GENE HUNT IN FOUR INHERITED DISEASES

by

Murat Çetinkaya

M.D., Marmara University School of Medicine, 2004

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

2010

GENE HUNT IN FOUR INHERITED DISEASES

APPROVED BY:

Prof. Aslıhan Tolun (Thesis Supervisor)	.....
Prof. A. Nazlı Başak	.....
Prof. Nur Buyru	.....
Prof. S. Hande Çağlayan	....
Prof. Ahmet Gül	.....

DATE OF APPROVAL: 06.07.2010

To my parents

ACKNOWLEDGMENTS

I would like to express my deepest respect and gratitude to my thesis advisor Prof. Aslıhan Tolun for her continuous guidance, extensive encouragement and valuable criticisms in every step of this project.

I would like to extend my thanks to Prof. A. Nazlı Başak, Prof. Nur Buyru, Prof. S. Hande Çağlayan and Prof. Ahmet Gül for devoting their time to evaluate this work.

I would like to acknowledge the microsatellite genome scan performed by NHLBI Mammalian Genotyping Service. This work was supported by TÜBİTAK (105T306 and 108S011), Boğaziçi University Scientific Research Fund (5186) and TÜBA.

I would like to express my gratitude to Dr. Hatice Karasoy, Dr Bülent Kara and Dr. Beyhan Tüysüz for providing patient samples and clinical evaluations. I thank all the families that participated in the study.

I would like to express my sincere gratitude to Dr. Sibel Uğur İşeri for sharing her experience during my graduate years, without any complaints. I also convey my appreciation to all present and previous colleagues for providing a productive and joyful environment. Nothing would have been possible without them. Thank you.

I wish to thank all of academic staff whose expertise and guidance added greatly to my graduate experience.

I would like to express my special thanks to Caroline Pirkevi, for her endless support, patience and love. Nothing would have been possible without her, thank you.

Last but not least, I am grateful to my parents. This work would not have been possible without their love, patience and encouragement.

ABSTRACT

GENE HUNT IN FOUR INHERITED DISEASES

Genetic linkage analysis is applied to identify loci that harbor gene or genes associated with a disease. It is possible to map disease loci by observing alleles the segregation of microsatellite or SNP markers with the disease in families, it is possible to map disease loci. Availability of densely spaced microsatellite marker maps and genome-wide SNP scans have made such studies feasible. In the framework of this thesis, genome-wide linkage data were evaluated with parametric linkage analysis software and haplotype segregation studies for four inherited disorders. Furthermore, candidate loci identified were further investigated by fine-mapping and candidate gene approach.

Autosomal Recessive Ataxia (ARA) is a subgroup of hereditary ataxias, and characterized by slowly progressive impaired coordination of gait, hands, eye movements and speech. Nine patients from two consanguineous families were included in the study. We found that *Aprataxin* gene was mutated in six patients. Three patients were homozygous for a known nonsense mutation, and three patients for a novel missense mutation. For the three remaining patients clinical reevaluation revealed that they were afflicted with another disease, but no responsible locus could be identified.

Autosomal recessive Larsen Syndrome (LRS) is characterized by congenital large-joint dislocations and craniofacial abnormalities. Seven consanguineous families were analyzed. In four families disease was mapped to 17q25.3 and homozygous mutations were found in *Calcium Activated Nucleotidase 1*. Three patients had the same known missense mutation. In one family, a novel missense mutation was identified. No common candidate region could be defined for the remaining three families.

Juvenile Parkinsonism (JP) is characterized by classical triad of parkinsonism signs, bradykinesia, rigidity and resting tremor and disease onset before the age of 40 years. A missense variant in *AK3L1* was identified in homozygous state in all patients but also in a

healthy individual in the family. This variant was not detected in any of the 130 control individuals. We propose that an undetermined second locus either contributes to the disease manifestation in patients or protects the homozygous healthy individual.

Lastly, autosomal recessive Congenital Cerebellar Hypoplasia (CCLH), characterized by nonprogressive congenital cerebellar ataxia with mental retardation, delayed motor development, disturbed coordination, hypotonia and cerebellar Hypoplasia was studied in two families. In one of the families we mapped the disease to 6q16.1-22.1. No candidate gene could be defined in the region. In the other family, several candidate loci at other chromosomal regions were found.

ÖZET

DÖRT KALITSAL HASTALIKTA GEN ARAMA

Genetik bağlantı analizi, belirli bir hastalıkla ilintili gen veya genleri saptamak için kullanılır. Hastalık bölgesinin belirlenmesi, mikrosatellit veya SNP belirteçleri alellerinin genotiplendirilmesi ve bunların ailedeki aktarımının incelenmesiyle mümkündür. Yoğun mikrosatellit belirteç haritaları ve genom çapında SNP taramalarının ulaşılabilirliği bu tarz çalışmaların önünü açmıştır. Bu çalışma çerçevesinde, dört kalıtsal hastalıkta genom tarama verileri parametrik bağlantı programları ve haplotip incelemeleri ile değerlendirilerek belirlenen anlamlı bölgeler ayrıntılı olarak haritalandırılmış ve aday genler incelemiştir.

Otozomal Resesif Ataksi (ARA) kalıtımsal ataksilerin bir alt grubudur ve yavaş ilerleyen yürüyüş, el, göz hareketlerinde koordinasyonsuzluk ve konuşma bozukluğu ile tanımlanmaktadır. Çalışmaya iki ailede tümü akraba evliliğine doğmuş dokuz hasta dahil edilmiştir. Altı hastada *Aprataxin* geninde mutasyon bulunmuştur. Üç bilinen, üç hastada ise literatürde ilk kez tanımlanan missens mutasyonlar bulunmuştur. Diğer üç hastanın kliniklerinin tekrar değerlendirmesi, bu kişilerin hastalıklarının farklı olduğunu göstermiştir, ama bir gen bölgesi bulunamamıştır.

Otozomal resesif Larsen Sendromu (LRS) büyük eklemlerdeki doğuştan çıkıklar ve yüz anomalileri ile tanımlanmaktadır. Çalışmada akraba evliliği yapmış olan yedi aile incelenmiştir. Dört ailede hastalık 17q25.3'e haritalandırılmış ve *Calcium Activated Nucleotidase 1* geninde homozigot mutasyonlar bulunmuştur. Üç hastada bilinen bir mutasyon, bir ailede ilk kez tanımlanan bir mutasyon belirlenmiştir. Diğer üç ailede ise ortak bir aday bölge belirlenememiştir.

Juvenil Parkinsonizm (JP), hareketlerde yavaşlama, katılık, dinlenme halinde titreme gibi klasik parkinson hastalığı belirtilerinin yanı sıra 40 yaşından önce başlaması

ile karakterizedir. *AK3L1* geninde bir homozigot missens değişim tüm hastalarda ve ayrıca ailedeki bir sağlıklı bireyde bulunmuştur. Toplumdan rastgele seçilen 130 kişide bu değişim bulunamamıştır. Ailede hastalığın ortaya çıkmasına etki eden veya değişimi homozigot olarak taşıyan bireyi koruyan ikinci bir bölge olduğunu düşünmekteyiz.

Son olarak, otosomal resesif Doğuştan Serebellar Hipoplazi (CCLH) hastalığı akraba evliliği yapmış iki ailede çalışılmıştır. Hastalık ilerleme göstermeyen doğuştan serebellar ataksi, motor gelişim geriliği, koordinasyon bozukluğu, hipotonus ve beyinciğin gelişmemesi ile tanımlanmaktadır. Bir ailede hastalık bölgesi 6q16.1-22.1'e haritalandırılmıştır. Bölgede herhangi bir aday gen saptanamamıştır. Diğer ailede ise bu bölgeyle örtüşmeyen birçok aday bölge gözlenmiştir.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS / ABBREVIATIONS

θ	Recombination Fraction
φ	Hydrophobic Residue
A	Adenine
AK3L1	Adenylate Kinase 3-Like 1
AMP	Adenosine Monophosphate
AOA	Oculomotor Apraxia and Hypoalbuminemia
APTX	Aprataxin
ARSACS	Autosomal Recessive Spastic Ataxia of Charlevoix-Sagunay
A-T	Ataxia-Telangiectasia
ATM	Ataxia-Telangiectasia Mutated Gene
ATP	Adenosine Triphosphate
ATV	Ataxia-Telangiectasia Variant
ATXN7	Ataxin 7
bp	Base Pair
C	Cytosine
CANT1	Calcium Activated Nucleotidase 1
CASK	Calcium/calmodulin-dependent serine protein kinase
CCLH	Congenital Cerebellar Hypoplasia
CHST3	Carbohydrate Sulfotransferase 3
cM	Centi Morgan
CMP	Cytosine Monophosphate
CNV	Copy Number Variation
CoQ10	Coenzyme Q10
DNA	Deoxyribonucleic Acid
DNAH17	Dynein Axonemal Heavy Chain 17
DSB	Double-Strand Break
EDTA	Ethylenediaminetetraacetate
EST	Expressed Sequence Tag
FLNB	Filamin Beta

FRDA	Friedreich Ataxia
FXN	Frataxin
G	Guanine
GTP	Guanosine Triphosphate
HCiE	Homozygosity Comparison in Excel
HIT	Histidine Triad
HSP	Heat Shock Protein
IBD	Identical by Descent
IQ	Intelligence Quotient
JP	Juvenile Parkinsonism
kb	Kilobase
Lod	Log of Odds
LRS	Larsen Syndrome
Mb	Mega Base
MLE	Maximum Likelihood Estimate
MLS	Maximum Lod Score
NCBI	National Center For Biotechnology Information
OPHN1	Oligopherin 1
p	Short Arm of a Chromosome
PAGE	Polyacrylamide Gel Electrophoresis
PANT	Polynucletiode Kinase-3’-Phosphotase (PNKP)-Aprataxin Amino Terminal Domain
PAR	Pseudoautosomal region
PCR	Polymerase Chain Reaction
PD	Parkinson’s Disease
PGS1	Phosphatidylglycerophosphatase Synthase 1
PINK1	PTEN Induced Putative Kinase 1
PMM2	Phosphomannomutase 2
q	Long Arm of a Chromosome
RELN	Reelin
RNA	Ribonucleic Acid
rpm	Revolution Per Minute
rs	Reference Sequence

SACS	Sacsin
SCA7	Spinocerebellar Ataxia 7
SDS	Sodium Dodecyl Sulphate
SETX	Senataxin
SNP	Single Nucleotide Polymorphism
SOCS3	Suppressor of Cytokine Signaling 3
SOD1	Superoxide Dismutase 1
SSCP	Single Strand Conformational Polymorphism
T	Thymine
Taq	Thermus aquaticus
TEMED	N, N, N, N’-Tetramethylethylenediamine
TIMP2	TIMP Metallopeptidase Inhibitor 2
U	Unit
UDP	Uridine Diphosphate
UTR	Untranslated Region
UV	Ultraviolet
VLDLR	Very Low Density Lipoprotein Receptor
W	Watt

1. INTRODUCTION

Molecular studies were performed in this thesis with the aim of mapping and identifying recessive genes responsible for families referred to us with the diagnosis of autosomal recessive ataxia, Larsen Syndrome, Juvenile Parkinsonism and Congenital Cerebellar Hypoplasia.

1.1. Autosomal Recessive Ataxia

Autosomal Recessive Ataxia (ARA) is a subgroup of hereditary ataxias that are a group of genetic disorders characterized by slowly progressive impaired coordination of gait and generally also with that of hands, speech and eye movements. Ataxia (from Greek, meaning lack of order) is lack of coordination of muscle movements. It is considered as a clinical manifestation of problems in the nervous system, especially in the cerebellum that coordinates movement. The incidence of hereditary ataxias is around 1 in 5,000. Mode of inheritance can be autosomal dominant, autosomal recessive or X-linked. Most of the autosomal dominant ataxias are spinocerebellar ataxias that are slowly progressive and associated with cerebellar atrophy. In contrast, in recessive ataxias, clinical signs and symptoms are more variable, but the age of onset is generally the first decade of life (Table 1.1).

Friedreich ataxia (FRDA, MIM 229300) is the most common autosomal recessive hereditary ataxia with an incidence of 3-4 cases per 100,000 individuals (Jörg *et al.*, 2009). The disorder usually manifests before adolescence. In addition to neuropathological disabilities such as ataxia, sensory loss and muscle weakness, common signs are scoliosis, foot deformity and hypertrophic cardiomyopathy (Pandolfo, 2009). The spinocerebellar tracts, dorsal columns, pyramidal tracts, cerebellum and medulla are involved in the disease, and this is responsible for the typical combination of signs and symptoms specific to FRDA.

Table 1.1. Autosomal recessive ataxias, their causative genes and clinical features. (Modified from Bird, 2009.)

Disease	MIM	Gene	Locus	Distinctive Features
Friedreich ataxia	229300	*FXN*	9q13	Hyporeflexia, Babinski responses, sensory loss, cardiomyopathy
Ataxia-telangiectasia	208900	*ATM*	11q22.3	Telangiectasia, immune deficiency, cancer, chromosomal instability
Ataxia with oculomotor apraxia type 1	208920	*APTX*	9p13.3	Oculomotor apraxia, choreoathetosis, mild mental retardation, hypoalbuminemia, CoQ10 deficiency
Ataxia with oculomotor apraxia type 2	606002	*SETX*	9q34	Cerebellar atrophy, choreathetosis, hypoalbuminemia, mild mental retardation
Autosomal recessive spastic ataxia of Charlevoix-Sagunay	270550	*SACS*	13q12	Spasticity, peripheral neuropathy, retinal striation

Disease progression is variable, and patients with mild disease may still be ambulatory decades after onset, while those with severe disease may become wheelchair-bound within a few years. Life expectancy is reduced to an average of 40 to 50 years of age.

Friedreich ataxia is caused by mutations in *Frataxin* (*FXN)* at 9q13-21.1. The gene has seven exons, and the protein product has a highly conserved N-terminal leader peptide that facilitates the protein to localize to mitochondria. The protein is involved in iron metabolism. Deletion of *frataxin* in yeast results in the accumulation of iron in mitochondria, leading to the loss of mitochondrial function and DNA (Babcock *et al.*, 1997, Pandolfo and Pastore, 2009). Oxidative stress is another consequence of disturbed iron metabolism; thus, cultured cells from patients with FRDA are highly sensitive to hydrogen peroxide (Wong *et al.*, 1999). The most common genetic defect is a trinucleotide (GAA) repeat expansion in the first exon (Al Mahdawi *et al.*, 2006). GAA repeat numbers <12 are considered normal. Genetic diagnosis is established whenever expanded GAA repeats (>66) are detected in both copies of the gene.

Ataxia-Telangiectasia (A-T, MIM 208900) is an autosomal recessive disorder that is characterized by an early onset progressive cerebellar ataxia, oculocutaneous telangiectasia and susceptibility to bronchopulmonary disease and lymphoid tumors (Boder *et al.*, 1985). Progressive ataxia in A-T patients starts in infancy, affecting first the head and trunk muscles. The disease later leads to inability to walk, difficulty in speaking, drooling and oculomotor apraxia (impaired ability to coordinate certain eye movements) by the age of 10 years. Telangiectasia is observed as small dilated blood vessels near the surface of the skin and mucous membranes. Patients develop telangiectasia around the age of 5 years. Telangiectasia is observed most generally in the face, extremities and conjunctiva. Another manifestation of the syndrome is a higher cancer predisposition as compared to the general population. Even though all types of cancers are observed with higher incidence in the patients, the most frequent ones are lymphoma and leukemia. Cancer treatment for an A-T patient is a great challenge because the cells are highly sensitive to ionizing radiation and radiation therapy, a classical treatment for several types of cancer, cannot be tolerated by A-T patients. Immunodeficiency is another feature of the syndrome. Genomic instability ranging from microdeletions to chromosomal breakages is quite common, due to defects in DNA repair mechanisms. Mental retardation is not a characteristic of the disease, but some patients have mild mental retardation, and intelligence quotient (IQ) scores may drop below the normal range as the age and disease progress. The frequency of A-T is reported as about 1 in 40,000 births, but the disease may be actually much more common, since many children with A-T die at young ages or are not correctly diagnosed.

Life expectancy and quality for A-T patients are very poor due to the nature of the disease. Most patients lose ambulation before adolescence. Cancer and recurrent respiratory infections are the primary causes of death. Only a few A-T cases were reported to live to their forties.

The gene responsible for A-T is *ataxia-telangiectasia mutated* gene (*ATM*) at 11q22-q23 (Savitsky *et al.*, 1995). It extends over 160 kb of genomic DNA, contains 66 exons and codes for PI3/PI4-kinase family member protein (Platzer *et al.*, 1997), an important cell cycle checkpoint kinase that functions as a regulator of a wide variety of downstream proteins such as p53, BRCA1, CHK2, RAD17, RAD9 and DNA repair protein NBS1. ATM protein is thought to be one of the master controllers of cell cycle checkpoint

signaling pathways that are required for cell response to DNA damage and for genome stability. The protein is held inactive in normal cells as a dimer until irradiation of the cell, which induces rapid intermolecular autophosphorylation that causes dimer dissociation and initiates cellular ATM kinase activity. Activation of the ATM kinase is an initiating event in cellular responses to irradiation (Bakkenist *et al.*, 2003). The kinase domain of ATM is located towards the C-terminus of the protein, and mutations in this domain disturb protein function and lead to A-T.

Another autosomal recessive ataxia is early onset ataxia with oculomotor apraxia syndrome (AOA). It has similar manifestations to A-T and has two subtypes: ataxia, early onset, with oculomotor apraxia and hypoalbuminemia (AOA1, EAOH; MIM 208920) and spinocerebellar ataxia, autosomal recessive 1 (AOA2, SCAR1; MIM 606002).

AOA1 is an early-onset cerebellar syndrome with peripheral axonal neuropathy, oculomotor apraxia (limitation of ocular movements on command) and hypoalbuminemia. It was first described by Aicardi *et al.* in 1988 as an autosomal recessive syndrome that closely resembled A-T but also differed from it in important aspects. They reported 14 patients with a neurological syndrome manifesting with oculomotor apraxia, ataxia and choreoathetosis (irregular, rapid, involuntary movement flowing randomly from one part of the body to another) but not with extraneurological features of A-T, such as sensitivity to ionizing radiation and cancer susceptibility. The onset of the disease was relatively late, age of 7 years, and also no immunologic component was defined.

Disease progression is usually slow for AOA1. Oculomotor apraxia is observed in about 80 per cent of the cases. Magnetic resonance imaging of patients is characterized by marked cerebellar atrophy.

Aprataxin (*APTX*) at 9p13.3 was found mutated in patients with AOA1 in a large study on seven unrelated families (Date *et al.*, 2001). *APTX* encodes a 342 amino acid protein that is a member of histidine triad (HIT) superfamily. The protein has a polynucletiode kinase-3’- phosphotase (PNKP)-aprataxin amino terminal domain (PANT), a histidine triad (HIT) motif and a zinc-finger domain. It is predicted to have a nuclear localization, based on the presence of a potential nuclear localization signal between

PANT and HIT domains and a zinc-finger domain that represents a putative DNA binding site (Moreira *et al.,* 2001). Some members of the HIT family have nucleotide-binding and diadenosine polyphosphate hydrolase activities; thus, protein products are thought to play a role in single-strand DNA repair, particularly resolving abortive DNA ligation intermediates (Rass *et al.*, 2007). The catalytic activity of APTX resides within the HIT domain, and the C-terminal zinc finger domain provides stabilizing contacts that lock the enzyme onto its high affinity DNA site (Ulrich *et al.,* 2007). In fact, cells producing a truncated product lacking HIT and zinc-finger domains are shown to be hypersensitive to artificially induced DNA single-strand breaks (Mosesso *et al.,* 2005). Mutations in the histidine triad (HφHφHφφ, φ representing a hydrophobic residue) of HIT domain are shown to abolish the catalytic activity (Ahel *et al.,* 2006). Another study showed that aprataxin specifically removes damaged 3'-ends and, through this action, it acts together with DNA polymerase and ligase to repair single-strand breaks. Disease-associated mutant forms of the protein generally lack the removal activity, strongly suggesting that the loss of such activity is closely linked to the pathogenesis (Takahashi *et al.,* 2007).

As of 2010 many mutations have been reported in AOA1, and one of those is p.W279X, that was discovered in a family afflicted with ataxia and CoQ10 deficiency (Quinzii *et al.,* 2005). This mutation leads to the premature termination of translation and thus to the truncation of the protein product, losing the zinc finger domain which is responsible for stabilizing the catalytic site, HIT domain, onto DNA target site. The report was the first to propose an association between CoQ10 deficiency and ataxia. It had analyzed 13 additional patients with CoQ10 deficiency, but no *APTX* mutation was found in those. Another study reported decreased muscle CoQ10 in five patients out of six with AOA1 (Le Ber *et al.,* 2007). Three of those patients were homozygous for mutation p.W279X and had the lowest CoQ10 levels. The study found no correlation between CoQ10 deficiency and disease duration or severity.

AOA2 is a subtype of AOA and characterized by cerebellar atrophy, early loss of reflexes, choreathetosis, mild mental retardation, peripheral neuropathy, hypoalbuminemia and increased alpha-fetoprotein. Oculomotor apraxia is only an occasional feature of AOA2. Prevelance of AOA2 is slightly less frequent than A-T. Progressive ataxia appears before the age of 15 years.

Senataxin (*SETX*) gene at 9q34.13 was found mutated in patients with AOA2. *SETX* encodes a large protein of 2,677 amino acids which mainly localizes to nucleus (Moreira *et al.,* 2004). The protein has a DNA-RNA helicase domain. It has been proposed to have a role in DNA damage response and double-strand break (DSB) repair, because cell lines from AOA2 patients showed an increased sensitivity to agents that generate cross-links in DNA and DSBs but lacked sensitivity to ionizing radiation (Suraweera *et al.*, 2007).

Autosomal recessive spastic ataxia of Charlevoix-Sagunay (ARSACS) is a neurodegenerative disorder characterized by cerebellar atrophy, progressive spastic ataxia and peripheral neuropathy. Retinal striation (increased visibility of the retinal nerve fibers) may be observed but is not a major diagnostic criterion. Age of onset is before 5 years.

Sacsin (*SACS*) at 13q12 was found mutated in patients with ARSACS. It encodes a protein with 4,579 amino acids and is predominantly expressed in neurons, especially in motor neurons. Its localization in cells is cytoplasmic and mitochondrial. A strong sequence similarity is identified between a region near sacsin C-terminus and HSP40, a heat shock protein. This highly conserved region is the defining feature of HSPs; therefore, sacsin may represent a direct link between a mechanism of ataxia and the HSP machinery (Parfitt *et al.,* 2009).

In this study two families were investigated, with a total of 9 patients and 29 healthy individuals. The initial diagnosis for the patients was autosomal recessive ataxia.

1.2. Juvenile Parkinsonism

Parkinson’s Disease (PD, MIM 168600) is the second most common neurodegenerative disorder after Alzheimer’s disease and affects 1-2 per cent of the population above the age of 60 years (Fahn *et al*., 2003). It is characterized by a classical triad of signs: bradykinesia, rigidity and resting tremor. Bradykinesia is the generalized slowness of movement, rigidity is resistance within the range of passive movement of a joint and resting tremor is the involuntary rhythmic oscillation of a body region that is produced by alternating contractions of reciprocally innervated muscles. PD is a result of selective degeneration of the dopamine producing cells in substantia nigra. The loss of

dopaminergic neurons is accompanied by the accumulation of Lewy bodies in the remaining substantia nigra neurons. Those bodies are fibrillar cytoplasmic inclusions consisting of protein aggregates (Mc Naught and Olanow, 2006).

Three mechanisms are proposed to be responsible for PD: defective ubiquitin-proteasome pathway, mitochondrial dysfunction and oxidative stress. The ubiquitin-proteasome pathway degrades proteins marked for destruction and is believed to have a central role in several neurodegenerative diseases. The first discovery of mitochondrial dysfunction in PD was the observation of Parkinsonism-like effects of some herbicides that disturbed mitochondrial electron transport chain, resulting in ATP depletion. Oxidative stress is very important in all neurons, especially in dopaminergic ones, because metabolism of dopamine is known to generate free radicals that will lead to inflammation and programmed cell death, as observed in the dopaminergic neurons in PD.

PD has autosomal dominant, autosomal recessive and sporadic forms. Seven genes have been implicated in the Mendelian forms (Table 1.2).

Table 1.2. Parkinson's Disease-associated genes, their mode of inheritance and age of onset. (Modified from Abou-Sleiman *et al*., 2006.)

Locus	MIM	Inheritance	Gene	Gene Function	Age of onset
PARK1/4 4q21	605543	Autosomal Dominant	*α-Synuclein*	Involved in vesicle synaptic formation	30 – 60 years
PARK2 6q25.2-q27	600116	Autosomal Recessive	*Parkin*	An E3 ligase	~ 30 years
PARK6 1p36	605909	Autosomal Recessive	*PINK1*	A mitochondrial kinase	30 – 50 years
PARK7 1p36.23	606324	Autosomal Recessive	*DJ1*	Involved in oxidative stress response	20 – 40 years
PARK8 12q12	607060	Autosomal Dominant	*LRRK2*	A protein Kinase	40 – 60 years

Juvenile Parkinsonism (JP) is parkinsonism with onset before the age of 40 years (Takahashi *et al*., 1994). The age of onset for JP is generally between 25 and 30 years, with a range from 10 to 45 years. Ishikawa and Tsuji, 1996 stated that patients present mild

classical PD symptoms and retropulsion (involuntary backward stepping when starting walking), dystonia (twisting and repetitive movements) of the feet and hyperreflexia (overreactive reflexes). Disease progression is usually slow. Patients respond well to levodopa but dopa-induced dyskinesias and wearing-off phenomena occurs frequently. Mode of inheritance is recessive except for α-synuclein-related dominant cases.

1.2.1. α-Synuclein

α-Synuclein gene encodes a cytosolic protein that is localized at presynaptic terminals and thought to be involved in synaptic function, modulating the dopamine neurotransmission (Abeliovich et al., 2000). The mutant α-synuclein forms stable β-pleated sheets and later fibrils in the cell, resulting in the generation of pathological inclusions known as Lewy bodies (Goldberg and Lansbury, 2000).

1.2.2. Parkin

Mutations in *Parkin* gene account for half of autosomal recessive JP and about 20 per cent of isolated PD cases (Lücking *et al.,* 2000). Parkin is a neuroprotective cytosolic protein and an E3 ubiquitin-ligase that is a part of ubiquitination system. Mutant forms of Parkin do not form aggregates.

Parkin-related PD cases have a slow progression but earlier age of onset of disease, present dystonia with classic triad of signs at onset and respond well to L-Dopa (Lohmann *et al.,* 2003).

1.2.3. DJ1

Oncogene DJ1 has been shown to act as a redox-sensitive molecular chaperone, preventing the aggregation of α-synuclein and thus exerting a protective function under oxidative stress conditions (Shendelman *et al.,* 2004).

Phenotypically, *DJ1* mutation cases are indistinguishable from *Parkin* and *Pink1*-linked PD cases (Klein *et al.,* 2007).

1.2.4. PINK1

PTEN induced putative kinase 1 (*PINK1*)-linked PD cases have similar phenotypes as *Parkin*-related cases and respond well to L-Dopa. Parkin and PINK1 is thought to function in the same mitochondrial signaling pathway (Poole *et al.,* 2008). It has also been reported that the mutant phenotype in PINK1 deficient drosophila can be rescued by Parkin expression but not vice versa, demonstrating that Parkin acts downstream of PINK1 (Clark *et al.,* 2006; Park *et al.,* 2006; Yang *et al.,* 2006).

Several genes in human genome are potential candidates to explain the familial PD not associated with known disease genes. One of them is *Adenylate Kinase 3-Like 1 (AK3L1),* a member of nucleoside monophosphate kinases that are ubiquitous enzymes, present both in prokaryotes and eukaryotes, and catalyze the reversible phosphoryl transfer between various nucleoside mono and triphosphates in the salvage pathway of nucleotide metabolism (Lee *et al.,* 1998). It is a poorly expressed mitochondrial matrix protein with highest expression levels in brain, kidney, liver and heart tissues (Yonada *et al.,* 1998). It phosphorylates AMP, dAMP, CMP and dCMP with ATP and GTP as phosphate donors. The protein has a mitochondrial targeting sequence mapped to first 11 amino acids and an ATP-binding loop that consists of amino acid residues 12-20 in the N-terminal part (Panayiotou *et al.,* 2010).

In this study genetic investigations were performed in a family with four affected members initially diagnosed as autosomal recessive Juvenile Parkinsonism.

1.3. Larsen Syndrome

Larsen Syndrome is characterized by congenital large-joint dislocations and craniofacial abnormalities. It occurs in about 1 in 100,000 births, with equal gender incidence (Larsen *et al.,* 1950). Additional cardinal features are short terminal phalanges leading to pseudoclubbing, dislocations of the hip, knee and elbow joints and foot deformities. Craniofacial anomalies include hypertelorism, prominent forehead, depressed nasal bridge and a flattened midface. Cleft palate and short stature as well are often

associated (Silverman, 1972). The severity of those features varies greatly among patients, even in those from the same family (Becker *et al.*, 2000). Mental capacity is not affected by the disease.

Diagnosis can be made as early as in the 16th week of gestation by ultrasonographic investigations in pregnancies at risk. The progress of the disease manifestations is rapid in the first two decades of life, due to the rapid growth of the body. Kyphosis (increased angle of upper spine) and scoliosis (abnormal lateral curvature of the spine) may develop and must be handled appropriately, as those can lead to respiratory problems or to spinal compressions, thus causing neurological deficits. Management of the disease is mainly based on surgically reducing the joint anomalies such as knee and hip dislocations and spinal deformities and physical therapy for alleviating the pain/problems.

The mode of inheritance for Larsen Syndrome can be either autosomal dominant (MIM 150250) or autosomal recessive (MIM 245600). Phenotypes of the two inheritance types are not distinct, but J. G. Hall in 1975 (MIM 245600) suggested that characteristic facial appearance is less striking and abnormalities such as syndactyly and severe short stature are more frequent in the recessive forms. In the dominant forms of the disease, the severity of symptoms may differ greatly among generations. For example, diagnosis of Larsen Syndrome in a child led to correct diagnosis in the parent who had escaped diagnosis due to an unexpectedly mild manifestation of the disease (Becker *et al.*, 2000). Germline mosaicism has also been reported (Petrella *et al.*, 1993). That family was originally reported by Bloch and Peck in 1965 as a possible example of the recessive form because the parents were unaffected and two sibs were affected. However, diagnosis was later corrected as autosomal dominant form with germline mosaicism when one of the affected sisters gave birth to an affected daughter (Petralla *et al.*, 1993).

Genetic heterogeneity has been described in both autosomal dominant and recessive forms, with great variability in clinical presentations (de Nazar, 1980). Two genes have been identified as mutated in Larsen Syndrome: *Filamin beta* (*FLNB*) at 3p14.3 and *Carbohydrate Sulfotransferase 3* (*CHST3*) at 10q22.1. The first gene to be identified was *FLNB*. Mutations in this gene cause human skeletal disorders such as autosomal dominant Larsen Syndrome, autosomal recessive spondylocarpotarsal syndrome (MIM 272460) and

autosomal dominant perinatal lethal atelosteogenesis I and III (MIM 108720 and 108721, respectively; Krakow *et al.,* 2004). This diversity in skeletal abnormalities caused by mutations in a single gene indicates a central role for filamin B in skeletal morphogenesis. *FLNB* is ubiquitously expressed and encodes a protein that is a member of the filamin family. Filamins are high molecular mass cytoplasmic proteins that organize filamentous actin in networks and stress fibers. They regulate the structure and activity of the cytoskeleton by cross-linking actin into three-dimensional networks, linking the cell membrane to the cytoskeleton and serving as scaffolds on which intracellular signaling and protein trafficking pathways responsible for skeletal development are organized (van der Flier and Sonnenberg, 2001). Flnb deficient mice developed phenotypic features similar to the human phenotype. They have shortened distal limbs and a small body size and develop fusion of the ribs and vertebrae, abnormal spinal curvatures and dysmorphic facial bones. Increased apoptosis along the bone periphery of the distal appendages and impaired progressive differentiation of chondrocyte precursors were also observed (Lu *et al.,* 2007).

The second gene reported as mutated in Larsen Syndrome is *CHST3*. Mutations in this gene result in three autosomal recessive diseases: spondylepiphyseal dysplasia (MIM 603799), humerospinal dysostosis (MIM 143095) and Larsen Syndrome (Thiele *et al.,* 2004, Hermanns *et al.,* 2008). *CHST3* encodes an enzyme that catalyzes sulfation of chondroitin, a proteoglycan found in the extracellular matrix. The protein is involved in cell migration and differentiation in tissues such as cartilage and skin (Raman *et al.,* 2005).

Calcium activated nucleotidase 1 (*CANT1*) is a gene that was reported as responsible for Larsen-like disease, Desbuquois Dysplasia (MIM 251450, Huber *et al.,* 2009). This gene is located at 17q25.3 and encodes a soluble UDP-preferring nucleotidase belonging to the apyrase family. The exact function of the protein is yet unknown, but its substrates are proposed to be involved in several major signaling pathways, such as calcium ion release. Five alternating positively and negatively charged residues (D114, E284, R300, E365 and K394) comprise a network of four salt bridges involved in the catalytic site of CANT1. Direct mutagenesis of R300 has been shown to disrupt the electrostatic interactions in the salt bridge and result in decreased enzyme activity without altering calcium binding or producing a conformational change (Dai *et al*., 2004).

Desbuquois Dysplasia is an autosomal recessive chondrodysplasia displaying multiple joint dislocations, severe perinatal growth retardation, joint laxity, short extremities and progressive scoliosis. The main radiological features are short long bones with metaphyseal splay (widening) and exaggerated trochanter (widening of bone growth plate ends) of proximal femur and advanced carpal and tarsal bone age with a delta (extra) phalanx (Huber *et al.*,2009). Hand anomalies such as phalangeal dislocations and extra ossification centers are not observed in all Desbuquois patients (Gillessen-Kaesbach *et al.*, 1995, Faivre *et al.*, 2004).

In this study genetic investigations were performed in seven families with eight patients initially diagnosed with autosomal recessive Larsen syndrome.

1.4. Congenital Cerebellar Hypoplasia

Congenital Cerebellar Hypoplasia (CCLH, MIM 213000) is an autosomal recessive disorder characterized by nonprogressive congenital cerebellar ataxia with mental retardation, delayed motor development, disturbed coordination and hypotonia associated with cerebellar hypoplasia (Schurig *et al.*, 1981). Cerebellar hypoplasia is referred to a small cerebellum with fissures (grooves dividing different parts of cerebellum) of normal size compared to folia (small leaflike laminae on cerebellar cortex; Barkovich, 2000). Cerebellar hypoplasia constitutes about 10 per cent of all cerebellar malformations and 3 per cent of all central nervous system malformations (Patel and Barkovich, 2002; Pinar *et al.*, 1998). All motor related symptoms, such as exaggerated deep tendon reflexes, gait abnormalities and very late or no walking, indicate a cerebellar pathology. Magnetic resonance imaging shows cerebellar atrophy affecting the vermis and/or hemispheric region of cerebellum. On histological examinations, degeneration of the granular cells as well as the Purkinje cells is observed (Ferrer *et al.*, 1987)

Differential diagnosis of congenital cerebellar hypoplasia is not simple, because several diseases of the nervous system such as spinocerebellar ataxias, pontocerebellar hypoplasias and Joubert Syndrome present with similar clinical manifestations.

As of May 2010 there are 7 genes reported as mutated in various forms of cerebellar hypoplasia: *VLDLR*, *RELN*, *OPHN1*, *CASK*, *EN2*, *POMT2* and *PMM2* (Table 1.3).

Table 1.3. Cerebellar hypoplasias, their causative genes and clinical features

Disease	MIM	Gene	Locus	Distinctive Features
Autosomal recessive				
Cerebellar ataxia, mental retardation and dysequlibrium syndrome	224050	*VLDLR*	9p24	Nonprogressive cerebellar ataxia, mental retardation, delayed motor development
Norman-Roberts type Lissencephaly Syndrome	257320	*RELN*	7q22	Lissencephaly, hyperreflexia, seizures, mental retardation
Congenital disorder of glycosylation, type 1a	212065	*PMM2*	16p13.3-p13.2	Peripheral neuropathy, cerebellar hypoplasia, mental and psychomotor retardation
X-linked				
Mental retardation and microcephaly with pontine and cerebellar hypoplasia	300749	*CASK*	Xp11.4	Pontine hypoplasia, severe mental retardation, microcephaly, sensorineural hearing loss
Mental retardation with cerebellar hypoplasia and distinctive facial appearance	300486	*OPHN1*	Xq12	Mental retardation, partial seizures, distinctive face, no ataxia

Mutations in *Reelin* (*RELN*) were identified in an autosomal recessive form of lissencephaly (lack of brain folds) associated with severe abnormalities of cerebellum, hippocampus and brainstem (Hong *et al*., 2000). RELN is a large secreted extracellular protein involved in a molecular pathway that regulates the migration of neurons along the radial glial fiber network. It thus controls cell-cell interactions critical for cell positioning and neuronal migration during brain development (Curran and D'Arcangelo, 1998).

Recessive mutations in the *Very low density lipoprotein receptor (VLDLR)* gene result in cerebellar hypoplasia (Boycott *et al., 2005*). The low density lipoprotein receptor gene family consists of cell surface proteins involved in receptor-mediated endocytosis of specific ligands. *VLDLR* protein plays important roles in low density lipoprotein-triglyceride metabolism and the reelin signaling pathway. Deletion of *VLDLR* homolog in

mice results similarly in a hypoplastic cerebellum and failure of assembly of Purkinje cells in a tight layer (Trommsdorf *et al.,* 1999).

Oligopherin 1 (*OPHN1*) gene mutations cause X-linked cerebellar hypoplasia with distinctive facial dysmorphism and severe mental retardation (Billuart *et al.*, 1998). The gene encodes a Rho-GTPase-activating protein that promotes GTP hydrolysis of Rho subfamily members. Rho proteins are important mediators of intracellular signal transduction that affects cell migration and cell morphogenesis. Inactivation of *Ophn1* in cultured mouse cells increased the density and proportion of immature dendritic spines, thus indicating that Ophn1 is required at all stages of development (Khelfaoui *et al*., 2007).

Mutations in *Calcium/calmodulin-dependent serine protein kinase* (*CASK*) gene result in congenital and marked postnatal microcephaly, severe mental retardation, sensorineural hearing loss, disproportionate pontine and cerebellar hypoplasia (Tarpey *et al*., 2009). The protein product is a member of membrane-associated guanylate kinase protein family. Proteins that are members of the family are scaffolds that are located at synapses in the brain.

Phosphomannomutase 2 (*PMM2*) gene mutations cause glycoprotein biosynthesis defects that manifest as carbohydrate-deficient glycoprotein syndrome type I (Matthijs *et al*., 1997). The protein is necessary for the synthesis of dolichol-P-oligosaccharides that play critical roles in metabolism and cell recognition, adhesion and migration (Marquardt and Denecke, 2003). Clinical features of the disease are severe encephalopathy, pronounced psychomotor retardation, peripheral neuropathy and cerebellar hypoplasia.

In this study genetic investigations were carried out in two families, each having 2 sibs diagnosed with Congenital Cerebellar Hypoplasia.

1.5. Linkage Analysis

Genetic linkage is the tendency of two loci on the same chromosome to be inherited together. Alternatively, a new combination of alleles could be created by a crossover between homologous chromosomes in meiosis. The recombination rate, or recombination

fraction (θ), is directly related to the distance between the loci. The maximum value for θ is 0.5, and it indicates that either the two loci are far apart from each other so that there is no linkage between them or they are on different chromosomes. Genetic markers on a chromosome are utilized to detect recombination events and to localize a disease gene near them.

The availability of various genetic markers, including tandem repeats, microsatellites and single nucleotide polymorphisms (SNP), and the accessibility to databases storing information on those genetic markers have made linkage analysis more efficient.

1.6. Lod Score Analysis

Linkage analysis is classified into two groups: parametric (model-based) and nonparametric (model-free). Parametric method necessitates various predetermined genetic parameters such as penetrance, disease-allele frequency and mutation rates to explain the mode of inheritance. The statistical method to evaluate parametric linkage analysis is lod (logarithm of odds) score analysis.

