of homozygosity (ROHs) were determined by PLINK and it was observed that CY7B1 gene (Chromosome 8: 64,587,763-64,798,761) co-localizes with the ROH identified on chromosome eight (Figure 5.4, 5.5). Columns in Figure 5.4 represent, chromosome number (A), initial (B) and final (C) positions of ROH on the chromosome, ROH length (D), number of SNPs identified in the ROH (E), SNP density (F), homozygosity (G) and heterozygosity (H) percentages of the SNPs in the ROH, respectively. This finding showed that PLINK might be used as a powerful tool to identify novel genes in our samples.

When PLINK was applied on WES data of other patients, ROHs identified were found to be consistent with the homozygous regions determined by the SNP genotyping data. Since SNP genotyping was performed for sib-pairs, it allowed the comparison of two siblings' data with that of PLINK data and differences in homozygosity were observed among siblings. To explain this situation, a ROH determined from the WES data of H6.1 was compared to SNP data of H6.2 for the corresponding locus as shown in Figures 1 and 2. For ROH between the nucleotide position 41486245 and 63270823 in chromosome one (third line of Figure 5.6), patient H6.1 was homozygous as expected (Figure 5.7, green highlighted region coded as AA), however, the affected brother H6.2 was heterozygous after the nucleotide position 57572121 (Figure 5.7, orange highlighted regions coded as AB). Therefore, ROH common to both affected siblings should end in the position 57572121. The same procedure explained for family H6 were also used for other families and the homozygous regions identified from

SNP genotyping data were compared in between the affected sibs. The limiting nucleotide positions that are common to both affected siblings were given in this section of the thesis for simplicity.

	A	B	C	D	E	F	G	H
1	CHR	POS1	POS2	KB	NSNP	DENSITY	PHOM	PHET
2	1	54666074	55474325	808.252	23	35.141	0.870	0.043
3	1	59127086	66102755	6.975.670	38	183.570	0.921	0.053
4	1	89637578	90459744	822.167	20	41.108	0.850	0.050
5	1	245861571	249211517	3.349.947	83	40.361	0.723	0.048
6	2	60687959	62081012	1.393.054	11	126.641	0.818	0.000
7	2	105885961	112614429	6.728.469	24	280.353	0.750	0.083
8	2	136575534	142888383	6.312.850	16	394.553	0.875	0.063
9	2	158157361	161174693	3.017.333	31	97.333	0.871	0.065
10	4	4199634	6623032	2.423.399	37	65.497	0.757	0.054
11	4	68436883	73012745	4.575.863	47	97.359	0.872	0.043
12	4	73942678	76957171	3.014.494	27	111.648	0.926	0.037
13	7	101194424	102854546	1.660.123	20	83.006	0.750	0.050
14	7	114332234	117824239	3.492.006	14	249.429	0.857	0.071
15	7	120610901	123672456	3.061.556	18	170.086	0.944	0.056
16	8	62626879	70978690	8.351.812	22	379.628	0.909	0.045
17	10	11505201	13214427	1.709.227	23	74.314	0.783	0.087
18	11	55110967	56086147	975.181	33	29.551	0.909	0.061
19	11	57563991	58207204	643.214	21	30.629	0.857	0.048
20	11	59382903	60635091	1.252.189	26	48.161	0.885	0.077
21	12	10241829	10746385	504.557	24	21.023	0.875	0.083
22	16	57071114	57937895	866.782	21	41.275	0.714	0.048

Figure 5.4. ROHs identified in the WES data of H98.3.

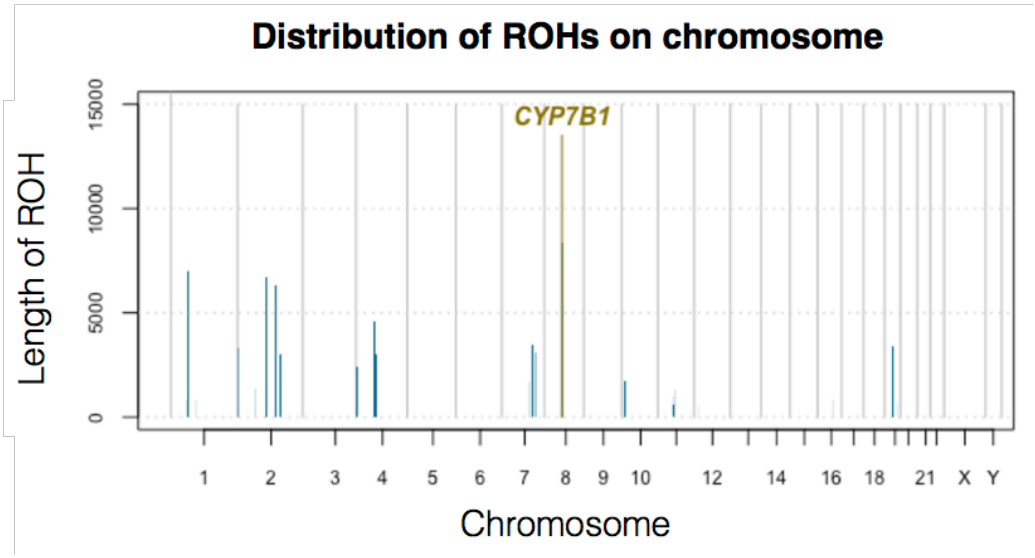
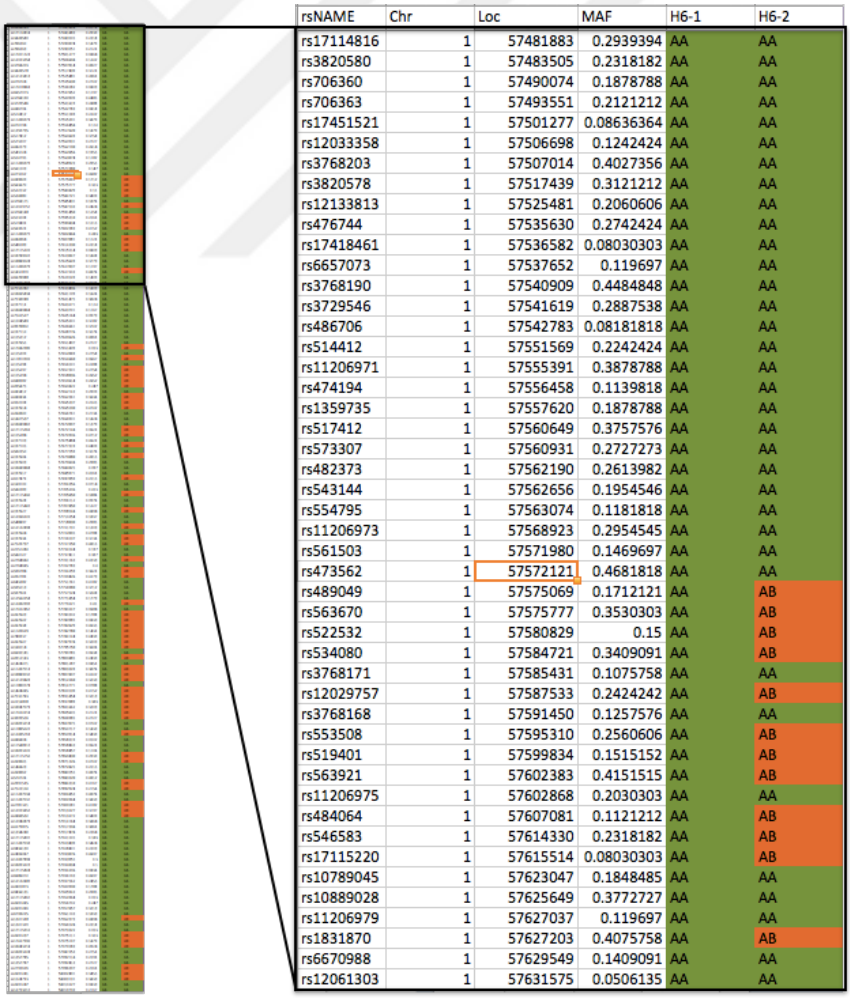


Figure 5.5. Distribution of ROHs determined on WES data of H98.3.

	A	B	C	D	E	F	G	H	I	J
1	CHR	POS1	POS2	KB	NSNP	DENSITY	PHOM	PHET	POS1-SNP	POS1-SNP
2	1	7906476	12251808	4345.333	67	64.856	0.642	0.045	10920000	12251808
3	1	41486245	63270823	21784.579	131	166.294	0.702	0.038	41486245	57572121
4	1	166958601	178483231	11524.631	73	157.872	0.726	0.055	170300000	171300000
5	2	207833983	216262621	8428.639	66	127.707	0.788	0.045	209000000	211900000
6	2	216283958	225244923	8960.966	82	109.28	0.512	0.037	217342357	220752532
7	8	139165272	144940230	5774.959	95	60.789	0.568	0.032	139405963	140497775
8	10	89720907	99991408	10270.502	94	109.261	0.777	0.043	89720907	92869911
9	11	1993975	5269572	3275.598	60	54.593	0.667	0.033	-	-
10	11	7071031	18267463	11196.433	123	91.028	0.813	0.033	7714238	7918294
11	12	105600935	110029186	4428.252	51	86.828	0.686	0.039	-	-
12	13	31233063	53282749	22049.687	144	153.123	0.771	0.021	34005964	34826347
13	14	20012850	21250124	1237.275	54	22.913	0.741	0.037	-	-
14	18	2771516	10696227	7924.712	50	158.494	0.72	0.06	2996922	3597864
15	20	39797465	49621149	9823.685	80	122.796	0.75	0.038	39797465	49621149

Figure 5.6. Update of ROH limiting positions after comparison of PLINK result to SNP genotyping data.



rsNAME	Chr	Loc	MAF	H6-1	H6-2
rs17114816	1	57481883	0.2939394	AA	AA
rs3820580	1	57483505	0.2318182	AA	AA
rs706360	1	57490074	0.1878788	AA	AA
rs706363	1	57493551	0.2121212	AA	AA
rs17451521	1	57501277	0.08636364	AA	AA
rs12033358	1	57506698	0.1242424	AA	AA
rs3768203	1	57507014	0.4027356	AA	AA
rs3820578	1	57517439	0.3121212	AA	AA
rs12133813	1	57525481	0.2060606	AA	AA
rs476744	1	57535630	0.2742424	AA	AA
rs17418461	1	57536582	0.08030303	AA	AA
rs6657073	1	57537652	0.119697	AA	AA
rs3768190	1	57540909	0.4484848	AA	AA
rs3729546	1	57541619	0.2887538	AA	AA
rs486706	1	57542783	0.08181818	AA	AA
rs514412	1	57551569	0.2242424	AA	AA
rs11206971	1	57555391	0.3878788	AA	AA
rs474194	1	57556458	0.1139818	AA	AA
rs1359735	1	57557620	0.1878788	AA	AA
rs517412	1	57560649	0.3757576	AA	AA
rs573307	1	57560931	0.2727273	AA	AA
rs482373	1	57562190	0.2613982	AA	AA
rs543144	1	57562656	0.1954546	AA	AA
rs554795	1	57563074	0.1181818	AA	AA
rs11206973	1	57568923	0.2954545	AA	AA
rs561503	1	57571980	0.1469697	AA	AA
rs473562	1	57572121	0.4681818	AA	AA
rs489049	1	57575069	0.1712121	AA	AB
rs563670	1	57575777	0.3530303	AA	AB
rs522532	1	57580829	0.15	AA	AB
rs534080	1	57584721	0.3409091	AA	AB
rs3768171	1	57585431	0.1075758	AA	AA
rs12029757	1	57587533	0.2424242	AA	AB
rs3768168	1	57591450	0.1257576	AA	AA
rs553508	1	57595310	0.2560606	AA	AB
rs519401	1	57599834	0.1515152	AA	AB
rs563921	1	57602383	0.4151515	AA	AB
rs11206975	1	57602868	0.2030303	AA	AA
rs484064	1	57607081	0.1121212	AA	AB
rs546583	1	57614330	0.2318182	AA	AB
rs17115220	1	57615514	0.08030303	AA	AB
rs10789045	1	57623047	0.1848485	AA	AA
rs10889028	1	57625649	0.3772727	AA	AA
rs11206979	1	57627037	0.119697	AA	AA
rs1831870	1	57627203	0.4075758	AA	AB
rs6670988	1	57629549	0.1409091	AA	AA
rs12061303	1	57631575	0.0506135	AA	AA

Figure 5.7. Comparison of ROH determined by PLINK with SNP genotyping data of H6.1 and H6.2.

5.2.2. Family P392

P392 was the index patient that was selected for WES analysis. Analysis revealed that there was a missense c.1388T>C variant in *SPG11* gene in heterozygous state. This variant is a known benign polymorphism (rs3759871) with a MAF value of 0.49%. Since the number of candidate variants determined in neurological disorder-genes after stringent filtering exceeded the threshold of five candidates, they were not further analyzed by direct sequencing.

Homozygous regions for this patient was determined by the PLINK software using the available WES data instead of SNP genotyping as the DNA sample of the affected brother was not available. With this software, it is possible to determine candidate variants that might be eliminated initially during filtration of WES data. By applying the parameters mentioned in the Table 4.1 in section 4.1.9, we ended up with 39 ROHs. As the numbers of ROHS was high to analyze via direct sequencing, the parameters were altered as follows; ROH being 1000 kb minimum, including 50 SNP at least, having at least one SNP in every 200 kb and if the distance between two ROHs located is higher than 4000 kb that should be considered as different ROHs. With new parameter settings, seven ROHs have been identified for P392 (Figure 5.8, 5.9).

	A	B	C	D	E	F	G	H
1	CHR	POS1	POS2	KB	NSNP	DENSITY	PHOM	PHET
2	3	170825905	186572089	15746.185	96	164.023	0.729	0.021
3	8	11996150	24771488	12775.339	117	109.191	0.744	0.034
4	12	99640557	109883374	10242.818	67	152.878	0.851	0.045
5	12	23915547	32908705	8993.159	77	116.794	0.805	0.039
6	2	43625184	50201110	6575.927	66	99.635	0.788	0.061
7	21	34925659	40969826	6044.168	57	106.038	0.754	0.053
8	19	48284356	53086598	4802.243	133	36.107	0.669	0.045

Figure 5.8. ROHs identified in the WES data of P392.

To determine the candidate variants in the regions given in the Figure 7, WES data was analyzed (Table 5.1).

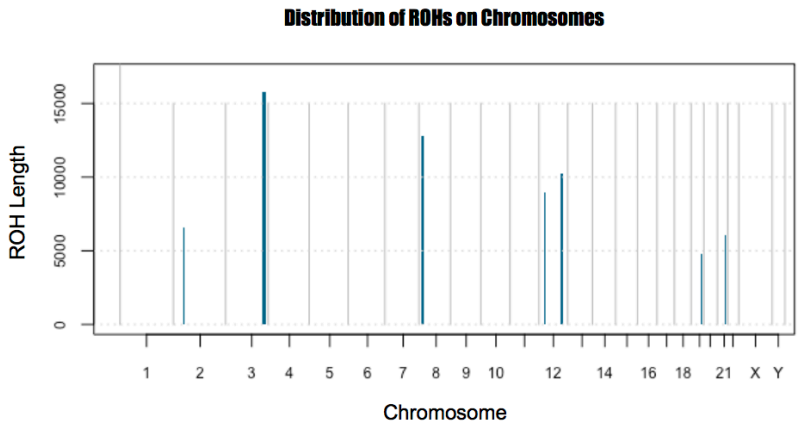


Figure 5.9. Distribution of ROHs identified for the patient P392.

Table 5.1. Candidate variants identified in the WES data of patient P392.

Gene	Variant	Position	Read Depth	Variant Type	SIFT	PolyPhen	dbSNP ID
KCNMB3	c.157G>A	3:178968634	37	Missense	D (0.05)	Pr_D (0.99)	rs7645550
SLC7A2	c.47G>A	8:17396380	70	Missense		Ps_D (0.453)	rs13259948
PDGFRL	c.980T>A	8:17500162	42	Missense	D (0)	Ps_D (0.772)	rs35346456
NAT2	c.590G>A	8:18258103	9	Missense	D (0.05)	Pr_D (0.992)	rs1799930
SCYL2	c.1070C>T	12:100708367	50	Missense	D (0.01)	Pr_D (0.95)	rs33968174
UTP20	c.1505C>G	12:101693534	18	Missense	D (0.01)	Ps_D (0.757)	rs4764643
MYBPC1	c.1443C>G	12:102045163	50	Missense	D (0.01)	Pr_D (0.999)	rs3817552
CAPRIN2	c.1906-6C>A	12:30877391	38	Splice site			
OVOS2	c.2783T>C	12:31285214	40	Missense			rs2287450
CCDC114	c.136C>T	19:48821757	27	Missense	D (0)	Pr_D (0.999)	rs16981988
ETFB	c.734C>T	19:51850290	47	Missense	D (0.02)	Ps_D (0.866)	rs1130426

D: deleterious, Pr_D:probably damaging, Ps_D:possibly damaging

Splice site variants were analyzed using Human Splicing Finder online tool to unravel their possible effect on gene expression. The splice site variants that are calculated not to have any effect on splicing were eliminated from further studies. The list of the eliminated splice site variants identified in the WES data of P392 is given as an example, in Table 5.2. Splice site variants dismissed in other families are not given in this report for the purpose of simplicity.

Table 5.2. The eliminated splice site variants identified in the WES data of P392.

Gene	Variant	Position	dbSNP ID
SCYL2	c.2145+4A>G	12:100731288	rs199970860
FGD4	c.1636-7T>C	12:32778581	rs11052113
BRWD1	c.2767-5G>C	21:40604429	rs2836950
ACPT	c.779-7T>C	19:51297138	rs45533637

Variants listed in Table 5.2, were analyzed in the family members by Sanger sequencing. Results for each variant and a table summarizing all data obtained by the sequencing were given below (Figure A.2-A.6, Table 5.3.).

Table 5.3. Sanger sequencing results for family P392.

Gene	P392.1 (index patient)	P392.2 (unaffected brother)	P392.3 (unaffected brother)	P392.4 (unaffected daughter)	P392.5 (unaffected daughter)	Conclusion
KCNMB3	var;var	+;+	+;+	+;var	+;var	Include
SLC7A2	var;var	+;var	var;var	+;var	+;var	Exclude
PDGFRL	var;var	var;var	var;var	var;var	var;var	Exclude
NAT2	var;var	+;var	var;var	+;var	+;var	Exclude
SCYL2	var;var	var;var	+;+	+;var	+;var	Exclude
UTP20	var;var	var;var	+;var	var;var	var;var	Exclude
MYBPC1	var;var	var;var	+;+	+;+	+;+	Exclude
CAPRIN2	var;var	var;var	+;+	+;var	+;var	Exclude
OVOS2	+;var	+;var	+;var	+;var	+;var	Exclude
CCDC114	var;var	+;var	+;+	+;var	+;var	Include
ETFB	var;var	+;var	var;var	+;var	+;var	Exclude

+: native allele, var: variant allele

According to these results, variants in *KCNMB3* and *CCDC114* genes were segregating with the disease in the family. However, since both variants have been reported in population databases in homozygous state, these variants were excluded and the analysis was finalized for this family.

5.2.3. Family P463

P463 was the index patient and WES analysis was performed for this patient. Analysis revealed a stop-gained c.1235C>G, p.S412X mutation in *SPG11* gene (Figure A.7). The mutation was previously reported (Stevanin *et al*., 2008). Sanger sequencing analysis revealed that both of the affected individuals are homozygous for the variant. Since patient has complicated HSP with the symptoms of pes cavus, mental retardation, atrophy in carpus callosum and cortical atrophy, the finding was compatible with the *SPG11* phenotype.

When the HSP genes were analyzed for compound heterozygous variants, two variants in *ZFYVE26* gene are observed (Table 5.4).

Table 5.4. Compound heterozygous variants in *ZFYVE26* gene for patient P463.

Gene	Variant	Position	Variant class	SIFT	PolyPhen
ZFYVE26	c.3794T>A	14:68250075	missense variant	D	PrD
ZFYVE26	c.985A>G	14:68273294	missense variant	D	PrD

D: deleterious; PrD: probably damaging.

Analysis of the variant c.3794T>A, p.L1265Q by Sanger sequencing revealed that all individuals in the family were homozygous for the native allele eliminating the possibility of compound heterozygosity (Figure A.8). It was not surprising to find out that c.3794T>A, p.L1265Q variant was a false positive because read depth of this variant was eight in the WED data. To minimize false positive data, it is required to filter out the variants with a read depth lower than 30, however, read depth filtering was not performed for HSP genes in this study not to miss any variants.

When the other candidate in *ZFYVE26* gene was analyzed by direct sequencing, it was observed that affected individuals were heterozygous for the variant (Figure A.9). Since the clinical findings in autosomal recessive *SPG15*-related patients reported in literature show similarities with affected individuals in P463 family, the possible additive effect of this variant in disease manifestation might be considered.

When all other possible candidate variants in the HSP genes were ruled out, 3D modelling algorithms were used to compare the native and mutant proteins to understand the effect of variant in *SPG11* gene. 3D prediction of SWISS-MODEL and RaptorX algorithms for native and mutant protein and their superimposition are given in Figure B.1 and Figure B.2. SWISS-MODEL prediction focuses on the protein sequence partially when sequence is longer than 1500 amino acids as observed in SPG11 protein. However, RaptorX protein can predict amino acid sequences up to 2500. The effect of the truncation on the protein was observed better in Raptor X result with a higher RMSD (root-mean-square deviation) value. As can be seen from the Figure 5.10, functional domain of the protein is conserved and truncated version of the protein results in the loss of C-terminal of the protein. It might easily be concluded that truncated protein might be involved in disease pathogenesis. According to ACMG-AMP standards and guidelines this variant falls into PVS1 (very strong pathogenicity) category indicating effect of the variant (Richards *et al*., 2015). The analysis was finalized by concluding that c.1235C>G, p.S412X mutation is associated with the disease in family P463.

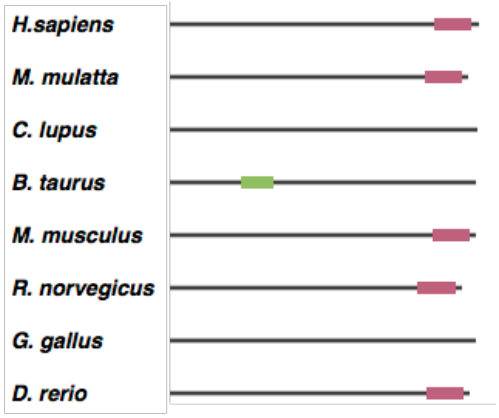


Figure 5.10. Evolutionary conservation of SPG11 protein.

5.2.4. Family P627

WES analysis was performed for the patient P627 who is the index patient. Analysis revealed that there was not any non-synonymous variant in known HSP genes neither in homozygous nor compound heterozygous state. Candidate variants selected in neurological disorder genes are listed in Table 9.

Table 5.5. Candidate variants for patient P627.

Gene	Variant	Position	Variant class	SIFT	PolyPhen	dbSNP ID
SACS	c.637G>A	13:23930114	missense	D	PsD	-
CES1	c.1087-3delT	16:55844924	splice region	-	-	-

D: deleterious; PsD: possibly damaging; PrD: probably damaging.

Sanger sequencing was performed for the index patient and c.1087-3delT variant was determined in heterozygous state. However, the other allele was found be carrying a c.1087-3delTT variant in heterozygous state (Figure A.10).

In silico tools that predict the potential splice-sites were used to analyze the effect of these variants. Only Maximum Entropy Modeling of Short Sequence Motifs with Applications to RNA Splicing Signals (Yoe and Burge, 2004) predicted the variants as a potential splicing motif. Sequence without any mutation had a potential splice site score of 7.32, while c.1087-3delT mutation lead to a score of 7.06. Moreover, c.1087-3delTT variant caused the lowest score as 5.66. Although the effects of the variants do not seem to cause a dramatic effect on the sequence, they might have an additive effect on disease manifestation.

The variant in *SACS* gene was verified in the P627 by Sanger sequencing (Figure A.11). Since DNA samples from other family members were not available, segregation analysis in the family could not be completed.

SACS gene is associated with Charlevoix-Saguenay type spastic ataxia, a childhood onset disease which is characterized cerebellar ataxia, peripheral neuropathy and pyramidal tract signs. Patients with late-disease-onset and peripheral neuropathy as the initial symptom

were also reported (Baets ve ark., 2010). P627 was initially diagnosed with peripheral neuropathy at the age of 30 and her phenotype included cerebellar findings, vibration sense loss and bilateral dysmetria. They indicated a possibility that she has an atypical Charlevoix-Saguenay type Spastic ataxia. With the conformation of the clinician, the variant identified in the *SACS* gene was regarded as the causative variant and the analysis was finalized in this family.

According to 3D model predictions performed by SWISS-MODEL, Phyre-2 (Kelley *et al*., 2015) and RaptorX, variant did not cause a structural difference on protein (Figure B.3). However, Glycine residue in 213th position is highly conserved (Figure 5.11) and conversion of this simplest amino acid to Arginine, a charged amino acid, might have an effect on interaction with other proteins or its ligands (Figure B.3).

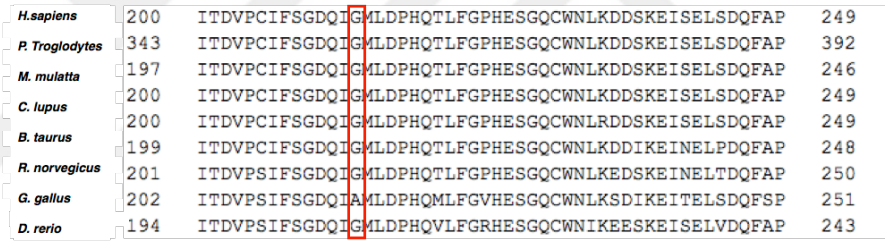


Figure 5.11. Conservation of Glycine amino acid in 213th position in SACS protein.

5.2.5. Family H6

Index patient was H6.1 and WES analysis of this patient revealed that there was not any non-synonymous variant in known HSP genes. Candidate variants selected in neurological disorders are listed in Table 5.6.

Table 5.6. Candidate variants in neurological disorders for patient H6.1.

Gene	Variant	Position	Variant class	SIFT	PolyPhen	dbSNP ID
SDHAP2	n.387A>G	3:195390505	Non coding exon variant	-	-	rs200352917
ANKRD20A4	c.1317-3C>T	9:69421850	Splice region	-	-	rs201612355

Sanger sequencing of the fourth exon of *SDHAP2* gene showed that n.387A>G was in heterozygous state in H6.1. Direct sequencing analysis of the c.1317-3C>T variant *ANKRD20A4* gene revealed that it was a false positive because affected individuals in the family were homozygous for the native allele. Therefore, this variant was excluded (Figure A.12).

Other candidate variants were determined after stringent filtering of the WES data in patient H6.1 (Table 5.7).

Table 5.7. Candidate variants for patient H6.1 after stringent filtering.

Gene	Variant	Position	Variant class	SIFT	PolyPhen	dbSNP ID
OR11H12	c.719T>G	14:19378312	Missense	D	PsD	rs201880193
RGPD8	c.4708G>A	2:113145814	Missense	D	PrD	rs200737546

D: deleterious; PsD: possibly damaging; PrD: probably damaging.

Sanger sequencing of the c.719T>G variant in *OR11H12* gene and c.4708G>A in *RGPD8* gene showed that affected individuals were heterozygous for these variants (Figure A.13). Therefore, theses variants were excluded for this family.

SNP genotyping and homozygosity mapping was performed for H6.1 and H6.2. Eleven ROHs shared by two affected siblings were determined (Figure 5.12).

	A	B	C	D	E	F	G	H
1	CHR	POS1	POS2	KB	NSNP	DENSITY	PHOM	PHET
2	1	10920000	12251808	4345.333	67	64.856	0.642	0.045
3	1	41486245	57572121	21784.579	131	166.294	0.702	0.038
4	1	170300000	171300000	11524.631	73	157.872	0.726	0.055
5	2	209000000	211900000	8428.639	66	127.707	0.788	0.045
6	2	217342357	220752532	8960.966	82	109.28	0.512	0.037
7	8	139405963	140497775	5774.959	95	60.789	0.568	0.032
8	10	89720907	92869911	10270.502	94	109.261	0.777	0.043
9	11	7714238	7918294	11196.433	123	91.028	0.813	0.033
10	13	34005964	34826347	22049.687	144	153.123	0.771	0.021
11	18	2996922	3597864	7924.712	50	158.494	0.72	0.06
12	20	39797465	49621149	9823.685	80	122.796	0.75	0.038

Figure 5.12. ROHs identified for the patients in family H6.

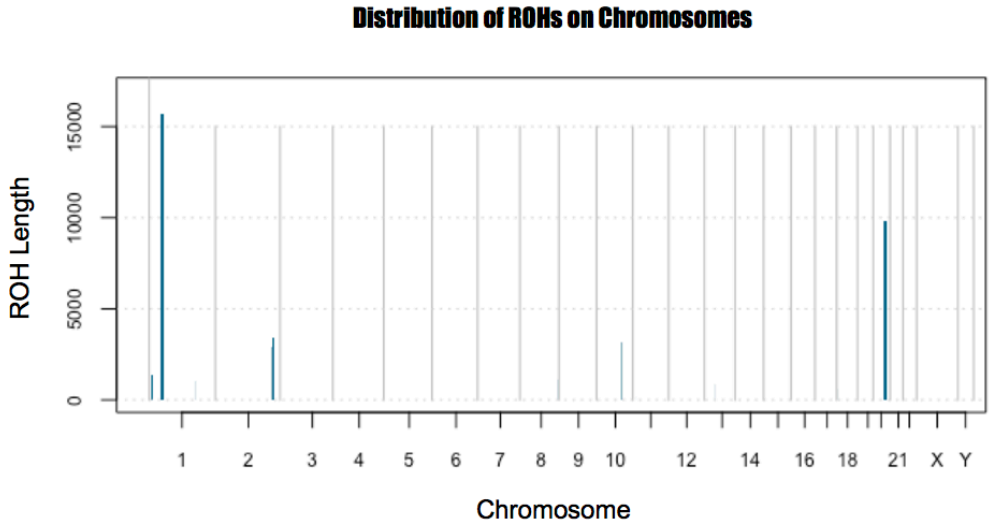


Figure 5.13. Distribution of ROHs identified for the patients in family H6.

WES data of H6.1 was analyzed for homozygous regions and the new candidate variants were determined (Table 5.8).

Table 5.8. Candidate variants identified in the WES data of H6.1.

Gene	Variant	Coordinate	Variant type	SIFT	PolyPhen	dbSNP ID
GUCA2A	c.20C>T	1:42630364	Missense	D (0.01)	Unknown	rs2071499
SLC5A9	c.1852G>T	1:48708228	Stop codon			rs850763
PRKAG3	c.1018C>T	2:219691801	Missense	D (0.02)	Ps_D (0.661)	rs33985460
SLC23A3	c.443-6T>C	2:220033830	Splice site			
RESP18	c.90G>T	2:220197388	Missense	D (0)	Ps_D (0.793)	rs2385404
IFIT5	c.5+6G>T	10:91174566	Splice site			
MYOM1	c.3576-5C>T	18:3100429	Splice site			
SPINT3	c.161T>C	20:44141400	Missense	D (0.01)	Pr_D (0.999)	rs73305953

D: deleterious; PsD: possibly damaging; PrD: probably damaging.

Segregation analyses were performed for new candidate variants and results are given in Figure A.14-A.17 and Table 5.9.

Table 5.9. Sanger sequencing results for family H6.

Gene	H6.1 (Index patient)	H6.2 (Aff. sister)	H6.3 (Unaf. mother)	H6.4 (Unaff. father)	H6.5 (Unaf. brother)	H6.6 (Unaf. sister)	Conclusion
GUCA2A	var;var	var;var	var;var	var;var	var;var	var;var	Exclude
SLC5A9	var;var	var;var	+;var	+;var	+;+	+;var	Include
PRKAG3	var;var	+;var	+;var	+;var	+;var	+;var	Exclude
SLC23A3	var;var	+;var	var;var	+;var	+;var	+;var	Exclude
RESP18	var;var	+;var	+;var	+;var	+;var	+;var	Exclude
IFIT5	var;var	+;var	+;var	+;var	+;+	+;+	Exclude
MYOM1	var;var	var;var	+;var	var;var	var;var	+;var	Exclude
SPINT3	var;var	var;var	+;var	+;var	+;var	+;+	Include

Unaf: unaffected, Aff: affected, +: native allele, var:variant allele

Variants in *SPINT3* and *SLC5A9* genes were segregating with the disease in the family. However, these variants could not be the main reason underlying disease pathogenesis because they were reported in population databases in homozygous state. Analyses was finalized for this family.

5.2.6. Family H16

H16.1 was selected for WES analysis as the index patient. After analysis of WES data, a frameshift c.6215_6219dupAGAT (p.Phe2074ArgfsTer15) insertion was determined in *SPG11* gene. Moreover, when the HSP genes are checked for compound heterozygous variants, two variants in *C19orf12* gene were observed (Table 5.10).

Table 5.10. Compound heterozygous variants in C19orf12 gene for patient H16.

Gene	Variant	Position	Variant class	SIFT	PolyPhen	dbSNP ID
C19orf12	c.*3745G>A	19:30189874	3'-UTR	-	-	rs140771696
C19orf12	c.424A>G	19:30193654	Missense	D	B	rs146170087

D: deleterious; B: benign.

Sanger sequencing analysis revealed that affected individuals H16.1 and H16.2 were homozygous for the micro-duplication in *SPG11* gene (Figure A.18).

Compound heterozygosity in *C19orf12* gene was verified in patient H16.1 in Sanger sequencing analysis by showing the heterozygous state of the c.*3745G>A variant and c.424A>G variant. However, relation of these variants to the disease was excluded by showing that patient H16.2 was homozygous for both candidate variants in *C19orf12* gene (Figure A.19).

The clinical phenotypes of the patients (TCC, cerebellar findings, hearing loss) completely fit in the *SPG11*-related HSP phenotype reported in the literature. Besides, according to ACMG-AMP guideline (Richards *et al*., 2015), this variant belongs to PVS1 category indicating that it might have high impact on disease pathogenesis. Variant was not found netiher in ExAC nor in 100G databases. 3D model algorithms did not reveal informative results because the templates used to build models did not match probably due to misconception of truncated version of the protein. The analyses were finalized for family H16 by considering c.6215_6219dupAGAT variant in *SPG11* as the disease-causing variant.

5.2.7. Family H28

Index patient H28.1 was selected for WES analysis. A missense c.1388T>C variant was determined in *SPG11* gene in heterozygous state (same variant observed in the P392). This variant is a known benign polymorphism (rs3759871) with a MAF value of 0.49%. When the HSP genes are checked for compound heterozygous variants, two variants in *AP5Z1* gene are observed. Both of the variants were observed in heterozygous state in all individuals (including affected and unaffected) analyzed in the family (Figure A.20). Therefore, the variants in *AP5Z1* gene were excluded.

SNP genotyping and homozygosity mapping was performed and 14 ROHs were determined for H28.1 and H28.2 (Figure 5.14). WES data of H28.1 was analyzed to determined candidate variants in ROHs (Table 5.11).

	A	B	C	D	E	F	G	H
1	CHR	POS1	POS2	KB	NSNP	DENSITY	PHOM	PHET
2	1	150531008	151108137	577.13	20	28.857	0.8	0.05
3	2	233723182	240029736	6306.555	95	66.385	0.663	0.042
4	6	31729359	32487398	758.04	74	10.244	0.824	0.041
5	6	79679577	80749972	1070.396	11	97.309	0.909	0.091
6	6	83949261	89388023	5438.763	25	217.551	0.8	0.04
7	6	111919108	119526049	7606.942	41	185.535	0.878	0.073
8	6	123130436	129612808	6482.373	26	249.322	0.692	0.077
9	10	86131059	90775291	4644.233	40	116.106	0.7	0.05
10	10	129249662	131506283	2256.622	22	102.574	0.818	0.045
11	13	75814018	79175788	3361.771	25	134.471	0.88	0.08
12	16	83065664	86371123	3305.46	51	64.813	0.667	0.039
13	21	45077865	45971023	893.159	36	24.81	0.722	0.056
14	21	46021179	46931002	909.824	41	22.191	0.707	0.049
15	21	46931004	47613004	682.001	36	18.944	0.583	0.056

Figure 5.14. ROHs identified for the patients in family H28.

Table 5.11. Candidate variants identified in the WES data of H28.1.

Gene	Variant	Position	Type of Variant	SIFT	PolyPhen	dbSNP ID
CERS2	c.411-4C>A	1:150940357	Splice site			rs6656535
CERS2	c.411-5G>A	1:150940358	Splice site			rs59988025
COL6A3	c.9206C>T	2:238243292	Missense		Unknown	rs1131296
COL6A3	c.3751G>A	2:238280909	Missense		Unknown	rs199646208
TNXB	c.8111G>A	6:32024395	Missense	D (0.01)	Pr_D (0.999)	rs146899254, rs10947230
TNXB	c.7235C>T	6:32029431	Missense	D (0)	Pr_D (1)	rs146464391, rs12524664
LAMA4	c.3239G>A	6:112460365	Missense	D (0)	Pr_D (0.999)	rs41289902
GPRC6A	c.431T>G	6:117130544	Missense	D (0)	Pr_D (0.972)	rs28360548
FAM35A	.730A>G	10:88911841	Missense	D (0.01)	Pr_D (0.937)	rs77879311
NUTM2D	c.103C>T	10:89118125	Missense	D (0)	Ps_D (0.862)	rs77153116
ATAD1	c.701T>C	10:89530788	Missense	D (0)	Pr_D (0.982)	
LIPF	c.511A>G	10:90429652	Missense	D (0)	Ps_D (0.86)	rs814628
HSDL1	c.980C>G	16:84158248	Missense	D (0.01)	Ps_D (0.493)	rs4378600

D: deleterious; PsD: possibly damaging; PrD: probably damaging.

Sanger sequencing was performed to analyze the candidate variants. Results were given in Figure A.21-A.26 and Table 5.12.

Table 5.12. Sanger sequencing results for family H28.

Gene	H28.1 (Index patient)	H28.2 (Affected brother)	H28.3 (Affected brother)	H28.4 (Unaffected sister)	H28.5 (Unaffected mother)	Conclusion
CERS2-1	var;var	var;var	+;var	+;var	+;var	Exclude
CERS2-2	var;var	var;var	+;var	+;var	+;var	Exclude
COL6A3-E41	var;var	+;var	var;var	var;var	+;var	Exclude
COL6A3-E9	var;var	+;+	+;var	var;var	+;var	Exclude
TNXB-E23	+;var	+;var	+;var	+;var	+;var	Exclude
TNXB-E21	var;var	+;var	+;var	+;var	+;var	Exclude
LAMA4	var;var	var;var	+;var	+;+	+;var	Exclude
GPRC6A	var;var	var;var	var;var	var;var	var;var	Exclude
FAM35A	+;+	+;+	+;+	+;+	+;+	Exclude
NUTM2D	var;var	var;var	var;var	+;var	+;var	Include
ATAD1	var;var	var;var	var;var	+;var	+;var	Include
LIPF	var;var	var;var	var;var	+;var	var;var	Exclude
HSDL1	var;var	var;var	+;+	+;var	+;var	Exclude

+: native allele, var:variant allele

According to direct sequencing results, variants in *NUTM2D* and *ATAD1* were segregating with the disease in the family. However, the variant in NUTM2D gene has been determined in population databases in homozygous state and it was also eliminated. Position of *ATAD1* might be observed on ROH distribution in Figure 5.15.

c.701T>C, p.Met234Thr variant in *ATAD1* (THORASE/ ATPase Family, AAA Domain Containing 1) gene was neither found in ExAc nor in 1000G databases. Methionine amino acid in position 234 is evolutionarily conserved and located in ATPase domain of the protein (Figure 5.16).

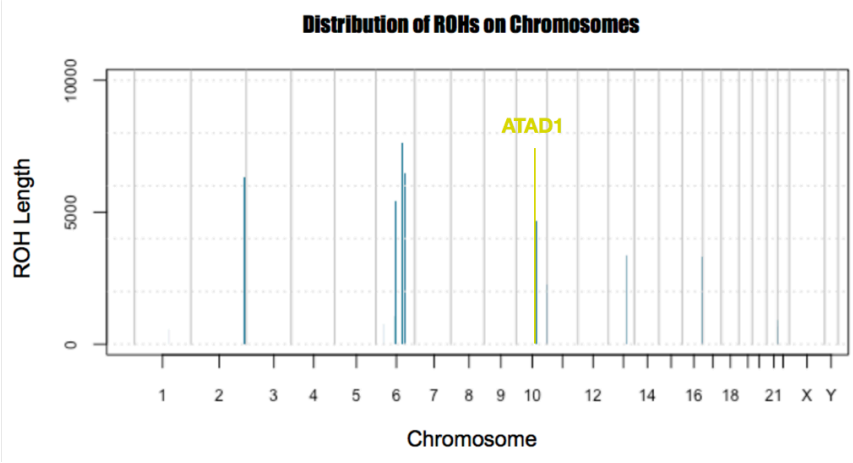


Figure 5.15. Distribution of ROHs identified for the patients in family H28 and position of ATAD1 gene.

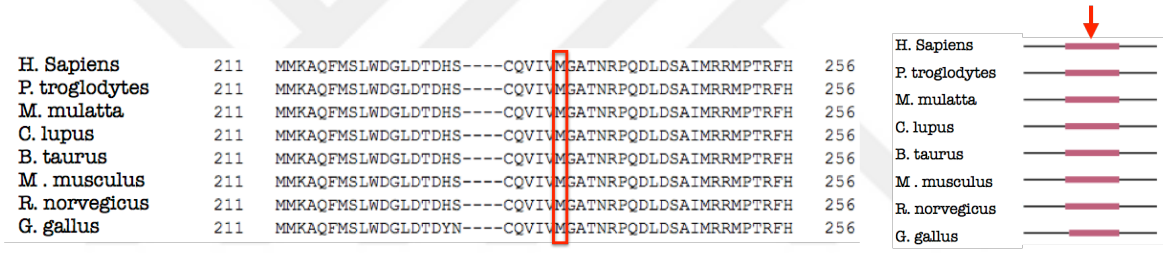


Figure 5.16. Evolutionary conservation of p.Met234 (left panel) and position on ATAD1 protein (right panel).

Variant did not result in a significant deviation in protein according to 3D model prediction algorithms (Figure B.4a,b,c). However, switch of methionine, a hydrophobic amino acid, to threonine, a non-charged hydrophilic might affect the interaction of protein (Figure B.5d). When the amino acid sequences of the proteins are used by iTASSER 3D model algorithm (Zhang, 2008) as a template to predict ATAD1 protein model, it was observed that methionine amino acid itself was not conserved among these proteins but all of the conserved amino acids in the very same position are hydrophobic including Leucine and Isoluecine (Figure B.5).

After updating pedigree information, X-linked recessive inheritance was considered for the family and WES data was analyzed (Table 5.13). Although *ATAD1* was determined as a disease-causing variant, XL candidates were sequenced to exclude this possibility.

Table 5.13. Candidate variants identified in X-linked genes in WES data of H28.1.

Gene	Variant	Coordinate	Consequence	SIFT	PolyPhen	dbSNP ID
ARSD	C>C/T	X:2833638	missense	D (0.01)	Pr_D (0.988)	
ARSD	C>C/T	X:2836184	missense	D (0.02)	Ps_D (0.778)	rs73632976
DGKK	A>AG/AG	X:50121146	frameshift			rs11415594, rs35830302
SSX7	C>T/T	X:52674516	missense	D (0.01)	Pr_D (1)	rs5986182
FAM104B	T>T/G	X:55172708	missense	D (0)	Ps_D (0.764)	rs1047034
ATP7A	G>C/C	X:77268502	missense	D (0.03)	Ps_D (0.435)	rs2227291
H2BFM	C>T/T	X:103294760	stop gained			rs2301384
IDH3G	C>CAG/CAG	X:153056311	splice region			rs17430

D: deleterious; PsD: possibly damaging; PrD: probably damaging.

Variants listed in Table 5.13, were analyzed in the family members via Sanger sequencing. Results for each variant were given in Figure A.27-A.30, Table 5.14.

Table 5.14. Sanger sequencing results for family H28 in X-linked genes.

Gene	H28.1 (Index patient)	H28.2 (Affected brother)	H28.3 (Affected brother)	H28.4 (Unaffected sister)	H28.5 (Unaffected mother)	Conclusion
ARSD-E6	+;Y	+;Y	+;Y	+;+	+;+	Exclude
ARSD-E5	+;Y	+;Y	+;Y	+;+	+;+	Exclude
ATP7A	var;Y	+;Y	var;Y	+;+	+;var	Exclude
DGKK	+;Y	+;Y	+;Y	+;+	+;+	Exclude
FAM104B	+;Y	+;Y	+;Y	+;+	+;+	Exclude
H2BFM	var;Y	var;Y	var;Y	+;var	+;var	Include
IDH3G	var;Y	+;Y	+;Y	+;+	+;var	Exclude
SSX7	+;Y	+;Y	+;Y	+;+	+;+	Exclude

+: native allele, var:variant allele

Segregation analyses revealed that variant in H2BFM gene was segregating with the disease. However, since the variant in H2BFM gene was reported in population databases in homozygous state, this variant was also excluded and ATAD1 gene was accepted as a novel candidate HSP gene.

5.2.8. Family H29

WES data analysis in the index patient H29.1 revealed a stop-gained c.4804C>T (p.Arg1602Ter) variant in *ZFYVE26* gene, however, it was found to be a false positive data by ssanger sequencing and excluded (Figure A.31).

In WES data, a heterozygous missense c.1168G>A, p.Gly390Ser variant was observed in *RTN2* gene that was known to be responsible for dominant HSP. Although pedigree suggests a recessive inheritance, considering the possibility of incomplete penetrance, this variant was analyzed by Sanger sequencing. It was excluded because the affected individual H29.2 was found to be homozygous for the native allele (Figure A.32).

SNP genotyping and homozygosity mapping were performed for two affected individuals from H29 family. Two affected siblings shared twelve ROHs (Figure 5.17, 5.18). New candidate variants in these ROHs were determined in WES data of H29.1 (Table 5.15).

	A	B	C	D	E	F	G	H
1	CHR	POS1	POS2	KB	NSNP	DENSITY	PHOM	PHET
2	1	53072454	57111169	4038.716	52	77.668	0.788	0.038
3	5	52856504	56207123	3350.62	36	93.073	0.833	0.056
4	7	149476666	150217248	740.583	30	24.686	0.7	0.067
5	8	17891259	19362768	1471.51	26	56.597	0.923	0.077
6	11	62766431	63997764	1231.334	29	42.46	0.828	0.069
7	12	66786073	71139664	4353.592	29	150.124	0.862	0.069
8	12	74932159	82792787	7860.629	40	196.516	0.825	0.025
9	14	61747644	65419293	3671.65	34	107.99	0.853	0.059
10	15	85487960	86814866	1326.907	26	51.035	0.885	0.038
11	16	103517	2059955	1956.439	186	10.518	0.715	0.038
12	17	39305775	39881289	575.515	58	9.923	0.81	0.052
13	18	72226539	77420703	5194.165	21	247.341	0.667	0.048

Figure 5.17. ROHs identified for the patients in family H29.

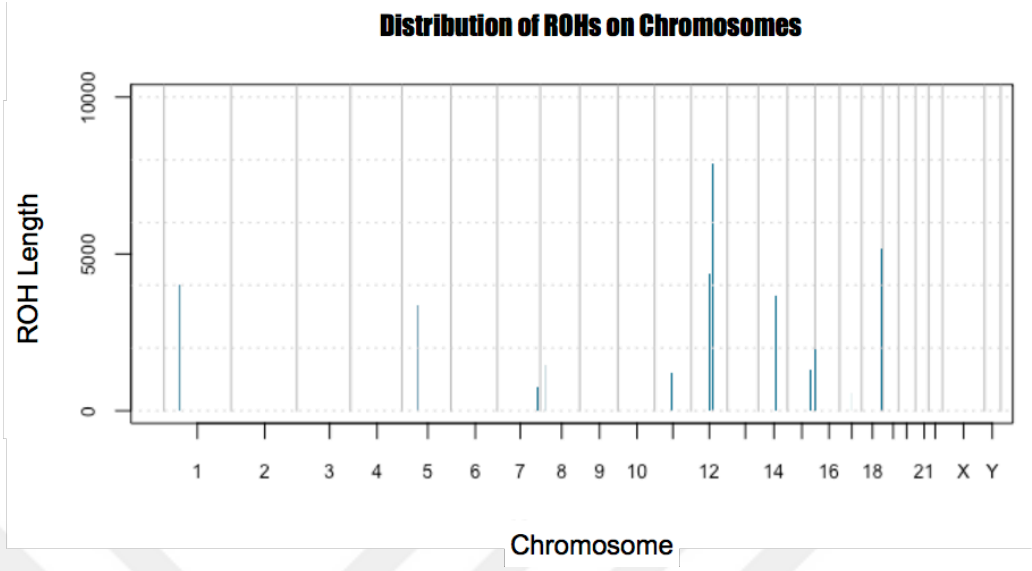


Figure 5.18. Distribution of ROHs identified for the patients in family H29.

Table 5.15. Candidate variants identified in the WES data of H29.1.

Gene	Variant	Position	Effect	SIFT	PolyPhen	dbSNP ID
TTC22	c.40C>G	1:55266797	Missense	D (0.01)	Unknown	rs671108
SSPO	c.15446-7G>A	7:149530894	Splice site			rs6960187
NAT2	c.590G>A	8:18258103	Missense	D (0.05)	Pr_D (0.992)	rs1799930
SLC22A10	c.1547C>T	11:63072310	Missense	D (0)	Pr_D (0.999)	rs1201559
KRR1	c.401G>A	12:75900382	Missense	D (0)	Pr_D (0.923)	rs11540407
AKAP13	c.1871G>T	15:86123170	Missense	D (0)	Ps_D (0.436)	rs745191
PDIA2	c.1144C>G	16:336377	Missense	D (0.01)	Pr_D (0.998)	rs45529833
WDR90	c.2594C>T	16:707882	Missense	D (0)	Pr_D (0.924)	
CLCN7	c.1798-8G>A	16:1498775	Splice site			rs35915213

D: deleterious; PsD: possibly damaging; PrD: probably damaging.

Variants listed in Table 5.15, were analyzed in the family members via Sanger sequencing. Results were given in Figure A.33-A.37, Table 5.16.

Table 5.16. Sanger sequencing results for family H29.

Gene	H29.1 (Index patient)	H29.2 (Aff. sister)	H29.3 (Unaf. sister)	H29.4 (Unaf. mother)	H29.5 (Unaf. father)	Conclusion
TTC22	var;var	var;var	var;var	var;var	var;var	Exclude
SSPO	var;var	+,var	+; +	+;var	+;var	Exclude
NAT2	var;var	var;var	var;var	+;var	+;var	Exclude
SLC22A10	var;var	var;var	+,var	+,var	+; +	Include
KRR1	var;var	+,var	+,var	+,var	var;var	Exclude
AKAP13	var;var	+; +	+,var	+,var	var;var	Exclude
PDIA2	var;var	var;var	var;var	+,var	var;var	Exclude
WDR90	var;var	+;var	+; +	+;var	+;var	Exclude
CLCN7	var;var	+;var	+; +	+;var	+;var	Exclude

Unaf: unaffected, +: native allele, var:variant allele

Direct sequencing excluded all candidate variants. Although the variant in *SLC22A10* was segregating with the disease, it was reported in population databases in homozygous state. Analysis was finalized for this family.

5.2.9. Family H38

H38.3 was the index patient. In WES analysis, the benign polymorphism (rs3759871) c.1388T>C in *SPG11* gene was determined in heterozygous state that was the same variant observed in patients P392, H28.1 and H29.1. Candidate variants selected in neurological disorder-genes are listed in Table 5.17.

Direct sequencing analysis of the c.39962C>T variant revealed that affected individual H38.1 was homozygous for the native allele, affected individual H38.6 was heterozygous, and unaffected individual H38.7 were homozygous for the variant. According to these findings, this variant was not segregating with the disease (Figure A.38).

Table 5.17. Candidate variants for patient H38.3.

Gene	Variant	Posiiton	Variant class	SIFT	PolyPhen	dbSNP
TTN	c.39962C>T	2:179514292	Missense	-	-	
TTN-AS1, TTN	c.63917G>A	2:179452021	Missense	-	-	rs202240487
ARNT	c.-29delG	1:150849071	5'-UTR	-	-	rs66776757, rs34415707

D: deleterious; PrD: probably damaging.

When segregation of c.63917G>A variant was analyzed in the family, affected individual H38.6 was found to be heterozygous and unaffected individual H38.7 was homozygous for the variant. Thus, this variant was not segregating with the disease (Figure A.39).

The c.-29delG variant identified in the 5'UTR region of the *ARNT* gene. According to Sanger sequencing analysis performed in the family members, affected individual H38.3 and unaffected individual H38.8 were homozygous, and affected individual H38.1 and unaffected individuals H38.2, H38.4, H38.5 and H38.7 were heterozygous for the variant (Figure A.40). These findings excluded the possibility that c.-29delG was the causative variant.

Since all candidate variants were eliminated SNP genotyping and homozygosity mapping were performed for this family. Twelve ROHs shared by affected individuals H38.4 and H38.7 were determined (Figure 5.19, 5.20).

	A	B	C	D	E	F	G	H
1	CHR	POS1	POS2	KB	NSNP	DENSITY	PHOM	PHET
2	1	17086052	21795388	4709.337	73	64.511	0.726	0.041
3	1	23560000	31095268	10473.99	168	62.345	0.792	0.042
4	1	152129101	154842199	2713.099	68	39.899	0.691	0.059
5	2	2580232	22546852	27606.13	172	160.501	0.86	0.035
6	2	30549428	37808745	9799.044	64	153.11	0.781	0.031
7	2	40069664	46386298	8977.615	51	176.032	0.863	0.039
8	2	170139346	188331704	18192.36	170	107.014	0.888	0.024
9	4	5798735	6451193	2443.732	55	44.431	0.8	0.055
10	5	50053015	58489384	8436.37	63	133.911	0.873	0.048
11	10	99799824	102331147	3764	57	66.035	0.842	0.035
12	13	28570608	30684241	13985.11	92	152.012	0.804	0.054
13	20	44996182	50636066	15576.48	89	175.017	0.831	0.034

Figure 5.19. ROHs identified for the patients in family H38.

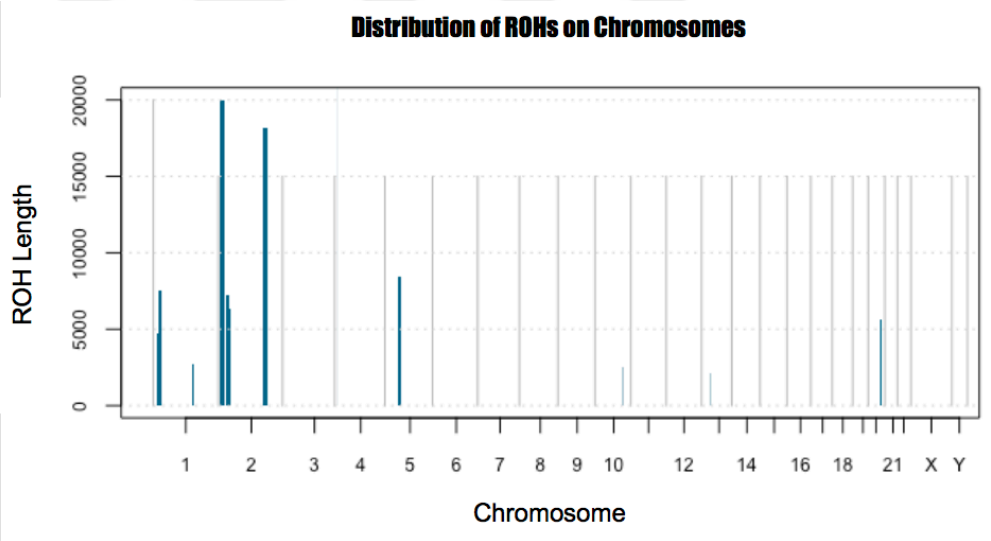


Figure 5.20. Distribution of ROHs identified for the patients in family H38.

To determine the candidate variants in the regions given in the Figure 5.18, WES data of H38.3 was analyzed. New candidate variants were listed in the Table 5.18.

Table 5.18. Candidate variants identified in the WES data of H38.3.

Gene	Variant	Position	Effect	SIFT	PolyPhen	dbSNP ID
EMC1	c.2128G>A	1:19553881	Missense	D (0)	Ps_D (0.889)	

Table 5.18. Candidate variants identified in the WES data of H38.3 (cont.).

Gene	Variant	Position	Effect	SIFT	PolyPhen	dbSNP ID
RNF186	c.67G>A	1:20141528	Missense	D (0.05)	Pr_D (0.958)	rs1541185
NBPF3	c.341A>G	1:21795388	Missense	D (0)	Pr_D (0.989)	rs1827293
MYOM3	c.3972-5T>C	124387577	Splice site			
TRAPPC12	c.901A>G	2:3392295	Missense	D (0.01)	Ps_D (0.779)	rs11686212
APOB	c.293C>T	2:21263900	Missense	D (0.02)	Ps_D (0.825)	rs1367117
CAPN14	c.847G>A	2:31420143	Missense	D (0)	Ps_D (0.893)	rs72797147
WFS1	c.1294C>G	4:6302816	Missense	D (0.01)	Pr_D (0.989)	rs35031397
LOXL4	c.1214A>C	10:100017453	Missense	D (0)	Pr_D (1)	rs1983864
PKD2L1	c.1414C>T	10:102054823	Missense	D (0.03)	Ps_D (0.553)	rs147426900
PREX1	c.4526+7G>A	20:47248808	Splice site			rs4616535

D: deleterious; PrD: probably damaging, PsD: possibly damaging.

Sanger sequencing was performed to analyze variants in the family members. Results were given in Figure A.41-A.45, Table 5.19.

Table 5.19. Sanger sequencing results for family H38.

Gene	H38.1 (Aff. sister)	H38.2 (Unaf.) broher)	H38.3 (Index patient)	H38.4 (Unaf. sister)	H38.5 (Unaf. brother)	H38.7 (Unaf. brother)	H38.8 (Unaf. brtoher)	Conclusion
EMC1	+;+	var;var	var;var	+;var	+;+	+;var	+;var	Exclude
RNF186	+;var	var;var	var;var	+;var	+;var	+;var	var;var	Exclude
NBPF3	var;var	var;var	var;var	+;var	+;var	+;var	var;var	Exclude
MYOM3	+;var	var;var	var;var	+;var	+;var	+;var	+;var	Exclude
TRAPPC 12	+;+	+;var	var;var	+;+	var;var	var;var	+;var	Exclude
APOB	+;var	+;var	var;var	var;var	+;var	var;var	var;var	Exclude

Table 5.19. Sanger sequencing results for family H38 (cont.).

Gene	H38.1 (Aff. sister)	H38.2 (Unaf.) broher)	H38.3 (Index patient)	H38.4 (Unaf. sister)	H38.5 (Unaf. brother)	H38.7 (Unaf. brother)	H38.8 (Unaf. brtoher)	Conclusion
CAPN14	+;var	+;var	var;var	var;var	+;var	+;var	var;var	Exclude
WFS1	+;+	+;var	var;var	var;var	+;var	+;var	+;var	Exclude
LOXL4	var;var	var;var	var;var	+;var	+;var	var;var	+;var	Exclude
PKD2L1	var;var	+;var	var;var	+;+	+;var	+;var	+;var	Exclude
PREX1	+;+	+;var	var;var	+;+	+;var	+;var	+;+	Exclude

Unaf:unaffected, +; native allele, var: variant allele

All candidate variants were eliminated for family H38. Analyses were finalized for this family.

5.2.10. Family H44

Patient H44.7 was selected for WES analysis. The *SPG11* c.1388T>C polymorphism observed in patients P392, H28.1, H29.1 and H38.3 was also present in this patient. When the HSP genes were filtered for compound heterozygosity, two variants in *ZFYVE26* gene were identified (Table 5.20).

