GENETIC FACTORS CONTRIBUTING TO ALS IN TURKEY: ATXN2 POLYQ REPEAT EXPANSIONS AND SNP/CNV ASSOCIATIONS

by

Özgür Ömür

B.S., Molecular Biology and Genetics, İzmir Institute of Technology, 2010

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

Graduate Program in Molecular Biology and Genetics

Boğaziçi University

2013

To the memory of Prof. Hilmi Özçelik

ACKNOWLEDGEMENTS

I am heartily grateful to my supervisor, Prof. A. Nazlı Başak, whose continuous guidance and valuable comments encouraged me, from the initial to the final level, in developing this thesis study. Her guidance helped me during research and writing this thesis.

Besides my supervisor, I would like to extend my best regards and gratitude to Prof. Georg Auburger, Suzana Gispert and all the Auburger Lab members for their warm hospitality and guidance during my stay at Experimental Neurology Center of Goethe Institute, Frankfurt, Germany.

I would also like to express my thanks to Prof. Georg Auburger and Cemalettin Bekpen for devoting their time to evaluate this thesis and I would like to acknowledge the Suna-İnan Kıraç Foundation for the financial support it provided throughout my study.

I would like to express my thanks to Aslıhan Özoğuz, A. Kadir Özkan, Ceren İskender, Nesli Ece Şen, Ceren Saygı, Ece Kartal, Selda Dağdeviren, Aslı Gündoğdu, İlknur Yıldız, Özge Demet Özçete, Irmak Şahbaz and previous lab members; Özgün Uyan, Gönenç Çobanoğlu, Sena Ağım, Pınar Deniz, for their sincere friendship and all the happy times we spent together. I would especially like to thank Suna Lahut for her patience and continuous support during my MSc. study and research.

I am also very grateful to my house-mates, Ahmet Cihangir and Mustafa Görece, for their support and friendship.

Last but not least, I owe my deepest gratitude and love to my devoted parents and little sisters.

ABSTRACT

GENETIC FACTORS CONTRIBUTING TO ALS IN TURKEY: ATXN2 POLYQ REPEAT EXPANSIONS AND SNP/CNV ASSOCIATIONS

ALS is the most frequent motor neuron disorder and so far no cure has been developed for this devastating disease. ALS is characterized by the selective loss of both the upper motor neurons of the motor cortex and the lower motor neurons extending through the brainstem and spinal cord to skeletal muscle. The majority of ALS patients (86.5-99 per cent) are sporadic and the remaining cases have a reported family history. With more than 25 genes and several loci identified so far, approximately 60 percent of familial and only a small proportion of sporadic ALS cases have been explained. Thus, identification of new genes contributing to disease is crucial. A comprehensive study combining yeast screen, fly/cell biology and human genetics led to the identification of the association of intermediate length (27-33) PolyQ expansions in the ATXN2 gene and ALS risk. In this study, the distribution of ATXN2 PolyQ sizes in a Turkish cohort including 236 patients and 420 healthy controls were investigated. Furthermore, to evaluate other possible genetic associations of this locus with ALS, SNP and haplotypes were analyzed. In addition to SNP and haplotype analyses, validation of candidate CNVs in the MAP4K3, HLA-B and EPHA3 genes, previously found to be disease-associated, was performed. In accordance with other studies, this thesis confirmed that >30 PolyQ repeats in the ATXN2 gene are associated with ALS risk in 1.7% of the Turkish ALS cohort under study. Additionally, a significant association of a 5-SNP haplotype block across the ATXN2 and SH2B3 genes was found in three per cent of the ALS cohort, which was not present in the controls ($p = 0.0011$). CNV validation studies revealed that EPHA3 is a possible protective factor against ALS in the Turkish ALS cohort under study. EPHA3 is one of the members of receptor tyrosine kinase (RTK) family and ATXN2 and SH2B3 encode proteins that both interact with growth receptor tyrosine kinases. Our novel observation suggests that proteins of RTK pathway and their interaction partners may be promising targets for therapeutic interventions.