Morton (1955) described lod score analysis as discriminating between two hypotheses: the null hypothesis of no linkage ($\theta = 0.5$) and the alternative hypothesis of linkage ($\theta < 0.5$). The statistical criterion for proving linkage between two traits is based on an observed odds ratio, namely "L" that is the ratio of the probability of observing the distributional model of the two traits in a family with linkage at θ to the same probability under the hypothesis of no linkage at $\theta = 0.5$ (Risch, 1992). This is a likelihood-based method that originated from the Neymann-Pearson lemma that states that if there is a best test for a given hypothesis, it takes the form of a likelihood ratio test. The logarithm of L, lod score, indicated by the function of $Z(\theta)$ is calculated at several values of θ, and the maximum test statistics of Z-maximum likelihood lod score is reported. A lod score of 3 is the limit for human linkage analysis to accept linkage (Terwilliger and Ott, 1994). The values between -2 and 3 are inconclusive for linkage analysis. A lod score of $\leq$-2 excludes linkage for the region

Basically there are two types of lod score analysis, two-point and multipoint. Two-point lod score analysis calculates lod scores by evaluating cosegregation of each marker with the disease locus. The analysis determines the recombination events in individuals and the inheritance of alleles throughout the pedigree. Multipoint lod score analysis allows analysis of more than two loci at a time. By this method linkage analysis can be more efficient, but it also becomes more cumbersome. It calculates lod scores for all possible positions on a given map of genetic markers and is especially useful for pedigrees that have limited informativeness for markers.

Complex traits result from two or more interacting genes and/or environmental factors. Due to these factors, model free, nonparametric methods are applied in studies on complex traits. The idea behind nonparametric methods is that affected individuals should be carrying a shared haplotype that is identical by descent in the region of a susceptibility gene, no matter what the mode of inheritance is.

Homozygosity mapping is another application of parametric analysis in consanguineous families and isolated populations. Affected individuals for an autosomal recessive trait in such groups have most likely received a shared haplotype from each parent and, thus, from a common ancestor and in the homozygous state (Lander and Botstein, 1987). Generally the frequency of a rare disease trait is so low that the disease can only arise in a consanguineous marriage. Such inbred families are of great value for linkage studies, as a single family alone can produce a significantly high lod score.

A variety of computer software solutions have been developed for statistical tests used in linkage analysis. SuperLink and FastLink are two such examples applied in two-point lod score analysis. GeneHunter and SimWalk are specifically designed for multipoint analysis. On the other hand, software such as Allegro and Merlin can be used for both of the above analyses as well as for nonparametric analyses. Since the input formats and commands for each of those software are different, a common platform, most generally easyLINKAGE, is used for automatically converting input formats and commands when needed (Hoffman and Lindner, 2005).

1.7. Homozygosity Comparison in Excel

Human Genome Project and HapMap Project facilitated the discovery of a vast number of single nucleotide polymorphisms (SNPs) in the last decade. SNPs are being used for linkage and association studies since then, due to their even distribution throughout the genome. However, an SNP has only two alleles, yielding a maximum heterozygote fraction of about 50 per cent in a population. The drawback of such low informativeness is compensated when clusters of SNPs are considered as units. The recent advancements in the hybridization technologies make it feasible to genotype hundreds of thousands of SNPs on a single chip. Unfortunately, the freely available software could not provide sufficient capacity to process the data generated with those SNPs. The huge data load generated from a large number of SNPs utilized in a single project necessitates the employment of a new analysis method. Within the scope of this study, Homozygosity Comparison in Excel (HCiE) was developed. This method uses formulas embedded in Microsoft Excel to identify homozygous loci for a patient and compare his/her genotype with other family members as well as with other patients'. The SNP names are on the first column, alleles for each individual are copied to consecutive columns on Sheet A and the formulas are written on Sheet B. The basic purpose of the formulas is to compare the contents of determined cells in a row of specific marker and report the result in binary format, "1"/"True" or "0"/"False". This binary format is then differentially colorized for easy visualization of the result. By using the program, SNP genome scan data can be analyzed and candidate regions determined for the next step: candidate gene approach.

1.8. Candidate Gene Approach

A genetic locus is identified by using the data generated from genome scan and later narrowed down by fine-mapping. Computer programs and haplotype inspection are applied to confirm linkage to the locus. The genes at the locus are evaluated to assess which of them are good candidates to be the disease gene. Four criteria to be considered in candidate gene approach are:

- Location: The gene must be localized within the maximal critical gene locus.
- Pathway: The gene product must be involved in mechanisms related to the pathogenesis observed in the disease phenotype.

- Expression: Gene expression in tissues and/or developmental stage must be compatible with disease phenotype.
- Model: A similar disease phenotype must be observed in animal models.

2. PURPOSE

In the first part of this study, the purpose was to map the gene loci for four rare autosomal recessive disorders using the genome scan data for the families and subsequent fine-mapping and statistical analyses. The disorders are Autosomal Recessive Ataxia (ARA), Juvenile Parkinsonism (JP), Autosomal Recessive Larsen Syndrome (LRS) and Congenital Cerebellar Hypoplasia (CCLH).

Optimization of existing methods and developing new ones to analyze the data obtained were aimed at the beginning of the study. Information gained from microsatellite and SNP genome scans and fine-mapping studies were analyzed with linkage software for haplotype segregation and lod scores. Additionally, Homozygosity Comparison in Excel was developed in our laboratory during the course of this thesis to analyze high density SNP data. The purpose was to identify the chromosomal regions that exhibit shared homozygosity in affected individuals in a pedigree.

In the second part of this study, candidate gene approach on identfied loci was conducted with the aim of finding the genes responsible for Autosomal Recessive Ataxia, Juvenile Parkinsonism and Autosomal Recessive Larsen Syndrome. Single strand conformational polymorphism, restriction enzyme studies and direct sequencing were applied with this intention. The purpose was to find a gene variant that was indicative of a pathological mutation and not a polymorphism.

3. MATERIALS

3.1. Subjects

Informed consent was obtained from all subjects participatad in this study. The study was approved by the Committee on Research with Human Participants at Boğaziçi University.

3.1.1. Autosomal Recessive Ataxia

Two families with Autosomal Recessive Ataxia were investigated for his study. Family 1 originated from Aegean region and Family 2 from Black Sea region (Figures 3.1 and 3.2). In total nine individuals were afflicted with the autosomal recessive syndrome. Clinical findings for the three cases in Family 1 were described by Dr. Hatice Karasoy at Ege University. Three of the five patients in Family 2 were evaluated by Dr. Bülent Kara at Kocaeli University. Patients were grouped by their clinical and neurological features (Table 3.1). Group 1 was composed of the three patients in Family 1 and 2-601, 2-605 and 2-608 in Family 2. Patients 2-504, 2-609 and 2-610 in Family 2 were considered as Group 2 (Figure 3.2).

Table 3.1. Clinical and neurological findings in patients with autosomal recessive ataxia.

Clinical Feature	Group 1	Group 2
Ataxia	+ (Progressive)	+
Intellectual disability	Mild	Severe
Age of onset	7 years	Birth
Communication	Understand and respond to verbal commands	Poor
Spasticity	-	+
Walking	With support	Without support
Nystagmus	+	-

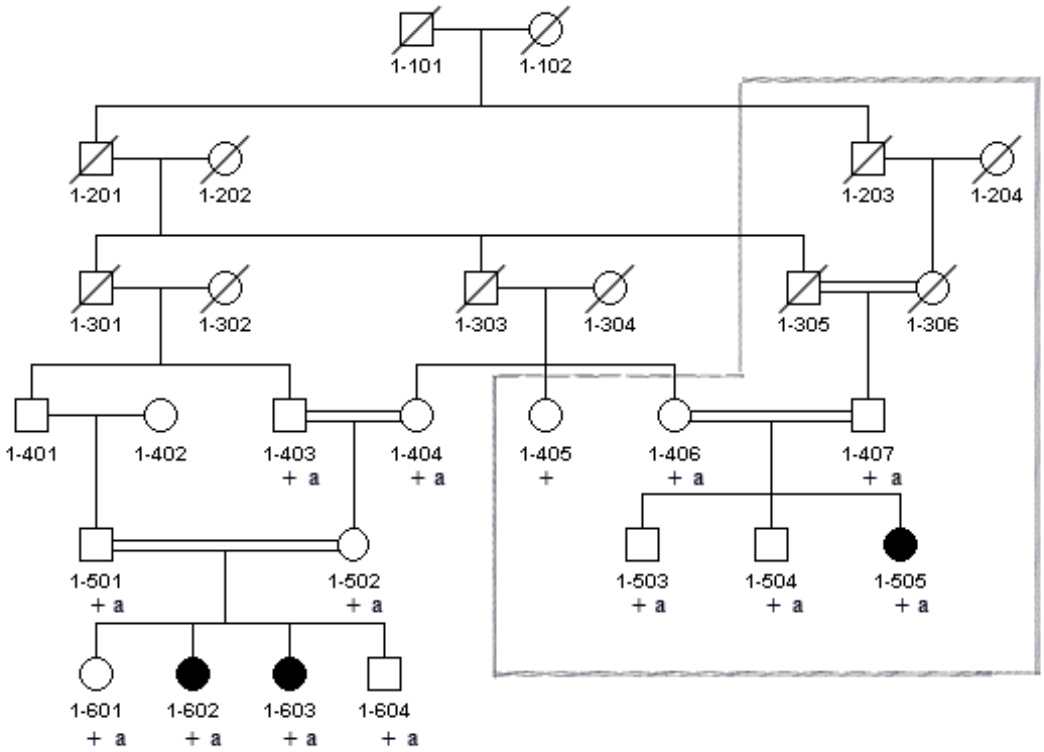


Figure 3.1. Pedigree diagram for ARA Family 1. The part of the family that was available after March 2008 is contained within shaded lines. DNA samples available for this study are marked with a + sign. DNA samples available for the Illumina 370K chip are depicted with an “a”.

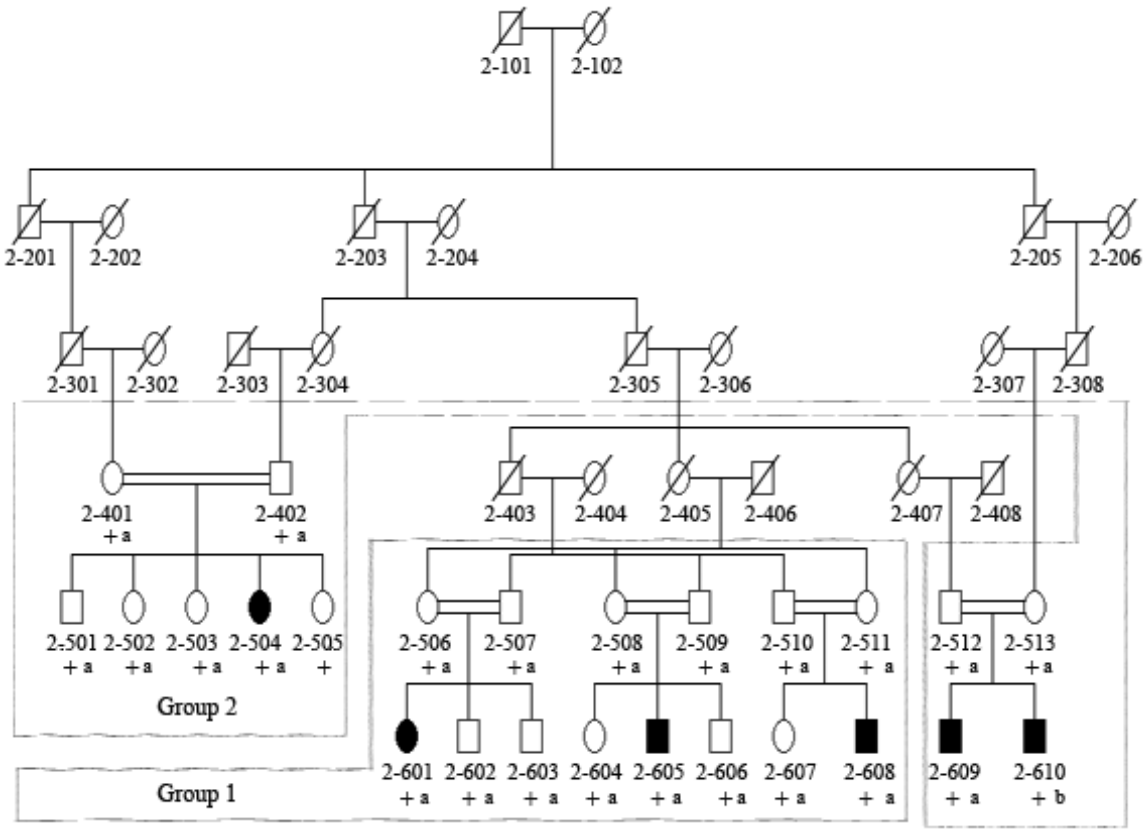


Figure 3.2. Pedigree diagram for ARA Family 2. Grouping according to clinical evaluation is contained within shaded lines. DNA samples available for this study are marked with a + sign. DNA samples available for the Illumina 370K chip and Illumina 1M chip are depicted with an “a” and a “b”, respectively.

3.1.2. Juvenile Parkinsonism

A large consanguineous family afflicted with Juvenile Parkinsonism was investigated (Figure 3.3). Blood samples from 16 individuals were supplied by Dr. Hatice Karasoy at Ege University. In May 2008, individual 505, previously reported healthy, presented disease symptoms at the age of 11 years. Individuals 405, 406 and 407 were not included in microsatellite or SNP genome scans.

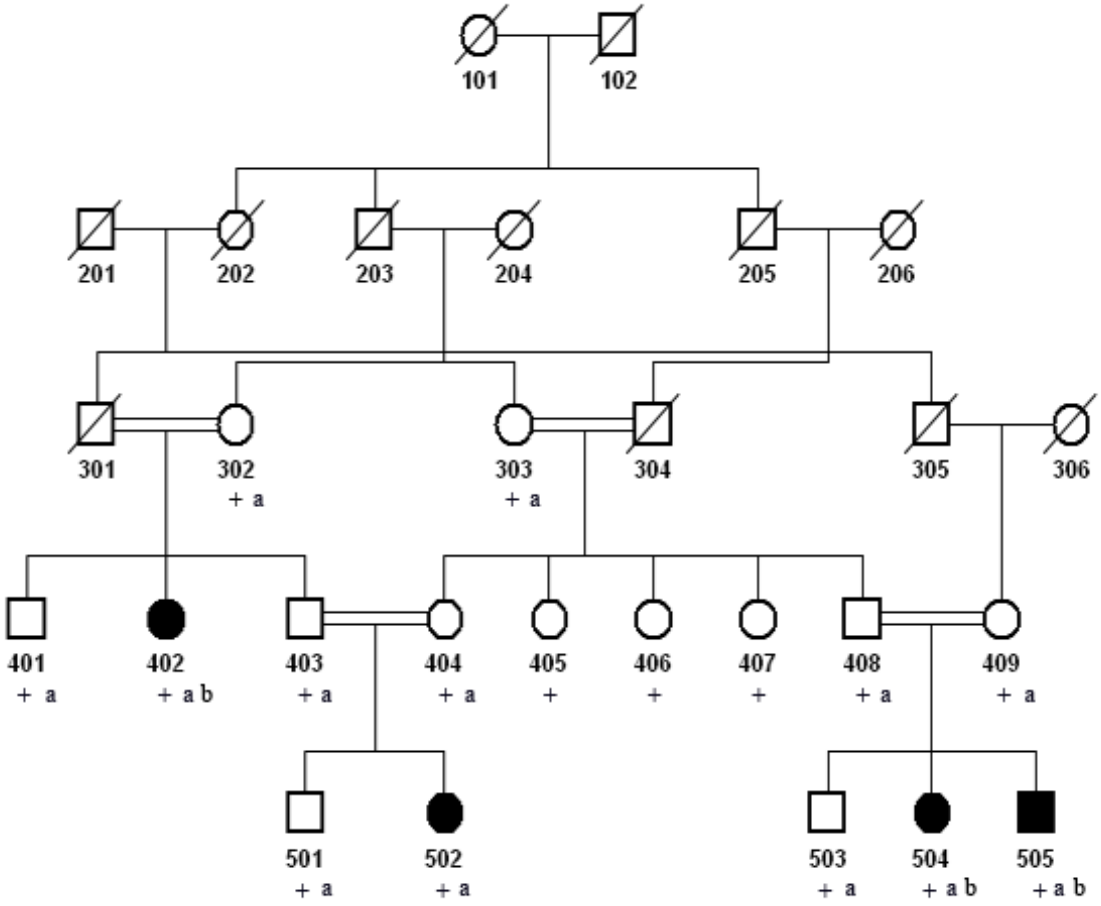


Figure 3.3. Pedigree diagram for JP family. DNA samples available for this study are marked with a + sign. DNA samples available for microsatellite genome scan, Illumina 370K chip and Illumina 1M chip are depicted with an “a” and a “b”, respectively.

3.1.3. Larsen Syndrome

Seven consanguineous families diagnosed with Larsen Syndrome were investigated in this study (Figure 3.4). Blood samples were supplied by Dr. Beyhan Tüysüz at Istanbul University and Dr. Davut Gül at Gülhane Military Academy of Medicine. Microsatellite genome scan was available for all patients from seven families and for healthy individuals of families LRS1, LRS2, LRS3 and LRS6. SNP genome scan in spring 2010 was carried out for patients in families LRS1, LRS2, LRS3, LRS4 and LRS 7, and in all members of family LRS6.

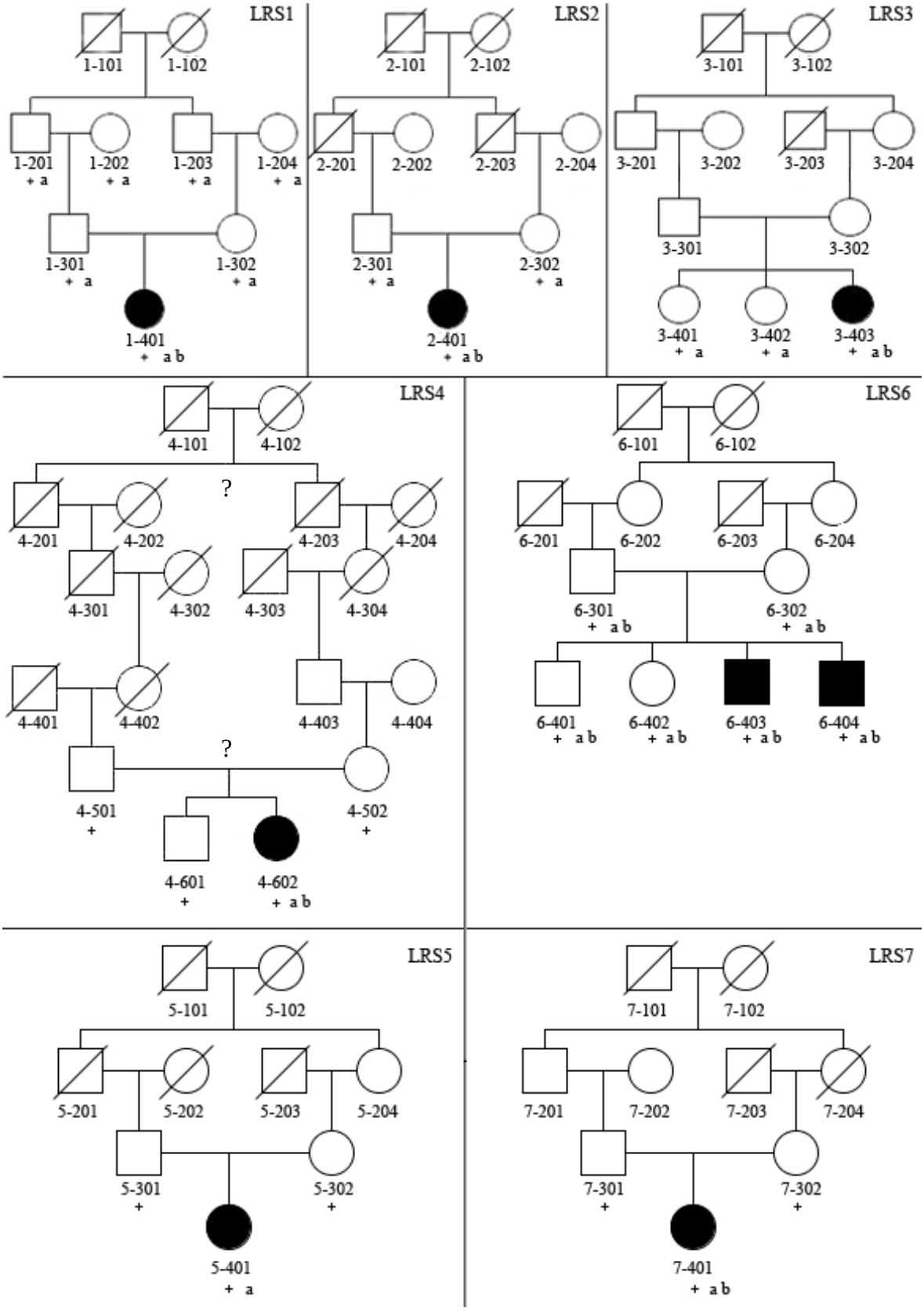


Figure 3.4. Pedigree diagrams for LRS families. DNA samples available for this study are marked with a + sign. DNA samples available for microsatellite genome scan and for Illumina 1M chip are depicted with an “a” and with a “b”, respectively.

3.1.4. Congenital Cerebellar Hypoplasia

Two consanguineous families afflicted with Congenital Cerebellar Hypoplasia were investigated in this study (Figure 3.5). Blood samples were obtained by Dr. Aslı Tolun. DNA samples from all members of two families were available for microsatellite genome scan. Later SNP genome scan also was carried out for patients 5-601 and 6-403.

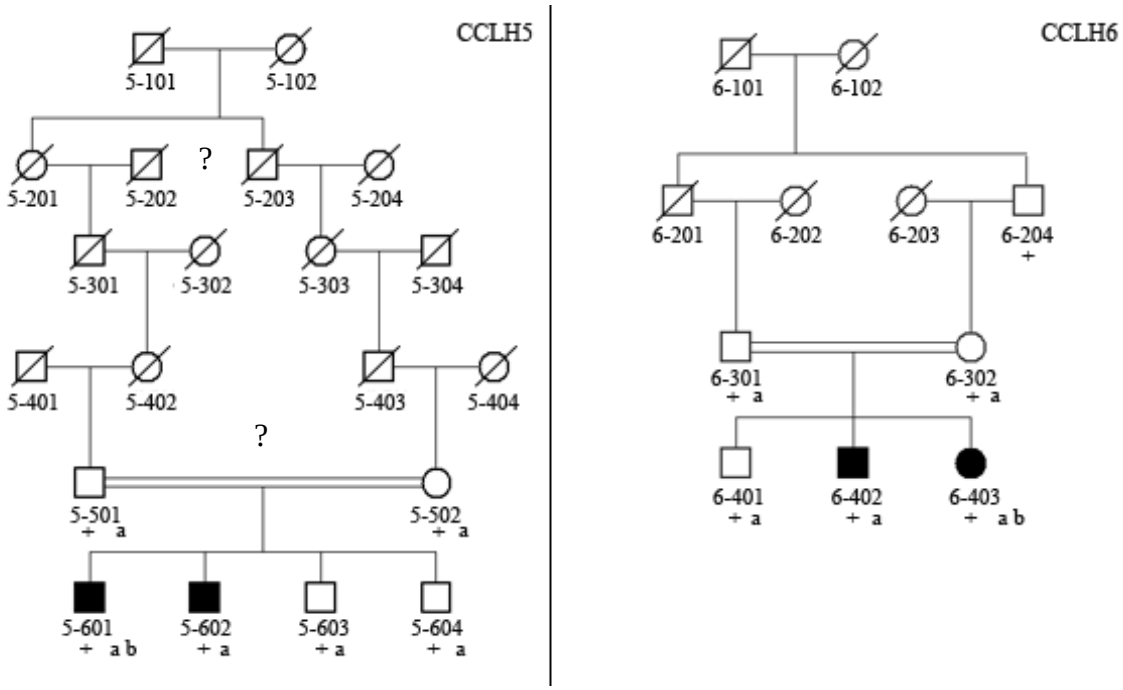


Figure 3.5. Pedigree diagrams for CCLH families. DNA samples available for this study are marked with a + sign. DNA samples available for microsatellite genome scan and for Illumina 670K chip are depicted with an “a” and a “b”, respectively.

3.2. Chemicals

Chemicals used in this study were purchased from Biochrom (Germany), Carlo Erba (Italy), Merck (Germany), Sigma (USA) and Riedel de-Häen (Germany) unless stated otherwise in the text. All solutions were prepared in dH_2O unless otherwise stated.

3.3. Buffers and Solutions

3.3.1. DNA Extraction from Whole Blood

Cell Lysis Buffer	:	155 mM NH_4Cl, 10 mM $KHCO_3$, 0.1 mM Na_2EDTA (pH 7.4)
Nucleus Lysis Buffer	:	400 mM NaCl, 2 mM Na_2EDTA, 10 mM Tris (pH 8.2)
Sodiumdodecylsulfate (SDS)	:	10 per cent SDS (w/v)
Proteinase K	:	20 mg/ml Proteinase K
Ammonium Acetate	:	7.5 M CH_3COONH_4
Ethanol	:	Absolute ethanol
TE Buffer	:	1 mM EDTA, 20 mM Tris-HCl (pH 8.0)

3.3.2. Polymerase Chain Reaction (PCR)

10X PCR Buffer A	:	20 mM $MgCl_2$, 500 mM KCl, 100 mM Tris-HCl (pH 8.3)
10X PCR Buffer B	:	20 mM $MgSO_4$, 100 mM KCl, 200 mM Tris-HCl (pH 8.8), 100 mM $(NH_4)_2SO_4$, 1 per cent Triton X-100, 1 mg/ ml BSA
$MgCl_2$	:	25 mM $MgCl_2$ (Roche, Germany)

dNTP	:	12.5 mM each of dATP, dTTP, dCTP and dGTP (Roche, Germany)
Betaine	:	5 M Betaine (Promega, USA)

3.3.3. Agarose Gel Electrophoresis

Agarose	:	2 per cent agarose in 0.5 X TBE buffer
10X TBE Buffer	:	0.89 M Trizma base, 0.89 M boric acid, 20 mM EDTA (pH 8.3)
6X Loading Buffer	:	10 mM Tris-HCl (pH 7.6), 50 per cent glycerol, 60 mM EDTA, 2.5 mg/ml bromophenol blue
Ethidium Bromide	:	10 mg/ml

3.3.4. Denaturing Polyacrylamide Gel Electrophoresis (PAGE)

40 per cent Acrylamide (stock)	:	40 per cent acrylamide-bisacrylamide (19:1)
8 per cent Instagel (denaturing)	:	8 per cent acrylamide-bisacrylamide (19:1), 8.3 M urea in 1X TBE Buffer
APS	:	10 per cent ammonium peroxidisulfate
TEMED	:	N,N,N,N-tetramethylethylenediamine
10X Sample Buffer	:	90 per cent formamide, 20 mM EDTA, 0.05 per cent bromophenol blue, 0.05 per cent xylene cyanol

3.3.5. Single Strand Conformational Polymorphism (SSCP) Gel Electrophoresis

40 per cent Acrylamide (stock)	:	40 per cent acrylamide-bisacrylamide (37.5:1)
8 per cent Acrylamide (non-denaturing)	:	8 per cent acrylamide-bisacrylamide (37.5:1) in 0.6X TBE Buffer
Glycerol	:	5 per cent glycerol in gel solution

3.3.6. Silver Staining

Staining Buffer	:	0.1 per cent $AgNO_3$
Developing Buffer	:	1.5 per cent NaOH, 0.01 per cent $NaBH_4$, 0.015 per cent formaldehyde

3.4. Enzymes

Taq DNA polymerase was purchased from Roche (Germany) or Qiagen (USA) ,supplied with Q-solution. Restriction enzyme *BseGI* and its buffer were supplied by Fermentas (Canada).

3.5. Kits

A list of commercial kits utilized in this study is given in Table 3.2.

Table 3.2. List of kits used in this study, their description and manufacturing company.

Name	Used for	Company
Qiagen Taq PCR Core Kit with Q solution	Amplification of difficult templates	Qiagen, USA
GC-Rich PCR System	Amplification of templates with high GC content	Roche, Germany
Agarose Gel DNA Extraction Kit	Elution of DNA fragments run in agarose gels	Roche, Germany
QIAquick PCR Purification Kit	PCR clean up for sequencing	Qiagen, USA
High Pure PCR Template Preparation Kit	DNA extraction from whole blood	Roche, Germany

3.6. Oligonucleotide Primers

Oligonucleotide primers used in this study were all custom synthesized. They were purchased from Iontek (Turkey) or Massachusetts General Hospital (MGH) DNA Synthesis Core (USA). Lypophilized primers were dissolved in 1000 µl dH_2O, and 10 µM dilutions were used for PCR.

3.7. DNA Molecular Weight Markers

Lambda DNA/*Hind*III and pUC19 DNA/*Msp*I markers were purchased from Fermantas (Canada), and 50 bp DNA ladder was purchased from Roche (Germany).

3.8. Equipment

The equipment used in this study and their models (producers, countries) are listed below.

Autoclave : Midas 55 (Prior Clave, UK)
AMB430T (Astell, UK)

Balance : Electronic Balance (Precisa, Switzerland)

Centrifuges	:	MiniSpin Plus (Eppendorf, Germany) Allegra X-22R (Beckman Coulter, USA) J2-MC (Beckman Coulter, USA) Universal 16R (Hettich, Germany)
Electrophoretic Equipment	:	Horizontal DNA Electrophoresis Gel Box (Bio-Rad, USA) Primo Minicell Horizontal Gel Sytem (Thermo Scientific, USA) Sequi-Gen Sequencing Cell (Bio-Rad, USA) DCode Universal Mutation Detection System (Bio-Rad, USA)
Documentation System	:	GelDoc Documentation System with Quantity One 1-D Analysis Software (BioRad, USA)
Deep Freezers	:	-80^{o}C Ultra Freezer (Thermo Scientific, USA) -20^{o}C (AEG, Turkey) -20^{o}C (Bosch, Germany)
Refrigerator	:	4^{o}C (Arçelik, Turkey)
Incubator	:	Orbital (Gallenkamp, Germany)
Magnetic Stirrer	:	MR3001 (Heidolph, Germany)
Micropipettes	:	Pipetman (Gilson, France)
Minishaker	:	Rotamax 120 (Heidolph, Germany)
Ovens	:	Heraus (Germany)

		EN 400 (Nüve, Turkey)
Power Supplies	:	Power Pac Model 3000 (Bio-Rad, USA) Fotoforce 250 Electrophoresis Power Supply (Fotodyne, USA) P250A Power Supply (Sigma-Aldrich, USA)
Spectrophotometers	:	8453 UV-Visible Spectrophotometer (Agilent, USA) NanoDrop 1000 (Thermo Scientific, USA)
Thermal Cyclers	:	MyCycler (Bio-Rad, USA) PTC-200 (MJ Research, USA) Techne (Progene, UK) LightCycler 480 (Roche, Germany)
Transilluminator	:	Fluorescent Table (Consort, Belgium)
Vortex	:	Reax vortexmixer (Heidolph, Germany) Lab Dancer Vario (Roth, Germany)
Waterbath	:	Grant LTD 6G Thermostatic Water Bath (Grant, Germany)
Water Purification System	:	Ultra Pure Water Purification system (Watech, Germany)

3.9. Electronic Databases

A list of electronix databases utilized in this study is given in Table 3.3.

Table 3.3. List of electronic databases used in this study

Name and Internet Address	**Description**
Ensembl http://www.ensembl.org/	Software system which produces and maintains automatic annotation on selected eukaryotic genomes
NCBI genome resources http://www.ncbi.nlm.nih.gov/genome/guide/human/	Supplies reference sequence and working draft assemblies for a large collection of genomes
Online Mendelian Inheritance in Man http://www.ncbi.nlm.nih.gov/Omim/	A guide to human genes and inherited disorders
RCSB Protein Databank http://www.rcsb.org/pdb/home/home.do/	Tools (including Protein Workbench software) and resources for studying the structures of biological macromolecules
UCSC Genome Browser http://genome.ucsc.edu/	Reference sequence and working draft assemblies for a large collection of genomes
Amino Acid Substitution Prediction Tools	
Molecular Modelling and Bioinformatics (MMB) http://mmb2.pcb.ub.es:8080/PMut/	Annotation and prediction of levels of pathology for mutations. Output prediction score of > 0.5 is pathological, otherwise neutral
MutDB http://www.mutdb.org/	Interactive structural analysis of mutation data
SIFT http://blocks.fhcrc.org/sift/SIFT.html/	Calculates position-specific scores for amino acids, using sequence homology to Output score ranges from 0 to 1, where 0 is damaging and 1 is neutral
SNPs3D http://snps3d.org/	Assigns functional effects of non-synonymous SNPs based on structure and sequence analysis. Output scores of <0 is damaging. Mutation on protein structure can be visualized.
Tools for Detection of Particular Patterns in Sequences	
Biology WorkBench http:// workbench.sdsc. edu/	Web-based package of tools for storing and analyzing input sequences. Includes Primer3, restriction analysis, etc.
Tandem Repeats Finder http://tandem.bu.edu/trf/trf.html/	Locates and displays tandem repeats in a given DNA sequence.

Table 3.3. List of electronic databases used in this study (continued)

Name and Internet Address	**Description**
Others	
Alternative Splicing Structural Genomics Project http://moult.umbi.umd.edu/human2004/	Modeling alternative splicing and its effects on protein folds
Laboratory of Statistical Genetics at Rockefeller University http://linkage.rockefeller.edu/	List of statistical tools for genetic linkage analysis
Mammalian Genotyping Service http://research.marshfieldclinic.org/genetics/	Genotyping maps and statistics
Rutgers map http://compgen.rutgers.edu/maps/	Second generation combined linkage and physical map

4. METHODS

4.1. DNA Extraction from Peripheral Blood Samples

Genomic DNA from patients and their family members were isolated from peripheral blood samples that had been collected into anticoagulant K_2EDTA containing sterile tubes. Thirty ml of cell lysis buffer was added to every 10 ml of blood and kept for 10 minutes at 4°C to lyse the plasma membrane. The samples were then centrifuged at 5000 revolution per minute (rpm) at 4degrees for 10 minutes. The supernatant was discarded, and the leukocyte nuclei containing pellet was washed by suspending in 10 ml of cell lysis buffer and centrifugation for 10 minutes. The supernatant was discarded, and nuclei were suspended by vortexing in 5 ml of nucleus lysis buffer. After the entire pellet had been dissolved, 50 µl of proteinase K (20 mg/ml) and 80 µl of 10 per cent SDS were added and mixed gently. The sample was incubated at 37°C for 14 hours to digest the nuclear proteins. The sample was shaken gently to salt out protein residues after adding 2.8 ml of NH_4Ac. The sample was centrifuged at 10,000 rpm for 25 minutes. The supernatant was transferred to a 50 ml tube. Two volumes of absolute ethanol was added to it to precipitate out DNA. DNA was fished out using a plastic pipette tip and transferred to a microtube. It was first air-dried, then dissolved in 500 µl of TE buffer and stored at -20°C until use.

If the amount of blood sample was small, then DNA was extracted instead using High Pure PCR Template Purification Kit.

4.2. Linkage Analysis

Whole-genome microsatellite scans had been conducted to search for the loci of genes responsible for JP, LRS and CCLH at the National Heart, Lung and Blood Institute (NHLBI) Mammalian Genotyping Service (Contract Number HV48141, Weber and Bronnan, 2001). Marshfield Screening Set 13 was used for CCLH, LRS1, LRS2 and LRS3 families while Set 16 was used for JP and the remaining LRS families. Screening sets used for the genome scan contained 411 and 402 microsatellite markers, respectively, that spanned both the autosomes and sex chromosomes with an average spacing of 10 cM. The

genotyping error rate of the service was 0.50 per cent, as estimated by blindly typing some control DNA samples. The genotyping service required 15-30 µg of DNA from each individual for the amplification of microsatellite loci with fluorescently labeled primers. The amplified products were later genotyped on a custom-built scanning fluorescence detector and allele sizes were assigned. The results were submitted to us in the LINKAGE file format.

Whole-genome single nucleotide polymorphism (SNP) scans were performed to localize the genes responsible for ARA, JP, LRS and CCLH at deCODE (Iceland), Macrogen (South Korea) and Yale University (United States of America). Illumina SNP genome chips used for this study are summarized in Table 4.1. DNA requirement for SNP chips varies from 200 ng to 300 ng, depending on the chip used.

Table 4.1. Illumina SNP genome scan chips, their SNP contents and families genotyped in this study.

Chip Name	Total SNP	CNV (approximate)	SNP spacing in kb (Mean/Median)	Family
Illumina 370Duo	370,405	74,000	7.7 / 5.0	ARA
Illumina 370Quad	373,339	74,000	7.7 / 5.0	JP
Illumina 610Quad	617,366	100,000	4.4 / 2.3	CCLH
Illumina 1MQuad	1,140,419	60,000	2.4 / 1.2	LRS, JP

The raw data obtained from NHLBI Mammalian Genotyping Service were reformatted for the software package easyLINKAGE in order to detect genotyping errors (PedCheck), calculating two (SuperLink v1.6) and multipoint (SimWalk v2.91) lod scores, and constructing haplotypes (Genehunter v2.1). The updated physical and genetic positions of the markers were obtained from the sequence tagged site (STS) map of GenBank (NCBI Build 36.3 and 37.1). The candidate loci deduced after computer analyses were fine-mapped by typing additional microsatellite markers flanking and/or within those loci in our laboratory. All of the statistical analyses were performed under a parametric model of recessive inheritance with full penetrance.

The raw data obtained from SNP genome scan were imported to Progeny Software in order to exclude the markers with Mendelian errors, reformat for HCiE and apply easyLINKAGE to calculate two-point (SuperLink v1.6) and multipoint (SimWalk v2.91) lod scores and construct haplotypes (Genehunter v2.1). All of the statistical analyses were performed under a parametric model of recessive inheritance with full penetrance.

Microsatellite markers that were reported in the databases were primarily selected for the genotyping analyses (Table 4.2). Sequences for primers used for amplification of those polymorphic repeat loci were obtained from GenBank. For regions where no markers had been reported, the DNA sequence of interest was introduced to Tandem Repeats Finder program and primer pairs were designed to amplify the repeats revealed by this program (Table 4.3).

PCR reaction to amplify a marker was carried out in a total volume of 11 µl, consisting of 1X PCR buffer, 0.3 µl of each primer pair, 0.2 mM of each dNTP, 20-100 ng of genomic DNA, 0.15 U Taq DNA polymerase and sufficient dH_2O to adjust the volume. The PCR conditions were as follows: an initial denaturation step at 95°C for 3 minutes, followed by 30 cycles of 30 seconds denaturation at 94°C, 30 seconds annealing at appropriate temperature and 1 minute elongation at 72°C, and a final extension step of 10 minutes at 72°C. In order to avoid shadow band synthesis in dinucleotide repeat amplifications, betaine together with DMSO were included in the PCR reaction, with final concentrations of 1.3 M and 1.3 per cent, respectively. Buffer B was preferred for the amplification of tri- and tetranucleotide repeats but was mostly not used for the amplification of dinucleotide repeats, as it enhances shadow band generation. Alleles of amplified microsatellite markers were resolved by PAGE and visualized after silver staining. Examples for gels showing marker alleles are given in Figure 4.1. Noninformative markers were left out in haplotype constructions.

Table 4.2. List of microsatellite markers used in this study.

ARA				
Chromosome 6				
D6S1622	D6S1659	D6S406		
Chromosome 9				
D9S1788	D9S1817	D9S1845	D9S304	D9S43

Table 4.2. List of microsatellite markers used in this study (continued).