Table 5.20. Candidate variants for patient H44.7.

Gene	Variant	Coordinate	Variant class	Sift	PolyPhen	dbSNP ID
ZFYVE26	c.7417-5G>C	14:68215361	Splice region	-	-	rs201771769
ZFYVE26	c.7607G>A	14:68215166	missense	D	B	-

D: deleterious, B: benign

Sanger sequencing analysis of the variant c.7417-5G>C showed that unaffected mother, affected individuals H44.3 and H44.8 were homozygous for the native allele. Unaffected father and affected individual H44.7 was heterozygous for this variant (Figure A.46). When segregation of the other variant, c.7607G>A, was analyzed in the family, it was

observed that unaffected father and affected individual H44.3 were homozygous for the native allele when unaffected mother, affected individuals H44.7 and H44.8 and unaffected individual H44.4 were heterozygous for the variant (Figure A.47). Although the index case bears both of heterozygous variants, other affected individuals did not show this inheritance suggesting that these variants do not segregate with the disease in family H44.

After homozygosity mapping, fifteen ROHs shared by two affected siblings were determined (Figure 5.21, 5.22).

	A	B	C	D	E	F	G	H
1	CHR	POS1	POS2	KB	NSNP	DENSITY	PHOM	PHET
2	1	69511	916549	847.039	20	42.352	0.6	0.05
3	1	74575182	78045369	3470.188	33	105.157	0.848	0.061
4	1	236208773	236978992	770.22	42	18.339	0.905	0.048
5	2	215674341	216808152	1133.812	26	43.608	0.923	0.077
6	2	219692527	220284779	592.253	27	21.935	0.741	0.074
7	3	121238749	122186188	947.44	24	39.477	0.917	0.042
8	5	129521126	132150132	2629.007	26	101.116	0.731	0.077
9	5	179779607	180661256	881.65	24	36.735	0.708	0.042
10	6	46822406	49411477	2589.072	32	80.909	0.906	0.063
11	11	55322298	56431216	1108.919	43	25.789	0.698	0.047
12	12	7992440	10560957	2568.518	59	43.534	0.881	0.034
13	12	10562025	11150046	588.022	39	15.077	0.821	0.051
14	13	24411772	25029218	617.447	20	30.872	0.7	0.05
15	14	37135752	39901157	2765.406	21	131.686	0.857	0.048
16	16	71163693	72088467	924.775	27	34.251	0.815	0.074

Figure 5.21. ROHs identified for the patients in family H44.

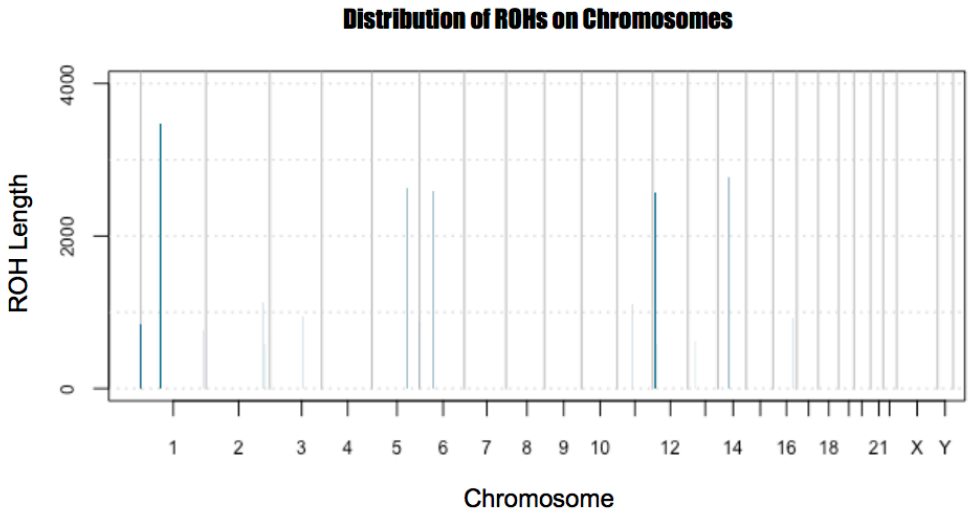


Figure 5.22. Distribution of ROHs identified for the patients in family H44.

One new candidate variant was determined in WES data of H44.7 (Table 5.21).

Table 5.21. Candidate variant identified in the WES data of H44.7.

Gene	Variant	Position	Effect	SIFT	PolyPhen	dbSNP ID
KLRB1	c.503T>C	12:9750669	Missense	D (0)	Ps_D (0.549)	rs1135816

D: deleterious, PsD: possibly damaging

c.503T>C variant was analyzed in the family members and found not to segregate with the disease in the family (Figure A.48). Since the only candidate variant was eliminated, analyses were finalized for this family.

5.2.11. Family H45

Index patient H45.3 was selected for WES analysis. In WES data, there was not any homozygous non-synonymous variant in known HSP genes. However, three heterozygous variants indicating compound heterozygosity were identified in SPG11 gene (Table 5.22).

Table 5.22. Candidate variants identified in *SPG11* gene for patient H45.3.

Variant	Position	Variant class	SIFT	PolyPhen	dbSNP ID
c.2245-1G>A	15:44914998	Splice acceptor			-
c.2057T>C	15:44920877	Missense	D	PrD	-
c.1499_1500delAG	15:44941165	Frameshift			-

D: deleterious; PrD:probably damaging.

Sanger sequencing was performed for variant c.2245-1G>A and mother H45.1, affected individuals H45.3 and H45.4 and unaffected individuals H45.5, H45.6 and H45.8 were found to be heterozygous for the variant (Figure A.49). Analyses of the variant c.2057T>C revealed that affected individuals H45.3 and H45.4 and unaffected individuals H45.1, H45.5, H45.6 and H45.8 were heterozygous for the variant (Figure A.50). The final variant observed in the *SPG11* gene (c.1499_1500delAG, p.E500Vfs57X) was analyzed by Sanger sequencing and affected individuals H44.3, H44.4 and H45.7 and unaffected

individuals H45.1 and H45.5 were found to be heterozygous for the variant (Figure A.51). When the Sanger sequencing data of this family were considered all together (Table 5.23), it was seen that unaffected individuals H45.1 and H45.5 carried all of the three candidate variants. Thus, the possible relation of compound heterozygosity of these variants to the disease was excluded.

Table 5.23. Sanger sequencing results in H45 family.

	H45.1 (Unaf. mother)	**H45.3 (Index patient)**	**H45.4 (Aff. brother)**	**H45.5 (Unaf. brother)**	**H45.6 (Unaf. brother)**	**H45.7 (Aff. sister)**	**H45.8 (Unaf. sister)**
c.2245-1G>A	+;var	+;var	+;var	+;var	+;var	+;var	+;var
c.2057T>C	+;var	+;var	+;var	+;var	+;var	+;var	+;var
c.1499_1500delAG	+;var	+;var	+;var	+;var		+;var	

Unaf:unaffected, Aff. Affected, +; native allele, var: variant allele

Since the number of candidate variants in neurological disorders is higher than five, homozygosity mapping was performed for two affected individuals from this family. Ten ROHs shared by two affected siblings were determined (Figure 5.23, 5.24).

	A	B	C	D	E	F	G	H
1	CHR	POS1	POS2	KB	NSNP	DENSITY	PHOM	PHET
2	1	22191454	22848972	657.519	25	26.301	0.84	0.08
3	2	120059988	121747429	1687.442	23	73.367	0.826	0.043
4	2	209033138	211342502	2309.365	24	96.224	0.667	0.083
5	3	158320597	159744440	1423.844	20	71.192	0.9	0.05
6	7	91503914	92147584	643.671	26	24.757	0.846	0.077
7	9	134020092	135152543	1132.452	24	47.186	0.667	0.083
8	11	49598178	51515912	1917.735	11	174.34	0.818	0
9	11	55322298	56431216	1108.919	47	23.594	0.745	0.064
10	11	71498562	72945341	1446.78	25	57.871	0.76	0.04
11	18	55364852	56246682	881.831	25	35.273	0.92	0.08

Figure 5.23. ROHs identified for the patients in family H45.

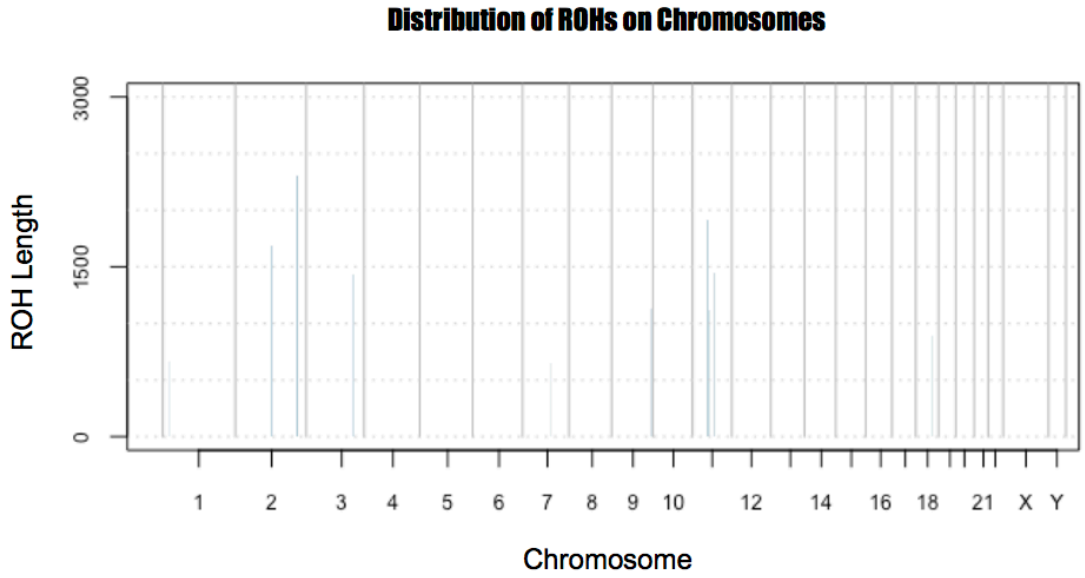


Figure 5.24. Distribution of ROHs identified for the patients in family H45.

To determine the candidate variants in the regions given in the Figure 5.23, WES data of H45.3 was analyzed. New candidate variants were listed in the Table 5.24.

Table 5.24. Candidate variants identified in the WES data of H45.3.

Gene	Variant	Position	Effect	SIFT	PolyPhen	dbSNP ID
PLCL1	c.1999G>A	2:198950240	Missense	D (0.01)	Pr_D (0.996)	rs1064213
ALPK2	c.4345C>T	18:56203074	Missense	D (0.01)	Pr_D (0.904)	rs3809982
ALPK2	c.2748T>A	18:56204671	Missense	D (0)	Ps_D (0.731)	rs4940404
ALPK2	c.2487A>C	18:56204932	Missense	D (0.01)	Ps_D (0.824)	rs3809973

D: deleterious, PrD: probably damaging, PsD: possibly damaging

For the variants listed in Table 24, segregation analyses were performed. Results for each variant and a table summarizing all data obtained by the sequencing were given in Figure A.52, A.53, Table 5.25.

Table 5.25. Sanger sequencing results for family H45.

Gene	H45.1 (Unaf. mother)	H45.3 (Index patient)	H45.4 (Aff. brother)	H45.5 (Unaf. brother)	H45.6 (Unaf. brother)	H45.7 (Aff. sister)	H45.8 (Unaf. sister)	Conclusion
PLCL1	var; var	var; var	var; var	var; var	var; var	var; var	var; var	Exclude
ALPK2 (c.2487A>C)	var; var	var; var	var; var	var; var	var; var	var; var	var; var	Exclude
ALPK2 (c.2748T>A)	var; var	var; var	var; var	var; var	var; var	var; var	var; var	Exclude
ALPK2 (c.4345C>T)	var; var	var; var	var; var	var; var	var; var	var; var	var; var	Exclude

Aff: affected, Unaf:unaffected, +; native allele, var: variant allele

According to Sanger sequencing results, the candidate variants were dismissed and analyses were finalized for family H45.

5.2.12. Family H52

Patient H52.5 was selected for WES analyses. Among known HSP genes from exome sequencing data, a c.1214G>A variant was identified in *SPG7* gene. Other family members were sequenced directly. However, variant was not observed in the affected brother, and was also absent in the unaffected father and sister. This variant is not listed as a polymorphism in Ensembl (Ensembl release 71, 2013) database. Since it was shown that *SPG7* gene can present a dominant inheritance, the variant identified in the H52.5 might have an additive effect on the severity of the disease (Sánchez-Ferrero *et al*., 2013).

WES data analysis revealed a homozygous C.901G>A (p.Arg301Gly) variant in *KIF1C* gene. Therefore, we performed direct sequencing in the other family members. The index patient H52.3 and unaffected mother were found to be heterozygous for the variant. Rebecca Schüle's Lab (Tübingen, Germany) analyzed a second family (THI26001) bearing another mutation in *KIF1C* gene (Figure 5.25a). All these three mutations are found in the highly conserved regions across species (Figure 5.25b). Kinesin motor domain of KIF1C have two functional parts, P-loop and switch-2 cluster, namely. These parts have roles in

ATP hydrolysis and effect conformational alterations in the ATP binding region and the microtubule-binding region. The mutations identified in these families resides in this functional domain. Argininde in 301st position is located in a conserved motif, that is essential for microtubule interaction (Figure 5.25c).

KIF1C belongs to the kinesin protein family (KIFs) that are kinesin-type microtubule-dependent motor proteins and that function as cargo molecules in eukaryotic cells. Since *KIF5A* gene that is associated to SPG10 also belongs to kinesin protein family, KIF1C became a strong HSP gene candidate.

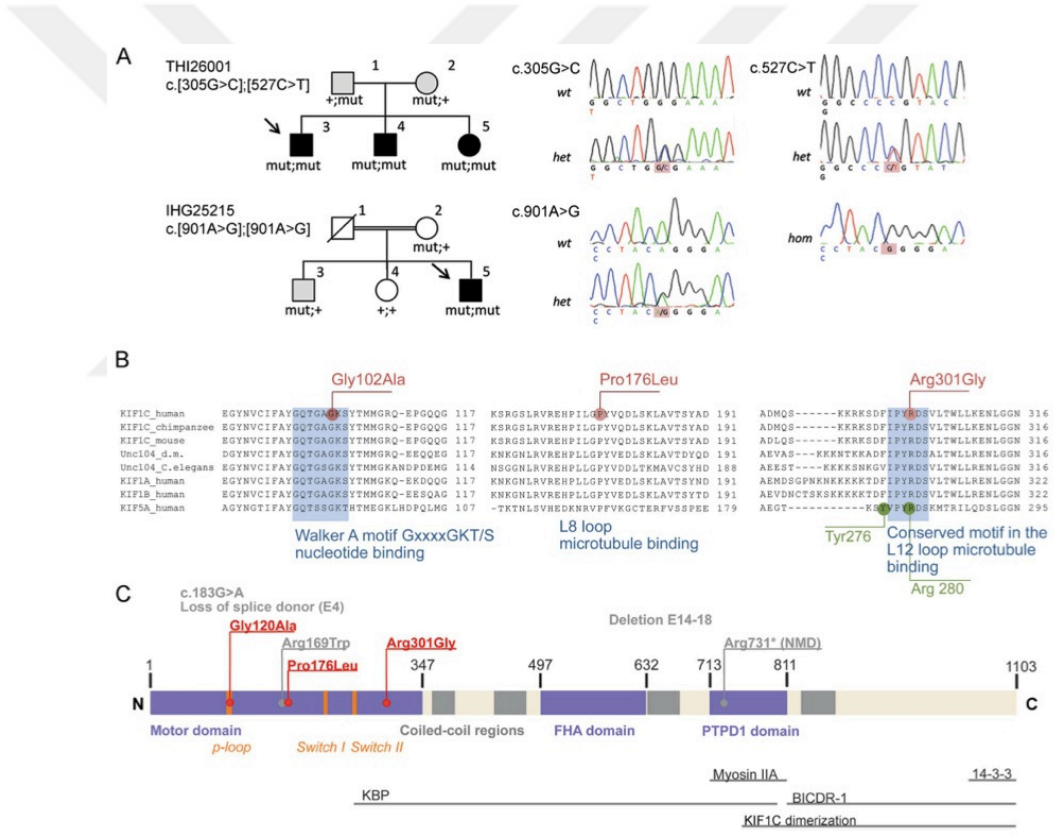


Figure 5.25. The KIF1C mutations c.[305G.C];[527C.T], p.[Gly102Ala];[Pro176Leu], and c.[901A.G];[901A.G], p.[Arg301Gly];[Arg301Gly] observed in families THI26001 and IHG25215 (H52), respectively (Caballero Oteyza *et al*., 2014).

According to 3D model predictions performed by SWISS-MODEL, there was not a structural difference on protein (Figure B.6a, b, c). However, conversion of a charged amino acid, arginine to glycine might alter protein interaction (Figure B.6d).

5.2.13. Family H53

DNA sample of the index patient H53.3 was exome sequenced. Analysis in known HSP genes did not reveal any variants. However, with additional analysis, a missense c.2239G>A (p.747R>W) mutation is determined in a *PLA2G6* gene. Other members of the family were analyzed by direct sequencing and variant was observed only in the affected individuals in a homozygous state (Figure A.55). *PLA2G6* mutations have been shown to be related to various neurodegenerative disorders including neurodegeneration with brain iron accumulation (NBIA), infantile neuroaxonal dystrophy (INAD), Karak syndrome and PARK14-linked young-onset dystonia-parkinsonism syndrome, that are grouped under the name of PLA2G6-associated neurodegeneration (PLAN) (Kurlan *et al*, 2008). In another study a mutation is reported in the same amino acid position in a patient with dystonia-parkinsonism (Paisan-Ruiz *et al*., 2009).

Arginine amino acid in the 747th position of the protein resides in the calmodulin-binding region. Although the function of the protein is not known thoroughly, the frequency of mutations in this domain points out that it is crucial for the proper functioning of PLA2G6. Prediction algorithms did not reveal a difference in 3D structure of the protein however conversion of polar arginine amino acid to non-polar aromatic tryptophan amino acid might alter the interaction dynamics of the protein (Figure B.7).

Since this gene has been previously associated with NBIA, we have analyzed our NBIA1 cohort including three families and three sporadic patients for the mutation, p.747R>W. One of the sporadic cases, HSS-2 was found to be homozygous for the mutation (Figure A.56).

5.2.14. Family H55

WES analysis was performed for H55.3. The missense c.1388T>C variant in *SPG11*, the same variant observed in P392, H28.1, H29.1, H38.3 and H44.7, was present in heterozygous state in the patient. The only strong candidate in neurological disease-genes was c.571C>T missense variant in *SHMT1* gene (Table 5.26).

Table 5.26. Candidate variant identified in SHMT1 gene.

Variant	Chr	Position	Variant class	SIFT	PolyPhen	dbSNP ID
c.571C>T	17	18244076	missense	D	PsD	rs148701087

D: deleterious; PsD: possibly damaging.

Sanger sequencing data result of the variant c.571C>T revealed that this variant was not segregating with the disease (Figure A.57).

After elimination of the variant in *SHMT1* gene, SNP genotyping and homozygosity mapping were performed. Six ROHs shared by two affected siblings were determined (Figure 5.26).

	A	B	C	D	E	F	G	H
1	CHR	POS1	POS2	KB	NSNP	DENSITY	PHOM	PHET
2	2	218695102	223423362	4728.261	68	69.533	0.559	0.059
3	3	42265634	47618953	5353.32	73	73.333	0.753	0.027
4	3	52237970	62431415	10193.446	117	87.123	0.761	0.026
5	17	39406427	43513433	4107.007	138	29.761	0.674	0.043
6	17	45266575	48452776	3186.202	52	61.273	0.692	0.038
7	18	19100758	30050356	10949.599	65	168.455	0.769	0.031

Figure 5.26. ROHs identified for the patients in family H55.

New candidate variants were determined in the WES data of H55.38 (Table 5.27).

Table 5.27. Candidate variants identified in the WES data of H55.3.

Gene	Variant	Position	Effect	SIFT	PolyPhen	dbSNP ID
ZNF142	c.3373C>T	219507866	Missense	D (0)	Ps_D (0.847)	rs201737913
CYP27A1	c.256-1G>C	219674299	Splice site			
SLC23A3	c.443-6T>C	220033830	Splice site			rs3795986
RESP18	c.90G>T	220197388	Missense	D (0)	Ps_D (0.793)	rs2385404
FAM198A	c.1615G>A	43097765	Missense	D (0.01)	Pr_D (0.998)	rs3732858

Table 5.27. Candidate variants identified in the WES data of H55.3 (cont.).

Gene	Variant	Position	Effect	SIFT	PolyPhen	dbSNP ID
NT5DC2	c.1348-7T>A	52558904	Splice site			rs34005367
ITIH1	c.1754A>T	52820981	Missense	D (0)	Pr_D (0.999)	rs678
CCDC66	c.1379G>A	56628031	Missense	D (0.01)	Pr_D (0.996)	rs7637449
DNAH12	c.3923A>G	57431945	Missense	D (0)	Unknown	rs62622492
KRT32	c.1106G>A	39619193	Missense	D (0.01)	Pr_D (0.972)	rs11078993
CDC27	c.77T>C	45258954	Missense	D (0)	Pr_D (0.997)	rs62077276
EFCAB13	c.806-1G>A	45447802	Splice site			rs76299620

D: deleterious, PrD: probably damaging, PsD: possibly damaging

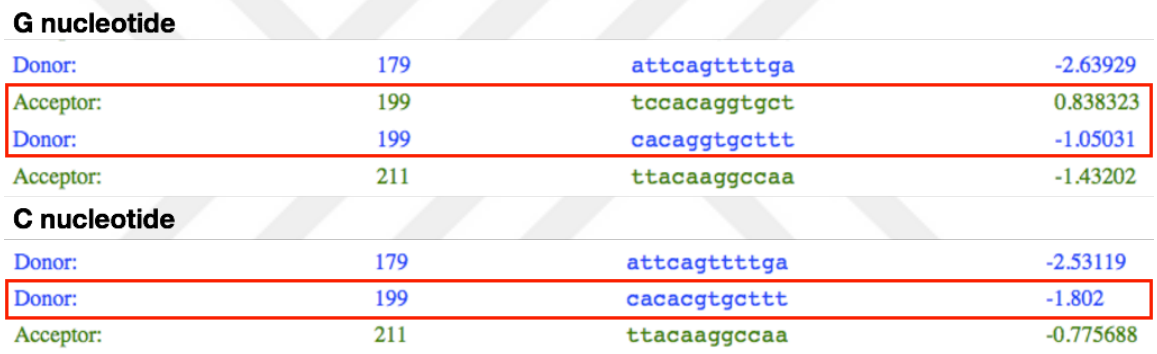
Segregation analyses were performed for new candidate variants (Figure A.58, A.63, Table 5.28).

Table 5.28. Sanger sequencing results for family H55.

Gene	H55.1 (Unaf. mother)	H55.2 (Unaf. father)	H55.3 (Index patient)	H55.4 (Aff. bro.)	H55.5 (Unaf. bro.)	H55.6 (Aff. bro.)	H55.7 (Unaf. sister)	Conclusion
ZNF142	+;var	+;var	var;var	var;var	+;var	var;var	+;var	Include
CYP27A1	+;var	+;var	var;var	var;var	+;var	var;var	+;var	Include
SLC23A3	var;var	var;var	var;var	var;var	var;var	var;var		Exclude
RESP18	+;var	+;var	var;var	var;var	+;var	var;var		Include
FAM198A	+;var	+;var	var;var	+;+	var;var	+;+		Exclude
NT5DC2	var;var	+;var	var;var	+;var	var;var	+;var		Exclude
ITIH1	var;var	+;var	var;var	+;var	var;var	+;var		Exclude
CCDC66	var;var	+;var	var;var	+;var	var;var	+;var		Exclude
DNAH12	+;var	+;var	var;var	+;+	+;+	+;+		Exclude
CDC27	+;+	+;+	+;+	+;+	+;+	+;+		Exclude
EFCAB13	+;var	+;var	var;var	+;+	+;var	+;+		Exclude

Aff: affected, Unaf:unaffected, bro: brother, +; native allele, var: variant allele

According to direct sequencing results, variants in *ZNF142, CYP27A1* and *RESP18* genes were found to be segregating with the disease. The variant in *RESP18* gene has been determined in population databases in homozygous state, thus was eliminated. The variant in *ZNF142* (c.3373C>T, p.Arg1125Trp) gene was observed only in six individuals in heterozygous state in ExAc database. It was predicted to be effective by SIFT and PolyPhen whereas the MutationTaster algorithm predicted this variant as polymorphism. The variant in *CYP27A1* was neither found in ExAC nor 1000G and it was predicted as disease causing by MutationTaster algorithm. When splice site prediction algorithm (http://spliceport.cbcb.umd.edu/cgi-bin/score_seq.cgi) is used to elucidate the effect of this variant, it was seen that it causes an acceptor site to be lost (Figure 5.27).



G nucleotide			
Donor:	179	attcagttttga	-2.63929
Acceptor:	199	tccacaggtgct	0.838323
Donor:	199	cacaggtgcttt	-1.05031
Acceptor:	211	ttacaaggccaa	-1.43202
C nucleotide			
Donor:	179	attcagttttga	-2.53119
Donor:	199	cacacgtgcttt	-1.802
Acceptor:	211	ttacaaggccaa	-0.775688

Figure 5.27. Alteration in acceptor site when c.256-1G>C variant was analyzed in splice site prediction algorithm.

ZNF142 gene have possibly a role in transcriptional regulation. Such transcriptional regulators bear repeated Cys2-His2 (C2H2) zinc finger domains and resembles the Drosophila Kruppel type zinc finger protein family. 3D prediction algorithms were not informative for this protein. Arginine at this position was conserved evolutionarily in higher eukaryotes, indicating its importance for ZNF142 protein function (Figure 5.28).

When the functions of protein products of *ZNF142* and *CYP27A1* genes were considered, both genes might be interpreted as causative for family H55. Positions of genes on distribution of ROHs are given in the Figure 5.29.

H. Sapiens	1114	QQERALRTHQIRGCPLEESGELHCSLCPFTAPAATALRLHQK--RRHPTA	1161
P. troglodytes	1139	QQERALRTHQIRGCPLEKSGELHCSLCPFTAPAAAALRLHQK--RRHPTA	1186
M. mulatta	1110	QQERALRTHQTRGCPLEESGELHCSLCPFTAPAASALRLHQK--RRHPTA	1157
C. lupus	1277	QQERALRTHQTRGCPLEESGELHCSLCSFTTPAAAALRLHQK--RRHPPA	1324
B. taurus	1266	QQERALRTHQTRGCPLEPSGELRCGLCPFTAPAPAALKLHQK--RRHPAA	1313
M . musculus	1172	QQERTLRTHQTQGCPLKS-GDLHCGLCPFTAPAAAALRLHQK--RRHPTA	1218
R. norvegicus	1101	QQERTLKTHQTQGCPLEP-GDLHCGLCPFTAPAPAALRLHQK--RRHPTA	1147
G. gallus	1436	HQERAMKTHKHRGCV--ALGEFRCASCPFTSKAAKALRLHRKLHRKH---	1480

Figure 5.28. Conservation of arginine amino acid in 1125th position of ZNF142 protein.

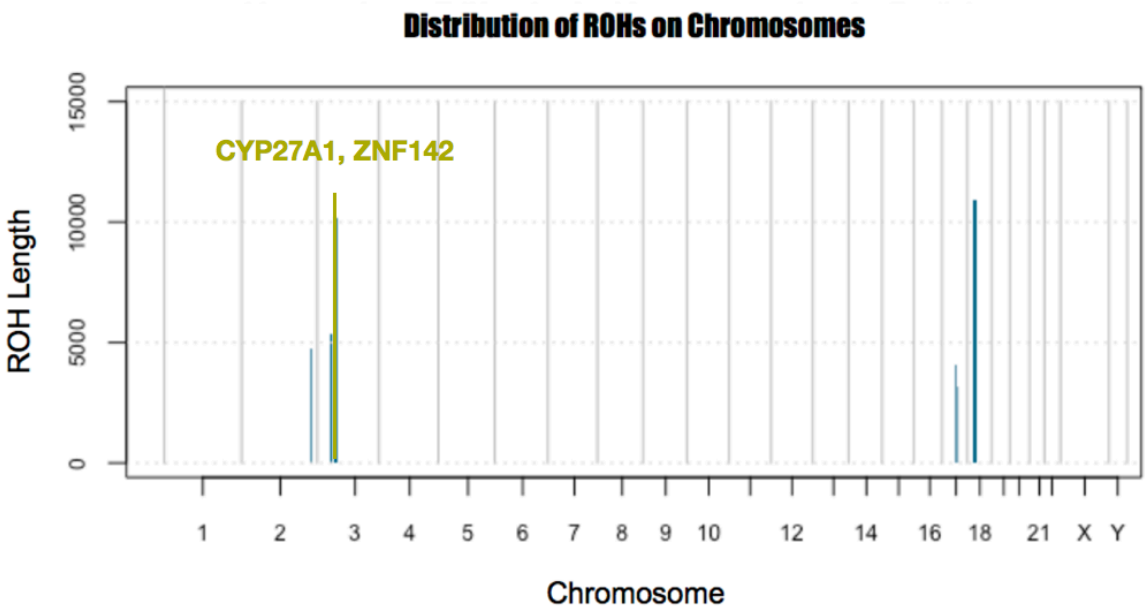


Figure 5.29. Distribution of ROHs identified for the patients in family H55.

During the analysis of the candidate variants, clinicians informed us about the diagnosis of the mother (H55.1) with ALS. With this information, the WES data of the H55.3 were re-analyzed for the variants in known ALS genes. None of the variants in these genes were good candidates as a disease causing variant. Therefore, the variants in neurological disease genes and that are in heterozygous state were analyzed to determine candidates for autosomal dominant inheritance. The candidate variants were listed in Table 5.29. Variants determined for AD inheritance were analyzed in the family members (Figure A.64-A.66, Table 5.30).

Table 5.29. Candidate variants determined for H55 for possible AD inheritance.

Gene	Varaint	Position	Consequence	Sift	PolyPhen	dbSNP ID
HEXB	c.1627G>A	5:74016956	missense	D(0)	Pr_D (0.988)	rs121907984
DST	c.208A>G	6:56569113	missense			rs112635185
FKBP14	c.589G>A	7:30054398	missense	D(0)	Pr_D (1)	
SYT12	c.175T>A	11:66802256	missense	D(0)	Pr_D (0.991)	
MYH4	c.452C>G	17:10368812	missense	D(0.01)	Ps_D (0.563)	
CTDP1	c.2395C>T	18:77477994	missense	D(0)	Ps_D (0.65)	rs148053048

D: deleterious, PrD: probably damaging, PsD: possibly damaging

Table 5.30. Sanger sequencing results for H55 for possible AD inheritance.

Gene	H55.1 (Unaf. mother)	H55.2 (Unaf. father)	H55.3 (Index patient)	H55.4 (Aff. Brother)	H55.5 (Unaf. Brother)	H55.6 (Aff. Brother)	Conclusion
HEXB	+;var	+;+	+;var	+;+	+;var	+;+	Exclude
DST	+;var	var;var	var;var	+;var	+;var	+;+	Exclude
FKBP14	+;var	+;var	+;+	+;var	+;var	var;var	Exclude
SYT12	+;+	+;+	+;+	+;+	+;+	+;+	Exclude
MYH4	+;+	+;var	+;var	+;+	+;+	+;+	Exclude
CTDP1	var;var	+;var	+;var	+;var	+;var	var;var	Exclude

Aff: affected, Unaf:unaffected, +; native allele, var: variant allele

According to Sanger sequencing results, all candidate variants listed in Table 5.29 were eliminated for family H55. ALS disease in the mother could be considered as an independent incidence instead of considering this diagnosis as a part of a spectrum arising from same mutation.

The protein product of *CYP27A1* gene is a member of the cytochrome P450 superfamily of enzymes which also includes *CYP7B1* (SPG5A) gene. This protein family have various functions in synthesis of cholesterol, steroids and other lipids, making *CYP27A1* protein important for overall cholesterol homeostasis. This gene is associated with

cerebrotendinous xanthomatosis (CTX), which is a rare lipid storage disease by progressive neurologic dysfunction. The symptoms might include cerebellar ataxia, systemic spinal cord involvement, premature atherosclerosis, and cataracts. Since CTX might be misdiagnosed as HSP in some cases (Nicholls *et al*., 2015), it was concluded that affected individuals in H55 family might be CTX patients. However, since patients have severe and variable phenotypes in the family, *ZNF142* might still be in the picture underlying disease pathogenesis.

5.2.15. Family H57

H57.3 was the index patient. There was not any homozygous variant determined in the HSP genes in WES data of this patient. However, when we filtered for the variants in AR-HSP genes in heterozygous state, four variants were observed (Table 5.31). *SACS* gene is not related to HSP directly but it is associated to Spastic Ataxia. Since Spinocerebellar Ataxias (SCA) can have similar symptoms with HSP, genes related to SCA are suggested to be screened for HSP patients (Lo Giudice *et al*., 2014).

Table 5.31. Heterozygous variants observed in the AR-HSP genes for the patient H57.3.

cDNA	Variant	Gene	Polyphen	LRT	SIFT	Mutation Taster
	3' -UTR	BICD2				
c.674C>A	missense	ERLIN2	0.964	D	D	D
c.8611T>C	missense	SACS	0.976	D	T	N
c.6062G>A	missense	ZFYVE26	1	D	T	D

D: deleterious, T: tolerated, N: neutral

The c.674C>A variant in the *ERLIN2* gene was not observed in any of the individuals in the family H57, including the index patient (Figure A.67). Sanger sequencing was also performed for variants in *ZFYVE26, SACS* and *BICD2* genes and it was observed that these variants were not segregating with the disease in the family (Figure A.67, A.68).

With further analysis of the data, some other candidate genes were determined. These variants were c.6130A>C in *CEP350* gene, c.1387C>G in *SH3TC1* gene, c.754G>A *PIP5K1B (FAM122A)* gene and c.79C>G in *COQ9* gene. These variants were also eliminated after Sanger sequencing analyses (Figure A.69, A.70).

After analysis of WES data with stringent filtering, four new candidates with highest reading quality and mutation prediction scores were determined (Table 5.32). Sanger sequencing was performed for each variant to analyze segregation in the family.

Table 5.32. Homozygous variants observed in the WES data of patient H57.3 after stringent filtering.

cDNA	Variant	Gene	Polyphen	SIFT	MutationTaster
c.1360C>G	missense	NCOA2	0.92	T	D
c.1373C>T	missense	NCOA2	1	T	D
c.2591G>A	missense	RBP3	0.997	D	D
c.533T>C	missense	SAMHD1	1	D	D

D: deleterious, T: tolerated, N:neutral

The c.533T>C (p.L178P) variant in *SAMHD1* was the only variant segregating with the disease in the family (Figure A.71, A.72). 3D model predictions did not indicate a significant alteration in protein model (Figure B.8). Although both leucine and proline have hydrophobic residues, conversion of leucine to proline that has a rigid ring structure might affect protein interactions (Figure 5.30).

SAMHD1 protein functions in the innate immune response and its expression increases as a response to viral infection. It also involves in control of intercellular dNTP levels. *SAMHD1* mutations are related to Aicardi-Goutiéres syndrome 5 (AGS5) and Chilblain lupus 2 (CHBL2).

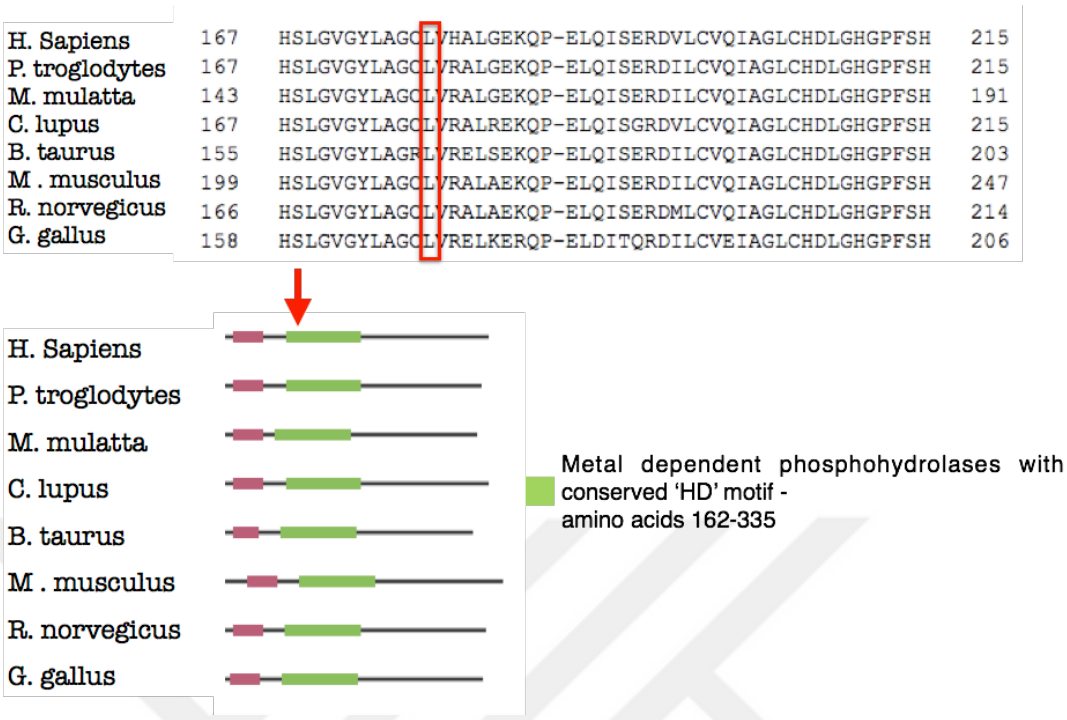


Figure 5.30. Conservation of leucine amino acid in 178th position of *SAMHD1* gene (upper panel) and position of variant on conserved HD motif of the protein (lower panel).

Although WES is a highly effective method to identify disease-causing variant for many patients, it requires a longer process in cases where a high number of strong candidates identified in the WES data. H57.3 has been a good example of such cases. Table 5.33 shows the candidates for H57 family and the corresponding direct sequencing results for affected individuals. Moreover, this table might be an indicator for the importance and effect of genetic background on disease manifestation. As the different variants within families revealed more, they might contribute to understanding the reasons of phenotypical differences.

Table 5.33. Candidate variants analyzed by direct sequencing in H57 family.

Variant	Protein_notation	Variant class	Gene	H57.3	H57.4
c.6130A>C	p.T2044P	missense	CEP350	+;var	+;var
c.1387C>G	p.P463A	missense	SH3TC1	+;var	var;var
c.754G>A	p.G252R	missense	PIP5K1B	+;var	+;var
c.79C>G	p.R27G	missense	COQ9	var;var	+;+

Table 5.33. Candidate variants analyzed by direct sequencing in H57 family (cont.).

Variant	Protein_notation	Variant class	Gene	H57.3	H57.4
		3' -UTR	BICD2	+;var	+;var
c.674C>A	p.A225D	missense	ERLIN2	+;+	+;+
c.8611T>C	p.S2871P	missense	SACS	+;var	+;+
c.6062G>A	p.R2021H	missense	ZFYVE26	+;var	+;+
c.1360C>G	p.Q454E	missense	NCOA2	+;var	+;+
c.1373C>T	p.P458L	missense	NCOA2	+;var	+;+
c.2591G>A	p.R864Q	missense	RBP3	var;var	+;var
c.533T>C	p.L178P	missense	SAMHD1	var;var	var;var

+: native allele; var: variant allele

5.2.16. Family H59

H59.3 was the index patient and he was selected for WES. Analysis revealed a c.325_326insTGTC insertion in *ALS2* gene for H59.3. After direct sequencing, it was observed that affected siblings H59.3 and H59.4 were homozygous and unaffected parents were heterozygous for the variant and unaffected sister was homozygous for the native allele. Patient H59.8, his unaffected parents and sister were also analyzed for this variant and the results were consistent with segregation of the insertion with the disease (Figure A.73).

The c.325_326insTGTC (p.N109RfsX34) variant was not reported previously but hypothetically causes expression of a truncated protein (Figure B.9). *ALS2* mutations were known to cause infantile ascending hereditary spastic paraplegia (IAHSP), juvenile primary lateral sclerosis (JPLS) and juvenile amyotrophic lateral sclerosis (JALS). ALS2 is a guanine exchange factor for GTPases and is localized to endosomal compartments functioning in modulation of endosomal dynamics. ALS2 gene was identified as the causative gene in the family.

5.2.17. Family H61

H61.3 was sequenced in WES experiment as the index patient. A c.4181G>A (p. W1394X) variant was observed in *Spastizin* gene (SPG15). Sequencing analysis showed that

the patient was homozygous and unaffected mother and brother were heterozygous for the variant (Figure A.74).

This variant was reported to be pathogenic, previously and listed in ClinVar database with accession number RCV000191149 (Yang *et al*., 2013; Posey *et al*., 2016). The patient had a complicated HSP phenotype with TCC, pheripheral neuropathy and mental retardation that was compatible with *SPG15* phenotype. Variant leads to truncated protein as predicted by Raptor-X algorithm (Figure B.10). SPG15 was determined as the causative variant in this family.

5.2.18. Family H65

Index patient H65.3 was analyzed by WES. Causative mutation could not be identified in the known HSP genes in the patient. With stringent analysis criteria a c.11264T>C (p.3754M>V) variant in *PLEC* gene and a missense c.14378C>A (p.T4793K) variant in *DST* gene were selected as candidate. Sanger sequencing was performed and it was shown that these variants were not segregating with the disease (Figure A.75, A.76, Table 5.34).

Table 5.34. Candidate variants analyzed by direct sequencing in family H65.

Gene	Variant	H65.2 (Aff. father)	H65.3 (Index patient)	H65.4 (Aff. sister)	H65.5 (Aff. brother)	H65.6 (Aff. sister)	H65.7 (Aff. sister)
PLEC	c.11264T>C	+;var	var;var	var;var	+;var	var;var	+;var
DST	c.14378C>A	+;var	+;+	+;var	+;+	+;+	+;var

Aff: affected, +: native allele, var:variant allele

Homozygosity mapping was performed for the individual H65.3 using WES data and ROHs were determined (Figure 5.31). Candidate variants from WES data of H65.3 were determined in the ROHs (Table 5.35).

	A	B	C	D	E	F	G	H
1	CHR	POS1	POS2	KB	NSNP	DENSITY	PHOM	PHET
2	1	68960551	70225692	1.265.142	1084	1.167	0.959	0.000
3	1	145816752	146317451	500.700	296	1.692	0.932	0.000
4	5	129521640	130498540	976.901	1195	0.817	0.983	0.000
5	8	3889768	6266591	2.376.824	1585	1.500	0.984	0.003
6	8	33424434	35406992	1.982.559	2305	0.860	0.988	0.000
7	8	135725334	136469883	744.550	761	0.978	0.974	0.000
8	12	73050516	75436747	2.386.232	1746	1.367	0.973	0.000
9	12	78592683	79689700	1.097.018	2552	0.430	0.990	0.002
10	17	10633272	11144513	511.242	603	0.848	0.967	0.000
11	17	69334998	70836414	1.501.417	3641	0.412	0.996	0.001

Figure 5.31. ROHs identified for the patients in family H55.

Table 5.35. Candidate variants identified in the WES data of H65.3.

Gene	**Position**	**Variant**	**Variant**	**dbSNP**
VPS13B	8:100520050	c.4211_4213delGTA	p.Cys1404_Thr1405delinsSer	rs758324072
EVPL	17:74003318	c.6034C>T	p.Arg2012Trp	rs140726801
BRI3BP	12:125509554	c.334T>G	p.Tyr112Asp	rs149696895

Segregation analyses were performed for the candidate variants. Results were given in Figure A.77-A.79, Table 5.36.

Table 5.36. Sanger sequencing results for family H65.

Gene	**H65.1 (Unaf. mother)**	**H65.2 (Affected father)**	**H65.3 (Index patient)**	**H65.4 (Affected sister)**	**H65.5 (Affected brother)**	**H65.6 (Affected sister)**	**H65.7 (Affected sister)**
VPS13B	+;V	+;V	V;V	V;V	+;V	+;V	+;+
EVPL	+;V	+;V	V;V	V;V	+;+	V;V	V;V
BRI3BP	+;V	+;V	V;V	+;V	+;V	V;V	V;V
Gene	**H65.8 (Unaf. sister)**	**H65.9 (Unaffected sister)**	**H65.10 (Unaffected brother)**	**H65.11 (Unaffected brother)**	**H65.12 (Unaffected sister)**	**H65.13 (Affected daughter)**	**H65.14 (Affected daughter)**
VPS13B	V;V	+;+	+;V	+;V	+;V	+;V	+;V
EVPL	+;+	+;V	+;+	+;+	V;V	+;V	+;V
BRI3BP	+;+	V;V	+;V	+;V	+;V	+;V	+;V

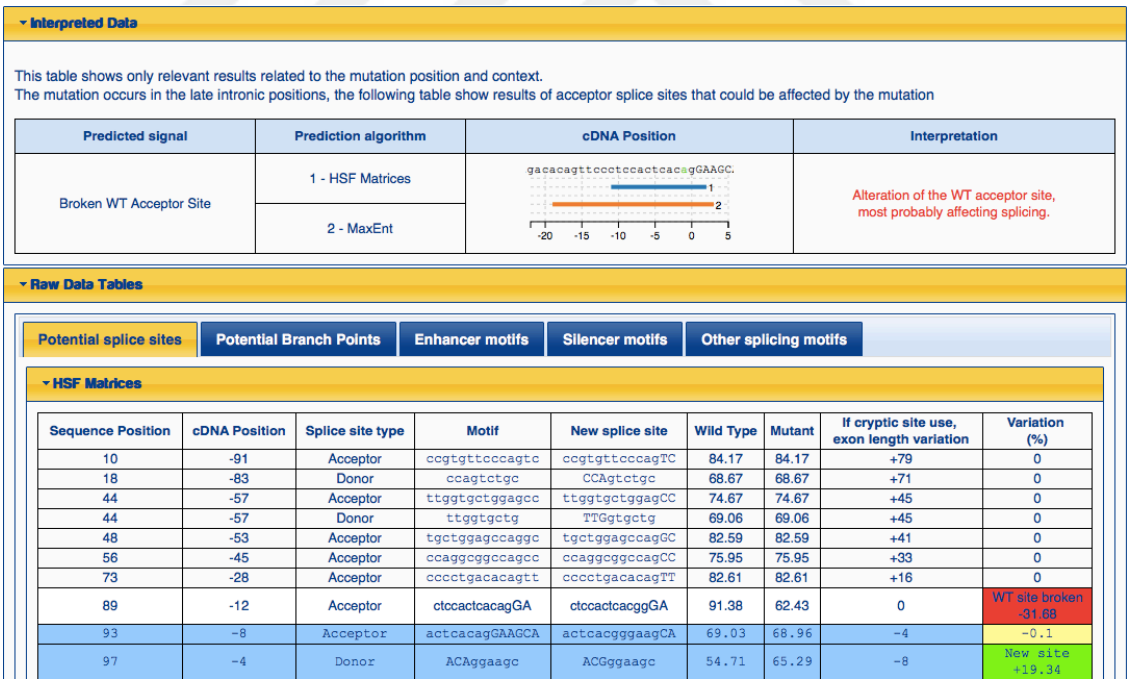
+: native allele, var: variant allele

According to Sanger sequencing results, all candidate variants were eliminated for family H65 and analysis was finalized for this family.

5.2.19. Family H72

WES analysis of the index patient H72.3 revealed a splice acceptor c.2104-2A>G variant in *SPG7* gene. Sanger sequencing analyses revealed that this variant was segregating with the disease in the family (Figure A.80).

The c.2104-2A>G variant was not listed in dbSNP, ExAC (Exome Aggregation Consortium) and HGMD (Human Genome Mutation Database) databases. When the variant was analyzed via Human Splicing Finder, it was calculated that an acceptor motif might be broken and splicing might be affected (Figure 5.32).



Interpreted Data

This table shows only relevant results related to the mutation position and context.
The mutation occurs in the late intronic positions, the following table show results of acceptor splice sites that could be affected by the mutation

Predicted signal	Prediction algorithm	cDNA Position	Interpretation
Broken WT Acceptor Site	1 - HSF Matrices	gacacagttccctccactcacagGAAGC	Alteration of the WT acceptor site, most probably affecting splicing.
	2 - MaxEnt		

Raw Data Tables

Potential splice sites | Potential Branch Points | Enhancer motifs | Silencer motifs | Other splicing motifs

HSF Matrices

Sequence Position	cDNA Position	Splice site type	Motif	New splice site	Wild Type	Mutant	If cryptic site use, exon length variation	Variation (%)
10	-91	Acceptor	ccgtgttcccagtc	ccgtgttcccagTC	84.17	84.17	+79	0
18	-83	Donor	ccagtctgc	CCAgtctgc	68.67	68.67	+71	0
44	-57	Acceptor	ttggtgctggagcc	ttggtgctggagCC	74.67	74.67	+45	0
44	-57	Donor	ttggtgctg	TTGgtgctg	69.06	69.06	+45	0
48	-53	Acceptor	tgctggagccaggc	tgctggagccagGC	82.59	82.59	+41	0
56	-45	Acceptor	ccaggcggccagcc	ccaggcggccagCC	75.95	75.95	+33	0
73	-28	Acceptor	cccctgacacagtt	cccctgacacagTT	82.61	82.61	+16	0
89	-12	Acceptor	ctccactcacagGA	ctccactcacggGA	91.38	62.43	0	WT site broken -31.68
93	-8	Acceptor	actcacagGAAGCA	actcacgggaagCA	69.03	68.96	-4	-0.1
97	-4	Donor	ACAggaagc	ACGggaagc	54.71	65.29	-8	New site +19.34

Figure 5.32. Scores calculated by Human Splicing Finder indicating the possible effect of c.2104-2A>G variant on splicing.

Phenotypic heterogeneity is very high among SPG7 patients. A wide spectrum of clinical variability including pure HSP and complicated HSP have been reported (Warnecke

et al., 2007). SNP genotyping had also been performed for this family. When the ROHs in the H72.3 and H72.3 determined, it is seen that position of one of the ROHs overlaps with the position of the SPG7 gene (Figure 5.33). Causative gene has been identified as *SPG7* for the family H72.

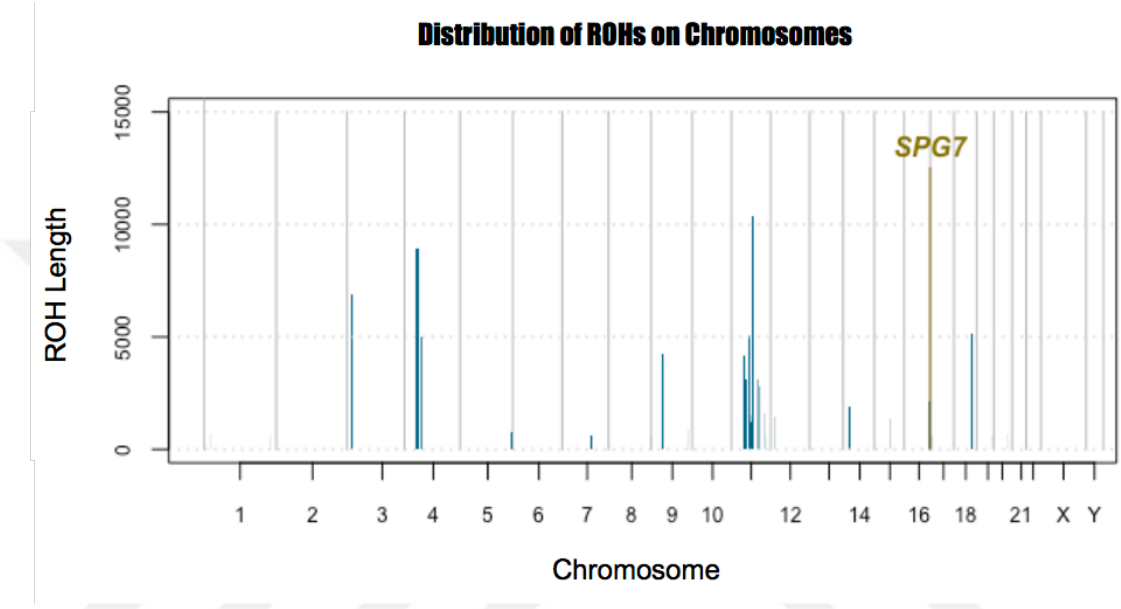


Figure 5.33. Distribution of ROHs identified for the patients in family H72 and location of *SPG7* on the distribution.

5.2.20. Family H77

WES analysis was performed for the index patient H77.3. A homozygous stop gained c.3036C>A (p.Tyr1012Ter) variant in *SPG11* gene was determined. According to Sanger sequencing analysis, unaffected mother H77.1 and unaffected mother H77.2 were heterozygous; affected individual H77.3 and unaffected individual H77.4 were homozygous for the variant. Unaffected individual was homozygous for the native allele (Figure A.81). Since the variant c.3036C>A was not listed in population databases and it was a strong candidate, a detailed clinical re-examination was required for H77.4 and it was found that this individual had MR findings similar to her affected sister. The variant leads to a truncated protein (Figure B.11), making it a PVS1 variant according to ACMG-AMP standards. It is concluded that c.3036C>A variant was segregating with the disease in the family.

5.2.21. Family H82

H82.3 was the index patient and WES analysis was performed for this patient revealed that there was not any homozygous non-synonymous variant in known HSP genes. Since the number of candidate variants was higher than five for this individual, SNP genotyping and homozygosity mapping were performed for the individuals H82.3 and H82.4. 23 ROHs shared by two affected siblings were determined (Figure 5.34, 5.35).

	A	B	C	D	E	F	G	H
1	CHR	POS1	POS2	KB	NSNP	DENSITY	PHOM	PHET
2	1	10566347	11655700	1089.353	64	17.015	0.82	0.18
3	2	131520663	132354365	833.703	26	32.066	0.808	0.077
4	3	100174722	101520537	1345.816	15	89.721	0.933	0.067
5	3	119305303	121973076	2667.774	34	78.464	0.853	0.059
6	3	123167249	125635515	2468.267	40	61.707	0.875	0.05
7	6	25420344	26200083	779.74	26	29.99	0.923	0.077
8	6	28891159	29910562	1019.404	70	14.563	0.686	0.071
9	6	46129345	51875250	5745.906	48	119.706	0.917	0.063
10	6	70866609	76660551	5793.943	26	222.844	0.923	0.077
11	6	79595097	80228673	633.577	10	63.358	1	0
12	12	10163379	11183576	1020.198	54	18.893	0.833	0.074
13	12	11244166	13028653	1784.488	21	84.976	0.857	0.048
14	12	13761741	21713402	7951.662	62	128.253	0.903	0.032
15	15	76673716	78758677	2084.962	41	50.853	0.78	0.049
16	17	64783081	67273882	2490.802	33	75.479	0.909	0.061
17	17	70943990	72741232	1797.243	30	59.908	0.867	0.067
18	19	36595436	38380494	1785.059	48	37.189	0.917	0.042
19	19	42880528	43773626	893.099	21	42.529	0.81	0.048
20	19	44057227	45028169	970.943	72	13.485	0.875	0.056
21	19	52130788	52888042	757.255	50	15.145	0.86	0.06
22	19	52888210	53553686	665.477	37	17.986	0.73	0.054
23	19	53612311	54214330	602.02	36	16.723	0.833	0.056
24	19	56565164	57723013	1157.85	29	39.926	0.897	0.069

Figure 5.34. ROHs identified for the patients in family H82.

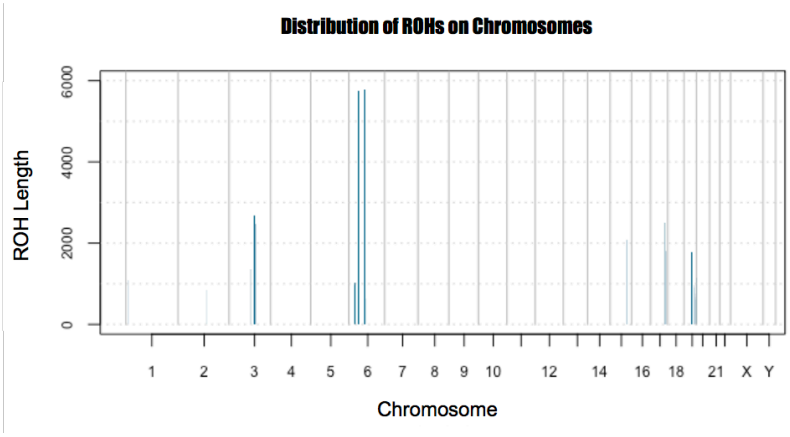


Figure 5.35. Distribution of ROHs identified for the patients in family H82.

WES data of H82.3 was analyzed and new candidate variants were determined and these candidates were given in Table 5.37.

Table 5.37. Candidate variants identified in the WES data of H82.3.

Gene	Position	Variant	Effect	SIFT	PolyPhen
POLQ-E29	3:121151812	c.7612C>G	Missense	D (0.01)	Pr_D (0.998)
ABCA10-E13	17:67145190	c.4515_4516delTC	Frameshift	-	-
ABCA10-E33	17:67149973	c.3964C>T	Splice site	-	-
ABCA10-E39	17:67190536	c.1331_1334delCTGT	Frameshift	-	-
ZNF781-E4	19:38160760	c.290C>T	Missense	D (0)	Pr_D (0.987)
ZNF880-E4-1	19:52887145	c.313delA	Frameshift	-	-
ZNF880-E4-2	19:52888245	c.1412A>G	Missense	D (0.04)	Pr_D (0.999)
VN1R4-E1	19:53770764	c.155C>T	Missense	D (0.03)	Ps_D (0.77)
ZSCAN5A-E7	19:56733425	c.1010G>T	Missense	D (0.01)	Ps_D (0.877)
ZSCAN5A-E4	19:56755560	c.149T>A	Missense	-	-

D: deleterious, PrD: probably damaging, PsD: possibly damaging

Segregation analyses were performed for each new candidate variant (Figure A.82-A.86, Table 5.38).

Table 5.38. Sanger sequencing results for family H82.

Gene	H82.1 (Unaf. mother)	H82.2 (Unaf. father)	H82.3 (Index patient)	H82.4 (Aff. sister)	H82.5 (Unaf. brother)	Conclusion
POLQ	+; var	+; var	var;var	+;+	+; var	Exclude
ABCA10-E13	var;var	+; var	var;var	+; var	var;var	Exclude
ABCA10-E33	var;var	+; var	var;var	+; var	var;var	Exclude
ABCA10-E39	var;var	+; var	var;var	+; var	var;var	Exclude
ZNF781	+; var	+; var	var;var	+; var	+; var	Exclude

Table 5.38. Sanger sequencing results for family H82 (cont.).

Gene	H82.1 (Unaf. mother)	H82.2 (Unaf. father)	H82.3 (Index patient)	H82.4 (Aff. sister)	H82.5 (Unaf. brother)	Conclusion
ZNF880-E4-1	+; var	+; var	var;var	var;var	var;var	Exclude
ZNF880-E4-2	+; var	+; var	var;var	var;var	var;var	Exclude
VN1R4	+; var	+; var	+; var	+; var	+; var	Exclude
ZSCAN5A-E7	+; var	+; var	var;var	var;var	var;var	Exclude
ZSCAN5A-E4	+;+	+;+	+;+	+;+	+;+	Exclude

Unaf: unaffected, Aff: affected, +:native allele, var: variant allele

All candidate variants were eliminated for family H82, after Sanger sequencing analyses. Analysis was finalized for this family.

5.2.22. Family H98

H98.3 was the index patient. In WES data, a homozygous stop gained c.825T>A (p.Tyr275Ter) variant in *CYP7B1* gene, was determined. This variant was previously reported to be pathogenic (Schüle *et al*., 2009). Further analysis indicated that it was segregating with the disease (Figure A.87).

The c.825T>A variant causes a truncated version of the CYP7B1 protein (Figure B.12). Since the variant was reported as pathogenic, it was considered as the disease causing variant for this family.