Chromosome 12				
D12S1025	D12S1036	D12S1040	D12S1047	D12S1052
D12S1061	D12S1291	D12S1347	D12S1649	D12S1684
D12S1702	D12S1709	D12S2586	D12S299	D12S303
D12S313	D12S329	D13S335	D12S337	D12S350
D12S371	D12S375	D13S376	D12S80	D12S868
D12S869				
JP				
Chromosome 1				
D1S198	D1S2137	D1S2806	D1S2825	D1S2866
D1S3467	D1S410			
Chromosome 4				
D4S1602	D4S2283	D4S2639	D4S2688	D4S403
Chromosome 6				
D6S1579				
Chromosome 7				
D7S2420	D7S2459	D7S2465	D7S496	
Chromosome 9				
D9S1162	D9S1690	D9S1784	D9S277	D9S53
D9S909				
Chromosome 16				
D16S285	D16S3022	D16S3044	D16S3068	D16S3105
D16S481	D16S490	D16S672	D16S682	D16S685
Chromosome 18				
D18S1009	D18S1092	D18S1095	D18S1115	D18S1122
D18S1125	D18S1131	D18S1157	D18S1161	D18S1269
D18S1361	D18S1367	D18S1374	D18S380	D18S386
D18S461	D18S466	D18S469	D18S485	D18S486
D18S488	D18S50	D18S501	D18S535	D18S541
D18S548	D18S61	D18S70	D18S812	D18S815
D18S826	D18S844	D18S847	D18S848	D18S850
D18S860	D18S865	D18S874	D18S879	D18S880
D18S979	RH4453			
Chromosome 22				
D22S1140	D22S1156	D22S1157	D22S1160	D22S1161
D22S272	D22S282	D22S283	D22S295	D22S417
D22S423	D22S426	D22S534	D22S928	
LRS				
Chromosome 1				
D1S1160	D1S1597	D1S160	D1S4612	D1S1635
D16S1677	D1S1679	D1S1768	D1S2133	D1S2145
D1S2346	D1S243	D1S2586	D1S2635	D16S2675
D1S2697	D1S2707	D2S2768	D1S2844	D1S3372
D1S402	D1S436	D1S468	D1S484	D1S489
D1S508	D1S548	D1S552		
Chromosome 2				
D2S1279	D2S1363	D2S21397	D2S2285	D2S2348
D2S336	D2S338	D2S395		

Table 4.2. List of microsatellite markers used in this study (continued).

Chromosome 3				
D3S1261	D3S1542	D3S1562	D3S1566	D3S1598
D3S1600	D3S1747	D3S2389	D3S2406	D3S2428
D3S2446	D3S3039	D3S3551	D3S3581	D3S3653
D3S4529	D3S4533	G46170		
Chromosome 4				
D4S1530	D4S1552	D4S2967	D4S415	
Chromosome 8				
D8S1130	D8S1145	D8S1729	D8S1836	D8S1925
D8S272	D8S346	D8S639		
Chromosome 9				
D9S1684	D9S1870	D9S921	D9S925	
Chromosome 10				
D10S1146	D10S1694	D10S616		
Chromosome 12				
D12S1660				
Chromosome 16				
D16S2618	D16S2622	D16S3024	D16S3065	16S3070
D16S3072	D16S3082	D16S3084	D16S3134	D16S3392
D16S3395	D16S423	D16S525	TTTA028	
Chromosome 17				
D17S1308	D17S1822	D17S184	D17S1874	D17S668
D17S674	D17S784	D17S836	D17S928	
Chromosome 20				
D20S100	D20S1085	D20S158	D20S430	D20S501
D20S845				
CCLH				
Chromosome 1				
D1S1678	D1S1726	D1S2615	D1S2738	D1S306
D1S373	D1S413	D1S510		
Chromosome 2				
D2S1374	D2S2112	D2S2113	D2S2397	D2S290
D2S291	D2S2977	D2S380	D2S393	D2S443
Chromosome 5				
D5S2094				
Chromosome 6				
D6S1060	D6S1594	D6S1603	D6S1647	D6S1671
D6S1679	D16S1717	D6S1958	D16S2418	D6S261
D6S283	D6S287	D6S301	D6S416	D6S432
D6S434	D6S468	D6S475		
Chromosome 9				
D9S126	D9S156	D9S157	D9S1813	D9S54
DD9S775	D9S939			
Chromosome 10				
D10S1213	D10S1216	D10S1473	D10S1683	D10S1705
D10S1725	D10S218	D10S2473	D10S2486	D10S527
D10S547				
Chromosome 13				
D13S305	D13S325	D13S328	D13S765	D13S893
D13S894				

Table 4.2. List of microsatellite markers used in this study (continued).

Chromosome 16				
D16S3026	D16S3040	D16S3049	D16S3051	D16S3106
D16S507	D16S515	D16S518	D16S684	D16S752
Chromosome 17				
D17S1152	D17S1303	D17S1353	D17S1528	D17S1529
D17S1533	D17S1537	D17S1541	D17S1566	D17S1574
D17S1584	D17S1789	D17S1796	D17S1798	D17S1810
D17S1828	D17S1832	D17S1840	D17S1854	D17S1881
D17S379	D17S513	D17S516	D17S520	D17S525
D17S559	D17S570	D17S578	D17S695	D17S786
D17S796	D17S804	D17S829	D17S906	D17S919
D17S938	D17S945	D17S960	D17S973	D17S974
Chromosome 18				
D18S1099	D18S1102	D18S363	D18S484	D18S487
D18S536	D18S970			

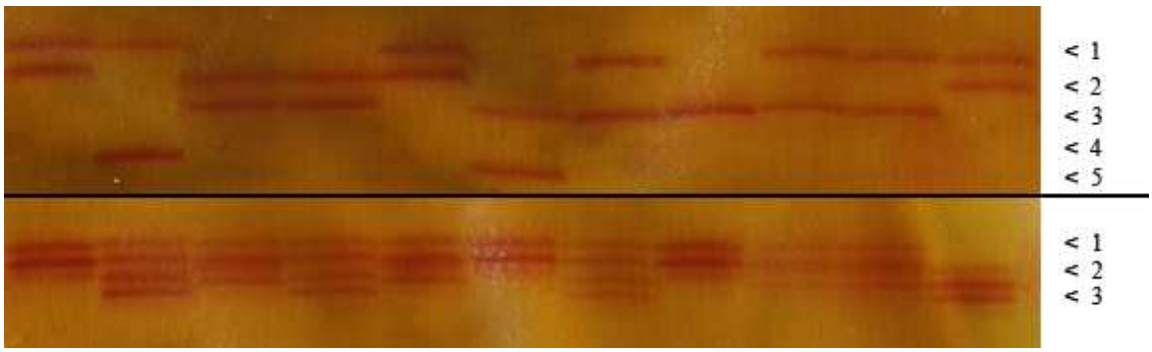


Figure 4.1. Examples of silver-stained denaturing PAGE gels. Resolved alleles of a tetra repeat marker (upper panel, RH4453) and a di repeat marker (lower panel, D18S469) are shown. Arrows indicate allele numbers according to ascending allele length.

Table 4.3. Primer sequences, properties and PCR conditions for microsatellite repeats designed in this study.

Marker	**Primer sequence (5' → 3')**	**Product size (bp)**	**Buffer Choice**	**Annealing Temperature (°C)**
ARA				
D12S74170kb	TGCCAGCAATCACTCCCTAT	289	Buffer A	56
	CCCAAACAACCTGAAAGGAA			
D12S75153kb	AGCCTATGTGACAGGGCAAG	241	Buffer A	55
	GAGGATATGGCTGGAACGAA			

Table 4.3. Primer sequences, properties and PCR conditions for microsatellite repeats designed in this study (continued).

Marker	Primer sequence (5' → 3')	Product size (bp)	Buffer Choice	Annealing Temperature (°C)
LRS				
D1S15988kb	GGCAACAAGAGCAAAACTCC	284	Buffer A	55
	TCATCCAGGTAACACCCCTC			
D1S160867kb	CCATGTTCTCCGTGTCCTTT	230	Buffer B	54
	GGGGCAGGAGATGTTTCATA			
D1S160921kb	ATTCAAACCAACCACCCAAA	186	Buffer B	56
	GGTAGCAATGGTGCTAAGGG			
D2S238234kb	AGGTCACAAAGGCTCT	258	Buffer A	54
	TTGCCCTCCAAGTCCTTCC			
D3S73924kb	GAATGTGGGACTTCTCTGCC	217	Buffer B	55
	GGCCTTACCCTGCAGAAATA			
D3S73988kb	TCATTGTGCTTGCTCTGGAC	212	Buffer A	55
	GGGAGGGAGGAAGAGAGAGA			
D3S74104kb	CTCACCTCACCTCCCAAAAA	215	Buffer A	55
	TATCTCAAATGGCTGCATGG			
D3S74771kb	ATTCGGCACACTCCTTGACT	299	Buffer A	56
	CTCCAGTTGCTCGGCATTAC			
D10S73486kb	GATCTCACCACTGCACTCCA	229	Buffer A	56
	CAGAGAGGTGAGGGGGTGTA			
D16S1715kb	CATTGAAGTGACAGCGCCTA	236	Buffer A	57
	ACGGCAGTATCAACCTCAGC			
D17S73820kb	TGCTAAAGCAGCATGAATGG	259	Buffer A	55
	CCAGAAGAGGAGTGAGCACC			
D17S74005kb	TCAGTGAACCGAGACTGTGC	197	Buffer A	54
	ATAGACCAGGAGCATGTGGG			
D17S7411m	GGAGAAGCTCTTTAGCGGGT	256	Buffer B	55
	CAGCAGGTCACAGATGCAGT			
D20S55048kb	CCCAAGTTCTTGCTCCACAT	279	Buffer A	55
	GCTCACATCCTGTCGTCTCA			
D20S55141kb	ACCCACTGACTTGGATCCTG	246	Buffer A	54
	CCCAGCCTATGTCGCTACTC			
D20S55241kb	AACATTTGTGGGCTCTTTGG	311	Buffer A	56
	TCACACTGTGGAGTGGCTTC			
D20S55278kb	ACAAACCCCCATTTAGCCTT	241	Buffer B	54
	CTCAAAGCCTTTTCTGGTGG			
D20S55980kb	TCAAAGCACCAGCACTTTTG	387	Buffer A	55
	GCTGACATTCGATGGATGG			
D20S55994kb	GGTGTGGTGCAATAGGCTTT	179	Buffer A	54
	TATTTTCTTCGGCCAGATCC			
CCLH				
D13S30977kb	TTTGGGAATTTCAAAGCTGC	188	Buffer B	56
	GCCTGGGAGACAAAGTGAGA			
D13S35923kb	TTCCAATGGATTCCAATGA	235	Buffer A	56
	TGAAACCGGTACCGGCCTTG			
D17S14870kb	TGCCACAGAACAAAGAGTGG	281	Buffer A	56
	AGCGAGACTCTGTCCCAAAA			

Table 4.3. Primer sequences, properties and PCR conditions for microsatellite repeats designed in this study (continued).

Marker	Primer sequence (5' → 3')	Product size (bp)	Buffer Choice	Annealing Temperature (°C)
D17S2273kb	CATGCAGACACCACAGGTTC	370	Buffer A	56
	GCAAAGCGTTAGGCTCTCAG			
D17S4318kb	AGGGGTTTTTGTTTTGGAGG	264	Buffer B	57
	GTTGCAGTGAGCTGAGACCA			
D17S4470kb	GAGAGAGAGAGGGAGGGAGG	325	Buffer A	54
	CAGGAGTAAGACGGGGTGAG			
D17S5470kb	CCAGTTTTCTTGGGGAAAAA	365	Buffer A	54
	CTTGAGGCTGGGAGTTTGAG			

4.2.1. Denaturing Polyacrylamide Gels

Sequi-Gen GT (BioRad, USA) nucleic acid electrophoresis system assembled with 0.4 mm spacers was used for resolution of marker alleles in fine mapping studies. Denaturing instagel was used for this purpose, which was initially pre-run for 20 min in order to facilitate the gel temperature to rise to 45-50°C. Samples to be subjected to electrophoresis were mixed in an equal ratio with 10X Stop buffer, denatured at 95°C for five minutes and chilled on ice before loading onto individual slots of a sharks tooth comb. Table 4.4 shows the systems used with required amounts of instagel, APS and TEMED and the power applied to run the gels.

Table 4.4. Gel electrophoresis systems used for genotyping.

Gel Cast size	Instagel	10 per cent APS	TEMED	Power
21 x 40 cm	35 ml	250 µl	25 µl	35 W
38 x 40 cm	50 ml	350 µl	35 µl	55 W

4.2.2. Silver Staining

Glass plates were separated so as to allow the gel to remain intact on one of them. The gel was then transferred to staining buffer by first placing a piece of filter paper on it and then separating it from the glass plate together with the filter paper. It was soaked in staining buffer for 10 min. The gel was subsequently incubated in developing buffer until

the bands appeared (Kavaslar *et al.*, 2000). When staining reaction was inadequate, the procedure was repeated after extensive washing with dH_2O in between.

4.3. Candidate Gene Approach

The mutation detection system used in this study was Single Strand Conformational Polymorphism (SSCP).

4.3.1. PCR Amplifications for the Analysis of Candidate Genes

In order to search forsequence variations in a candidate gene, the parts to be analyzed were first amplified by PCR. Intronic exon-flanking primers were designed by Primer 3 software to amplify the exons and associated splice junctions. Primer specificity was checked with in silico PCR at UCSC database. Unless otherwise stated in the text, PCR reactions were carried out in a total volume of 25 µl, consisting of 1X PCR buffer, 0.2 mM of each dNTP, 400 nM of each primer pair, 20-50 ng of genomic DNA, 0.2 U of Taq DNA polymerase and sufficient dH_2O to adjust the volume. The cycling conditions for amplifications were as follows: an initial denaturation step at 95°C for 3 minutes, followed by 30 cycles of 30 seconds denaturation at 94°C, 30 seconds annealing at appropriate temperature and 1 minute elongation at 72°C, and a final extension step for 10 minutes at 72°C. Some short exons of *APTX* and *PDE4B* were screened for mutations first by SSCP. All fragments for candidate genes were sequenced, irrespective of observing an aberrant SSCP pattern or not. The primer sequences and PCR conditions optimized for the candidate genes are given in Table 4.5.

APTX, *ATXN7*, *PDE4B*, *AK3L1*, *PRKAR2B*, *GAL3ST2*, *SOCS3*, *TIMP2* and *CANT1* were analyzed in relevant patients using this procedure.

4.3.2. Analysis of PCR Products

To check the extent of the amplification and whether any nonspecific fragments have been amplified as well, five µl aliquot of the PCR product was mixed with one µl of 6X loading buffer, loaded on two per cent agarose gel containing 10 µg ethidium bromide, and

subjected to electrophoresis in 0.5X TBE at 130 volts for 10 minutes. The bands were subsequently visualized over a UV light transilluminator.

4.3.3. SSCP Analysis

Selected exons of *APTX* and *PDE4B* were initially screened by SSCP using the PCR products of three individuals: an affected individual, one of his/her parents, and an unrelated individual randomly chosen from the population. The analyses were carried out on nondenaturing gel matrices with crosslinking ratios of eight per cent and with and without five per cent glycerol. Whenever an aberrant SSCP pattern was observed in the affected individual, that region of the gene was analyzed by SSCP also in other members of the family as well as in at least 55 unrelated individuals, in order to assign the sequence variation as a mutation responsible for the relevant disorder or just a polymorphism. Furthermore, the region was subjected to DNA sequence analysis to determine the nature of the variation.

4.3.4. SSCP Gel Electrophoresis

The gel plates were 20 X 20 cm in size and assembled using 0.75 mm spacers. Six ml of 40 per cent acrylamide stock was mixed with 1.8 ml of 10X TBE, and the volume was adjusted to 30 ml by dH_2O. Three hundred µl of 10 per cent APS and 30 µl of TEMED were added, and the solution was poured between the glass plates. A 15 or 20-well comb was inserted, and the gel was left to polymerize for at least an hour.

Electrophoresis was carried out in 0.6X TBE buffer at 4°C. Samples were mixed in equal ratios with 10X Stop buffer, denatured at 95°C for 5 minutes and immediately chilled on ice. Eight µl of each sample mix was loaded and subjected to electrophoresis in BioRad DCode Universal Mutation detection System at a constant power of 10 W for 10 to 15 hours, depending on the length of the products. Subsequently, the gel was silver stained to visualize the samples.

4.3.5. DNA Sequence Analysis

PCR-amplified fragments were purified from primers, nucleotides, polymerase and salts using QIAquick PCR cleanup columns (QIAGEN) and were sequenced on an ABI PRISM Genetic Analyzer (Applied Biosystems) at Iontek (Istanbul, Turkey), DNA Oligonucleotide Synthesis Core (Cambridge, USA) or Macrogen (Seoul, South Korea).

4.3.6. Restriction Enzyme Analysis

APTX exon 8 was screened for the presence of restriction enzyme *BseGI* cutting site in ARA Family 2 members and population controls. *BseGI* enzyme recognizes sequence 5' GGATG 3' and cleaves DNA at 5' GGATGNN^ 3'. *BseGI* digestion was carried out in a total volume of 10 µl, consisting of 3.3 µl of amplified product, 1X Tango restriction buffer, 1 U *BseGI* and sufficient dH_2O to adjust the volume. The digestions were incubated for 14 hours at 55°C, run on a 2 per cent agarose gel and visualized under UV light.

Table 4.5. Sequences, PCR product sizes and PCR conditions for primers designed for candidate gene analysis.

Exon	Primer sequence (5' → 3')	Product size (bp)	Buffer Choice	Annealing Temperature (°C)
ARA				
APTX				
1-1	GCTGAAACAGCCCAAGGTAG	393	-	-
	TCCAGGATCCTCAGTTGGAC			
1-2	CCAAGGACTGCTTCTTCTGG	640	Q Solution	55
	TCCAGGATCCTCAGTTGGAC			
2	Not analyzed, no expression	-	-	-
3	CCCTGGTGTTGGCTAAAAGA	477	Buffer A	56
	TTTTCACAACAGCCTCCTGA			
4	GAGTCAATGCTGGTGGGATT	407	Buffer A	56
	CCGCCTGTAGTGAAGCTTGT			
5	ATCCAAATGATTTTCACGCC	565	Buffer A	57
	ACATAAGGCAAGGTGCTTGG			
6	ATTGCCAATGTGAAAAGCCT	400	Buffer B	56
	AAGGCAATGGAGTGAAGCAG			
7	GTCAGGCAGAGAGGTGGAAG	494	Buffer B	56
	CATTTGCACAAGACTGTGGG			
8	CACAGGGCTCGCTTTATAGG	479	Buffer A	56
	AAAGTCTCAGTGCCTGGAACA			

Table 4.5. Sequences, PCR product sizes and PCR conditions for primers designed for candidate gene analysis (continued).

Exon	Primer sequence (5' → 3')	Product size (bp)	Buffer Choice	Annealing Temperature (°C)
APTX				
9	CTGAAGATGGTGGGTCAGGT	688	Buffer A	58
	GCCAACCAGAGGTAAAGCAC			
ATXN7				
1	GACTAGGGGTACCGAGGGG	790	GC-rich PCR Kit	57
	GGGGAAAACAGGGAGAATGT			
2	GGGCAGGTAAAACCAACAGA	416	Buffer A	56
	ATCCTAGCATTTTGGGAGGC			
3	CCCAGGAGGCTAACTCTGTG	604	GC-rich PCR kit	56
	AACAAATTTGGAGGGGATGTT			
4	TCCCATTCATGCTTTTGACA	454	Q Solution	57
	CTTCTGCTCAGAGGACCTGG			
5-1	GGAGTCGAAAGCGAAAGCTA	676	Buffer A	55
	CCCAGCAAACGATAGCATCT			
5-2	GTCCTCTGAGCAGAAGCAGG	944	Buffer A	55
	GACTCACCTTCCCGACAGAG			
6	GATGCTATCGTTTGCTGGGT	287	Buffer B	56
	GTATGGTGGGCTGATGTTCC			
7	TGGTGGTGTACCACATCAGTT	363	Buffer A	57
	AAAGCCTACTTCCTGGTGGG			
8	GTCTGTGGGGCCTAGATGAG	499	Buffer A	56
	GGCACTCTCACATGCCACTA			
9	CTCTGGCTCACCATTCACCT	410	Buffer A	58
	ATTGCAGTTGATTTGCTCCC			
10	ATGTTGCTGCTTTGTTGCAG	316	Q Solution	58
	ACATCCCAGTCACCCATGTT			
11	GGGAGTGGTGTTTTGGGATA	460	Buffer A	56
	TCTCCACAGTGAGCCTGAGA			
12	GTGCTAGAGGTGGAACCCAG	441	Buffer B	56
	TCTCTCACACCTTGACCGC			
13	CTTCACAGCCTCCGGTACTC	390	Buffer A	55
	CAAGGCACCTTCTTCTCAGG			
14a	TGAATGGCTGTGGTCAGTGT	790	Q Solution	58
	GAATGAATGGCCCAAGAGAA			
14b	CCCCACTGTTGGTTCACTCT	621	Q Solution	58
	ACCAACCCAACCAAACACAT			
15	AGTGGCTAAGTGGTTGTGGG	281	Buffer A	56
	AAGACGAGGAAGGGGAAAAA			
16	GCTGACTGTTCCTGGTGTCA	334	Buffer B	56
	AGCTTCAGGGAAATCCTCGT			
JP				
PDE4B				
1	GGGGTTCAAACCAGGATTTT	326	Buffer A	55
	AACAATGTGACCGTATCGCA			
2	TTGGCTTGGGAAATTATGATG	346	Buffer A	56
	AGTGGTGTGTTAAATGCCCC			

Table 4.5. Sequences, PCR product sizes and PCR conditions for primers designed for candidate gene analysis (continued).

Exon	Primer sequence (5' → 3')	Product size (bp)	Buffer Choice	Annealing Temperature (°C)
3	TTCTTTCCCTGTGGTTCCTC	396	Buffer A	56
	GGGAAACGAATTCTCAAAAT			
4	TTTGTTTTGGATGGTGAAAGC	421	Buffer A	57
	TTCATCTGACCTGGTCTCCA			
5	GCAAAGCCAGCCTGATAAAG	306	Q Solution	58
	TCTGACCTGGTCTCCATTCA			
6-8	CCACACGTTGGTTCTCAGTG	696	Q Solution	58
	AACTGGAAGCCAGACAGAGC			
9	GTGGAATGGGCAGTTTTGAT	220	Buffer A	56
	TTTGAGAGCCACAATATGCAA			
10	ACCAAAGGCTCTTTCCAGGT	386	Buffer B	56
	AGGCATCGTTAGGTGTCAGG			
11	GCAGAGCACCACTGTGATTT	399	Buffer A	57
	GTTGGTAGCATTCCCTGGTG			
12	TCCAACCTCTTGATTCCTGG	299	Q Solution	57
	TGACTCAACTCCATCCCAAA			
13	TTGAATATTGCAGTGGATTCT	432	Buffer A	56
	CAACCTTCCTCAAATGACAGC			
14	ACCCGACCTCCTAGGTCATT	298	Buffer A	56
	TGGAAAGCCAGTATTCACTCC			
15	CCAATTGGATGGCTCTTTATG	381	Buffer A	56
	GCCTCCCGAGTAGCTAGGAC			
16	CAATCCCTCTCCCAACAAGA	328	Buffer A	57
	GAAATGGAAGGATTGCCTCA			
17	TGTAGAGGGCCACTGCTTCT	357	Buffer B	57
	GAGAGAATGGCAGTAAGTGA			
18	CAGCGTCAGACACTCAGGAA	300	Buffer B	58
	GCTTAGTTTGGTTTGGTTTTGT			
19	CATGAGATGTTCCGGCAGTA	381	Buffer B	56
	TGGAACAAAATGCGTAGTTGT			
20	GAGTGCCCAAATCATGGAAT	664	Buffer B	55
	TTGCTTCCTGGTCTTGCTTT			
AK3L1				
1	TTCCCTTCCCTTGCAGTAGA	358	Buffer A	57
	GCCACCAAAAGCAATCAAAT			
2	ATTTGATTGCTTTTGGTGGC	418	Buffer A	57
	CCTCCAGGGGTCAGAGGT			
3	GCTTCAGCAGCTTTGCGT	558	Buffer A	57
	GTACGAGCCTGCACGACC			
4	TTCTTCCCACATCAGGAAGG	501	Buffer A	56
	GGATGGCTTACGGAAACAAA			
5	TGCAGGTTTGTGACATGGAT	529	Buffer A	56
	AAATGTCCCTTCCCTGACAA			
6	TGCATGCAATTTTAGGCTGA	572	Buffer A	56
	CAGCTCCAGCATTTCATCTG			

Table 4.5. Sequences, PCR product sizes and PCR conditions for primers designed for candidate gene analysis (continued).

Exon	Primer sequence (5' → 3')	Product size (bp)	Buffer Choice	Annealing Temperature (°C)
AK3L1				
7	GCCATTGTATGTTGGGCTTT	390	Buffer A	56
	CTTGTGATGGGCTTGAGACA			
PRKAR2B				
1	GGAGGTTGCCATGGTTTC	807	GC-rich PCR Kit	57
	GATGGGAAGGTAGGGAAAGG			
2	TGAGCCTTTTCAGGGTTTTG	390	Buffer A	56
	TGGGCTGAGAGTCTCCATTT			
3	GGGGGTAGACTGCTTCTCCT	379	Buffer A	57
	TCTGCAAGACTCACAGCAGG			
4	TGAATCTCAGCACCAACCTG	318	Buffer A	57
	AGGGCAAAACTCAAGACCCT			
5	TCCATTTCTTTGGACTTGTTCA	382	Buffer A	56
	CCACTGTGGTTTCACTGCAT			
6	GGGGATAACAGGCTGGAAAT	344	Buffer A	56
	CATGCGGCAATTTGTAACAC			
7	TTCAATCATTGTTTCTTGTAGG	551	Buffer B	56
	AGGAAGAGGCCAGAGAGGAC			
8	CTGCTCGATTTATTTGGCCT	241	Buffer A	57
	CCACTACAGGAACCTTTTGGA			
9-10	CAGTCTGGGCAACAGAGTGA	655	Buffer A	57
	ACCACAAAACCTGTGAACCC			
11	TGCAGAGTGTGGTCCAAGAG	744	Buffer B	56
	TTTAATTGCCAAGTGCCCTC			
LRS				
SOCS3				
1	AGCCTTTCTCTGCTGCGAGT	640	Buffer A	56
	GTAGGGAGGGGACAGGAGAA			
2a	TCCTGGAGACCTAACTTCCG	790	Buffer A	56
	CGGAGGAGGGTTCAGTAGGT			
2b	CAGCTCCAAGAGCGAGTACC	700	Buffer A	56
	CTCTCCTGGTTGGCTTCTTG			
TIMP2				
1	GGGAGGAGGAGCAGAAAATC	635	Buffer A	56
	ACTTGCAGGATTCGAGAAGG			
2	GCGTGTTTCAGGAACTCTCC	295	Q Solution	58.4
	TCGGATCTATTTGCAGCTTG			
3	ACTTCCTGGTCCTCTTGGGT	298	Buffer A	57
	TAGGAACAGCCCCACTTCTG			
4	ACTTTTCCCTCTCCCTGAGC	336	Buffer B	58.4
	CCAGCCAGACAGCCTTACTC			
5	TCCCTGCACATGACAGGTAG	399	Buffer A	56
	TTTCCAGGAAGGGATGTCAG			
GAL3ST2				
1	CAGACTCCCTCAGCCAGTTC	423	Buffer A	57
	TCAGGGAAGCCTGAGTAGGA			

Table 4.5. Sequences, PCR product sizes and PCR conditions for primers designed for candidate gene analysis (continued).

Exon	Primer sequence (5' → 3')	Product size (bp)	Buffer Choice	Annealing Temperature (°C)
2	CGTTACAGAGGTGAGGGGAG	334	GC-rich PCR Kit	56
	TCCTCTTCCAGAGGGAGACA			
3	GCATGGTTACGGGAGACAGT	586	Buffer A	56
	ACTCCTTCTCTCCAGCCTCC			
4	CTCGGCAGATGTGAGGATG	1108	Buffer A	57
	GTGCTCTTCACTGGGGTCTC			
CANT1				
1	AAACACAAACCACAGCTCCC	495	GC-rich PCR Kit	55
	TAATGGTCCCCACGCTAGTC			
2	GGATGAGGTGGCATTTGAGT	1270	GC-rich PCR Kit	56
	CAGTAGGCTATGCTCCTGGC			
3	TGGGGTTACAGCTTGAGGAG	696	GC-rich PCR Kit	56
	TGGTGGTGCACACCTGTAGT			
4	ACGAGAACTGGGTGTCCAAC	1124	GC-rich PCR Kit	56
	AAATGCAGGAGGAAGCAGGT			

5. RESULTS

Gene localization and candidate gene approach studies performed on four diseases are summarized below.

5.1. Autosomal Recessive Ataxia

In this study 9 patients and 29 healthy individuals from two families were investigated. Family pedigrees are given in Figure 3.1 and Figure 3.2. The initial diagnosis for the patients was Ataxia-Telangiectasia variant.

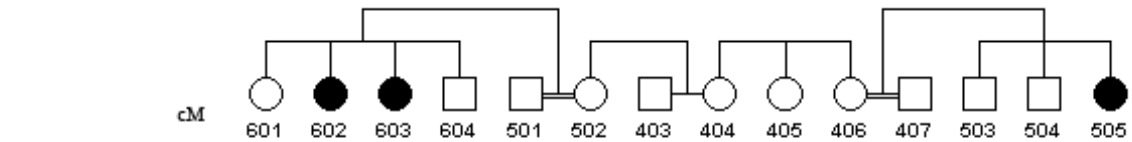
5.1.1. Analysis of Locus 12q21.1-21.2 in Family 1

Locus 12q21.1-21.2 had been determined to harbor the disease gene by Suna Önengüt and Sibel Aylin Uğur in our laboratory. They had narrowed down the locus to a 6.96-Mb region between markers D12S1722 and D12S75153kb by haplotype investigation.

In this study, the family was genotyped again with six of the previously used microsatellite markers, namely, D12S868, D12S1291, D12S375, D12S1036, D12S1025 and D12S80, to verify the linkage. The region was also investigated additionally with D12S329, D12S1649, D12S869 and D12S1061; however, it could not be narrowed down. A maximum two-point lod score of 1.60 at D12S375, D12S80, D12S1047 and D12S869, and multipoint lod score of 2.04 at D12S376 were obtained.

We were informed in March 2008 that a cousin of the patients (1-505 in Figure 3.1) was found to present the disease symptoms. Blood samples from the new patient and her family were obtained. The new part of Family 1 was investigated for linkage to the candidate locus by genotyping with previously employed markers D12S1649, D12S1702, D12S868, D12S1291, D12S1025, D12S80, D12S299, D12S1047, D12S303, D12S376, D12S869, D12S1052, D12S74170kb, D12S337, D12S75153kb, D12S1061 and D12S326.

Haplotypes were constructed using the data generated. Patient 1-505 was found heterozygous at the locus (Figure 5.1).



	cM	601	602	603	604	501	502	403	404	405	406	407	503	504	505
D12S1726	73.28	0 0	1 2	0 2	1 0	0 1	2 0	2 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0
D12S329	74.58	1 2	2 1	1 1	2 2	1 2	1 2	1 3	2 1	0 0	2 0	0 0	0 2	0 2	0 2
D12S1649	76.36	2 3	3 1	2 1	3 3	2 3	1 3	1 4	3 3	3 5	3 5	3 1	1 3	3 3	1 3
D12S1702	77.25	4 3	1 1	1 1	1 3	4 1	1 3	1 1	3 2	1 3	3 3	4 1	4 3	4 3	1 3
D12S868	77.64	1 2	2 2	2 2	2 2	1 2	2 2	2 2	2 2	2 1	2 1	1 2	1 2	1 1	2 2
D12S1291	78.12	2 3	3 4	3 4	3 3	2 3	4 3	4 3	3 1	3 3	3 3	2 4	2 3	2 3	4 3
D12S56	78.13	1 2	1 3	1 3	1 2	1 1	3 2	3 0	2 0	0 0	2 0	0 0	0 2	0 0	0 2
D12S335	78.14	0 0	1 2	1 2	1 0	0 1	2 0	2 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0
D12S313	79.93	1 0	1 2	1 2	0 0	0 1	2 0	2 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0
D12S375	80.52	2 1	2 2	2 2	3 1	3 2	2 1	2 3	1 3	0 0	1 0	0 0	0 0	0 0	0 1
D12S1036	80.53	2 1	2 3	2 3	1 1	1 2	3 1	3 0	1 0	0 0	1 0	0 0	0 0	0 0	0 1
D12S1043	81.16	1 1	1 3	1 3	2 1	2 1	3 1	3 2	1 1	0 0	1 0	0 0	0 0	0 0	0 1
D12S1722	82.02	1 1	1 2	1 2	3 1	3 1	2 1	2 0	1 0	0 0	1 0	0 0	0 0	0 0	0 1
D12S1025	82.09	2 3	2 2	2 2	2 3	2 2	2 3	2 4	3 1	2 2	3 2	3 2	3 2	3 2	3 3
D12S80	83.18	4 1	2 2	2 2	2 1	4 2	2 1	2 4	1 4	2 4	1 4	3 2	3 4	3 4	3 1
D12S299	83.21	1 1	1 1	1 1	1 2	4 1	1 2	1 2	2 4	1 2	2 2	3 1	3 2	3 2	3 2
D12S1047	83.29	1 1	2 2	2 2	2 1	1 2	2 1	2 1	1 1	2 1	1 1	1 2	1 1	1 1	1 1
D12S303	83.41	1 2	1 1	1 1	3 2	3 1	1 2	1 2	2 1	2 2	2 2	2 1	2 2	2 2	2 2
D12S828	83.54	2 1	2 2	2 2	1 1	1 2	2 1	2 2	1 3	0 0	1 0	0 0	0 0	0 0	0 1
D12S376	83.73	1 2	1 1	1 1	1 2	3 1	1 2	1 1	2 2	2 3	2 3	2 1	2 3	2 3	2 2
D12S869	83.79	2 1	3 3	3 3	3 1	2 3	3 1	3 2	1 2	3 2	1 2	2 3	2 2	2 2	2 1
D12S1052	83.82	3 2	2 2	2 2	3 2	3 2	2 2	2 2	2 4	2 1	2 1	2 2	2 1	2 1	2 2
D12S74170kb	84.55	1 1	1 1	1 1	1 1	1 1	1 1	1 3	1 1	3 2	1 2	3 1	3 2	3 2	3 1
D12S337	84.81	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 2	1 2	1 2	1 1	1 2	1 2	1 1
D12S75153kb	85.08	2 1	3 1	2 3	3 1	2 3	3 1	3 2	1 3	3 1	1 1	3 3	3 1	3 1	3 1
D12S1061	86.21	3 3	3 3	3 3	2 3	3 2	3 3	3 1	3 2	3 4	3 4	3 3	3 4	3 4	3 4
D12S326	86.41	2 3	2 3	2 2	2 2	2 2	2 3	2 0	3 1	3 1	3 1	2 2	2 1	2 1	2 1
D12S1679	153.19	1 0	1 2	1 2	0 2	1 0	2 0	2 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0

Figure 5.1. Haplotypes of Ataxia Family 1 at 12q21.1-21.2.

5.1.2. SNP Genome Scan of Autosomal Recessive Ataxia Families

SNP and CNV data generated for the available members of Family 1 and Family 2 were used for lod score analyses. We were informed that there was another patient in Family 2 (2-610 in Figure 3.2) in March 2010 so it was not included in the calculations in this section. The SNP genome scan data and related studies for the new patient were carried on in spring 2010. Multipoint lod scores in ataxia families were calculated assuming autosomal recessive inheritance with full penetrance (Figure 5.2). Due to the huge amount of SNP data and the internal restrictions of linkage programs regarding family size and complexity arising from consanguinity, the families were divided into nuclear families for the calculations, and the results were summed to obtain the final lod scores. Positive multipoint lod scores were obtained only at 9p21.1-13.3, between

rs10491884 and rs942480 (28,562,149 bp and 36,005,428 bp, respectively), with the highest value 4.49 at rs943926 (31,873,011 bp). However, haplotype analysis revealed that 2-504 and 2-609 were heterozygous at the region, indicating that the region was not a candidate locus for all of the patients. Thus, a shared candidate locus could not be identified. Two-point and multipoint lod scores on pseudoautosomal regions (PARs) were calculated separately, and no significant locus was identified.

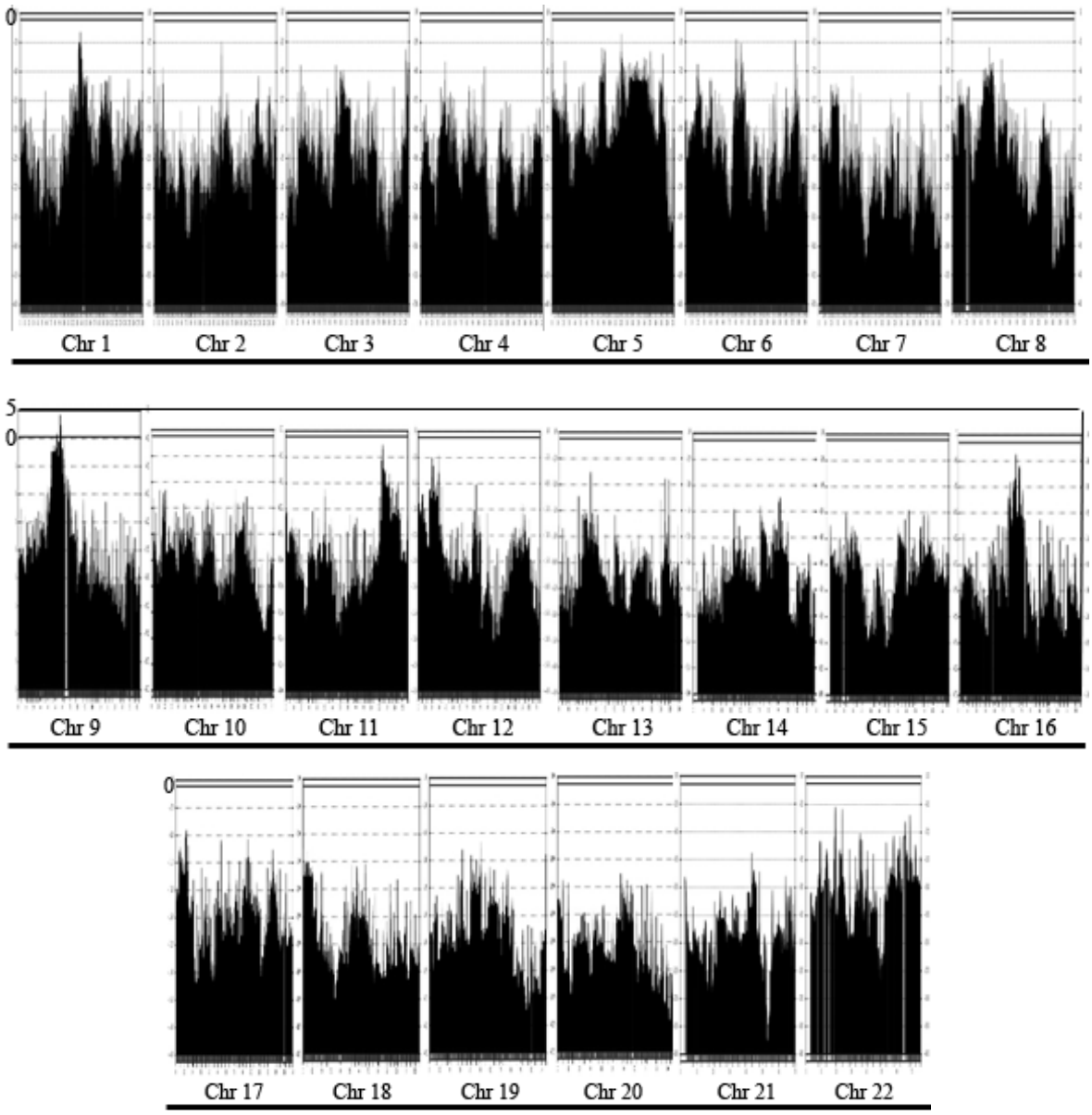


Figure 5.2. Total multipoint lod score graphics for SNP genome scan data for ataxia families 1 and 2 together. An autosomal recessive inheritance was assumed, and 2-610 was not included.

Initial genome scan data was also analyzed to search for homozygosity regions shared by all patients by using program Homozygosity Comparison in Excel (HCiE). The aim was to identify all homozygous loci for each patient and compare his/her genotype with those of other family members and other patients. No region where all patients shared a homozygous genotype was found; thus, no candidate locus for all patients together was identified by this method, either.

Failure to find a candidate locus for all patients together indicated that possibly not all patients were afflicted with the same disease. Accordingly, the patients were grouped by their clinical features. Group 1 was composed of the three patients in Family 1 and 2-601, 2-605 and 2-608 in Family 2. Patients 2-504 and 2-609 in Family 2 were considered as Group 2. Patient 2-610 was added to Group 2 in spring 2010. Each group was separately subjected to multipoint lod score analysis and homozygosity comparison.

Homozygosity comparison in Group 1 using HCiE resulted in the identification of a single strong candidate locus at 9p21.1-p13.3, the same locus at which the only positive total multipoint lod scores were obtained for all patients together (Figure 5.3). Shared maximum homozygosity was between rs10511886 and rs10971505 (31,836,555 bp and 33,429,322 bp, respectively) for patients in Family 1 and between rs2383208 and rs7047872 (22,132,076 bp and 37,406,391 bp, respectively) for patients 2-601, 2-605 and 2-608 in Family 2. The maximum total multipoint lod score for Group 1 was 7.5 (rs943926, 31,873,011 bp). Haplotypes were constructed, and the patients in Family 1 were found to share the same homozygous genotype, while the three patients in Family 2 shared another homozygous haplotype. The databases reported 23 genes in the 1.59-Mb region between rs10511886 and rs10971505. The strongest candidate gene was *APTX* (32,972,608 bp – 33,001,626 bp), known to be responsible for early-onset ataxia with oculomotor apraxia and hypoalbuminemia 1 (AOA1, MIM 208920, Date *et al.*, 2001). All patients in Group 1 had been diagnosed with early-onset ataxia but oculomotor apraxia was observed only in Family 2 patients.

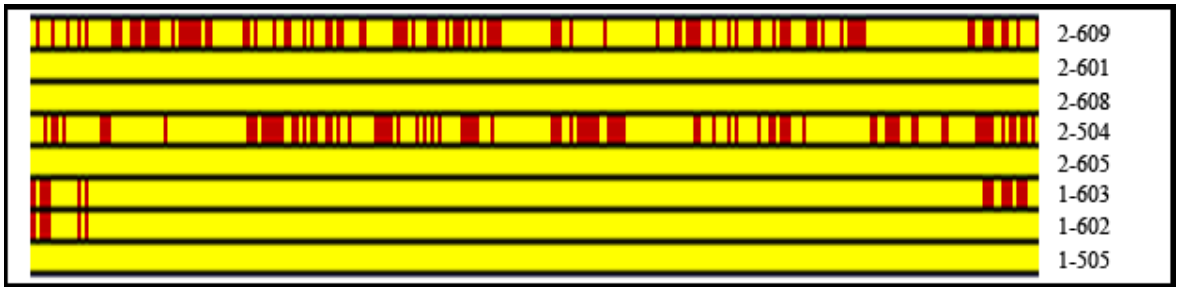
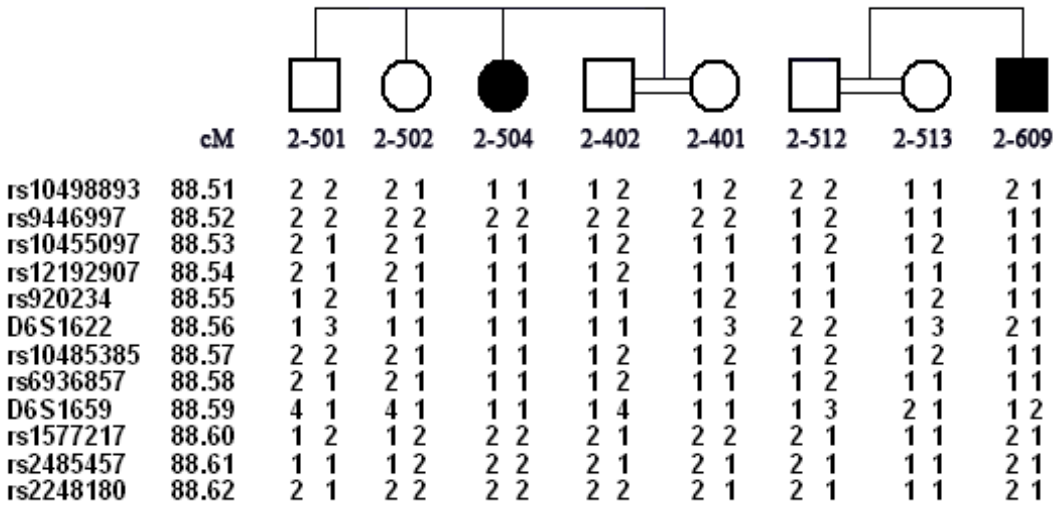


Figure 5.3. HCiE for ataxia patients at 9p21.1-p13.3 between rs2044401 and rs3739680. Red bars represent heterozygosity. Data for individual 2-610 are not shown.

HCiE in Group 2 patients failed to give a strong candidate locus. High bit number due to the distant relationship between the two patients and the high number of consanguineous marriages in the pedigree gave rise to problems in calculating lod scores using linkage programs. To overcome the difficulty posed by the high bit number, the pedigree was divided into two to form two pedigrees that were suitable for processing with linkage software programs. One pedigree included two patients and their parents, who were first cousins. The second pedigree included only Patient 2-504 and her family. The strategy was to compare the lod score graphics in search of a common, previously unnoticed shared homozygosity. With this approach, three new candidate loci on chromosomes 3, 6 and 10 were determined. The maximum multipoint lod scores for those loci were 4.19 for the pedigree with two patients and 2.17 for the second one. The next step for verifying the loci was to prepare haplotypes to investigate whether the genotypes were the same, and, if they were, to delineate the common homozygosity regions. Regions where patients shared homozygous common genotypes were determined as 3p14.1 (rs11919532 – rs1835898; 63,772,361 bp – 64,121,290 bp), 6q13 (rs9446997 – rs2485457; 74,432,663 bp – 75,451,413 bp) and 10q25.3 (rs4751902 – rs10787548; 116,177,976 bp – 116,658,106 bp). 6p13 was studied with two microsatellite markers, and patient 2-609 was found heterozygous at the locus, excluding the locus as harboring the disease gene (Figure 5.4). At 10q25.3, none of the 5 genes were assessed as having a possible causative role for the disease. At 3p14.1 *ATXN7* (63,850,233 bp – 63,989,138 bp) was selected for sequencing (Figure 5.5). Patient 2-610 was found heterozygous with SNP genome scan data at the region in spring 2010.



	cM	2-501	2-502	2-504	2-402	2-401	2-512	2-513	2-609
rs10498893	88.51	2 2	2 1	1 1	1 2	1 2	2 2	1 1	2 1
rs9446997	88.52	2 2	2 2	2 2	2 2	2 2	1 2	1 1	1 1
rs10455097	88.53	2 1	2 1	1 1	1 2	1 1	1 2	1 2	1 1
rs12192907	88.54	2 1	2 1	1 1	1 2	1 1	1 1	1 1	1 1
rs920234	88.55	1 2	1 1	1 1	1 1	1 2	1 1	1 2	1 1
D6S1622	88.56	1 3	1 1	1 1	1 1	1 3	2 2	1 3	2 1
rs10485385	88.57	2 2	2 1	1 1	1 2	1 2	1 2	1 2	1 1
rs6936857	88.58	2 1	2 1	1 1	1 2	1 1	1 2	1 1	1 1
D6S1659	88.59	4 1	4 1	1 1	1 4	1 1	1 3	2 1	1 2
rs1577217	88.60	1 2	1 2	2 2	2 1	2 2	2 1	1 1	2 1
rs2485457	88.61	1 1	1 2	2 2	2 1	2 1	2 1	1 1	2 1
rs2248180	88.62	2 1	2 2	2 2	2 2	2 1	2 1	1 1	2 1

Figure 5.4. Haplotypes of ataxia Group 2 patients at 6q13.

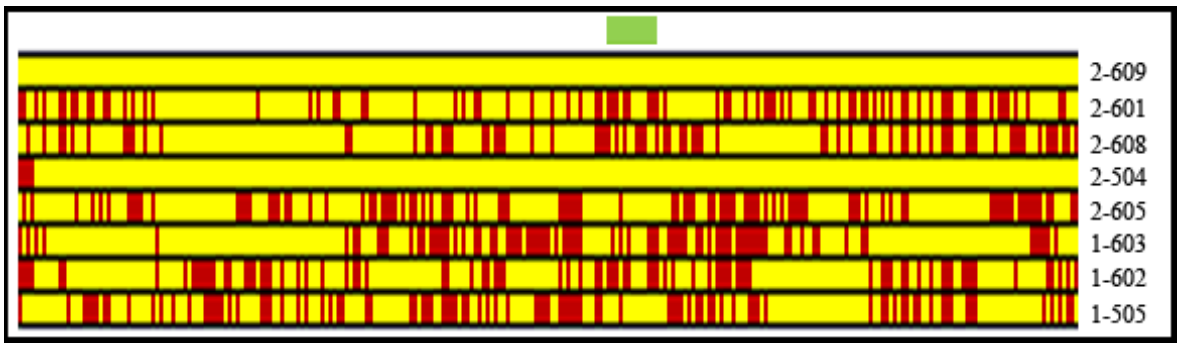


Figure 5.5. HCiE for ataxia patients at 3p14.1 between rs1447445 and rs1443532. Approximate location of *ATXN7* is shown with a green bar. Red bars represent heterozygosity, and data for individual 2-610 are not shown.

Analysis of CNV data for patients 2-504 and 2-609 were completed, and no homozygous deletion or duplication common to the two patients was identified.

5.1.3. Analysis of *APTX*

Coding exons and flanking sequences of *APTX* were sequenced in patients 1-505 and 2-605 (Table 5.1). The reason for analyzing a patient from each family was that they were found to have different genotypes at the region. Transition c.838G→A in exon 8 in the homozygous state was identified in patient 1-505 (Figure 5.6). Other patients from Family 1 were also sequenced for the exon and found to carry the transition also in the

homozygous state. This variant results in the conversion of the tryptophan codon to a termination codon (p.W279X). It leads to the premature termination of translation and thus to the truncation of the protein product, losing the zinc finger domain which is responsible for stabilizing the catalytic site of the protein, HIT domain, onto DNA target site. The mutation has been reported in both homozygous and compound heterozygous states in other AOA1 patients (Moreira *et al.*, 2001, Tranchant *et al.*, 2003).

Table 5.1. The extent of sequencing of *APTX* in ataxia patients 1-505 and 2-605.

Primer	PCR Product (bp)	Exon (bp)	Sequence Read (bp)	
APTXex1-1	No amplification			
APTXex1-2	640	62	-325	+115
APTXex2	Not analyzed, no expression			
APTXex3	477	137	-117	+131
APTXex4	407	51	-216	+79
APTXex5	565	303	-102	+99
APTXex6	400	60	-100	+127
APTXex7	494	227	-105	+91
APTXex8	479	104	-50	+200
APTXex9	688	155	-82	+73

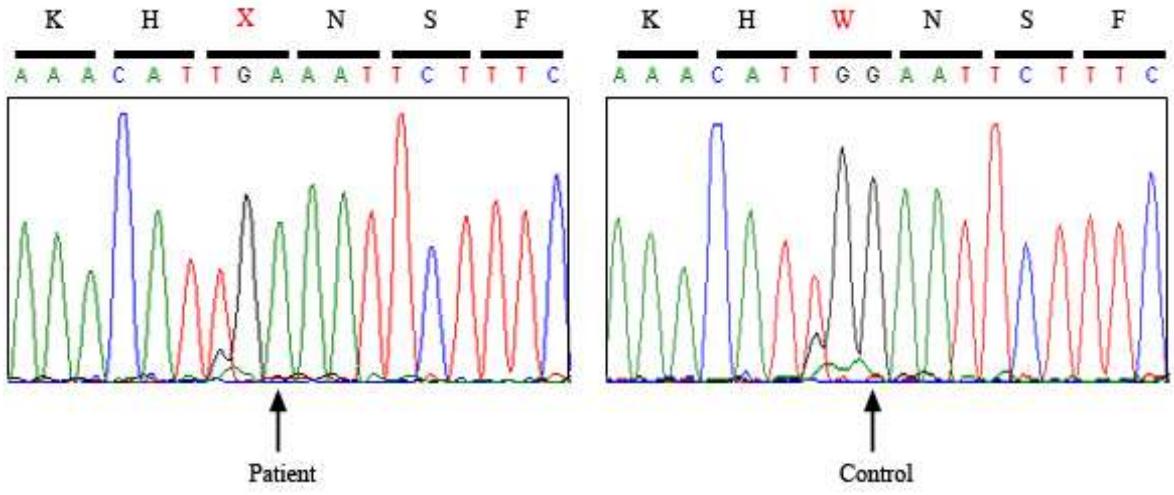


Figure 5.6. Chromatograms showing c.838G→A (p.W279X) in ataxia patient 1-505 and the reference sequence.

Transition c.782T→C in exon 8 was identified in the homozygous state in patient 2-605 (Figure 5.7). It results in the substitution of nonpolar hydrophobic proline for aliphatic

nonpolar hydrophobic leucine (p.L261P). This variant was not reported in SNP or mutation databases. p.L261P was predicted to be damaging for protein structure and function with high reliability by online tools predicting effects of amino acid substitutions, MMB, SNPs3D and SIFT (Table 5.2). In fact, p.L261 hydrophobic residue is part of the ultra-conserved active site with a histidine triad motif (HφHφHφφ, φ representing a hydrophobic residue) at HIT domain, which is responsible for catalytic activity of the protein (Figure 5.8). Substitution of ring structured proline for leucine in triad motif is predicted to interfere with the function of the protein, as the peptide would have less allowable degrees of rotation and reduced the stability. 2-601, 2-608 and their parents were sequenced for the exon, and the patients were found homozygous and parents heterozygous for the variant. The next step was to find out whether the novel variant was a rare polymorphism in the population or not. The standard method in our laboratory for population screening is SSCP. Unfortunately, it would not have been straightforward to interpret the results of this method for this particular mutation, due to the presence of three nearby polymorphisms (rs4878522, rs34600530 and rs62542135). Those SNPs, alone or in combination, were expected to give several aberrant SSCP bands. As an alternative to SSCP, restriction enzyme cleavage was utilized. First, the region of interest was searched for restriction enzyme recognition sites. The variant was found to create a recognition site for *BseGI*. The PCR product carrying the variant and treated with *BseGI* yielded bands of 241 bp and 238 bp, while the normal sequence was not cleaved by the enzyme and remained a single band of 479 bp.

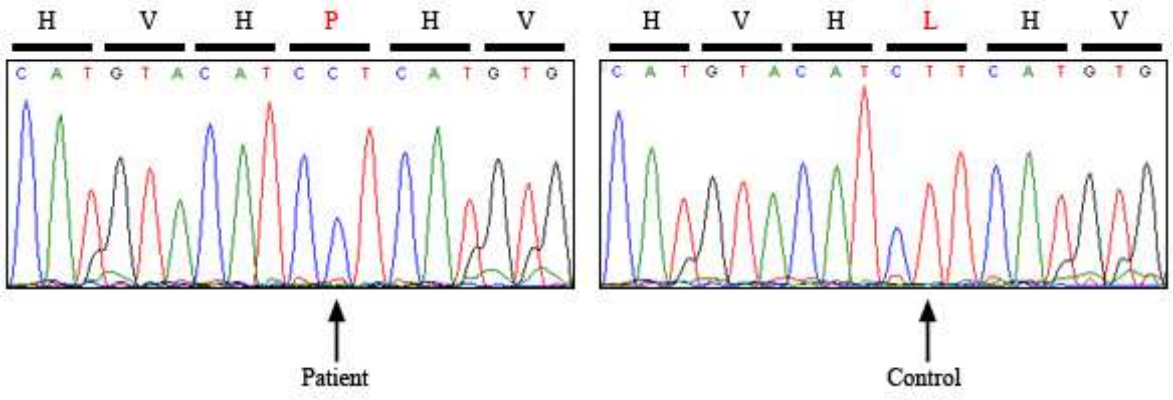


Figure 5.7. Chromatograms showing c.782T→C (p.L261P) in ataxia patient 2-605 and the reference sequence.

Table 5.2.. The effect of p.L261P on APTX protein function as predicted by tools online. Protein Accession number was NP_778239.

Database	MMB	SNPs3D	SIFT
Score	0.83	-2.69	0
Effect	Pathological with high reliability	Intolerable amino acid change	Affects Protein function

```
                  Active histidine triad motif
                                  L261
                              HφHφHφφ
Human    242 KLRFRLGYHAIPSMSHVHLHVISQDFDSPCLKNKKHWNSFNTEYFLESQAVIEMV 297
Chimp    256 KLRFRLGYHAIPSMSHVHLHVISQDFDSPCLKNKKHWNSFNTEYFLESQAVIEMV 311
Dog      242 KLRFRLGYHAIPSMSHVHLHVISQDFDSPCLKNKKHWNSFNTEYFLESQAVIEMV 297
Cattle   256 KFRFRLGYHAIPSMSHVHLHVISQDFDSPCLKNKKHWNSFNTEYFLESQAVIEMV 311
Mouse    235 KLRFRLGYHAIPSMSHVHLHVISQDFDSPCLKNKKHWNSFNTEYFLESQAVIKMV 290
Rat      229 KLRFRLGYHAIPSMSHVHLHVISQDFDSPCLKNKKHWNSFNTEYFLESQAVIKMV 284
Chicken  165 SLEFRLGYHAIPSMSQLHLHVISQDFDSPALKTKKHWNSFTTEYFLNSEEVIEMV 220
Z.fish   224 KLSFRLGYHAIPSMSHVHLHVISQDFDSPCLKNKKHWNSFTTDYFVESQDVISML 279
```

Figure 5.8. Evolutionary conservation of APTX p.L261 and the histidine triad motif localization. The region shown is ultra conserved in mammals.

PCR amplified samples of patients 2-601, 2-605 and 2-608 and their families were treated with *BseGI*, and the segregation of c.782T→C with disease trait was ascertained (Figure 5.9).

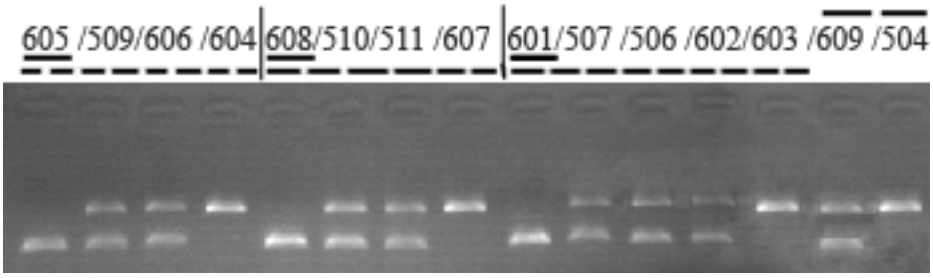


Figure 5.9. Agarose gel showing *BseGI* cleavage results. Only patients 2-601, 2-605 and 2-608 were homozygous for c.782T→C. Families are indicated with dotted underlines, Group 1 Family 2 patients with understrikes and Group 2 patients with overstrikes.

In total 108 control individuals randomly chosen from the population were assayed with *BseGI* restriction enzyme, and the variant was not found in any of them. c.782T→C was concluded to be a novel mutation causing early-onset autosomal recessive ataxia.

5.1.4. Analysis of *ATXN7*

The CAG repeat expansion at the 5'-translated region of *ATXN7* gene causes spinocerebellar ataxia 7 (SCA7, David *et al.*, 1997). Because the clinical symptoms of patients 2-504 and 2-609 were similar to SCA7 and the patients shared a homozygous genotype at the locus, a mild form of SCA7 was hypothesized. The repeat numbers in our patients were kindly assessed at NDAL (Nörodejenerasyon Araştırma Laboratuarı) and compared to negative and positive controls. They were determined to be in the normal range. By sequencing, the exact repeat number was identified as 13, within the normal range of 7-17.

As the next step, all exons of the gene were sequenced in patient 2-504 (Table 5.3). No sequence variant indicative of a disease-causing mutation was identified. Patient 2-504 was found homozygous for the rare alleles of rs1053338, rs6779372, rs3733120 and rs3733121.

Table 5.3. The extent of sequencing of *ATXN7* in ataxia patient 2-504.

Primer Name	PCR product (bp)	Exon (bp)	Sequence Read (bp)	
ATXN7e1	790	440	-80	+135
ATXN7e2	416	98	-70	+40
ATXN7e3	604	99	-180	+193
ATXN7e4	454	253	-58	+59
ATXN7e5-1	No amplification			
ATXN7e5-2	944	336	-261	+236
ATXN7e6	287	68	-72	+70
ATXN7e7	363	104	-82	+90
ATXN7e8	499	253	-150	+22
ATXN7e9	410	260	-26	+48
ATXN7e10	316	83	-56	+45
ATXN7e11	460	266	-65	+64
ATXN7e12	441	199	-110	+44
ATXN7e13	390	121	-67	+125
ATXN7e14a	790	979	-25	614
ATXN7e14b	621	979	508	+75
ATXN7e15	281	67	-108	+32
ATXN7e16	334	17	-57	+70

5.1.5. Analysis for Group 2 Patients after the Addition of the Third Patient

Patient 2-610 was included in our study in March 2010. The DNA sample was immediately sent for SNP genome scan. Group 2 patients were reevaluated including the data generated for 2-610. The regions found significant previously, namely 3p14.1, 6q13 and 10q25.3, were eliminated due to heterozygosity in 2-610. As a result, no strong candidate locus was found in Group 2 patients for further investigation.

5.1.6. Summary of Results for Ataxia Telangiectasia Variant

In summary, candidate locus 9p21.1-13.3 was identified, and then homozygous mutations were detected in *APTX* in both families. One of the mutations was novel.

Two patients in Family 2 did not link to the locus. Three significant loci at 3p14.1, 6q13 and 10q25.3 were identified for these patients. Exons of *ATXN7* at 3p14 were sequenced but no mutation was found. Microsatellite genotyping at 6p13 was heterozygous for patient 2-609. Candidate gene approach at 10q25.3 failed to point out any gene for further investigation. After the addition of 2-610 to Group 2, reevaluation resulted in the elimination of all of the three loci, and no other significant locus was identified.

5.2. Juvenile Parkinsonism

In this study 4 patients and 12 healthy individuals from a single consanguineous family afflicted with Juvenile Parkinsonism were studied (Figure 3.3). The initial genetic analyses on the family became invalid in May 2008 when sibling 505, previously reported healthy, presented disease symptoms at the age of 11 years.

5.2.1. Microsatellite Genome Scan Data Analysis Assuming 505 as Healthy

The initial microsatellite genome scan data were subjected to two- and multipoint lod score analyses assuming autosomal recessive inheritance. The maximum two-point lod score obtained was 2.5 at ATA82B02 (18q22.3, 106.81 cM, Figure 5.10). Multipoint lod score was also highest (1.66) at the same marker. In addition, multipoint lod scores were

rather high at two other loci: 1.10 at interval CATA002 – GATA71H05 (16p12.1, 48.35 cM) and 0.95 at interval GATA104 – GATA189C06 (7q35-36.1, 157.47 cM, Figure 5.11). The next highest lod score was 0.2. Because multipoint lod scores did not reach the significant value (3.00), haplotypes were constructed to investigate descent from a common ancestor. Haplotype segregation analysis indicated possible linkage to six loci, namely, 4p16.1-16.33, 7q35-36.1, 9q31.1-31.2, 16p13.13-q12.1, 18q12.3-23 and 22q13.1-13.33. Linkage to the known Juvenile Parkinsonism loci was also ruled out by haplotype inspection and lod score analysis of microsatellite genome scan data. Two-point and multipoint lod scores on pseudoautosomal regions (PARs) were calculated separately and no significant locus was identified.

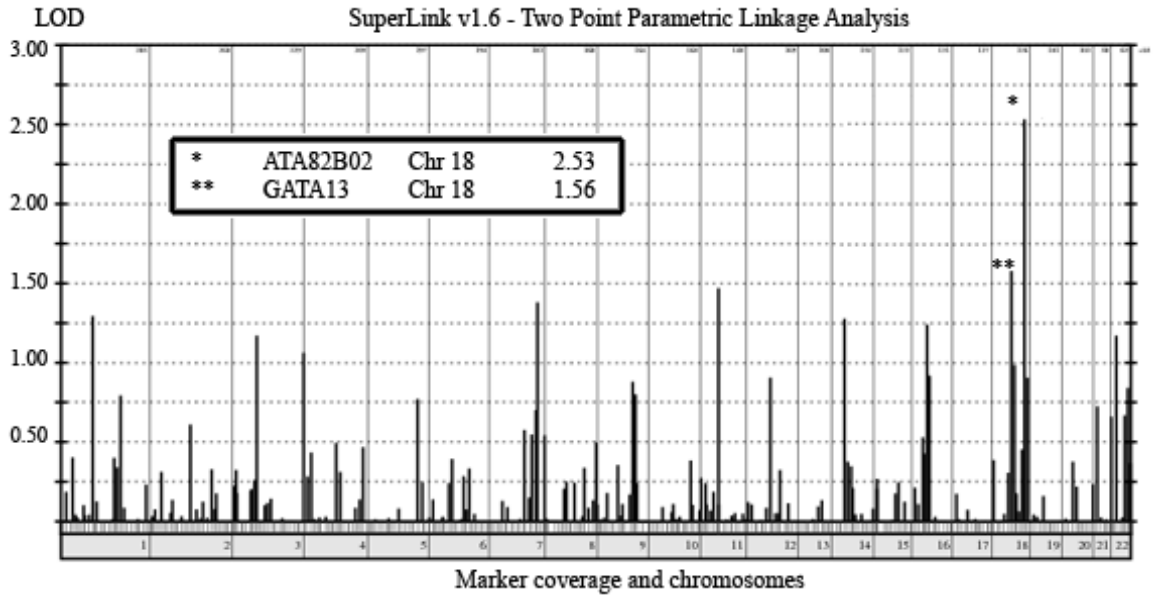


Figure 5.10. Two-point lod scores of initial microsatellite genome scan data of Family JP. Only scores >0 are shown. Loci yielding >1.5 are depicted with asterisks.

Linkage to the 6 loci (4p16.1-16.33, 7q35-36.1, 9q31.1-31.2, 16p13.13-q12.1, 18q12.3-23 and 22q13.1-13.33) with the highest multipoint lod scores obtained by initial microsatellite genome scan haplotype segregation analysis, was not supported by the results of additional genotyping with a total of 63 markers and haplotype segregation analysis at the loci. Highest multipoint lod scores at the loci were -0.34, 0.69, -1.1, 1.29, -1.89 and 1.49, respectively.

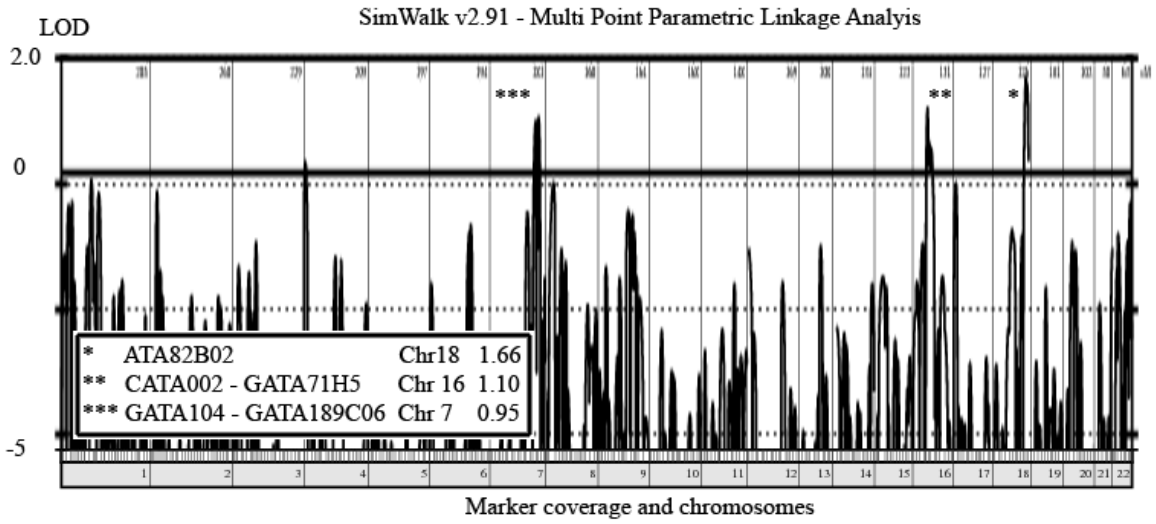


Figure 5.11. Multipoint lod scores of initial microsatellite genome scan data in Family JP. Only scores >-5 are shown. Lod scores > 0.2 are depicted with asterisks.

5.2.2. SNP Genome Scan Data Analysis Assuming 505 as Healthy

Because the gene locus could not be identified using microsatellite genome scan data, a denser genome scan analysis was necessary. SNP data from 13 members of the family were generated and used for both lod score analysis and to search for homozygous regions shared by the three known patients using HCiE. The two analyses together pointed out to a 305-kb region between rs2881210 and rs4569285 (25,623,912 bp and 25,928,335 bp respectively). The region was in the intron of gene *HS3ST4*. Databases reported 16 expressed sequence tags (ESTs) in the region: M78418, BF943172Q2, BF963411Q5, AW875848Q2, BM007749Q2, AW875701, BF892428, DA533203Q2, BF961920Q2, BX0908844Q2, DA947185, CX947185, N71057, W01887, CV378420Q2 and T49785. They were sequenced in a patient, but no mutation was found.

5.2.3. Microsatellite and SNP Genome Scan Data Analysis Assuming 505 as Affected

Upon being informed that disease symptoms were observed in 505, the initial microsatellite genome scan data together with the subsequent microsatellite genotyping results were subjected to lod score and haplotype analyses accordingly (Figure 5.12). No candidate locus with a lod score >3 was identified. Two-point and multipoint lod scores on

pseudoautosomal regions (PARs) also were calculated separately, and no significant locus was identified.

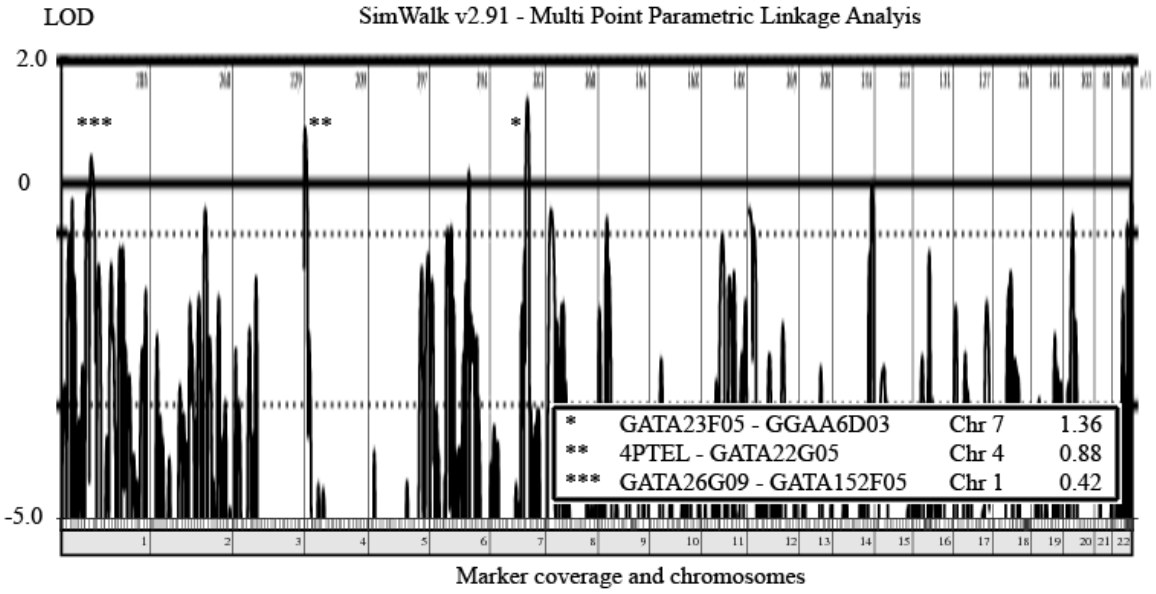


Figure 5.12. Multipoint lod score graphics for Family JP. Final microsatellite genotyping data was used. Only scores >-5 are shown. Lod scores >0.2 are depicted with asterisks.

The data from the SNP genome scan were subjected to multipoint lod score. In the calculation of multipoint lod scores using SNP data, core families were used to reduce data complexity to the limit of linkage software. Individual 303 was not included in the calculations. Two loci were chosen for further studies: 1p31.3-31.2 between rs12047389 and rs747253 (65,107,137 bp and 68,842,414 bp, respectively) and 7q22.1-31.1 between rs6465953 and rs10251582 (103,643,666 bp and 109,677,042 bp, respectively). The maximum multipoint lod scores were calculated as 3.6 at 1p31.3-31.2 (rs121116575; 67,912,259 bp) and 1.67 at 7q22.1-31.1 (rs13232692; 107,317,710 bp). The data was analyzed also with HCiE. At 1p31.3-2 healthy individual 303 was also found to share the homozygous genotype with the patients (Figure 5.13). At 7q22.1-31.1 patient 402 was found by HCiE to share the same homozygous genotype only in a few short regions. The locus was first analyzed by haplotype segregation using SNP data. The region between rs1860803 and rs10242548 (106,779,840 bp and 107,462,149 bp, respectively) was chosen for further studies because patient 402 was observed to have the longest homozygosity here, as was also verified by genotyping with two microsatellite markers (Figure 5.14). At 4pter, patients did not share the homozygous genotypes as judged using HCiE.

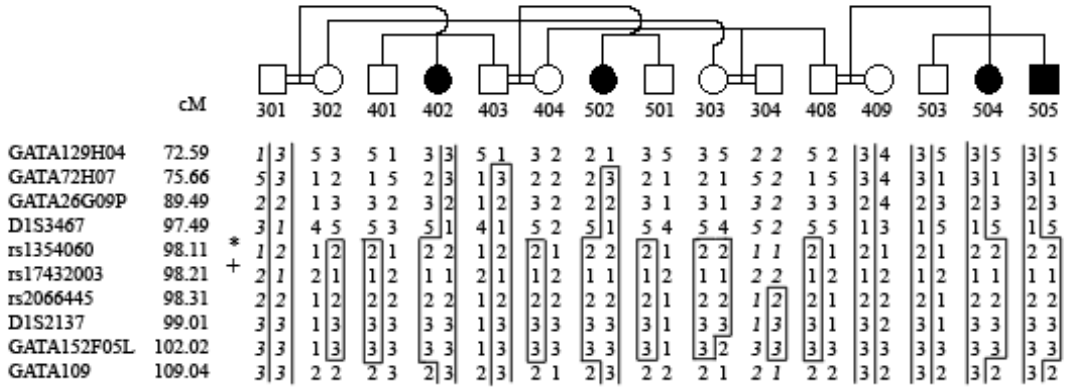


Figure 5.13. Haplotypes of Family JP at 1p31.3-31.2. Approximate location of *AK3L1* and *PDE4B* are indicated by an asterisk and a + sign, respectively.

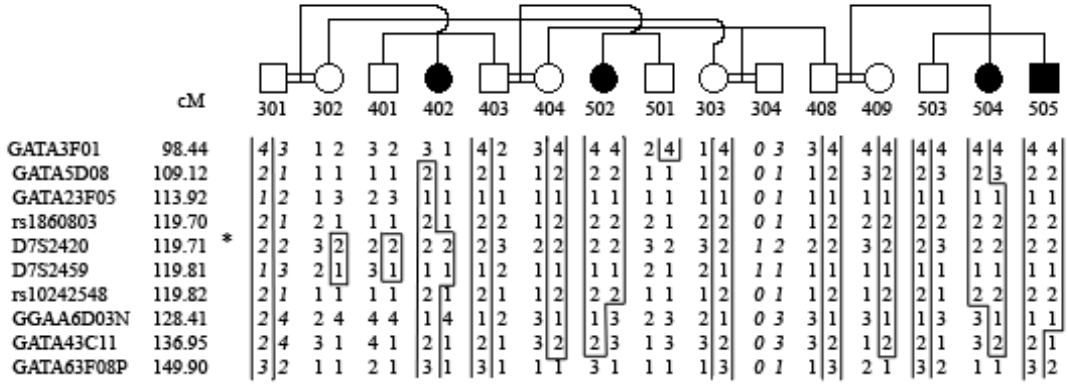


Figure 5.14. Haplotypes of Family JP at 7q22.1-31.1. Asterisk indicates approximate location of *PRKAR2B*.

CNV data from SNP genome scan were analyzed and no homozygous deletion or duplication common to all patients could be identified at the two candidate loci. Among the 40 genes at the candidate locus 1p31.3-31.2, *AK3L1* (65,613,232 bp – 65,693,177 bp) and *PDE4B* (66,258,193 bp – 66,840,262 bp) and among 15 genes at 7q22.1-31.1 *PRKAR2B* (106,685,178 bp – 106,802,256 bp) were assigned as candidates by function and tissue specificity.

5.2.4. Analysis of *PDE4B*

All *PDE4B* exons and flanking sequences were analyzed by sequencing in patient 505, but no mutation was found (Table 5.4). The rare allele of one polymorphism, rs783036, was detected in exon 16.

Table 5.4. The extent of sequencing of *PDE4B* in JP patient 505.

Primer Name	PCR Product (bp)	Exon (bp)	Sequence read (bp)	
PDE4Bex1	326	117	-100	+87
PDE4Bex2	346	112	-120	+20
PDE4Bex3	396	239	-20	+71
PDE4Bex4	421	236	-27	+130
PDE4Bex4b	594	236	-110	+130
PDE4Bex5	306	195	-23	+25
PDE4Bex6-8	696	248	-32	+120
PDE4Bex9	220	50	-45	+25
PDE4Bex10	386	50	-110	+32
PDE4Bex11	399	231	-55	+51
PDE4bex12	299	94	-75	+32
PDE4Bex13	432	183	-50	+140
PDE4Bex14	298	99	-102	+25
PDE4Bex15	381	165	-80	+20
PDE4Bex16	328	100	-31	+110
PDE4Bex17	357	155	-99	+35
PDE4Bex18	300	123	-66	+35
PDE4Bex19	381	183	-50	+78
PDE4Bex20	664	366	-85	+145

5.2.5. Analysis of *AK3L1*

All exons of *AK3L1* were sequenced in patient 505 (Table 5.5). In exon 2 rare alleles of rs7550993 and rs7550994 were detected and c.182A→T transition in exon 4 was identified, all in the homozygous state (Figure 5.15). Other patients, parents and unaffected 303 were analyzed for the variant by direct sequencing. All patients and 303 were found homozygous for the variant while all parents were heterozygous. As control, 130 individuals randomly chosen from the population were screened by SSCP, and variant

c.182A→T was confirmed to be unique for the family (Figure 5.16). This variant was not reported in SNP or mutation databases. It results in the substitution of polar hydrophobic cysteine for polar hydrophilic serine (p.S61C). The residue and the following 13 residues were ultra conserved in chimpanzee (Figure 5.17). The human protein has the exact amino acid sequence with that of chimpanzee's, but the latter has 131 amino acids more at the carboxyl terminus of the protein (5' end of the gene). The residue is aliphatic neutral glycine in all other species. p.S61C was predicted to affect the protein structure and function, by online SIFT tool and SNPs3D. MMB reported that it was neutral with relatively low reliability (Table 5.6). The side chain on cysteine is thiol, which is easily oxidized and involved in covalent disulfide bonds. Substitution of cysteine for serine can disturb protein structure by forming a strong disulfide bond.