5.2.23. Family H99

WES was performed for H99.3. It was determined that there was not any homozygous non-synonymous variant in known HSP genes. When the HSP genes were checked for compound heterozygous variants, three variants in *ZFYVE26* gene are observed (Table 5.39).

Table 5.39. Compound heterozygous variants in *ZFYVE26* gene for patient H99.3.

Variant	Position	Variant class	SIFT	PolyPhen	dbSNP ID
c.7417-5G>C	14:68215361	splice_region variant	-	-	rs201771769
c.5806T>A	14:68233149	Missense variant	D	PrD	-
c.1844C>T	14:68265135	Missense variant	T	PrD	rs117228915

D: deleterious; T: tolerated; PrD: probably damaging.

Sanger sequencing data results were given in Figure A.88 and listed in Table 5.40. Since the affected individual H99.6 was homozygous for the native allele for all three candidates, these variants were concluded as not being segregated with the disease.

Table 5.40. Sanger sequencing results in H99 family for the variants in *ZFYVE26* gene.

Variant	99.3 (Index patient)	99.4 (Affected son)	99.5 (Affected son)	99.6 (Affected son)	Conclusion
c.7417-5G>C	+;var	+;var	+;var	+;+	Exclude
c.5806T>A		+;+	+;+	+;+	Exclude
c.1844C>T	+;var	+;var	+;var	+;+	Exclude

+: native allele, var: variant allele

When the WES data analyzed for other candidates, it was shown that the number of candidate variants was higher than five for this individual. SNP genotyping and homozygosity mapping were performed for the individuals H99.4 and H99.5. Six ROHs that are common to H99.3, H99.4 and H99.5 have been identified by comparing the ROHs identified by PLINK in WES data of H99.3 with ROHs obtained by SNP genotyping (Figure 5.36, 5.37).

	A	B	C	D	E	F	G	H
1	CHR	POS1	POS2	KB	NSNP	DENSITY	PHOM	PHET
2	1	9562830	10743363	1180.533	ROHs identified only in the SNP data			
3	1	25932186	27289431	1357.245				
4	5	110784822	112196949	1412.128	19	74.323	0.895	0.105
5	7	149545056	150324825	779.77	19	41.041	0.789	0.105
6	8	56364382	56986417	622.036	18	34.558	0.667	0.056
7	14	46784407	64681193	17896.786	16	56.049	0.938	0.063

Figure 5.36. ROHs identified for the patients in family H99.

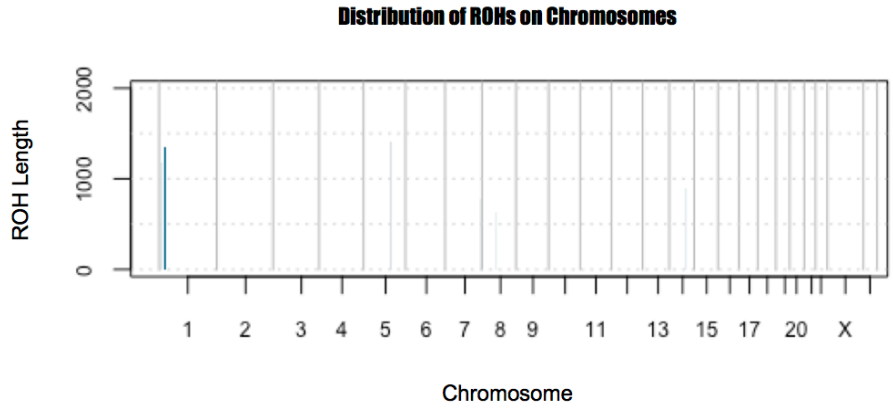


Figure 5.37. Distribution of ROHs identified for the patients in family H99.

New candidate variants were determined in the WES data of H99.3 (Table 5.41).

Table 5.41. Candidate variants identified in the WES data of H99.3.

Gene	Konum	Variant	dbSNP ID	Effect	SIFT	PolyPhen
FUCA1	1:24194773	c.4C>T	rs2070955	Missense	D (0)	Ps_D (0.765)
SSPO	7:149506211	c.9211_9212insC	rs71194663	Frameshift	-	-
GIMAP7	7:150217309	c.247C>T	rs3735080	Missense	D (0)	Pr_D (0.971)

D: deleterious

According to Sanger sequencing results (Figure A.89, Table 5.42), all candidate variants were eliminated for family H99. Analysis was finalized for this family.

Table 5.42. Sanger sequencing results for family H99.

Gene	99.3 (Index patient)	99.4 (Affected son)	99.5 (Affected son)	99.6 (Affected son)	Conclusion
FUCA1	var;var	+;var	+;var	+;var	Exclude
SSPO	var;var	var;var	var;var	var;var	Exclude
GIMAP7	var;var	+;var	+;var	+;var	Exclude

+: native allele, var: variant allele

5.2.24. Family H108

H108.3 was the index patient and he was selected for WES analysis. Analysis revealed a c.*1020C>A variant in 3'UTR of *C19orf12* gene, a heterozygous missense c.820G>A variant in *ZFYVE27* gene and a heterozygous c.1343+6T>G variant in splice region of *CPT1C* gene. The latter two genes are associated to dominant HSP.

Sanger sequencing analysis performed for the c.*1020C>A variant in *C19orf12* gene revealed that affected individuals H108.3, H018.4, H108.5 and unaffected individuals H108.1 and H108.6 were homozygous for the variant (Figure A.90). Therefore, this variant was not segregating with the disease. According to direct sequencing results of the c.1343+6T>G variant in *CPT1C* gene, unaffected mother H108.1, unaffected father H108.2, affected individuals H108.3, H108.4 and H108.5 were homozygous for the native allele (Figure A.91). Thus, this variant was also excluded from being a disease-causing variant. Direct sequencing analysis performed for the c.820G>A variant in *ZFYVE27* (SPG33) gene revealed that all individuals analyzed in the family were homozygous for the native allele (Figure A.92). This variant was excluded for this family.

SNP genotyping and homozygosity mapping were performed for two affected individuals from this family. Twenty ROHs shared by two affected siblings were determined (Figure 5.38, 5.39).

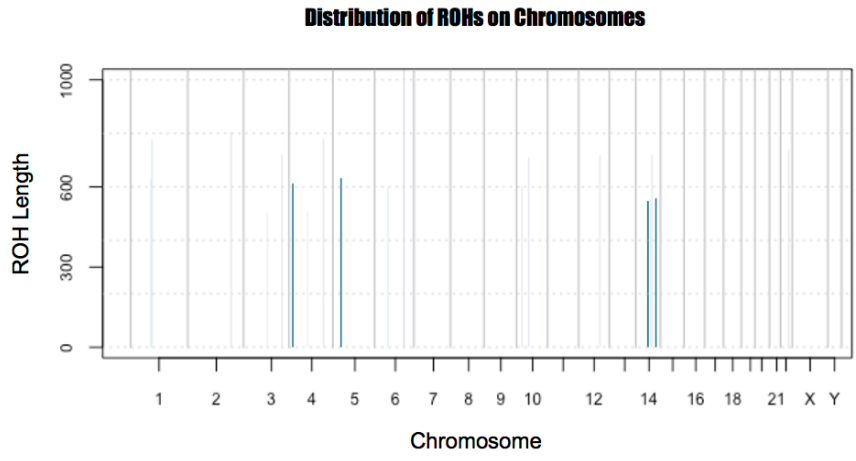


Figure 5.38. Distribution of ROHs identified for the patients in family H108.

	A	B	C	D	E	F	G	H
1	CHR	POS1	POS2	KB	NSNP	DENSITY	PHOM	PHET
2	1	84551960	85177256	625.297	45	13.895	0.444	0.044
3	1	92292407	93068407	776.001	41	18.927	0.512	0
4	2	186625770	187428110	802.341	41	19.569	0.659	0.049
5	3	100621426	101122120	500.695	25	20.028	0.44	0.04
6	3	164766808	165482411	715.604	28	25.557	0.536	0.071
7	4	15780162	16392683	612.522	36	17.015	0.472	0.056
8	4	74736144	75243193	507.05	23	22.046	0.391	0.043
9	4	144502160	145279673	777.514	48	16.198	0.5	0.021
10	5	33637891	34190630	552.74	33	16.75	0.515	0.061
11	5	34918415	35550372	631.958	27	23.406	0.519	0.037
12	6	53215440	53813385	597.946	24	24.914	0.417	0.083
13	6	126334041	127685893	1351.853	23	58.776	0.435	0.043
14	10	20605495	21207548	602.054	22	27.366	0.455	0.045
15	10	50824586	51532528	707.943	22	32.179	0.5	0.045
16	12	89356428	90073585	717.158	23	31.181	0.391	0.043
17	14	52798266	53345378	547.113	21	26.053	0.476	0.048
18	14	60484931	61038944	554.014	31	17.871	0.419	0.032
19	14	68280848	68999331	718.484	26	27.634	0.423	0.038
20	14	87840865	88400353	559.489	30	18.65	0.433	0.067
21	22	31600460	32340188	739.729	26	28.451	0.462	0.077

Figure 5.39. ROHs identified for the patients in family H108.

New candidate variants determined in WES data of H108.3 after homozygosity mapping were listed in the Table 5.43.

Table 5.43. Candidate variants identified in the WES data of H108.3.

Gene	Konum	Variant	dbSNP ID	Effect	SIFT	PolyPhen
SLC45A2	5:33951693	c.1122G>C	rs16891982	Missense	D (0)	Pr_D (0.986)
AMACR	5:34004707	c.524G>A	rs10941112	Missense	D (0.04)	Ps_D (0.76)
NEBL	10:21134282	c.1132G>C	rs41277370	Missense	D (0.01)	Pr_D (0.944)
C10orf53	10:50916542	c.353C>T	rs4838554	Missense	-	Unknown

D: deleterious, PrD: probably damaging

Candidate variants were analyzed in the family members by Sanger sequencing. Results for each variant and a table summarizing all data obtained by the sequencing were given below (Figure A.93, A.94, Table 5.44).

Table 5.44. Sanger sequencing results for family H108.

Gene	108.1 (Unaf. mother)	108.2 (Unaf. father)	108.3 (Index patient)	108.4 (Aff. brother)	108.5 (Aff. brother)	108.6 (Unaf. brother)	Conclusion
SLC45A2	+;var	+;var	+;+	var;var	+;+	+;var	Exclude
AMACR	+;var	+;var	+;var	+;var	+;var	var;var	Exclude
NEBL	+;var	+;var	var;var	var;var	+;var	+;var	Exclude
C10orf53	+;var	var;var	+;var	var;var	var;var	+;var	Exclude

Unaf: unaffected, Aff: affected, +: native allele, var: variant allele

According to Sanger sequencing results, all candidate variants were dismissed for family H108. Analysis was finalized for this family.

5.2.25. Family H110

In WES analysis of the index patient H110.3, it was observed that there was not any non-synonymous variant in known HSP genes. When variants in neurological disease-genes were analyzed, a missense c.12166A>C (p.Thr4056Pro) variant was observed in *SACS* gene, which is related to spastic ataxia. The index patient was found to be homozygous for the variant (Figure A.95). Patient H110.3 was diagnosed with pure HSP and the phenotype of the patient is not consistent with the phenotype of *SACS* related disorders. Therefore, WES data of patient H110.3 was analyzed with PLINK program and 12 ROHs were determined in this patient (Figure 5.40, 5.41).

	A	B	C	D	E	F	G	H
1	CHR	POS1	POS2	KB	NSNP	DENSITY	PHOM	PHET
2	1	248185827	248845499	659.673	33	19.99	0.818	0.03
3	3	239555	5241309	5001.755	32	156.305	0.906	0.063
4	6	89790559	91266349	1475.791	22	67.081	0.773	0.091
5	6	152647681	155575607	2927.927	22	133.088	0.864	0.045
6	10	11299735	12940455	1640.721	21	78.13	0.857	0.048
7	11	48373986	51516036	3142.051	20	157.103	0.9	0.05
8	11	122659875	123894402	1234.528	26	47.482	0.808	0.077
9	11	126139100	129321066	3181.967	24	132.582	0.875	0.042
10	13	25020863	25670945	650.083	20	32.504	0.9	0.05
11	13	37393912	42803341	5409.43	28	193.194	0.786	0.071
12	14	94088587	95058462	969.876	21	46.185	0.857	0.048
13	19	36595436	38056763	1461.328	34	42.98	0.824	0.059

Figure 5.40. ROHs identified for the patient H110.

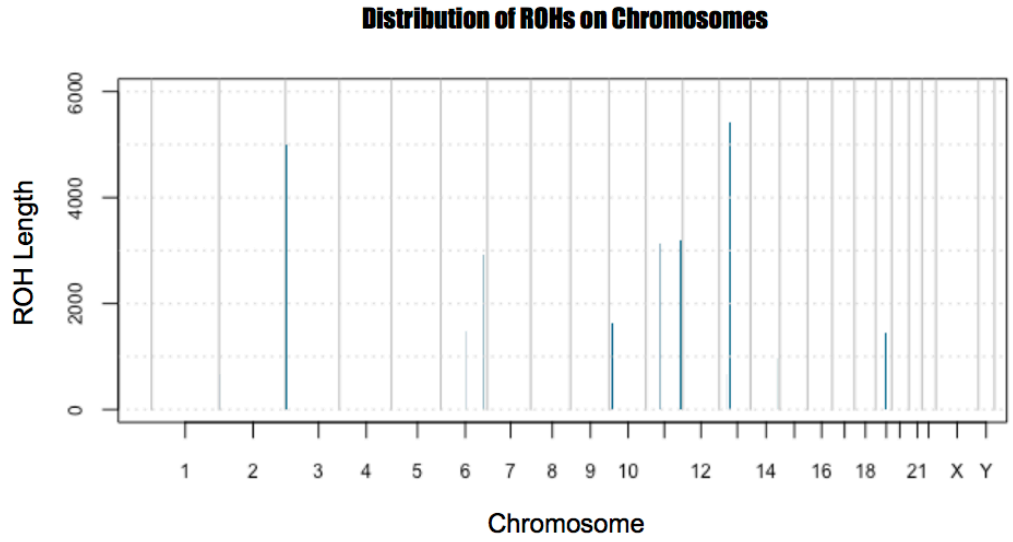


Figure 5.41. Distribution of ROHs identified for the patient H110.

To determine the candidate variants in the regions given in the Figure 5.40, WES data of H110.3 was analyzed. New candidate variants were listed in the Table 5.45.

Table 5.45. Candidate variants identified in the WES data of H6.1.

	Gene	Position	Variant	Effect	SIFT	PolyPhen	dbSNP ID
1	CNTN4	3:3080611	c.2093-6G>C	Splice site	-	-	rs10510251
2	LRRN1	3:3887313	c.988C>T	Missense	D (0)	Ps_D (0.728)	rs146290783
3	MDN1	6:90408742	c.9010G>A	Missense	D (0.01)	Pr_D (0.961)	rs12530146
4	TRIM64C	11:49076963	c.762-6_762-5dupTT	Splice site	-	-	rs34650085
5	TP53AIP1	11:128805548	c.313C>T	Missense	-	Ps_D (0.862)	rs11601431
6	TP53AIP1	11:128805557	c.304A>G	Missense	-	Ps_D (0.438)	rs56235610
7	TP53AIP1	11:128807650	c.63dupG	Frameshift	-	-	rs141395772
8	PABPC3	13:25670945	c.609T>G	Missense	D (0.03)	Pr_D (0.919)	-

D: deleterious, PrD: probably damaging, PsD: possibly damaging

New candidate variants were analyzed in H110.3 via Sanger sequencing (Figure A.96). According to Sanger sequencing results, H110 were found to be homozygous for all candidate variants. Since we could not obtain the DNA samples from other family members, the analyses cannot be pursued for this patient.

5.2.26. Family H133

H133 was the index patient and was selected for WES analysis. Analysis revealed that there was not any non-synonymous variant in known HSP genes. When variants in neurological disease-genes were analyzed, a missense c.437A>T (p.Glu146Val) variant was observed in *CCT5* gene, which is related to Hereditary Sensory Neuropathy with Spastic Paraplegia. This variant was observed in homozygous state in the index patient (Figure A.97). However, the variant was observed in ExAc database in homozygous state and it was eliminated as being a polymorphism.

To analyze other possibilities, WES data of patient H133 was analyzed with PLINK program and 12 ROHs were determined in this patient (Figure 5.42, 5.43).

	A	B	C	D	E	F	G	H
1	CHR	POS1	POS2	KB	NSNP	DENSITY	PHOM	PHET
2	2	73651967	74300717	648.751	22	29.489	0.909	0.045
3	3	54930794	58081888	3151.095	24	131.296	0.875	0.083
4	3	124174212	124692689	518.478	23	22.543	0.913	0.043
5	4	68338199	71008910	2670.712	42	63.588	0.81	0.048
6	4	75937971	77675505	1737.535	21	82.74	0.81	0.048
7	5	94994339	96332085	1337.747	28	47.777	0.893	0.071
8	8	6264130	7215694	951.565	17	55.974	0.941	0.059
9	8	68864728	72981327	4116.6	20	205.83	0.95	0.05
10	11	18108492	18741260	632.769	21	30.132	0.905	0.048
11	13	38237872	43148313	4910.442	40	122.761	0.825	0.05
12	14	104194196	105167807	973.612	20	48.681	0.85	0.05
13	15	100340375	101120963	780.589	27	28.911	0.852	0.037

Figure 5.42. ROHs identified for the patient H133.

New candidate variants were determined in the WES data of H133 and these candidates were given in Table 5.46.

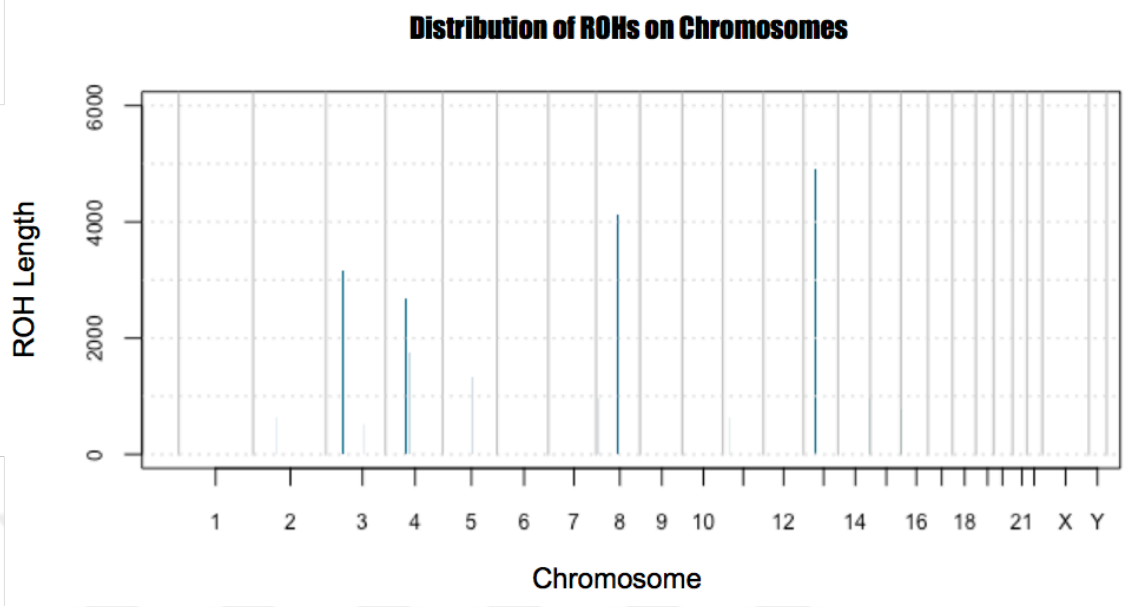


Figure 5.43. Distribution of ROHs identified for the patient H133.

Table 5.46. Candidate variants identified in the WES data of H133.

	Gene	Position	Variant	Effect	SIFT	PolyPhen	dbSNP ID
1	ALMS1	2:73679956	c.6299C>T	Missense	D (0.01)	Ps_D (0.796)	rs28730854
2	CCDC66	3:56628031	c.1379G>A	Missense	D (0.01)	Pr_D (0.996)	rs7637449
3	KALRN	3:124174212	c.3729+6C>T	Splice site	-	-	-
4	UGT2A1, UGT2A2	4:70462042	c.949G>A	Missense	D (0)	Pr_D (0.99)	rs4148301
5	C4orf26	4:76489345	c.133C>T	Missense	-	Ps_D (0.795)	rs2306175
6	SDAD1	4:76878716	c.1724C>G	Missense	D (0.04)	Ps_D (0.519)	rs2242471
7	SAA2-SAA4,SAA2	11:18267597	c.92-7_92-3dupTTTCC	Splice site	-	-	rs202167679, rs34472946
8	HPS5	11:18319180	c.1249C>A	Missense	D (0.01)	Pr_D (1)	rs7128017
9	IGSF22	11:18735480	c.2014C>T	Missense	D (0)	Pr_D (0.998)	rs111397851

Table 5.46. Candidate variants identified in the WES data of H133 (cont.).

	Gene	Position	Variant	Effect	SIFT	PolyPhen	dbSNP ID
10	IGSF22	11:18741260	c.696+3A>G	Splice site	-	-	rs61886896
11	VWA8	13:42385446	c.1978A>G	Missense	D (0)	Pr_D (0.997)	rs9562353

D: deleterious, PrD: probably damaging, PsD: possibly damaging

According to Sanger sequencing results, H133 were found to be homozygous for all candidate variants, except variants 7 and 8 for which she was heterozygous (Figure A.98). Since we could not obtain the DNA samples from other family members, the analyses cannot be pursued for this patient.

5.2.27. Family H142

WES was performed for was the index patient H142.3. There was not any non-synonymous variant identified in known HSP genes. Candidate variants in neurological disease-genes are listed in Table 5.47.

Table 5.47. Candidate variants in neurological disease-genes for patient H142.

Gene	Variant	Position	Variant class	SIFT	PolyPhen	dbSNP ID
SDHAP2	n.1482G>A	3:195400848	Non coding exon variant	-	-	-
C14orf23	n.450_451ins AAC	14:29261309	non_coding exon variant	-	-	rs56025822
SDHAF2	c.103T>C	11:61205163	Missense	D	PrD	-

D: deleterious; T: tolerated; PrD: probably damaging.

The n.450_451insAAC insertion in *C14orf23* gene was analyzed by Sanger sequencing in the family members and it was found that not only index case but also unaffected individuals were homozygous for the variant (Figure A.99). Therefore, this variant was not segregating with the disease in this family. The c.103T>C variant in *SDHAF2* gene and n.1482G>A variant in *SDHAP2* gene were analyzed by Sanger

sequencing and were found not to be segregating with the disease in this family (Figure A.100). After elimination of candidate variants, 19 ROHs shared by two affected siblings were determined by homozygosity mapping (Figure 5.44, 5.45). New candidate variants were determined in the WES data of H142.3 (Table 5.48).

	A	B	C	D	E	F	G	H
1	CHR	POS1	POS2	KB	NSNP	DENSITY	PHOM	PHET
2	1	94341267	95709939	1368.673	15	91.245	0.733	0.067
3	1	214542819	220091716	5548.898	25	221.956	0.88	0.08
4	2	196673533	201468732	4795.2	28	171.257	0.857	0.036
5	2	208811097	212578379	3767.283	25	150.691	0.84	0.08
6	3	184104495	186952035	2847.541	23	123.806	0.783	0.087
7	7	11022230	16131286	5109.057	25	204.362	0.84	0.04
8	7	80290369	83606518	3316.15	17	195.068	0.882	0.059
9	7	86569438	97822210	11252.773	40	281.319	0.95	0.05
10	7	101916647	103155864	1239.218	20	61.961	0.9	0.05
11	8	17827260	20069273	2242.014	20	112.101	0.85	0.05
12	9	106896809	108456888	1560.08	12	130.007	0.917	0
13	9	114359624	116116626	1757.003	20	87.85	0.95	0.05
14	11	5080181	5602968	522.788	56	9.336	0.821	0.054
15	11	49059362	49974371	915.01	11	83.183	0.727	0.091
16	11	55135876	56086147	950.272	34	27.949	0.824	0.059
17	11	57513280	59480951	1967.672	36	54.658	0.806	0.056
18	11	67414354	69462910	2048.557	28	73.163	0.786	0.071
19	11	76868372	78279790	1411.419	25	56.457	0.88	0.04
20	14	94776221	100135306	5359.086	35	153.117	0.8	0.057

Figure 5.44. ROHs identified for the patients in family H142.

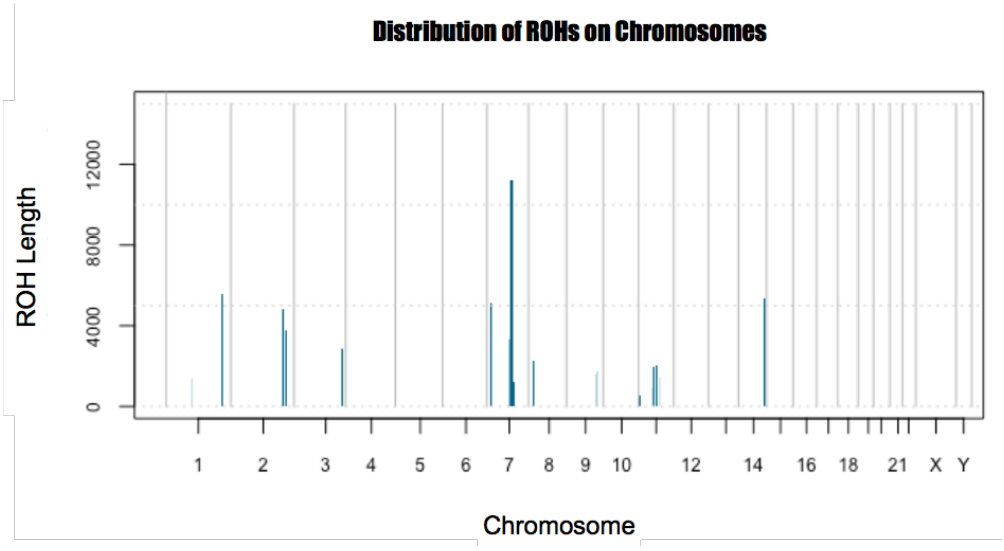


Figure 5.45. Distribution of ROHs identified for the patients in family H142.

Table 5.48. Candidate variants identified in the WES data of H142.3.

Gene	Position	Variant	Effect	SIFT	PolyPhen	dbSNP ID
GPATCH2	1:217604654	c.1420A>G	Missense	D (0.01)	Pr_D (0.979)	rs35737297
DNAH7	2:196756535	c.4897-9_4897-8delTT	Splice site	-	-	rs79407338
PLCL1	2:198950240	c.1999G>A	Missense	D (0.01)	Pr_D (0.996)	rs1064213
DGKB	7:14775824	c.169-5T>G	Splice site	-	-	rs80024169
SEMA3C	7:80418690	c.1286A>G	Missense	D (0)	Pr_D (0.921)	-
KIAA1324L	7:86569438	c.742-7T>C	Splice site	-	-	rs185737704
PRPF4	9:116048977	c.812-8T>C	Splice site	-	-	rs10981714
UBQLNL	11:5536852	c.820T>C	Missense	D (0.02)	Ps_D (0.608)	rs2017434
UNC93B1	11:67765163	c.887dupC	Frameshift	-	-	rs5792426
C11orf24	11:68030173	c.290G>T	Missense	D (0.03)	Unknown	rs3802746
AQP11	11:77314646	c.665C>T	Missense	D (0)	Pr_D (1)	-
USP35	11:77909014	c.706G>A	Missense	D (0.01)	Unknown	rs2510044
TCL1B	14:96157679	c.*15+6T>A	Splice site	-	-	rs34993719

D: deleterious, PrD: probably damaging.

Variants listed in Table 5.48 were analyzed in the family members via Sanger sequencing. Results for each variant and a Table summarizing all data obtained by the sequencing were given below (Figure A.101-A.107, Table 5.49).

Table 5.49. Sanger sequencing results for family H142.

Gene	H142.1 (Unaffected mother)	H142.3 (Index patient)	H142.4 (Affected brother)	H142.5 (Unaffected brother)	H142.6 (Unaffected brother)	Conclusion
DNAH7	var;var	var;var	+;var	+;var	var;var	Exclude
PLCL1	var;var	var;var	var;var	var;var	var;var	Exclude

Table 5.49. Sanger sequencing results for family H142 (cont.).

Gene	H142.1 (Unaffected mother)	H142.3 (Index patient)	H142.4 (Affected brother)	H142.5 (Unaffected brother)	H142.6 (Unaffected brother)	Conclusion
DGKB	var;var	var;var	var;var	var;var	var;var	Exclude
SEMA3C	+;var	var;var	var;var	var;var	+;var	Include
KIAA1324L	+;var	var;var	var;var	var;var	+;var	Include
PRPF4-	+;var	var;var	+;var	+;var	+;var	Exclude
UBQLNL	+;var	var;var	+;var	+;var	var;var	Exclude
UNC93B1	+;+	+;+	+;+	+;+	+;+	Exclude
C11orf24	+;var	var;var	+;var	+;+	+;+	Exclude
AQP11	+;var	var;var	+;var	+;+	+;+	Exclude
USP35	+;var	var;var	+;var	+;+	+;+	Exclude
TCL1B	+;var	var;var	+;var	+;+	+;var	Exclude

+: native allele, var: variant allele

According to direct sequencing results, variants in *SEMA3C*, and *KIAA1324L* genes were found to be segregating with the disease. The variant in *KIAA1324L* gene has been determined in population databases in homozygous state and this variant was eliminated. The variant in SEMA3C gene was observed only in one individual in heterozygous state in ExAc database. This gene locates to longest ROH determined in this family (Figure 5.46).

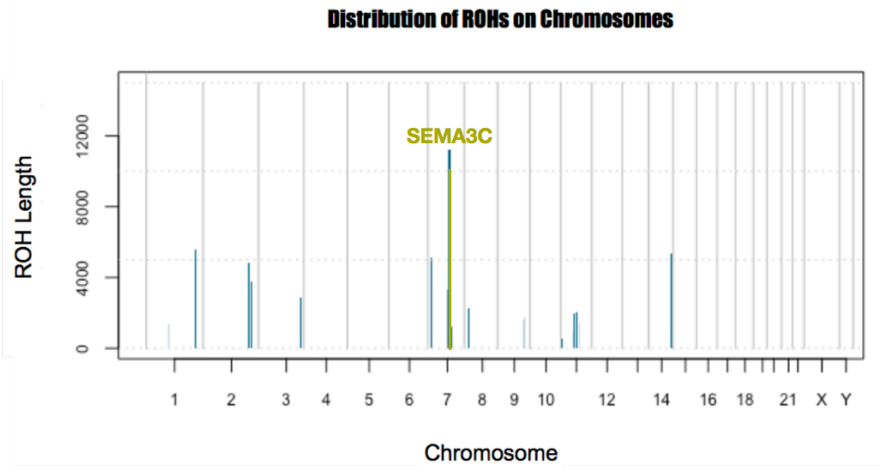


Figure 5.46. Position of SEMA3C on distribution of ROHs on chromosomes.

The c.1286A>G (p.Tyr429Cys) variant in *SEMA3C* was predicted to be effective by SIFT, MutationTaster and PolyPhen. According to 3D prediction algorithms, this variant did not lead to an effective change on protein form. Both tyrosine and cysteine amino acids are polar residues which might not affect the charge of the protein structure. However, ring structure of tyrosine is lost due to mutation which might affect the proximity of the nearby amino acids (Figure B.13). p.Tyr429 is conserved evolutionarily indicating its importance for SEMA3C protein (Figure 5.47).

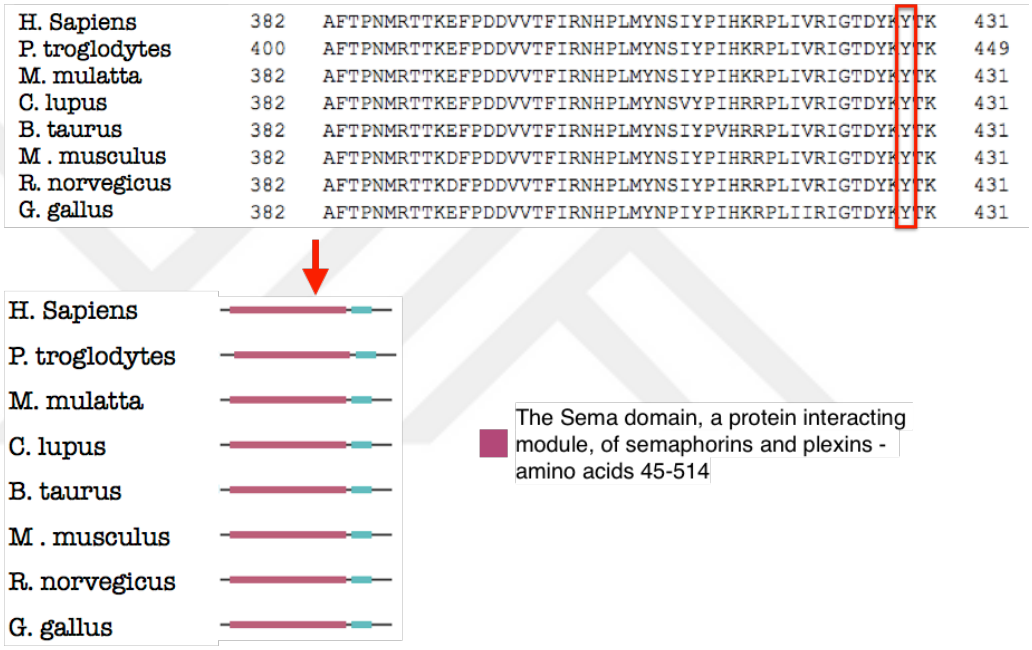


Figure 5.47. Evolutionary conservation of tyrosine amino acid (upper panel) and position on SEMA3C protein (lower panel).

SEMA3C (Semaphorin 3C) belongs to the semaphorin class 3 family and codes for a secreted glycoprotein. The protein contains a sema domain, immunoglobulin-like domain, and a C-terminal basic domain. SEMA3C might have a role in axon guidance and might be involved in in cancer cell invasion and adhesion. The gene is associated with Hirschsprung disease, also known as aganglionic megacolon. As the function of the SEMA3C is considered, c.1286A>G variant in this gene was interpreted as the causative variant for family H142.

5.2.28. Family H143

H143.3 was the index patient and she was selected for WES analysis. She has early-onset (20 years old) complicated HSP with symptoms including TCC and mental retardation. Analysis revealed a c.3904_3907delTCTA (p.Ser1302AsnfsTer2) deletion in *SPG11* gene. When the HSP genes were checked for compound heterozygous variants, three variants in *KIF1A* gene are observed (Table 5.50). Sanger sequencing results for variants in *KIF1A* gene were given Figure A.108, and Table 50. According to these analyses, the variants in KIF1A genes were eliminated.

Table 5.50. Compound heterozygous variants in *KIF1A* gene for patient H143.3.

Variant	Chr	Position	Variant class	SIFT	PolyPhen	dbSNP ID
c.767A>C	2	241723187	Missense variant	D	B	-
c.883-3C>T	2	241715346	splice_region variant	-	-	rs111507743
c.1261A>C	2	241710468	Missense variant	T	PsD	-

D: deleterious; T: tolerated; B:benign; PsD: possibly damaging.

Table 5.51. Sanger sequencing results for H143 family for the variants in *KIF1A* gene.

Variant	H143.1 (Unaffected mother)	H143.2 (Unaffected father)	H143.3 (Index patient)	H143.4 (Unaffected brother)	H143.5 (Unaffected sister)
c.767A>C	+;+	+;+	+;+	+;+	+;+
c.883-3C>T	+;var	+;+	+;var	+;+	+;+
c.1261A>C	+;+	+;+	+;+	+;+	+;+

+: native allele, var: variant allele

Variant identified in the *SPG11* gene was found to be segregating with the disease (Figure A.109). This variant was not listed in ExAc or in 1000G and it was calculated as ‘effective’ by prediction algorithms. The variant leads to a truncated protein as predicted by 3D modelling algorithms (Figure B.14). This variant belongs to the PVS1 category according to ACMG-AMP standards and it was considered as the disease causing variant.

5.2.29. Population screening

The variants identified as causative variants in this study were either not reported in the population databases such as ExAc and 1000G or reported in heterozygous state in very small frequencies. In many cases, small populations do not diverge from the global MAF ratios in remarkable amounts. However, there is still a possibility of observing these variants as polymorphisms in our population. To eliminate this possibility, these variants were analyzed in the WES data of 420 individuals from Turkey. The WES data was kindly shared by Prof. Nazlı Başak (NDAL). In this analysis, it was determined that only the variant in *ZNF142* was observed in one individual in heterozygous state. The other variants were not detected in the NDAL-WES data base (Table 5.52). Therefore, these variants were not polymorphic in our population in accordance with global population findings.

Table 5.52. Frequencies of variants in NDAL-WES database.

Family	Gene	Variant	dbSNP ID	Frequency of variants in NDAL-WES database
P463	SPG11	c.1235C>G, p.Ser412Ter	rs312262723	0
P627	SACS	c.637G>A, p.Gly213Arg	-	0
H16	SPG11	c.6215_6219dupAGAT, p.Phe2074ArgfsTer15	-	0
H28	ATAD1	c.701T>C, p.Met234Thr	-	0
H52	KIF1C	c.901A>G, p.Arg301Gly		0
H53	PLA2G6	c.2239C>T, p.747R>W		0
H55	ZNF142	c.3373C>T, p.Arg1125Trp	rs201737913	One individual in heterozygous state
H55	CYP27A1	c.256-1G>C	-	0
H57	SAMHD1	c.533T>C, p.L178P	-	0
H59	ALS2	c.325_326insTGTC, p.N109RfsX34	-	0
H61	SPG15	c.4181C>T, p.A1394X	-	0
H72	SPG7	c.2104-2A>G	-	0
H77	SPG11	c.3036C>A, p.Tyr1012Ter	-	0
H98	CYP7B1	c.825T>A, p.Tyr275Ter	rs121908613	0

Table 5.52. Frequencies of variants in NDAL-WES database (cont.).

Family	Gene	Variant	dbSNP ID	Frequency of variants in NDAL-WES database
H142	SEMA3C	c.1286A>G, p.Tyr429Cys	-	0
H143	SPG11	c.3904_3907delTCTA, p.Ser1302AsnfsTer2	-	0

5.2.30. Summary of genetic findings

The causative variants that were identified in fifteen families are listed in Table 5.53.

Table 5.53. The variants identified as causative in the families.

Family	Gene	Variant	HGVSc /HGVSp
P463	SPG11	c.1235C>G, p.Ser412*	NM_025137.3 / NP_079413.3
P627	SACS	c.637G>A, p.Gly213Arg	NM_014363.4 / NP_055178.3
H16	SPG11	c.6215_6219dupAGAT, p.Phe2074Argfs*15	NM_025137.3 / NP_079413.3
H28	ATAD1	c.701T>C, p.Met234Thr	NM_032810.2 / NP_116199.2
H52	KIF1C	c.901G>A, p.Arg301Gly	NM_006612.5 / NP_006603.2
H53	PLA2G6	c.2239G>A, p.Arg747>Trp	NM_001349864.1 / NP_001336793.1
H55	CYP27A1 / ZNF142	c.256-1G>C / c.3373C>T, p.Arg1125Trp	NM_000784.3, NM_001105537.1 / NP_001099007.1
H57	SAMHD1	c.533T>C, p.L178P	NM_015474.3 / NP_056289.2
H59	ALS2	c.325_326insTGTC, p.Asn109Argfs*34	NM_020919.3 / NP_065970.2
H61	SPG15	c.4321C>T, p. Ala1394*	NM_015346.3 / NP_056161.2
H72	SPG7	c.2104-2A>G	NM_003119.2
H77	SPG11	c.3036C>A, p.Tyr1012*	NM_025137.3 / NP_079413.3
H98	CYP7B1	c.825T>A, p.Tyr275*	NM_004820.3 / NP_004811.1
H142	SEMA3C	c.1286A>G, p.Tyr429Cys	NM_006379.3 / NP_006370.1
H143	SPG11	c.3904_3907delTCTA, p.Ser1302Asnfs*2	NM_025137.3 / NP_079413.3

5.3. Effect of Rare Variants on Disease Manifestation

As can be seen in the previous parts, analysis of the WES data revealed more than one strong candidate for the exome-sequenced-individuals from each family. Gonzaga-Jaureugi *et al*., (2015) showed the effect of variants in neuropathy-related genes in a large group of peripheral neuropathy patients from US population and 23 CMT patients from Turkey although these variants are not the primary cause of the disease. Although the rare variants they observed in their cohort were not segregating with the disease, they suggested that these variants might explain the phenotypic heterogeneity observed in the patients. To prove this suggestion, they suppressed six different neuropathy-related genes separately and in pair-wise combinations in a zebrafish model. In this model, axons of motor neurons were analyzed to the see the outcome. With this experiment, it was shown that pair-wise suppression had an additive effect on the phenotype.

To identify whether there is a similar situation in our HSP cohort, WES data for nine individuals from nine families were re-analyzed. GEMapp is used to determine rare variants in the data. With this aim, various filtrations were used. First of all, quality of the reads was very important and GQ (genotype quality) should be at least 30. To see the effect of all rare variants, non-synonymous missense, nonsense, splice-donor, splice-acceptor, 3'-UTR, 5'-UTR, noncoding exonic variants and inframe/frameshift INDELs were chosen from the data. MAFs of the variants were set to be less than 1 % in Asian, European and African populations according to HapMap database. To limit the effect of rare variants into neurodenegeration/neuropathy related genes, variants in the genes responsible for the HSP, SMA, dHMN, CMT, ALS, Ataxia, Muscular Dystrophy and genes listed in Mitochondrial Disease Locus Specific Database were chosen.

GEMapp is a WES data analysis tool that also harbors the WES data of individuals with hereditary motor neuron diseases (ALS, HMN) and disorders of extended axons in the central and peripheral nervous system (CMT, HSP). It gives the opportunity to see if a specific variant is observed in any other individual in the database. Thus, to increase rate of the rareness, the number of families with the same variant were set to be less than 20. Table 5.54 shows the number of variants identified in the patients after applying these criteria.

Table 5.54. List of number of variants identified in 'Effect of Rare Variant' Analysis

Patient	HPMV*	Total # of variants	Variants according to inheritance model			
			XLR/ XLD	AR	CH**	AD (including CH)
H17.1	PLP1	88	2	-	4	86
H31.1	SPAST	38	2	1	2	35
H32.2	SPAST	37	-	-	-	37
H52.5	KIF1C	62	1	2	4	59
H53.3	PLA2G6	85	5	2	12	78
H57.3	Not determined	40	1	3	3 (same gene)	36
H59.3	ALS2	51	3	2	9	46
H61.3	ZFYVE26	44	1	-	-	43
H65.4	Not determined	55	2	2	10	41

*HPMV: highly penetrant Mendelizing variant indicating the predominant variant segregating with the disease.

**CH: Compound Homozygous

Since the number of variants observed in each individual is very high in the initial phase, some prediction algorithms were used to eliminate the variants that might have less or no effect on disease manifestation. In our analysis, we chose three algorithms and filtered the variants that are predicted as 'effective' (deleterious variants or highly conserved regions) by all of these three algorithms. These algorithms are explained below.

LRT (Likelyhood ratio test) for amino acid conservation within the human proteome. It gives p values to show significance level of the alteration on amino acid sequence. D is deleterious (p value<0.001), N is neutral and U is unknown.

SIFT uses amino acid homologies to Sort Intolerant From Tolerant (SIFT) amino acid substitutions. It shows the predictions within the range of 1-0, where lower scores indicate higher probability of pathogenicity. SIFT Prediction is based on the SIFT score. According to this T is tolerated (SIFT score > 0.05), D is damaging (SIFT score <= 0.05).

MutationTaster Score gives probability value for disease- causing potential of DNA sequence alterations. Range is between 0 and 1. MutationTaster Prediction is based on MutationTaster and N stands for polymorphism, P shows polymorphism automatic; D is disease causing and A is disease causing automatic. The MutationTaster algorithm predicts N and D whereas P and A are identified by external information. Categories A and D are considered deleterious (Schwarz *et al*., 2010).

Table 5.55 shows the number of variants identified for the same patients after applying prediction algorithms to the variants mentioned in Table 9.

Table 5.55. List of variants identified after LRT, SIFT and MutationTaster predictions.

Patient	**Number of variants**	**Gene**	**Variant**	**Related Disease / Function**
H17.1	2	COL11A1	c.3980G>A (p.G1327E)	Fibrochondrogenesis, Marshall syndrome, Stickler syndrome- II
		NNT	c.2078G>A (p.R693H)	Glucocorticoid deficiency 4
H31.1	2	FAHD1	c.320A>G (p.D107G)	Mitochondria-related
		NTRK1	c.1252C>G (p.D418H)	Insensitivity to pain, congenital, with anhidrosis
H31.1	2	FAHD1	c.320A>G (p.D107G)	Mitochondria-related
		NTRK1	c.1252C>G (p.D418H)	Insensitivity to pain, congenital, with anhidrosis
H32.2	3	HSCB	c.218A>G (p.Y73C)	Mitochondria-related
		NDUFAF1	c.638G>A (p.R213H)	Mitochondrial complex I deficiency
		SCN11A	c.3473T>C (p.L1158P)	Neuropathy, hereditary sensory and autonomic-type VII
H52.5	11	AKR7A2	c.592G>A (p.G198S)	Succinic semialdehyde dehydrogenase deficiency
		ARMC4	**c.74C>G (p.P25R)**	Ciliary dyskinesia
		GLDC	c.2291G>T (p.G764V)	Glycine encephalopathy
		AKR7A2	c.592G>A (p.G198S)	Succinic semialdehyde dehydrogenase deficiency

Table 5.55. List of variants identified after LRT, SIFT and MutationTaster predictions (cont.).

Patient	Number of variants	Gene	Variant	Related Disease / Function
H52.5	11	**ARMC4**	**c.74C>G (p.P25R)-HR**	Ciliary dyskinesia
		GLDC	c.2291G>T (p.G764V)	Glycine encephalopathy
		GPD1	c.329C>T (p.S110L)	Hypertriglyceridemia, transient infantile
		HADHB	c.1285C>A (p.P429T)	Trifunctional protein deficiency
		KIF21A	c.2719C>T (p.R907C)	Fibrosis of extraocular muscles, congenital
		MYO7A	c.562C>G (p.Q188E)	Usher syndrome, Deafness
		OPA1	c.2720A>G (p.E907G)	Optic atrophy
		POLR3B	c.599A>C (p.E200A)	Leukodystrophy, hypomyelinating, 8,
		SARDH	c.2152G>A (p.A718T)	Sarcosinemia
		TMEM237	c.97C>T (p.R33C)	Joubert syndrome 14
H53.3	6	GAMT	c.362C>A (p.A121E)	Cerebral creatine deficiency syndrome
		GNPAT	c.652C>G (p.L218V)	Chondrodysplasia punctata, rhizomelic
		HK1	c.734T>C (p.I245T)	Neuropathy, hereditary motor and sensory, Russe type
		SPG20	c.1358C>T (p.A453V)	Troyer Syndrome
		TCIRG1	c.758G>A (p.R253H)	Osteopetrosis, autosomal recessive 1
		THEM5	c.476G>C (p.G159A)	Mitochondria-related
H57.3	3	ERLIN2	c.674C>A (p.A225D)	HSP
		MRPL33	c.128T>G (p.L43R)	Mitochondria-related
		SAMHD1	**c.533T>C (p.L178P)-HR**	Aicardi-Goutieres syndrome, Chilblain lupus
H59.3	7	CPS1	c.23C>T (p.T8M)	Carbamoylphosphate synthetase I deficiency
		DYSF	**c.5296G>A (p.V1766M)-HR**	Muscular dystrophy, myopathy
		IMPDH1	c.1436C>T (p.T479M)	Leber congenital amaurosis, Retinitis pigmentosa

Table 5.55. List of variants identified after LRT, SIFT and MutationTaster predictions (cont.).

Patient	Number of variants	Gene	Variant	Related Disease / Function
H59.3	7	MRPL37	c.907C>T (p.R303C)	Mitochondria-related
		MYO7A	c.2122A>G (p.M708V)	Usher syndrome, Deafness
		NEB	c.18751A>G (p.M6251V)	Nemaline myopathy 2
		OGG1	c.254A>G (p.Y85C)	Mitochondria-related, Renal cell carcinoma
H61.3	6	AASS	c.421G>T (p.D141Y9	Hyperlysinemia, Saccharopinuria
		ALKBH7	c.598G>A (p.V200M)	Mitochondria-related
		DDHD2	c.1972C>T (p.R658C)	HSP
		IMPDH1	c.1247G>T (p.C416F)	Leber congenital amaurosis, Retinitis pigmentosa
		PLEKHG5	c.202C>A (p.P68T)	CMT, Spinal muscular atrophy
		UNG	c.235C>T (p.R79C)	Immunodeficiency
H65.4	9	ABCF2	c.386C>G (p.S129C)	energy-dependent transport
		CDH23	*c.694G>A (p.R232W)-CH*	Usher syndrome, Deafness
		ETFA	c.200T>C (p.L67P)	Glutaric acidemia
		GAA	c.1048G>A (p.V350M)	Glycogen storage disease
		GBE1	c.1388C>T (p.T463M)	Glycogen storage disease, Polyglucosan body disease
		HK3	c.835G>A (p.D279N)	Hexokinase family such as HK1 (Neuropathy, hereditary motor and sensory, Russe type)
		HSD17B8	c.74G>A (p.R25Q)	Fatty-acid metabolism
		IDH1	c.820G>T (p.G274C)	Glioma
		USH2A	c.6365C>A (p.A2122E)	Usher Syndrome, Retinitis pigmentosa

As can be seen from the Table 5.55, there are 49 rare variants in nine individuals (5.4 variant per individual) and the HPMV for the individuals were not included in the list after

applying algorithm predictions. This shows us that our stringent criteria eliminated even the variants that are predicted to be the main cause of the disease (because they segregate with the disease). The Tables 5.56-5.62 shoe the prediction algorithms results for HPMV (shaded) in grey) and for the variants in each patient listed in the Table 5.55.

Table 5.56. Prediction algorithm scores for the HPMV and rare variants in patient H17.1.

Gene	LRT	SIFT	Mutation Taster	LRT Score	SIFT Score	Mutation taster Score
PLP1				-	-	-
COL11A1	D	D	D	0	0	0.99
NNT	D	D	D	9.9e-07	0	0.99

Table 5.57. Prediction algorithm scores for the HPMV and rare variants in patient H31.1.

Gene	LRT	SIFT	Mutation Taster	LRT Score	SIFT Score	Mutation taster Score
SPAST	N	D	D	1.9e-06	0	1
FAHD1	D	D	D	0	0.01	1
NTRK1	D	D	D	0.0001	0	0.99

Table 5.58. Prediction algorithm scores for the HPMV and rare variants in patient H32.2.

Gene	LRT	SIFT	Mutation Taster	LRT Score	SIFT Score	Mutationtaster Score
SPAST	**-**	**-**	**D**	**-**	**-**	**1**
HSCB	D	D	D	0.0005	0	0.99
NDUFAF1	D	D	D	0	0.019	1
SCN11A	D	D	D	$9.9e^{-07}$	0	1

Table 5.59. Prediction algorithm scores for the HPMV and rare variants in patient H52.5.

Gene	LRT	SIFT	Mutation Taster	LRT Score	SIFT Score	Mutation taster Score
KIF1C		D	D	-	0	0.99
AKR7A2	D	D	D	0	0.02	1
ARMC4	D	D	D	$4E^{-06}$	0	1
GLDC	D	D	D	0	0	1
GPD1	D	D	D	0	0	1
HADHB	D	D	D	0	0	1
KIF21A	D	D	D	0.0001	0	1
MYO7A	D	D	D	$1E^{-06}$	0	1
OPA1	D	D	D	0	0	1
POLR3B	D	D	D	$1E^{-06}$	0.02	1
SARDH	D	D	D	$8E^{-06}$	0.01	1
TMEM237	D	D	D	$3.6E^{-05}$	0.029	0.99

Table 5.60. Prediction algorithm scores for the HPMV and rare variants identified in patient H53.3.

Gene	LRT	SIFT	Mutation Taster	LRT Score	SIFT Score	Mutation taster Score
PLA2G6	N	D	A	0.0041	0	1
GAMT	D	D	D	0.0001	0.01	0.99
GNPAT	D	D	D	$2.3E^{-05}$	0	1
HK1	D	D	D	$1E^{-06}$	0.01	1
SPG20	D	D	D	0	0.02	0.99
TCIRG1	D	D	D	0	0	0.99
THEM5	D	D	D	$1E^{-06}$	0	0.98

Table 5.61. Prediction algorithm scores for the HPMV and rare variants in patient H59.3.

Gene	LRT	SIFT	MutationTaster	LRT Score	SIFT Score	Mutationtaster Score
ALS2				-	-	-
CPS1	D	D	D	$6.5E^{-05}$	0	1
DYSF	D	D	D	$1E^{-05}$	0.01	0.999
IMPDH1	D	D	D	0	0	1
MRPL37	D	D	D	$1E^{-06}$	0.02	0.961
MYO7A	D	D	D	$4.1E^{-05}$	0.039	0.999
NEB	D	D	D	0	0.02	1
OGG1	D	D	D	$3E^{-06}$	0	1

Table 5.62. Prediction algorithm scores for the HPMV and rare variants in patient H61.3.

Gene	LRT	SIFT	MutationTaster	LRT Score	SIFT Score	Mutationtaster Score
ZFYVE26	N	T	A	0.070	0.10	1
AASS	D	D	D	0	0	1
ALKBH7	D	D	D	$7.9e^{-06}$	0.009	0.99
DDHD2	D	D	D	0	0	1
IMPDH1	D	D	D	0	0	1
PLEKHG5	D	D	D	0.0002	0.009	0.537
UNG	D	D	D	0.0005	0	1

These findings implicate that the number of variants would significantly increase by changing algorithm predictions to moderate levels or by searching for the variants that are predicted effective (deleteriousness, conservation) by only one algorithm. To show possible contribution of the rare variants to disease, WGS data of 15 individuals from Turkish Genome Project (TGP) (Alkan *et al*., 2014) were analyzed by applying the same criteria explained above. Variants that are common in more than three individuals were excluded to identify rare variants in healthy Turkish individuals. The average number of rare variants in TGP individuals was 4.53 and this value was 7.16 in ARHSP patients-analyzed. Thus, the number of rare variants in ARHSP patients is significantly higher than that of the healthy individuals (p=0.027) (Figure 5.48).

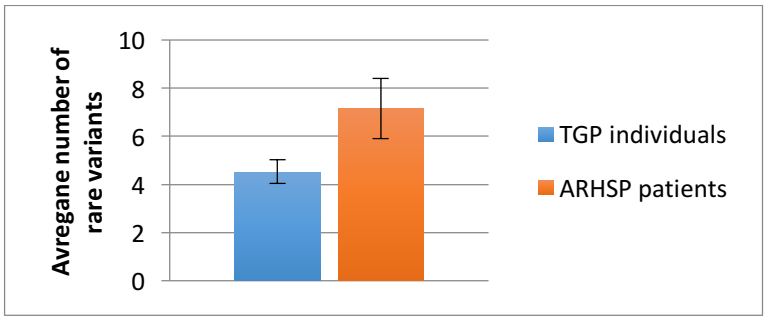


Figure 5.48. Graph showing the number of rare variants in ARHSP patietns and TGP individuals (*p=0.027).

The approach to "end the genetic analysis after determining the disease-causing variant" should be questioned because in this approach other variants that might contribute to disease formation are underestimated. These rare variants observed with the HPMV might be the explanation for inter- and intra-familial heterogeneity in phenotype.

5.4. *In vitro* Analyses

In this part of the study, we aimed to provide further evidence for the involvement of the novel genes in HSP pathogenesis. For this purpose, the expression levels of ATAD1, SAMHD1, and SEMA3C were investigated in immortalized cell lines obtained from corresponding families. CYP27A1 and ZNF142 were also included to those *in vitro* analyses to understand which one of these genes might be causative in disease pathogenesis.

5.4.1. Analyses of Immortalized B-Lymphocytes

To analyze the effect of variants, lymphoblastoid cell lines were obtained by immortalizing B-Lymphocytes obtained from family members. Individuals H28.1 and H28.2 from family H28; individuals H55.1, H55.2, H55.3, H55.4 and H55.6 from family H55; individuals H57.1, H57.2, H57.3 and H57.4 from family H57, and individuals H142.1, H142.3, H142.4, H142.5, H142.6, and H142.7 from family H142 volunteered to donate blood samples for this part of the experiment. Lymphoblastoid cell lines were analyzed by quantitative RT-PCR, western-blot and immunocytochemical analyses to determine the efftec of variant on mRNA expressin, on protein level and on localization, respectively.

5.4.1.1. Family H28 and *ATAD1* Gene. RT-PCR experiment was performed to monitor the mRNA levels of the *ATAD1* gene. There was not any significant difference between the mRNA levels of control individuals and individuals that are homozygous for the variant (Figure 5.49).

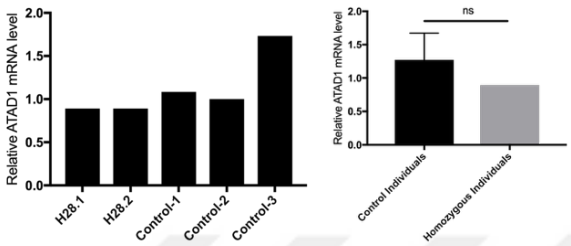


Figure 5.49. RT-PCR analysis of mRNA levels of *ATAD1* gene for H28 family (left panel) and comparison of mRNA levels in affected individuals and healthy controls (right panel).

Western- blot analysis performed for H28.1 and H28.1 revealed a significant decrease in ATAD1 protein level in affected individuals compared to healthy controls. Student t-test analyses requires at least three different values and this limit of biological replicate data is not reached because only two affected individuals were analyzed. Thus, technical replicate results were combined to obtain a statistically meaningful comparison (right panel in the Figure 5.50). Technical replicate results are given in Figure C.1, C.2.

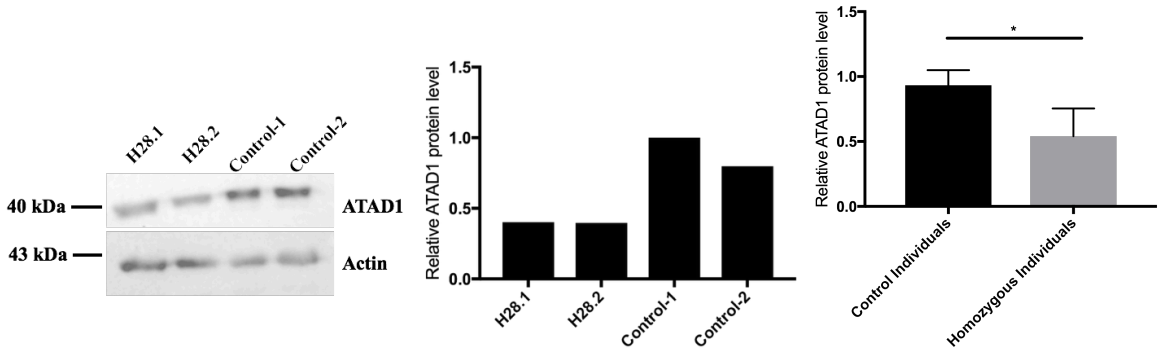


Figure 5.50. Western blot analyses of ATAD1 protein performed for H28.1 and H28.2 (left and middle panels) and comparison protein levels in affected individuals and healthy controls (right panel) (*p=0.0365).

Immunocytochemical assays were performed for ATAD1 protein. It is known that ATAD1 is localized to the nucleus, and mitochondria. There was any localization alteration in immunostaining images of the lymphoblasotoid cell lines of affected individuals compared to carries and control individuals (Figure 5.51). Since mitochondrial staining experiments were not informative due to technical problems in immunostaining of this organelle, we could not analyze morphological changes in mitochondrial if there is any. Piard *et al*., did not observed any alterations mitochondrial morphology and respiration however they determined that AMPAR trafficking was altered in *Atad1-/-* primary cortical neuron cultures (2018). Since it is known that ATAD1 plays role in AMPAR internalization (Zhang *et al*., 2011), this might be suggested as the underlying reason causing HSP in the patients in family H28.