Table 5.5. The extent of sequencing of *AK3L1* in JP patient 505.

Primer Name	PCR Product (bp)	Exon (bp)	Sequence Read (bp)	
AK3L1ex1	358	116	-157	+25
AK3L1ex2	418	228	-105	+26
AK3L1ex3	558	350	-40	+50
AK3L1ex4	501	120	-90	+215
AK3L1ex5	529	173	-95	+150
AK3L1ex6	572	119	-178	+140
AK3L1ex7	390	115	-20	+170

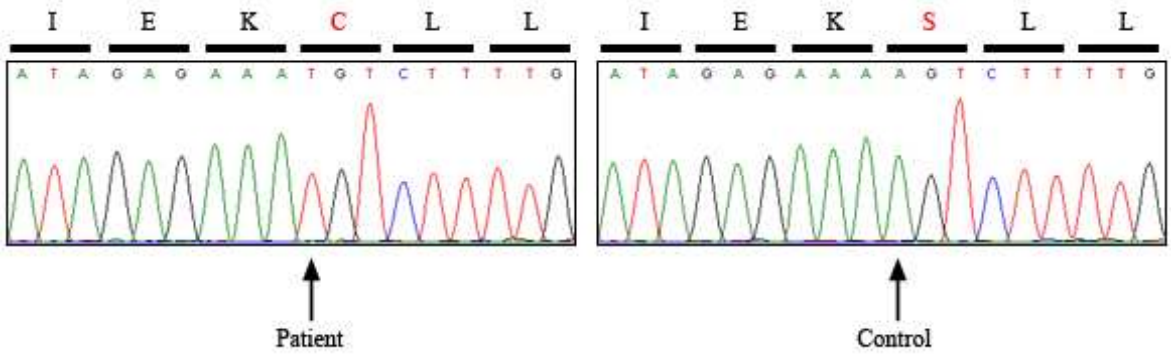


Figure 5.15. Chromatograms showing transition c.182A→T (p.S61C) in JP patient 505 and the reference sequence.

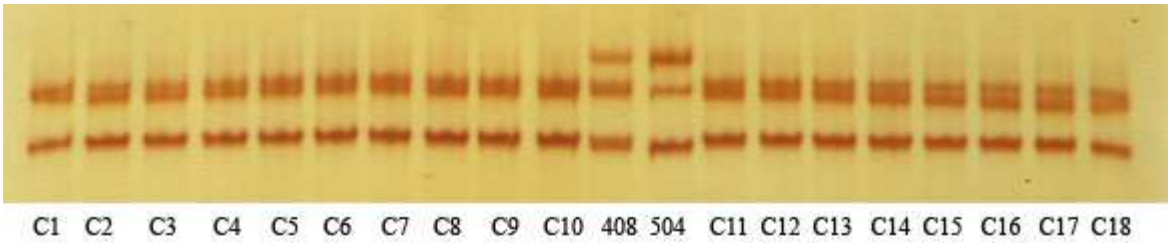


Figure 5.16. An example of the SSCP results for population screening for *AK3L1* c.182A→T. C1-18 are population samples, 408 is heterozygous and 504 homozygous.

```
                                                  S61
Human          28 NFGLQHLSSGHFLRENIKASTEVGEMAKQYIEKSLLVPDHVITRLMM  75
Chimpanzee    158 NFGLQHLSSGHFLRENIKASTEVGEMAKQYIEKSLLVPDHVITRLMM 205
Orangutan      28 NFGLQHLSSGHFLRENIKANTEVGEMAKQYIEKSLLVPDHVITRLMM  75
Macaque        28 NFGLQHLSSGHFLRENIKANTEVGEMAKQYIERSLLVPDHVITRLMM  75
Dog            28 SFGLQHLSSGHFLRENIRANTEVGDMAKQYIEKGLLVPDHVITRLMM  75
Cattle         28 NFGLQHLSSGHFLRENIKANTEVGDMAKQYIEKGLLVPDHVITRLML  75
Wild Boar      28 NLGLQHLSSGHFLRENIKANTEVGDMAKQYIEKGLLVPDHLITRLML  75
Mouse          28 NFGLQHLSSGHLLRENLKTGTEVGDVAKQYLEKGLLVPDHVITRLMM  75
Rat            28 NFGLQHLSSGHLLRENLKTNTEVGDVAKQYLEKGLLVPDHVITRLMM  75
Chicken        15 ----------------------EVGVLAKQYLERGLLVPDHVITRVMM  38
Zebrafish      27 NFGLKHLSSGDFVRENISSKTDAGVLAKTYINKGLLVPDHVMTRLLL  74
```

Figure 5.17. Evolutionary conservation of AK3L1 S61.

Table 5.6. Effect of p.S61C on AK3L1 protein function as predicted by online tools.

Protein Accession is NP_001005353.

Database	MMB	SNPs3D	SIFT
Score	0.25	-3.09	0
Effect	Neutral with low reliability	Intolerable amino acid change	Affects Protein function

5.2.6. Analysis of *PRKAR2B*

All exons of *PRKAR2B* were sequenced in patient 505, but no mutation was found (Table 5.7). The rare allele of one polymorphism, rs3729876, was detected in exon 6.

Table 5.7. The extent of sequencing of *PRKAR2B* in JP patient 505.

Primer Name	PCR Product (bp)	Exon (bp)	Sequence Read (bp)	
PRKAR2Bex1	807	566	-45	+115
PRKAR2Bex2	390	36	-55	+214
PRKAR2Bex3	379	53	-124	+128
PRKAR2Bex4	318	84	-85	+65
PRKAR2Bex5	382	107	-76	+105
PRKAR2Bex6	344	154	-153	+58
PRKAR2Bex7	551	102	-41	+338
PRKAR2Bex8	241	75	-48	+40
PRKAR2Bex9-10	655	205	-169	+19
PRKAR2Bex11	744	134	-161	+322

5.2.7. Analysis of Patients 402, 404 and 405 with Illumina 1M Quad Chip

DNA samples of patients 402, 404 and 505 were reanalyzed, this time using a denser chip (Illumina 1M Quad). The data generated by the scan was imported to Illumina Genome Viewer and two homozygosity mappings, one with 250 SNP filter and the other with 100 SNP filter, were completed. No significant locus other than the previously determined two loci was found. No additional narrowing down of the two previous candidate loci was possible.

CNV analysis was done, and no homozygous deletion/duplication that could possibly contribute to the etiology of the disease was found in patients.

5.2.8. Summary of Results for Juvenile Parkinsonism

In summary, two significant loci, 1p31.3-2 and 7q22.1-31.1, were identified. *PDE4B* and *AK3L1* at 1p31.3-2 and *PRKAR2B* at 7q22.1-31.1 were analyzed. Transition c.182A→T in *AK3L1* exon 4 was identified in the homozygous state in all patients as well as in unaffected 303. One hundred and thirty control samples randomly chosen from the population were screened by SSCP, and the mutation was shown to be unique for the family.

5.3. Larsen Syndrome

In this study 8 patients and 21 healthy individuals from 7 families were genetically investigated (Figure 3.4). One of the families (LRS6) was the most informative one, as it had two affected and two healthy sibs, while the other families had single cases only.

5.3.1. Microsatellite Genome Scan Data Analysis

The initial microsatellite genome scan data for seven Larsen Syndrome families were subjected to two-point and multipoint lod score analyses, assuming autosomal recessive inheritance. Two-point lod scores >2 were obtained at three loci: 4.62 at GATA178G09 (2q37.3, 251.94 cM), 4.59 at GGAA5F09 (1q23.3, 170.84 cM) and 2.35 at CATC015 (1p36.32, 8.01 cM, Figure 5.18). Multipoint lod scores >1 were obtained at two of those loci (1.3 at CATC015 and 1.01 at GATA178G09) plus 1.5 at interval 044xg3 – AAT095 (17q25.3, 119.53 cM, Figure 5.19). Haplotype segregation analysis was also performed for each family and evaluated together with multipoint lod scores for Family LRS6 to identify candidate loci (Figure 5.20). Two-point and multipoint lod scores on pseudoautosomal regions (PARs) were calculated separately and no significant locus was identified.

Linkage to *FLNB* (3p14.3, 57,994,127 bp – 58,157,978 bp), that is known as mutated in autosomal dominant form of the disease, was ruled out by genome scan data analysis, assuming an autosomal recessive inheritance model. Mutations in *CHST3* (10q22.1, 73,724,120 bp – 73,773,322 bp) were reported to cause autosomal recessive Larsen syndrome (Hermanns *et al.* 2008). Linkage to this locus was also ruled out, by the data obtained from flanking markers D10S218, D10S72486kb and D10S1146 (Figure 5.25).

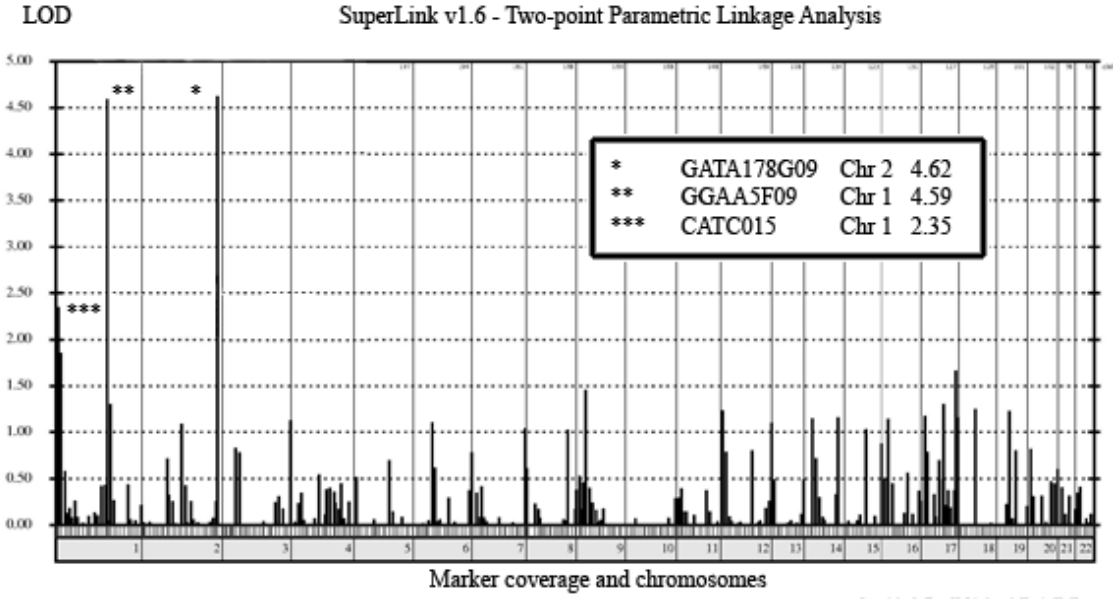


Figure 5.18. Two-point lod score graphics of initial microsatellite genome scan data for all Larsen families together. Only the scores >0 are shown. Loci yielding >2 are depicted with asterisks.

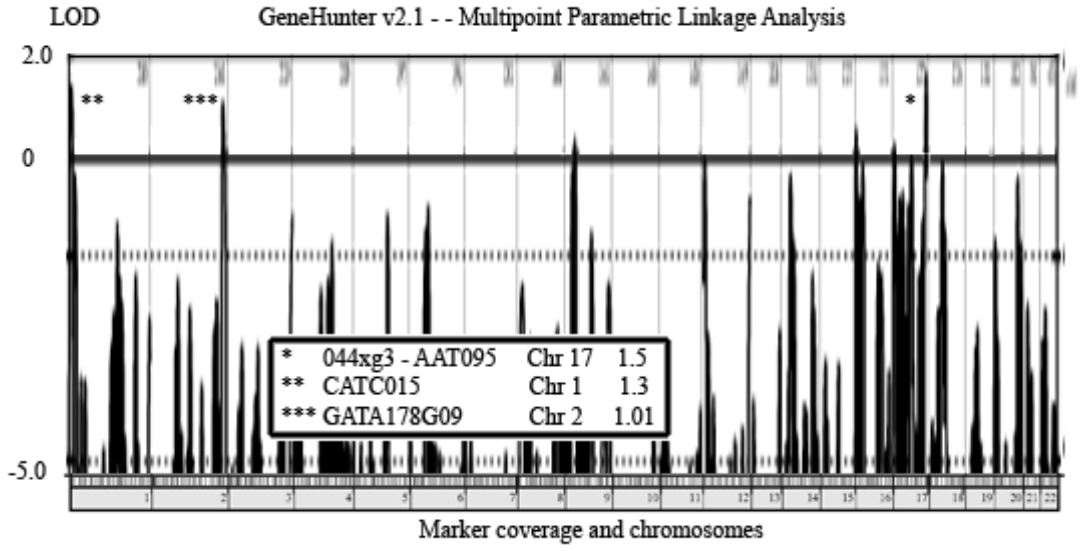


Figure 5.19. Multipoint lod score graphics of initial microsatellite genome scan data of all Larsen families together. Only scores >-5 are shown. Loci yielding >0.2 are depicted with asterisks.

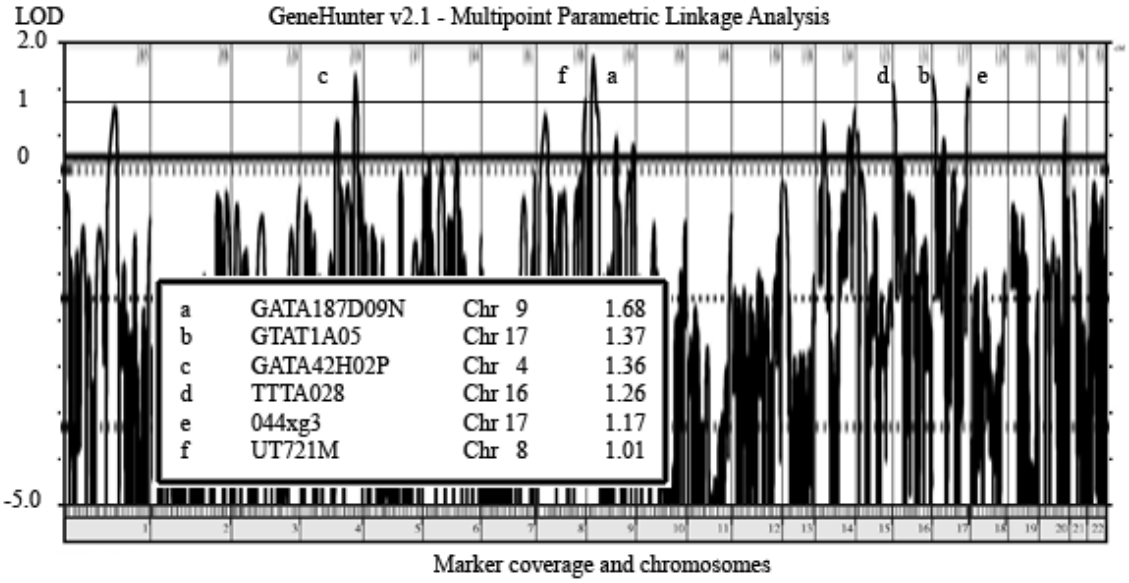


Figure 5.20. Multipoint lod score graphics of initial microsatellite genome scan data for Family LRS6. Only scores >-5 are shown. Loci yielding >1 are depicted with letters.

Eleven loci that were assessed as significant and selected for further linkage studies as well as *FLNB* (3p14.3) and *CHST3* (10q22.1) loci are shown in Table 5.8. Maximum two-point and total multipoint lod scores obtained by SuperLink and GeneHunter are also included.

Table 5.8. Loci studied with microsatellites and linkage analysis in Larsen families.

Locus	**Family**	**Position (Mb)**	**Maximum lod score**	
			Two-point	**Multipoint**
1p36.33-36.13	LRS1-7	2 – 17	3.22	-1.04
1q21.1-23.3	LRS1-7	144 – 162	4.75	-8.28
2q37.2-qter	LRS1-7	237- 243	4.62	-3.97
3p14.1-12.1	LRS1-7	68 – 86	0.3	-5.51
4q26.31-27.1	LRS6	175 – 185	1.4	-0.19
8p23.1-21.3	LRS1-7	9 – 20	0.2	-5.9
8q24.22-24.3	LRS6	136 – 146	1.4	0.57
9p23-21.3	LRS6	10 – 23	1.29	-0.51
10q22.1	LRS1-7	73 – 74	-15.9	-9.8
16pter-13.3	LRS6	0 – 6	1.51	0.99

Table 5.8. Loci studied with microsatellites and linkage analysis in Larsen families (continued).

Locus	Family	Position (Mb)	Maximum lod score	
			Two-point	Multipoint
17p13.3-13.2	LRS6	0.6 – 6	0.44	1.35
17q25.3	LRS1-7	74 – 79	3.4	2.88
20q13.13-13.32	LRS1-7	47 – 57	1.06	-3.28

5.3.2. Fine-Mapping and Linkage Studies for All Larsen Families

All loci that yielded high lod scores in all families together (Table 5.5) were studied further by genotyping in all families.

Two loci were investigated on chromosome 1. The first one was at 1p36.33-36.13. Individuals were genotyped with 18 microsatellite markers at the region (Figure 5.21). Maximum two-point lod score was calculated as 3.22 at D1S1160 (20.61 cM) and total multipoint lod score as -1.04 at 25.58 cM. The second locus was at 1q21.1-23.3. Data from 10 additional microsatellite markers were used for lod score and haplotype analysis (Figure 5.21). Maximum two-point lod score was obtained as 4.75 at D1S1679 (162,361,915 bp) and total multipoint lod score as -8.28 at the same marker. D1S1679 was found to be homozygous and informative for patients of 6 families. The nearby gene *DDR2* (160,602,228 bp – 162,750,255 bp) was investigated for homozygosity in these patients. All patients except 4-602 were found heterozygous at the designed microsatellite markers D1S160867kb and D1S160921kb, located within the exons of *DDR2* (Figure 5.21). Linkage to both loci was excluded.

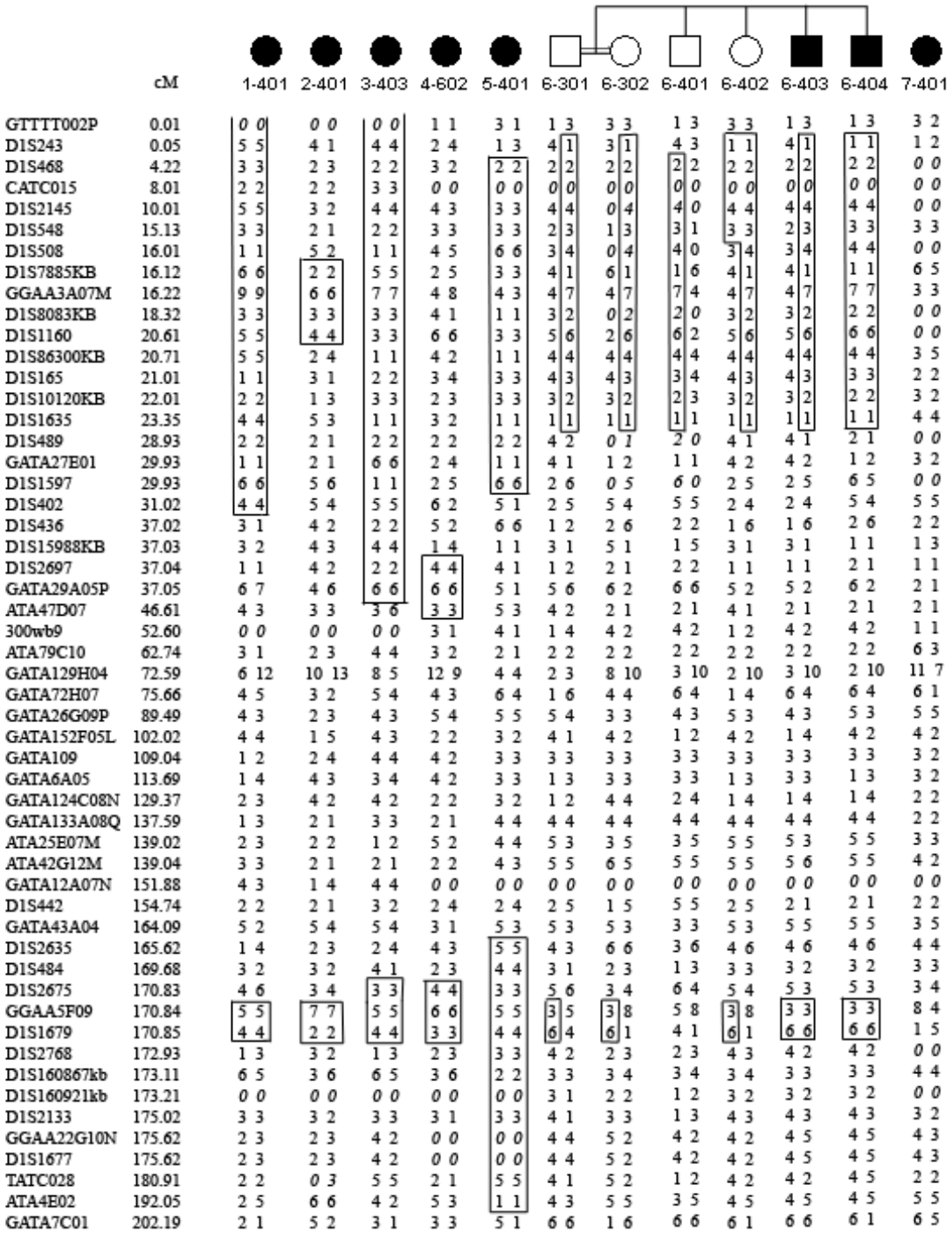


Figure 5.21. Haplotypes of Larsen patients and Family LRS6 at 1p36.33-36.13 and 1q21.1-23.3.

Six microsatellite markers were used for fine mapping at 2q37.2-qter. At GATA178G09 (251.94 cM) the maximum two-point lod score was calculated as 4.62 and total multipoint lod score as -3.97. *GAL3ST2* (238,716,240 bp – 238,743,902 bp) that is at

the homozygous region for families LRS1, LRS3 and LRS5, was selected as a good candidate by candidate gene approach (Figure 5.22).

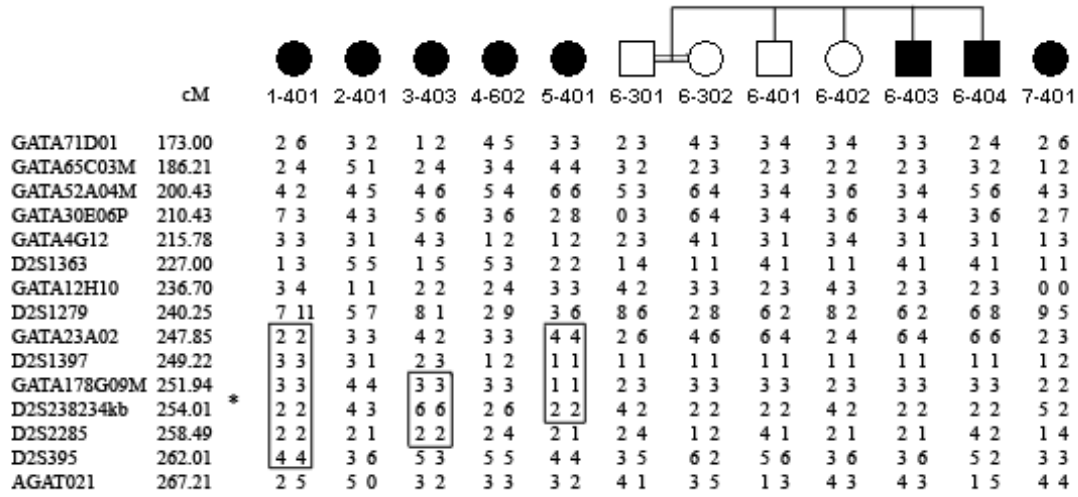


Figure 5.22. Haplotypes of Larsen patients and Family LRS6 at 2q37.2-qter. Asterisk indicate *GAL3ST2*.

Locus 3p14.1-12.1 was fine-mapped with 13 additional microsatellite markers (Figure 5.23). The maximum two-point lod score was 0.3 at D3S3581 (104.3 cM) and total multipoint lod score -5.51 at 117.12 cM (GATA128C02 – GATA84B12). Linkage to this locus was excluded.

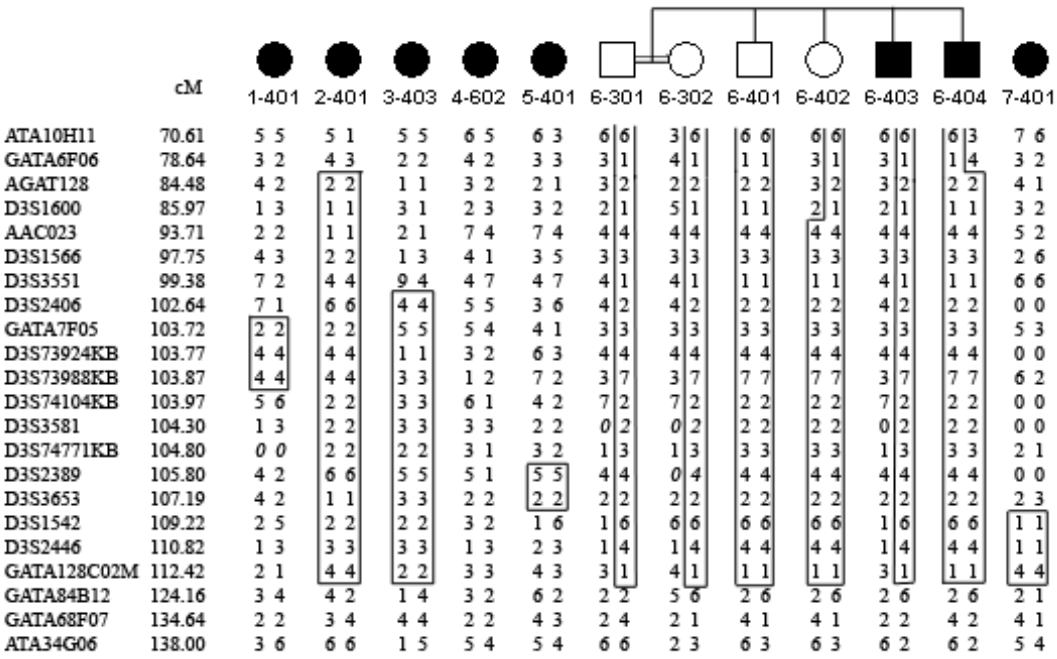


Figure 5.23. Haplotypes of Larsen patients and Family LRS6 at 3p14.1-12.1.

Three microsatellite markers were genotyped at 8p23.1-21.3 (Figure 5.24). The maximum two-point lod score was calculated as 0.2 at 15 Mb and total multipoint lod score as -5.9 at 17 Mb. Linkage to this locus was excluded.

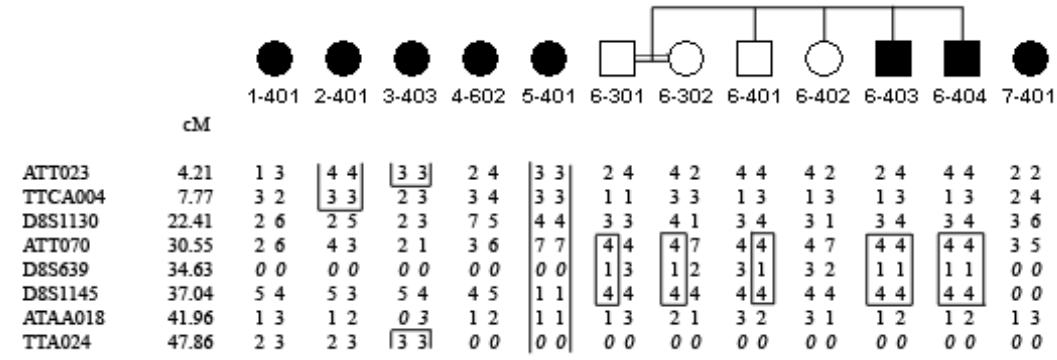
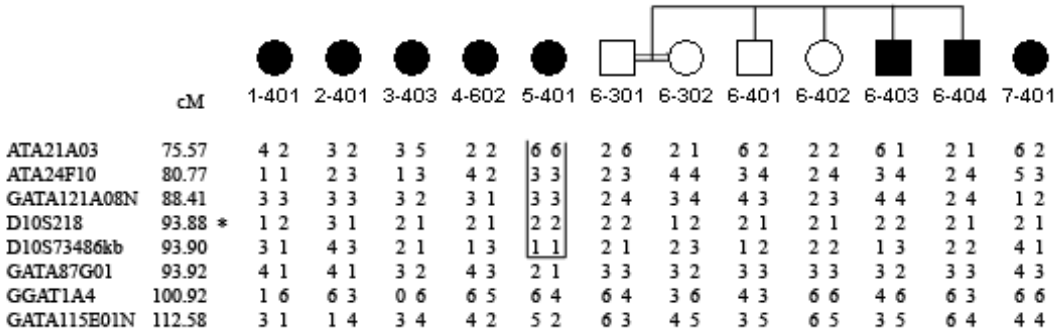


Figure 5.24. Haplotypes of Larsen patients and Family LRS6 at 8p23.1-21.3.

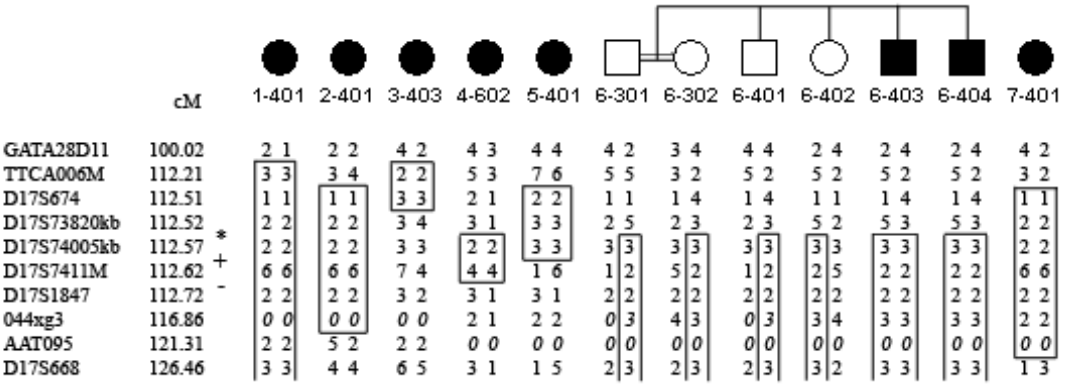
At 10q22.1 *CHST3* locus (73,724,120 bp – 73,773,322 bp) was investigated with 2 microsatellite markers (Figure 5.25). Linkage to this region was ruled out for all families except LRS5 with haplotype segregation studies.



	cM		1-401	2-401	3-403	4-602	5-401	6-301	6-302	6-401	6-402	6-403	6-404	7-401
ATA21A03	75.57		4 2	3 2	3 5	2 2	6 6	2 6	2 1	6 2	2 2	6 1	2 1	6 2
ATA24F10	80.77		1 1	2 3	1 3	4 2	3 3	2 3	4 4	3 4	2 4	3 4	2 4	5 3
GATA121A08N	88.41		3 3	3 3	3 2	3 1	3 3	2 4	3 4	4 3	2 3	4 4	2 4	1 2
D10S218	93.88	*	1 2	3 1	2 1	2 1	2 2	2 2	1 2	2 1	2 1	2 2	2 1	2 1
D10S73486kb	93.90		3 1	4 3	2 1	1 3	1 1	2 1	2 3	1 2	2 2	1 3	2 2	4 1
GATA87G01	93.92		4 1	4 1	3 2	4 3	2 1	3 3	3 2	3 3	3 3	3 2	3 3	4 3
GGAT1A4	100.92		1 6	6 3	0 6	6 5	6 4	6 4	3 6	4 3	6 6	4 6	6 3	6 6
GATA115E01N	112.58		3 1	1 4	3 4	4 2	5 2	6 3	4 5	3 5	6 5	3 5	6 4	4 4

Figure 5.25. Haplotypes of Larsen patients and Family LRS6 at 10q22.1. Asterisk indicates approximate location of *CHST3*.

At 17q25.3, the total multipoint lod score of 1.5 was obtained using the initial genome scan data at interval 044xg3 – AAT095. The region was studied with 6 additional markers (Figure 5.26). Maximum two-point lod score was obtained as 3.40 at D17S74005kb and total multipoint 2.88 at 044xg3. The locus could be only partially narrowed down because the families were not informative for the microsatellite markers used for genotyping. The region between 76,308,659 bp and 76,494,162 bp was defined to be homozygous for most patients. Three genes; *suppressor of cytokine signaling 3* (*SOCS3*), *phosphatidylglycerophosphatase synthase 1* (*PGS1*) and *dynein, axonemal, heavy chain 17* (*DNAH17*) are located in this region. *SOC3* (76,352,859 bp – 76,356,158 bp) was chosen for further studies after candidate gene approach. Additionally, *TIMP metallopeptidase inhibitor 2* (*TIMP2*, 76,921,472 bp – 76,849,059 bp) at the neighborhood of the region was also selected for sequencing.



	cM		1-401	2-401	3-403	4-602	5-401	6-301	6-302	6-401	6-402	6-403	6-404	7-401
GATA28D11	100.02		2 1	2 2	4 2	4 3	4 4	4 2	3 4	4 4	2 4	2 4	2 4	4 2
TTCA006M	112.21		3 3	3 4	2 2	5 3	7 6	5 5	3 2	5 2	5 2	5 2	5 2	3 2
D17S674	112.51		1 1	1 1	3 3	2 1	2 2	1 1	1 4	1 4	1 1	1 4	1 4	1 1
D17S73820kb	112.52		2 2	2 2	3 4	3 1	3 3	2 5	2 3	2 3	5 2	5 3	5 3	2 2
D17S74005kb	112.57	*	2 2	2 2	3 3	2 2	3 3	3 3	3 3	3 3	3 3	3 3	3 3	2 2
D17S7411M	112.62	+	6 6	6 6	7 4	4 4	1 6	1 2	5 2	1 2	2 5	2 2	2 2	6 6
D17S1847	112.72	-	2 2	2 2	3 2	3 1	3 1	2 2	2 2	2 2	2 2	2 2	2 2	2 2
044xg3	116.86		*0 0*	*0 0*	*0 0*	2 1	2 2	*0* 3	4 3	*0* 3	3 4	3 3	3 3	2 2
AAT095	121.31		2 2	5 2	2 2	*0 0*	*0 0*	*0 0*	*0 0*	*0 0*	*0 0*	*0 0*	*0 0*	*0 0*
D17S668	126.46		3 3	4 4	6 5	3 1	1 5	2 3	2 3	2 3	3 2	3 3	3 3	1 3

Figure 5.26. Haplotypes of Larsen patients and Family LRS6 at 17q25.3. *SOCS3*, *TIMP2* and *CANT1* are shown by an asterisk, an + sign and a – sign respectively.

Locus 20q13.13-13.32 was investigated with 13 additional markers (Figure 5.27). Maximum two-point lod score was 1.06 at D20S100 and total multipoint lod score as -3.28 between D20S451 and D20S164. Linkage to this locus was excluded.

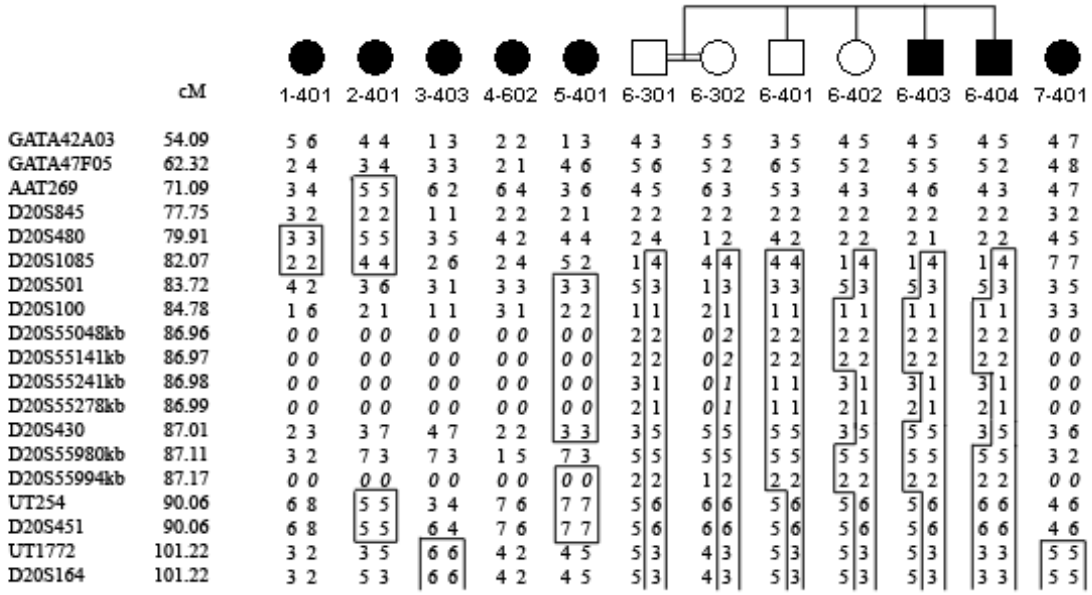
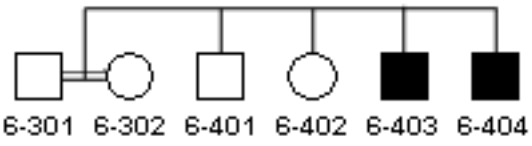


Figure 5.27. Haplotypes of Larsen patients and Family LRS6 at 20q13.13-13.32.

5.3.3. Fine Mapping and Linkage Studies for Family LRS6

Three of the loci given at Table 5.5 were further studied by genotyping only in LRS6 family.

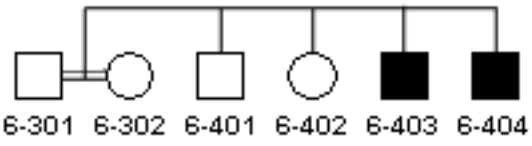
At 4q26.31-27.1, Family LRS6 was studied with three markers between GGAA19H07 and 165zf8ZP (Figure 5.28). Maximum two-point lod score was calculated as 1.4 at GATA42H02 (181.94 cM) and multipoint lod score as -0.19 at 179.29 cM (GGAA19H07 – D4S415). Linkage to this locus was excluded for the family with lod scores and haplotype data.



	cM	6-301	6-302	6-401	6-402	6-403	6-404
GATA8A05	157.99	2 1	1 1	1 1	2 1	1 1	1 1
GATA27G03	167.55	3 1	2 2	1 2	3 2	1 2	1 2
GGAA19H07	176.19	3 2	2 1	3 2	3 2	2 2	2 1
D4S415	181.35	1 4	2 3	1 2	1 2	4 2	4 2
GATA42H02P	181.93	1 2	2 1	1 2	1 2	2 2	2 2
D4S2967	181.94	3 2	4 1	3 4	3 4	2 4	2 4
D4S1530	185.28	3 3	2 1	3 2	3 2	3 2	3 2
165zf8ZP	192.37	2 3	4 1	2 1	2 4	3 4	3 4
GATA5B02M	207.62	2 1	4 3	2 3	2 4	1 3	2 4

Figure 5.28.ACGER. Haplotypes of Larsen Family LRS6 at 4q26.31-27.1

At 8q24.22-24.23, Family LRS6 was studied with four additional markers (Figure 5.29). Maximum two-point lod score was obtained as 1.4 at D8S346 (158.94 cM) and multipoint lod score as 0.57 at UT721 (164.48 cM). Linkage to this locus was excluded later by SNP analysis because no homozygosity for the patients was observed (Figure 5.30).



	cM	6-301	6-302	6-401	6-402	6-403	6-404
GATA7G07	135.08	4 3	1 3	4 3	3 3	4 3	4 1
GATA21C12	139.53	3 2	6 3	3 3	2 3	3 6	3 6
GATA50D10	154.02	4 2	5 4	4 4	2 4	4 5	4 5
D8S272	154.03	2 1	3 1	2 1	1 1	2 3	2 3
D8S346	158.94	1 2	1 2	1 2	2 1	1 1	1 1
UT721	164.47	3 3	3 4	3 4	3 3	3 3	3 3
D8S1836	165.93	3 2	1 3	3 3	2 1	3 1	3 1
D8S1925	167.90	2 3	1 1	2 1	3 1	2 1	2 1

Figure 5.29. Haplotypes of Larsen Family LRS6 at 8q24.22-24.23.

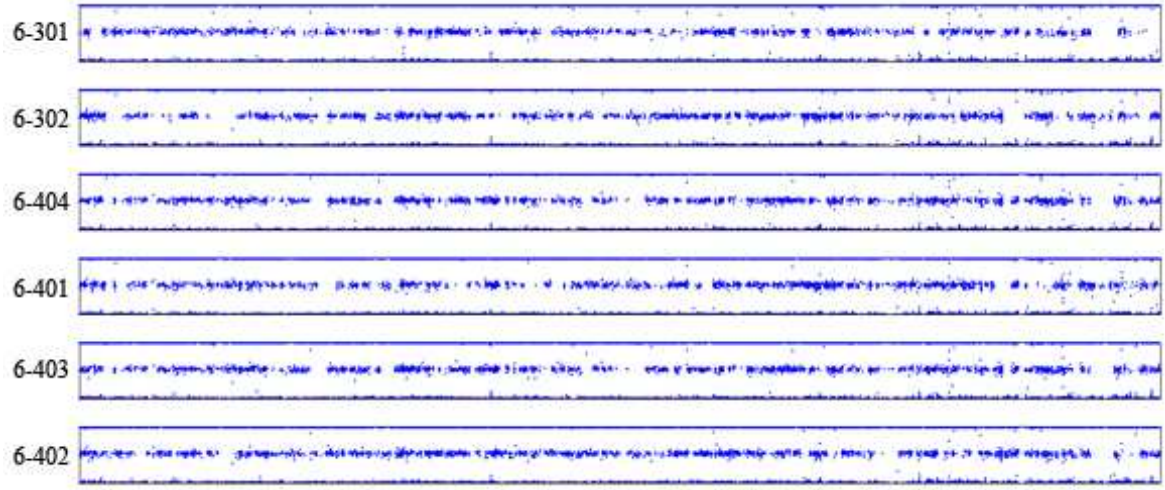
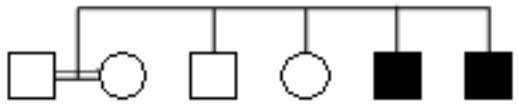


Figure 5.30. Family LRS6 SNP genome scan at 8q24.22-24.23.

Four additional markers were studied for Family LRS6 at 9p23-21.3 (Figure 5.31). Maximum two-point lod score was calculated as 1.29 at GATA187D09 and multipoint lod score as -0.51 between D9S921 and AGAT142P. Linkage to this locus was excluded for the family.

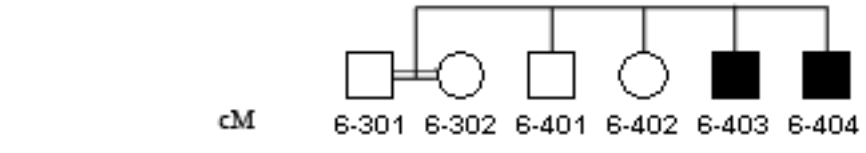


	cM	6-301	6-302	6-401	6-402	6-403	6-404
MFD455-AAT052	0.11	2 1	2 4	2 4	2 4	1 2	2 2
aaaac001	10.25	5 2	5 2	5 2	5 2	2 5	5 5
GATA62F03M	14.23	3 2	3 2	3 2	3 2	2 3	3 3
GATA187D09N	21.88	5 5	5 7	5 7	5 7	5 5	5 5
D9S921	21.91	1 5	8 9	1 9	1 9	5 8	1 8
AGAT142P	31.24	3 3	3 2	3 2	3 3	3 3	3 3
D9S925	32.24	3 1	1 2	1 2	3 1	1 1	3 1
D9S1684	34.42	2 2	2 1	2 1	2 2	2 2	2 2
D9S1870	37.58	3 2	1 3	2 3	3 1	2 1	3 1
GATA87E02N	44.28	6 1	6 4	1 4	6 6	1 6	6 6
GATA5E06P	52.76	7 2	6 3	2 3	7 6	2 6	7 6
GATA7D12	66.32	2 8	1 8	8 8	2 1	8 1	2 8

Figure 5.31. Haplotypes of Larsen Family LRS6 at 9p23-21.3.

At 16pter-13.3, Family LRS6 was studied with 6 microsatellite markers (Figure 5.32). The maximum two-point lod score was calculated as 1.51 at D16S3024 and multipoint lod score as 0.99 at TTTA028. A homozygous region for patients was identified between rs12719809 and rs12102635 (542,727 bp and 800,821 bp, respectively) but the

father was also observed to be homozygous at the region (Figure 5.33). Linkage to this locus was excluded for the family.



	cM	6-301	6-302	6-401	6-402	6-403	6-404
TTTA028	1.01	2 2	2 1	2 1	2 1	2 2	2 2
D16S3024	5.04	1 2	2 3	1 3	2 3	2 2	2 2
D16S1715kb	5.50	1 1	2 1	1 1	1 1	1 2	1 2
D16S3395	6.14	2 1	1 2	2 2	1 2	2 1	1 1
D16S2.68mb	7.24	2 2	1 1	2 1	2 1	2 1	2 1
D16S3082	8.14	2 3	1 3	2 3	2 3	2 1	2 1
D16S3.26mb	9.24	1 2	2 1	1 1	1 1	1 2	1 2
ATA41E04	11.46	3 3	2 1	3 1	3 1	3 2	3 2
ATA3A07	22.65	4 1	3 2	1 3	1 2	4 3	4 3
TTAT023Z	37.41	2 1	1 3	1 3	1 3	1 1	2 1
ATT001	42.31	2 3	1 3	3 3	3 3	3 1	2 1
CATA002Z	43.12	3 1	2 1	1 1	1 1	1 2	3 2

Figure 5.32. Haplotypes of Larsen Family LRS6 at 16pter-13.3.

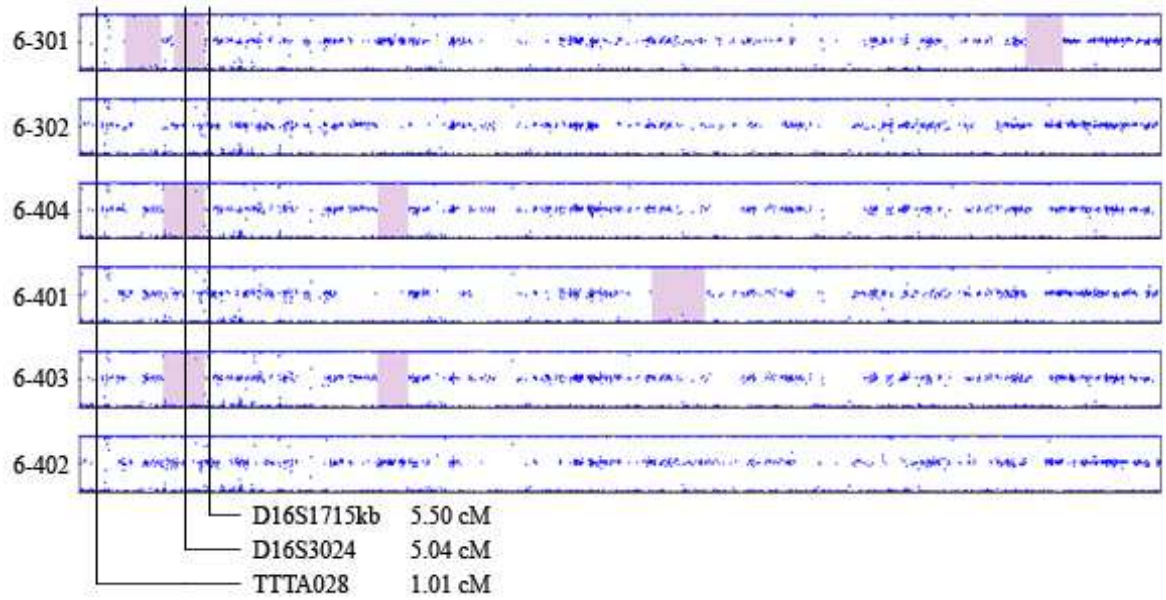
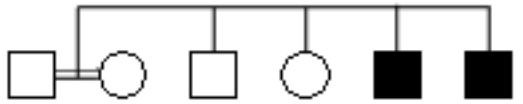


Figure 5.33. Family LRS6 SNP genome scan at 16pter-13.3.

At 17p13.3-13.2, Family LRS6 was studied with two additional markers (Figure 5.34). Maximum two-point lod score was calculated as 0.44 and multipoint lod score as 1.35 at GTAT1A05 (0.63 cM). Linkage to this locus was excluded for the family.



	cM	6-301	6-302	6-401	6-402	6-403	6-404
GTAT1A05	0.63	1 1	1 1	1 1	1 1	1 1	1 1
D17S1840	1.29	1 1	3 2	1 2	1 2	1 3	1 3
D17S919	9.92	2 3	1 3	3 3	3 3	2 1	2 1
GATA158H04	14.65	3 2	4 1	2 1	2 1	3 4	3 4
GATA8C04	22.23	2 3	3 1	3 3	3 1	2 3	2 3
ATA78D02N	31.96	2 2	3 1	2 3	2 1	2 3	2 3
GATA185H04N	44.62	1 2	1 3	2 1	2 3	2 1	1 1

Figure 5.34. Haplotypes of Larsen Family LRS6 at 17p13.3-13.2.

5.3.4. SNP Analysis in Patients with Larsen Syndrome

Illumina 1M Quad SNP chip was used to analyze patients 1-401, 2-401, 3-403, 4-602, 6-403, 6-404, 7-401 as well as healthy individuals 6-301, 6-302, 6-401, 6-402 in Family LRS6. The genome scan data was imported to Illumina Genome Viewer and homozygosity mapping analysis was completed. Using the data from SNP genome scan, the strongest candidate region was again 17q25.3. Patients 1-401, 2-401, 6-403, 6-404 and 7-401 were found homozygous between rs691094 and rs7226158 (76,425,176 bp and 77,390,725 bp, respectively; Figure 5.35). Patients 1-401, 2-401 and 7-401 were found to share the same homozygous genotype, while 6-403 and 6-404 shared another homozygous genotype. Eleven genes were located at the region. *Calcium activated nucleotidase 1* (*CANT1*, 76,987,799 bp – 77,005,838 bp) was chosen for further studies after candidate gene approach. Mutations in the *CANT1* gene were reported to cause an autosomal recessive chondrodysplasia, Desbuquois Dysplasia. Larsen Syndrome and Desbuquois Dysplasia exhibit similar clinical manifestations (Huber *et al.*, 2009). Patient 3-403 and 4-602 were heterozygous at gene sequences.

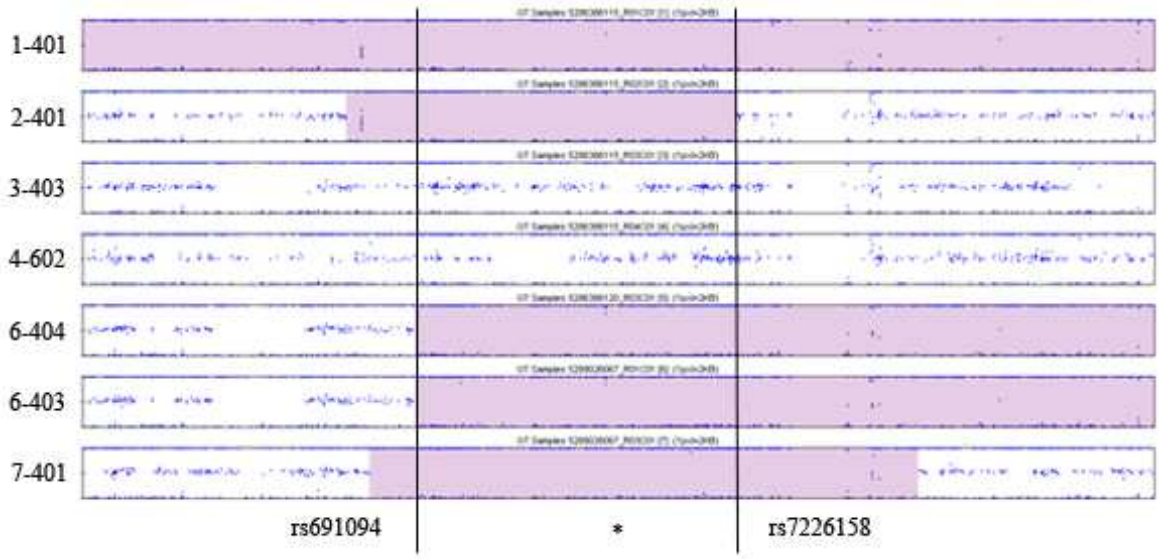


Figure 5.35. Homozygous regions in Larsen Syndrome patients at 17q25.3. Pink areas show homozygous regions. Asterisk indicates an approximate location for *CANT1*.

Besides 17q25.3, two loci that were shared in more than two patients were noted. A common homozygous region for patients 1-401, 2-401, 3-403 and 7-401 was identified at 5q31.1 between rs153025 and rs11242115 (131,540,884 bp and 131,826,413 bp, respectively). Patients 6-403, 6-404 and one healthy sibling, 6-401 were also homozygous at the region but genotype of 6-403 was different from 6-404 and 6-401; thus, 5q31.1 was not considered a candidate region for Family LRS6. Another homozygous region was identified for patients 1-401, 6-602 and 7-401 at 20q11.33 between rs2104417 and rs2425209 (34,127,871 bp and 34,860,332 bp, respectively).

Three candidate loci were found restricted to Family LRS6 only. The first one was 2q21.2, between rs6750271 and rs2163654 (133,271,657 and 134,259,612 bp, respectively). The second locus was a 7-Mb region at 14q32.31-14qter, beginning from rs1953392 and extending to q terminus (101,621,566 bp to 107,350,000 bp). The last locus was at 17q25.3-17qter, beginning from rs691094 (76,425,176 bp to 81,195,000 bp).

CNV analysis was done and no homozygous deletion or duplication that could possibly contribute to the etiology of the disease was identified in patients.

5.3.5. Analysis of *TIMP2*

All exons of the *TIMP2* gene were sequenced in patient 6-403 (Table 5.9). No mutation was identified. Rare alleles of rs7503726 in exon 1 and rs2277698 in exon 3 were detected in the homozygous state.

Table 5.9. The extent of sequencing of *TIMP2* in Larsen patient 6-403.

Primer Name	PCR Product (bp)	Exon (bp)	Sequence Read (bp)	
Timp2ex1	635	432	-35	+98
Timp2ex2	295	101	-35	+76
Timp2ex3	298	109	-51	+45
Timp2ex4	336	125	-35	+99
Timp2ex5	399	198	-58	+65

5.3.6. Analysis of *SOCS3*

Exons of *SOCS3* gene were sequenced for patient 6-403 (Table 5.10). No mutation was identified.

Table 5.10. The extent sequencing of *SOCS3* in Larsen patient 6-403.

Primer Name	PCR Product (bp)	Exon (bp)	Sequence Read (bp)	
SOCS3ex1	640	328	-110	+180
SOCS3ex2a	790	766	-225	455
SOCS3ex2b	700	766	175	+85

5.3.7. Analysis of *GAL3ST2*

Exons of *GAL3ST2* gene were sequenced in patient 5-401 (Table 5.11). No mutation was identified.

Table 5.11. The extent of sequencing of *GAL3ST2* in Larsen patient 5-401.

Primer Name	PCR Product (bp)	Exon (bp)	Sequence Read (bp)	
GAL3ST2ex1	423	160	-100	+71
GAL3ST2ex2	334	90	-104	+63
GAL3ST2ex3	586	256	-169	+72
GAL3ST2ex4	1108	864	-25	+35

5.3.8. Analysis of *CANT1*

Exons of *CANT1* were sequenced in patients 1-401, 4-602, 5-401, 6-403, 6-604, 7-401 (Table 5.12). c.835G→A transition in exon 3 was identified in homozygous state in patients 6-403 and 6-604 (Figure 5.36). This variant was not reported in SNP or mutation databases. It results in the substitution of polar hyrodphilic serine for aliphatic neutral glycine (p.G279S). The residue was observed to be evolutionary conserved in mammals and zebrafish but not in chicken (Figure 5.37). p.G279S was predicted to affect the protein structure and function, by online SIFT tool and SNPs3D. MMB reported that it was neutral low reliability (Table 5.13).

Table 5.12. The extent of sequencing of *CANT1* in LRS patient 6-403.

Primer Name	PCR Product (bp)	Exon (bp)	Sequence Read (bp)	
CANT1ex1	495	93	-50	+120
CANT1ex2	1270	972	-65	+57
CANT1ex3	696	204	-153	+180
CANT1ex4	1124	371	-285	+253

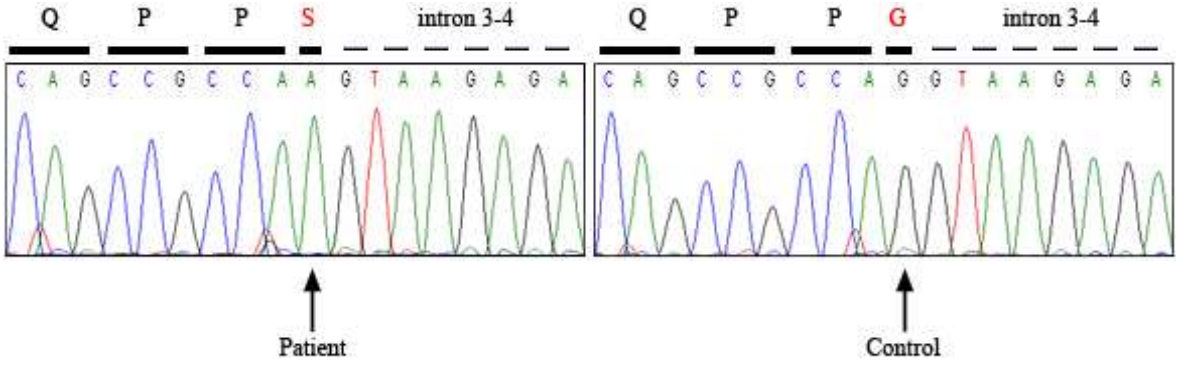


Figure 5.36. Chromatograms showing c.835G→A (p.G279S) in LRS patient 6-403 and the reference sequence.

```
                                  G279
Human         261  NWVSNYNALRAAAGIQPPGYLIHESACWSDTLQRWFFLPRRASQERYSEK  310
Chimpanzee    261  NWVSNYNALRAAAGIRPPGYLIHESACWSDTLQRWFFLPRRASQERYLPR  310
Macaque       261  NWVSNYNALRAAAGIRPPGYLIHESACWSDTLQRWFFLPRRASQERYSEK  310
Mouse         260  NWVSSYNALRAAAGIRPPGYLIHESACWSDTLQRWFFLPRRASHERYSEK  309
Cattle        261  NWVSSYNALRAAAGIRPPGYLIHESACWSDTLQRWFFLPRRASHERYSEK  310
Chicken       257  NWVVNYNALRAAAGIRPP------------------------------  279
Zebrafish     264  NWVPTYNSLRSAAGISPPGYLIHESAAWSDTLQRWFFLPRRASSERYDET  313
```

Figure 5.37. Evolutionary conservation of CANT1 G279.

Table 5.13. Effect of p.G279S on CANT1 protein function as predicted by online tools. Protein Accession is NP_620148.

Database	MMB	SNPs3D	SIFT
Score	0.43	-2.53	0
Effect	Neutral with low reliability	Intolerable amino acid change	Affects Protein function

c.898C→T transition was identified in homozygous state in patients 1-401 and 7-401 in exon 4 (Figure 5.38). This variant results in the substitution of polar hydrophobic cysteine for polar hydrophilic arginine (p.R300C). The variant has been reported to cause Desbuquois Dysplasia (Huber *et al.*, 2009).

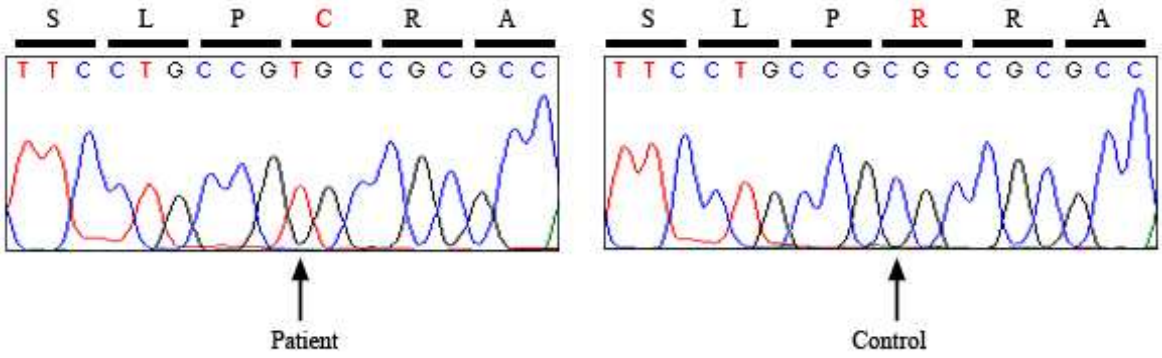


Figure 5.38. Chromatograms showing c.898C→T (p.R300C) in LRS patient 1-401 and the reference sequence in a control sample.

5.3.9. Summary of Results for Larsen Syndrome

No homozygous region common for all Larsen Syndrome families could be identified.

At 2q37.2-qter patients 1-401, 3-403 and 5-401 were found homozygous. *GAL3ST2* at the region was sequenced, but no mutation was identified.

At 17q25.3 a common homozygous region was detected. The borders of the homozygous region could not be delineated because the microsatellite markers that had been employed for genotyping were not sufficiently informative. *SOCS3* and *TIMP2* were sequenced but no mutation was identified. The homozygosity at the region could be delineated using the SNP genome scan data. *CANT1*, which was previously reported as mutated in Desbuquois Dysplasia, is located at the locus where patients 1-401, 2-401, 6-403, 6-404 and 7-401 were homozygous. Homozygous c.835G→A was identified in patients 6-403 and 6-404. Homozygous c.898C→T was identified in patients 1-401 and 7-401. Patients 1-401, 2-401 and 7-401 were found to share the same homozygous genotype at the loci, thus patient 2-401 was expected to carry c.898C→T.

5.4. Congenital Cerebellar Hypoplasia

5.4.1. Microsatellite Genome Scan Data Analysis

Genome scan data from two Congenital Cerebellar Hypoplasia (CCLH) families were subjected to two- and multipoint lod score analyses assuming autosomal recessive inheritance (Figure 3.5). Total two-point lod scores >1.5 were found as 1.99 at GGAA6D03 (7q31.32, 128.41 cM), 1.93 at GATA5D08 (7q12.3, 109.12), 1.81 at GATA63F09 (11q11, 58.41 cM), 1.81 at GATA23C03 (13q12.13, 8.87 cM) and 1.53 at ATA18A07 (9q22.33, 104.48 cM, Figure 5.39). Total multipoint lod scores were calculated as 2.23 at GATA23C03, at which one of the highest two-point lod scores was obtained, 1.86 between GATA48B01 and GATA124F08 (1q32.1-32.2, 217.93 cM), 1.64 at interval AAT243 – GATA22D12 (1q43, 263.63 cM), 1.16 at GGAA6D03 and 1.04 at interval GTAT1A05 – GATA158H04 (17p13.3-p13.2, 9.04 cM, Figure 5.40). Haplotype segregation analysis was also performed together with lod scores.

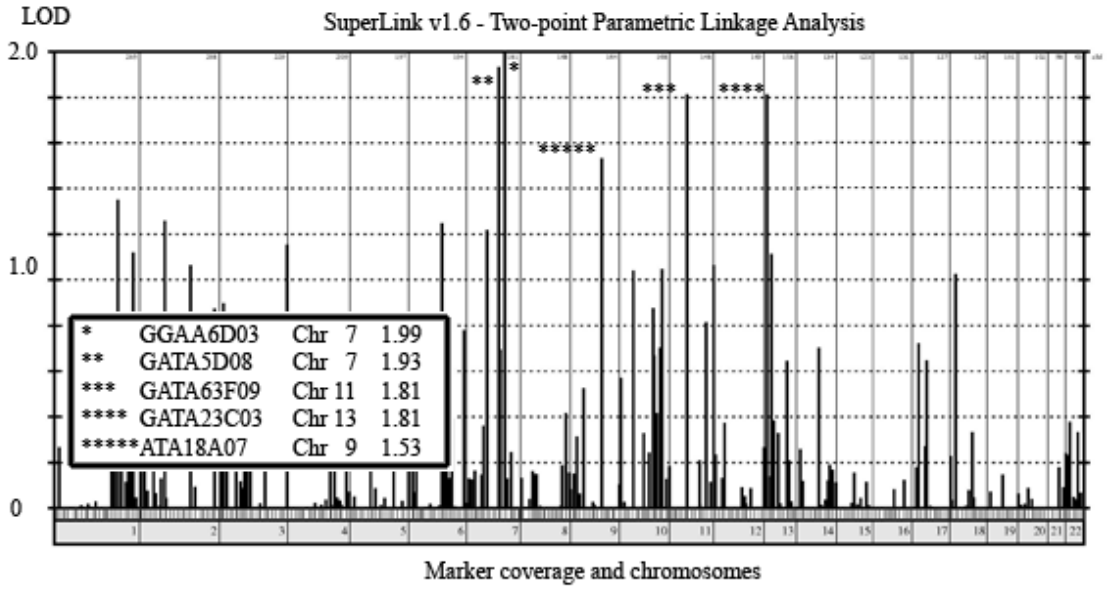


Figure 5.39. Two-point lod score graphics of initial microsatellite genome scan data for CCLH families. Only scores >0 are shown. Loci yielding >1.5 are depicted with asterisks.

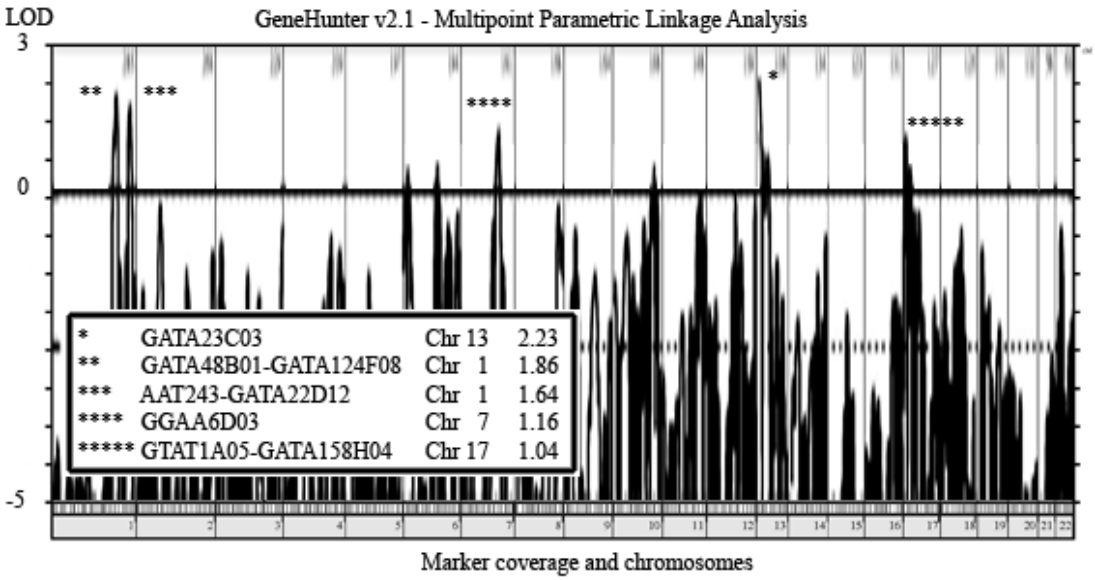


Figure 5.40. Multipoint lod score graphics of initial microsatellite genome scan data for CCLH families. Only scores > -5 are shown. Loci yielding >0.5 are depicted with asterisks.

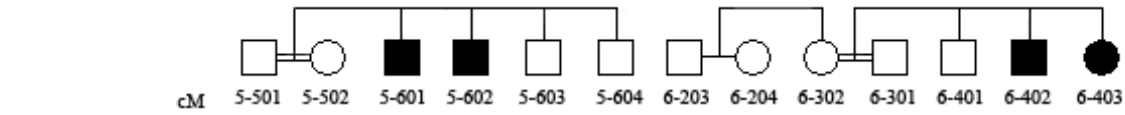
Nine loci that were assessed as significant and selected for further studies as well as *VLDLR* gene (9p24, 2,621,793 bp – 2,654,485 bp) that was reported to cause autosomal recessive cerebellar hypoplasia are shown in Table 5.14 (Özçelik *et al.* 2008). Linkage to this locus was excluded for each family with additional markers (Figure 5.42).

Table 5.14. Loci studied with microsatellites and linkage analysis in CCLH families

Locus	**Family**	**Position (Mb)**	**Maximum lod score**	
			Two-point	**Multipoint**
1q32.1	CCLH5-6	198 – 203	1.5	2.2
2p21-13.1	CCLH6	50 – 73	1.67	1.92
6q16.1-23.1	CCLH5	94 – 131	1.47	3.24
9p24.2-21.3	CCLH5-6	2 – 25	0.26	-2.05
10p14-12.33	CCLH5-6	9 – 14	1.16	-1.15
10q24.32-26.2	CCLH5-6	103 – 129	1.09	-0.42
13q13.1-21.1	CCLH5-6	31 – 55	2.2	1.2
16q21-23.3	CCLH6	60 – 82	1.67	1.92
17pter-q21.31	CCLH5-6	0 - 42	2.24	-0.4
18q12.1-21.2	CCLH5	26 – 50	1.47	1.01

5.4.2. Linkage Studies for Congenital Cerebellar Hypoplasia at Common Loci

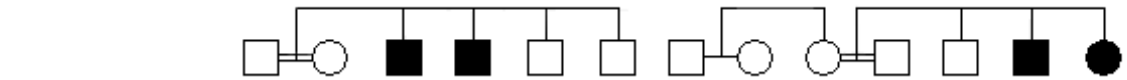
Assuming that the same gene was responsible for the disorder in both families, loci possibly candidate for both were analyzed further. At 1q32.1 six additional microsatellite markers were utilized to genotype (Figure 5.41). The maximum two-point lod score was calculated as 1.5 at D1S510 and total multipoint lod score as 2.02 at 212.43 cM between D1S413 and GATA48B01. Linkage to this locus was excluded. A common homozygosity for patients of Family CCLH6 was shown between GATA7C01 and D1S373.



	cM	5-501	5-502	5-601	5-602	5-603	5-604	6-203	6-204	6-302	6-301	6-401	6-402	6-403
TATC028	180.91	3 3	3 1	3 1	3 3	3 1	3 1	*3 0*	2 3	3 3	1 3	3 3	1 3	1 3
ATA4E02	192.05	3 3	2 4	3 4	3 2	3 4	3 4	*2 0*	3 4	2 4	3 1	1 2	1 4	3 4
GATA7C01	202.19	3 5	2 4	3 4	3 2	5 4	5 4	*1 0*	3 4	1 4	3 5	5 1	5 4	3 4
D1S413	212.43	1 3	3 1	1 3	1 3	3 1	3 1	*1 0*	3 2	1 3	3 1	1 1	3 3	3 3
GATA48B01	212.44	1 1	2 3	1 2	1 2	1 3	1 3	*3 0*	2 1	3 2	2 1	1 3	2 2	2 2
D1S373	214.08	4 1	2 5	4 2	4 2	1 5	1 5	*3 0*	7 5	3 7	5 6	6 3	5 7	5 7
D1S2738	215.17	1 2	2 3	1 2	1 2	2 3	2 3	*2 0*	2 2	2 2	2 2	2 2	2 2	2 2
D1S306	215.27	4 1	1 3	4 1	4 1	1 3	1 3	*2 0*	3 1	2 3	1 1	1 2	1 3	1 3
D1S2615	215.37	2 3	4 3	2 4	2 4	3 3	3 3	*4 0*	1 3	4 1	3 3	3 4	3 1	3 1
D1S510	218.46	3 3	3 1	3 3	3 3	3 1	3 1	*2 0*	4 5	2 4	5 2	2 2	5 4	5 4
GATA124F08	226.16	4 3	4 2	4 4	4 4	3 2	4 2	*1 0*	5 2	1 5	2 4	4 5	2 5	2 1
GATA4H09	239.66	1 1	4 1	1 4	1 4	1 1	1 4	*3 0*	2 3	3 2	3 2	2 2	3 2	3 3
ATA29C07L	247.23	4 1	3 3	4 3	4 3	1 3	4 3	*2 0*	1 5	2 1	1 3	3 1	1 1	1 2

Figure 5.41. Haplotypes of Family CCLH5 and CCLH6 at 1q32.1.

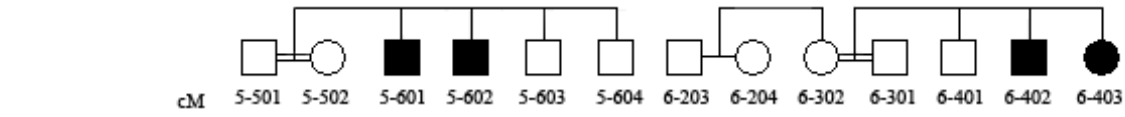
9p24.2-21.3 where *VLDLR* is localized was further analyzed with 7 additional markers (Figure 5.42). Linkage to this locus was excluded with both the haplotype data and multipoint lod score. The maximum two-point lod score was 0.26 at GATA187D09 (21.99 cM), and total multipoint lod score was -2.05 at interval AGAT142 – D9S1870 (33.78 cM).



	cM		5-501	5-502	5-601	5-602	5-603	5-604	6-203	6-204	6-302	6-301	6-401	6-402	6-403
MFD455-AAT052	0.11		3 2	4 2	2 4	2 4	2 2	3 2	2 0	2 1	2 1	2 4	4 2	4 1	2 2
D9S2572kb	4.01	*	2 1	2 2	1 2	1 2	1 2	1 2	2 0	2 2	2 2	2 2	2 2	2 2	2 2
D9S2622kb	4.41		2 2	3 2	2 3	2 3	2 2	2 3	1 0	2 1	1 1	1 2	1 1	2 1	1 1
D9S1813	9.83		4 4	1 1	4 1	4 1	4 1	4 1	2 0	1 1	2 1	3 1	3 2	1 1	3 2
aaaac001	10.25		2 4	4 3	4 4	4 4	4 3	4 4	3 0	3 4	3 4	3 1	3 3	1 4	3 3
GATA62F03M	14.23		3 2	1 2	2 1	2 1	2 2	2 1	1 0	4 3	1 3	4 2	4 1	2 3	2 1
D9S775	21.82		2 4	3 3	4 3	4 3	2 3	4 3	1 0	4 4	1 4	2 3	2 1	3 4	3 1
GATA187D09N	21.88		3 3	3 2	3 3	3 3	3 2	3 3	3 0	3 1	3 1	2 2	2 3	2 1	2 3
D9S921	21.91		1 3	1 4	3 1	3 1	1 4	3 1	5 0	0 2	5 2	1 6	1 5	6 2	6 5
AGAT142P	31.24		2 2	2 3	2 2	2 2	2 3	2 3	3 0	1 2	3 2	4 4	4 2	4 2	4 3
D9S1870	37.58		1 1	2 5	1 2	1 2	1 5	1 2	4 0	3 2	4 2	1 2	1 2	2 4	2 4
GATA87E02N	44.28		3 1	1 1	1 1	1 1	3 1	1 1	3 0	3 2	3 2	3 2	3 2	3 3	2 3
D9S126	44.58		3 1	2 3	1 2	1 3	3 3	1 2	1 0	2 1	1 1	1 1	1 1	1 1	1 1
GATA5E06P	52.76		5 1	3 6	1 3	1 6	5 6	1 3	5 0	2 5	5 5	2 4	2 5	2 5	4 5
GATA7D12	66.32		4 4	5 3	4 5	4 3	4 3	4 5	2 0	1 4	2 4	1 1	1 2	1 2	1 2
GATA89A11	75.88		4 2	5 1	2 5	2 1	4 1	2 5	3 0	3 4	3 3	3 3	3 3	3 3	3 3

Figure 5.42. Haplotypes of Family CCLH5 and CCLH6 at 9p24.2-21.3. Asterisk indicates the approximate location of *VLDLR*.

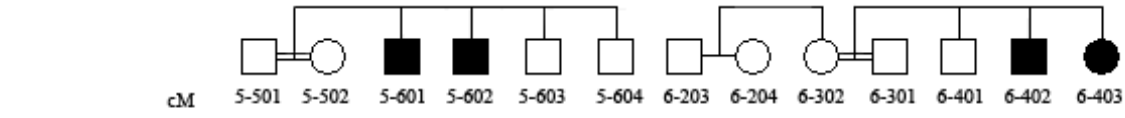
Two loci were studied on chromosome 10 (Figure 5.43). Four microsatellite markers were used for genotyping at 10p14-12.33. The maximum two-point lod score was obtained as 1.16 at GATA70E11 and total multipoint lod score as -1.15 at interval D10S1725 – GATA70E11 (44.2 cM). At the second locus 10q24.32-26.2, three microsatellite markers were used. The total maximum scores were calculated as 1.09 at ATA29C03 for two-point and -0.42 at 129.34 cM between GATA64A09 and D10S1683 for multipoint. Linkage to both loci was excluded. A common homozygosity for Family CCLH6 patients was detected between GATA87G01 and D10S2473.



	cM	5-501	5-502	5-601	5-602	5-603	5-604	6-203	6-204	6-302	6-301	6-401	6-402	6-403
GATA88F09	6.32	4 2	2 3	2 2	2 2	4 2	2 3	*1 0*	2 4	1 4	4 4	4 1	4 1	4 1
ATCC001	17.21	1 1	2 1	1 2	1 2	1 2	1 1	*1 0*	1 1	1 1	1 2	2 1	2 1	1 1
ATA31G11P	25.20	2 5	1 1	2 1	5 1	2 1	2 1	*3 0*	1 1	3 1	3 4	4 3	3 3	3 3
D10S1216	30.00	2 3	1 4	3 1	3 1	2 4	2 4	*4 0*	4 2	4 2	5 4	4 4	5 4	5 4
D10S1705	32.80	2 2	1 2	2 1	2 1	2 1	2 2	*0 0*	*0 0*	*0 0*	*0 0*	0 0	*0 0*	*0 0*
GATA84C01	33.18	2 2	1 4	2 1	2 1	2 1	2 4	*3 0*	5 6	3 6	1 4	4 3	1 3	1 3
D10S527	33.48	2 1	3 1	1 3	1 3	2 3	2 1	*2 0*	2 2	2 2	1 2	2 2	1 2	1 2
D10S1725	36.06	1 3	2 1	3 2	3 2	1 2	1 1	*1 0*	1 1	1 1	3 1	1 1	3 1	3 1
GATA70E11	46.23	3 1	4 2	1 4	1 4	3 4	3 2	*2 0*	2 1	2 1	2 2	2 1	2 2	2 2
GATA73E11	59.03	2 2	3 2	2 3	2 3	2 3	2 3	*1 0*	2 3	1 3	2 3	3 3	2 1	2 3
ATA5A04N	63.30	3 3	2 1	3 1	3 2	3 2	3 2	*2 0*	2 3	2 3	3 2	2 2	3 2	3 3
ATA21A03	75.57	1 1	3 2	1 2	1 3	1 3	1 3	*3 0*	3 3	3 3	1 1	1 3	1 3	1 3
ATA24F10	80.77	2 1	1 3	1 3	1 1	2 1	2 1	*3 0*	1 1	3 1	3 3	3 1	3 1	3 1
GATA121A08N	88.41	1 1	3 3	1 3	1 3	1 3	1 3	*3 0*	3 3	3 3	2 3	3 3	2 3	2 3
GATA87G01	93.92	5 2	3 2	2 2	5 3	5 3	5 3	*1 0*	1 3	1 3	4 3	3 3	4 3	4 3
GGAT1A4	100.92	2 1	2 2	1 2	2 2	2 2	2 2	*1 0*	2 2	1 2	2 2	2 2	2 2	2 2
GATA115E01N	112.58	6 3	4 5	3 4	6 4	6 4	6 5	*6 0*	4 1	6 1	1 2	2 1	1 1	1 1
GGAA2f11N	117.42	3 2	3 4	2 3	3 3	3 3	3 4	*2 0*	1 3	2 3	3 3	3 2	3 3	3 3
GATA64A09	125.41	1 3	2 5	3 2	1 2	1 2	1 5	*2 0*	5 2	2 2	2 4	4 2	2 2	2 2
D10S1683	135.24	3 5	4 3	3 4	3 4	3 4	3 3	*2 0*	1 1	2 1	1 5	5 2	1 1	1 1
GATA71C09	136.67	1 1	2 1	1 2	1 2	1 2	1 1	*3 0*	1 1	3 1	1 1	1 3	1 1	1 1
D10S2473	138.47	1 3	2 1	1 2	1 2	1 2	1 1	*1 0*	1 4	1 4	1 2	2 1	1 4	1 4
ATA29C03	142.78	3 2	1 2	3 1	3 1	3 2	3 2	*2 0*	1 3	2 3	1 2	2 2	1 3	1 3
D10S1213	148.17	3 4	5 5	3 5	4 5	3 5	3 5	*1 0*	4 2	1 2	4 5	5 1	4 2	4 2
ATA22D02	156.27	3 3	3 3	3 3	3 3	3 3	3 3	*1 0*	1 2	1 2	1 3	3 1	1 2	3 2
GGAA23C05N	165.27	3 3	2 1	3 2	3 1	3 2	3 1	*4 0*	3 4	4 4	3 4	4 4	3 4	4 4
ATGT006Z	167.35	2 4	3 3	2 3	4 3	2 3	4 3	*1 0*	2 2	1 2	2 1	1 1	2 2	1 2

Figure 5.43. Haplotypes of Family CCLH5 and CCLH6 at 10p14-12.33 and 10q24.32-26.2.