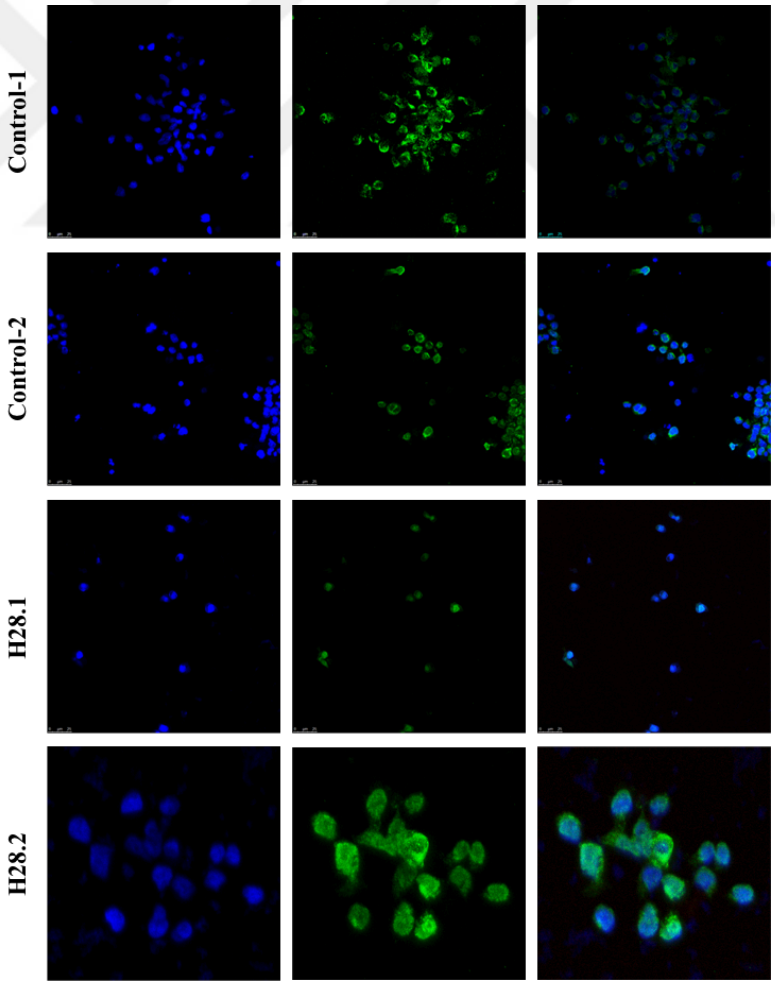


Figure 5.51. Immunostaining for ATAD1 protein (blue:DAPI, green:ATAD1).

5.4.1.2. Family H55 and *CYP27A1* and *ZNF142* Genes. The mRNA levels of the *CYP27A1* and *ZNF142* genes were monitored by RT-PCR experiment. Control individuals did not vary significantly from the individuals that are heterozygous or homozygous for the variant (Figure 5.52, 5.53).

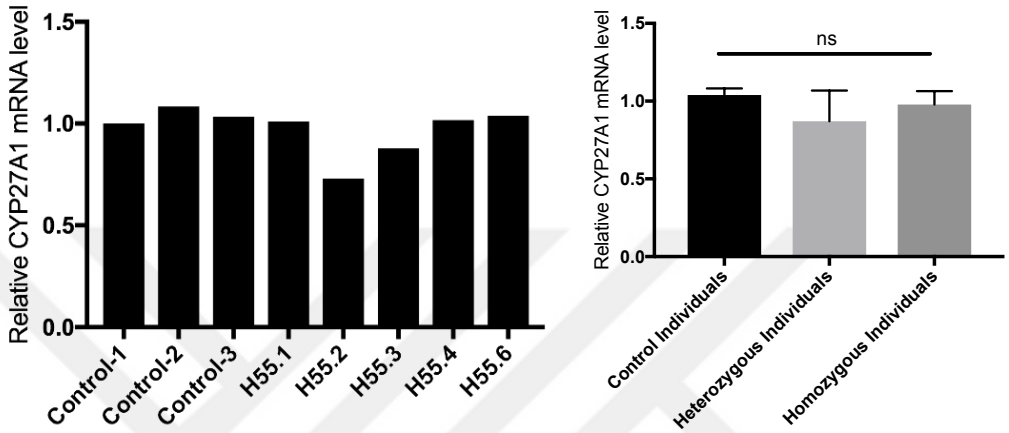


Figure 5.52. RT-PCR analysis of mRNA levels of *CYP27A1* gene for H55 family (left panel) and comparison of mRNA levels in carrier and affected individuals and healthy controls (right panel).

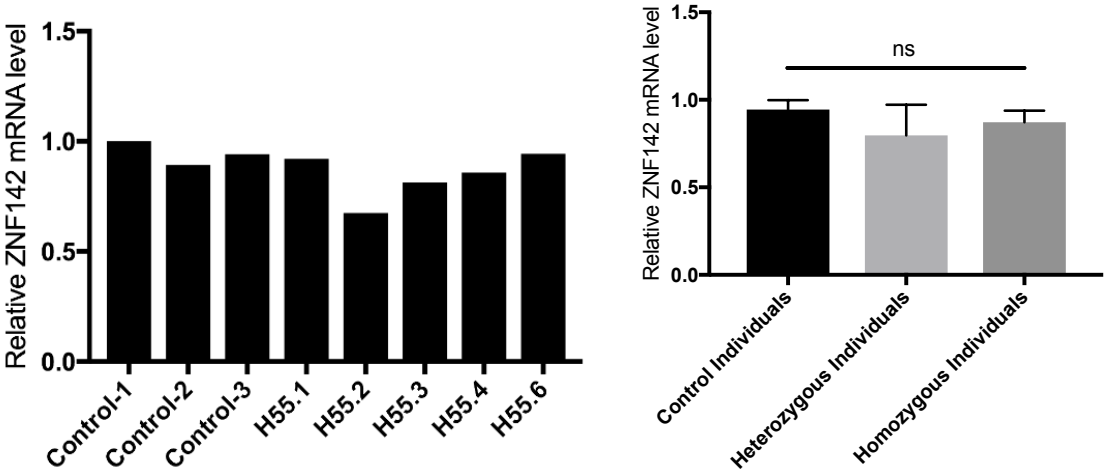


Figure 5.53. RT-PCR analysis of mRNA levels of *ZNF142* gene for H55 family (left panel) and its comparison in carrier and affected individuals to healthy controls (right panel).

CYP27A1 protein levels was analyzed in family H55. A significant decrease in affected individuals was observed compared to individuals that are heterozygous for the variant (Figure 5.54). Technical replicate result is given in Figure C.3.

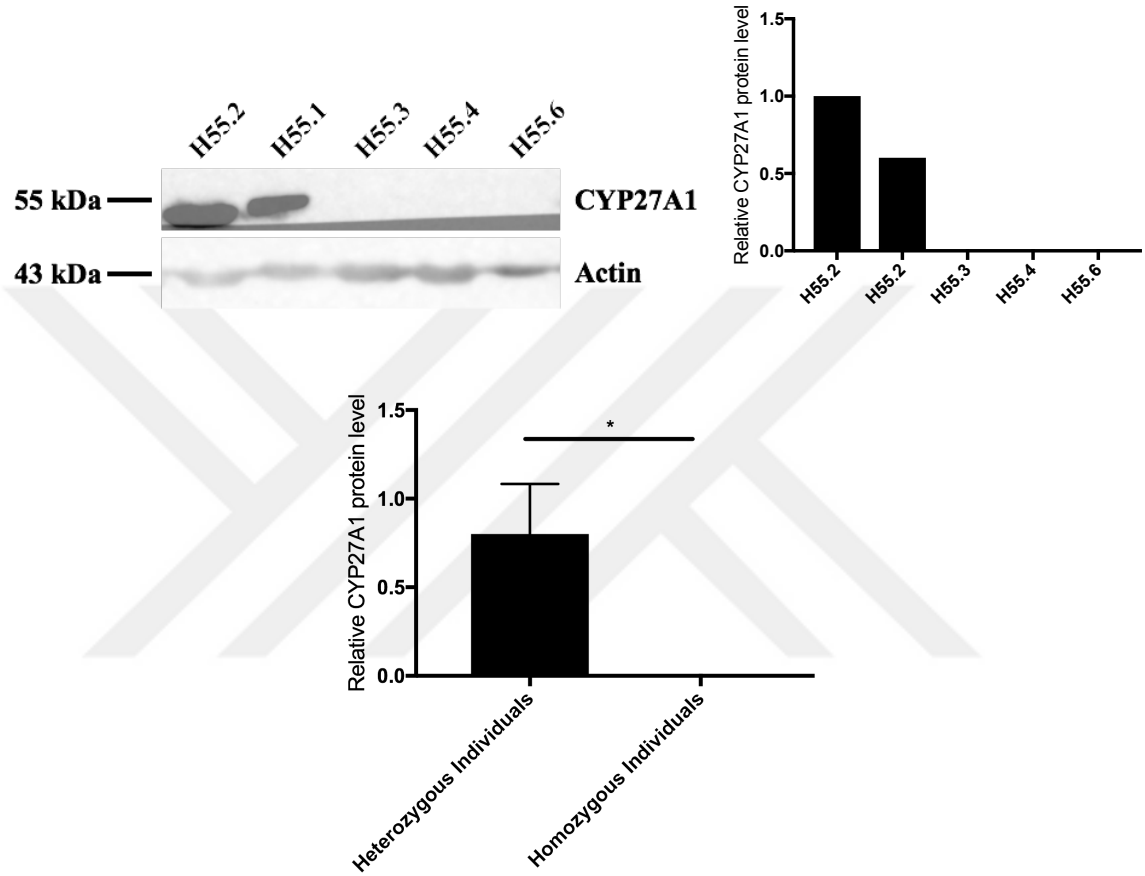


Figure 5.54. Western blot analyses of CYP27A1 protein performed for H55 family (upper panels) and its comparison in carrier to affected individuals (lower panel) (*p=0.0125).

Western-blot analysis and immunostaining experiments could not be performed for ZNF142 protein due to difficulties in optimization of ZNF142 antibody.

In immunostaining experiments performed for CYP27A1 protein, which has mitochondrial localization. Affected individuals H55.3 and H55.6 had less protein compared to their father H55.2 in accordance with western-blot experiment (Figure 5.55). Mitochondrial staining did not reveal informative result therefore it was possible to compare the mitochondria of carriers and affected individuals.

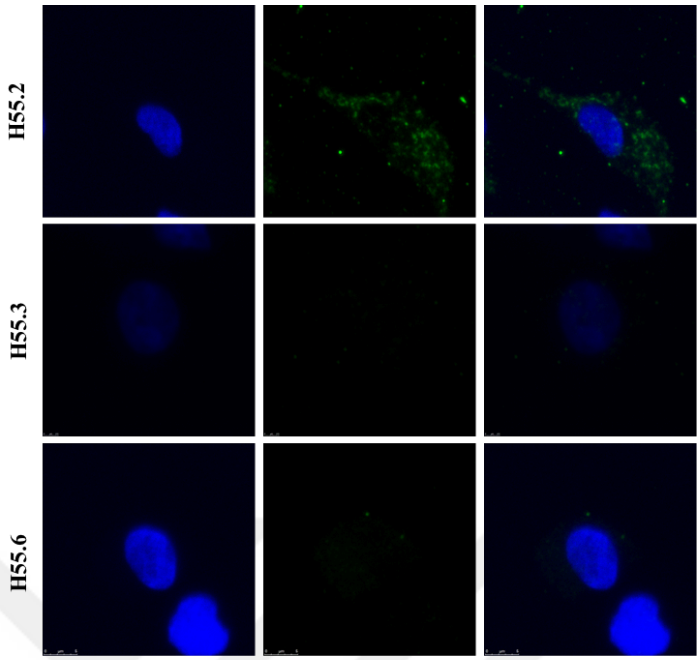


Figure 5.55. Immunostaining for CYP27A1 protein (blue:DAPI, green:CYP27A1).

5.4.1.3. Family H57 and *SAMHD1* Gene. RT-PCR experiments were performed using cDNAs obtained from lymphoblastoid cell lines of the H57 family members. There was a significant decrease in the SAMHD1 mRNA level of control individuals compared to family members (Figure 5.56). However, there was not any difference among heterozygous (H57.1, H57.2) and homozygous individuals (H57.3, H57.4).

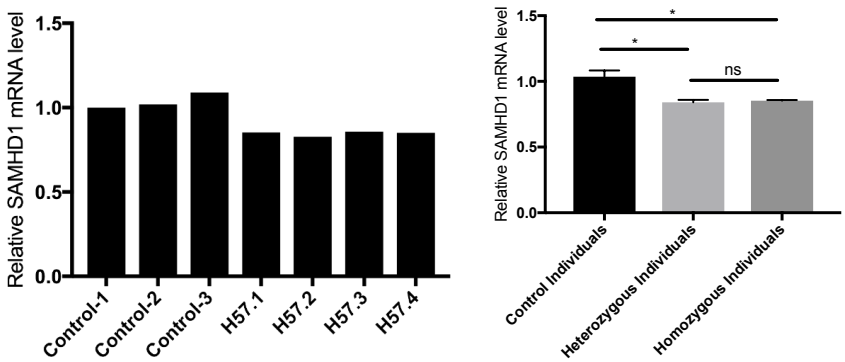


Figure 5.56. RT-PCR analysis of mRNA levels of *SAMHD1* gene for H57 family (left panel) and its comparison in carrier and affected individuals to healthy controls (right panel) (p=0.0137).

To understand how decrease in mRNA level effected protein level, western-blot analyses was performed. Individuals that are heterozygous for the variant had a significantly lower protein level compared to healthy controls. Moreover, protein levels of affected individuals were significantly less than protein level of the carrier individuals (Figure 5.57). Technical replicate results are given in Figure C.4, C.5.

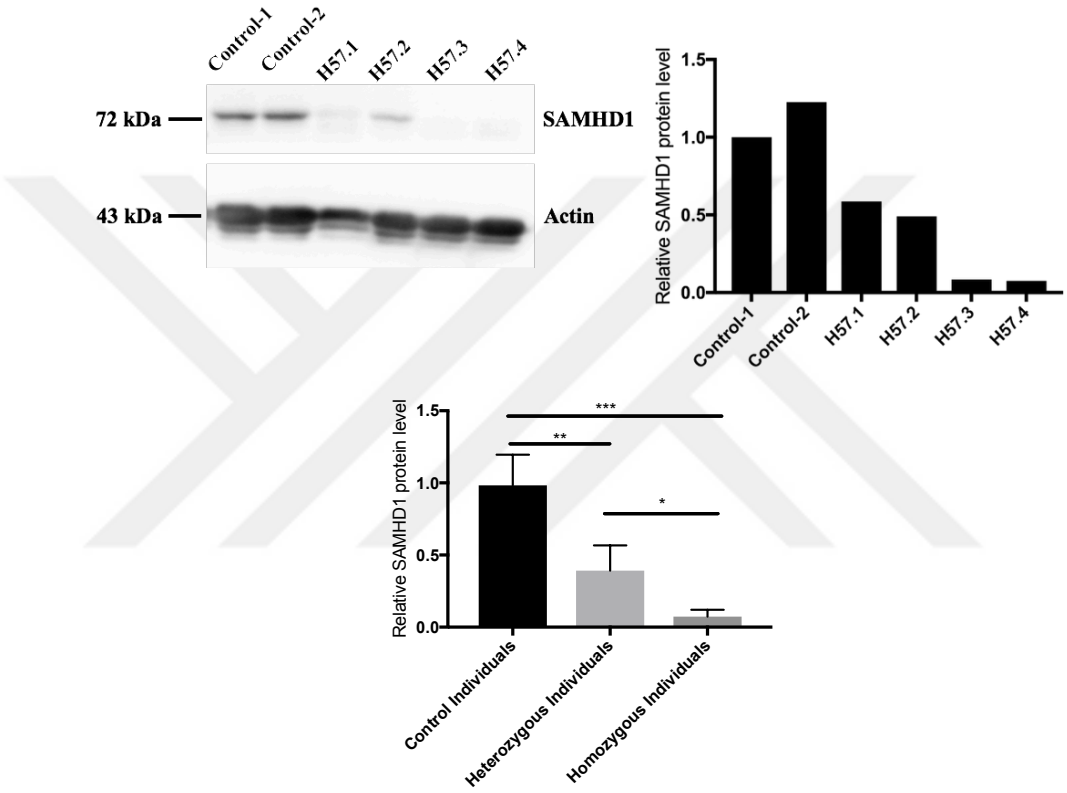


Figure 5.57. Western blot analyses of SAMHD1 protein performed for H57 family (left anf middle panels) and its comparison in carrier to affected individuals (lower panel) (*p=0.0126, **p=0,0051, ***p=0,0002).

Immunostaining experiment was performed for family H57. SAMHD1 is mainly localized to the nucleoplasm. Localization of the protein to the plasma membrane is also known. When the immunostaining images of family members and control individual was analyzed, there was not any mislocalization observed in the affected individual (figure 5.58, 5.59). Although mislocalization of the protein seems not to be affected by this mutation, might affect the protein function in other ways such as altering enzymatic activity or protein binding affinity.

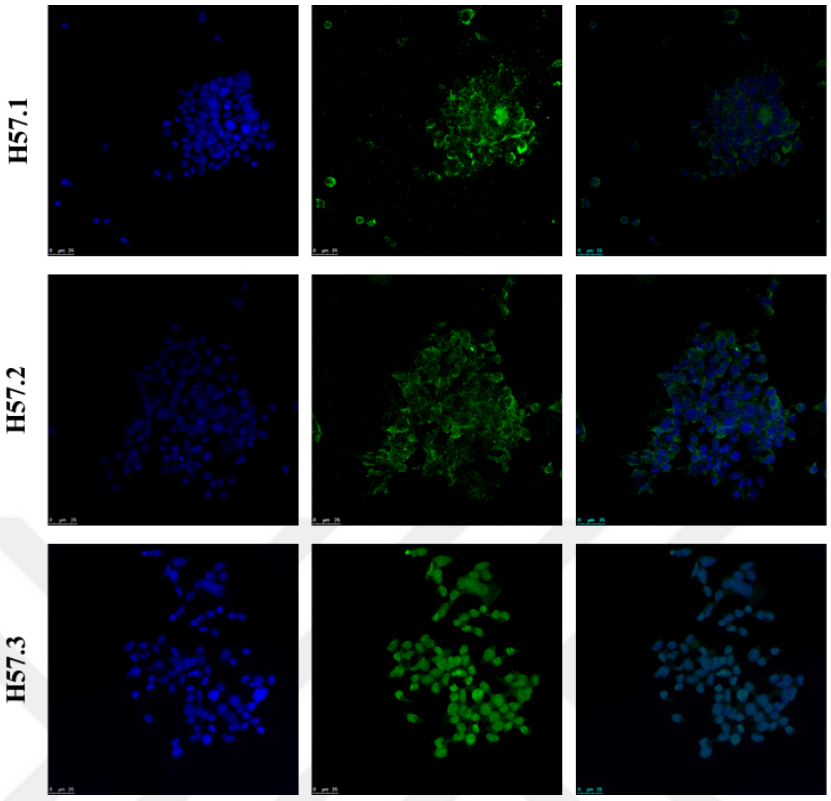


Figure 5.58. Immunostaining for SAMHD1 protein for family H57 (blue:DAPI, green: SAMHD1).

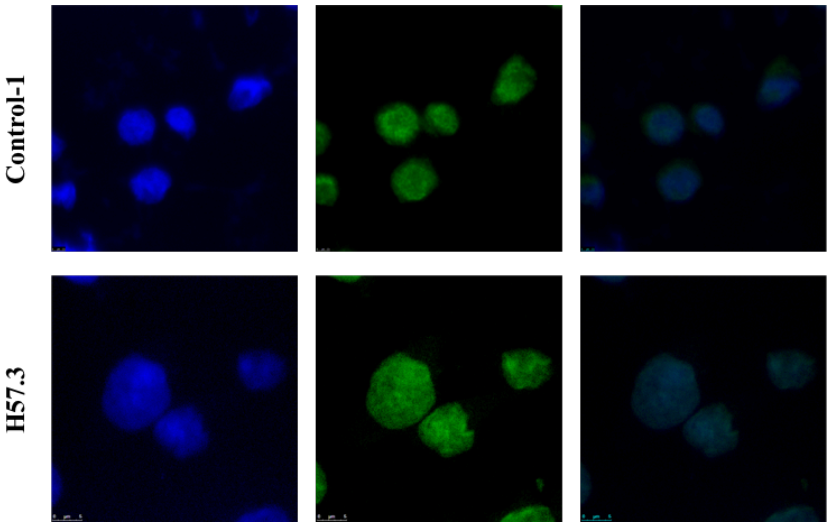


Figure 5.59. Immunostaining for SAMHD1 protein for H57.3 and control individual (blue:DAPI, green: SAMHD1).

5.4.1.4. Family H142 and *SEMA3C* Gene. The mRNA levels of *SEMA3C* gene was monitored by quantitative RT-PCR experiment. There was not any difference in mRNA expression level of *SEMA3C* gene in affected individuals (H142.3, H142.4, H142.5) compared to carriers (H142.1, H142,6) and family members compared to healthy controls (Figure 5.60).

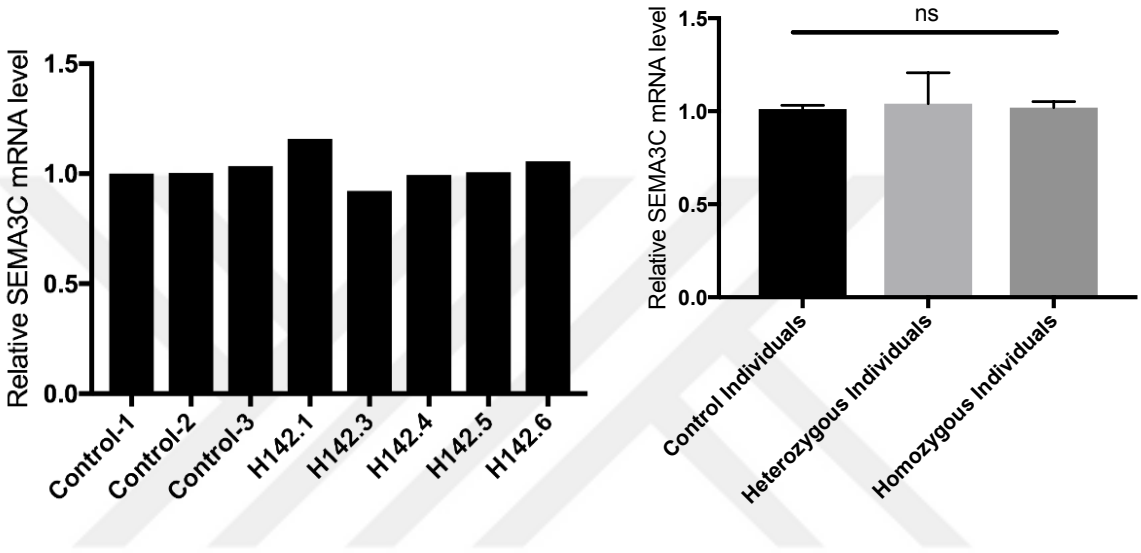


Figure 5.60. RT-PCR analysis of mRNA levels of *SEMA3C* gene for H142 family (left panel) and comparison of carrier and affected individuals to healthy controls (right panel).

Protein level of SEMA3C could not be monitored in family H142 by Western-blot analyses due to troubles in optimization of SEMA3C antibody.

In immonostaining experiments performed for family H142, there was not any difference observed among affected individuals H142.3, H142.5 and unaffected individual H142.7 (Figure 5.57). This situation might arise from a technical problem such as unspecific staining by antibody. It is known that SEMA3C is expressed in all tissues and it localizes to Golgi, cytosol and plasma membrane. Thus, the protein is expected to be expressed in lymphoblastoid cell lines. However, the real effect of any variant on SEMA3C, which functions in axonal guidance, might not be directly observed in a blood-driven cell line.

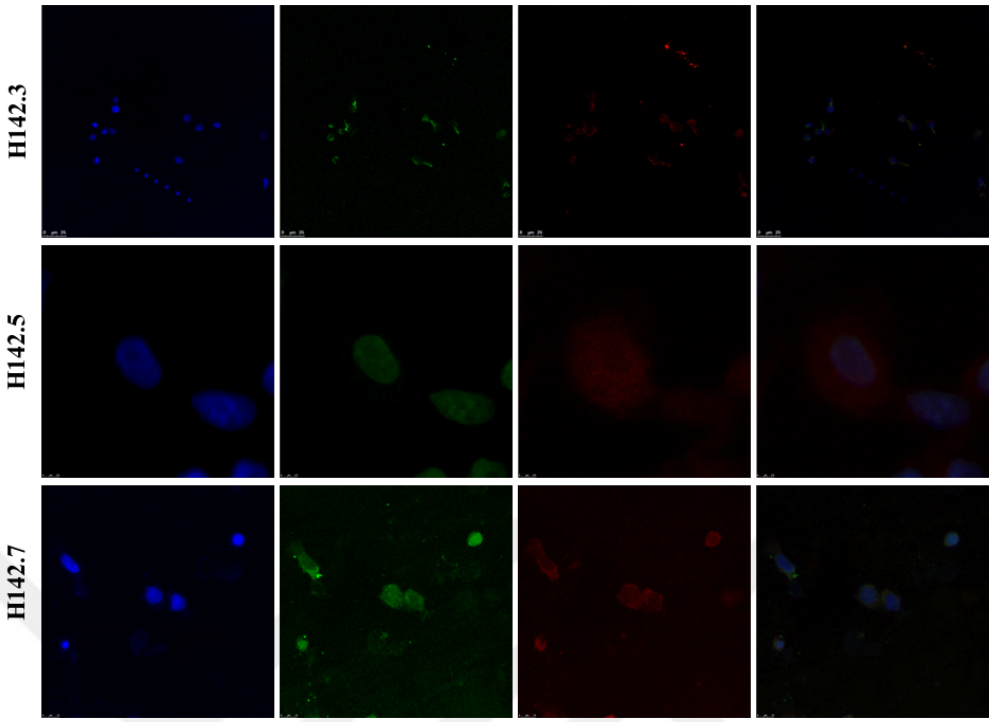


Figure 5.61. Immunostaining for SEMA3C protein (blue:DAPI, green: Golgi apparatus, red:SEMA3C).

6. DISCUSSION

In this study, twenty-seven AR-HSP families were analyzed by whole exome sequencing with the aim of identifying novel genes and mutations causing HSP. Analyses revealed causative variants in *SPG11* gene in four families. In another four families, SPG5A-associated *CYP7B1, SPG7, SPG15* and IAHSP-associated *ALS2* gene mutations were found to be causative. Besides, in two families, differential diagnosis was provided by identifying mutations in *SACS* and *CYP27A1* genes, which causes Autosomal Recessive Spastic Ataxia of Charlevoix-Saguenay (ARSACS) and Cerebrotendinous xanthomatosis (CTX), respectively. In the family with *CYP27A1* gene mutation, *ZNF142* gene was also considered as candidate gene. *KIF1C* gene was associated with disease in one family and this finding was reported as SPG58 gene in the scope of a partnership with Prof. Rebecca Schüle (University of Tubingen, Germany) and Prof. Stefan Zühcner (University of Miami, USA) (Oteyza *et al*., 2014). In still one more family, *PLA2G6* gene was identified as a novel HSP gene and it was reported as a gene that expands the clinical spectrum of PLA2G6-associated neurodegeneration in a way to include HSP (Ozes *et al.,* 2017). Moreover, *ATAD1, SAMHD1,* and *SEMA3C* genes were identified as promising novel HSP gene candidates. Taken together, causative variants were identified in fifteen among twenty-seven families (55%).

Most of the genes identified in the autosomal recessive diseases in recent years were determined in the studies that make use of WES. Moreover, combination of WES and homozygosity mapping is widely used in genetic diagnosis of AR diseases. Although, these useful genetic tools were applied in this study, the causative gene was not identified in half of the patients. Indeed, this ratio was not surprising because in a study reporting findings of 608 HSP patients, the genetic diagnosis rate was given as 50% (Schüle *et al*., 2016). Half of the cases without genetic diagnosis indicates presence of further genetic heterogeneity in HSP and unidentified HSP genes. Technical limitations of WES might also explain some of the undiagnosed cases. WES covers approximately 2% of the human genome and although this region contains 85% of the pathogenic variants, 15% of this region cannot be captured by WES (Chilamakuri *et al*., 2014). Therefore, some of the causative variants might not be

detected due to low coverage rate of WES. Although, an epigenetic effect was not associated with HSP previously to our knowledge, these type of alterations might explain some of unexplained cases. Incomplete penetrance is common in HSP families that may cause misinterpretation of pedigrees. These families with dominant inheritance may be recognized as autosomal recessive and their WES data may be analyzed from a wrong perspective. As a precaution, AD-HSP genes were also analyzed in our families, however, only AR inheritance was considered for identification of novel genes.

Among fifteen families that received a genetic diagnosis, eight families had mutations in known HSP genes and four of them had *SPG11* mutations. 20% of AR-HSP cases and 59% of the AR-HSP-TCC cases can be explained by *SPG11* mutations in literature (Finsterer *et al*.,2012; Stevanin *et al*., 2008). In this study, 52% (fourteen families) of the 27 families analyzed had complicated HSP and seven families had TCC, additionally. All four families with *SPG11* mutations had AR-HSP-TCC. Thus, in accordance with literature, *SPG11* mutations explained 15% of AR-HSP cases and in 57% of AR-HSP-TCC cases in our cohort. *SPG15* is the second highly mutated gene in AR-HSP-TCC cases and this gene was found to be mutated in one of our families with AR-HSP-TCC, as expected. *CYP7B1* (SPG5A) is the second highly mutated gene in AR-HSP cases with a ratio of 8% and it causes 16% of pure AR-HSP cases (Biancheri *et al*., 2009). In our cohort, *CYP7B1* was identified in one family. *SPG7* gene that was identified in one family in our cohort, is among the highly mutated genes with 1,5-7% rate (Schüle and Schlös, 2011). The frequencies of other HSP genes are very low compared to aforementioned genes. As might be observed from our literature-compatible findings, clinical findings might accelerate genetic diagnosis. Initial screening of *CYP7B1* gene in pure AR-HSP cases, and *SPG11* gene in AR-HSP-TCC cases, would be a time and cost-effective way.

Identification of *SACS* as the causative gene in one family clarified the diagnosis of the patients as spastic ataxia. This finding indicates the importance of genetic analyses in differential diagnosis. *SACS* gene codes for the *Sacsin* protein and it is associated with autosomal recessive Spastic Ataxia of Charlevoix-Saguenay (ARSACS). Although the function of Sacsin is not known thoroughly, it is expressed in central nervous system and it might have a role in ubiquitin-proteasome system and Hsp70 chaperone function. This chaperone regulates especially neurodegeneration-related mutant proteins that form

aggregates inside the cell (Parfitt *et al*., 2009). ARSACS develops mainly in the early-childhood and cerebellar ataxia, pyramidal findings and peripheral neuropathy might be observed in the patients. Cases with peripheral neuropathy as the initial symptom that emerge during adulthood were also reported (Baets *et al*., 2010). Age of onset in patient P627 was 30 years and she had peripheral neuropathy, initially. Other findings in the patient including proximal weakness in lower extremities, spinal cord involvement, cerebellar findings, vibration sense loss and bilateral dysmetria indicates the possibility of ARSACS for P627. After consulting the clinician of the patient, it was concluded that the patient has ARSACS. Genes such as *KIF1C* and *GBA2* that are common to HSP and ataxia have been reported previously. There is still the possibility that *SACS* might be one of those genes common to HSP and ataxia.

Differential diagnosis was also provided to family H55 with the identification of a *CYP27A1* mutation. Cerebrotendinous xanthomatosis (CTX) is caused by mutations of this gene and it has symptoms compatible to HSP. CTX is inherited in autosomal recessive manner and it is a rare lipid storage disease by progressive neurologic dysfunction. The symptoms might include diarrhea, cataracts, spastic paraplegia, pes cavus, epilepsy, cerebellar ataxia, systemic spinal cord involvement, peripheral neuropathy, premature atherosclerosis, dementia, and psychiatric disturbances. *CYP27A1* gene encodes for the mitochondrial enzyme sterol 27-hydroxylase, a member of the cytochrome P450 superfamily of enzymes as *CYP7B1* (SPG5A) gene. This protein is important for overall cholesterol homeostasis. CYP27A1 deficiency negatively affects the conversion of cholesterol to bile acids and leads to accumulation of intermediate products in this metabolic pathway (Gallus *et al*., 2006; Sasamura *et al.,* 2017).

Since the signs and symptoms of CTX might be observed in other neurological diseases one by one or in combinations, diagnosis of CTX becomes very complicated and patients get diagnosis of HSP and multiple sclerosis in some cases (Nicholls *et al.,* 2015). This information might explain the diagnosis of HSP in this family. One important outcome of this finding comes from the information that when patients are treated with chenodeoxycholic acid (CDCA), a natural form of bile acid, progress of the neurological symptoms might be decelerated (Rosafio *et al*., 2016; Sekijima *et al*., 2018). Therefore, the clinician was informed about the differential diagnosis obtained by this study which

hopefully be beneficial for the family. Lymphoblastoid cell lines of the family members were analyzed and Western blot and immunostaining experiments showed that protein levels of the affected individuals are significantly reduced compared to unaffected individuals. Therefore, further evidence was obtained that the splice region variant identified in the family H55 was the disease-causing variant. The other candidate variant identified for the H55 family is in the *ZNF142* gene. Given the severe and complex neurodegeneration seen in patients and the function of the *ZNF142* and *CYP27A1* genes, both genes were predicted as a possible cause of pathogenesis in the family. Although the function of the *ZNF142* gene is not known in detail, it is suggested to be involved in transcriptional regulation like other zinc finger protein family members and have effect in various steps including monitoring cell growth, proliferation, differentiation, and apoptosis (Dey *et al*., 2015). This protein family carry Cys2-His2 (C2H2) zinc finger domains. According to analyses performed on lymphoblastoid cell lines, mRNA level of the ZNF142 was not altered in the patients and the Western blot analyses of ZNF142 was not informative for this protein. Since the supporting evidence for the effect of ZNF142 was not obtained and the variant in *CYP27A1* had an important effect on protein level, it might be concluded the variant identified *CYP27A1* gene is a highly penetrant Mendelizing variant (HPMV) and the variant in *ZNF142* might have an additive effect on disease pathogenesis.

The main aim of our project was to identify new AR-HSP genes and in accordance with this aim three genes, that were not associated with HSP previously, were identified as causative genes in three families. These three genes were *KIF1C, PLA2G6* and *SAMHD1* that were directly identified by analysis of the WES data in families H52, H53 and H57, respectively. In two further families, H28 and H142, not only WES but homozygosity mapping allowed us to identify *ATAD1* and *SEMA3C* genes as candidates. The significance of the variants identified in these genes was supported by population screening, gene prioritization algorithms and / or three-dimensional modeling studies.

KIF1C gene identified in family H52 was reported with another family bearing the same gene mutation that was identified by Prof. Rebecca Schüle (University of Tubingen, Germany). Mutations in KIF1C, a kinesin-type microtubule-dependent motor protein, are defined as the cause of SPG58 (Caballero Oteyza *et al*., 2014). To underline the interesting finding in this family, genotypes and phenotypes of the individuals should be considered. In

clinical examination, unsteady gait starting with the adolescence was observed in both patients however; disease progression was very different. While H52.5, that is homozygous for the variant, lost his ability to walk around age of 40, the older brother H52.3, that is heterozygous for the variant, had mild spastic gait at age 46 and his daily activities was not interfered with this symptom. Lower limb spastic paraparesis was observed in both patients (mild for H52.3 and severe for H52.5). Mild cervical dystonia and loss of vibration sense was observed for H52.5. Mother was heterozygous for the variant and she was reported normal after neurological examination Manifesting a mild phenotype in heterozygous mutation state and a severe phenotype when both alleles are mutated showed that this gene can cause AD and AR inheritance. Moreover, phenotype of mother suggests incomplete penetrance in this family. This genotype-phenotype correlation implicated presence of AD and AR inheritance and incomplete penetrance in the same family that reflects how complex HSP inheritance can be in some families.

The other HSP gene identified in this study was *PLA2G6*. The c.2239G>A (p.747R>W) variant identified, was reported in a patient with dystonia-parkinsonism (Paisan-Ruiz *et al*., 2009) and in another patient with early-onset Parkinsonism (Giri *et al*., 2016). Patient H53.3 had progressive spasticity and ataxia suggesting a milder phenotype compared to his brother and the patient reported by Paisan-Ruiz *et al.,* (2009). Cranial nerves were intact and he had normal intellectual performance. The patient was able to walk spastically without any assistance. The symptoms in his brother H53.4 had started at the age of 21 and progressed more rapid compared to the index case. He had hyperactive deep tendon reflexes, achilles clonus, bilaterally positive babinski sign and a wide based spastic gait with marked ataxia. The patient was unable to walk in his recent examination at the age of 32. With our finding, HSP was added to the list of PLA2G6-associated neurodegeneration (PLAN) including Parkinsonism, neurodegeneration with brain iron accumulation (NBIA), and infantile neuroaxonal dystrophy (INAD). PLA2G6 is not the first gene that is common to HSP and NBIA. Another HSP gene *FA2H* (SPG35) was also associated with a group of neurodegenerative disorders including demyelinating leukodystrophy, and NBIA (Kruer *et al*., 2010; Dick *et al.,* 2010). Moreover, recently identified HSP gene *C19ORF12*, is related to neurodegeneration with brain iron accumulation-4 (NBIA4) (Landoure *et al.,* 2013). Therefore, this mutation identified in PLA2G6 gene in an HSP patient makes contribution to the list of genes that are common in neurodegenerative diseases (Ozes *et al.,* 2017).

The *ATAD1* gene identified in family H28 is localized to SPG27 locus (10q22.1-q24.1). HSP was associated with SPG27 locus in a French family, previously but the causative gene was unknown. The affected individuals in this French family were reported to have pure HSP with age of onset between 25-45 years (Meijer *et al.,* 2004). Clinical findings of patients in family H28 indicated pure HSP and the age at onset was over 25 years. Therefore, it is highly probable that *ATAD1* is the gene that causes SPG27. If this suggestion is confirmed, the goal of identifying genes for SPG loci, for which the causative gene was not identified, will be reached.

ATAD1 (THORASE / ATPase Family, AAA Domain Containing 1) is a member of AAA + ATPase family that functions in protein transport, membrane dynamics and protein complex formation and dissociation, and plays a role in the regulation of surface expression of α-amino-3-hydroxy-5-methylisoxazole-4-proprionate receptors (AMPAR). Mutations in ATAD1 gene lead to impairments in internalization of AMPARs causing an increase in postsynaptic receptor density which finally result in neuronal death (Zhang *et al*., 2011). Another function of the ATAD1 was identified in a study by Chen *et al,* (2014). In this study, it was shown that ATAD1 is required for degradation of tail-anchored (TA) proteins that are mislocalized to mitochondria. In knock-out mice and HeLa cell lines, mitochondrial morphology was altered affecting the function of the organelle. Detection of localization of ATAD1 to peroxisome suggest a similar function in both organelles (Okreglak *et al*., 2015; Grimm *et al.,* 2015). As mentioned in the introduction part (Section 1.5), mitochondrial impairment is one of the major reasons causing HSP. SPG7 gene (paraplegia), which is frequently mutated in AR-HSP patients, and SPG4 (spastin), which is most frequently mutated in AD-HSP, also belong to the AAA protease family and this functional similarity strengthens our findings regarding *ATAD1* as a novel HSP gene.

In a recent study, a loss of function mutation (p.E276X) in *ATAD1* gene was identified in a Kuwaiti family, diagnosed with a neurological disorder with encephalopathy, hypertonia, and seizures. Disease causes respiratory failure and death in affected individuals (Ahrens-Nicklas *et al.,* 2017). In this study, perampanel, an AMPAR antagonist, was used to reverse the effect of the ATAD1 mutation. In ATAD1 knockout mice, the average life-span was prolonged with the use of perampanel. When perampanel was given to two affected individuals from the family, deceleration in disease symptoms was observed, however,

progression of the disease was not inhibited completely. Another study reported p.His357Argfs*15 variant in *ATAD1* as causative gene in family diagnosed with lethal encephalopathy and congenital stiffness (Piard *et al.,* 2018). In this study, defect in AMPAR internalization was identified as the molecular mechanism underlying the symptoms observed in the patients. Although they were not able detect a morphological alteration in mitochondria, it was suggested that combination of impairments in AMPAR recycling and mitochondrial function might affect the severity of the disease. Compared to the reported families, H28 family has a different phenotype. However, this finding is not surprising because it is known that there are overlapping genes between HSP and other neurological diseases. For example, it is known that the *SPG11* gene is responsible for ALS Type 5 and CMT Type 2X (Klebe *et al.,* 2015). Patients' reduced protein level observed by Western-blot analyses supports our suggestion of *ATAD1* being a novel HSP gene. Overall, when the function of ATAD1 and its relation to neurological disorders are considered *ATAD1* is a good candidate for HSP pathogenesis.

The causative gene identified in family H57 was *SAMHD1* (SAM And HD Domain Containing Deoxynucleoside Triphosphate Triphosphohydrolase 1). The gene codes for sterile a motif and histidine-aspartic domain (HD) containing protein 1. This protein catalyzes hydrolysis of the intercellular dNTPs producing deoxynucleosides and inorganic triphosphates (Rossi, 2014). As Goldstone et al., reported SAMHD1 mainly functions on dGTPs making the protein a deoxyguanosine triphosphate triphosphohydrolase (2011). It is suggested that suppressing intercellular dNTP levels, blocks both function of endogeneous reverse transcription (RT) and effect of RT due to viral infection (Goldstone *et al.,* 2011). In case of an inhibition of SAMHD1 activity due to either a loss of function mutation or presence of Vpx, which is a lentiviral protein targeting SAMHD1 for proteosomal degradation, the dNTP levels are increased leading to autoimmune response or viral infection, respectively.

SAMHD1 mutations are related to Aicardi-Goutiéres syndrome 5 (AGS5), Chilblain lupus 2 (CHBL2), HIV-1 infection and chronic lymphocytic leukemia (CLL) (Rice *et al.,* 2009; Ravenscroft *et al.,* 2011; Laguette *et al.,* 2011; Clifford *et al.,* 2014). Chilblain lupus is defined as a rare skin-related form of systemic lupus erythematosus (SLE). Bluish-red

swellings and nodules on the hands, feet, ears, and nose are mainly observed in the patients. AGS is an inflammatory encephalopathy starting at infancy and findings include cerebral atrophy, chronic CSF lymphocytosis, leukodystrophy, intracranial calcifications, increased interferon alpha (IFNα) levels in cerebrospinal fluid (CSF), spasticity, and chilblains (Ali *et al.,* 2006; Stephenson, 2008; Orcesi *et al.,* 2009). AGS is a genetically heterogeneous disease. Genes and corresponding ASG types are TREX1/AGS1, RNASEH2B/AGS2, RNASEH2C/AGS3, RNASEH2A/AGS4, ADAR1/AGS6, and IFIH1/AGS7. There are 17 families with *SAMHD1* mutations reported (Rice *et al.,* 2009; Dale *et al.,* 2010; Leshinsky-Silver, *et al.,* 2011). While majority of the patients display the main symptoms explained above, two siblings reported by Dale *et al.,* (2010) represent the mild phenotype of Aicardi-Goutiéres syndrome spectrum. Older sibling had mildly delayed psychomotor development, postnatal microcephaly, and chilblains and mild learning difficulties while her brother had lower limb spasticity, hyperreflexia, and presence of chilblains but he neither has microcephaly nor learning disabilities. The phenotype of the affected brother might resemble to the phenotypes of H57.3 and H57.4 however the presence of chilblains seems to be the differentiating factor in diagnosis. In 2015, Crow *et al.,* reported that they identified three patients with SAMHD1 mutations that were diagnosed with non-syndromic spastic paraplegia (2015). However, the authors did not give the identified mutations and make a clear correlation with phenotype in their paper. The affected siblings are directed to our laboratory with the pure HSP diagnosis. Age of onset in the affected individuals was around 2 years old. Brisk tendon reflexes in lower extremities and bilateral Babinski were also observed in the patients. Although their spinal MRI (magnetic resonance imaging) was normal, periventricular nonspecific T2 hyperintensities were observed in brain MRI showed. All symptoms were milder in H55.3. H55.4 had myopathic changes in lower extremity muscles. It is known that there is not a clear cut in the diagnosis of some neurological diseases due to overlapping symptoms. For family H57, it could be suggested that identification of *SAMHD1* mutation as causative variant in this family might expand the clinical spectrum of *SAMHD1*-related diseases to HSP with additional findings such as myopathy. The phenotype of patients in H28 family might correlate with the patients reported in the study by Crow *et al*. (2015). However, this study is the first one presenting *SAMHD1* as a novel HSP gene. The western analysis indicating a decreased protein level in affected individuals further supports our finding of *SAMHD1* being causative gene in this family.

Another strong candidate was *SEMA3C* gene identified in the family H142, SEMA3C belongs to the Semaphorin class 3 protein family. In semaphorin family, there are at least seven different protein classes and SEMA3C (Semaphorin 3C) belongs to the secreted class 3 and it codes for a secreted glycoprotein that acts as a chemoattractive in axon guidance (Gonthier *et al*., 2007; Hernandez-Montiel, *et al.,* 2007). SEMA3C also functions in endothelial cell branching and binds to neuropilin1, which takes role in various signaling pathways controlling cell migration (Reidy and Tufro, 2011). The protein was suggested to function in adhesion and invasion of cancer cells (Herman and Meadows, 2007; Malik *et al.,* 2016). It contains a sema domain, an immunoglobulin-like domain, and a C-terminal basic domain. *SEMA3C* gene variants were reported to increase the risk of Hirschsprung disease that cause abnormally enlarged colon in the patients (Kapoor *et al.,* 2015). In the affected individuals, intrinsic ganglion cells are not formed in the myenteric and submucosal plexuses of the gastrointestinal tract (Gunadi *et al*., 2016). The findings of Hirschsprung's disease do not overlap with HSP, but it is known that the same gene may cause different phenotypes. In our experiments performed to support our findings, there were no significant differences in mRNA levels and immunostaining analyzes between unaffected and affected individuals' immortalized B-lymphocytes. When the function of the SEMA3C and the conservation of amino acid where the variant resides are considered, *SEMA3C* is still a good novel HSP gene candidate.

The phenomenon of pleiotropy, observing different phenotypes with same mutated gene and even with same mutation, is reported in several studies and it was also observed in our study. One explanation of this situation might be that different variants of the same gene might alter different down-stream pathways. Type of mutation might also affect the phenotype. For instance, a frameshift deletion causing a truncated protein might have a higher impact on disease pathogenesis while a point mutation might lead to a mild phenotype. The cases for which the same mutation ends up in different phenotypes might be explained by 'differential genetic backgrounds' of the patients. In other words, variants in other genes might affect the phenotype although they cannot lead to disease manifestation on their own (Pang *et al*., 2017). The effect of genetic background might bring an explanation for some of the genetically undiagnosed cases. It might be possible that combination of the variants might lead to disease in family members and a single strong variant that follow expected inheritance might not be identified in some cases. The

significant difference in the number of modifier variants in affected individuals compared to healthy controls that has been identified in this study can be considered as a support for this hypothesis.

As a result, in accordance with the project aim, further evidence was provided that there are still unidentified genes responsible for AR-HSP, the importance of genetic analysis in differential diagnosis was emphasized, and five new candidate HSP genes, one of them residing in SPG27 locus, were identified. Our findings highlighted the requirement of initial screening of *SPG11* gene in AR-HSP-TCC cases and *CYP7B1* gene in pure AR-HSP, in our population, as in other populations.

7. CONCLUSION

In 1986, Kenwrick *et al*., reported their findings about first HSP locus (1986) and in the last 32 years, 78 loci and 63 HSP genes were identified. The number of AR-HSP loci was 32 and that of AR-HSP genes was 23 at the end of 2013, when this study was proposed. In five years' time, these numbers were increased to 53 and 47, respectively. This big step in the genetics of AR-HSP, mainly taken with the propellant power of whole exome sequencing and homozygosity mapping.

With this study, we have accomplished applying these world-wide-accepted approaches to our AR-HSP cohort and identified five novel HSP genes in a relatively small cohort of 27 patients (Table 7.1). This contribution to HSP genetics, is important not only for HSP studies but also for other neurological diseases because each novel gene identified in these diseases fills a gap in the mechanisms underlying neurodegeneration. With the novel genes associated to HSP, we also contributed to the HSP genetics of Turkish population. As the HSP genes and their frequencies in our population characterized further, a realistic genetic diagnostic panel can be designed that meet the requirements of our population. Moreover, this study carried the genetic findings one step further by analyzing the effect of candidate variants *in vitro*.

Table 7.1. Five novel HSP genes identified in this study.

Family	**Gene**	**Variant**
H28	ATAD1	c.701T>C, p.Met234Thr
H52	KIF1C	c.901G>A, p.Arg301Gly
H53	PLA2G6	c.2239G>A, p.Arg747>Trp
H57	SAMHD1	c.533T>C, p.L178P
H142	SEMA3C	c.1286A>G, p.Tyr429Cys

As WES data of both affected and unaffected individuals accumulate, a clear difference in the number of effective variants among these two groups has emerged. The

rare variants that predominate the affected individuals' WES data should be taken up comprehensively because without understanding their effects, it might not be possible to explain variances on phenotypes. WES and following genetic analyses of AR-HSP data are more straightforward compared to that of AD-HSP. First reason is that the number of homozygous variants are smaller than the number of heterozygous variants making segregation analyses faster. Other reason is the possibility of analyzing another affected individual to confirm segregation of variant. These reasons highlight the 'WES studies on AR families' as the most feasible method to identify effective rare variants in the families.

One drawback of WES studies comes out when a novel gene is identified in a single family because to provide a valid link between gene and disease a second family with a causative mutation in the same gene is required. Such a drawback can be crossed out by searching other families in open-source WES databases. The number of such databases are on rise world-wide, however, a WES database for both affected and unaffected individuals from our population would serve greatly to analyze the situation of a target variant in our population.

REFERENCES

Abou Jamra, R., O. Philippe, A. Raas-Rothschild, S. H. Eck, E. Graf, R. Buchert, ..., L. Colleaux, 2011, "Adaptor protein complex 4 deficiency causes severe autosomal-recessive intellectual disability, progressive spastic paraplegia, shy character, and short stature", *American Journal of Human Genetics,* Vol. 88, No. 6, pp. 788-795.

Adzhubei, I. A., S. Schmidt, L. Peshkin, V. E. Ramensky, A. Gerasimova, P. Bork, A. S. Kondrashov and S. R. Sunyaev, 2010, "A method and server for predicting damaging missense mutations", *Nature Methods*, Vol. 7, No. 4, pp. 248-249.

Ahrens-Nicklas, R. C., G. K. Umanah, N. Sondheimer, M. A. Deardorff, A. B. Wilkens, L. K. Conlin, ..., E. D. Marsh, 2017, "Precision therapy for a new disorder of AMPA receptor recycling due to mutations in ATAD1", *Neurology Genetics*, Vol. 3, No. 1, pp. e130.

Ali, M., L. J. Highet, D. Lacombe, C. Goizet, M. D. King, U. Tacke, ..., Y. J. Crow, 2006, "A second locus for Aicardi-Goutieres syndrome at chromosome 13q14-21", *Journal of Medical Genetics*, Vol. 43, No. 5, pp. 444-450.

Alkan, C., P. Kavak, M. Somel, O. Gokcumen, S. Ugurlu, C. Saygi, ..., C. Bekpen, 2014, "Whole genome sequencing of Turkish genomes reveals functional private alleles and impact of genetic interactions with Europe, Asia and Africa", *BMC Genomics*, Vol. 15, No. pp. 963.

Antonicka, H., E. Ostergaard, F. Sasarman, W. Weraarpachai, F. Wibrand, A. Pedersen, ..., E. Shoubridge, 2010, "Mutations in C12orf65 in patients with encephalomyopathy and a mitochondrial translation defect", *American Journal of Human Genetics*, Vol. 87, No. 1, pp. 115-122.

Baets, J., T. Deconinck, K. Smets, D. Goossens, P. Van den Bergh, K. Dahan, ..., P. De Jonghe, 2010, "Mutations in SACS cause atypical and late-onset forms of ARSACS", *Neurology*, Vol. 75, No. 13, pp. 1181-1188.

Bakowska, J. C., H. Jupille, P. Fatheddin, R. Puertollano and C. Blackstone, 2007, "Troyer syndrome protein spartin is mono-ubiquitinated and functions in EGF receptor trafficking", *Molecular Biology of the Cell*, Vol. 18, No. 5, pp. 1683-1692.

Battini, R., A. Fogli, D. Borghetti, A. Michelucci, S. Perazza, F. Baldinotti, ..., G. Cioni, 2011, "Clinical and genetic findings in a series of Italian children with pure hereditary spastic paraplegia", *European Journal of Neurology*, Vol. 18, No. 1, pp. 150-157.

Bechara, A., H. Nawabi, F. Moret, A. Yaron, E. Weaver, M. Bozon, K. ..., V. Castellani, 2008, "FAK-MAPK-dependent adhesion disassembly downstream of L1 contributes to semaphorin3A-induced collapse", *EMBO Journal*, Vol. 27, No. 11, pp. 1549-1562.

Bernard-Marissal, N., J. J. Medard, H. Azzedine and R. Chrast, 2015, "Reply: Is SIGMAR1 a confirmed FTD/MND gene?", *Brain*, Vol. 138, No. Pt 11, pp. e394.

Biancheri, R., M. Ciccolella, A. Rossi, A. Tessa, D. Cassandrini, C. Minetti and F. M. Santorelli, 2009, "White matter lesions in spastic paraplegia with mutations in SPG5/CYP7B1", *Neuromuscular Disorders*, Vol. 19, No. 1, pp. 62-65.

Biasini, M., S. Bienert, A. Waterhouse, K. Arnold, G. Studer, T. Schmidt, F. ..., T. Schwede, 2014, "SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information", *Nucleic Acids Research*, Vol. 42, No. Web Server issue, pp. W252-258.

Blackstone, C., 2012, "Cellular pathways of hereditary spastic paraplegia", *Annual Review of Neuroscience*, Vol. 35, No. pp. 25-47.

Burgos, P. V., G. A. Mardones, A. L. Rojas, L. L. daSilva, Y. Prabhu, J. H. Hurley and J. S. Bonifacino, 2010, "Sorting of the Alzheimer's disease amyloid precursor protein mediated by the AP-4 complex", *Developmental Cell*, Vol. 18, No. 3, pp. 425-436.

Caballero Oteyza, A., E. Battaloglu, L. Ocek, T. Lindig, J. Reichbauer, A. P. Rebelo, M. A. ..., R. Schule, 2014, "Motor protein mutations cause a new form of hereditary spastic paraplegia", *Neurology*, Vol. 82, No. 22, pp. 2007-2016.

Casari, G. and E. Rugarli, 2001, "Molecular basis of inherited spastic paraplegias", *Current Opinion in Genetics and Development*, Vol. 11, No. 3, pp. 336-342.

Chen, Y. C., G. K. Umanah, N. Dephoure, S. A. Andrabi, S. P. Gygi, T. M. Dawson, V. L. Dawson and J. Rutter, 2014, "Msp1/ATAD1 maintains mitochondrial function by facilitating the degradation of mislocalized tail-anchored proteins", *EMBO Journal*, Vol. 33, No. 14, pp. 1548-1564.

Chilamakuri, C. S., S. Lorenz, M. A. Madoui, D. Vodak, J. Sun, E. Hovig, O. Myklebost and L. A. Meza-Zepeda, 2014, "Performance comparison of four exome capture systems for deep sequencing", *BMC Genomics*, Vol. 15, No. pp. 449.

Clifford, R., T. Louis, P. Robbe, S. Ackroyd, A. Burns, A. T. Timbs, ..., A. Schuh, 2014, "SAMHD1 is mutated recurrently in chronic lymphocytic leukemia and is involved in response to DNA damage", *Blood*, Vol. 123, No. 7, pp. 1021-1031.

Connell, J. W., C. Lindon, J. P. Luzio and E. Reid, 2009, "Spastin couples microtubule severing to membrane traffic in completion of cytokinesis and secretion", *Traffic*, Vol. 10, No. 1, pp. 42-56.

Crosby, A. H. and C. Proukakis, 2002, "Is the transportation highway the right road for hereditary spastic paraplegia?", *American Journal of Human Genetics*, Vol. 71, No. 5, pp. 1009-1016.

Crow, Y. J., D. S. Chase, J. Lowenstein Schmidt, M. Szynkiewicz, G. M. Forte, H. L. Gornall, ..., G. I. Rice, 2015, "Characterization of human disease phenotypes associated with mutations in TREX1, RNASEH2A, RNASEH2B, RNASEH2C, SAMHD1, ADAR, and IFIH1", *American Journal of Medical Genetics. Part A*, Vol. 167A, No. 2, pp. 296-312.

Dale, R. C., H. Gornall, D. Singh-Grewal, M. Alcausin, G. I. Rice and Y. J. Crow, 2010, "Familial Aicardi-Goutieres syndrome due to SAMHD1 mutations is associated with chronic arthropathy and contractures", *American Journal of Medical Genetics. Part A*, Vol. 152A, No. 4, pp. 938-942.

de Brito, O. M. and L. Scorrano, 2008, "Mitofusin 2 tethers endoplasmic reticulum to mitochondria", *Nature*, Vol. 456, No. 7222, pp. 605-610.

de Souza, P. V. S., W. B. V. de Rezende Pinto, G. N. de Rezende Batistella, T. Bortholin and A. S. B. Oliveira, 2017, "Hereditary Spastic Paraplegia: Clinical and Genetic Hallmarks", *Cerebellum*, Vol. 16, No. 2, pp. 525-551.

Depienne, C., G. Stevanin, A. Brice and A. Durr, 2007, "Hereditary spastic paraplegias: an update", *Current Opinion in Neurology*, Vol. 20, No. 6, pp. 674-680.

Derivery, E. and A. Gautreau, 2010, "Assaying WAVE and WASH complex constitutive activities toward the Arp2/3 complex", *Methods in Enzymology*, Vol. 484, No. pp. 677-695.

Dey, K. K., I. Pal, R. Bharti, G. Dey, B. N. Kumar, S. Rajput, ..., M. Mandal, 2015, "Identification of RAB2A and PRDX1 as the potential biomarkers for oral squamous cell carcinoma using mass spectrometry-based comparative proteomic approach", *Tumour Biology*, Vol. 36, No. 12, pp. 9829-9837.

Dick, K. J., M. Eckhardt, C. Paisan-Ruiz, A. A. Alshehhi, C. Proukakis, N. A. Sibtain, ..., A. H. Crosby, 2010, "Mutation of FA2H underlies a complicated form of hereditary spastic paraplegia (SPG35)", *Human Mutation*, Vol. 31, No. 4, pp. E1251-1260.

DiMauro, S. and E. A. Schon, 2008, "Mitochondrial disorders in the nervous system", *Annual Review of Neuroscience*, Vol. 31, No. pp. 91-123.

Edgar, J. M., M. McLaughlin, D. Yool, S. C. Zhang, J. H. Fowler, P. Montague, ..., I. R. Griffiths, 2004, "Oligodendroglial modulation of fast axonal transport in a mouse model of hereditary spastic paraplegia", *Journal of Cell Biology*, Vol. 166, No. 1, pp. 121-131.

Esteves, T., A. Durr, E. Mundwiller, J. L. Loureiro, M. Boutry, M. A. Gonzalez, ..., F. Darios, 2014, "Loss of association of REEP2 with membranes leads to hereditary spastic paraplegia", *American Journal of Human Genetics*, Vol. 94, No. 2, pp. 268-277.

Fei, W., X. Du and H. Yang, 2011, "Seipin, adipogenesis and lipid droplets", *Trends in Endocrinology and Metabolism*, Vol. 22, No. 6, pp. 204-210.

Ferreirinha, F., A. Quattrini, M. Pirozzi, V. Valsecchi, G. Dina, V. Broccoli, ..., E. I. Rugarli, 2004, "Axonal degeneration in paraplegin-deficient mice is associated with abnormal mitochondria and impairment of axonal transport", *Journal of Clinical Investigation*, Vol. 113, No. 2, pp. 231-242.

Fink, J. K., 2003, "Advances in the hereditary spastic paraplegias", *Experimental Neurology*, Vol. 184 Suppl 1, No. pp. S106-110.

Finsterer, J., W. Loscher, S. Quasthoff, J. Wanschitz, M. Auer-Grumbach and G. Stevanin, 2012, "Hereditary spastic paraplegias with autosomal dominant, recessive, X-linked, or maternal trait of inheritance", *Journal of the Neurological Sciences*, Vol. 318, No. 1-2, pp. 1-18.

Gallus, G. N., M. T. Dotti and A. Federico, 2006, "Clinical and molecular diagnosis of cerebrotendinous xanthomatosis with a review of the mutations in the CYP27A1 gene", *Neurological Sciences*, Vol. 27, No. 2, pp. 143-149.

Giri, A., G. Guven, H. Hanagasi, A. K. Hauser, N. Erginul-Unaltuna, B. Bilgic, ..., J. Simon-Sanchez, 2016, "PLA2G6 Mutations Related to Distinct Phenotypes: A New Case with Early-onset Parkinsonism", *Tremor and Other Hyperkinetic Movements (New York, N.Y.)*, Vol. 6, No. pp. 363.

Goizet, C., A. Boukhris, E. Mundwiller, C. Tallaksen, S. Forlani, A. Toutain, ..., A. Brice, 2009, "Complicated forms of autosomal dominant hereditary spastic paraplegia are frequent in SPG10", *Human Mutation*, Vol. 30, No. 2, pp. E376-385.

Goizet, C., C. Depienne, G. Benard, A. Boukhris, E. Mundwiller, G. Sole, ..., G. Stevanin, 2011, "REEP1 mutations in SPG31: frequency, mutational spectrum, and potential association with mitochondrial morpho-functional dysfunction", *Human Mutation*, Vol. 32, No. 10, pp. 1118-1127.

Goldstone, D. C., V. Ennis-Adeniran, J. J. Hedden, H. C. Groom, G. I. Rice, E. Christodoulou, ..., M. Webb, 2011, "HIV-1 restriction factor SAMHD1 is a deoxynucleoside triphosphate triphosphohydrolase", *Nature*, Vol. 480, No. 7377, pp. 379-382.

Gonthier, B., C. Nasarre, L. Roth, M. Perraut, N. Thomasset, G. Roussel, D. Aunis and D. Bagnard, 2007, "Functional interaction between matrix metalloproteinase-3 and semaphorin-3C during cortical axonal growth and guidance", *Cerebral Cortex*, Vol. 17, No. 7, pp. 1712-1721.

Gonzalez, M., S. Nampoothiri, C. Kornblum, A. C. Oteyza, J. Walter, I. Konidari, ..., R. Schule, 2013, "Mutations in phospholipase DDHD2 cause autosomal recessive hereditary spastic paraplegia (SPG54)", *European Journal of Human Genetics*, Vol. 21, No. 11, pp. 1214-1218.

Gonzaga-Jauregui, C., T. Harel, T. Gambin, M. Kousi, L. B. Griffin, L. Francescatto, ..., J. R. Lupski, 2015, "Exome Sequence Analysis Suggests that Genetic Burden Contributes to Phenotypic Variability and Complex Neuropathy", *Cell Reports*, Vol. 12, No. 7, pp. 1169-1183.

Gonzalez, M. A., R. F. Lebrigio, D. Van Booven, R. H. Ulloa, E. Powell, F. Speziani, ..., S. Zuchner, 2013, "GEnomes Management Application (GEM.app): a new software tool for large-scale collaborative genome analysis", *Human Mutation*, Vol. 34, No. 6, pp. 842-846.

Grimm, I., R. Erdmann and W. Girzalsky, 2016, "Role of AAA(+)-proteins in peroxisome biogenesis and function", *Biochimica et Biophysica Acta*, Vol. 1863, No. 5, pp. 828-837.

Gruenenfelder, F. I., G. Thomson, J. Penderis and J. M. Edgar, 2011, "Axon-glial interaction in the CNS: what we have learned from mouse models of Pelizaeus-Merzbacher disease", *Journal of Anatomy*, Vol. 219, No. 1, pp. 33-43.

Guizetti, J., L. Schermelleh, J. Mantler, S. Maar, I. Poser, H. Leonhardt, T. Muller-Reichert and D. W. Gerlich, 2011, "Cortical constriction during abscission involves helices of ESCRT-III-dependent filaments", *Science*, Vol. 331, No. 6024, pp. 1616-1620.

Gunadi, A. Makhmudi, N. Agustriani and Rochadi, 2016, "Effects of SEMA3 polymorphisms in Hirschsprung disease patients", *Pediatric Surgery International*, Vol. 32, No. 11, pp. 1025-1028.

Hadano, S., R. Kunita, A. Otomo, K. Suzuki-Utsunomiya and J. E. Ikeda, 2007, "Molecular and cellular function of ALS2/alsin: implication of membrane dynamics in neuronal development and degeneration", *Neurochemistry International*, Vol. 51, No. 2-4, pp. 74-84.

Hapala, I., E. Marza and T. Ferreira, 2011, "Is fat so bad? Modulation of endoplasmic reticulum stress by lipid droplet formation", *Biologie Cellulaire*, Vol. 103, No. 6, pp. 271-285.

Herman, J. G. and G. G. Meadows, 2007, "Increased class 3 semaphorin expression modulates the invasive and adhesive properties of prostate cancer cells", *International Journal of Oncology*, Vol. 30, No. 5, pp. 1231-1238.

Hernandez-Montiel, H. L., E. Tamariz, M. T. Sandoval-Minero and A. Varela-Echavarria, 2008, "Semaphorins 3A, 3C, and 3F in mesencephalic dopaminergic axon pathfinding", *Journal of Comparative Neurology*, Vol. 506, No. 3, pp. 387-397.

Hirst, J., L. D. Barlow, G. C. Francisco, D. A. Sahlender, M. N. Seaman, J. B. Dacks and M. S. Robinson, 2011, "The fifth adaptor protein complex", *PLoS Biology*, Vol. 9, No. 10, pp. e1001170.

Hurley, J. H. and P. I. Hanson, 2010, "Membrane budding and scission by the ESCRT machinery: it's all in the neck", *Nature Reviews: Molecular Cell Biology*, Vol. 11, No. 8, pp. 556-566.

Jo, Y., P. V. Sguigna and R. A. DeBose-Boyd, 2011, "Membrane-associated ubiquitin ligase complex containing gp78 mediates sterol-accelerated degradation of 3-hydroxy-3-methylglutaryl-coenzyme A reductase", *Journal of Biological Chemistry*, Vol. 286, No. 17, pp. 15022-15031.

Kallberg, M., H. Wang, S. Wang, J. Peng, Z. Wang, H. Lu and J. Xu, 2012, "Template-based protein structure modeling using the RaptorX web server", *Nature Protocols*, Vol. 7, No. 8, pp. 1511-1522.

Kapoor, A., Q. Jiang, S. Chatterjee, P. Chakraborty, M. X. Sosa, C. Berrios and A. Chakravarti, 2015, "Population variation in total genetic risk of Hirschsprung disease from common RET, SEMA3 and NRG1 susceptibility polymorphisms", *Human Molecular Genetics*, Vol. 24, No. 10, pp. 2997-3003.

Kelley, L. A., S. Mezulis, C. M. Yates, M. N. Wass and M. J. Sternberg, 2015, "The Phyre2 web portal for protein modeling, prediction and analysis", *Nature Protocols*, Vol. 10, No. 6, pp. 845-858.

Kenwrick, S., V. Ionasescu, G. Ionasescu, C. Searby, A. King , M. Dubowitz and K. E. Davies, 1986, "Linkage studies of X-linked recessive spastic paraplegia using DNA probes", *Human genetics,* Vol. 73, pp. 264-6.

Klebe, S., G. Stevanin and C. Depienne, 2015, "Clinical and genetic heterogeneity in hereditary spastic paraplegias: from SPG1 to SPG72 and still counting", *Revue Neurologique*, Vol. 171, No. 6-7, pp. 505-530.

Krols, M., G. van Isterdael, B. Asselbergh, A. Kremer, S. Lippens, V. Timmerman and S. Janssens, 2016, "Mitochondria-associated membranes as hubs for neurodegeneration", *Acta Neuropathologica*, Vol. 131, No. 4, pp. 505-523.

Kruer, M. C., C. Paisan-Ruiz, N. Boddaert, M. Y. Yoon, H. Hama, A. Gregory, ..., S. J. Hayflick, 2010, "Defective FA2H leads to a novel form of neurodegeneration with brain iron accumulation (NBIA)", *Annals of Neurology*, Vol. 68, No. 5, pp. 611-618.

Kumar, P., S. Henikoff and P. C. Ng, 2009, "Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm", *Nature Protocols*, Vol. 4, No. 7, pp. 1073-1081.

Laguette, N., N. Rahm, B. Sobhian, C. Chable-Bessia, J. Munch, J. Snoeck, ..., M. Benkirane, 2012, "Evolutionary and functional analyses of the interaction between the myeloid restriction factor SAMHD1 and the lentiviral Vpx protein", *Cell Host & Microbe*, Vol. 11, No. 2, pp. 205-217.

Leshinsky-Silver, E., G. Malinger, L. Ben-Sira, D. Kidron, S. Cohen, S. Inbar, ..., T. Lerman-Sagie, 2011, "A large homozygous deletion in the SAMHD1 gene causes atypical Aicardi-Goutieres syndrome associated with mtDNA deletions", *European Journal of Human Genetics*, Vol. 19, No. 3, pp. 287-292.

Lo Giudice, T., F. Lombardi, F. M. Santorelli, T. Kawarai and A. Orlacchio, 2014, "Hereditary spastic paraplegia: clinical-genetic characteristics and evolving molecular mechanisms", *Experimental Neurology*, Vol. 261, No. pp. 518-539.

Lynes, E. M. and T. Simmen, 2011, "Urban planning of the endoplasmic reticulum (ER): how diverse mechanisms segregate the many functions of the ER", *Biochimica et Biophysica Acta*, Vol. 1813, No. 10, pp. 1893-1905.

Malik, M. F., L. K. Satherley, E. L. Davies, L. Ye and W. G. Jiang, 2016, "Expression of Semaphorin 3C in Breast Cancer and its Impact on Adhesion and Invasion of Breast Cancer Cells", *Anticancer Research*, Vol. 36, No. 3, pp. 1281-1286.

Manzini, M. C., A. Rajab, T. M. Maynard, G. H. Mochida, W. H. Tan, R. Nasir, ..., C. A. Walsh, 2010, "Developmental and degenerative features in a complicated spastic paraplegia", *Annals of Neurology*, Vol. 67, No. 4, pp. 516-525.

Meijer, I. A., P. Cossette, J. Roussel, M. Benard, S. Toupin and G. A. Rouleau, 2004, "A novel locus for pure recessive hereditary spastic paraplegia maps to 10q22.1-10q24.1", *Annals of Neurology*, Vol. 56, No. 4, pp. 579-582.

Montenegro, G., A. P. Rebelo, J. Connell, R. Allison, C. Babalini, M. D'Aloia, ..., S. Zuchner, 2012, "Mutations in the ER-shaping protein reticulon 2 cause the axon-degenerative disorder hereditary spastic paraplegia type 12", *Journal of Clinical Investigation*, Vol. 122, No. 2, pp. 538-544.

Moreno-De-Luca, A., S. L. Helmers, H. Mao, T. G. Burns, A. M. Melton, K. R. Schmidt, ..., C. L. Martin, 2011, "Adaptor protein complex-4 (AP-4) deficiency causes a novel autosomal recessive cerebral palsy syndrome with microcephaly and intellectual disability", *Journal of Medical Genetics*, Vol. 48, No. 2, pp. 141-144.

Murmu, R. P., E. Martin, A. Rastetter, T. Esteves, M. P. Muriel, K. H. El Hachimi, ..., G. Stevanin, 2011, "Cellular distribution and subcellular localization of spatacsin and spastizin, two proteins involved in hereditary spastic paraplegia", *Molecular and Cellular Neurosciences*, Vol. 47, No. 3, pp. 191-202.

Nicholls, Z., E. Hobson, J. Martindale and P. J. Shaw, 2015, "Diagnosis of spinal xanthomatosis by next-generation sequencing: identifying a rare, treatable mimic of hereditary spastic paraparesis", *Practical Neurology*, Vol. 15, No. 4, pp. 280-283.

Okreglak, V. and P. Walter, 2014, "The conserved AAA-ATPase Msp1 confers organelle specificity to tail-anchored proteins", *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 111, No. 22, pp. 8019-8024.

Orcesi, S., R. La Piana and E. Fazzi, 2009, "Aicardi-Goutieres syndrome", *British Medical Bulletin*, Vol. 89, No. pp. 183-201.

Orthmann-Murphy, J. L., E. Salsano, C. K. Abrams, A. Bizzi, G. Uziel, M. M. Freidin, ..., D. Pareyson, 2009, "Hereditary spastic paraplegia is a novel phenotype for GJA12/GJC2 mutations", *Brain*, Vol. 132, No. Pt 2, pp. 426-438.

Ozes B, and E. Battaloglu, 2011, "Herediter spasitk paraparezi: genetikten patagoneze" *Turkiye klinikleri journal neurology- special topics,* Vol. 4, No. 2, pp. 93 -102.

Ozes, B., N. Karagoz, R. Schule, A. Rebelo, M. J. Sobrido, F. Harmuth, ..., E. Battaloglu, 2017, "PLA2G6 mutations associated with a continuous clinical spectrum from neuroaxonal dystrophy to hereditary spastic paraplegia", *Clinical Genetics*, Vol. 92, No. 5, pp. 534-539.

Paisan-Ruiz, C., K. P. Bhatia, A. Li, D. Hernandez, M. Davis, N. W. Wood, ..., S. A. Schneider, 2009, "Characterization of PLA2G6 as a locus for dystonia-parkinsonism", *Annals of Neurology*, Vol. 65, No. 1, pp. 19-23.

Pang, S. Y., K. C. Teo, J. S. Hsu, R. S. Chang, M. Li, P. C. Sham and S. L. Ho, 2017, "The role of gene variants in the pathogenesis of neurodegenerative disorders as revealed by next generation sequencing studies: a review", *Transl Neurodegener*, Vol. 6, No. pp. 27.

Parfitt, D. A., G. J. Michael, E. G. Vermeulen, N. V. Prodromou, T. R. Webb, J. M. Gallo, ..., J. P. Chapple, 2009, "The ataxia protein sacsin is a functional co-chaperone that protects against polyglutamine-expanded ataxin-1", *Human Molecular Genetics*, Vol. 18, No. 9, pp. 1556-1565.

Park, S. H., P. P. Zhu, R. L. Parker and C. Blackstone, 2010, "Hereditary spastic paraplegia proteins REEP1, spastin, and atlastin-1 coordinate microtubule interactions with the tubular ER network", *Journal of Clinical Investigation*, Vol. 120, No. 4, pp. 1097-1110.

Pearce, M. M., D. B. Wormer, S. Wilkens and R. J. Wojcikiewicz, 2009, "An endoplasmic reticulum (ER) membrane complex composed of SPFH1 and SPFH2 mediates the ER-associated degradation of inositol 1,4,5-trisphosphate receptors", *Journal of Biological Chemistry*, Vol. 284, No. 16, pp. 10433-10445.

Posey, J. E., J. A. Rosenfeld, R. A. James, M. Bainbridge, Z. Niu, X. Wang, ..., S. E. Plon, 2016, "Molecular diagnostic experience of whole-exome sequencing in adult patients", *Genetics in Medicine*, Vol. 18, No. 7, pp. 678-685.

Potter, K. A., M. J. Kern, G. Fullbright, J. Bielawski, S. S. Scherer, S. W. Yum, ..., H. Hama, 2011, "Central nervous system dysfunction in a mouse model of FA2H deficiency", *Glia*, Vol. 59, No. 7, pp. 1009-1021.

Rainier, S., M. Bui, E. Mark, D. Thomas, D. Tokarz, L. Ming, ..., J. K. Fink, 2008, "Neuropathy target esterase gene mutations cause motor neuron disease", *American Journal of Human Genetics*, Vol. 82, No. 3, pp. 780-785.

Ravenscroft, J. C., M. Suri, G. I. Rice, M. Szynkiewicz and Y. J. Crow, 2011, "Autosomal dominant inheritance of a heterozygous mutation in SAMHD1 causing familial chilblain lupus", *American Journal of Medical Genetics. Part A*, Vol. 155A, No. 1, pp. 235-237.

Read, D. J., Y. Li, M. V. Chao, J. B. Cavanagh and P. Glynn, 2009, "Neuropathy target esterase is required for adult vertebrate axon maintenance", *Journal of Neuroscience*, Vol. 29, No. 37, pp. 11594-11600.

Reid, E., 1999, "The hereditary spastic paraplegias", *Journal of Neurology*, Vol. 246, No. 11, pp. 995-1003.

Reid, E., M. Kloos, A. Ashley-Koch, L. Hughes, S. Bevan, I. K. Svenson, ..., D. A. Marchuk, 2002, "A kinesin heavy chain (KIF5A) mutation in hereditary spastic paraplegia (SPG10)", *American Journal of Human Genetics*, Vol. 71, No. 5, pp. 1189-1194.

Reidy, K. and A. Tufro, 2011, "Semaphorins in kidney development and disease: modulators of ureteric bud branching, vascular morphogenesis, and podocyte-endothelial crosstalk", *Pediatric Nephrology*, Vol. 26, No. 9, pp. 1407-1412.

Renvoise, B., R. L. Parker, D. Yang, J. C. Bakowska, J. H. Hurley and C. Blackstone, 2010, "SPG20 protein spartin is recruited to midbodies by ESCRT-III protein Ist1 and participates in cytokinesis", *Molecular Biology of the Cell*, Vol. 21, No. 19, pp. 3293-3303.

Rice, G. I., J. Bond, A. Asipu, R. L. Brunette, I. W. Manfield, I. M. Carr, ..., Y. J. Crow, 2009, "Mutations involved in Aicardi-Goutieres syndrome implicate SAMHD1 as regulator of the innate immune response", *Nature Genetics*, Vol. 41, No. 7, pp. 829-832.

Richards, S., N. Aziz, S. Bale, D. Bick, S. Das, J. Gastier-Foster, ..., A. L. Q. A. Committee, 2015, "Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology", *Genetics in Medicine*, Vol. 17, No. 5, pp. 405-424.

Ropers, F., E. Derivery, H. Hu, M. Garshasbi, M. Karbasiyan, M. Herold, ..., A. Rajab, 2011, "Identification of a novel candidate gene for non-syndromic autosomal recessive intellectual disability: the WASH complex member SWIP", *Human Molecular Genetics*, Vol. 20, No. 13, pp. 2585-2590.

Rosafio, F., F. Cavallieri, P. Guaraldi, F. Taroni, P. F. Nichelli and J. Mandrioli, 2016, "The wide spectrum of cerebrotendinous xanthomatosis: Case report of a rare but treatable disease", *Clinical Neurology and Neurosurgery*, Vol. 143, No. pp. 1-3.

Rose, K. A., G. Stapleton, K. Dott, M. P. Kieny, R. Best, M. Schwarz, ..., R. Lathe, 1997, "Cyp7b, a novel brain cytochrome P450, catalyzes the synthesis of neurosteroids 7alpha-hydroxy dehydroepiandrosterone and 7alpha-hydroxy pregnenolone", *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 94, No. 10, pp. 4925-4930.

Rossi, D., 2014, "SAMHD1: a new gene for CLL", *Blood*, Vol. 123, No. 7, pp. 951-952.

Ruano, L., C. Melo, M. C. Silva and P. Coutinho, 2014, "The global epidemiology of hereditary ataxia and spastic paraplegia: a systematic review of prevalence studies", *Neuroepidemiology*, Vol. 42, No. 3, pp. 174-183.

Sagona, A. P., I. P. Nezis, N. M. Pedersen, K. Liestol, J. Poulton, T. E. Rusten, ..., H. Stenmark, 2010, "PtdIns(3)P controls cytokinesis through KIF13A-mediated recruitment of FYVE-CENT to the midbody", *Nature Cell Biology*, Vol. 12, No. 4, pp. 362-371.

Salinas, S., R. E. Carazo-Salas, C. Proukakis, G. Schiavo and T. T. Warner, 2007, "Spastin and microtubules: Functions in health and disease", *Journal of Neuroscience Research*, Vol. 85, No. 12, pp. 2778-2782.

Sasamura, A., S. Akazawa, A. Haraguchi, I. Horie, T. Ando, N. Abiru, ..., A. Kawakami, 2018, "A Case of Late-onset Cerebrotendinous Xanthomatosis with a Novel Mutation in the CYP27A1 Gene", *Internal Medicine*, Vol. No. pp.

Schule, R., E. Brandt, K. N. Karle, M. Tsaousidou, S. Klebe, S. Klimpe, ..., C. Beetz, 2009, "Analysis of CYP7B1 in non-consanguineous cases of hereditary spastic paraplegia", *Neurogenetics*, Vol. 10, No. 2, pp. 97-104.

Schule, R. and L. Schols, 2011, "Genetics of hereditary spastic paraplegias", *Seminars in Neurology*, Vol. 31, No. 5, pp. 484-493.

Schule, R., S. Wiethoff, P. Martus, K. N. Karle, S. Otto, S. Klebe, ..., L. Schols, 2016, "Hereditary spastic paraplegia: Clinicogenetic lessons from 608 patients", *Annals of Neurology*, Vol. 79, No. 4, pp. 646-658.

Schwarz, J. M., C. Rodelsperger, M. Schuelke and D. Seelow, 2010, "MutationTaster evaluates disease-causing potential of sequence alterations", *Nat Methods*, Vol. 7, No. 8, pp. 575-576.

Sekijima, Y., S. Koyama, T. Yoshinaga, M. Koinuma and Y. Inaba, 2018, "Nationwide survey on cerebrotendinous xanthomatosis in Japan", *Journal of Human Genetics*, Vol. 63, No. 3, pp. 271-280.

Slabicki, M., M. Theis, D. B. Krastev, S. Samsonov, E. Mundwiller, M. Junqueira, ..., F. Buchholz, 2010, "A genome-scale DNA repair RNAi screen identifies SPG48 as a novel gene associated with hereditary spastic paraplegia", *PLoS Biology*, Vol. 8, No. 6, pp. e1000408.

Stephenson, J. B., 2008, "Aicardi-Goutieres syndrome (AGS)", *European Journal of Paediatric Neurology*, Vol. 12, No. 5, pp. 355-358.

Stevanin, G., H. Azzedine, P. Denora, A. Boukhris, M. Tazir, A. Lossos, ..., S. consortium, 2008, "Mutations in SPG11 are frequent in autosomal recessive spastic paraplegia with thin corpus callosum, cognitive decline and lower motor neuron degeneration", *Brain*, Vol. 131, No. Pt 3, pp. 772-784.

Tesson, C., M. Nawara, M. A. Salih, R. Rossignol, M. S. Zaki, M. Al Balwi, ..., G. Stevanin, 2012, "Alteration of fatty-acid-metabolizing enzymes affects mitochondrial form and function in hereditary spastic paraplegia", *American Journal of Human Genetics*, Vol. 91, No. 6, pp. 1051-1064.

Tsaousidou, M. K., K. Ouahchi, T. T. Warner, Y. Yang, M. A. Simpson, N. G. Laing, ..., A. H. Crosby, 2008, "Sequence alterations within CYP7B1 implicate defective

cholesterol homeostasis in motor-neuron degeneration", *American Journal of Human Genetics*, Vol. 82, No. 2, pp. 510-515.

Varga, R. E., R. Schule, H. Fadel, I. Valenzuela, F. Speziani, M. Gonzalez, ..., C. Beetz, 2013, "Do not trust the pedigree: reduced and sex-dependent penetrance at a novel mutation hotspot in ATL1 blurs autosomal dominant inheritance of spastic paraplegia", *Human Mutation*, Vol. 34, No. 6, pp. 860-863.

Verny, C., N. Guegen, V. Desquiret, A. Chevrollier, A. Prundean, F. Dubas, ..., V. Procaccio, 2011, "Hereditary spastic paraplegia-like disorder due to a mitochondrial ATP6 gene point mutation", *Mitochondrion*, Vol. 11, No. 1, pp. 70-75.

Warnecke, T., T. Duning, A. Schwan, H. Lohmann, J. T. Epplen and P. Young, 2007, "A novel form of autosomal recessive hereditary spastic paraplegia caused by a new SPG7 mutation", *Neurology*, Vol. 69, No. 4, pp. 368-375.

Windpassinger, C., M. Auer-Grumbach, J. Irobi, H. Patel, E. Petek, G. Horl, ..., K. Wagner, 2004, "Heterozygous missense mutations in BSCL2 are associated with distal hereditary motor neuropathy and Silver syndrome", *Nature Genetics*, Vol. 36, No. 3, pp. 271-276.

Wu, Z., K. O. Martin, N. B. Javitt and J. Y. Chiang, 1999, "Structure and functions of human oxysterol 7alpha-hydroxylase cDNAs and gene CYP7B1", *Journal of Lipid Research*, Vol. 40, No. 12, pp. 2195-2203.

Yagi, T., D. Ito, Y. Nihei, T. Ishihara and N. Suzuki, 2011, "N88S seipin mutant transgenic mice develop features of seipinopathy/BSCL2-related motor neuron disease via endoplasmic reticulum stress", *Human Molecular Genetics*, Vol. 20, No. 19, pp. 3831-3840.

Yang, D., N. Rismanchi, B. Renvoise, J. Lippincott-Schwartz, C. Blackstone and J. H. Hurley, 2008, "Structural basis for midbody targeting of spastin by the ESCRT-III

protein CHMP1B", *Nature Structural & Molecular Biology*, Vol. 15, No. 12, pp. 1278-1286.

Yang, Y., D. M. Muzny, J. G. Reid, M. N. Bainbridge, A. Willis, P. A. Ward, ..., C. M. Eng, 2013, "Clinical whole-exome sequencing for the diagnosis of mendelian disorders", *New England Journal of Medicine*, Vol. 369, No. 16, pp. 1502-1511.

Yeo, G. and C. B. Burge, 2004, "Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals", *Journal of Computational Biology*, Vol. 11, No. 2-3, pp. 377-394.

Zhang, J., Y. Wang, Z. Chi, M. J. Keuss, Y. M. Pai, H. C. Kang, ..., V. L. Dawson, 2011, "The AAA+ ATPase Thorase regulates AMPA receptor-dependent synaptic plasticity and behavior", *Cell*, Vol. 145, No. 2, pp. 284-299.

Zhang, Y., 2008, "I-TASSER server for protein 3D structure prediction", *BMC Bioinformatics*, Vol. 9, No. pp. 40.

APPENDIX A: CHROMATOGRAMS

All chromatograms are given below.

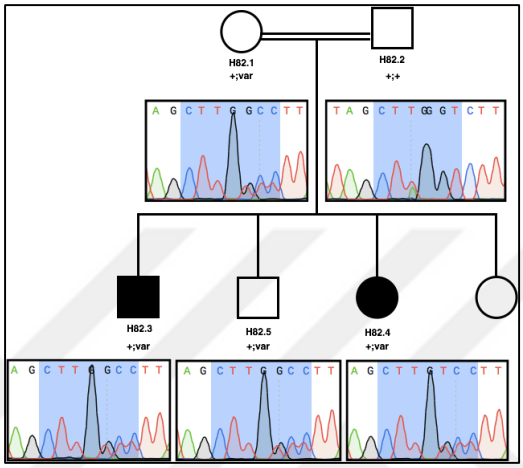


Figure A.1. H82 pedigree and chromatograms showing c.1222-38_1222-37del variant intron 10 of *SPAST* gene.

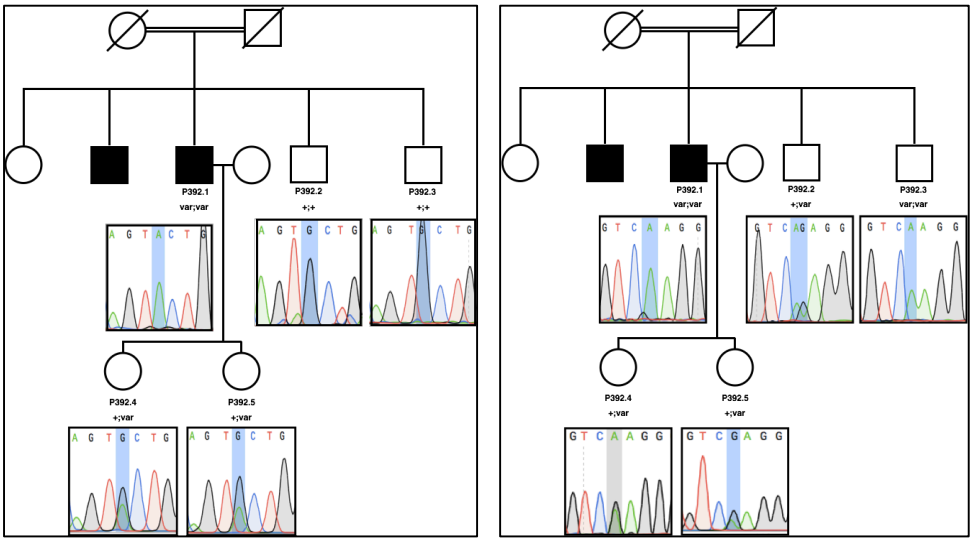


Figure A.2. P392 pedigree and chromatograms showing the c.157G>A variant in KCNMB3 gene (left panel) and the c.47G>A variant in SLC7A2 gene (right panel) (+: native allele; var: variant allele).

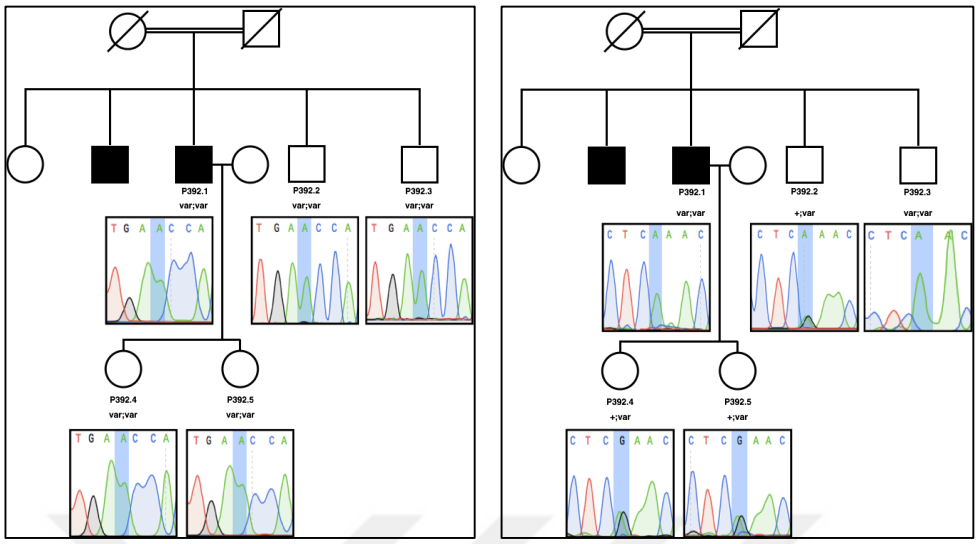


Figure A.3. P392 pedigree and chromatograms showing the c.980T>A variant in PDGFRL gene (left panel) and the c.590G>A variant in NAT2 gene (right panel) (+: native allele; var: variant allele).

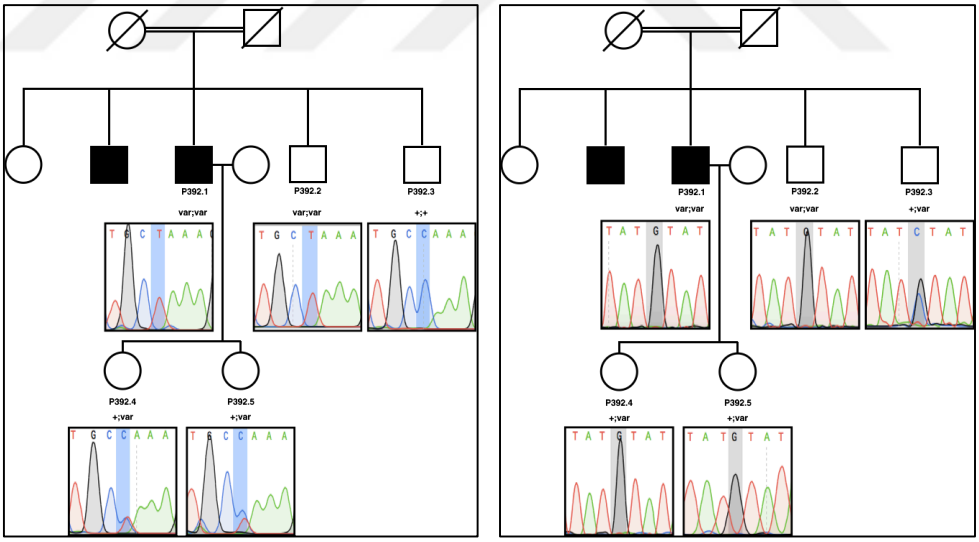


Figure A.4. P392 pedigree and chromatograms showing the c.1070C>T variant in SCYL2 gene (left panel) and the c.1505C>G variant in UTP20 gene (right panel) (+: native allele; var: variant allele).

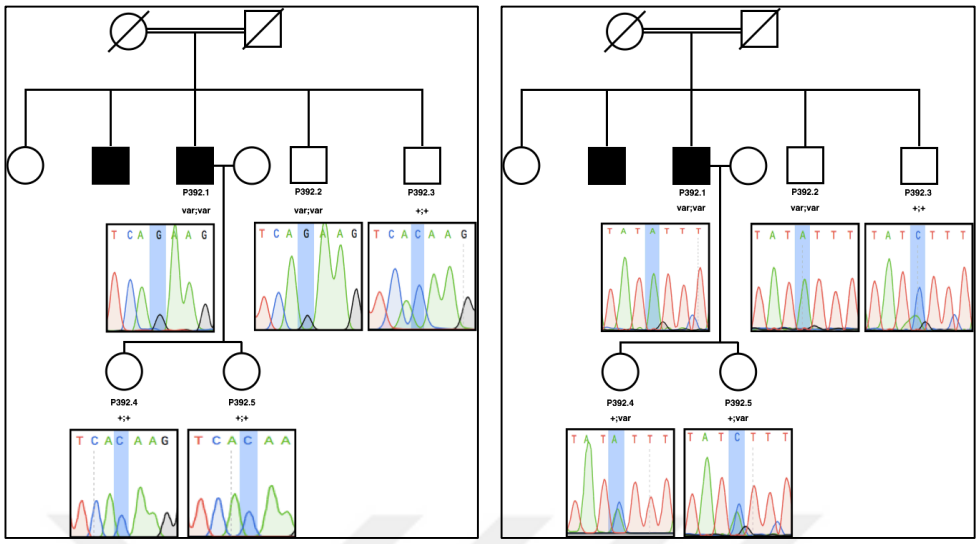


Figure A.5. P392 pedigree and chromatograms showing the c.1443C>G variant in MYBPC1 gene (left panel) and c.1906-6C>A variant in CAPRIN2 gene (rigth panel) (+: native allele; var: variant allele).

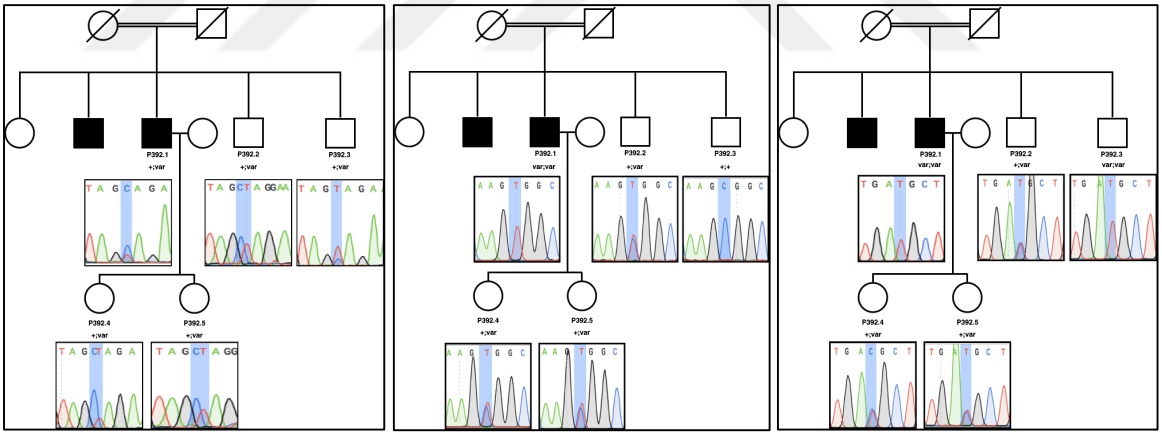


Figure A.6. P392 pedigree and chromatograms showing the c.2783T>C variant in OVOS2 gene (left panel) c.136C>T variant in CCDC114 gene (middle panel) and the c.734C>T variant in ETFB gene (right panel) (+: native allele; var: variant allele).

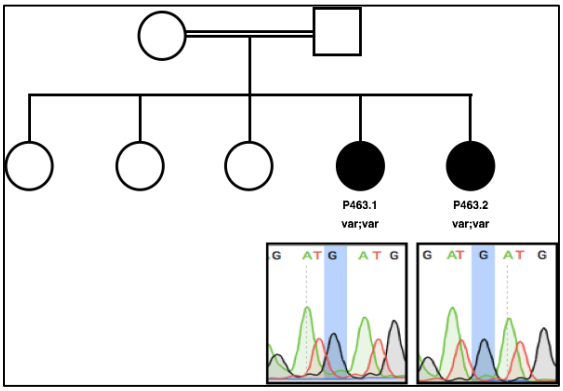


Figure A.7. P463 pedigree and chromatograms showing c.1235C>G, p.S412X variant in SPG11 gene in affected individuals (var: variant allele).

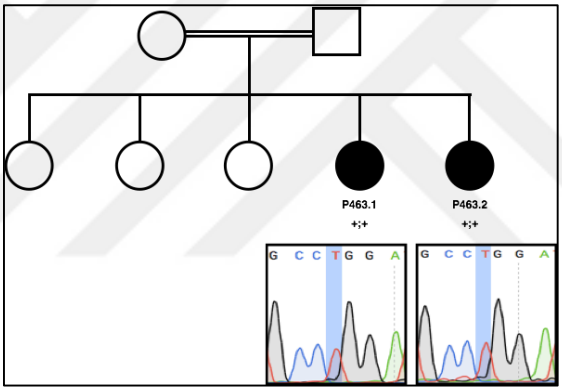


Figure A.8. P463 pedigree and chromatograms showing c.3794T>A, p.L1265Q variant in ZFYVE26 gene in affected individuals (+: native allele).

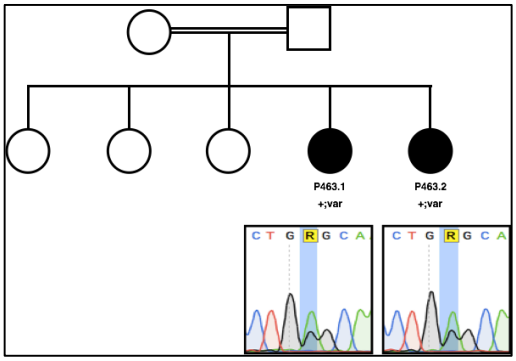


Figure A.9. P463 pedigree and chromatograms showing c.985A>G, p.S329G variant in ZFYVE26 gene in affected individuals (+: native allele, var: variant allele).

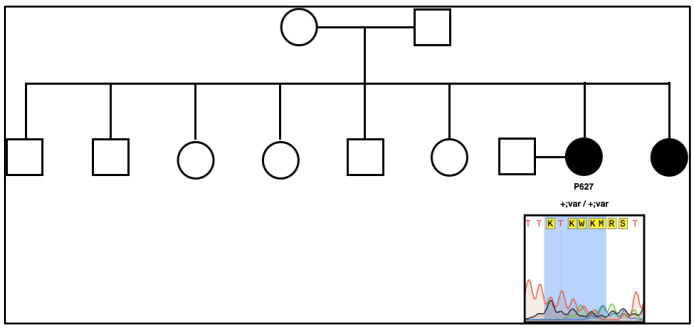


Figure A.10. P627 pedigree and chromatogram showing c.1087-3delT and c.1087-3delTT variants in CES1 gene, in patient P627.

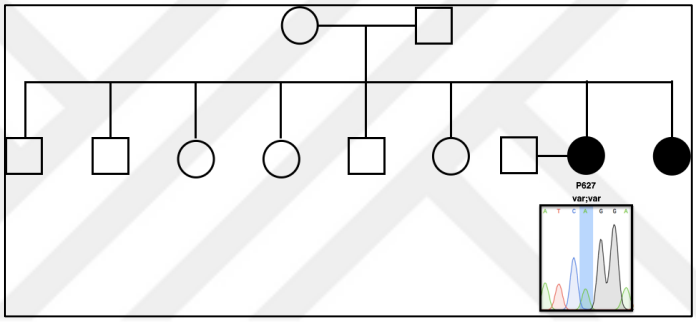


Figure A.11. P627 pedigree and chromatogram showing c.637G>A variant in SACS gene, in patient P627.

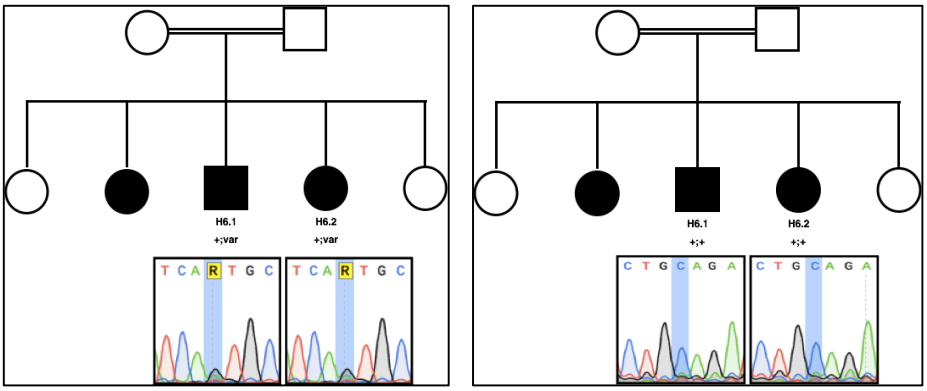


Figure A.12. H6 pedigree and the chromatograms showing heterozygous n.387A>G variant in SDHAP2 gene (left panel) and absence of c.1317-3C>T variant in ANKRD20A4 gene (right panel) (+: native allele; var: variant allele).

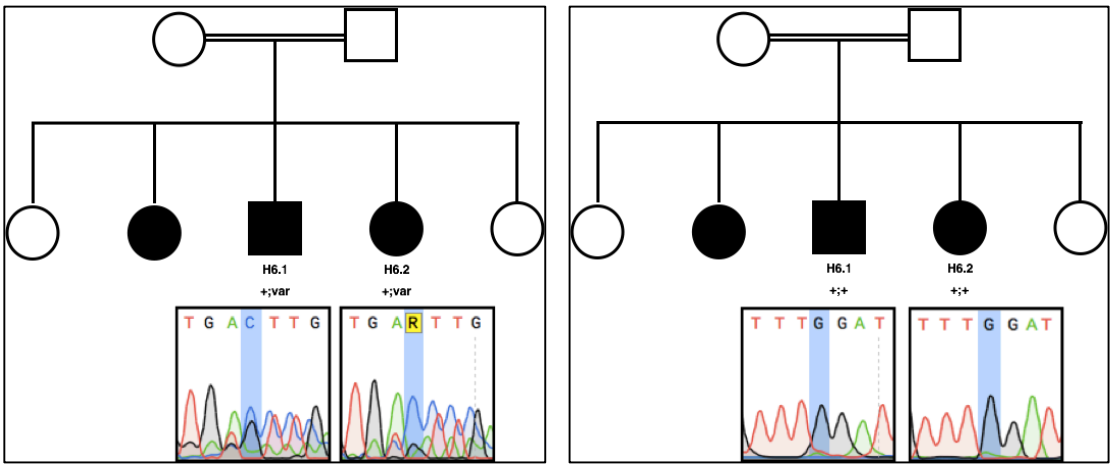


Figure A.13. H6 pedigree and chromatograms showing the c.719T>G variant in OR11H12 gene (left panel) and c.4708G>A variant in RGPD8 gene (right panel) in patients H6.1 and H6.2 (+: native allele; var: variant allele).

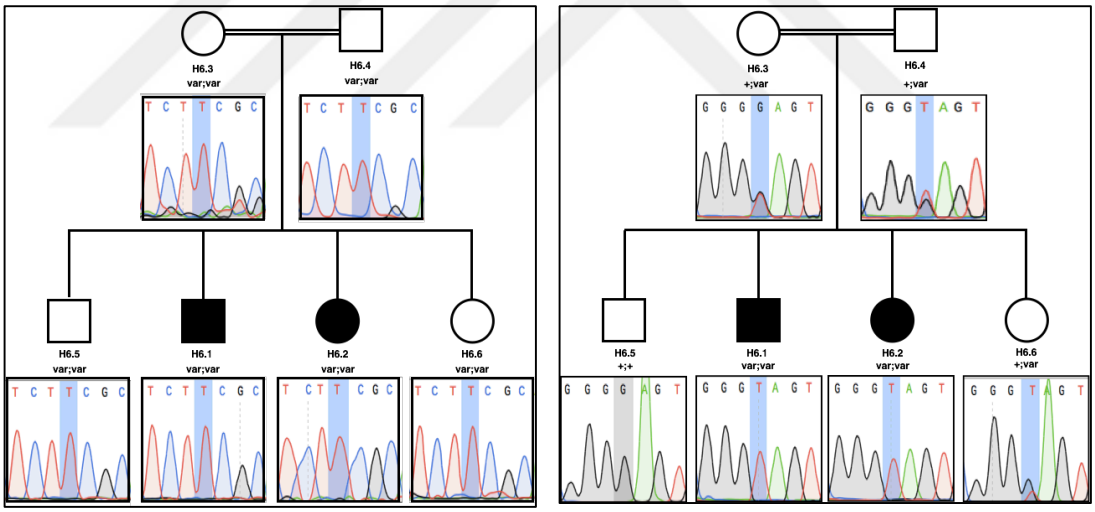


Figure A.14. H6 pedigree and chromatograms showing the c.20C>T variant in GUCA2A gene (left panel) and the c.1852G>T variant in SLC5A9 gene (right panel) (+: native allele; var: variant allele).

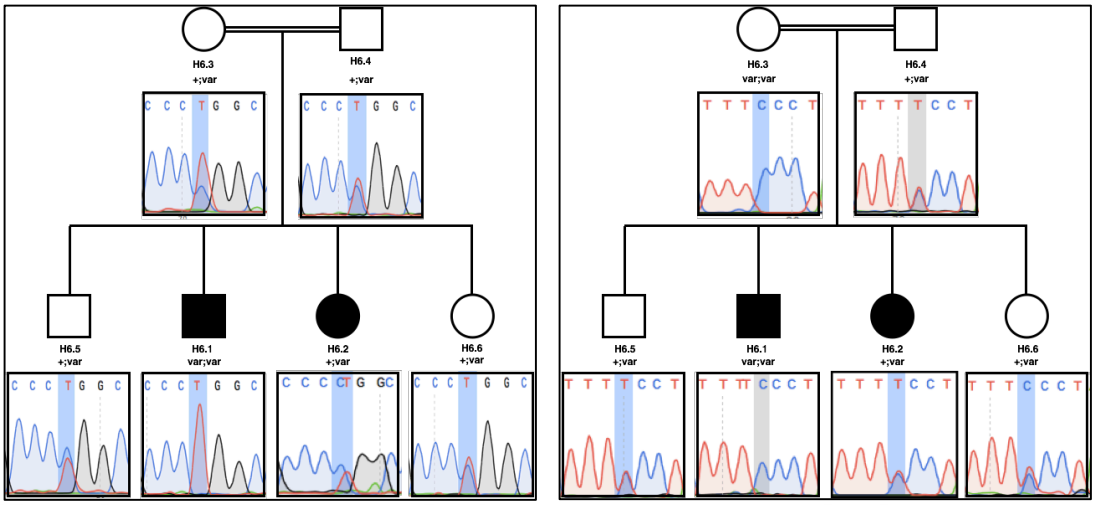


Figure A.15. H6 pedigree and chromatograms showing the c.1018C>T variant in PRKAG3 gene (left panel) and the c.443-6T>C variant in SLC23A3 gene (right panel) (+: native allele; var: variant allele).

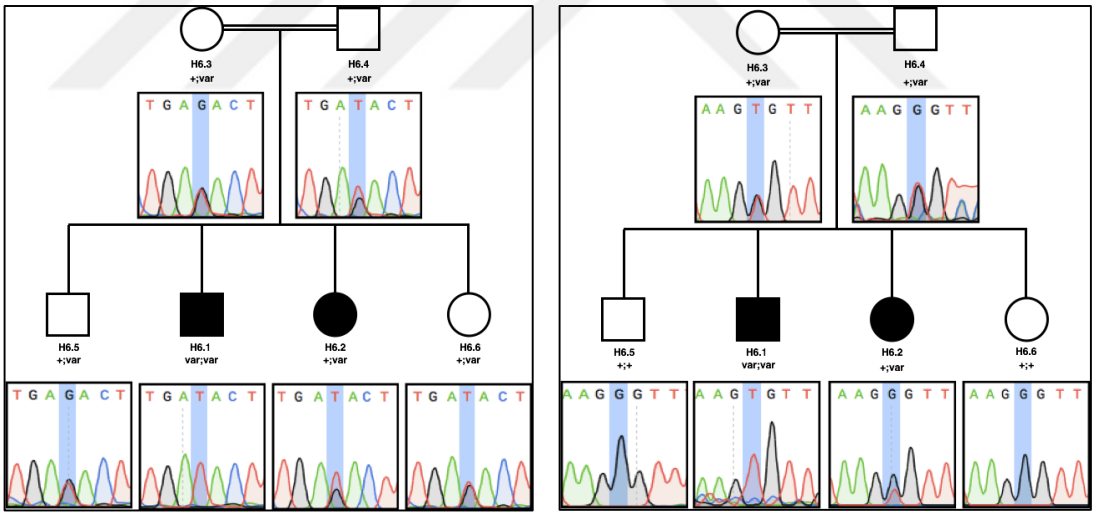


Figure A.16. H6 pedigree and chromatograms showing the c.90G>T variant in RESP18 gene (left panel) and the c.5+6G>T variant in IFIT5 gene (right panel) (+: native allele; var: variant allele).

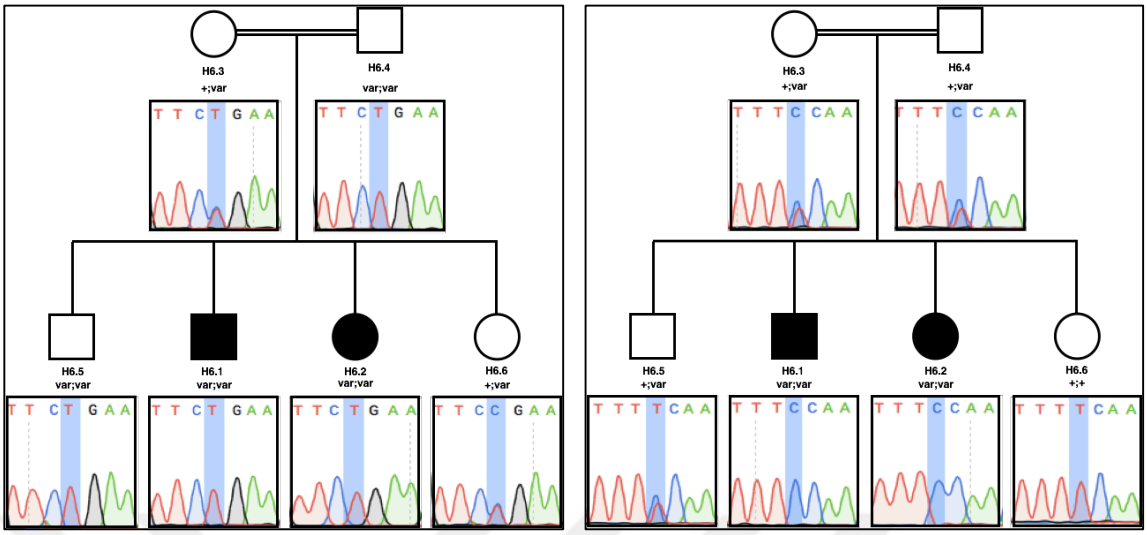


Figure A.17. H6 pedigree and chromatograms showing the c.3576-5C>T variant in MYOM1 gene (left panel) and the c.161T>C variant in SPINT3 gene (right panel) (+: native allele; var: variant allele).

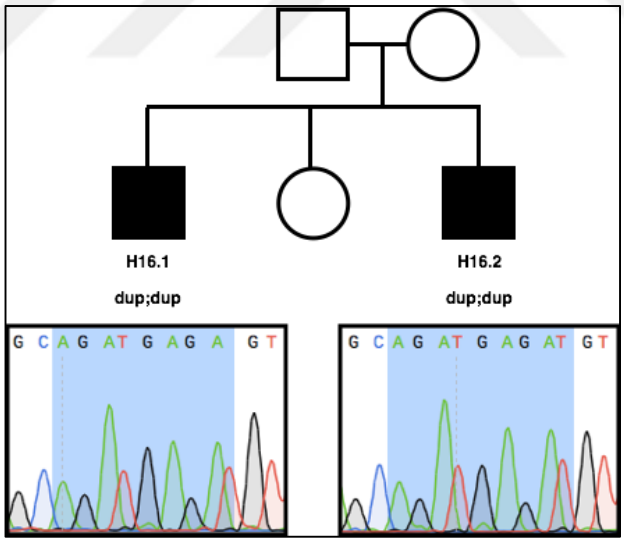


Figure A.18. H16 pedigree and chromatograms showing c.6215_6219dupAGAT, p.Phe2074ArgfsTer15 insertion in SPG11 gene (dup:duplicated allele).

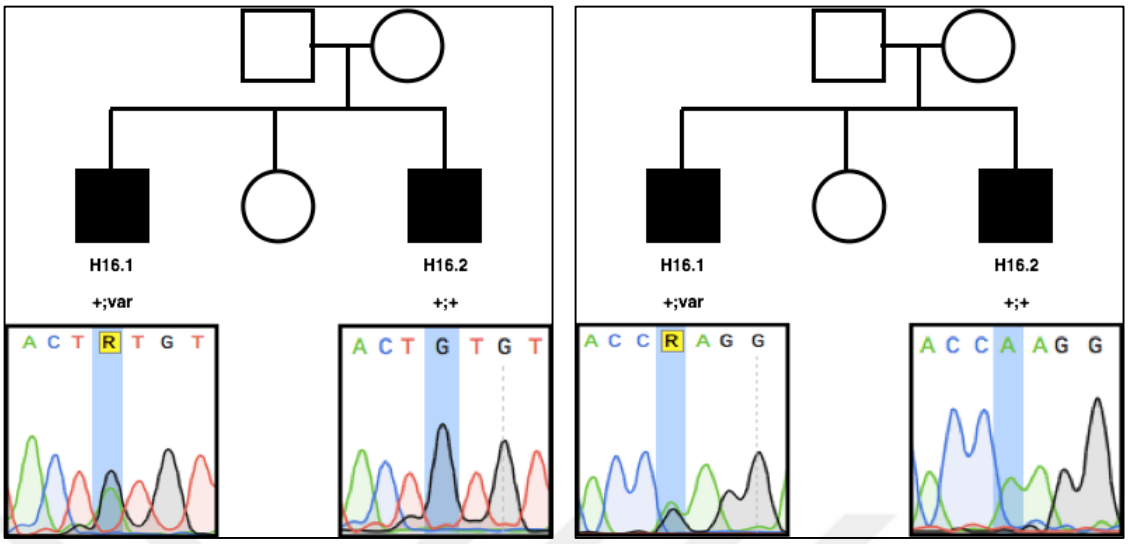


Figure A.19. H16 pedigree and chromatograms showing the sequences including c.*3745G>A (left panel) and c.424A>G (right panel) in C19orf12 gene (+: native allele; var: variant allele).

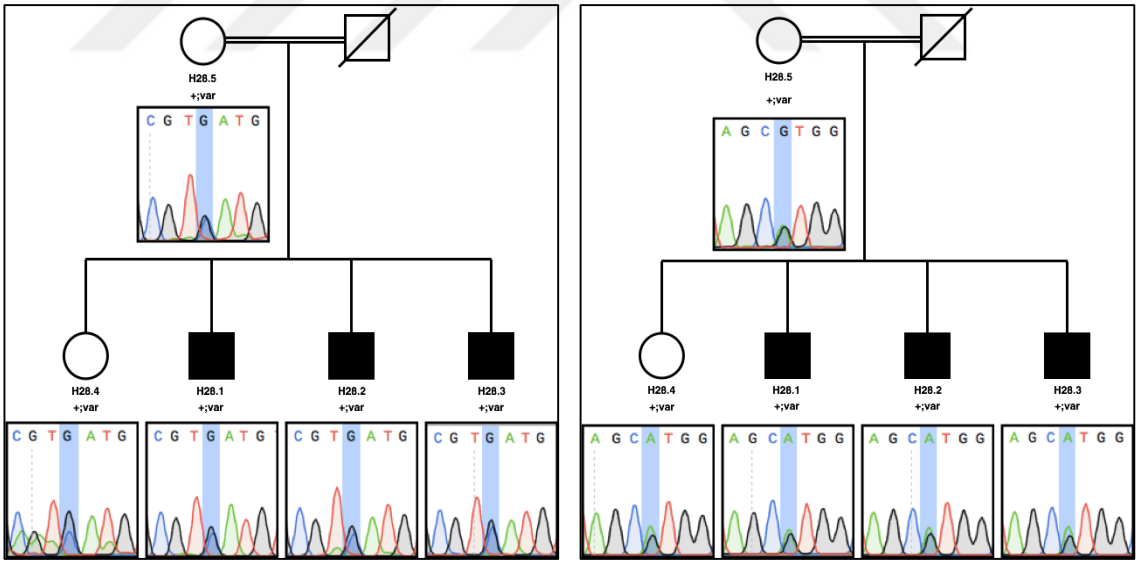


Figure A.20. H28 pedigree and chromatogram showing c.-1G>C (left panel) and c.2287G>A (right panel) variants in AP5Z1 gene in patient H28.1 and H28.2 (+: native allele; var: variant allele).

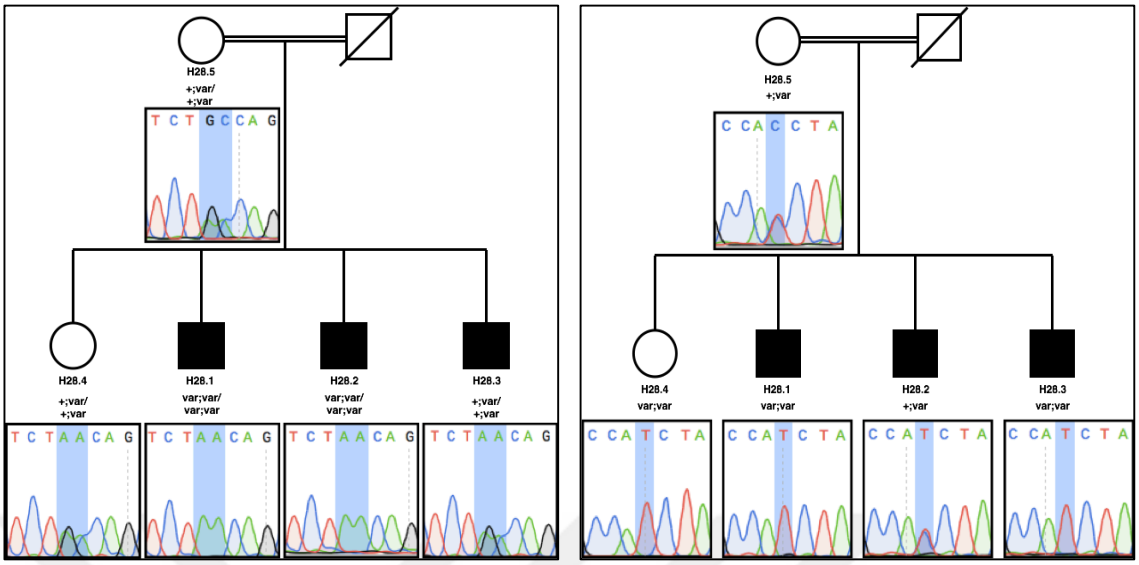


Figure A.21. H28 pedigree and chromatograms showing the c.411-4C>A and c.411-5G>A variants in CERS2 gene (left panel) and the c.9206C>T variant in 41st exon of COL6A3 gene (right panel) (+: native allele; var: variant allele).