Six microsatellite markers were used for genotyping at 13q13.1-21.1 (Figure 5.44). Maximum two-point lod score of 2.22 at D13S893 and total multipoint lod score of 1.2 at 34.59 cM, between GATA86H01 and D13S765, was obtained. Linkage for the region was excluded. A shared homozygous region between ATA5A09N and GATA11C08 was determined for Family CCLH6 patients.



	cM	5-501	5-502	5-601	5-602	5-603	5-604	6-203	6-204	6-302	6-301	6-401	6-402	6-403
GATA23C03P	8.87	3 3	3 2	3 3	3 3	3 3	3 2	*3 0*	3 3	3 3	3 1	1 3	3 3	3 3
ATA5A09N	20.44	1 2	2 3	2 2	2 3	1 2	1 3	*4 0*	3 3	4 3	3 3	3 4	3 3	3 4
D13S893	22.01	*0* 2	2 2	2 2	2 2	2 2	*0 2*	*3 0*	1 2	3 1	1 2	2 3	1 1	1 1
D13S30977kb	24.01	1 2	3 4	2 3	2 4	2 3	2 4	*6 0*	5 3	6 5	5 3	3 6	5 5	5 5
GGAA29H03N	25.80	1 3	2 2	1 2	1 2	3 2	3 2	*6 0*	5 4	6 5	5 5	5 6	5 5	5 5
D13S305	28.04	2 1	4 4	2 4	2 4	1 4	1 4	*3 0*	3 3	3 3	3 3	3 3	3 3	3 3
D13S35923kb	30.01	3 5	4 1	3 4	3 1	5 4	5 1	*7 0*	6 2	7 6	6 1	1 7	6 6	6 6
GATA86H01	32.90	2 2	1 1	2 1	2 1	2 1	2 1	*2 0*	1 2	2 1	1 1	1 2	1 1	1 1
D13S765	35.01	1 2	2 3	1 2	1 2	2 2	2 3	*3 0*	1 2	3 1	1 4	4 3	1 1	1 1
D13S325	38.96	1 1	4 3	1 4	1 3	1 4	1 3	*1 0*	1 5	1 1	5 2	2 1	5 1	2 1
GATA11C08P	47.19	2 2	2 2	2 2	2 2	2 2	2 2	*1 0*	3 1	1 3	1 3	3 1	1 3	3 3
GATA64F08	55.31	3 3	1 2	3 1	3 2	3 1	3 1	*3 0*	2 4	3 2	4 5	5 2	5 2	5 2
GATA7G10	63.90	1 5	4 3	1 3	1 3	5 4	5 4	*0 0*	2 2	*0 2*	2 6	6 2	6 2	6 2

Figure 5.44. Haplotypes of Family CCLH5 and CCLH6 at 13q13.1-21.1.

At 17pter-q21.31, 39 additional microsatellite markers were used (Figure 5.45). Highest lod scores were calculated as 2.24 for two-point at D17S804 and -0.4 for multipoint at 53 cM at interval D17S1294 – GATA25A04. Lod score and haplotype of healthy individual 5-603 together were used to rule out linkage.

	cM	5-501	5-502	5-601	5-602	5-603	5-604	6-203	6-204	6-302	6-301	6-401	6-402	6-403
D17S1308	0.63	2 2	1 1	2 1	2 1	2 1	2 1	*1 0*	1 2	1 1	2 2	2 1	2 1	2 1
D17S1840	1.29	3 4	3 2	3 2	4 2	3 3	3 2	*1 0*	1 3	1 1	4 5	5 1	4 1	5 1
D17S1533	3.99	6 4	1 4	6 4	4 4	6 1	6 4	*3 0*	*2 0*	3 2	2 5	5 3	2 3	5 2
D17S14870KB	4.31	4 3	1 3	4 3	3 3	4 1	4 3	*2 0*	1 1	2 1	1 3	3 2	1 2	3 1
D17S525	5.37	1 1	1 1	1 1	1 1	1 1	1 1	*2 0*	1 1	2 1	1 1	1 2	1 2	1 1
D17S1528	6.60	1 2	5 4	1 4	2 4	1 5	1 4	*1 0*	2 5	1 2	2 3	3 1	2 1	3 2
D17S2273kb	7.08	3 1	3 2	3 2	1 2	3 3	3 2	*2 0*	1 1	2 1	1 2	2 2	1 2	2 1
D17S1566	7.30	3 1	1 3	3 3	3 3	3 1	3 3	*2 0*	3 2	2 3	2 2	2 2	2 2	2 3
D17S919	9.92	3 1	2 3	3 3	3 3	3 2	1 3	*2 0*	3 3	2 3	2 3	3 2	3 2	3 3
D17S1828	10.02	2 1	1 3	2 3	2 3	2 1	2 3	*0 0*	*0 0*	*0 0*	*0 0*	*0 0*	*0 0*	*0 0*
D17S1584	10.72	1 1	1 1	1 1	1 1	1 1	1 1	*1 0*	1 1	1 1	1 1	1 1	1 1	1 1
D17S4318KB	11.11	1 4	1 3	1 3	1 3	1 1	1 3	*2 0*	2 5	2 2	2 3	3 2	3 2	3 2
D17S4470KB	11.32	3 2	2 2	3 2	3 2	3 2	2 2	*1 0*	2 1	1 2	1 1	1 1	1 1	1 2
D17S559	11.59	2 2	1 2	2 2	2 2	2 1	2 2	*2 0*	4 4	2 4	3 1	1 2	1 2	1 4
D17S5470KB	12.61	3 2	3 1	3 1	3 1	3 3	2 1	*2 0*	3 3	2 3	2 2	2 2	2 2	2 3
D17S1537	13.12	1 2	1 3	1 3	1 3	1 1	2 3	*2 0*	1 2	2 1	2 1	1 2	1 2	1 1
D17S513	13.15	3 1	1 1	3 1	3 1	3 1	1 1	*3 0*	1 1	3 1	2 3	3 1	3 1	3 1
D17S1832	13.17	1 1	3 1	1 1	1 1	1 3	1 1	*2 0*	1 2	2 1	2 3	3 1	3 1	3 2
D17S938	14.62	2 5	4 6	2 6	2 6	2 4	5 6	*5 0*	6 1	5 6	4 3	3 6	3 6	3 5
D17S796	14.64	3 4	2 1	3 1	3 1	3 2	4 1	*1 0*	1 5	1 1	5 3	3 1	3 1	3 1
G10693	14.65	3 4	1 3	3 3	3 3	3 1	4 3	*2 0*	4 3	2 4	3 2	2 2	2 2	2 4
D17S570	14.66	1 1	1 1	1 1	1 1	1 1	1 1	*1 0*	1 1	1 1	1 2	2 1	2 1	2 1
D17S1881	14.67	1 3	2 3	1 3	1 3	1 2	3 3	*0 0*	*0 0*	*0 0*	*0 0*	*0 0*	*0 0*	*0 0*
D17S516	14.68	1 1	1 1	1 1	1 1	1 1	1 1	*1 0*	1 1	1 1	1 1	1 1	1 1	1 1
D17S906	14.69	4 4	3 6	4 6	4 6	4 3	4 6	*7 0*	5 1	7 5	1 2	2 7	2 7	2 5
D17S578	14.71	1 2	1 1	1 1	1 1	1 1	2 1	*0 0*	*0 0*	*0 0*	*0 0*	*0 0*	*0 0*	*0 0*
D17S1353	15.77	2 3	1 3	2 3	2 3	3 1	2 3	*1 0*	4 2	1 4	2 1	1 1	1 1	1 4
D17S1796	15.78	1 1	4 1	1 1	1 1	1 1	1 4	*2 0*	3 2	2 3	3 3	3 2	3 2	3 3
D17S786	17.38	1 2	2 2	2 2	2 2	2 2	1 2	*3 0*	2 2	3 2	3 4	4 3	4 3	4 2
D17S973	18.12	2 2	2 2	2 2	2 2	2 2	2 2	*2 0*	2 2	2 2	2 1	1 2	1 2	1 2
D17S1541	20.52	2 3	3 3	3 3	3 3	3 3	2 3	*2 0*	3 2	2 3	2 1	1 2	1 2	1 3
D17S945	21.01	4 2	3 1	2 3	2 3	2 1	4 3	*0 0*	*0 0*	*0 0*	*0 0*	*0 0*	*0 0*	*0 0*
D17S804	21.03	3 1	1 2	1 1	1 1	1 2	3 1	*1 0*	1 3	1 1	2 1	1 1	1 1	1 1
D17S520	22.21	5 1	3 2	1 3	1 3	1 2	5 3	*1 0*	4 4	1 4	3 2	2 1	2 1	2 4
D17S974	22.24	1 3	3 2	3 3	3 3	3 2	1 3	*2 0*	3 3	2 3	2 3	3 2	3 2	3 3
D17S1303	23.56	1 1	3 1	1 3	1 3	1 1	1 3	*3 0*	1 3	3 1	0 2	2 3	2 3	2 1
ATA78D02N	31.96	5 3	3 2	5 3	3 3	3 2	5 3	*1 0*	4 3	1 4	4 1	1 1	1 1	1 4
D17S2196	44.62	2 3	3 4	2 3	3 3	3 3	2 3	*3 0*	1 5	3 1	4 3	3 3	3 1	3 1
GGAA9D03	49.61	1 1	1 1	1 1	1 1	1 1	1 1	*2 0*	3 4	2 3	2 2	2 2	2 3	2 3
D17S1294	50.74	1 1	1 1	1 1	1 1	1 1	1 1	*2 0*	3 4	2 3	2 2	2 2	2 3	2 3
GATA25A04	62.01	2 3	3 2	2 2	3 3	3 3	2 2	*3 0*	2 1	3 2	2 2	2 3	2 2	2 3
D17S1789	63.09	2 3	1 4	3 4	3 4	3 1	2 4	*3 0*	2 3	3 2	3 3	3 3	3 3	3 2
AAT245	76.57	4 2	4 5	4 5	2 5	2 4	4 5	*5 0*	3 1	5 3	4 5	5 5	4 5	5 3
GATA49C09N	82.00	2 3	4 2	2 2	3 2	3 4	2 2	*5 0*	1 5	5 1	5 6	6 5	5 5	6 1
ATA43A10M	89.32	1 4	1 4	1 4	4 4	4 4	1 4	*4 0*	4 2	4 4	3 2	2 4	3 4	2 4

Figure 5.45. Haplotypes of Family CCLH5 and CCLH6 at 17pter-q21.31.

No common homozygous locus could be identified for the two families together after microsatellite genome scan and additional genotyping studies; therefore, each CCLH family was analyzed separately.

5.4.3. Linkage Studies for Family CCLH5

Two additional loci 6q16.1-23.1 and 18q12.1-21.2 were studied for Family CCLH5 (Figure 5.46).

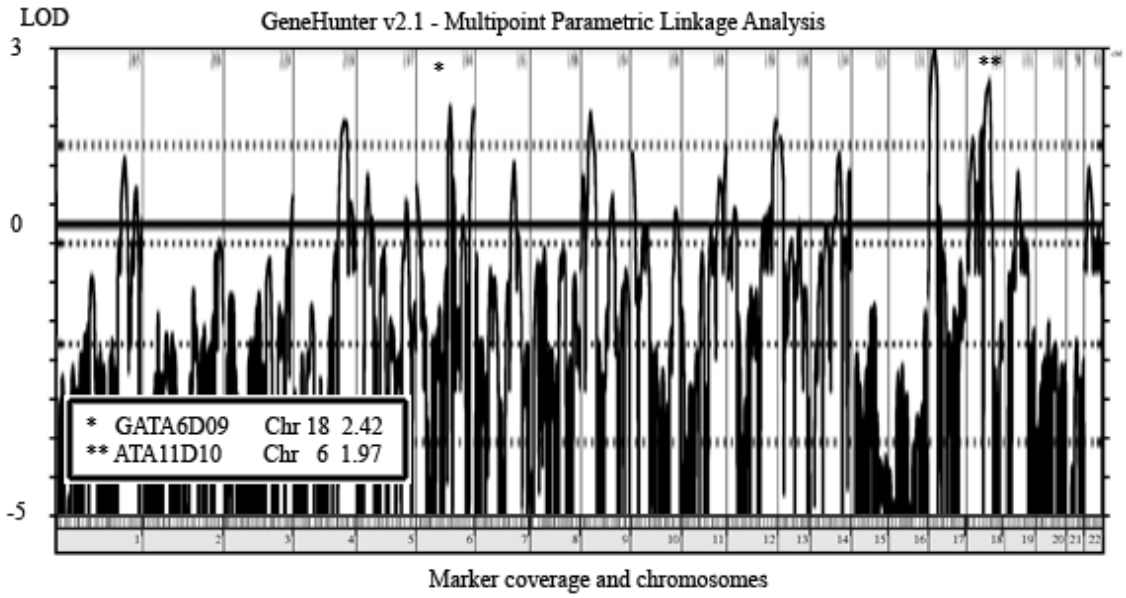
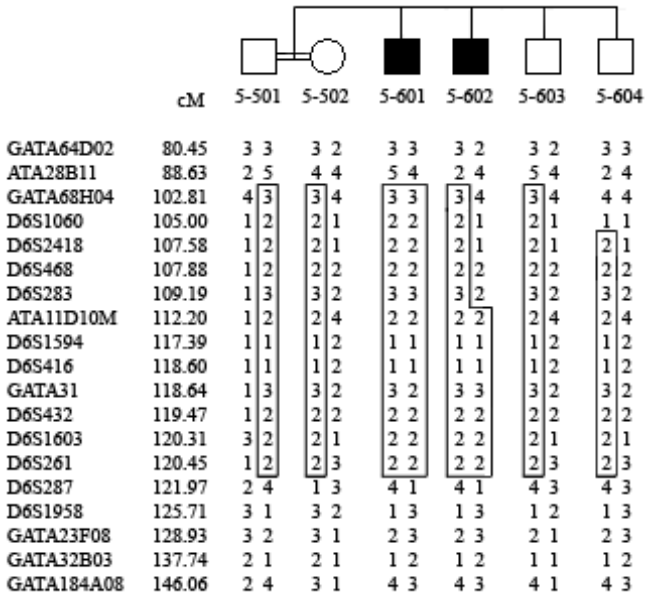


Figure 5.46. Multipoint lod score graphics of initial microsatellite genome scan of Family CCLH5. Only scores >-5 are shown. Loci chosen for further studies are depicted with asterisks.

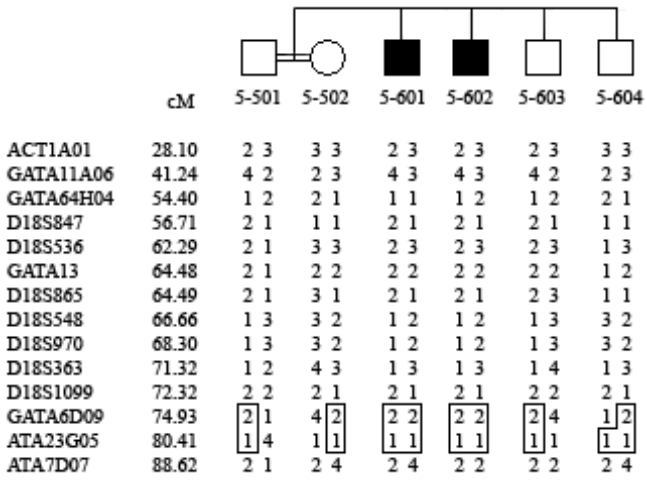
17pter at where highest lod score was obtained was genotyped previously (Figure 5.45). At 6q16.1-23.1, eleven microsatellite markers were utilized to genotype Family CCLH5 (Figure 5.47). Maximum two-point lod score of 1.47 at D6S1603 (120.31 cM) and multipoint lod score of 3.24 at D6S416 (118.6 cM) was calculated. The region between D6S283 (102,359,770 bp) and D6S287 (119,553,773 bp) was identified as a common homozygous region for the patients. The region spans 17.2 Mb and contains a very high number of genes, 154 in total. Candidate gene approach failed to point out a very strong candidate for further analysis from among the 154 genes.



	cM	5-501	5-502	5-601	5-602	5-603	5-604
GATA64D02	80.45	3 3	3 2	3 3	3 2	3 2	3 3
ATA28B11	88.63	2 5	4 4	5 4	2 4	5 4	2 4
GATA68H04	102.81	4 3	3 4	3 3	3 4	3 4	4 4
D6S1060	105.00	1 2	2 1	2 2	2 1	2 1	1 1
D6S2418	107.58	1 2	2 1	2 2	2 1	2 1	2 1
D6S468	107.88	1 2	2 2	2 2	2 2	2 2	2 2
D6S283	109.19	1 3	3 2	3 3	3 2	3 2	3 2
ATA11D10M	112.20	1 2	2 4	2 2	2 2	2 4	2 4
D6S1594	117.39	1 1	1 2	1 1	1 1	1 2	1 2
D6S416	118.60	1 1	1 2	1 1	1 1	1 2	1 2
GATA31	118.64	1 3	3 2	3 2	3 3	3 2	3 2
D6S432	119.47	1 2	2 2	2 2	2 2	2 2	2 2
D6S1603	120.31	3 2	2 1	2 2	2 2	2 1	2 1
D6S261	120.45	1 2	2 3	2 2	2 2	2 3	2 3
D6S287	121.97	2 4	1 3	4 1	4 1	4 3	4 3
D6S1958	125.71	3 1	3 2	1 3	1 3	1 2	1 3
GATA23F08	128.93	3 2	3 1	2 3	2 3	2 1	2 3
GATA32B03	137.74	2 1	2 1	1 2	1 2	1 1	1 2
GATA184A08	146.06	2 4	3 1	4 3	4 3	4 1	4 3

Figure 5.47. Haplotypes of Family CCLH5 at 6q16.1-23.1.

At 18q12.1-21.2, seven markers were analyzed for Family CCLH5 (Figure 5.48). Highest lod scores were found as 1.47 at AGAT060 (19.31 cM) and GATA6D09 (74.93 cM) for two-point and 1.01 at GATA6D09 for multipoint analysis. Linkage to this region was excluded with lod scores.



	cM	5-501	5-502	5-601	5-602	5-603	5-604
ACT1A01	28.10	2 3	3 3	2 3	2 3	2 3	3 3
GATA11A06	41.24	4 2	2 3	4 3	4 3	4 2	2 3
GATA64H04	54.40	1 2	2 1	1 1	1 2	1 2	2 1
D18S847	56.71	2 1	1 1	2 1	2 1	2 1	1 1
D18S536	62.29	2 1	3 3	2 3	2 3	2 3	1 3
GATA13	64.48	2 1	2 2	2 2	2 2	2 2	1 2
D18S865	64.49	2 1	3 1	2 1	2 1	2 3	1 1
D18S548	66.66	1 3	3 2	1 2	1 2	1 3	3 2
D18S970	68.30	1 3	3 2	1 2	1 2	1 3	3 2
D18S363	71.32	1 2	4 3	1 3	1 3	1 4	1 3
D18S1099	72.32	2 2	2 1	2 1	2 1	2 2	2 1
GATA6D09	74.93	2 1	4 2	2 2	2 2	2 4	1 2
ATA23G05	80.41	1 4	1 1	1 1	1 1	1 1	1 1
ATA7D07	88.62	2 1	2 4	2 4	2 2	2 2	2 4

Figure 5.48. Haplotypes of Family CCLH5 at 18q12.1-21.2.

5.4.4. Linkage Studies for Family CCLH6

Two additional loci, 2q14-13.3 and 16q22.1-23.3, were studied for Family CCLH6 (Figure 5.49).

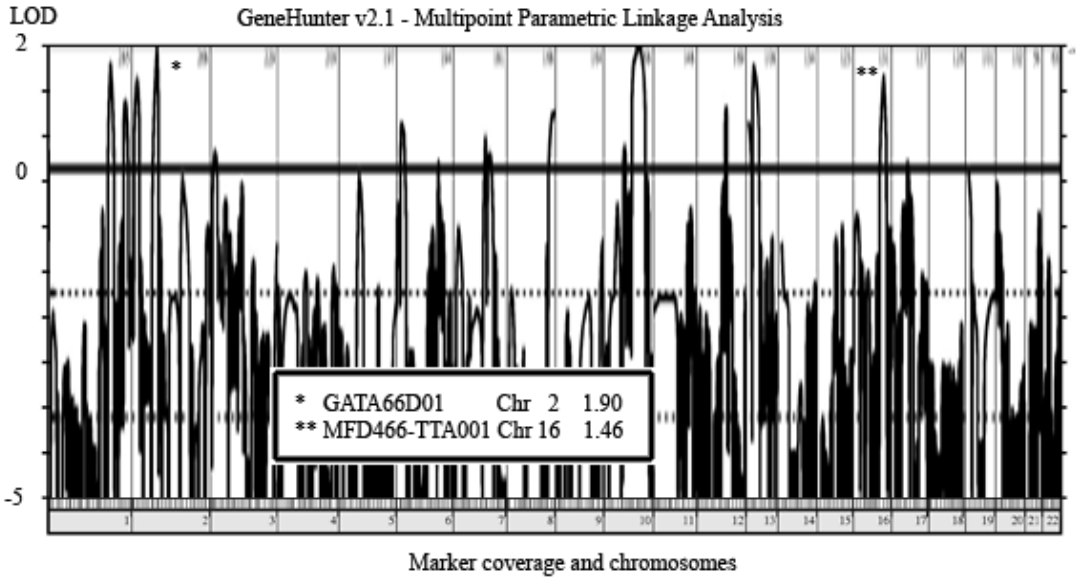
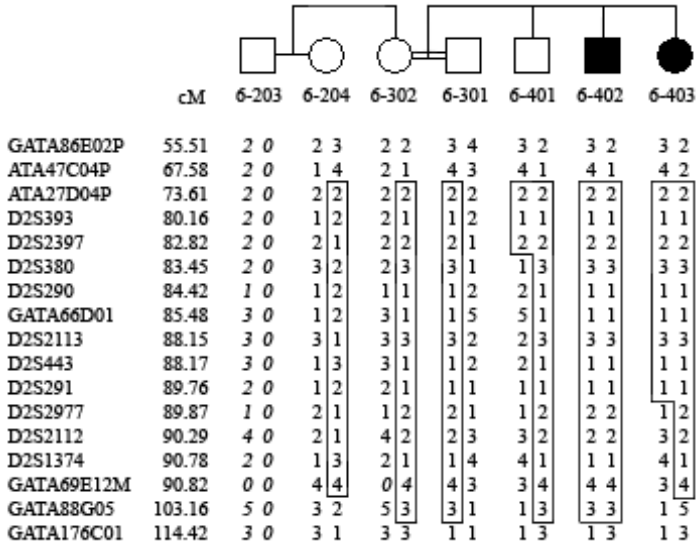


Figure 5.49. Multipoint lod score graphics of initial microsatellite genome scan of Family CCLH6. Only scores >-5 are shown. Loci chosen for Family CCLH6 are depicted with asterisks.

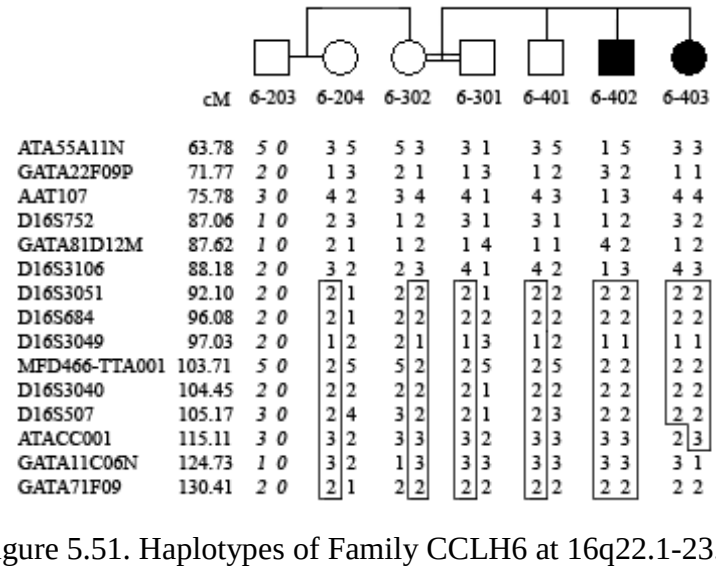
Chromosome 1 and 10 were previously genotyped (Figures 5.41 and 5.43). Data generated by genotyping Family CCLH6 with 10 additional microsatellite markers at 2q14-13.3 were used (Figure 5.50). At GATA66D01 (85.48 cM), maximum two-point lod score of 1.67 and multipoint lod score of 1.92 were calculated. A shared homozygosity in patients was shown on region between ATA47C04P and D2S2977. However, healthy sib 6-401 was also observed homozygous between ATA47C04P and D2S2397. A candidate region was identified between D2S2397 and D2S2977 (64,636,792 bp – 72,123,252 bp, respectively). The 7.5-Mb region contains 92 genes.



	cM	6-203	6-204	6-302	6-301	6-401	6-402	6-403
GATA86E02P	55.51	2 0	2 3	2 2	3 4	3 2	3 2	3 2
ATA47C04P	67.58	2 0	1 4	2 1	4 3	4 1	4 1	4 2
ATA27D04P	73.61	2 0	2 2	2 2	2 2	2 2	2 2	2 2
D2S393	80.16	2 0	1 2	2 1	1 2	1 1	1 1	1 1
D2S2397	82.82	2 0	2 1	2 2	2 1	2 2	2 2	2 2
D2S380	83.45	2 0	3 2	2 3	3 1	1 3	3 3	3 3
D2S290	84.42	1 0	1 2	1 1	1 2	2 1	1 1	1 1
GATA66D01	85.48	3 0	1 2	3 1	1 5	5 1	1 1	1 1
D2S2113	88.15	3 0	3 1	3 3	3 2	2 3	3 3	3 3
D2S443	88.17	3 0	1 3	3 1	1 2	2 1	1 1	1 1
D2S291	89.76	2 0	1 2	2 1	1 1	1 1	1 1	1 1
D2S2977	89.87	1 0	2 1	1 2	2 1	1 2	2 2	1 2
D2S2112	90.29	4 0	2 1	4 2	2 3	3 2	2 2	3 2
D2S1374	90.78	2 0	1 3	2 1	1 4	4 1	1 1	4 1
GATA69E12M	90.82	0 0	4 4	0 4	4 3	3 4	4 4	3 4
GATA88G05	103.16	5 0	3 2	5 3	3 1	1 3	3 3	1 5
GATA176C01	114.42	3 0	3 1	3 3	1 1	1 3	1 3	1 3

Figure 5.50. Haplotypes of Family CCLH6 at 2q14-13.3.

At 16q22.1-23.3, 7 markers were used for genotyping Family CCLH6 (Figure 5.51). Maximum two-point lod score 1.67 was obtained at D16S507 (105.17 cM). Highest multipoint lod score was 1.92 at D16S3040 (103.71 cM). A shared homozygosity was narrowed down to between D16S3106 and ATACC001. This region spans 10 Mb and contains 82 genes.



	cM	6-203	6-204	6-302	6-301	6-401	6-402	6-403
ATA55A11N	63.78	5 0	3 5	5 3	3 1	3 5	1 5	3 3
GATA22F09P	71.77	2 0	1 3	2 1	1 3	1 2	3 2	1 1
AAT107	75.78	3 0	4 2	3 4	4 1	4 3	1 3	4 4
D16S752	87.06	1 0	2 3	1 2	3 1	3 1	1 2	3 2
GATA81D12M	87.62	1 0	2 1	1 2	1 4	1 1	4 2	1 2
D16S3106	88.18	2 0	3 2	2 3	4 1	4 2	1 3	4 3
D16S3051	92.10	2 0	2 1	2 2	2 1	2 2	2 2	2 2
D16S684	96.08	2 0	2 1	2 2	2 2	2 2	2 2	2 2
D16S3049	97.03	2 0	1 2	2 1	1 3	1 2	1 1	1 1
MFD466-TTA001	103.71	5 0	2 5	5 2	2 5	2 5	2 2	2 2
D16S3040	104.45	2 0	2 2	2 2	2 1	2 2	2 2	2 2
D16S507	105.17	3 0	2 4	3 2	2 1	2 3	2 2	2 2
ATACC001	115.11	3 0	3 2	3 3	3 2	3 3	3 3	2 3
GATA11C06N	124.73	1 0	3 2	1 3	3 3	3 3	3 3	3 1
GATA71F09	130.41	2 0	2 1	2 2	2 2	2 2	2 2	2 2

Figure 5.51. Haplotypes of Family CCLH6 at 16q22.1-23.3.

5.4.5. SNP Genome Scan Data Analysis

SNP genome scan using 670K Illumina chip was completed for patient 5-601 and 6-403 after the microsatellite genotyping studies failed to point out a single candidate locus. The aim was if possible to determine a single, previously undetected locus or exclude some of the candidate loci established by microsatellite data, as well as to search for homozygous deletion or duplication at the candidate loci. Our previous microsatellite data with no common locus for the two CCLH families was confirmed with SNP genome scan. The results of SNP genome scan data combined with microsatellite genotyping data for CCLH families are given in Table 5.15.

Table 5.15. Combined results of SNP and microsatellite data for CCLH families.
(Shaded cells represent the borders of maximum homozygosity)

Family	Chromosome	Microsatellite Genotyping Marker (Location in bp)		SNP Genotyping SNP (Location in bp)	
CCLH5	6	D6S283 (102,359,770)	D6S287 (119,553,773)	rs711263 (91,190,244)	rs590572 (117,084,761)
CCLH6	1	GATA7C01 (187,550,288)	D1S373 (200,254,486)	rs12139999 (191,065,967)	rs1036332 (199,012,478)
	2	D2S2397 (64,636,792)	D2S2977 (72,123,252)	rs11903097 (50,623,839)	rs1006049 (71,176,660)
	10	GATA87G01 (74,659,396)	D10S2473 (120,169,137)	rs11203042 (90,989,109)	rs10748913 (108,047,595)
	13	ATA5A09 (31,072,681)	GATA11C08 (55,106,096)	rs4943320 (31,474,177)	rs1319281 (40,965,706)
	16	D16S3106 (72,187,760)	ATACC001 (82,249,557)	rs3751765 (77,910,418)	rs12599282 (80,424,384)

5.4.6. Copy Number Variation Analysis

CNV analysis was done, and no homozygous deletion/duplication that could possibly contribute to the etiology of the disease was identified in patients.

5.4.7. Summary of the Results for Congenital Cerebellar Hypoplasia

Genome-wide lod score analysis together with haplotype constructions were used for Congenital Cerebellar Hypoplasia families. No common candidate locus could be identified. Candidate loci for each family could be identified by analyzing each separately.

One candidate locus was identifed for family CCLH5, at 6q16.1-22.1 between D6S283 and rs590572 (102,359,770 bp and 117,084,761 bp, respectively). Candidate gene approach failed to point out a very strong candidate for further analysis from among the 130 genes.

Five candidate loci were identified for family CCLH6:

- 1p31.2-32.1, interval rs12139999 – rs1036332 (46 genes)
- 2q14-13.3, interval D2S2397 – rs1006049 (78 genes)
- 10q23.31-25.1, interval rs11203042 – rs10748913 (248 genes)
- 13q13.1-14.11, interval rs4943320 – rs1319281 (60 genes)
- 16q23.1-23.2, interval rs3721765 – rs12599282 (8 genes)

CNV analysis was done, and no homozygous deletion/duplication that could possibly contribute to the etiology of the disease was identified in patients.

6. DISCUSSION

Cost effective genotyping with high density microsatellite and/or SNP marker maps facilitate effective mapping of genes that are associated with human diseases. Mammalian Genotyping Service, funded by the National Heart, Lung and Blood Institute (NHLBI), had been performing genome scans with microsatellite markers to support worldwide studies of mapping genes for 12 years until December, 2006. The service accepted three genotyping projects from our laboratory and analyzed a total of 293 samples, with approximately 400 markers for each individual.

SNP based genome scans became more accessible in the last five years with the recent advancements in chip technologies. The large number of the SNPs distributed evenly throughout the genome made this new technology more attractive, especially for the studies that are in need of a high density mapping.

This study describes locus and gene analysis in three inherited disorders (JP, LRS and CCLH) submitted to the Mammalian Genotyping Service and one additional inherited disorder (ARA) submitted to SNP genotyping service after the preliminary studies.

Extended families with several affected members as a consequence of high incidence of consanguineous marriages in Turkey have provided an initiation point for this study. In all of the disorders within the framework of this study, the disease alleles could be assumed homozygous by descent and homozygosity mapping could be applied. Afterwards, in the identified loci, if possible, genes were selected by candidate gene approach and analyzed in relevant patients.

6.1. Autosomal Recessive Ataxia

Studies on autosomal recessive ataxia were initially started by Suna Önengüt and later continued by Sibel Aylin Uğur. A candidate locus was mapped at 12q21.1-21.2. In the present project, the family was genotyped again with some of the previously studied microsatellite markers to verify the locus. On March 2008, we were informed that a cousin

of the patients had been recently diagnosed with the disease. Locus 12q21.1-21.2 was excluded after genotyping of that patient (1-505) and her family. Later, SNP genome scan was performed in the two ARA families in an attempt to find a new candidate locus for the disease. Lod score analysis and HCiE studies were completed, but no candidate locus for all patients was identified. Then we changed the strategy and divided the patients into two groups according to the clinical findings.

For Group 1, we identified the disease locus at 9p21.1-p13.3. After candidate gene approach, *APTX*, known to be mutated in early-onset ataxia with oculomotor apraxia and hypoalbuminemia 1 (AOA1), was chosen for further studies. Early-onset ataxia was observed in all patients in Group 1; but oculomotor apraxia was only reported in patients in Family 2. We identified a different mutation in each family after sequencing *APTX*.

c.838G→A transition in exon eight was identified in patients of Family 1. This change results in the conversion of a tryptophan codon to a termination codon, p.W279X, and leads to premature termination of translation. The truncated protein lacks the zinc finger domain that is responsible for stabilizing the catalytic site of the protein, HIT domain, onto DNA target site. Clinical examination of Family 1 patients showed no oculomotor apraxia. This feature of the disease is observed in 80 per cent of the cases having mutations in the gene and is not essential for diagnosis, despite its being a part of the disease name. The mutation had been already described by Moreira *et al.,* 2001 and also reported to be responsible for low muscle CoQ10 enzyme levels in some patients (Le Ber *et al.,* 2007). We learned that CoQ10 levels had not been tested yet due to the high cost of the assay.

c.782T→C transition in *APTX* exon eight was identified in Group 1 Family 2 patients. It results in the substitution of nonpolar hydrophilic proline for aliphatic nonpolar hydrophobic leucine (p.L261P). We screened 108 control individuals randomly chosen from the general population and did not find this variant. With the assumption of a polymorphism frequency of one per cent, the 216 control chromosomes gave power greater than 80 per cent. The variant was not reported in SNP or mutation databases and was predicted to be damaging for protein structure and function with high reliability by three online tools, MMB, SNPs3D and SIFT. L261 residue is part of the catalytic HIT domain of

the protein and was found evolutionarily ultra-conserved across species. c.782T→C was concluded to be a novel mutation causing early-onset autosomal recessive ataxia. Oculomotor apraxia was reported for the patients with a clinical examination completed after the identification of the mutation. Muscle CoQ10 enzyme levels for those patients could not be assayed due to high cost.

For Group 2, three loci were identified by using both multipoint lod scores and HCiE. Regions where patients shared homozygosity were determined as 3p14.1, 6p13 and 10q25.3. The loci were inspected further by constructing haplotypes. Patient 2-609 was found heterozygous for two microsatellite markers at 6p13, thus excluding the locus as harboring the disease gene. At 10q25.3, none of the five genes were assessed as having a possible causative role for the disease. At 3p14.1, *ATXN7* was selected for sequencing but no sequence variant indicative of a disease-causing mutation was identified. In spring 2010, Patient 2-610 was found heterozygous at all three loci above after a SNP genome scan. The SNP genome scan chips for Group 2 patients were composed of different SNPs, thus a unified linkage input file could not be generated and lod scores including patient 2-610 could not be calculated. No shared homozygosity could be identified for Group 2 patients using HCiE and Illumina Genome Viewer.

CNV analysis was done for Group 2 patients. A high number of homozygous and heterozygous deletions scattered throughout the genome were detected in patients, but none of those were common to all patients. The increased number of deletions could be the result of impaired DNA repair mechanisms. Many diseases manifesting with ataxia such as A-T, AOA1 and AOA2 are known to be caused by mutations in genes that have various roles in DNA repair mechanisms.

ARA families presented with an autosomal recessive inheritance, and *ATXN7* is known to be mutated in autosomal dominant SCA7. We nevertheless sequenced all exons of the gene, as some genes can be responsible for similar diseases with alternative recessive or dominant inheritance patterns (Jonsson *et al*., 2002). Mutations in *Superoxide Dismutase 1* (*SOD1*) in Amyotrophic Lateral Sclerosis is an example for such cases. While *SOD1* mutations are presented with an autosomal dominant inheritance, one specific mutation p.D90A can be inherited as with an autosomal recessive (Jonsson *et al*., 2009).

Group 2 patients can be reevaluated with multipoint lod score analysis again in the future. This analysis may reveal previously undetected candidate regions.

6.2. Juvenile Parkinsonism

In this study, a family with an autosomal recessive neurological condition diagnosed as Juvenile Parkinsonism was investigated. Data from microsatellite genome scan were analyzed, using the initial pedigree structure composed of 3 patients and 13 healthy family members; however, no candidate locus could be identified by fine-mapping studies; therefore, a denser genome scan was necessary. SNP genome scan data pointed to a region that did not include any known exons. Meanwhile, one individual that was previously reported healthy was hospitalized with disease symptoms. We updated the pedigree and reanalyzed the microsatellite genome scan data and fine mapping results together with SNP genome scan data. Candidate gene approach was performed at two significant loci, 7q22.1-31.1 and 1p31.3-31.2, the former being stronger. Three genes *PDE4B*, *AK3L1* and *PRKAR2B* were selected for sequencing.

c.182A→T variant in *AK3L1* exon four resulting in p.S61C was identified in all four patients. One healthy individual, 303, was also found homozygous for the variant. She is about 70 years old and does not present the disease. SSCP analysis completed for 260 control chromosomes showed that the variant was unique for the family. Online tools SIFT and SNPs3D predicted that this change disrupts protein structure and function. Another tool, MMB, suggested that it is neutral for the protein but with relatively low reliability. S61 is conserved in hominoids, and the surrounding 14 residues are conserved in all mammals.