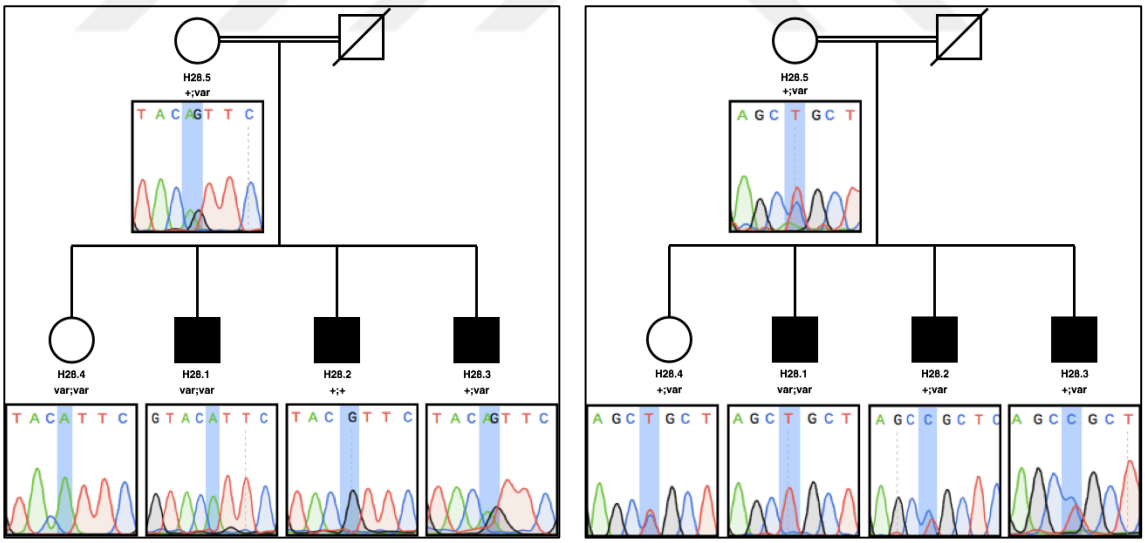


Figure A.22. H28 pedigree and chromatograms showing the c.3751G>A variant in exon nine of COL6A3 gene (left panel) and the c.7235C>T variant in 21st exon of TNXB gene (right panel) (+: native allele; var: variant allele).

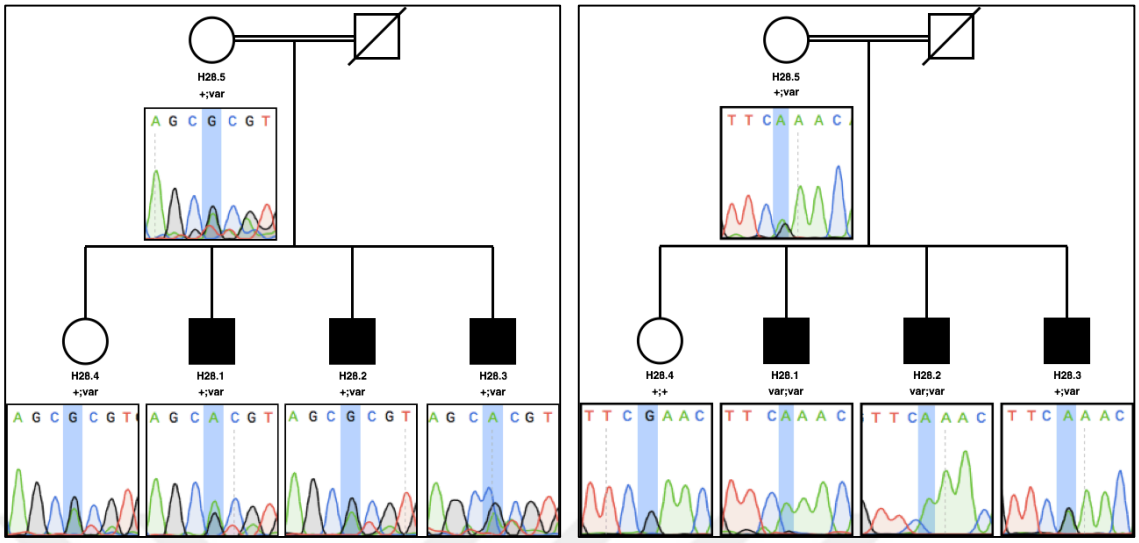


Figure A.23. H28 pedigree and chromatograms showing the c.8111G>A variant in 23rd exon of TNXB gene (left panel) and c.3239G>A variant in LAMA4 gene (rigth panel) (+: native allele; var: variant allele).

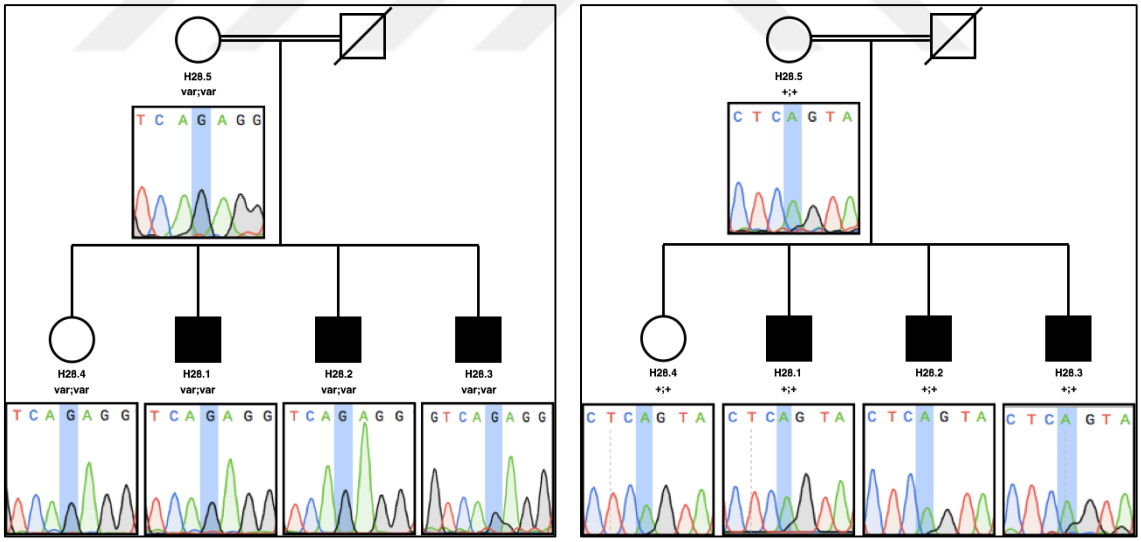


Figure A.24. H28 pedigree and chromatograms showing the c.431T>G variant in GPRC6A gene (left panel) and the c.730A>G variant in FAM35A gene (right panel) (+: native allele; var: variant allele).

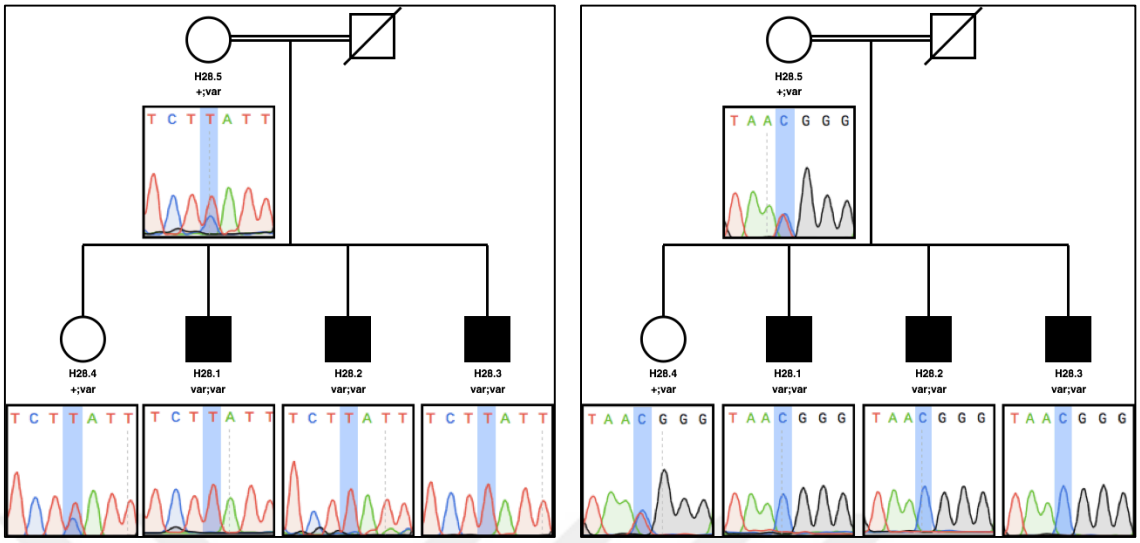


Figure A.25. H28 pedigree and chromatograms showing the the c.103C>T variant in NUTM2D gene (left panel) and the c.701T>C variant in ATAD1 gene (right panel) (+: native allele; var: variant allele).

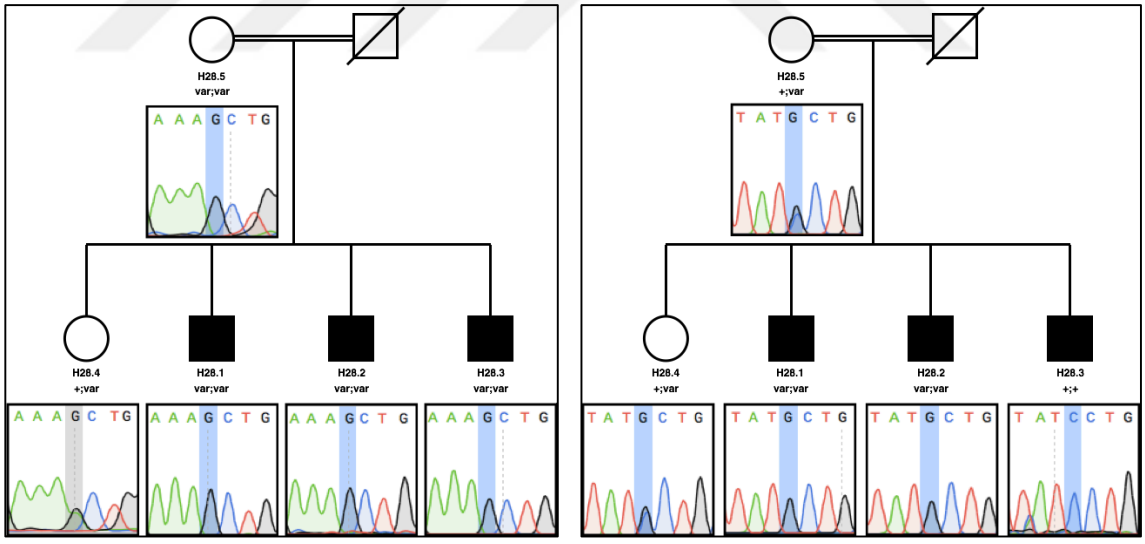


Figure A.26. H28 pedigree and chromatograms showing the c.511A>G variant in LIPF gene (left panel) and the c.980C>G variant in HSDL1(right panel) (+: native allele; var: variant allele).

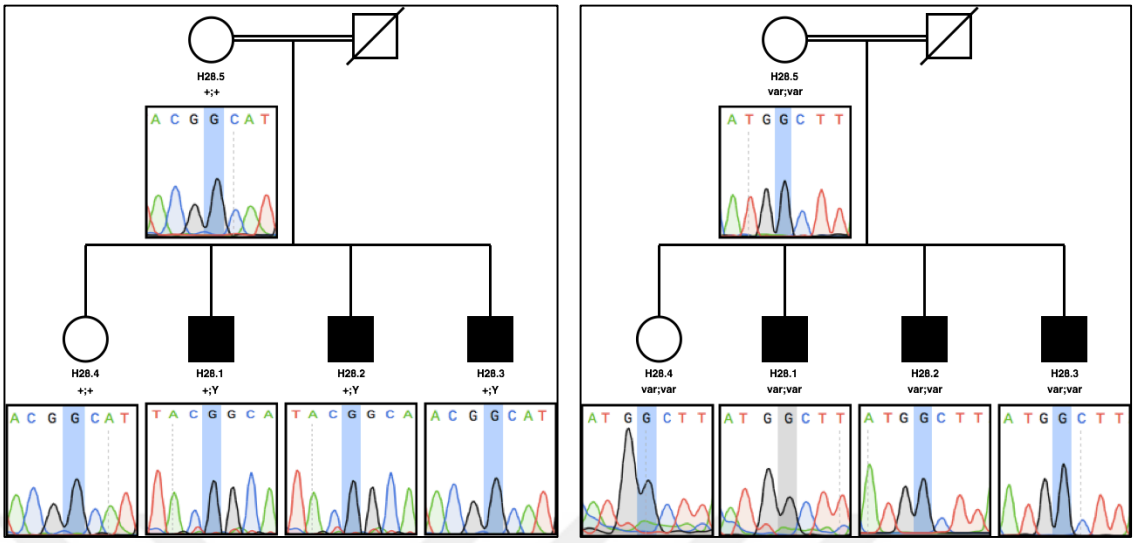


Figure A.27. H28 pedigree and chromatograms showing the c.524G>A variant in 5th exon of ARSD gene (left panel) and the c.959G>A variant in 6th exon ARSD gene (right panel) (+: native allele; var: variant allele).

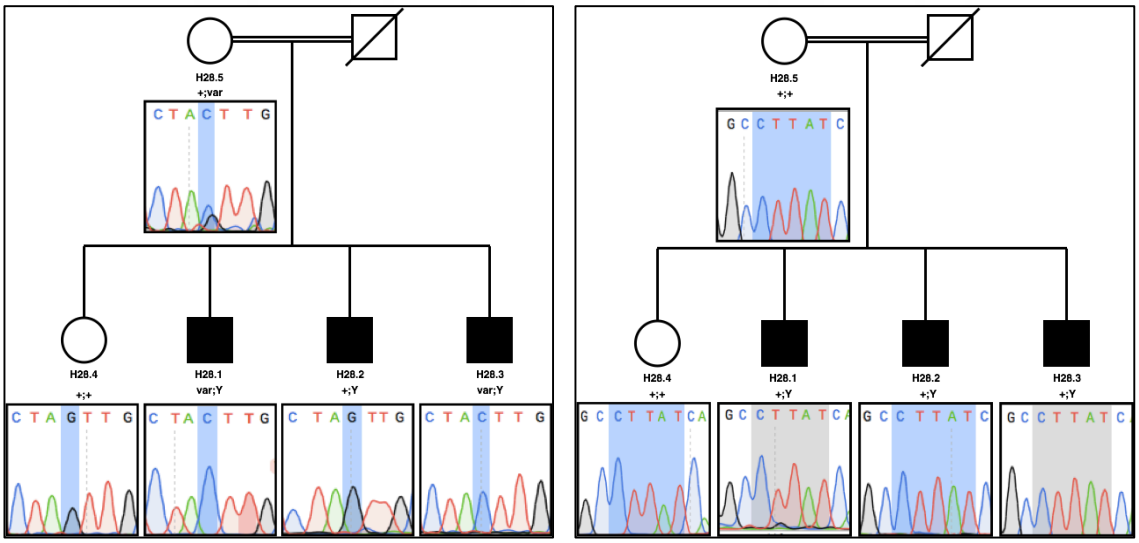


Figure A.28. H28 pedigree and chromatograms showing the c.2299G>C variant in ATP7A gene (left panel) and the c.3060dupC variant in DGKK gene (right panel) (+: native allele; var: variant allele).

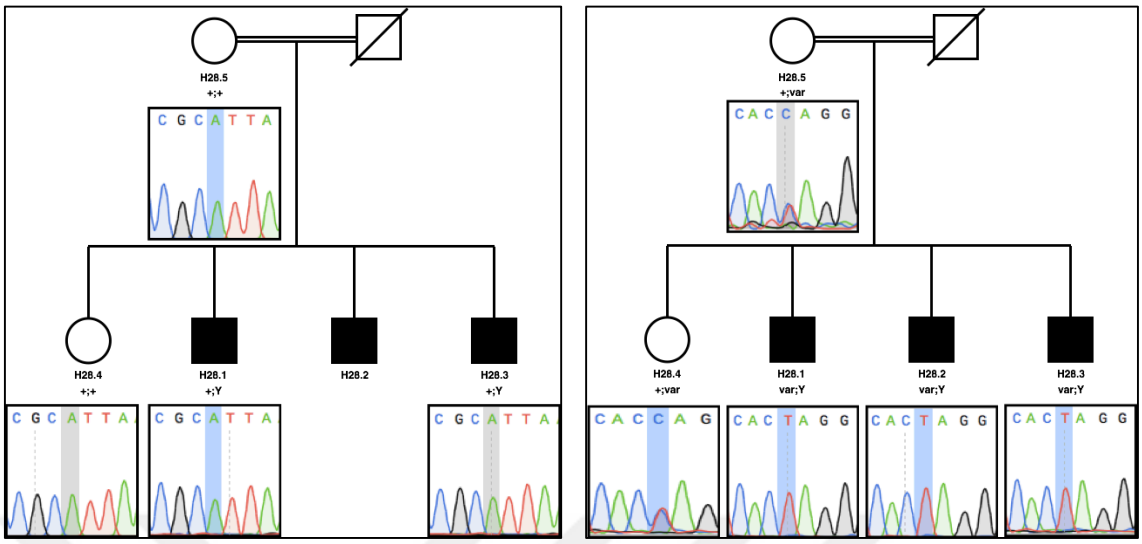


Figure A.29. H28 pedigree and chromatograms showing the c.160A>C variant in FAM104B gene (left panel) and the c.217C>T variant in H2BFM gene (right panel) (+: native allele; var: variant allele).

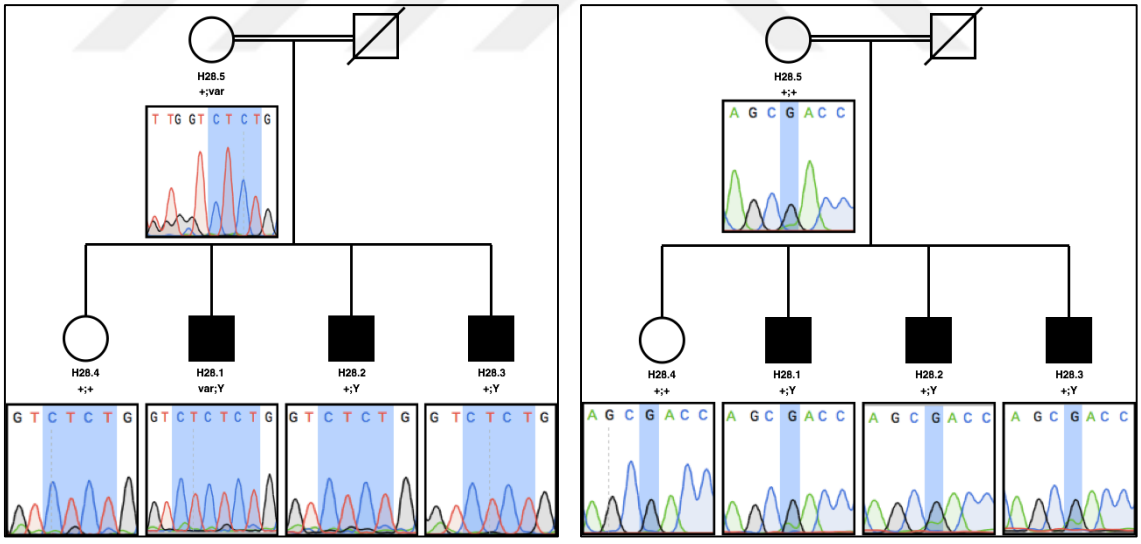


Figure A.30. H28 pedigree and chromatograms showing the c.82-7_82-6dupCT variant in IDH3G gene (left panel) and the c.544G>A variant in SSX7 gene (right panel) (+: native allele; var: variant allele).

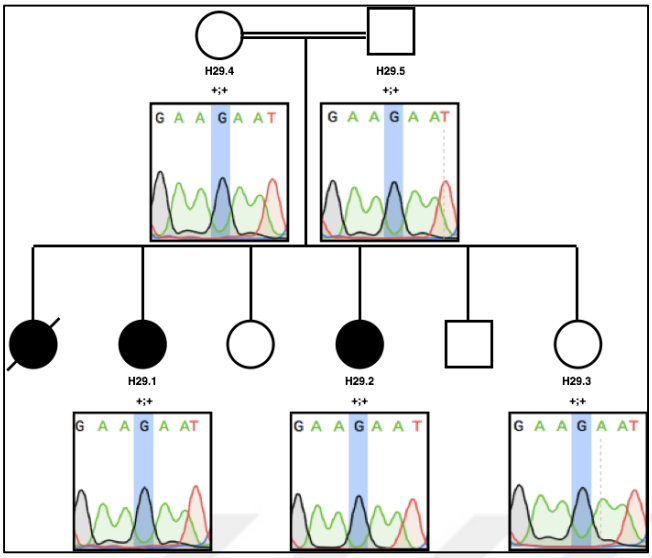


Figure A.31. H29 pedigree and chromatograms showing c.4804C>T variant in ZFYVE26 gene (+: native allele).

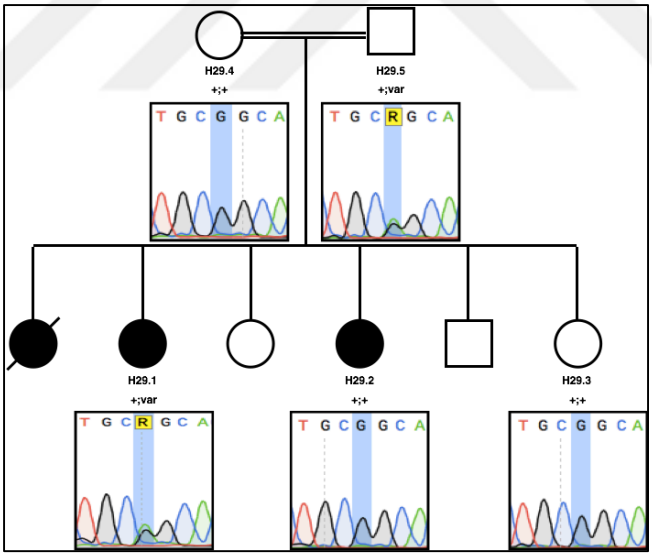


Figure A.32. H29 pedigree and chromatograms showing c.1168G>A, p.Gly390Ser variant in RTN2 gene (+: native allele; var: variant allele).

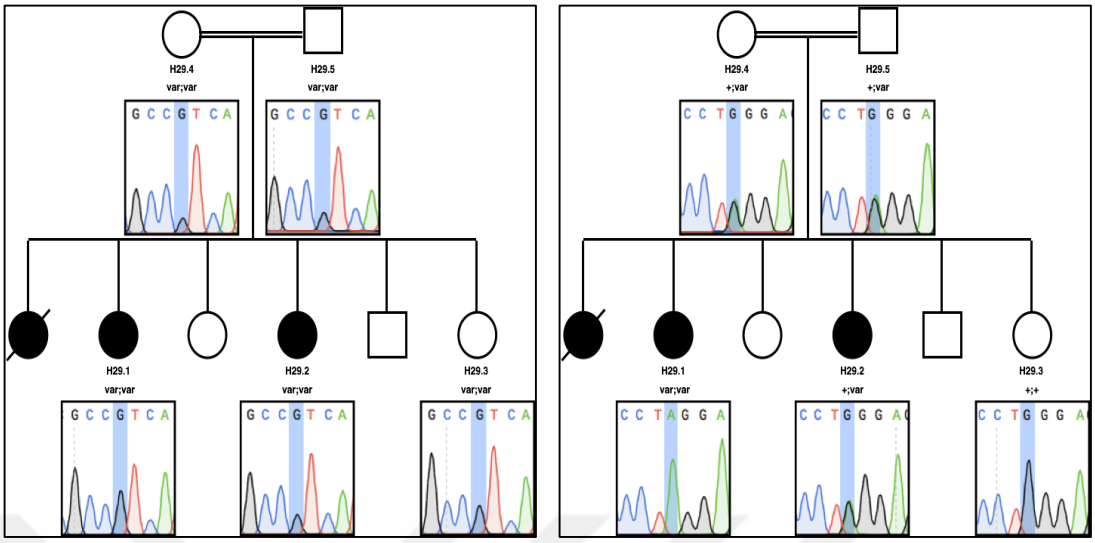


Figure A.33. H29 pedigree and chromatograms showing the c.40C>G variant in TTC22 gene (left panel) and the c.15446-7G>A variant in SSPO gene (right panel) (+: native allele; var: variant allele).

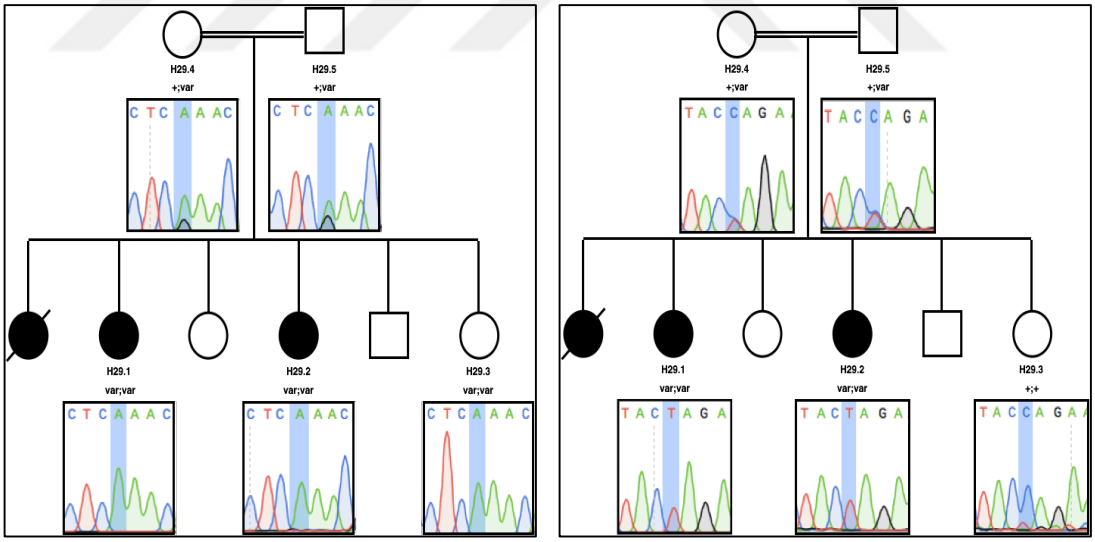


Figure A.34. H29 pedigree and chromatograms showing the c.590G>A variant in NAT2 gene (left panel) and the c.1547C>T variant in SLC22A10 gene (right panel) (+: native allele; var: variant allele).

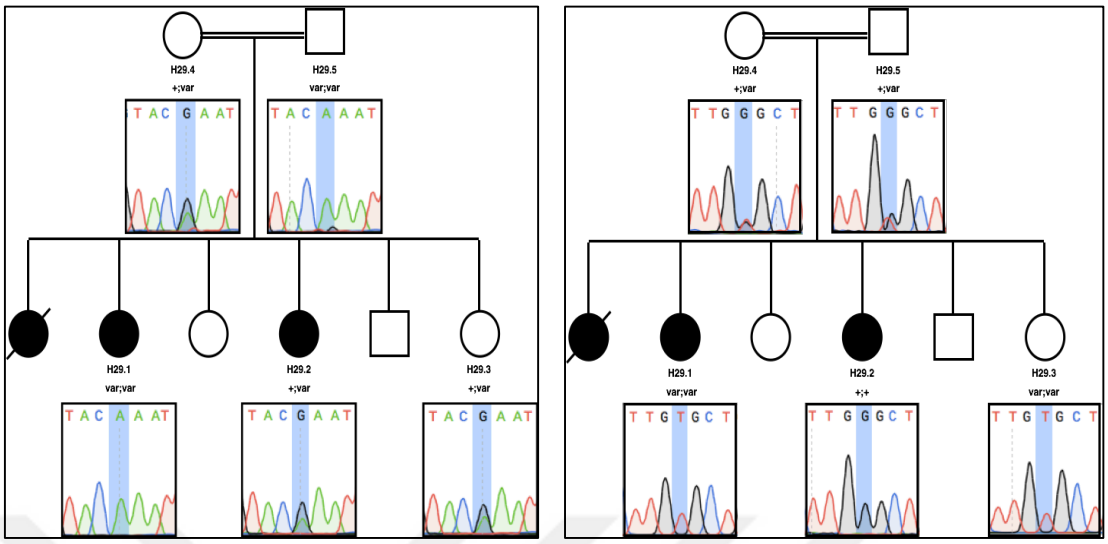


Figure A.35. H29 pedigree and chromatograms showing the variant c.401G>A in KRR1 gene (left panel) and the c.1871G>T variant in AKAP13 gene (right panel) (+: native allele; var: variant allele).

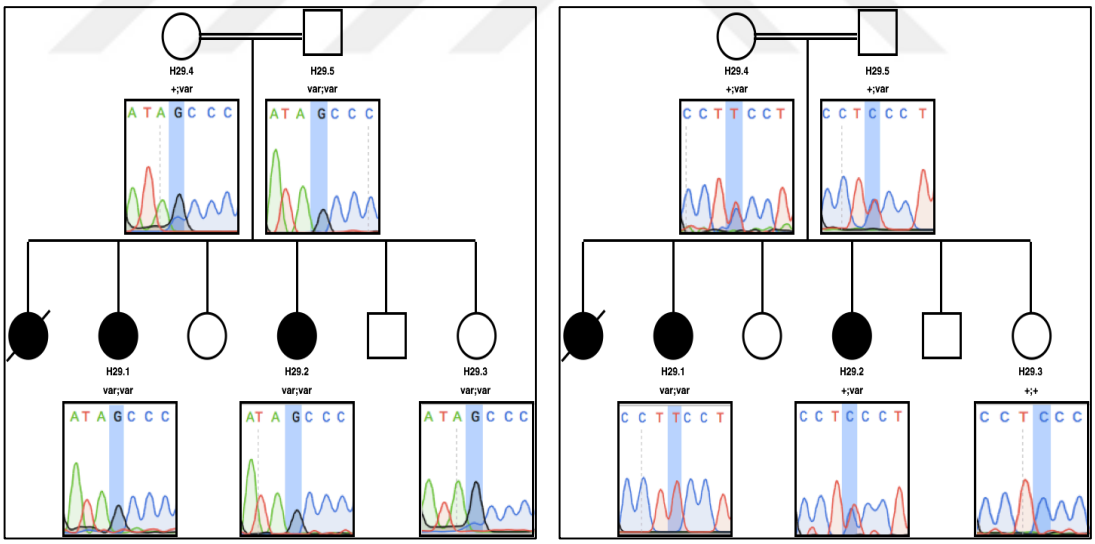


Figure A.36. H29 pedigree and chromatograms showing the the c.1144C>G variant in PDIA2 gene (left panel) and the c.2594C>T variant in WDR90 gene (right panel) (+: native allele; var: variant allele).

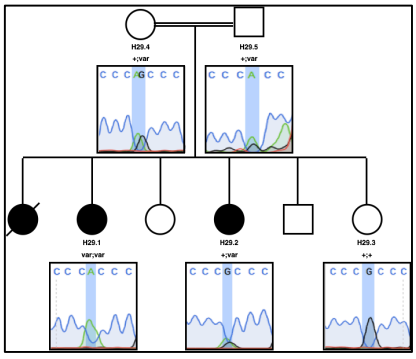


Figure A.37. H29 pedigree and chromatograms showing the c.1798-8G>A variant in CLCN7 gene (+: native allele; var: variant allele).

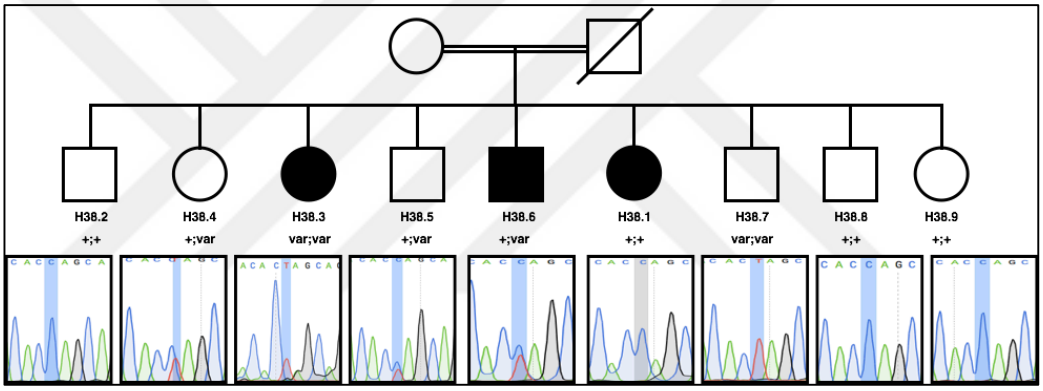


Figure A.38. H38 pedigree and chromatograms showing the c.39962C>T variant in TTN gene (+: native allele; var: variant allele).

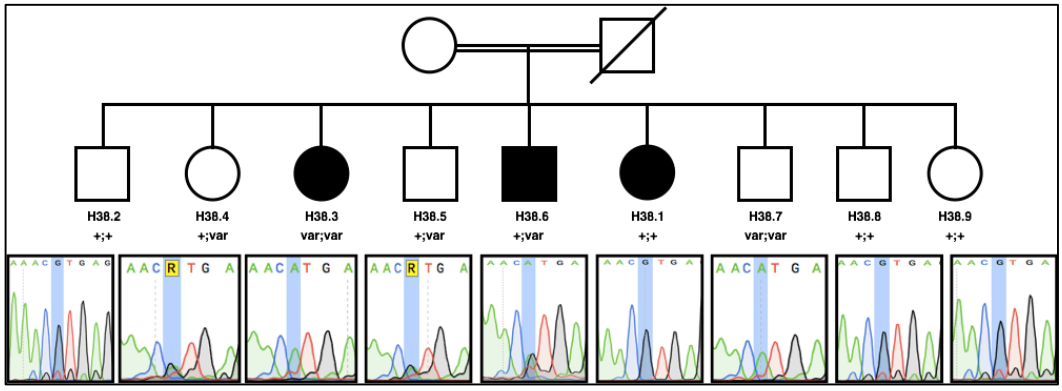


Figure A.39. H38 pedigree and chromatograms showing the c.63917G>A variant in TTN-AS1, TTN gene (+: native allele; var: variant allele).

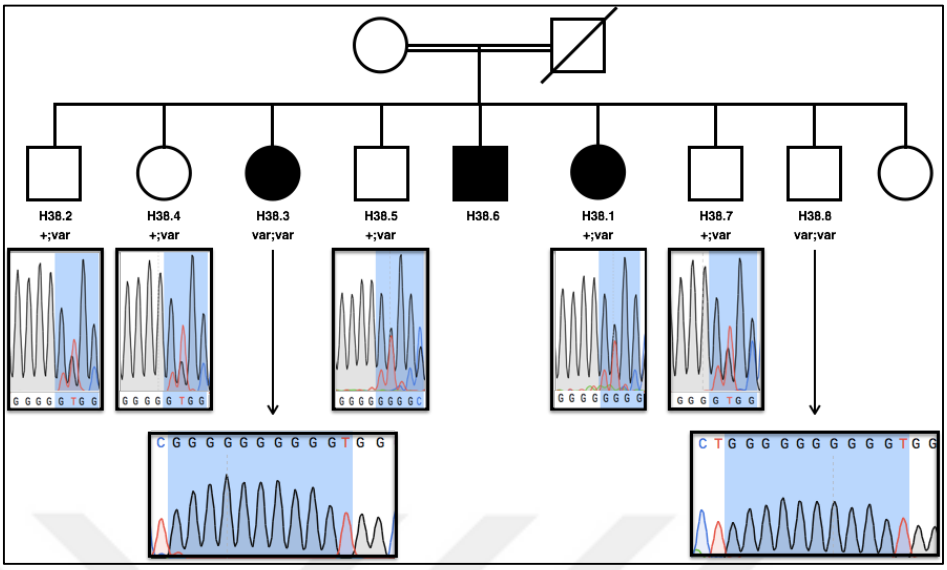


Figure A.40. H38 pedigree and chromatograms showing the c.-29delG variant in ARNT gene (+: native allele; var: variant allele).

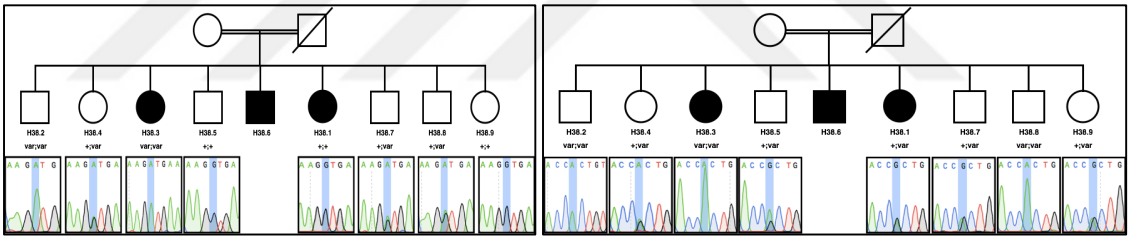


Figure A.41. H38 pedigree and chromatograms showing the c.2128G>A variant in EMC1 gene (left panel) and the c.67G>A variant in RNF186 gene (right panel) (+: native allele; var: variant allele).

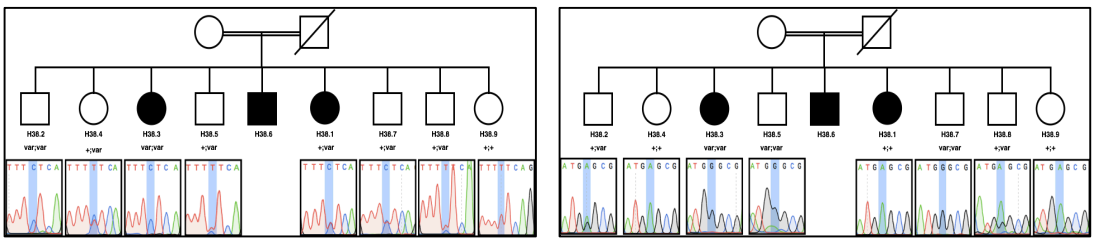


Figure A.42. H38 pedigree and chromatograms showing the c.3972-5T>C variant in MYOM3 gene (left panel) and c.901A>G variant in TRAPPC12 gene (right panel) (+: native allele; var: variant allele).

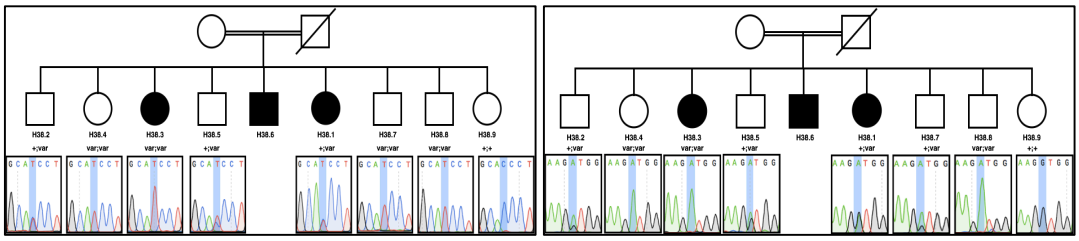


Figure A.43. H38 pedigree and chromatograms showing the c.293C>T variant in APOB gene (left panel) and the c.847G>A variant in CAPN14 gene (right panel) (+: native allele; var: variant allele).

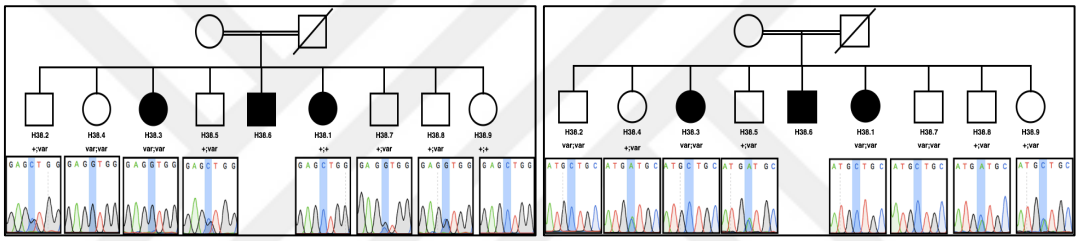


Figure A.44. H38 pedigree and chromatograms showing the c.1294C>G variant in WFS1 gene (left panel) the c.2128G>A variant in LOXL4 gene (right panel) (+: native allele; var: variant allele).

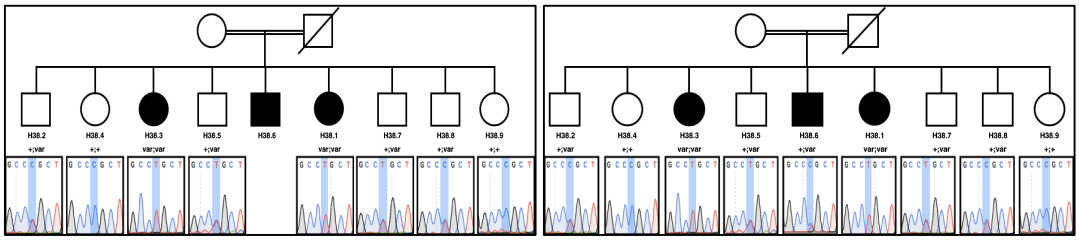


Figure A.45. H38 pedigree and chromatograms showing the c.67G>A variant in PKD2L1 gene (left panel) and the c.2128G>A variant in PREX1 gene (+: native allele; var: variant allele).

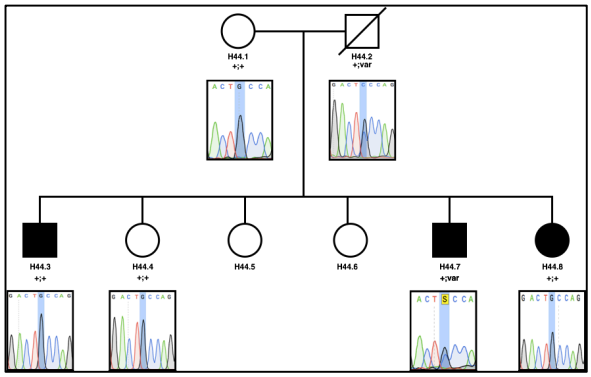


Figure A.46. H44 pedigree and chromatograms showing c.7417-5G>C variant in ZFYVE26 gene (+: native allele; var: variant allele).

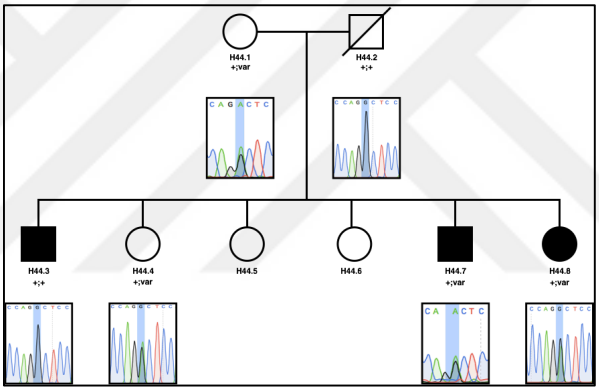


Figure A.47. H44 pedigree and chromatograms showing c.7607G>A variant in ZFYVE26 gene (+: native allele; var: variant allele).

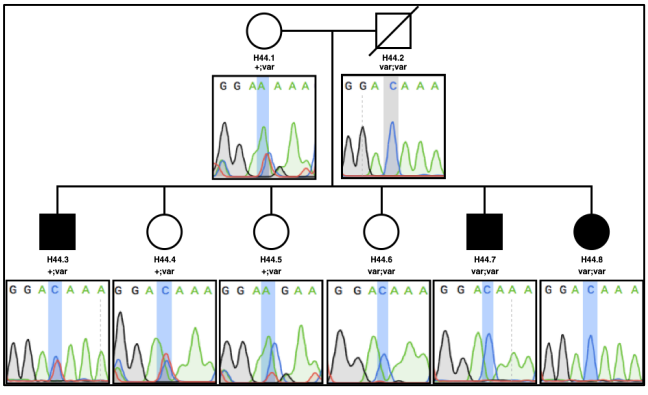


Figure A.48. H44 pedigree and chromatograms showing c.503T>C variant in KLRB1 gene (+: native allele, var: variant allele).

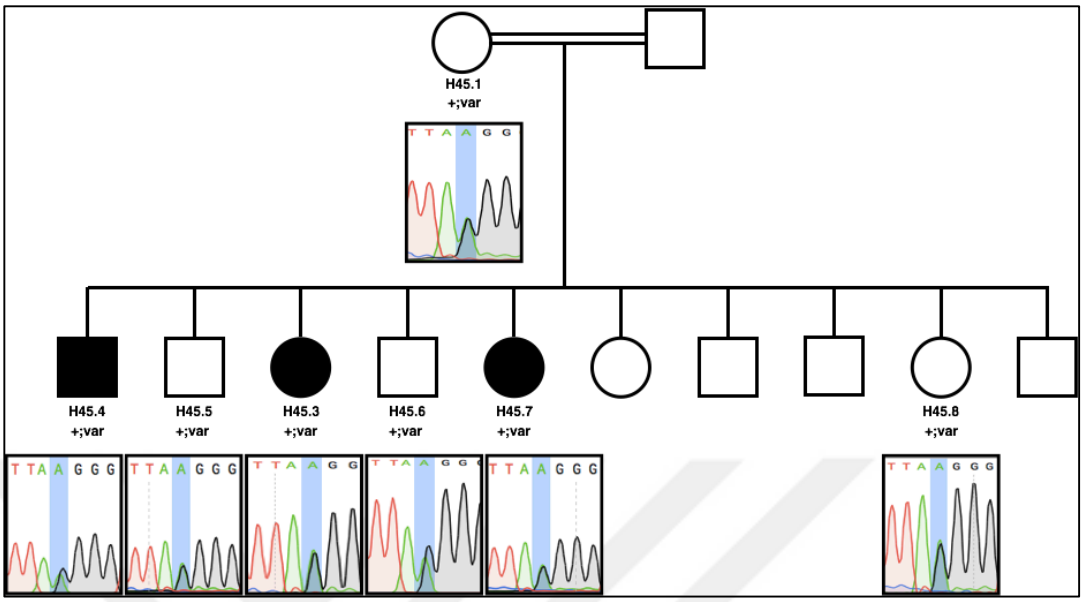


Figure A.49. H45 pedigree and chromatograms showing c.2245-1G>A variant in SPG11 gene (+: native allele, var: variant allele).

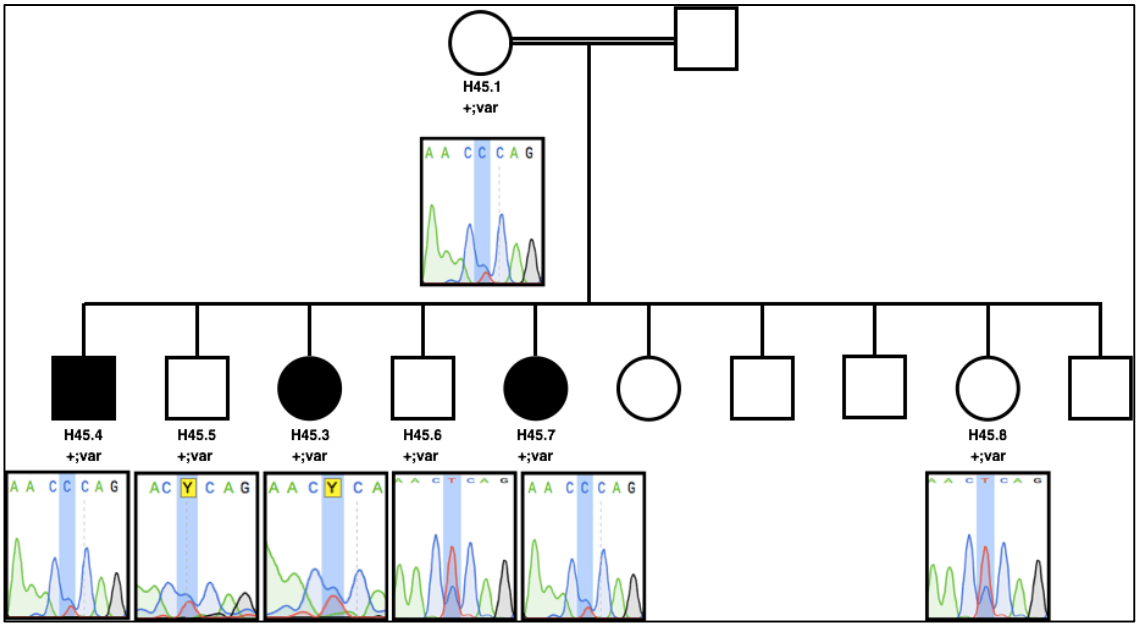


Figure A.50. H45 pedigree and chromatograms showing c.2057T>C variant in SPG11 gene (+: native allele, var: variant allele).

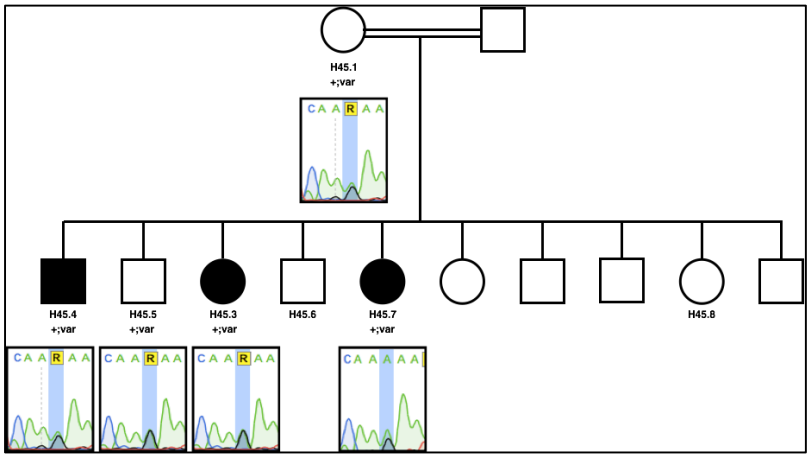


Figure A.51. H45 pedigree and chromatograms showing c.1499_1500delAG, p.E500Vfs57X variant in SPG11 gene (+: native allele; var: variant allele).

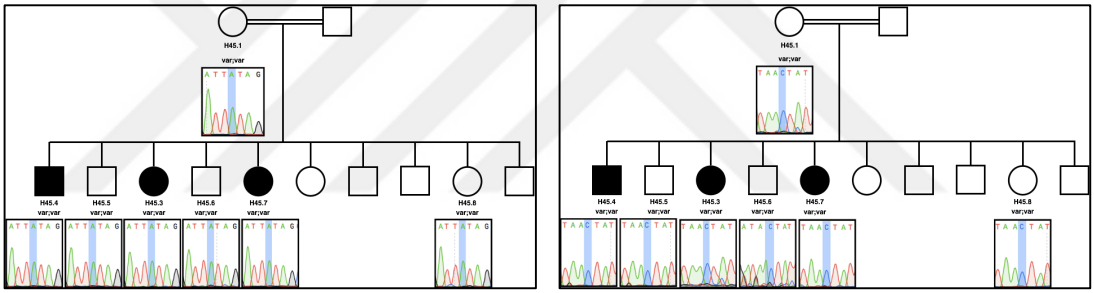


Figure A.52. H45 pedigree and chromatograms showing the c.1999G>A variant in PLCL1 gene (left panel) and the c.2487A>C variant in ALPK2 gene (right panel) (+: native allele; var: variant allele).

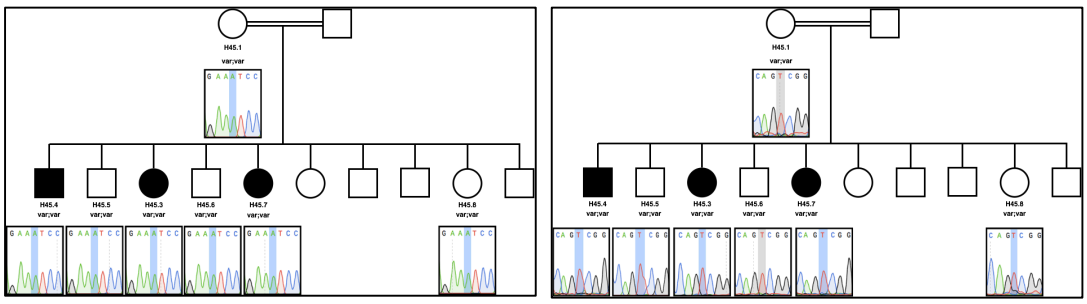


Figure A.53. H45 pedigree and chromatograms showing c.2748T>A variant (left panel) and the c.4345C>T variant in ALPK2 gene (right panel) (+: native allele; var: variant allele).

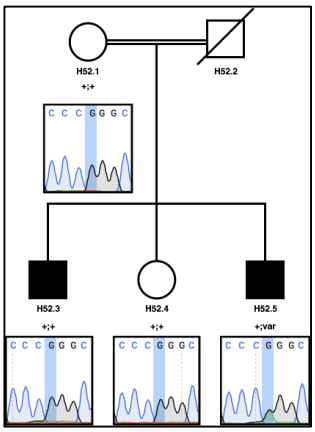


Figure A.54. H52 pedigree and chromatograms showing c.1214G>A variant in SPG7 gene.

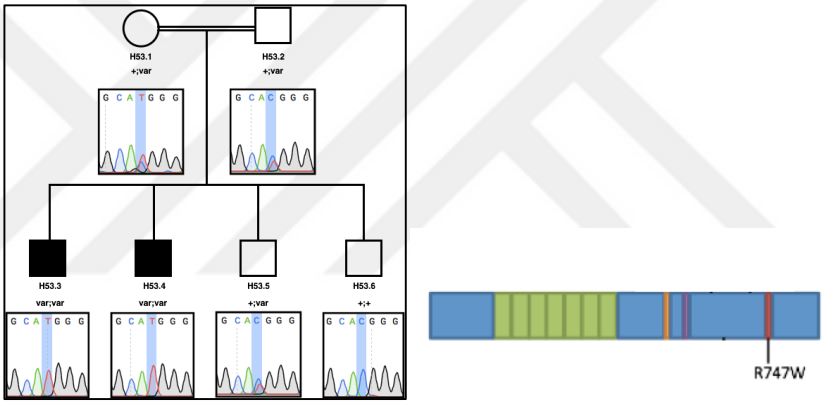


Figure A.55. H53 pedigree and chromatograms showing c.2239G>A variant (left panel) and domain structure of the PLA2G6 gene indicating the position of the mutations identified in this study (modified from Engel *et al*. 2010) (right panel).

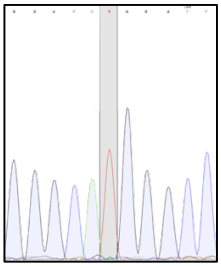


Figure A.56. Chromatogram showing c.2239G>A variant in patient HSS-2.

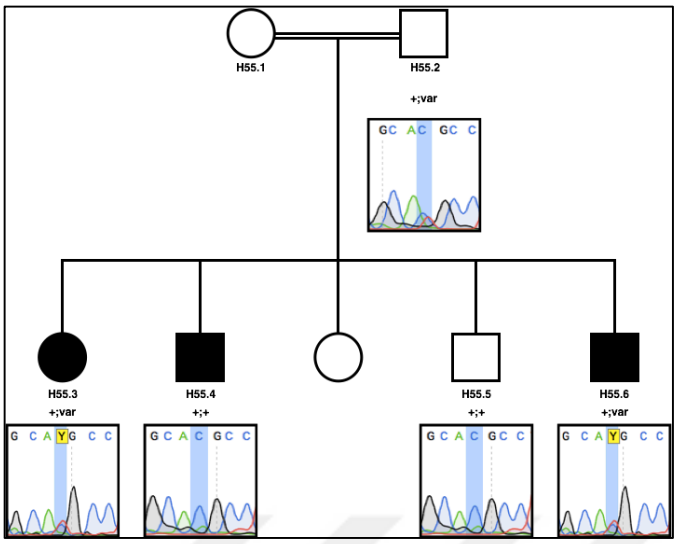


Figure A.57. H55 pedigree and the chromatograms showing c.571C>T variant in SHMT1 gene (+:native allele, var: variant allele).

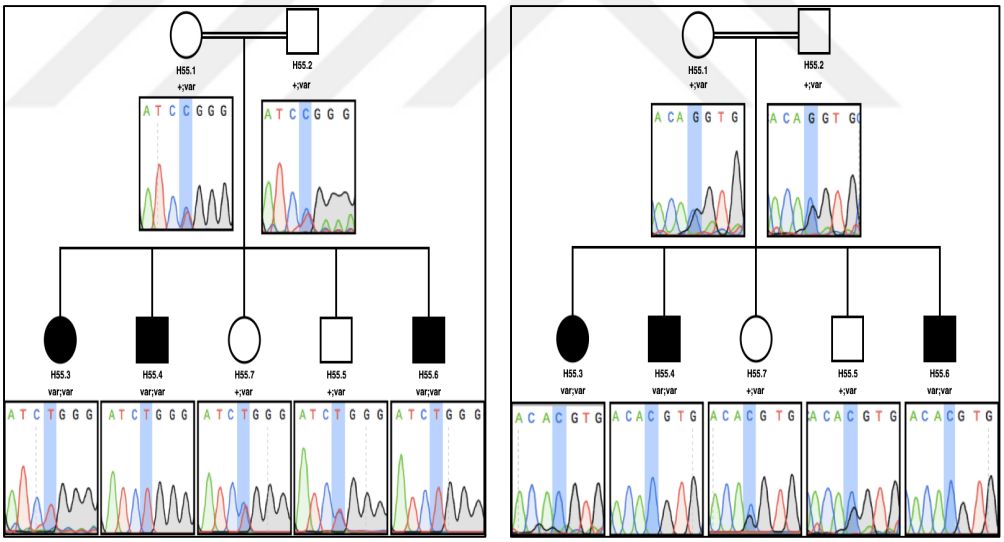


Figure A.58. H55 pedigree and chromatograms showing c.3373C>T variant in ZNF142 gene (left panel) and the c.256-1G>C variant in CYP27A1 gene (right panel) (+: native allele; var: variant allele).

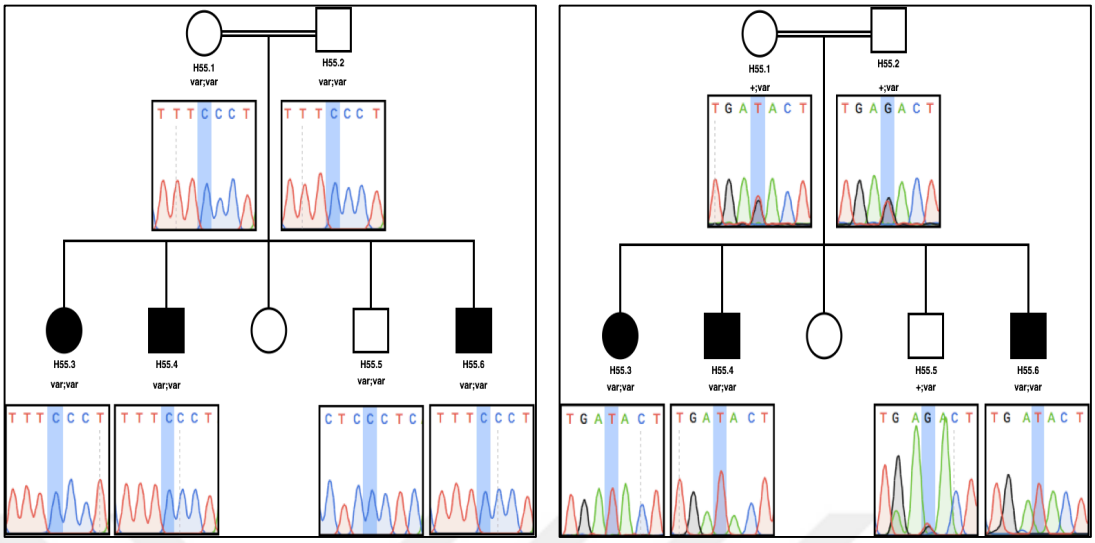


Figure A.59. H55 pedigree and chromatograms showing c.443-6T>C variant in SLC23A3 gene (left panel) and the c.90G>Tvariant in RESP18 gene (right panel) (+: native allele; var: variant allele).

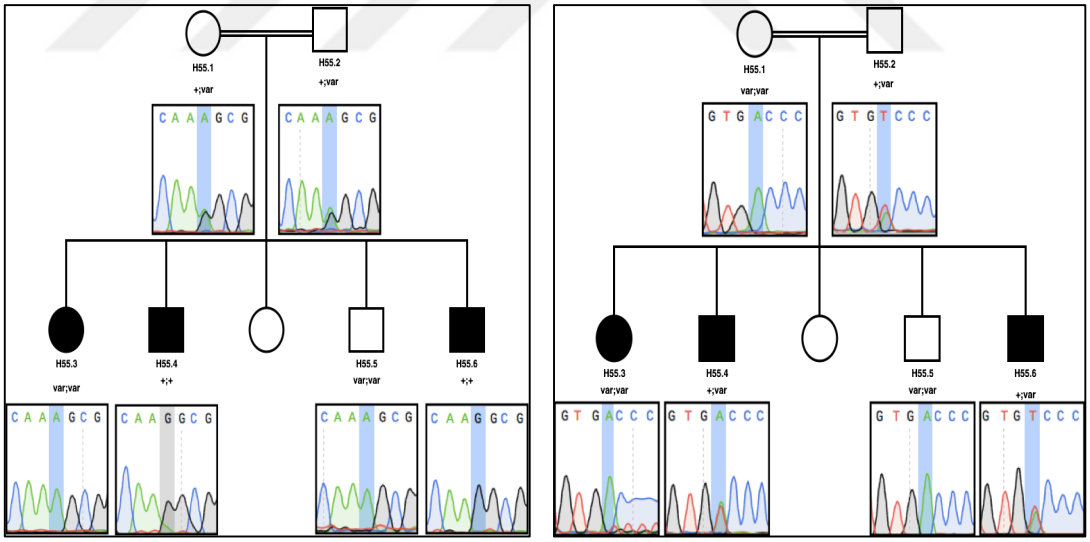


Figure A.60. H55 pedigree and chromatograms showing c.1615G>A variant in FAM198A gene (left panel) and the c.1348-7T>A variant in NT5DC2 gene (right panel) (+: native allele; var: variant allele).

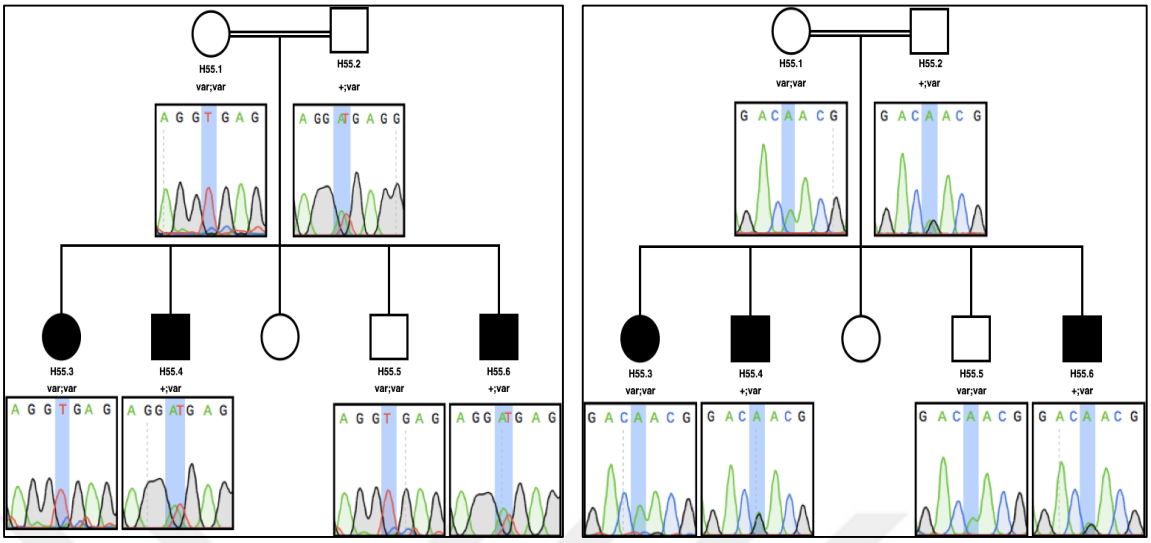


Figure A.61. H55 pedigree and chromatograms showing c.1754A>T variant in ITIH1 gene (left panel) and the c.1379G>A variant in CCDC66 gene (right panel) (+: native allele; var: variant allele).

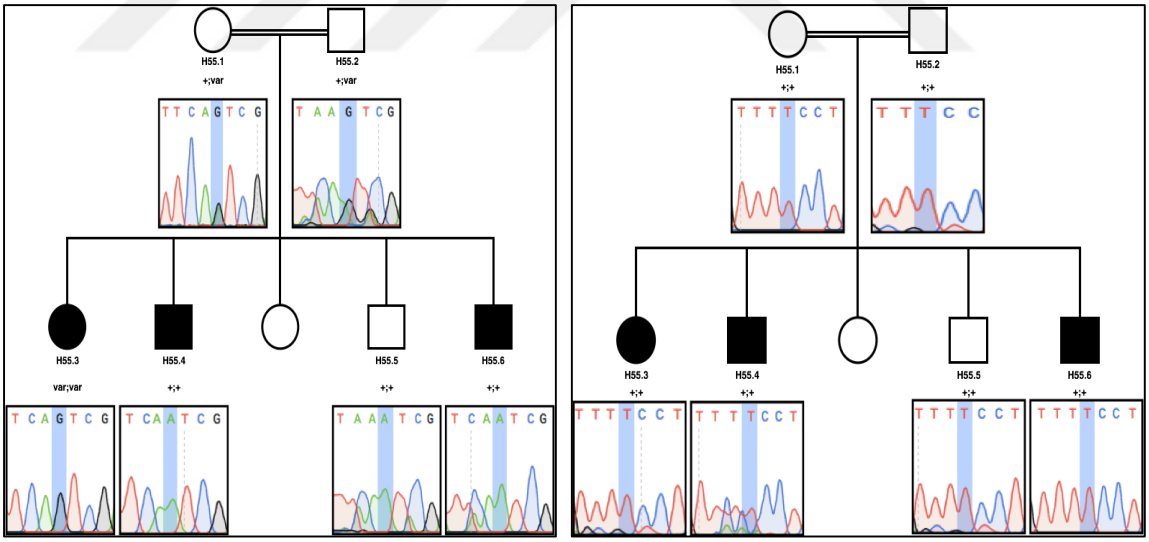


Figure A.62. H55 pedigree and chromatograms showing c.3923A>G variant in DNAH12 gene (left panel) and the c.77T>C variant in CDC27 gene (right panel) (+: native allele; var: variant allele).

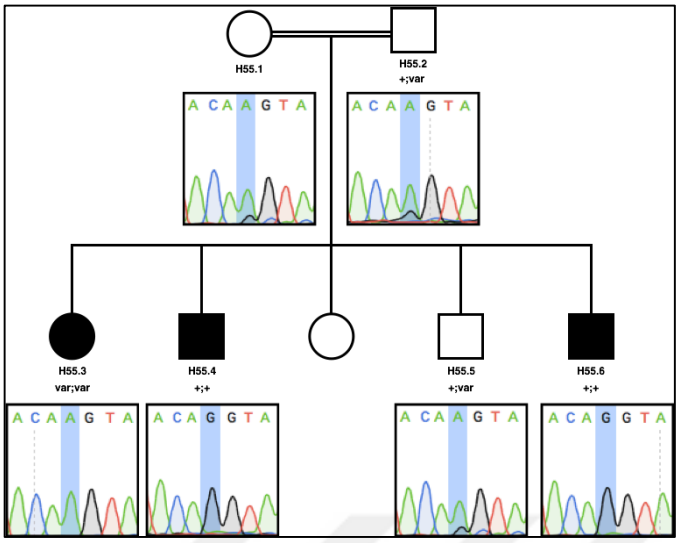


Figure A.63. H55 pedigree and chromatograms showing c.806-1G>A variant in EFCAB13 gene (+: native allele; var: variant allele).

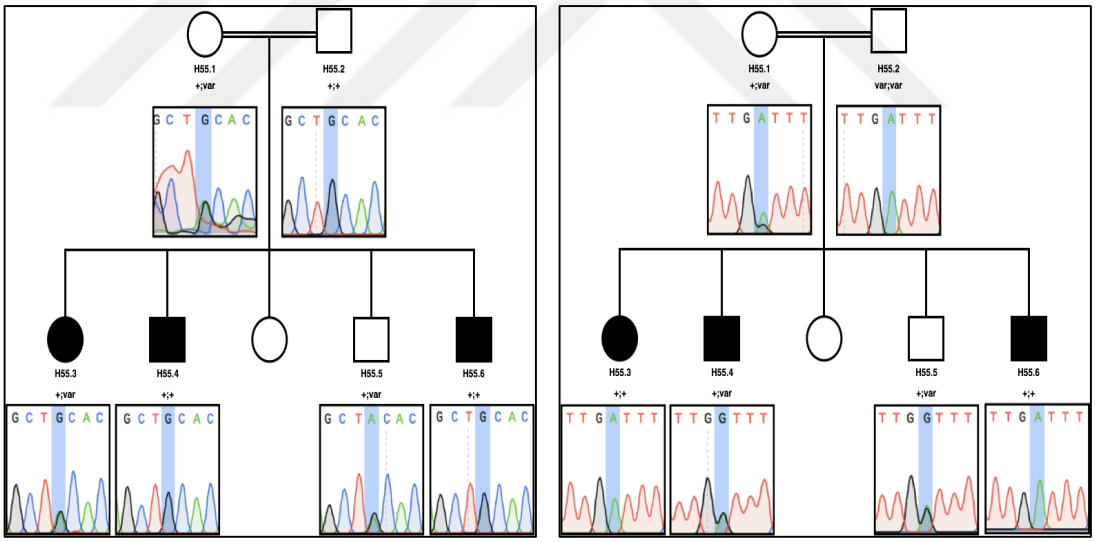


Figure A.64. H55 pedigree and chromatograms showing c.1627G>A variant in HEXB gene (left panel) and the c.208A>G variant in DST gene (right panel) (+: native allele; var: variant allele).

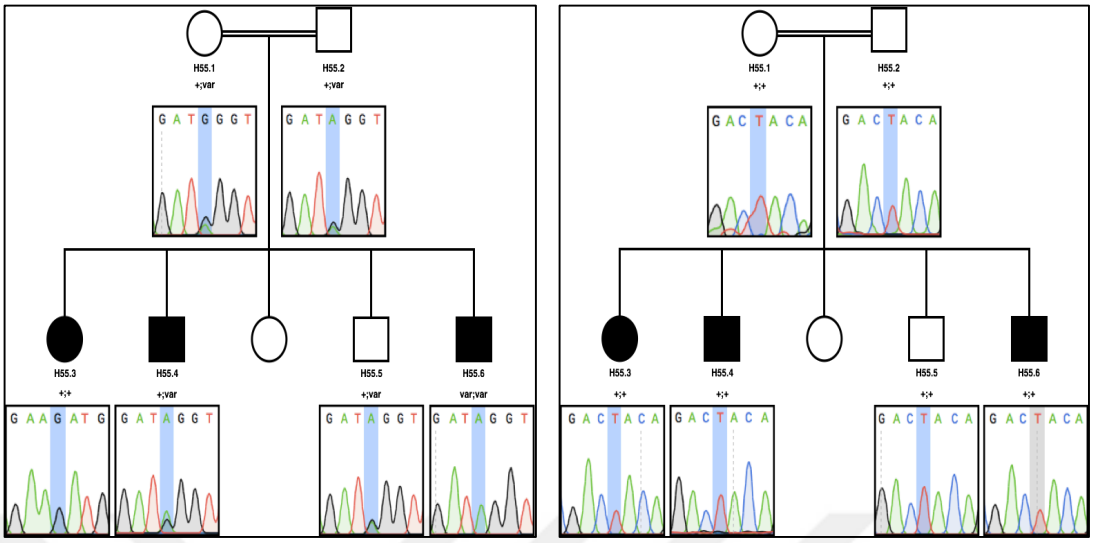


Figure A.65. H55 pedigree and chromatograms showing c.589G>A variant in FKBP14 gene (left panel) and the c.175T>A variant in SYT12 gene (right panel) (+: native allele; var: variant allele).

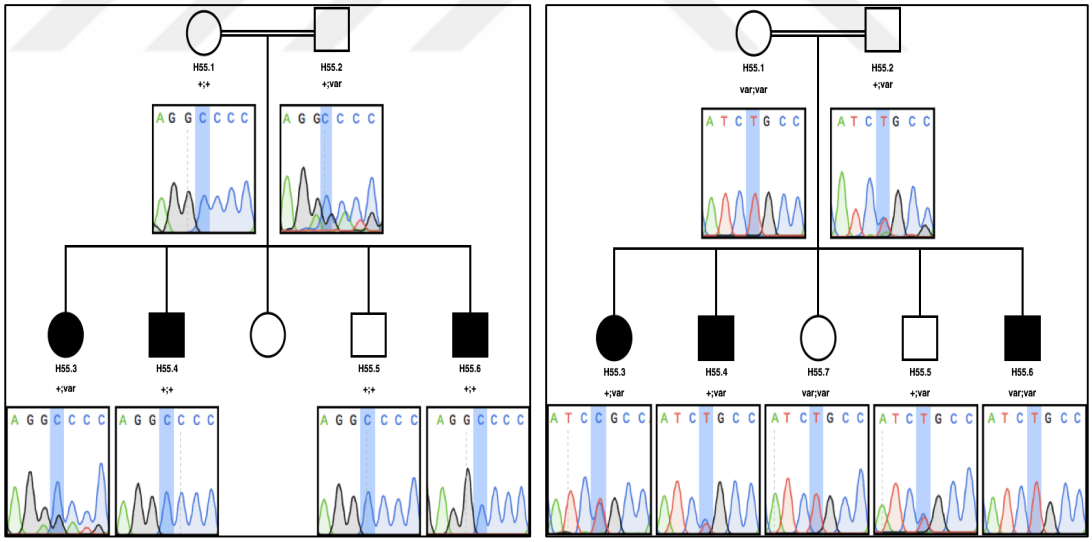


Figure A.66. H55 pedigree and chromatograms showing c.452C>G variant in MYH4 gene (left panel) and the c.2395C>T variant in CTDP1 gene (right panel) (+: native allele; var: variant allele).

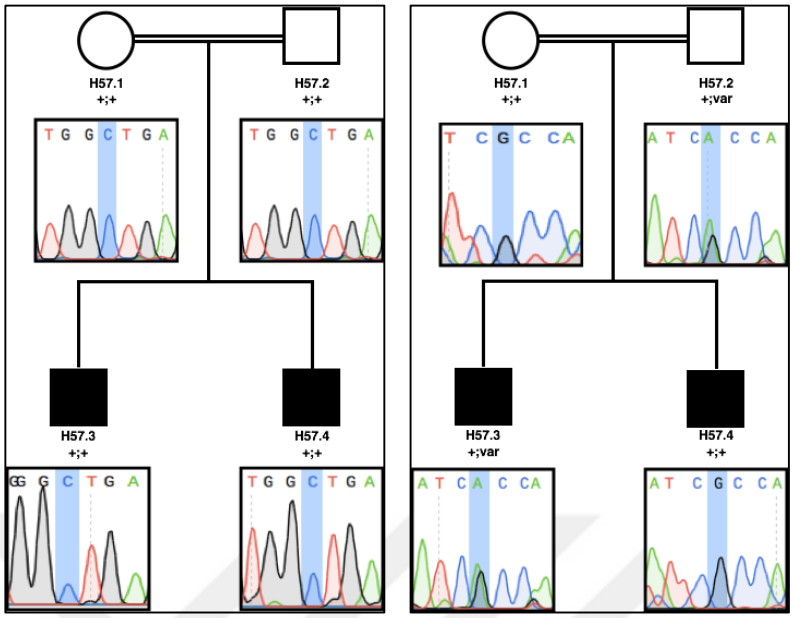


Figure A.67. H57 pedigree and chromatograms showing c.674C>A variant in ERLIN2 gene (left panel) and c.6062G>A variant in ZFYVE26 (right panel).

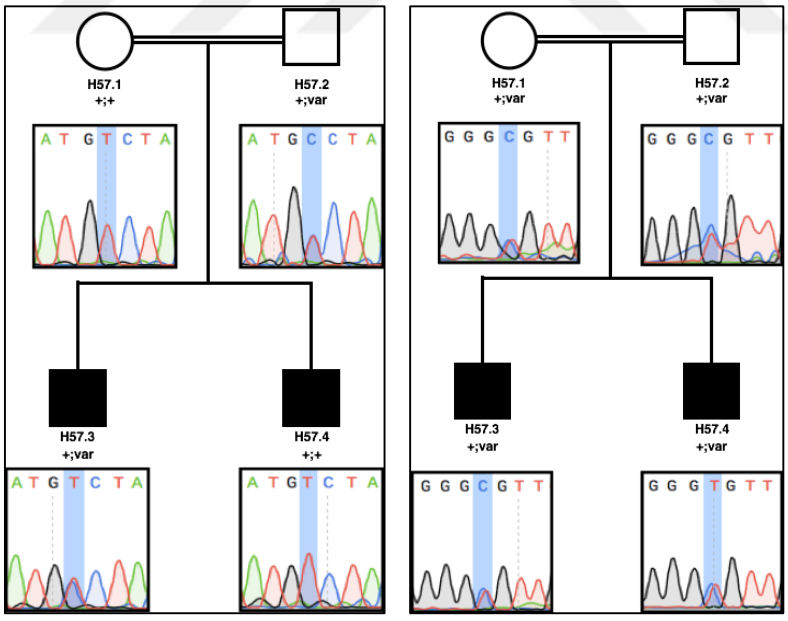


Figure A.68. H57 pedigree and chromatograms showing c.8611T>C variant in SACS gene (left panel) and variant in position chr9:92715016 in BICD2 gene (right panel).

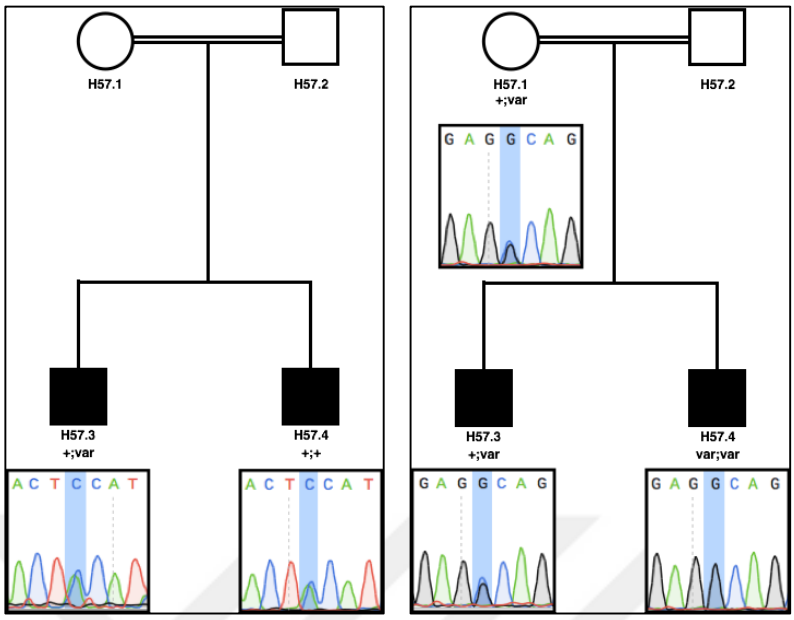


Figure A.69. Chromatograms showing variants in genes CEP350 (left panel) and SH3TC1 (right panel) for H57.3

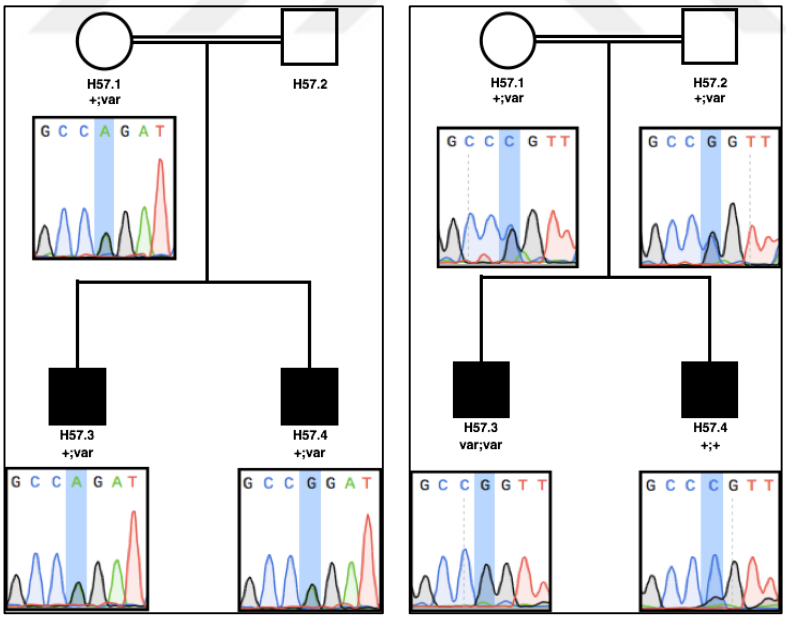


Figure A.70. H57 pedigree and chromatograms showing c.754G>A variant in PIP5K1B (FAM122A) gene (left panel) and c.79C>G variant in COQ9 gene (right panel).

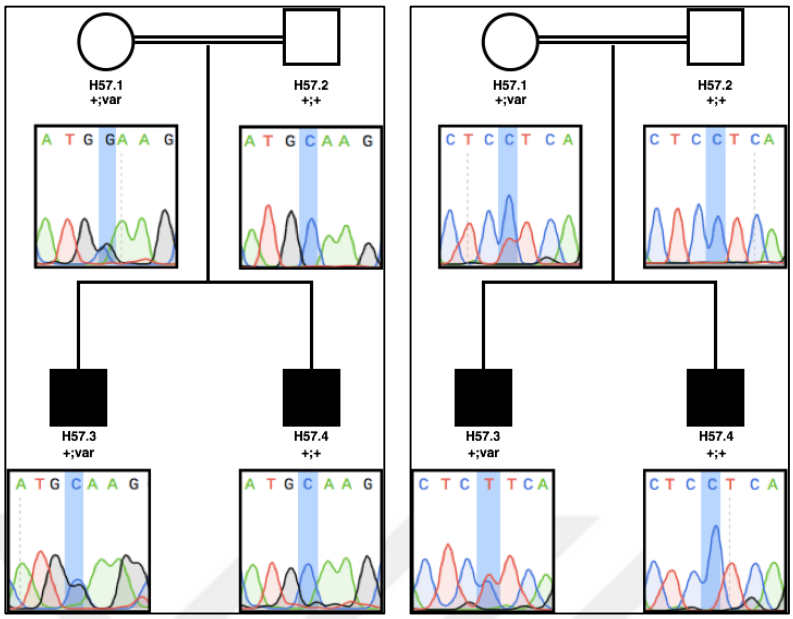


Figure A.71. H57 pedigree and chromatograms showing c.1360C>G (left panel) and c.1373C>T (right panel) variants in NCOA2 gene and.

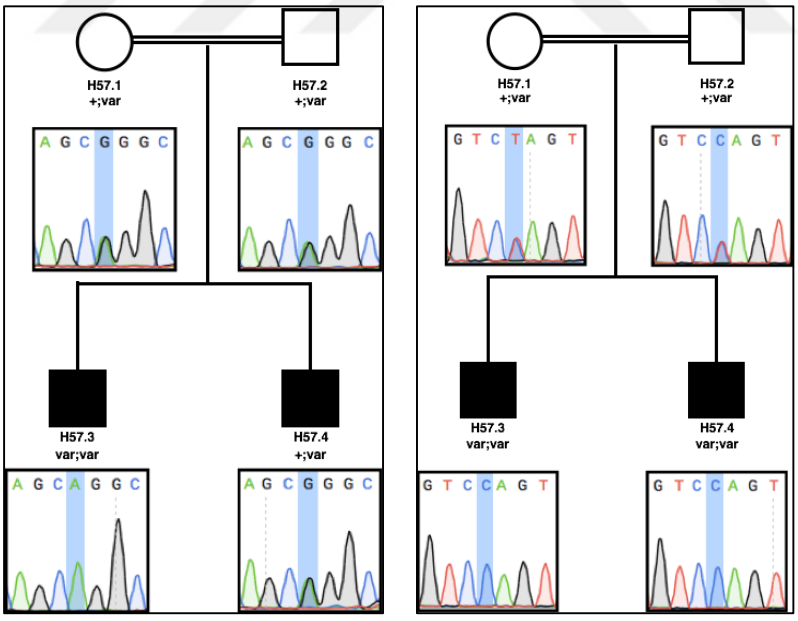


Figure A.72. H57 pedigree and chromatograms showing c.2591G>A variant in RBP3 gene (left panel) and c.533T>C (p.L178P) variant in SAMHD1 gene (right panel).

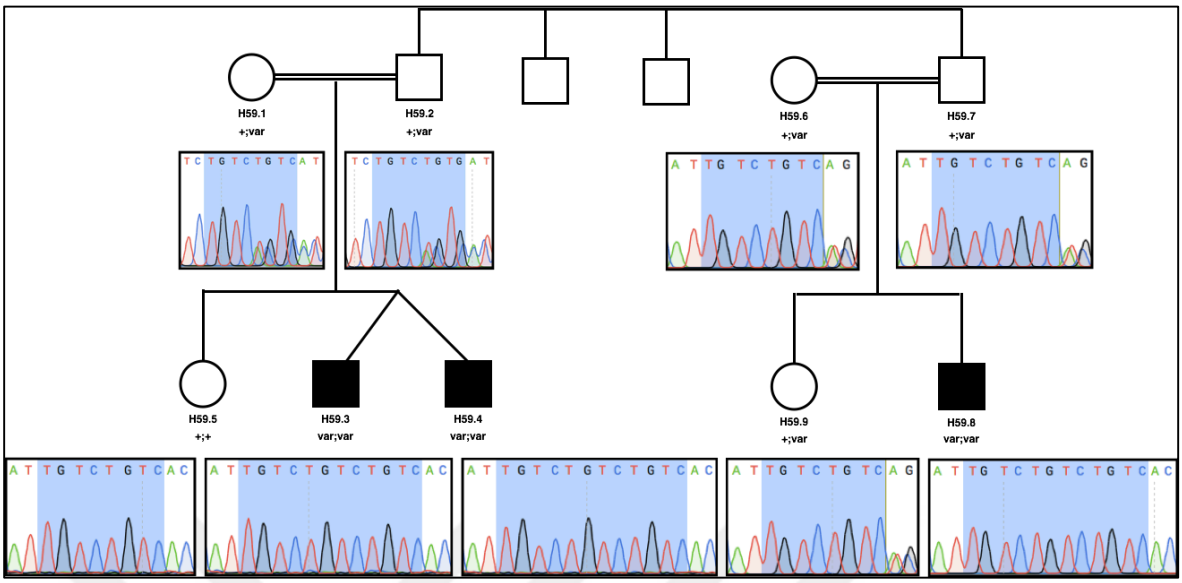


Figure A.73. H59 pedigree and chromatograms showing c.325_326insTGTC insertion in ALS2 gene.

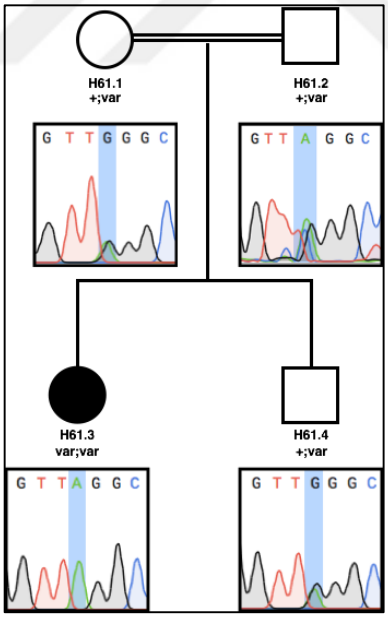


Figure A.74. H61 pedigree and chromatograms showing c.4321C>T variant in Spastizin gene.

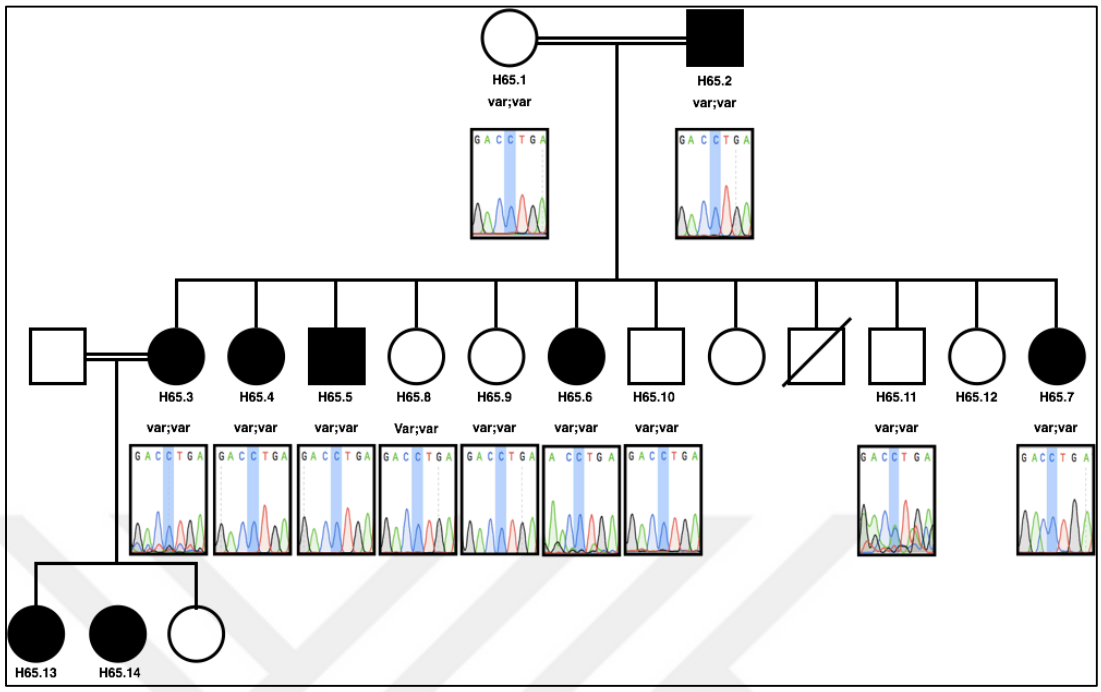


Figure A.75. H65 pedigree and chromatograms showing c.11264T>C variant in PLEC gene.

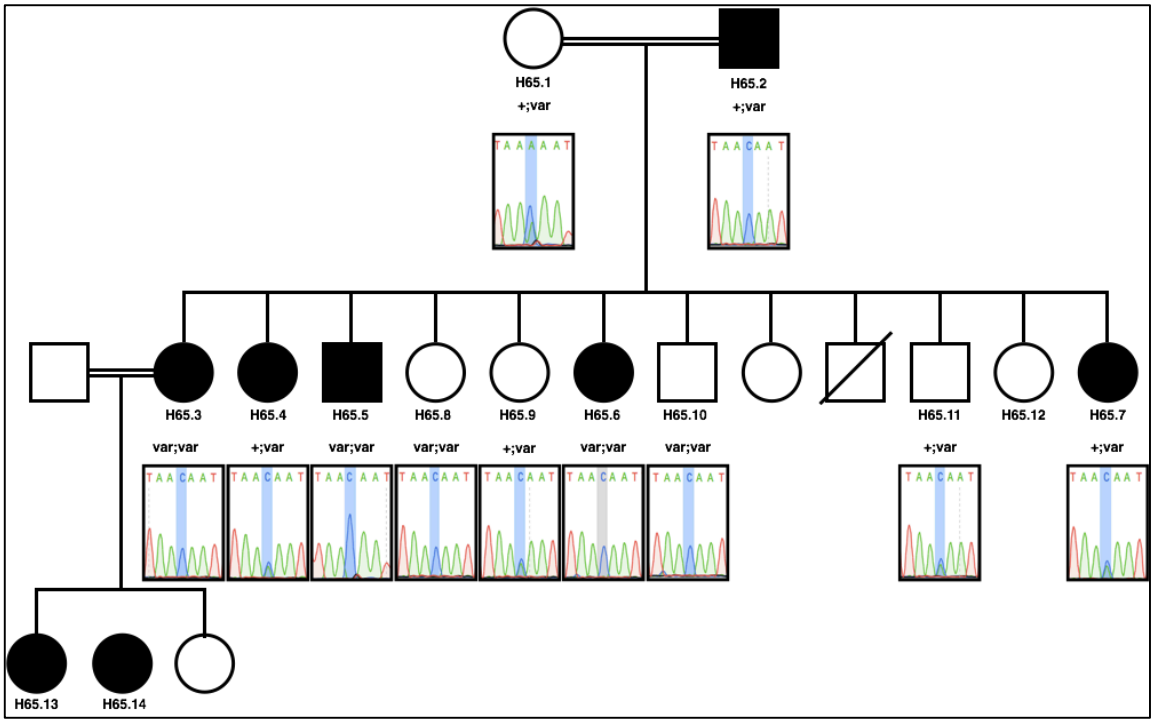


Figure A.76. H65 pedigree and chromatograms showing c.14378C>A variant in DST gene.

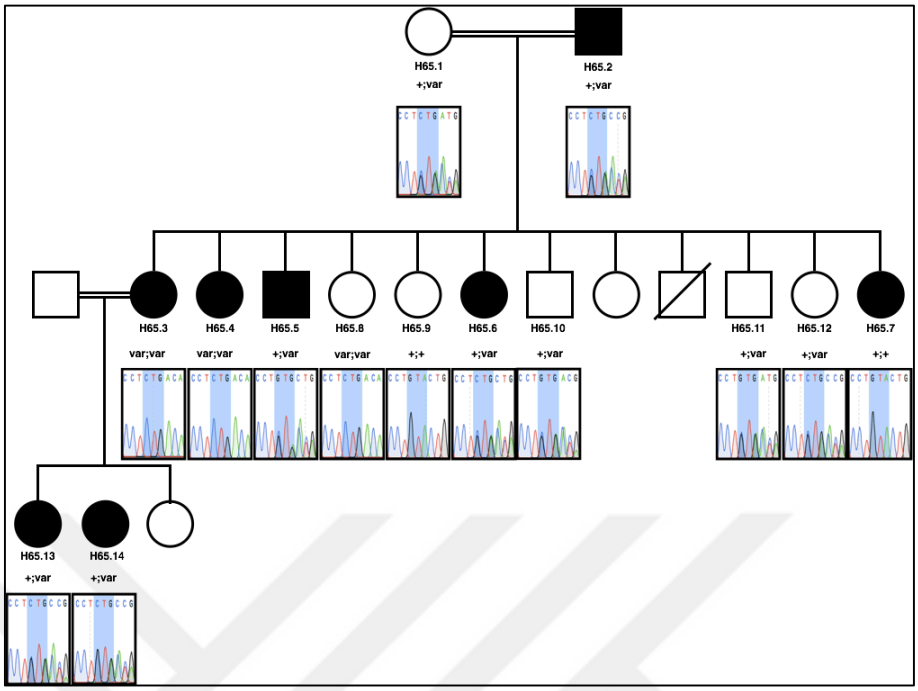


Figure A.77. H65 pedigree and chromatograms showing the c.4211_4213delGTA variant in VPS13B gene (+: native allele; var: variant allele).

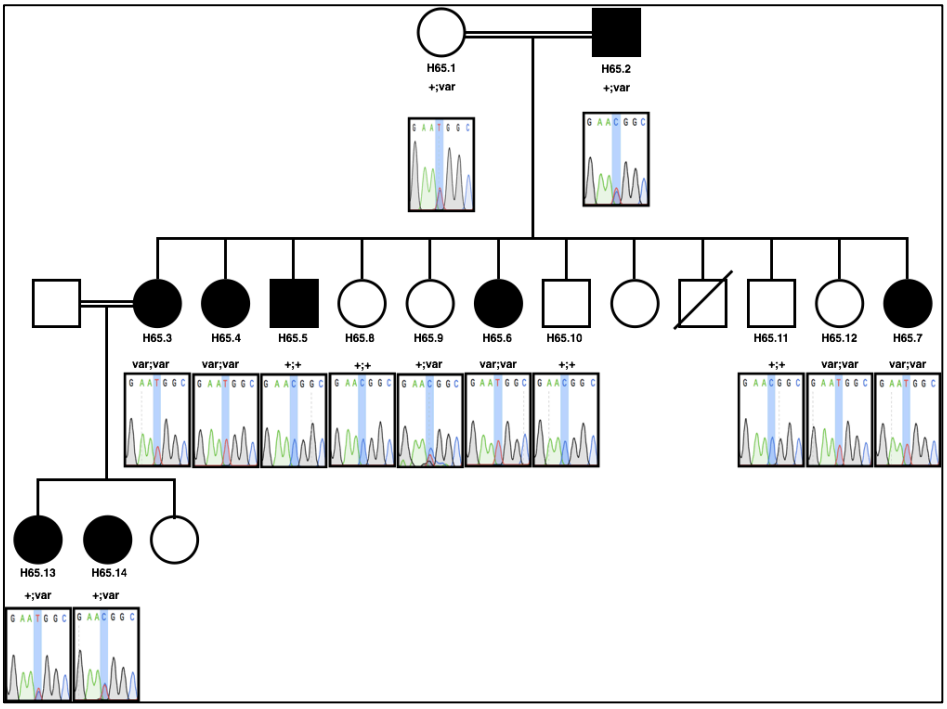


Figure A.78. H65 pedigree and chromatograms showing the c.6034C>T variant in EVPL gene (+: native allele; var: variant allele).

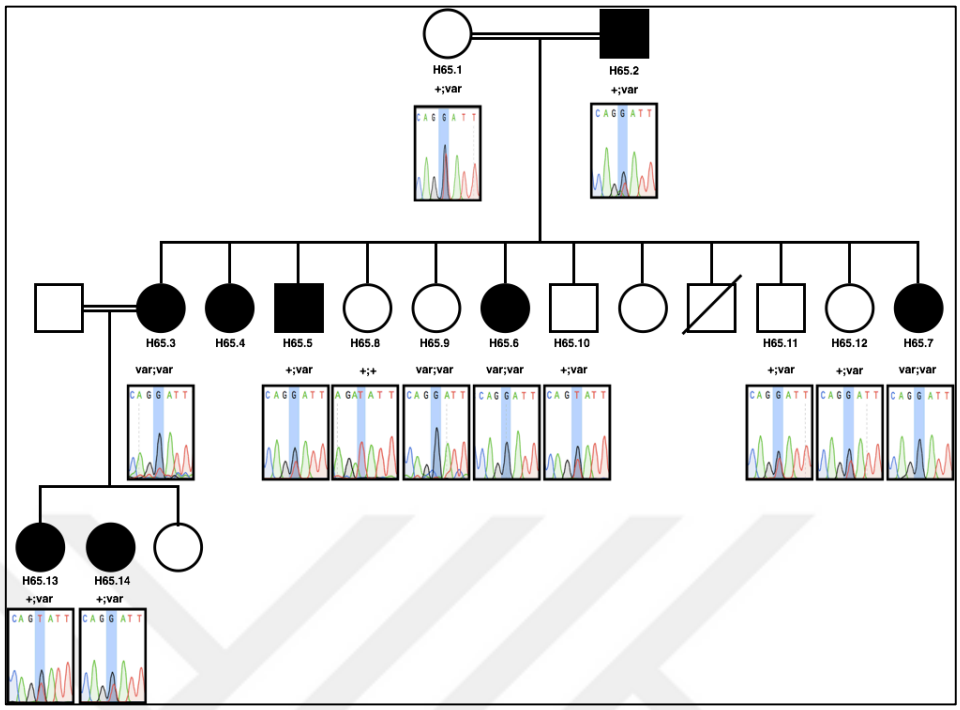


Figure A.79. H65 pedigree and chromatograms showing the c.334T>G variant in BRI3BP gene (+: native allele; var: variant allele).

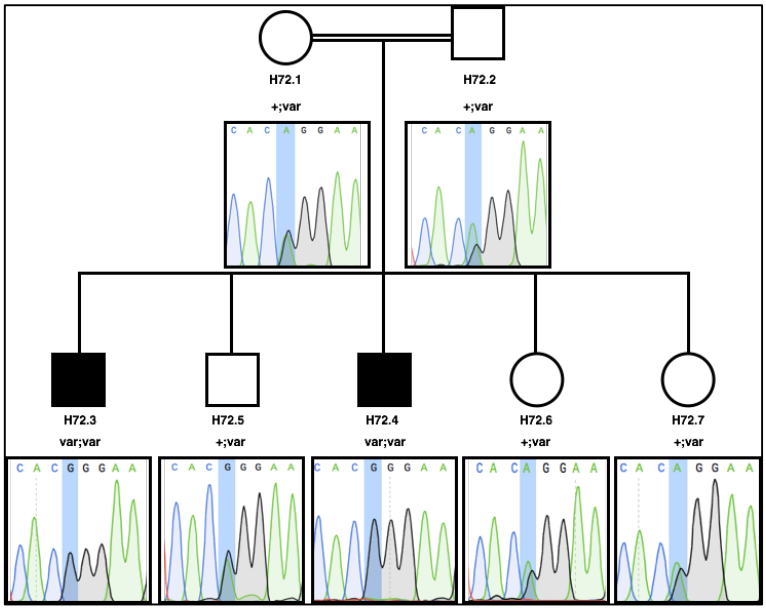


Figure A.80. H72 pedigree and chromatograms showing c.2104-2A>G variant in SPG7 gene (+: native allele, var: variant allele).

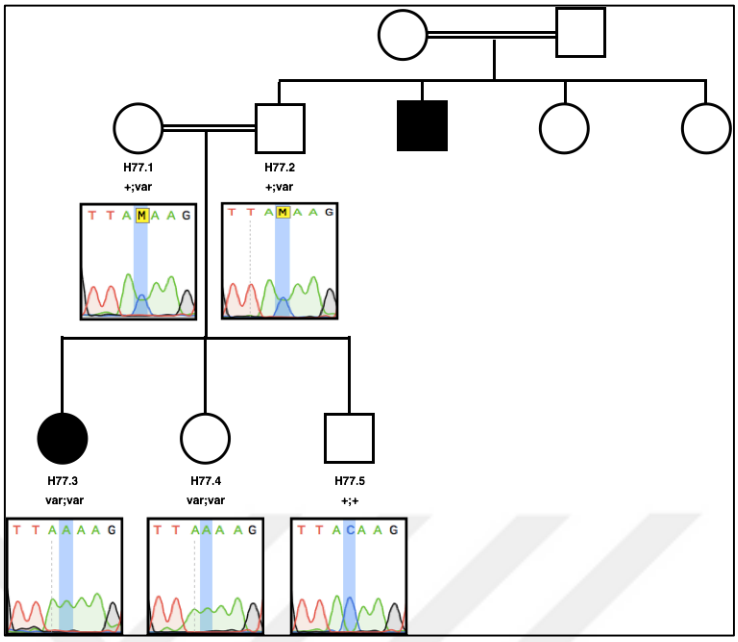


Figure A.81. H77 pedigree and chromatograms showing c.3036C>A (p.Tyr1012Ter) variant in SPG11 gene (+: native allele, var: variant allele).

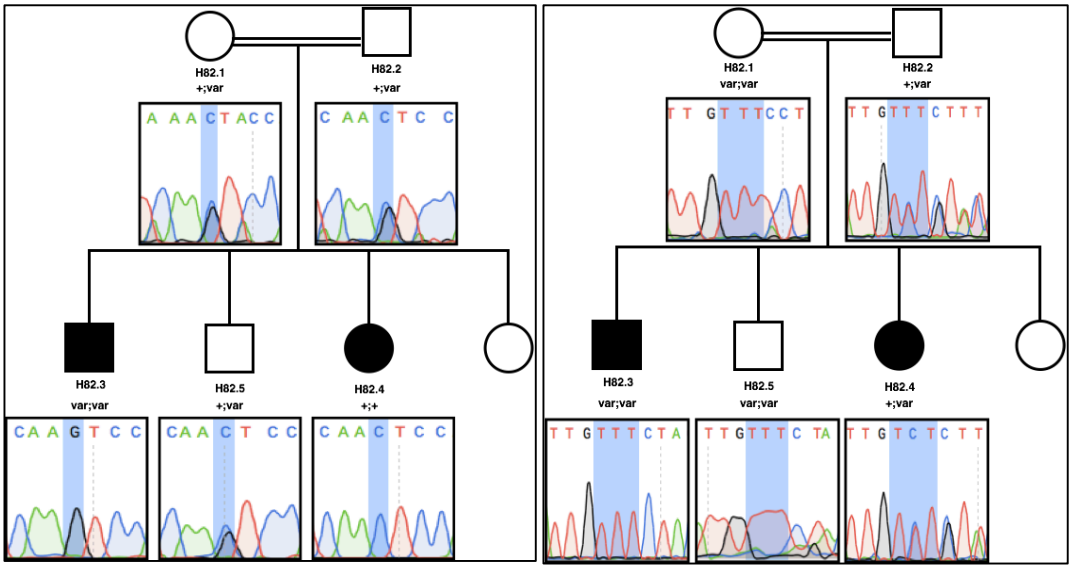


Figure A.82. H82 pedigree and chromatograms showing the c.7612C>G variant in POLQ gene (left panel) and the c.4515_4516delTC variant in 13th exon of ABCA10 gene (right panel) (+: native allele; var: variant allele).

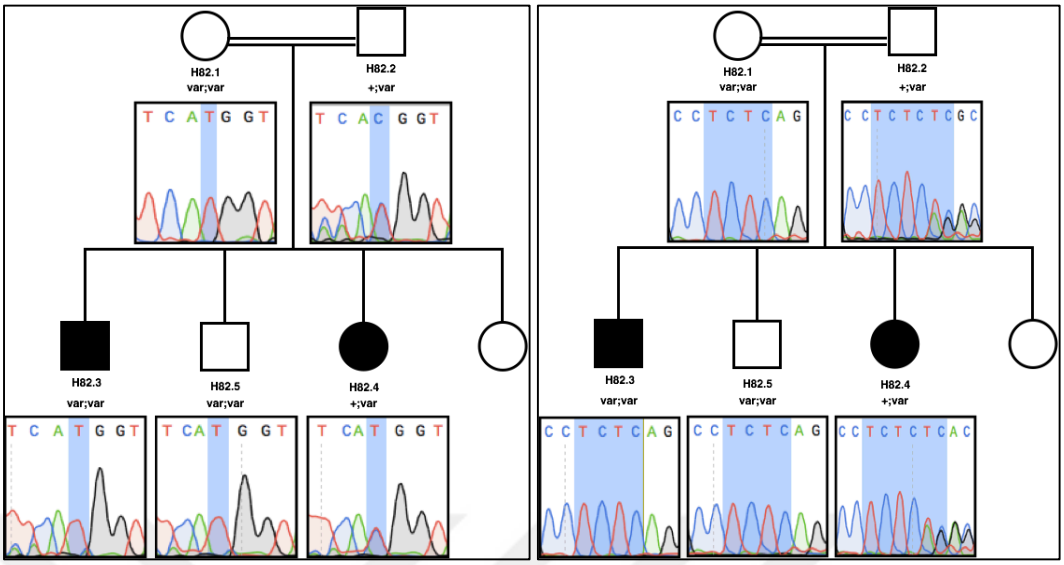


Figure A.83. H82 pedigree and chromatograms showing the c.3964C>T variant in 33rd exon of ABCA10 gene (left panel) and the c.1331_1334delCTGT variant in 39th exon of ABCA10 gene (right panel) (+: native allele; var: variant allele).

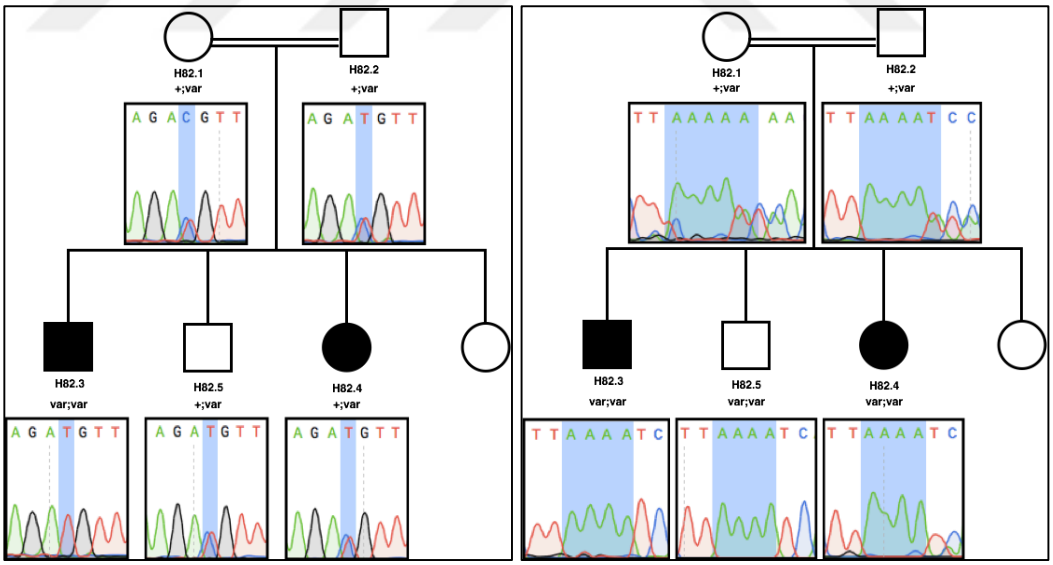


Figure A.84. H82 pedigree and chromatograms showing the c.290C>T variant in ZNF781 gene (left panel) and the c.313delA variant in 4th exon of ZNF880 gene (right panel) (+: native allele; var: variant allele).

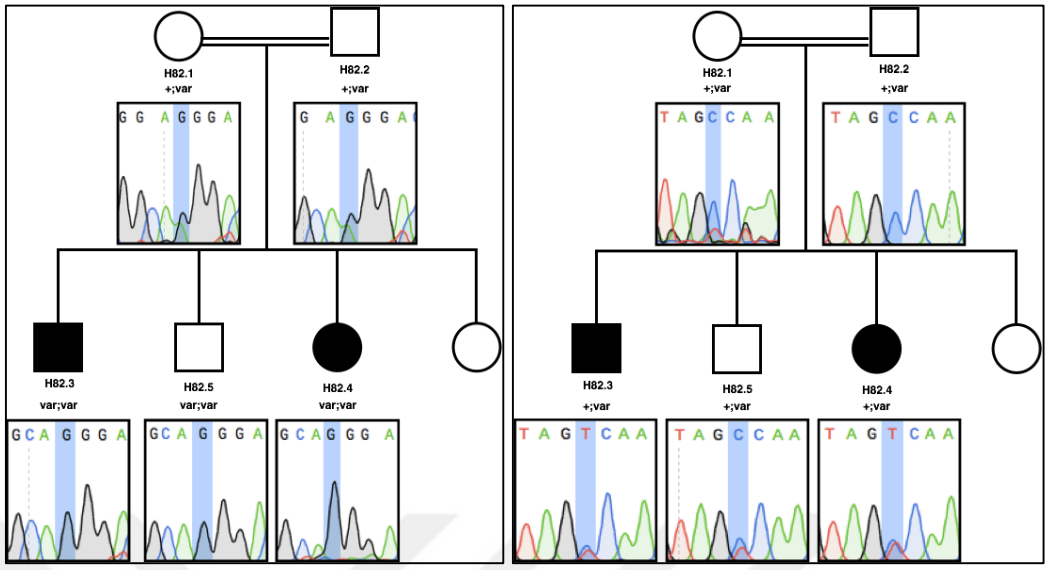


Figure A.85. H82 pedigree and chromatograms showing the c.1412A>G variant in 4th exon of ZNF880 gene (left panel) and the c.155C>T variant in VN1R4 gene (right panel) (+: native allele; var: variant allele).

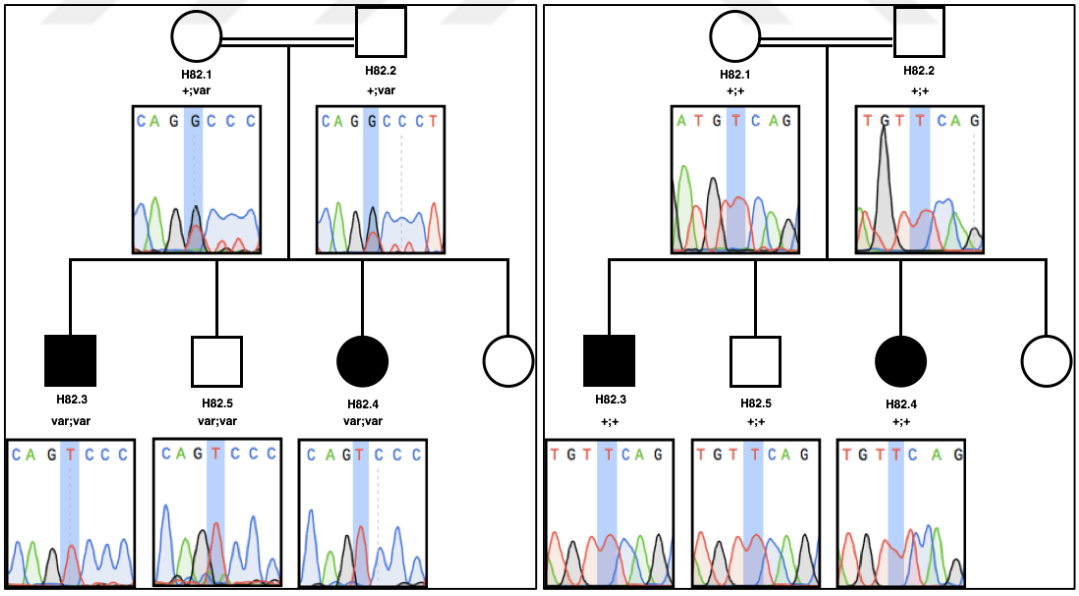


Figure A.86. H82 pedigree and chromatograms showing the c.1010G>T variant in 7th exon (left panel) the c.149T>A variant in 4th exon of ZSCAN5A gene (right panel) (+: native allele; var: variant allele).

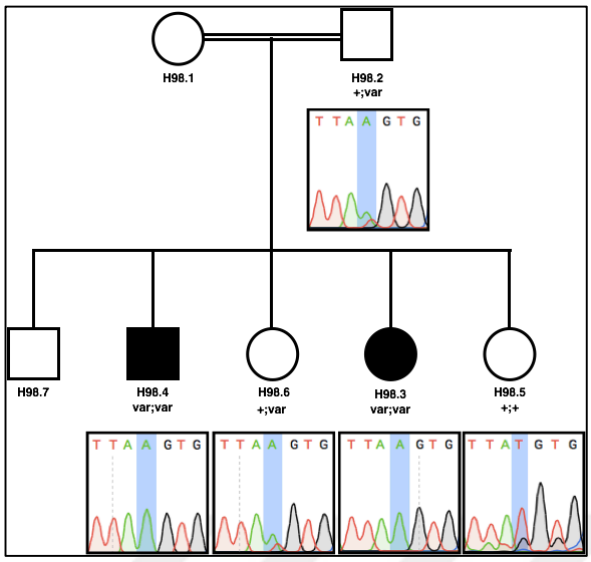


Figure A.87. H98 pegidree and chromatograms showing c.825T>A (p.Tyr275Ter) variant in CYP7B1 gene (+: native allele; var: variant allele).

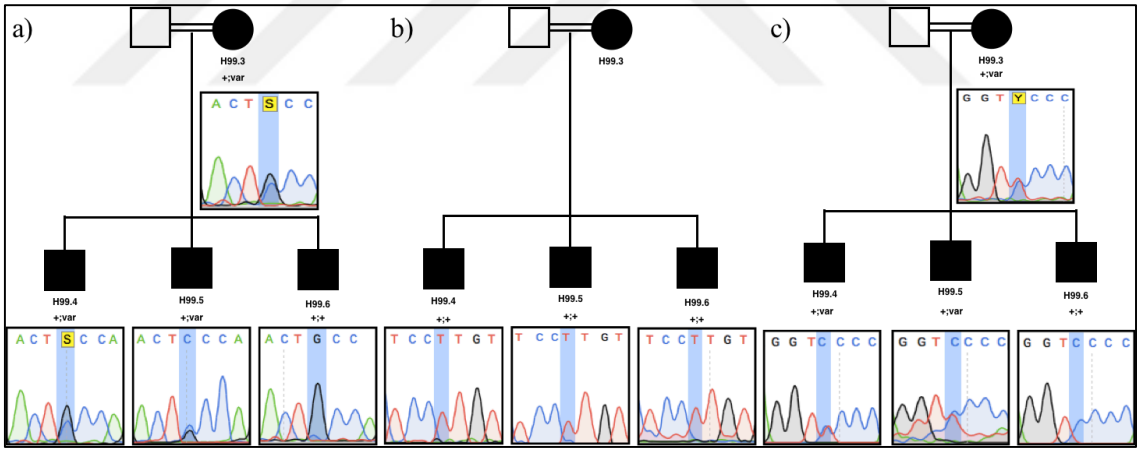


Figure A.88. H99 pedigrees and chromatograms showing a) c.7417-5G>C, b) c.5806T>A (p.Leu1936Met) and c) c.1844C>T (p.Ser615Phe) variants in ZFYVE26 gene (+: native allele; var: variant allele).

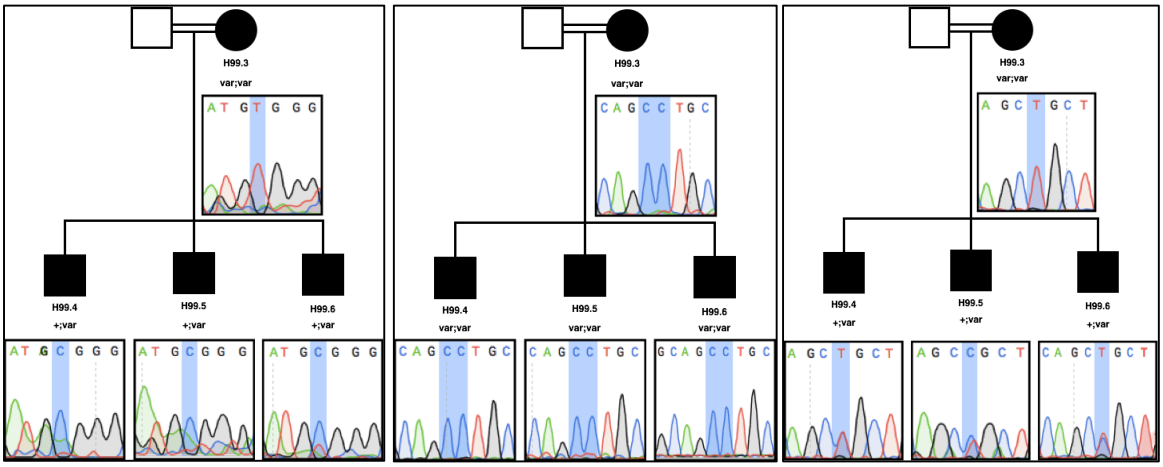


Figure A.89. H99 pedigree and chromatograms showing the c.4C>T variant FUCA1 gene (left panel), the c.9211_9212insC variant in SSPO gene (middle panel) and the c.247C>T variant in GIMAP7 gene (right panel) (+: native allele; var: variant allele).