Three mechanisms are hypothesized to lead to PD: ubiquitin-proteasome pathway, mitochondrial dysfunction and oxidative stress. *AK3L1* is a member of adenylate kinase family of enzymes. Its protein product is localized to mitochondrial matrix and plays an important role in the regulation of adenine and guanine nucleotide compositions within a cell. Disturbance of *AK3L1* function in mitochondria may cause a phenotype similar to PD by disrupting the relative ratios of the energy nucleotides in mitochondria and the cell. However, if the variant was a fully penetrant and the sole disease-causing mutation in the

family, individual 303 would have been expected to present the disease already; thus, the variant in *AK3L1* alone was not enough to explain the etiology of JP in the family. Alternatively, there may be a protective locus that prevented a pathological condition in individual 303. Spinal muscular atrophy (SMA) is an example for such a model. The disease results from mutations in the *survival motor neuron 1* (*SMN1*) gene. Some families having homozygous *SMN1* deletions have been shown to present low penetrance for the disease. This is partly explained by the increased *Survival motor neuron 2* (*SMN2*) gene copy number in those families or high expression of *Plastin 3* (*PLS3*) gene (Lorson *et al.*, 1999; Oprea *et al.*, 2008).

Another possibility is that the disease may be exhibiting a complex pattern of inheritance, such as a digenic model. Digenic inheritance mechanism results from the interaction of two non-homologous genes. Bardet-Biedl Syndrome (BBS, MIM 209900) is a digenic and triallelic disease that is caused by a total of three mutations in two genes: two mutations in one and one in the other. Neither two mutant alleles at one locus nor one mutant allele alone at the other is sufficient for disease manifestation (Gropman and Adams, 2007).

Parkinsonism is generally observed at elderly people and progresses slowly; however, disease severity and age of onset in the family may be modified by other genes. Some of the phenotypes associated with cystic fibrosis (CF) is known to be modified by other genes. It is caused by mutations in *Cystic Fibrosis Transmembrane Conductance Regulator* (*CFTR*). A broad spectrum of disease severity is observed in patients with CF, and *CFTR* alone does not explain this diversity. It has been shown that there are many CF modifiers including *mannose-binding lectin, beta2-adrenergic receptor*, *tumor necrosis factor-alpha*, and *transforming growth factor-beta1* (Merlo and Boyle, 2003).

The possibility of linkage to the known Parkinson's Disease loci was excluded by both haplotype segregation studies and HCiE (Table 1.1).

CNV analysis was done, and no homozygous deletion/duplication that could possibly contribute to the etiology of the disease was identified in patients.

The presence of a protective/modifier locus or complex pattern of inheritance is possible for the family; however, examples for such cases are not common. In the future Family JP data must be reevaluated with other inheritance models to clarify the likelihood of those models. Status of p.S61C variant in *AK3L1* must be further verified by increasing the number of individuals in population screening. Also, genes on 7q22.1-31.1 must be reevaluated to determine new candidate genes.

6.3. Larsen Syndrome

The third disorder analyzed was Larsen Syndrome with autosomal recessive inheritance. We analyzed DNA samples of eight patients in seven families. Microsatellite genome scan was completed and followed by further genotyping for fine-mapping. Family LRS6 was taken as a starting point for candidate locus search studies since it was the most informative family. We could not identify a single locus for all patients together. The strongest locus was at 17q25.3 where 5 patients had a common homozygous region. Due to low informativeness of the microsatellite markers used for genotyping at the locus, we could not determine the exact borders at first. We sequenced *SOCS3* and *TIMP2* at the region, but no mutation was found. A denser genome scan was necessary, so Illumina 1M Quad analysis was ordered for all patients except 5-401 whose DNA quality was inadequate. SNP genome scan results confirmed the candidate locus and also determined the borders of the homozygosity for five patients from four families. The morbid map search and candidate gene approach revealed that *CANT1*, mutated in Desbuquois Dysplasia, was located in the region. Clinical features of Larsen Syndrome and Desbuquois Dysplasia are known to be very similar (Huber *et al.*, 2009). Therefore, *CANT1* was chosen for sequencing in patients.

c.898C→T in exon four, a known missense mutation in patients with Desbuquois Dysplasia from Turkey and Iran, was detected in patients 1-401 and 7-401. Patients 1-401, 2-401 and 7-401 were found to share the same homozygous genotype at the loci; thus, patient 2-401 also was expected to carry the mutation. In fact, expanding the history of those families revealed that they all originated from the same district of Anatolia, despite the fact that they had not described any relationship with each other. Having the same

mutation together with the same haplotype at the locus indicates that they likely have a common ancestor.

c.835G→A in exon three resulting in p.G279S was identified in Family LRS6. This variant was not reported in SNP or mutation databases. The residue was on the splice site of the protein and observed to be evolutionary conserved in mammals. It was predicted to affect the protein structure and function, by online SIFT tool and SNPs3D. MMB reported that it was neutral with low reliability. Additionally, it may lead to aberrant splice products. Population screening is required to confirm the state of variation, whether it is a new disease-causing mutation or a rare SNP.

Genetic heterogeneity is a known entity for Larsen syndrome. This may be the explanation for patients 3-403, 4-602 and 5-401 who do not have linkage to *CANT1* locus. However, we could not identify any common homozygous locus for those patients. We also analyzed data by grouping those patients in pairs but again this approach did not identify any candidate locus. Consanguineous parents with a single affected sib do not yield a significant lod score to point out a candidate locus separately.

Linkage to known Larsen Syndrome genes, *FLNB* and *CHST3* was eliminated for all patients except 5-401 with haplotype inspection of microsatellite and SNP genome scan data. Patient 5-401 was found homozygous at *CHST3* locus. Sequencing of this gene most likely will uncover whether the gene indeed is responsible for the disease in this family.

CNV analysis was completed for those patients with SNP genome scan, and no homozygous deletion/duplication that could possibly contribute to the etiology of the disease was identified in patients.

CANT1 p.R300C mutation has been reported in five families (Huber *et al.,* 2009). Population screening may give data for the frequency of the mutation. Also we identified that three families carrying the mutation have the same haplotype at the gene locus, thus a founder effect is present. Haplotyping of previously published other three patients from Turkey and those two Iranian patients may reveal whether all p.R300C mutations are descending from the same common founder.

Population screening for *CANT1* p.G279S variant must be completed to identify the state of the variant.

6.4. Congenital Cerebellar Hypoplasia

The last disorder analyzed was Congenital Cerebellar Hypoplasia (CCLH) in two families with the assumption of a common disease locus. Fine-mapping studies after initial microsatellite genome scan failed to provide linkage to a single common locus for the two families; therefore, we analyzed each family separately.

A SNP genome scan using 1M Illumina Quad chip was applied to one patient from each family with the purpose of eliminating or narrowing down the previously determined regions and analyzing the CNVs. Samples of other patients and family members could not be processed because either the quantity was not sufficient or the quality was not good enough.

The microsatellite genotyping studies singled out 6q16.1-23.1 as a candidate locus for Family CCLH5. SNP genome scan in 5-601 further narrowed down the region by about 2 Mb. The maximum shared homozygosity region between D6S283 and rs590572 at 6q16.1-23.1 was defined as the sole locus for Family CCLH5. A maximum two-point lod score of 1.47 at D6S1603 (120.31 cM) and multipoint lod score of 3.24 at D6S416 (118.6 cM) were obtained. Spanning 15 Mb and harboring 130 genes, candidate gene approach at the region failed to point out a gene for further analysis. Those genes can be reevaluated using the morbid database to continue with the study.

Family CCLH6 did not yield a single candidate locus by the microsatellite based linkage studies because it was not sufficiently large. At five loci patients were found to share homozygosity. SNP scan of 6-403 did not rule out any of the previously determined loci but facilitated the narrowing down of the loci (Table 5.15).

CNV analysis was done, and no homozygous deletion/duplication that could possibly contribute to the etiology of the disease was identified in patients.

We also excluded the possibility of linkage in both families to seven genes that were known to be responsible for cerebellar hypoplasia by lod scores and haplotypes segregation studies (Table 1.3).

The initiation point of the CCLH part of the study was the assumption of two families having the same disease locus. During the course of the thesis we showed that families do not share the same disease locus. Separating the data for each family decreased the possibility of identifying a significant candidate locus for Family CCLH6 because parents were first degree cousins and family size was not big enough to give enough data to point a single candidate locus. However, in Family CCLH5, non-consanguineous marriage and two healthy sibs were two advantageous points for determining a candidate locus.

7. CONCLUSION

In the framework of this study, microsatellite, SNP genome scans and fine mapping studies were carried on four autosomal recessive diseases: Autosomal Recessive Ataxia, Juvenile Parkinsonism, Larsen Syndrome and Congenital Cerebellar Hypoplasia.

Disease genes were localized and mutations were found in two disease genes: *APTX* for Autosomal Recessive Ataxia and *CANT1* for Autosomal Recessive Larsen Syndrome. Novel variants were identified in *APTX* and *CANT1*.

A novel variant strongly related to Juvenile Parkinsonism was found in *AK3L1*; however, this variant solely could not explain the disease. A second candidate locus was identified at 7q22.1-31.1. Candidate genes at the locus must be analzyed in the patients for mutations.

The number of controls must be increased in order to validate the nature of these three novel variants and to assess them as mutations.

A candidate locus for Congenital Cerebellar Hypoplasia has been identified.

REFERENCES

Abeliovich, A., Y. Schmitz, I. Fariñas, D. Choi-Lundberg, W.H. Ho, P.E. Castillo, N. Shinsky, J.M. Verdugo, M. Armanini, A. Ryan, M. Hynes, H. Phillips, D. Sulzer and A. Rosenthal, 2000, "Mice Lacking Alpha-synuclein Display Functional Deficits in the Nigrostriatal Dopamine System", *Neuron*, Vol.25(1), pp.239-252.

Ahel, I., U. Rass, S.F. El-Khamisy, S. Katyal, P.M. Clements, P.J. McKinnon, K.W. Caldecott and S.C. West, 2006, "The Neurodegenerative Disease Protein Aprataxin Resolves Abortive DNA Ligation Intermediates", *Nature*, Vol.443(7112), pp.713-716.

Aicardi, J., C. Barbosa, E. Andermann, F. Andermann, R. Morcos, Q. Ghanem, Y. Fukuyama, Y Awaya and P. Moe, 1988, "Ataxia-Ocular Motor Apraxia: a Syndrome Mimicking Ataxia-Telangiectasia", *Annals of Neurology*, Vol.24(4), pp.497-502.

Al-Mahdawi, S., R.M. Pinto, D. Varshney, L. Lawrence, M.B. Lowrie, S. Hughes, Z. Webster, J. Blake, J.M. Cooper, R. King and M.A. Pook, 2006, "GAA Repeat Expansion Mutation Mouse Models of Friedreich Ataxia Exhibit Oxidative Stress Leading to Progressive Neuronal and Cardiac Pathology", *Genomics*, Vol.88(5), pp.580-590.

Amouri, R., M.C. Moreira, M. Zouari, G. El Euch, C. Barhoumi, M. Kefi, S. Belal, M. Koenig and F. Hentati, 2004, "Aprataxin Gene Mutations in Tunisian Families", *Neurology*, Vol. 63(5), pp.928-929.

Babcock, M., D. de Silva, R. Oaks, S. Davis-Kaplan, S. Jiralerspong, L. Montermini, M. Pandolfo, J. Kaplan, 1997, "Regulation of Mitochondrial Iron Accumulation by Yfh1p, a Putative Homolog of Frataxin", *Science*, Vol.276, pp.1709-1712.

Bakkenist, C.J. and M.B. Kastan, 2003, “DNA Damage Activates ATM through Intermolecular Autophosphorylation and Dimer Dissociation”, *Nature*, Vol.421(6922), pp.499-506.

Barkovich, A.J., 2000, “Concepts of Myelin and Myelination in Neuroradiology”, *American Journal of Neuroradiology*, Vol.21(6), pp.1099-1109.

Becker, R., R.D. Wegner, J. Kunze, S. Runkel, M. Vogel and M. Entezami, 2000, “Clinical Variability of Larsen Syndrome: Diagnosis in a Father After Sonographic Detection of a Severely Affected Fetus”, *Clinical Genetics*, Vol.57(2), pp.148-150.

Billuart, P., T. Bienvenu, N. Ronce, V. des Portes, M.C. Vinet, R. Zemni, H. Roest Crollius, A. Carrié, F. Fauchereau, M. Cherry, S. Briault, B. Hamel, J.P. Fryns, C. Beldjord, A. Kahn, C. Moraine and J. Chelly, 1998, “Oligophrenin-1 Encodes a rhoGAP Protein Involved in X-linked Mental Retardation”, *Nature*, Vol.392(6679), pp.923-926.

Bird, T.D., 1998, “Hereditary Ataxia Overview”, in R.A. Pagon, T.C. Bird, C.R. Dolan and K. Stephens (eds), GeneReviews [Internet], University of Washington, Seattle.

Bloch, C. and H.M. Peck, 1965, “Bilateral Congenital Dislocation of the Knees”, *Mount Sinai Journal of Medicine*, Vol.32, pp.607-614.

Boder, E, 1985, “Ataxia-Telangiectasia: an Overview”, *Kroc Foundation Series*, Vol.19, pp.1-63.

Boycott, K.M., S. Flavelle, A. Bureau, H.C. Glass, T.M. Fujiwara, E. Wirrell, K. Davey, A.E. Chudley, J.N. Scott, D.R. McLeod and J.S. Parboosingh, 2005, “Homozygous Deletion of the Very Low Density Lipoprotein Receptor Gene Causes Autosomal Recessive Cerebellar Hypoplasia with Cerebral Gyral Simplification”, *American Journal of Human Genetics*, Vol.77(3), pp.477-483.

Clark, I.E., M.W. Dodson, C. Jiang, J.H. Cao, J.R. Huh, J.H. Seol, S.J. Yoo, B.A. Hay and M. Guo, 2006, “Drosophila PINK1 is Required for Mitochondrial Function and Interacts Genetically with Parkin”, *Nature*, Vol. 441, pp.1162-1166.

Curran, T. and G. D'Arcangelo, 1998, “Role of Reelin in the Control of Brain Development”, *Brain Research Reviews*, Vol.26(2-3), pp.285-294.

Dai, J., J. Liu, Y. Deng, T.M. Smith and M. Lu, 2004, “Structure and Protein Design of a Human Platelet Function Inhibitor”, *Cell*, Vol.116, pp.649-659.

Date, H., O. Onodera, H. Tanaka, K. Iwabuchi, K. Uekawa, S. Igarashi, R. Koike, T. Hiroi, T. Yuasa, Y. Awaya, T. Sakai, T. Takahashi, H. Nagatomo, Y. Sekijima, I. Kawachi, Y. Takiyama, M. Nishizawa, N. Fukuhara, K. Saito, S. Sugano and S. Tsuji, 2001, “Early-onset Ataxia with Ocular Motor Apraxia and Hypoalbuminemia is Caused by Mutations in a New HIT Superfamily Gene”, *Nature Genetics*, Vol. 29(2), pp.184-188.

David, G., N. Abbas, G. Stevanin, A. Dürr, G. Yvert, G. Cancel, C. Weber, G. Imbert, F. Saudou, E. Antoniou, H. Drabkin, R. Gemmill, P. Giunti, A. Benomar, N. Wood, M. Ruberg, Y. Agid, J.L. Mandel and A. Brice, 1997, “Cloning of the SCA7 Gene Reveals a Highly Unstable CAG Repeat Expansion, *Nature Genetics*, Vol.17(1), pp.65-70.

Fahn, S., 2003, “Description of Parkinson's Disease as a Clinical Syndrome”, *Annals of the New York Academy of Sciences*,Vol.991, pp.1-14.

Faivre, L., M. Le Merrer, K. Zerres, M. Ben Hariz, D. Scheffer, I.D. Young, P. Maroteaux, A. Munnich and V. Cormier-Daire, 2004, “Clinical and Genetic Heterogeneity in Desbuquois Dysplasia”, *American Journal of Medical Genetics*, Vol.128A(1), pp.29-32.

Ferrer, I., J. Sirvent, J.M. Manresa, E. Galofre, E. Fernandez-Alvarez and M. Pineda, 1987, "Primary Degeneration of the Granular Layer of the Cerebellum (Norman Type): a Golgi Study", *Acta Neuropathologica*, Vol.75, pp.203-208.

Gillessen-Kaesbach, G., P. Meinecke, M.G. Ausems, M. Nöthen, B. Albrecht, F.A. Beemer and K. Zerres, 1995, "Desbuquois Syndrome: Three Further Cases and Review of the Literature", *Clinical Dysmorphology*, Vol.4(2), pp.136-144.

Goldberg, M.S. and P.T. Jr. Lansbury, 2000, "Is There a Cause-and-Effect Relationship Between Alpha-synuclein Fibrillization and Parkinson's Disease?", *Nature Cell Biology*, Vol.2(7): E115-119.

Gropman, A.L. and D.R. Adams, 2007, "Atypical Patterns of Inheritance", *Seminars in Pediatric Neurology*, Vol.14(1), pp.34-45.

Hermanns, P., S. Unger, A. Rossi, A. Perez-Aytes, H. Cortina, L. Bonafé, L. Boccone, V. Setzu, M. Dutoit, L. Sangiorgi, F. Pecora, K. Reicherter, G. Nishimura, J. Spranger, B. Zabel and A. Superti-Furga, 2008, "Congenital Joint Dislocations Caused by Carbohydrate Sulfotransferase 3 Deficiency in Recessive Larsen Syndrome and Humero-Spinal Dysostosis", *American Journal of Human Genetics*, Vol.82(6), pp.1368-1374.

Hoffmann, K. and T. H. Lindner, 2005, "easyLINKAGE-Plus-Automated Linkage Analyses Using Large-Scale SNP Data", *Bioinformatics*, Vol. 21(17), pp.3565-3567.

Hong, S.E., Y.Y. Shugart, D.T. Huang, S.A. Shahwan, P.E. Grant, J.O. Hourihane, N.D. Martin and C.A. Walsh, 2000, "Autosomal Recessive Lissencephaly with Cerebellar Hypoplasia is Associated with Human RELN Mutations", *Nature Genetics*, Vol.26(1), pp.93-96.

Huber, C., B. Oulès, M. Bertoli, M. Chami, M. Fradin, Y. Alanay, L.I. Al-Gazali, M.G. Ausems, P. Bitoun, D. P. Cavalcanti, A. Krebs, M. Le Merrer, G. Mortier, Y. Shafeghati, A. Superti-Furga, S.P. Robertson, C. Le Goff, A.O. Muda, P. Paterlini-

Bréchot, A. Munnich and V. Cormier-Daire, 2009, "Identification of CANT1 Mutations in Desbuquois Dysplasia", *American Journal of Human Genetics*, Vol.85(5), pp.706-710.

Ishikawa, A. and S. Tsuji, 1996, "Clinical Analysis of 17 Patients in 12 Japanese Families with Autosomal-Recessive Type Juvenile Parkinsonism", *Neurology*, Vol.47(1), pp.160-166.

Jackson, S.P., 2001, "Detecting, Signalling and Repairing DNA Double-Strand Breaks", *Biochemical Society Transactions*, Vol.6, pp.655-661.

Jonsson, P.A., K.S. Graffmo, P.M. Andersen, S.L. Marklund and T. Brännström, 2009, "Superoxide Dismutase in Amyotrophic Lateral Sclerosis Patients Homozygous for the D90A Mutation", *Neurobiology of Disease*, Vol.36(3), pp.421-424.

Jonsson P.A., A. Bäckstrand, P.M. Andersen, J. Jacobsson, M. Parton, C. Shaw, R. Swingler, P.J. Shaw, W. Robberecht, A.C. Ludolph, T. Siddique, V.I. Skvortsova and S.L. Marklund, 2002, "CuZn-Superoxide Dismutase in D90A Heterozygotes from Recessive and Dominant ALS Pedigrees", *Neurobiology of Disease*, Vol.10(3), pp.327-333.

Kavaslar, G.N., S. Önengüt, O. Derman, A. Kaya and A. Tolun, 2000, "The Novel Genetic Disorder Microhydranencephaly Maps to Chromosome 16p13.3-12.1", *American Journal of Human Genetics*, Vol.66(5), pp.1705-1709.

Khelfaoui, M., C. Denis, E. van Galen, F. de Bock, A. Schmitt, C. Houbron, E. Morice, B. Giros, G. Ramakers, L. Fagni, J. Chelly, M. Nosten-Bertrand and P. Billuart, 2007, "Loss of X-linked Mental Retardation Gene Oligophrenin1 in Mice Impairs Spatial Memory and Leads to Ventricular Enlargement and Dendritic Spine Immaturity", *The Journal of Neuroscience*, Vol.27(35), pp.9439-9450.

Klein, C. and K. Lohmann-Hedrich, 2007, "Impact of Recent Genetic Findings in Parkinson's Disease", *Current Opinion in Neurology*, Vol.20, pp.453-464.

Krakow, D., S.P. Robertson, L.M. King, T. Morgan, E.T. Sebald, C. Bertolotto, S. Wachsmann-Hogiu, D. Acuna, S.S. Shapiro, T. Takafuta, S. Aftimos, C.A. Kim, H. Firth, C.E. Steiner, V. Cormier-Daire, A. Superti-Furga, L. Bonafe, J.M. Jr. Graham, A. Grix, C.A. Bacino, J. Allanson, M.G. Bialer, R.S. Lachman, D.L. Rimoin and D.H. Cohn, 2004, "Mutations in the Gene Encoding Filamin B Disrupt Vertebral Segmentation, Joint Formation and Skeletogenesis", *Nature Genetics*, Vol.36(4), pp.405-410.

Lander, E. S. and D. Botstein, 1987, "Homozygosity Mapping: A Way to Map Human Recessive Traits with the DNA of Inbred Children", *Science*, Vol.236(4808), pp.1567-1570.

Larsen, L.J., E.R. Schottstaedt and F.C. Bost, 1950, "Multiple Congenital Dislocations Associated with Characteristic Facial Abnormality", *Journal of Pediatrics*, Vol.37(4), pp.574-581.

Lavin, M.F., 2008, "Ataxia-Telangiectasia: from a Rare Disorder to a Paradigm for Cell Signalling and Cancer", *Nature Reviews Molecular Cell Biology*, Vol.9(10), pp.759-769.

Lavin, M.F., G. Birrell, P. Chen, S. Kozlov, S. Scott and N. Gueven, 2005, "ATM Signaling and Genomic Stability in Response to DNA Damage", *Mutation Research*, Vol.569(1-2), pp.123-132.

Le Ber, I., O. Dubourg, J.F. Benoist, C. Jardel, F. Mochel, M. Koenig, A. Brice, A. Lombès and A. Dürr, 2007, "Muscle Coenzyme Q10 Deficiencies in Ataxia with Oculomotor Apraxia 1", *Neurology*, Vol.68(4), pp.295-297.

Lee, Y., J.W. Kim, S.M. Lee, H.J. Kim, K.S. Lee and C. Park, 1998, "Cloning and Expression of Human Adenylate Kinase 2 Isozymes: Differential Expression of Adenylate Kinase 1 and 2 in Human Muscle Tissues, *The Journal of Biochemistry*, Vol.123(1), pp.47-54.

Lohmann, E., M. Periquet, V. Bonifati, N.W. Wood, G. De Michele, A.M. Bonnet, V. Fraix, E. Broussolle, M.W. Horstink, M. Vidailhet, P. Verpillat, T. Gasser, D. Nicholl, H. Teive, S. Raskin, O. Rascol, A. Destée, M. Ruberg, F. Gasparini, G. Meco, Y. Agid, A. Durr, A. Brice, 2003, "How much phenotypic variation can be attributed to parkin genotype?", *Annals of Neurology*, Vol.54(2):176-185.

Lorson, C.L., E. Hahnen, E.J. Androphy and B. Wirth, 1999, "A Single Nucleotide in the SMN Gene Regulates Splicing and is Responsible for Spinal Muscular Atrophy", *Proceedings of the National Academy of Sciences USA*, Vol.96(11), pp. 6307-6311.

Lu, J., G. Lian, R. Lenkinski, A. De Grand, R.R. Vaid, T. Bryce, M. Stasenko, A. Boskey, C. Walsh and V. Sheen, 2007, "Filamin B Mutations Cause Chondrocyte Defects in Skeletal Development", *Human Molecular Genetics*, Vol.16(14), pp.1661-1675.

Lücking, C.B., A. Dürr, V. Bonifati, J. Vaughan, G. De Michele, T. Gasser, B.S. Harhangi, G. Meco, P. Denèfle, N.W. Wood, Y. Agid and A. Brice, 2000, "Association Between Early-Onset Parkinson's Disease and Mutations in the Parkin Gene", *New England Journal of Medicine*, Vol.342(21), pp.1560-1567.

Marquardt, T. and J. Denecke, 2003, "Congenital Disorders of Glycosylation: Review of Their Molecular Bases, Clinical Presentations and Specific Therapies", *European Journal of Pediatrics*, Vol.162(6), pp.359-379.

Matthijs, G., E. Schollen, E. Pardon, M. Veiga-Da-Cunha, J. Jaeken, J.J. Cassiman and E. Van Schaftingen, 1997, "Mutations in PMM2, a Phosphomannomutase Gene on Chromosome 16p13, in Carbohydrate-Deficient Glycoprotein Type I Syndrome (Jaeken Syndrome)", *Nature Genetics*, Vol.16(1), pp.88-92.

McNaught, K.S. and C.W. Olanow, 2006, "Protein Aggregation in the Pathogenesis of Familial and Sporadic Parkinson's Disease", *Neurobiology of Aging*, Vol.27(4), pp.530-545.

Merlo, C.A. and M.P. Boyle, 2003, “Modifier Genes in Cystic Fibrosis Lung Disease”, *Journal of Laboratory and Clinical Medicine*, Vol.141(4), pp.237-241.

Moreira, M.C., C. Barbot, N. Tachi, N. Kozuka, E. Uchida, T. Gibson, P. Mendonça, M. Costa, J. Barros, T. Yanagisawa, M. Watanabe, Y. Ikeda, M. Aoki, T. Nagata, P. Coutinho, J. Sequeiros and M. Koenig, 2001, “The Gene Mutated in Ataxia-Ocular Apraxia 1 Encodes the New HIT/Zn-finger Protein Aprataxin”, *Nature Genetics*, Vol.29(2), pp.189-193.

Moreira, M.C., S. Klur, M. Watanabe, A.H. Németh, I. Le Ber, J.C. Moniz, C. Tranchant, P. Aubourg, M. Tazir, L. Schöls, M. Pandolfo, J.B. Schulz, J. Pouget, P. Calvas, M. Shizuka-Ikeda, M. Shoji, M. Tanaka, L. Izatt, C.E. Shaw, A. M’Zahem, E. Dunne, P. Bomont, T. Benhassine, N. Bouslam, G. Stevanin, A. Brice, J. Guimarães, P. Mendonça, C. Barbot, P. Coutinho, J. Sequeiros, A. Dürr, J.M. Warter and M. Koenig, 2004, “Senataxin, the Ortholog of a Yeast RNA Helicase, is Mutant in Ataxia-Ocular Apraxia 2”, Nature Genetics, Vol.36(3), pp.225-227.

Mosesso, P., M. Piane, F. Palitti, G. Pepe, S. Penna and L. Chessa, 2005, “The Novel Human Gene Aprataxin is Directly Involved in DNA Single-Strand-Break Repair”, *Cellular and Molecular Life Sciences*, Vol.62(4), pp.485-491.

Noma, T., 2005, “Dynamics of Nucleotide Metabolism as a Supporter of Life Phenomena”, *Journal of Investigative Medicine*, Vol.52(3-4), pp.127-136.

Oprea, G.E., S. Kröber, M.L. McWhorter, W. Rossoll, S. Müller, M. Krawczak, G.J. Bassell, C.E. Beattie and B. Wirth, 2008, “Plastin 3 is a Protective Modifier of Autosomal Recessive Spinal Muscular Atrophy”, *Science*, Vol.320(5875), pp. 524-527.

Ozcelik, T., N. Akarsu, E. Uz, S. Caglayan, S. Gulsuner, O.E. Onat, M. Tan and U. Tan, 2008, “Mutations in the Very Low-Density Lipoprotein Receptor VLDLR Cause Cerebellar Hypoplasia and Quadrupedal Locomotion in Humans”, *The Proceedings of the National Academy of Sciences USA*, Vol.105(11), pp.4232-4236.

Paglia, L.L., A. Laugé, J. Weber, J. Champ, E. Cavaciuti, A. Russo, J.L. Viovy and D. Stoppa-Lyonnet, 2010, "ATM Germline Mutations in Women with Familial Breast Cancer and a Relative with Haematological Malignancy", *Breast Cancer Research and Treatment*, Vol.119(2), pp.443-452.

Panayiotou, C., N. Solaroli, M. Johansson and A. Karlsson, 2010, "Evidence of an Intact N-terminal Translocation Sequence of Human Mitochondrial Adenylate Kinase 4", *The International Journal of Biochemistry and Cell Biology*, Vol.42(1), pp.62-69.

Pandolfo, M. and A. Pastore, 2009, "The Pathogenesis of Friedreich Ataxia and the Structure and Function of Frataxin", *Journal of Neurology*, Vol.256(l), pp.9-17.

Pandolfo, M., 2009, "Friedreich Ataxia: the Clinical Picture", *Journal of Neurology*, Vol.256(1), pp.3-8.

Parfitt, D.A., G.J. Michael, E.G.M. Vermeulen, N.V. Prodromou, T.R. Webb, J.M. Gallo, M.E. Cheetham, W.S. Nicoll, G.L Blatch, and 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(9), pp.1556-1565.

Park, J., S.B. Lee, S. Lee, Y. Kim, S. Song, S. Kim, E. Bae, J. Kim, M. Shong, J.M. Kim, and J. Chung, 2006, "Mitochondrial Dysfunction in Drosophila PINK1 Mutants is Complemented by Parkin", *Nature*, Vol.441, pp.1157-1161.

Patel, S. and A.J. Barkovich, 2002, "Analysis and Classification of Cerebellar Malformations", *American Journal of Neuroradiology*, Vol.23(7), pp.1074-1087.

Petrella, R., J.G. Rabinowitz, B. Steinmann and K. Hirschhorn, 1993, "Long-term Follow-up of Two Sibs with Larsen Syndrome Possibly Due to Parental Germ-line Mosaicism", *American Journal of Medical Genetics*, Vol.47(2), pp.187-197.

Pfeiffer, P, 1998, “The Mutagenic Potential of DNA Double-Strand Break Repair”, *Toxicology Letters*, Vol.96-97, pp.119-129.

Pinar, H., N. Tatevosyants and D.B. Singer, 1998, “Central Nervous System Malformations in Perinatal/Neonatal Autopsy Series”, *Pediatric and Developmental Pathology*, Vol.1, pp.42-48.

Platzer, M., G. Rotman, D. Bauer, T. Uziel, K. Savitsky, A. Bar-Shira, S. Gilad, Y. Shiloh and A. Rosenthal, 1997, “Ataxia-Telangiectasia Locus: Sequence Analysis of 184 kb of Human Genomic DNA Containing the Entire ATM gene”, *Genome Research*, Vol.7(6), pp.592-605.

Poole, A.C., R.E. Thomas, L.A. Andrews, H.M. McBride, A.J. Whitworth and L.J. Pallanck, 2008, “The PINK1/Parkin Pathway Regulates Mitochondrial Morphology”, *The Proceedings of the National Academy of Sciences USA*, Vol.105, pp.1638-1643.

Quinzii, C.M., A.G. Kattah, A. Naini, H.O. Akman, V.K. Mootha, S. DiMauro and M. Hirano, 2005, “Coenzyme Q Deficiency and Cerebellar Ataxia Associated with an Aprataxin Mutation”, *Neurology*, Vol.64(3), pp.539-541.

Raman, R., V. Sasisekharan and R. Sasisekharan, 2005, “Structural Insights into Biological Roles of Protein-Glycosaminoglycan Interactions”, *Chemistry and Biology*, Vol.12(3), pp.267-277.

Rass, U., I. Ahel and S.C. West, 2007, “Actions of Aprataxin in Multiple DNA Repair Pathways”, *The Journal of Biological Chemistry*, Vol.282(13), pp.9469-9474.

Risch, N., 1992, “Genetic Linkage: Interpreting Lod Scores”, *Science*, Vol.255(5046), pp.803-804.

Sano, Y., H. Date, S. Igarashi, O. Onodera, M. Oyake, T. Takahashi, S. Hayashi, M. Morimatsu, H. Takahashi, T. Makifuchi, N. Fukuhara and S. Tsuji, 2004, “Aprataxin,

the Causative Protein for EAOH is a Nuclear Protein with a Potential Role as a DNA Repair Protein", *Annals of Neurology*, Vol. 55(2), pp. 241-249.

Savitsky, K., A. Bar-Shira, S. Gilad, G. Rotman, Y. Ziv, L. Vanagaite, D.A. Tagle, S. Smith, T. Uziel, S. Sfez, M. Ashkenazi, I. Pecker, M. Frydman, R. Harnik, S.R. Patanjali, A. Simmons, G.A. Clines, A. Sartiel, R.A. Gatti, L. Chessa, O. Sanal, M.F. Lavin, N.G. Jaspers, A.M. Taylor, C.F. Arlett, T. Miki, S.M. Weissman, M. Lovett, F.S. Collins and Y. Shiloh, 1995, "A Single Ataxia Telangiectasia Gene with a Product Similar to PI-3 kinase, *Science*, Vol. 268(5218), pp. 1749-1753.

Schulz, J.B., S. Boesch, K. Bürk, A. Dürr, P. Giunti, C. Mariotti, F. Pousset, L. Schöls, P. Vankan and M. Pandolfo, 2009, "Diagnosis and Treatment of Friedreich Ataxia: a European Perspective", *Nature Reviews Neurology*, Vol.5(4), pp.222-234.

Schurig, V., A.V. Orman and P. Bowen, 1981, "Nonprogressive Cerebellar Disorder with Mental Retardation and Autosomal Recessive Inheritance in Hutterites", *American Journal of Medical Genetics*, Vol.9(1), pp.43-53.

Shendelman, S., A. Jonason, C. Martinat, T. Leete and A. Abeliovich, 2004, "DJ1 is a Redox-dependent Molecular Chaperone that Inhibits Alpha-synuclein Aggregate Formation", *PloS Biology*, Vol. 2, e362.

Shevell, M.I. and A. Majnemer, 1996, "Clinical Features of Developmental Disability Associated with Cerebellar Hypoplasia", *Pediatric Neurology*, Vol.15(3), pp.224-229.

Silverman, F.N. "Larsen's Syndrome: Congenital Dislocation of the Knees and Other Joints, Distinctive Facies, and, Frequently, Cleft Palate", 1972, *Annales de Radiologie*, Vol.15(3), pp.297-328.

Sonoda, E., T. Matsusaka, C. Morrison, P. Vagnarelli, O. Hoshi, T. Ushiki, K. Nojima, T. Fukagawa, I.C. Waizenegger, J.M. Peters, W.C. Earnshaw and S. Takeda, 2001,

"Scc1/Rad21/Mcd1 is Required for Sister Chromatid Cohesion and Kinetochore Function in Vertebrate Cells", *Developmental Cell*, Vol.1(6), pp.759-770.

Suraweera, A., O.J. Becherel, P. Chen, N. Rundle, R. Woods, J. Nakamura, M. Gatei, C. Criscuolo, A. Filla, L. Chessa, M. Fusser, B. Epe, N. Gueven and M.F. Lavin, 2007, "Senataxin, Defective in Ataxia Oculomotor Apraxia Type 2, is Involved in the Defense Against Oxidative DNA Damage", *The Journal of Cell Biology*,Vol.177(6), pp.969-979.

Takahashi, H., E. Ohama, S. Suzuki, Y. Horikawa, A. Ishikawa, T. Morita, S. Tsuji and F. Ikuta, 1994, "Familial Juvenile Parkinsonism: Clinical and Pathologic Study in a Family", *Neurology*, Vol.44, pp.437-441.

Takahashi, T., M. Tada, S. Igarashi, A. Koyama, H. Date, A. Yokoseki, A. Shiga, Y. Yoshida, S. Tsuji, M. Nishizawa, O. Onodera, 2007, "Aprataxin, Causative Gene Product for EAOH/AOA1, Repairs DNA Single-Strand Breaks with Damaged 3'-phosphate and 3'-phosphoglycolate Ends", *Nucleic Acids Research*, Vol.35(11), pp.3797-3809.

Téoule, R., 1987, "Radiation-Induced DNA Damage and Its Repair, International Journal of Radiation Biology and Related Studies in Physics, *Chemistry and Medicine*, Vol.51(4), pp.573-589.

Terwilliger, J. D. and J. Ott, 1994, Handbook of Human Genetic Linkage, Johns Hopkins University, Baltimore.

Thiele, H., M. Sakano, H. Kitagawa, K. Sugahara, A. Rajab, W. Höhne, H. Ritter, G. Leschik, P. Nürnberg and S. Mundlos, 2004, "Loss of Chondroitin 6-O-sulfotransferase-1 Function Results in Severe Human Chondrodysplasia with Progressive Spinal Involvement", *The Proceedings of the National Academy of Sciences USA*, Vol.101(27), pp.10155-10160.

Thompson, D., S. Duedal, J. Kirner, L. McGuffog, J. Last, A. Reiman, P. Byrd, M. Taylor and D.F. Easton, 2005, "Cancer Risks and Mortality in Heterozygous ATM Mutation Carriers", *Journal of the National Cancer Ins*titute, Vol.97(11),pp.813-822.

Tissir, F. and A.M. Goffinet, 2003, "Reelin and Brain Development", *Nature Reviews Neuroscience*, Vol.4(6), pp.496-505.

Tranchant, C., M. Fleury, M.C. Moreira, M. Koenig and J.M. Warter, 2003, "Phenotypic variability of aprataxin gene mutations", *Neurology*, Vol.60(5), pp.868-870.

Trommsdorff, M., M. Gotthardt, T. Hiesberger, J. Shelton, W. Stockinger, J. Nimpf, R.E. Hammer, J.A. Richardson and J. Herz, 1999, "Reeler/Disabled-like Disruption of Neuronal Migration in Knockout Mice Lacking the VLDL Receptor and ApoE Receptor 2", *Cell*, Vol.97(6), pp.689-701.

Van der Flier, A. and A. Sonnenberg, 2001, "Structural and Functional Aspects of Filamins", *Biochimica et Biophysica Acta*, Vol.1538(2-3), pp.99-117.

Wong, A., J. Yang, P. Cavadini, C. Gellera, B. Lonnerdal, F. Taroni and G. Cortopassi, 1999, "The Friedreich's Ataxia Mutation Confers Cellular Sensitivity to Oxidant Stress which is rescued by Chelators of Iron and Calcium and Inhibitors of Apoptosis", *Human Molecular Genetics*, Vol.8, pp.425-430.

Yang, Y., S. Gehrke, Y. Imai, Z. Huang, Y. Ouyang, J.W. Wang, L. Yang, M.F. Beal, H. Vogel and B. Lu, 2006, "Mitochondrial Pathology and Muscle and Dopaminergic Neuron Degeneration Caused by Inactivation of Drosophila PINK1 is Rescued by Parkin", *The Proceedings of the National Academy of Sciences USA*, Vol. 103, pp.10793-10798.

Yoneda, T., M. Sato, M. Maeda and H. Takagi, 1998, "Identification of a Novel Adenylate Kinase System in the Brain: Cloning of the Fourth Adenylate Kinase", *Molecular Brain Research*, Vol.62(2), pp.187-195.

Zhang, L., M. Shimoji, B. Thomas, D.J. Moore, S.W. Yu, N.I. Marupudi, R. Torp, I.A. Torgner, O.P. Ottersen, T.M. Dawson and V.L. Dawson, 2005, “Mitochondrial Localization of the Parkinson’s Disease Related Protein DJ1: Implications for Pathogenesis”, *Human Molecular Genetics*, Vol. 14, pp.2063-2073.