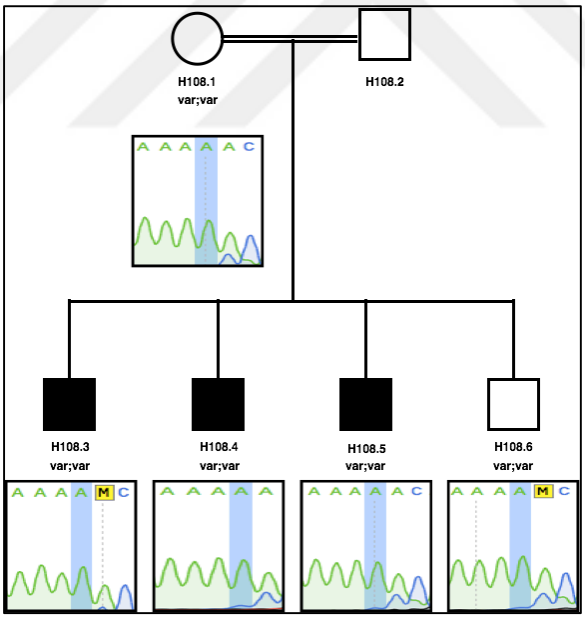


Figure A.90. H108 pedigree and chromatograms showing c.*1020C>A variant in C19orf12 gene (var: variant allele).

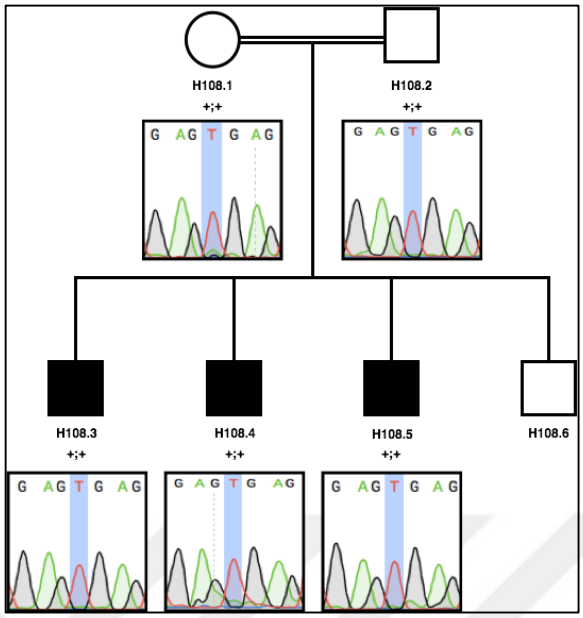


Figure A.91. H108 pedigree and chromatograms showing c.1343+6T>G variant in CPT1C gene (var: variant allele).

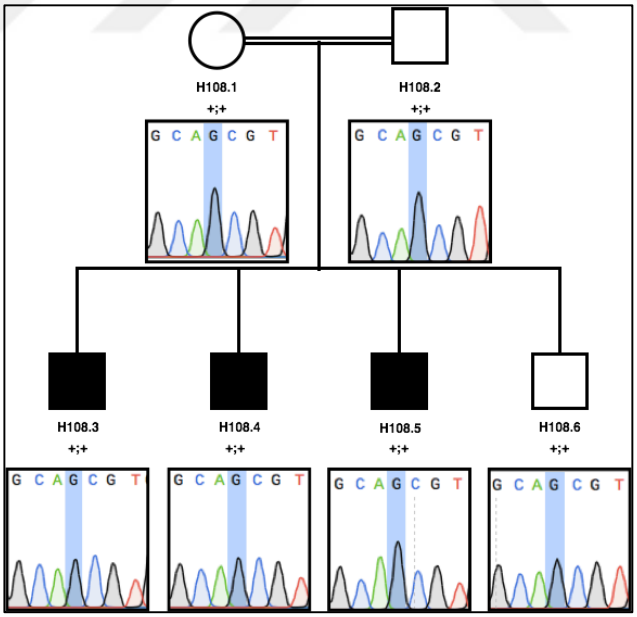


Figure A.92. H108 pedigree and chromatograms showing c.820G>A variant in ZFYVE27 (SPG33) gene (var: variant allele).

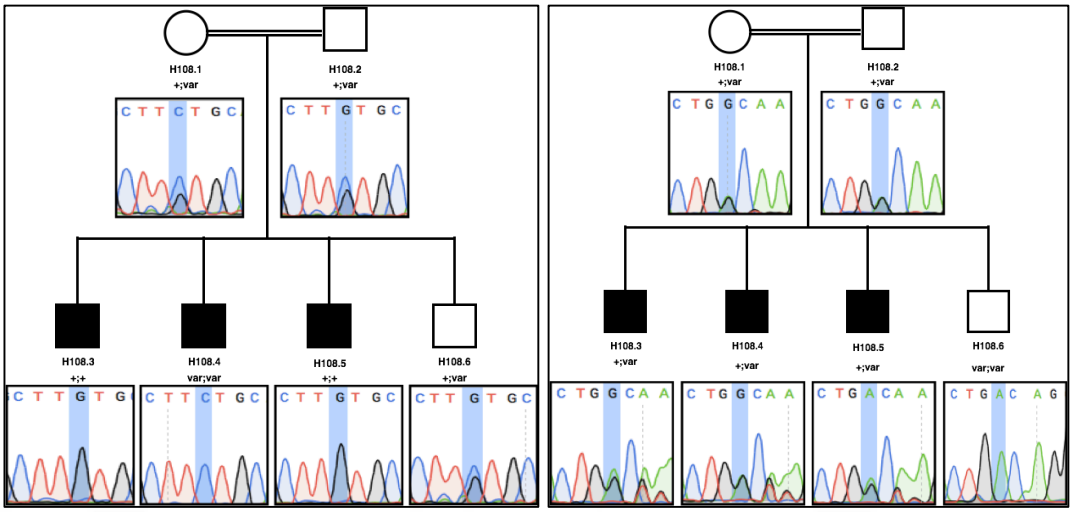


Figure A.93. H108 pedigree and chromatograms showing the c.1122G>C variant in SLC45A2 gene (left panel), and the c.524G>A variant in AMACR gene (right panel) (+: native allele; var: variant allele).

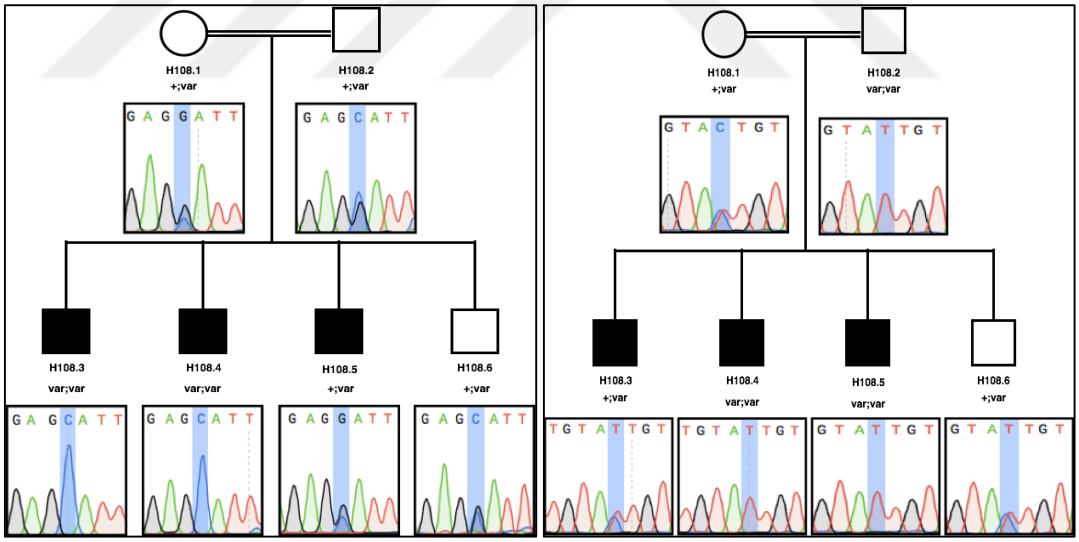


Figure A.94. H108 pedigree and chromatograms showing the c.1132G>C variant in NEBL gene (left panel), and the c.353C>T variant in C10orf53 gene (right panel) (+: native allele; var: variant allele).

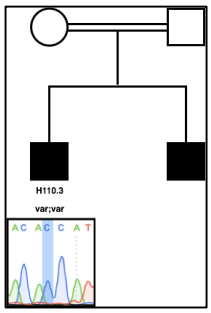


Figure A.95. H110 pedigree and chromatogram showing c.12166A>C variant in SACS gene (var: variant allele).

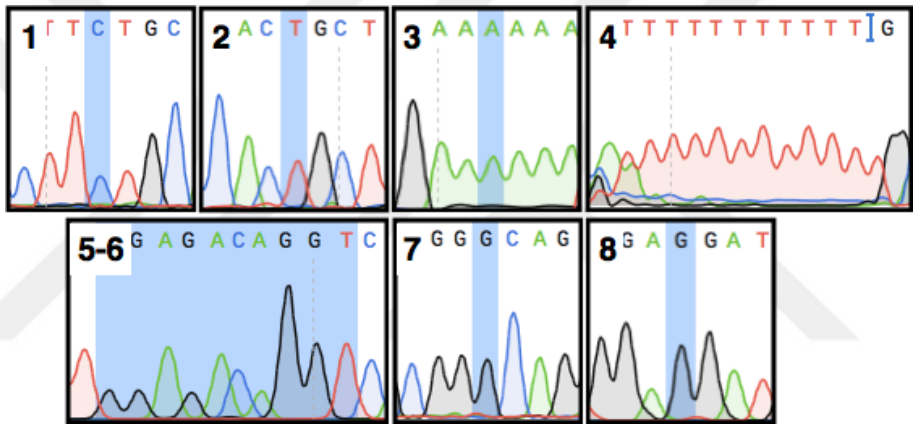


Figure A.96. Chromatograms showing the variants listed in Table 34.

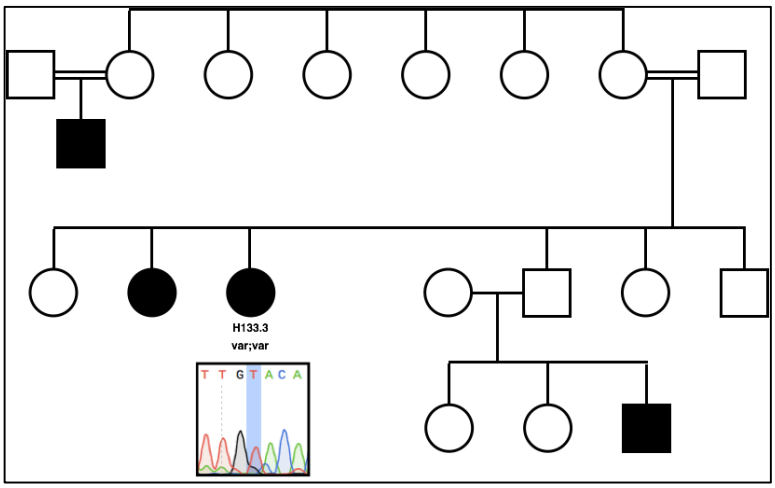


Figure A.97. H133 pedigree and chromatogram showing c.437A>T variant in CCT5 gene (var: variant allele).

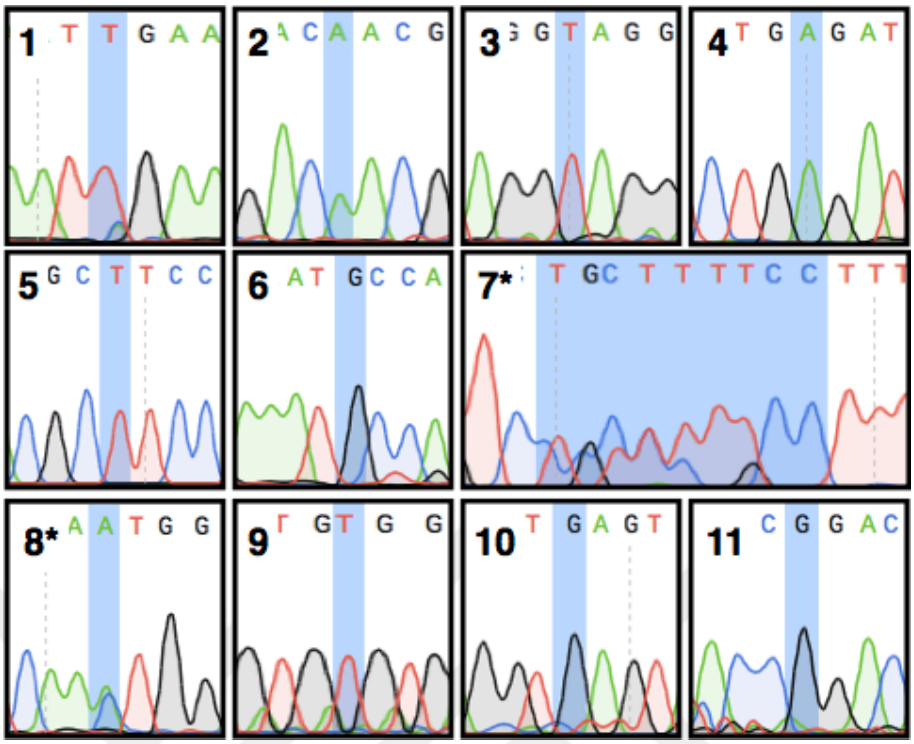


Figure A.98. Chromatograms showing the variants listed in Table 35.

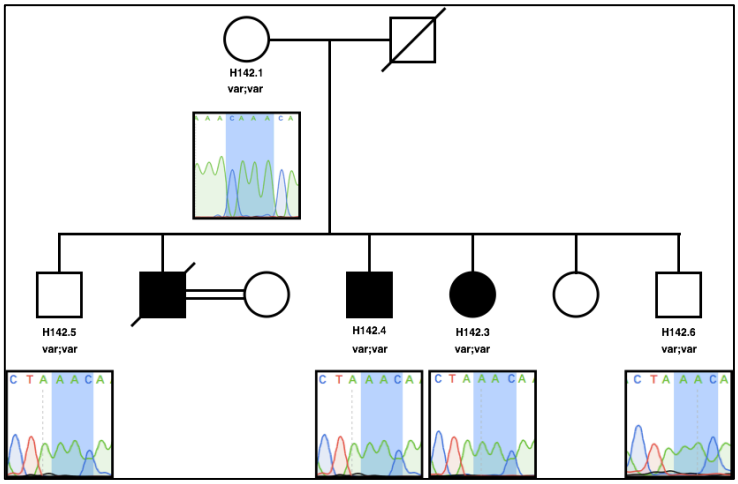


Figure A.99. H142 pedigree and chromatograms showing n.450_451insAAC insertion in C14orf23 gene (var: variant allele).

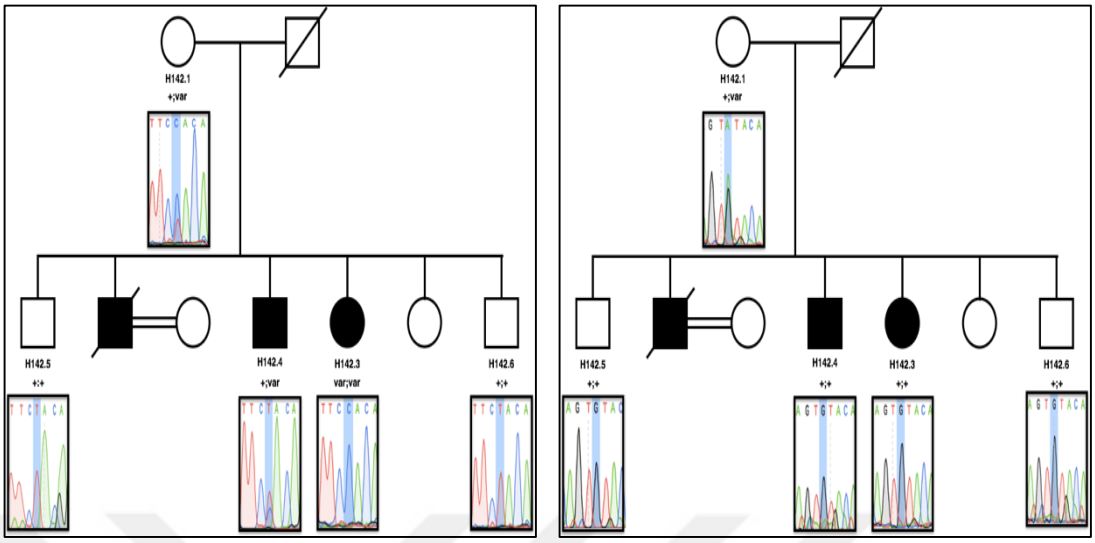


Figure A.100. H142 pedigree and chromatograms showing c.103T>C variant in SDHAF2 gene (left panel) and n.1482G>A variant in SDHAP2 gene (right panel) (+: native allele, var: variant allele).

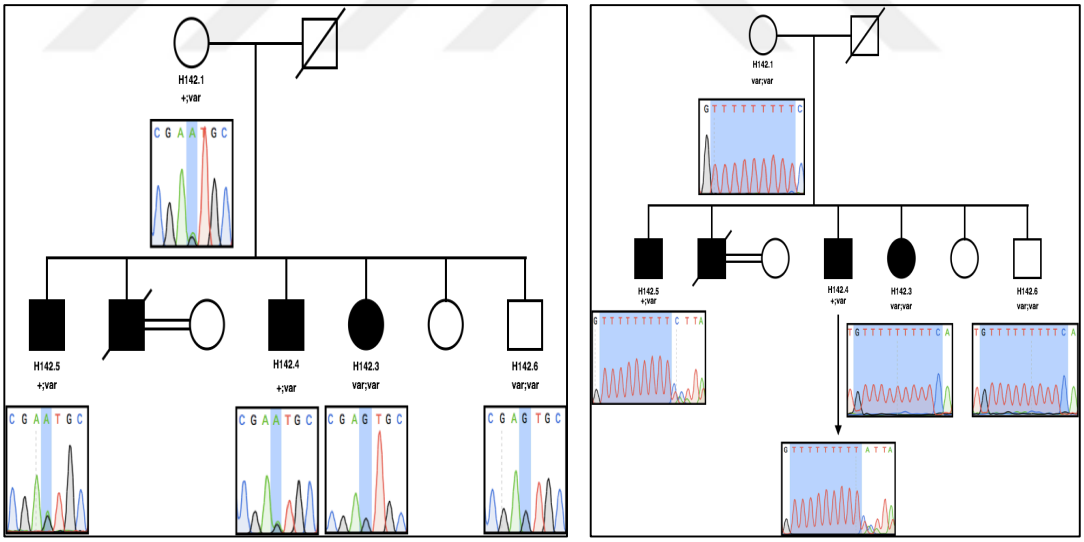


Figure A.101. H142 pedigree and chromatograms showing the c.1420A>G variant in GPATCH2 gene (left panel), and the c.4897-9_4897-8delTT variant in DNAH7 gene (right panel) (+: native allele; var: variant allele).

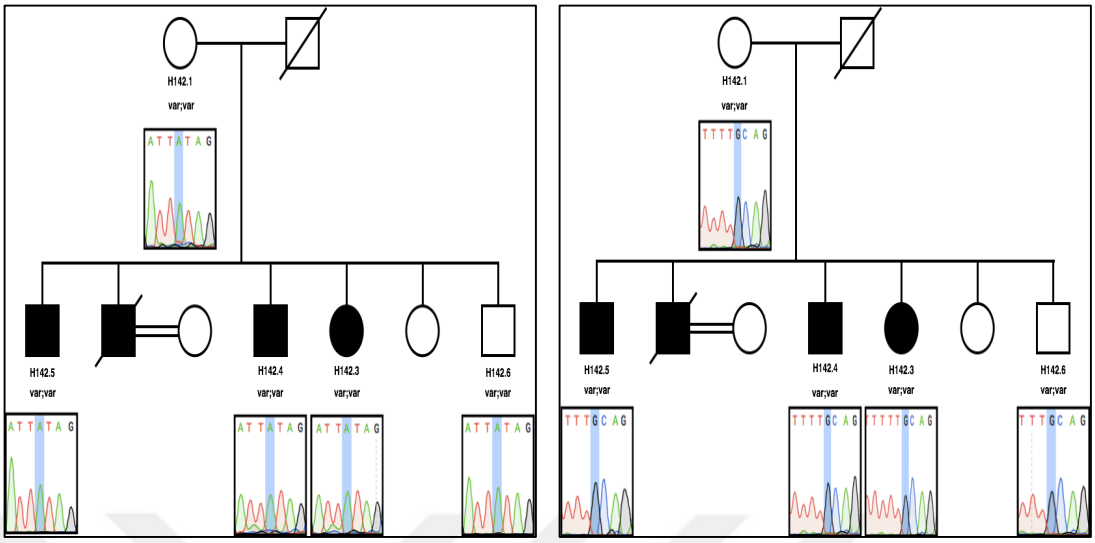


Figure A.102. H142 pedigree and chromatograms showing the c.1999G>A variant in PLCL1 gene (left panel), and the c.169-5T>G variant in DGKB gene (right panel) (+: native allele; var: variant allele).

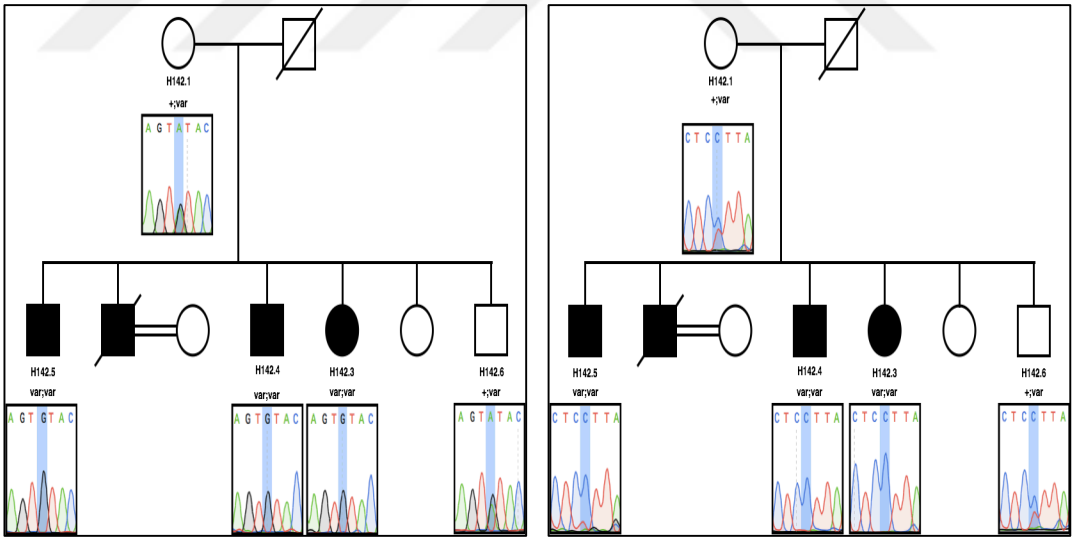


Figure A.103. H142 pedigree and chromatograms showing the c.1286A>G variant in SEMA3C gene (left panel), and the c.742-7T>C variant in KIAA1324L gene (right panel) (+: native allele; var: variant allele).

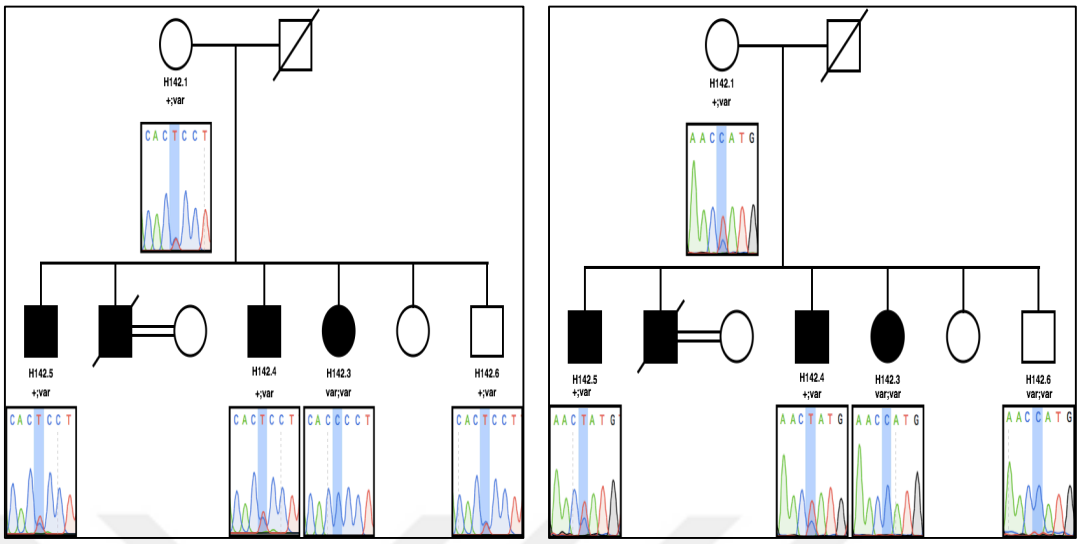


Figure A.104. H142 pedigree and chromatograms showing the c.812-8T>C variant in PRPF4 gene (left panel), and the c.820T>C variant in UBQLNL gene (right panel) (+: native allele; var: variant allele).

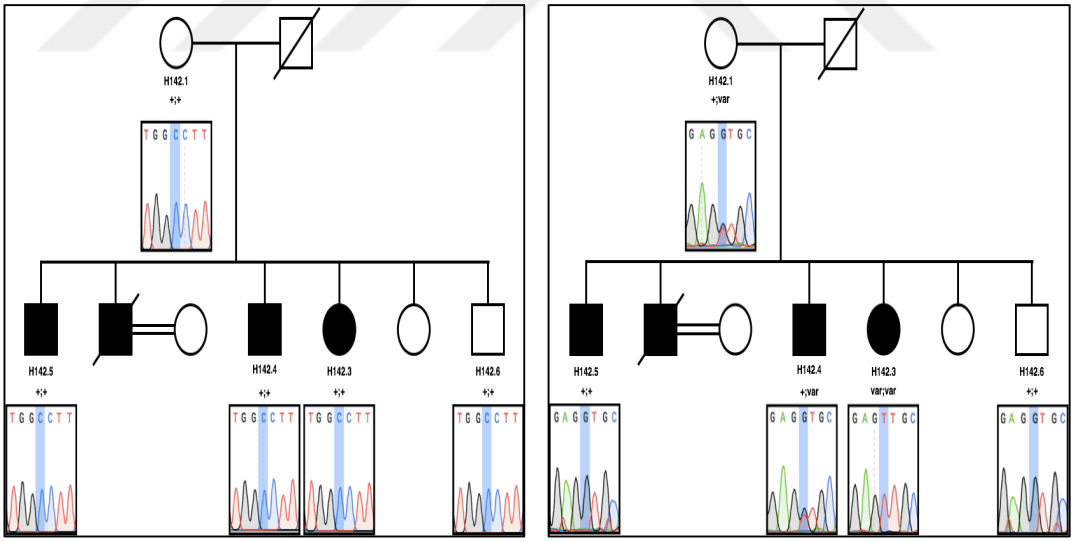


Figure A.105. H142 pedigree and chromatograms showing the c.887dupC variant in UNC93B1 gene (left panel), and the c.290G>T variant in C11orf24 gene (right panel) (+: native allele; var: variant allele).

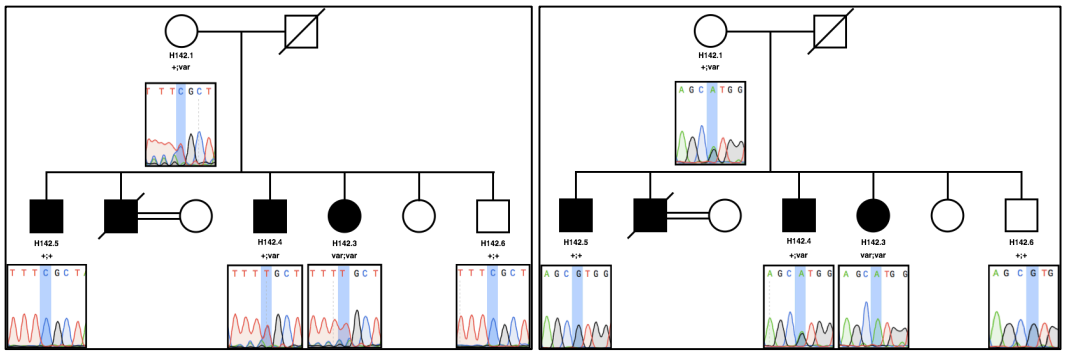


Figure A.106. H142 pedigree and chromatograms showing the c.665C>T variant in AQP11 gene (left panel), and the c.706G>A variant in USP35 gene (right panel) (+: native allele; var: variant allele).

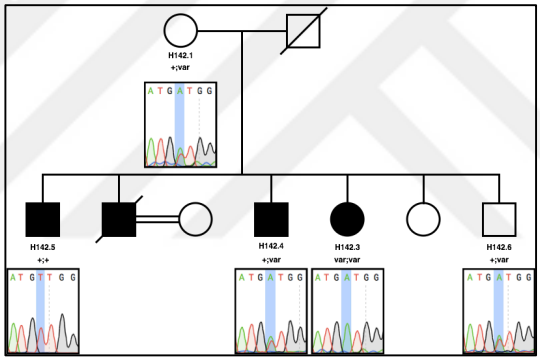


Figure A.107. H142 pedigree and chromatograms showing the c.*15+6T>A variant in TCL1B gene (+: native allele; var: variant allele).

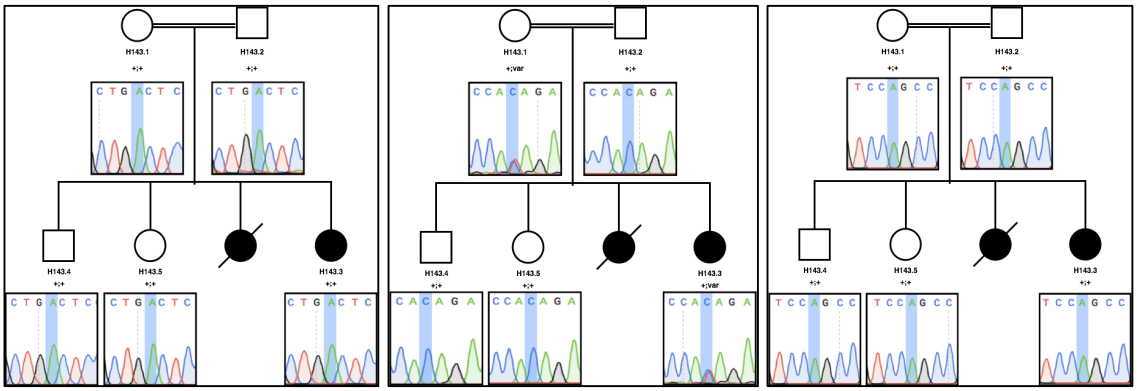


Figure A.108. H143 pedigree and chromatograms showing the c.767A>C, c.883-3C>T and c.1261A>C variants in KIF1A gene (+: native allele; var: variant allele).

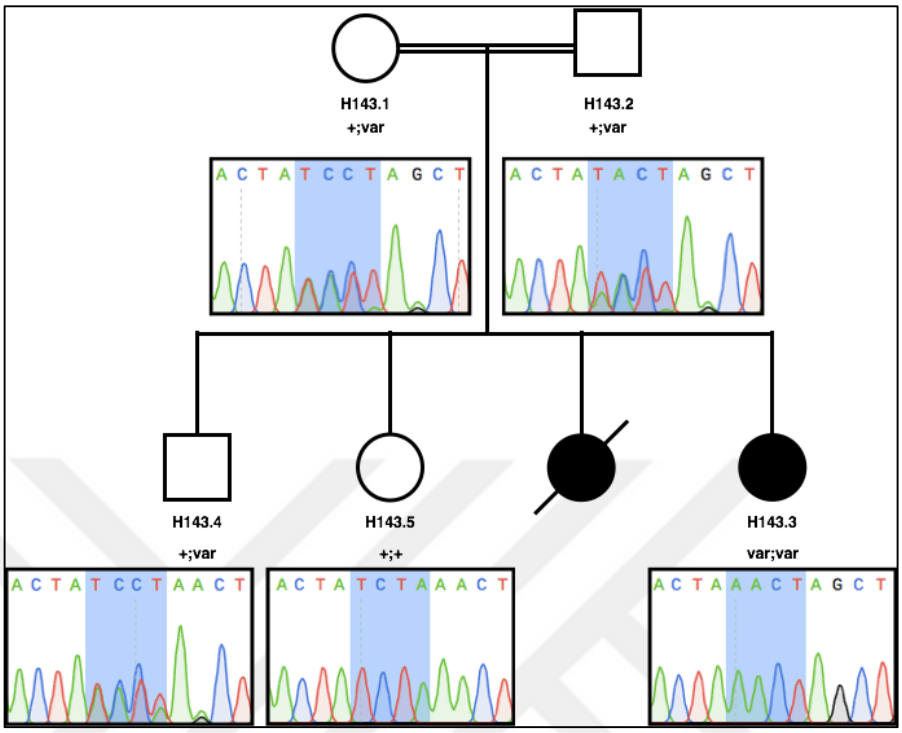


Figure A.109. H143 pedigree and chromatograms showing the c.3904_3907delTCTA variant in SPG11 gene (+: native allele; var: variant allele).

APPENDIX B: 3D PROTEIN MODELS

SWISS-MODEL, RaptorX, Pyhre2 and iTasser algorithms were used to obtain 3D of model prediction for the variants identified in this study.

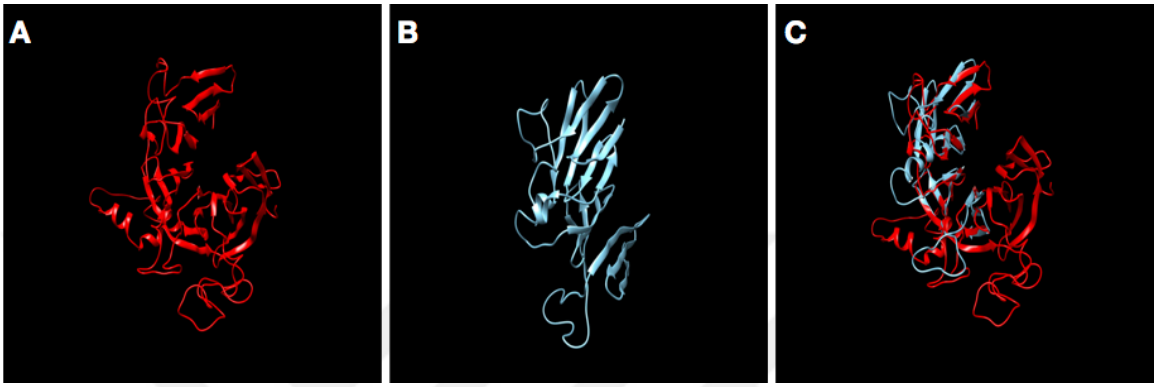


Figure B.1. 3D models for SPG11 protein predicted by SWISS-MODEL algorithm. Native protein (left panel), protein with variant identified in P463 (middle panel) and superimposition of models (right panel) (RMSD:9.444).

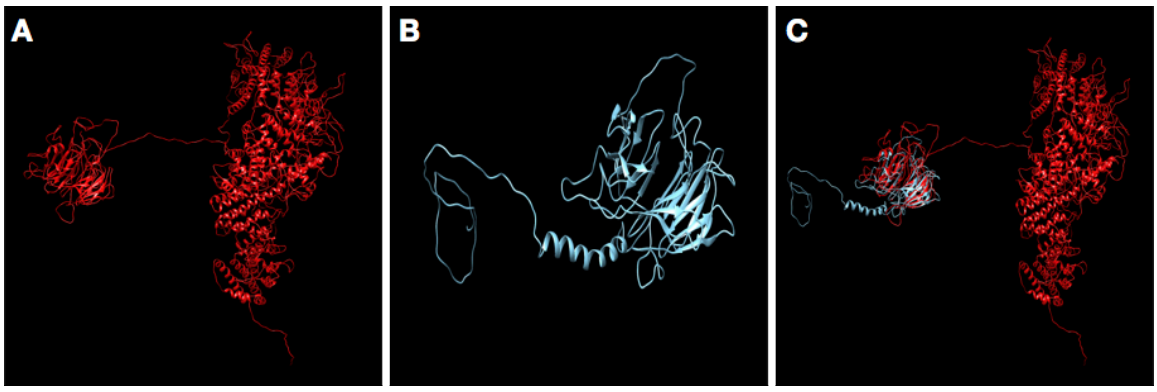


Figure B.2. 3D models for SPG11 protein predicted by RaptorX algorithm. Native protein (left panel), protein with variant identified in P463 (middle panel) and superimposition of models (right panel) (RMSD: 34.872).

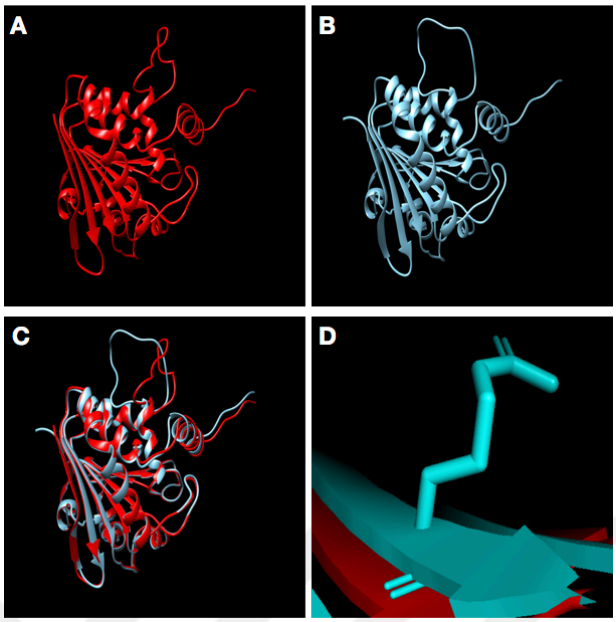


Figure B.3. 3D models for SACS protein predicted by SWISS-MODEL algorithm. Native protein (A), protein with variant identified in P627 (B) and superimposition of models (C), 213th position in SACS protein (B) (RMSD: 0.522).

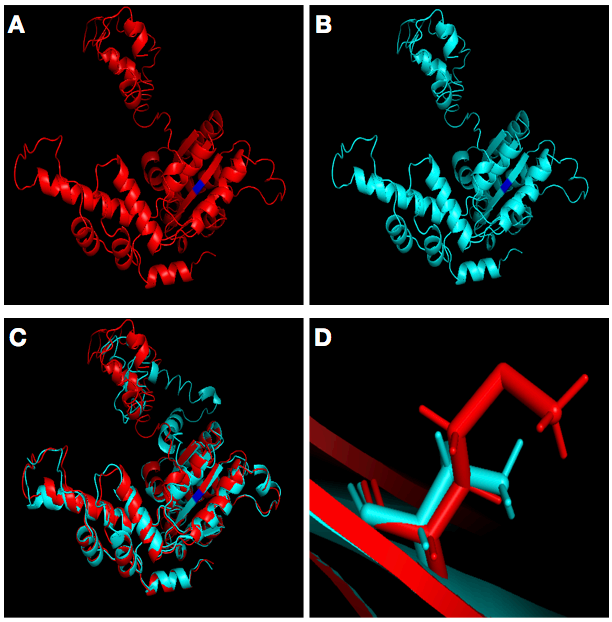


Figure B.4. 3D models for ATAD1 protein predicted by iTASSER algorithm. Native protein (A), protein with variant identified in H28 (B) and superimposition of models (C), 234th position in ATAD1 protein (D). (RMSD: 0.948).

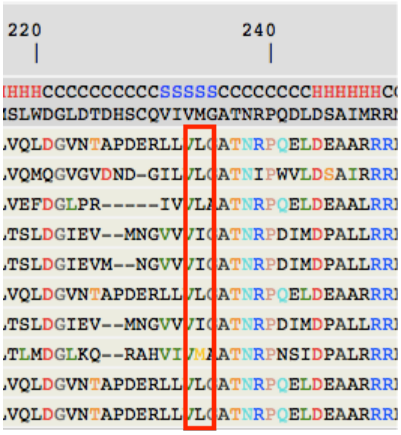


Figure B.5. Conservation of hydrophobic amino acids in the very same position of p.234Met in ATAD1 gene.

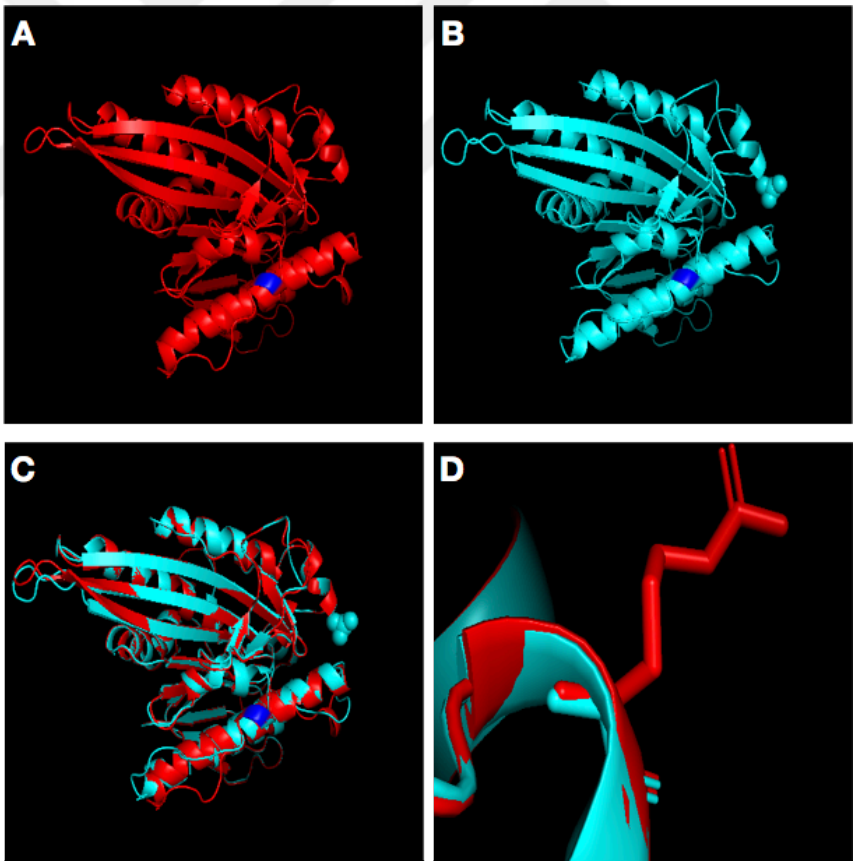


Figure B.6. 3D models for KIF1C protein predicted by SWISS-MODEL algorithm. Native protein (A), protein with variant identified in H52 (B) and superimposition of models (C), 301st position in KIF1C protein (D). (RMSD: 0.195).

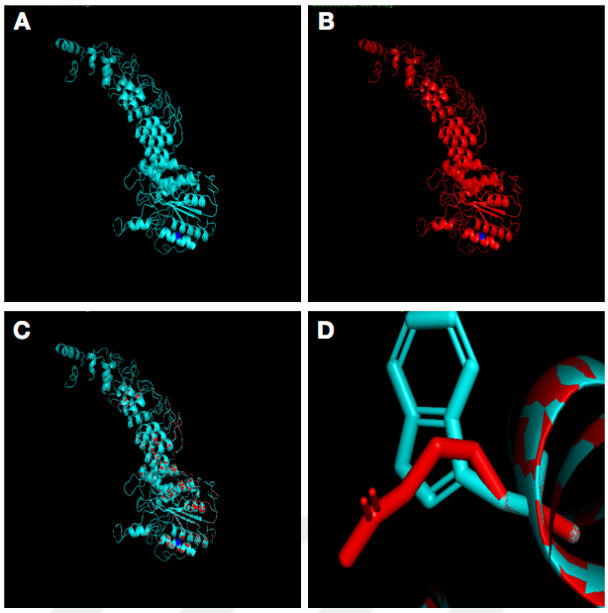


Figure B.7. 3D models for PLA2G6 protein predicted by SWISS-MODEL algorithm. Native protein (A), protein with variant identified in H53 (B) and superimposition of models (C), 747th position in PLA2G6 protein (D). (RMSD: 0.0).

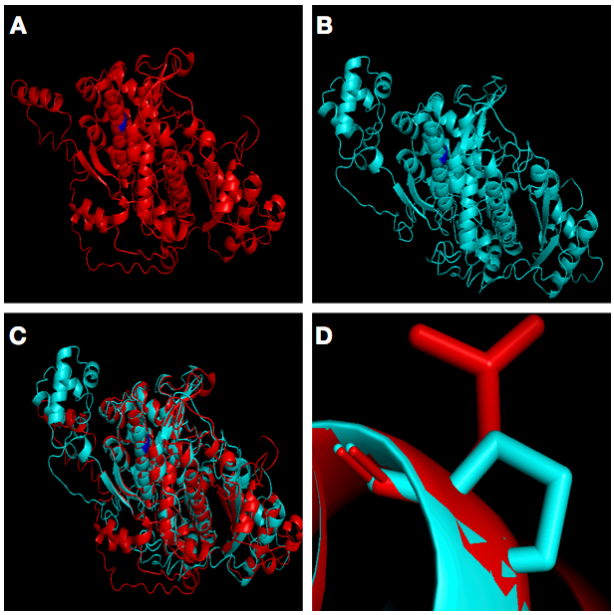


Figure B.8. 3D models for SAMHD1 protein predicted by iTASSER algorithm. Native protein (A), protein with variant identified in H57 (B) and superimposition of models (C), 178th position in SAMHD1 protein (D) (RMSD:2.123).

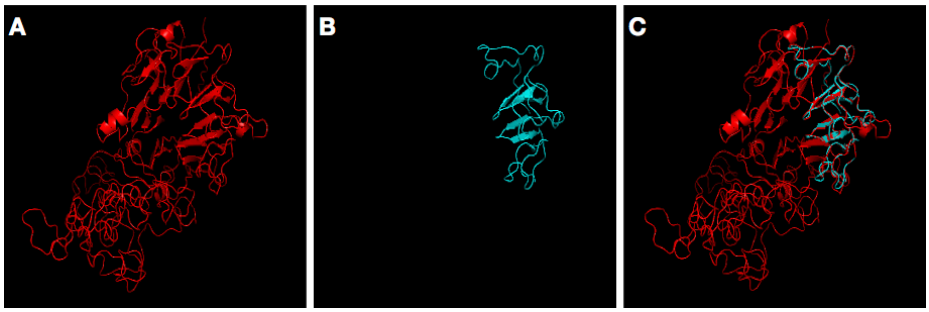


Figure B.9. 3D models for ALS2 protein predicted by SWISS-MODEL algorithm. Native protein (A), protein with variant identified in H59 (B) and superimposition of models (C). (RMSD: 1.136).

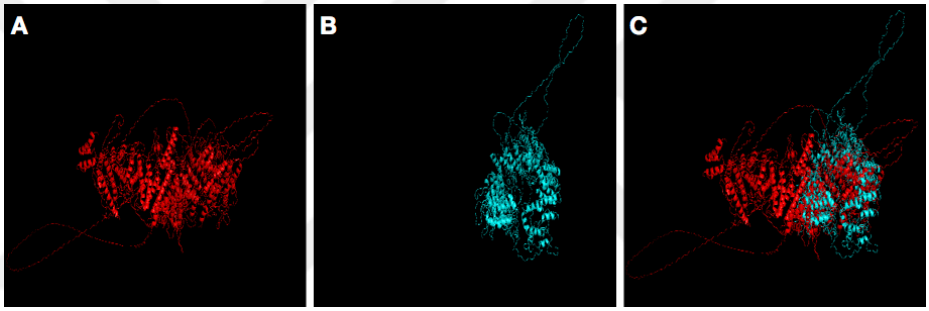


Figure B.10. 3D models for SPG15 protein predicted by Raptor-X algorithm. Native protein (A), protein with variant identified in H61 (B) and superimposition of models (C). (RMSD: 27,557).

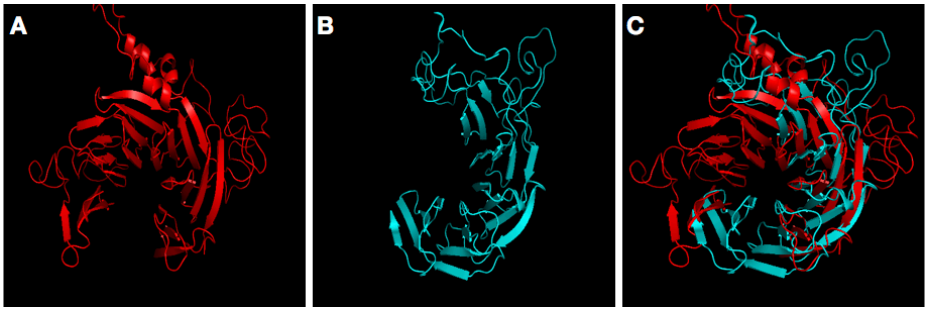


Figure B.11. 3D models for SPG11 protein predicted by SWISS-MODEL algorithm. Native protein (A), protein with variant identified in H77 (B) and superimposition of models (C). (RMSD: 12,019).

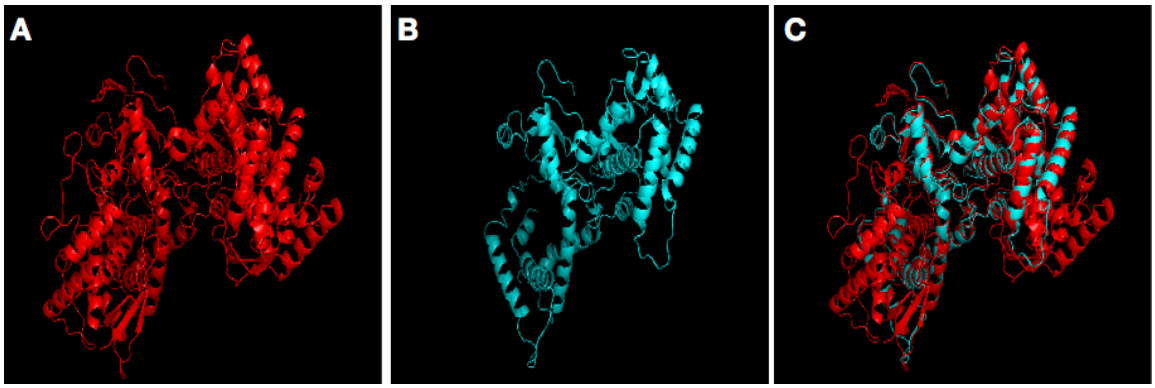


Figure B.12. 3D models for CYP7B1 protein predicted by SWISS-MODEL algorithm. Native protein (A), protein with variant identified in H98 (B) and superimposition of models (C). (RMSD: 1,104).

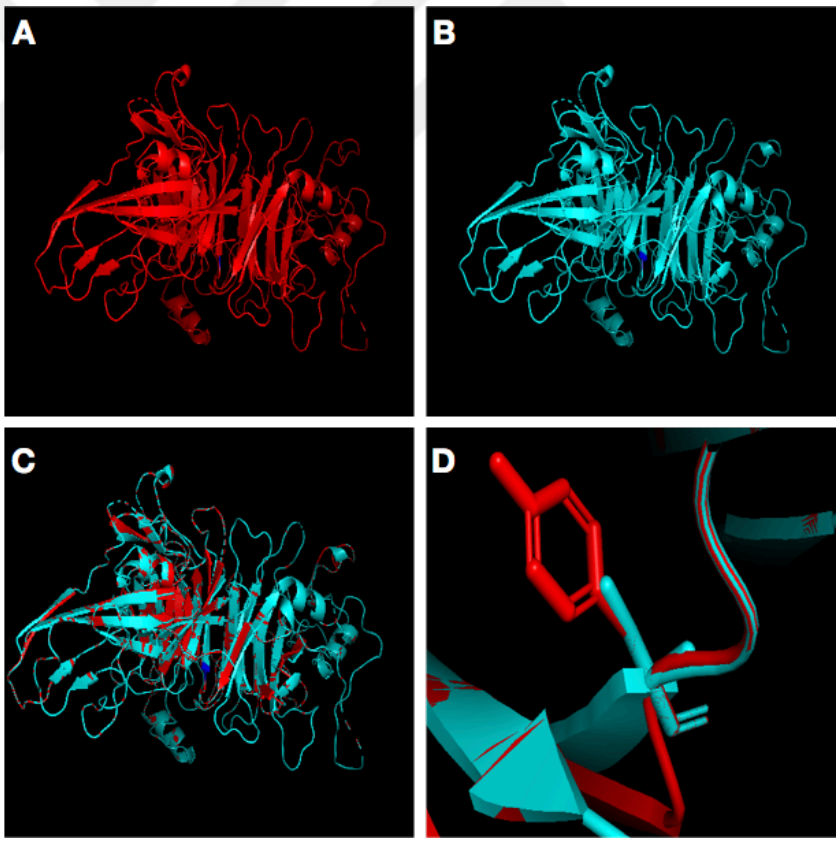


Figure B.13. 3D models for SEMA3C protein predicted by SWISS-MODEL algorithm. Native protein (A), protein with variant identified in H142 (B) and superimposition of models (C), 429th position in SEMA3C protein (D) (RMSD:0.052).

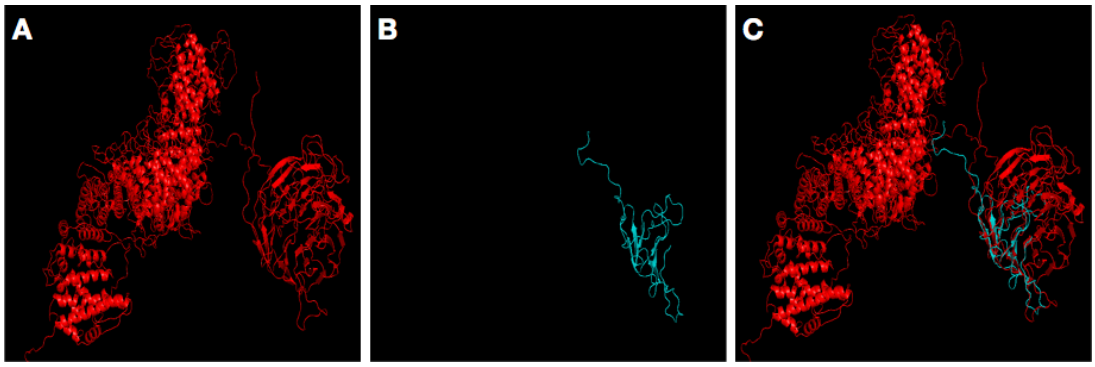


Figure B.14. 3D models for SPG11 protein predicted by Raptor-X algorithm. Native protein (A), protein with variant identified in H143 (B) and superimposition of models (C). (RMSD: 4,224).

APPENDIX C: RESULTS OF TECHNICAL REPLICATES

Experimental replicates were performed to ensure the validity of the results.

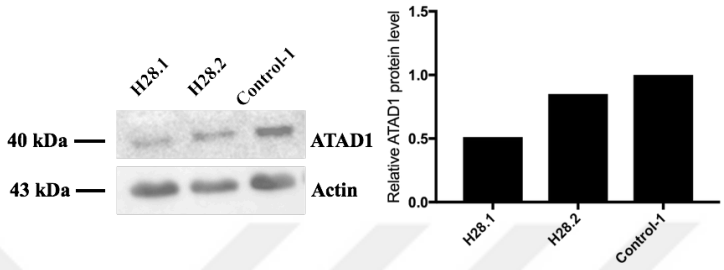


Figure C.1. Western blot analyses of ATAD1 protein performed for H28.1 and H28.2 (Replicate #2).

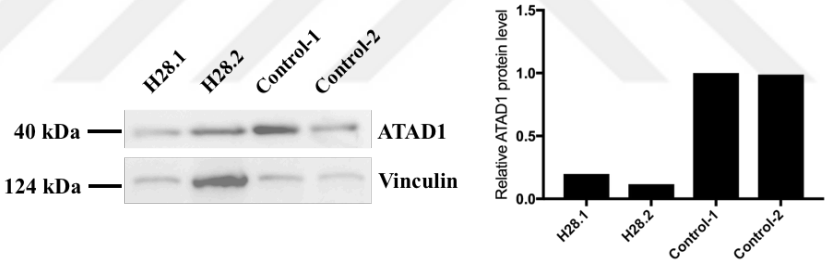


Figure C.2. Western blot analyses of ATAD1 protein performed for H28.1 and H28.2 (Replicate #3).

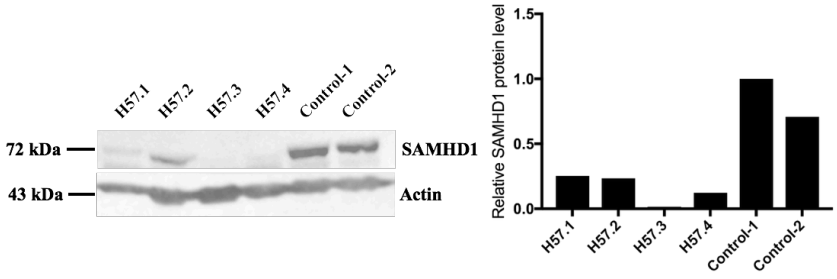


Figure C.3. Western blot analyses of CYP27A1 protein performed for family H55 (Replicate #2).

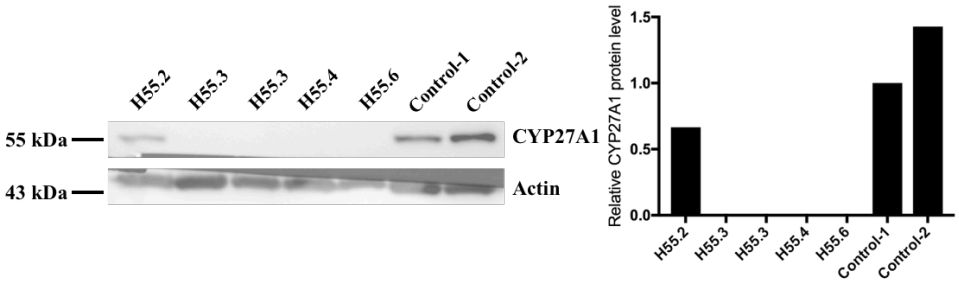


Figure C.4. Western blot analyses of SAMHD1 protein performed for family H57 (Replicate #2).

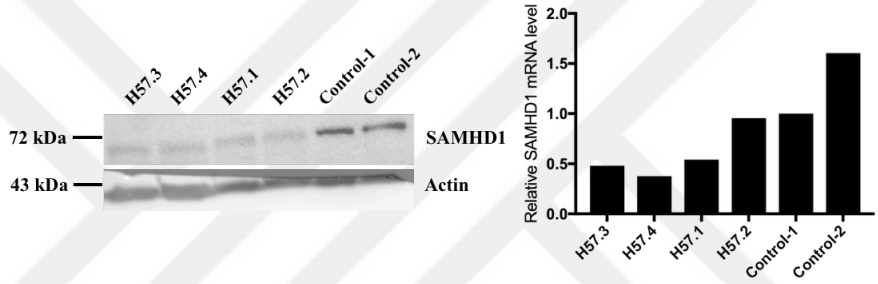


Figure C.5. Western blot analyses of SAMHD1 protein performed for family H57 (Replicate #3).