ÖZET

TÜRKİYE'DE ALS'YE ETKİ EDEN GENETİK FAKTÖRLER: ATXN2 POLİGLUTAMİN TEKRAR ARTIŞLARI VE SNP/CNV İLİŞKİLENDİRME ÇALIŞMALARI

ALS en yaygın motor nöron hastalığıdır, bu ölümcül hastalığa yönelik bir tedavi yöntemi halen geliştirilmemiştir. ALS, motor korteksin üst motor nöronları ile beyin sapından omuriliğe, oradan da iskelet kaslarına doğru uzanan alt motor nöronların özgün kaybıyla karakterizedir. ALS olgularının büyük çoğunluğu (%85-90) sporadiktir, geri kalan kısmında ailesel geçiş öyküsü mevcuttur. Bugüne kadar tanımlanmış 20'yi aşkın gen ve birçok lokus, ailesel olguların yüzde 60'ını, sporadik olguların ise çok az bir kısmını açıklamaktadır. Bu nedenle, hastalığa etki eden genlerin bulunması büyük önem taşımaktadır. Maya, sinek/hücre biyolojisi ve insan genetiği gibi yöntemlerden yararlanılarak yapılan bir araştırma, ATXN2 genindeki poliglutamin (PolyQ) bölgesinin normalden fazla uzamasının (27-33) ALS riskiyle ilişkili olduğunu göstermiştir. Bu çalışmada, 236 hasta ve 420 sağlıklı birey içeren Türk hasta grubunda ATXN2 genindeki PolyQ tekrar sayısının araştırılması hedeflenmiştir. Ayrıca, bu bölgede ALS ile ilişkili olabilecek diğer genetik varyasyonlar da araştırılmıştır. SNP ve haplotip analizlerinin yanısıra, daha önce ALS ile ilişkili olabileceği gösterilen MAP4K3, HLA-B ve EPHA3 genlerindeki CNV'lerin doğrulaması da yapılmıştır. Literatürle uyumlu olarak, bu tezde de, ATXN2 genindeki >30 PolyQ tekrar sayısının ALS riski taşıdığı hastaların %1.7'sinde gösterilmiştir. Bunun yanında, ATXN2 ve SH2B3 genlerinde bulunan 5-SNP'lik bir haplotip blokunun da ALS ile istatistiksel olarak anlamlı bir şekilde (hastaların %3'ünde mevcut, kontrollerde yok) ilişkili olduğu tanımlanmıştır (p=0.0011). Tezin ikinci aşamasını oluşturan CNV validasyon çalışmaları da, EPHA3 genindeki delesyonun incelenen Türk hasta grubunda, ALS'ye karşı koruyucu bir faktör olduğunu göstermiştir. EPHA3, reseptör tirozin kinaz (RTK) ailesine dahildir, ATXN2 ve SH2B3 genleri de bu aileden olan *growth receptor tyrosine kinase* ile etkileşen proteinleri kodlamaktadır. Bu bulgular ışığında çalışmamız, RTK yolağında rol alan ve onlarla etkileşen proteinlerin ALS'ye terapi yaklaşımlarında önemli yeni hedeflere işaret etmektedir.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

#	Number
%	per cent
$(NH_4)_2SO_4$	Ammonium Sulfate
*	Asterisk
<	Smaller than
>	Larger than
±	Plus Minus
Ca^{+2}	Calcium Ions
Cu	Cupper
dH_2O	Distilled Water
H_2O_2	Hydrogen Peroxide
HCl	Hydrogen Chloride
KCl	Potassium Chloride
$MgCl_2$	Magnesium Chloride
O_2^-	Superoxide
^oC	Centigrade degree
Q	Glutamine
Zn	Zinc
Δ	Delta
ε	Epsilon
σ	Sigma

LIST OF ACRONYMS/ABBREVIATIONS

µg	Microgram
µl	Microliter
A	Adenine
aa	Amino Acid
AD	Alzheirmer's Disease
ALS	Amyotrophic Lateral Sclerosis
ALS2	Alsin
ANG	Angiogenin
ApoE-ε4	Apolipo Protein E
AR	Autosomal Recessive
ATXN	Ataxin
bp	Basepair
C	Cytosine
C9ORF72	Chromosome 9 Open Reading Frame 72
CNV	Copy Number Variation
C_T	Threshold Cycle
DNA	Deoxiribonucleic Acid
dNTP	Deoxiribonucleotide Triphosphate
DPP6	Dipeptidyl-peptidase 6
EDTA	Ethylenediaminetetraacetic Acid
Eph	Ephrin
EPHA	Ephrin Receptor A
ER	Endoplasmic Reticulum
EtBr	Ethidium Bromide
EU	Europe
fALS	Familial Amyotrophic Lateral Sclerosis
FAM	Fluorescein Amidite
FIG4	Polyphosphoionisite Phosphatase 4
FTD	Frontotemporal Dementia
FUS	Fused in Sarcoma

GWAS	Genome Wide Association Study
HLA-B	Major Histocompatibility Complex, Class 1, B
HWE	Hardy-Weinberg Disequilibrium
ID	Identity
IFNK	Interferone Kappa
InDels	Insertion Deletion
kb	Kilobase
LD	Disequilibrium
LMN	Lower Motor Neuron
Lsm	Like SM
LsmAD	Lsm-associated
M	Molar
MAF	Minor Allele Frequency
MAP4K3	Mitogen-activated Protein Kinase Kinase Kinase Kinase 3
ml	Mililiter
mM	Milimolar
MOBKL2B	MOB Kinase Activator 2B
mRNA	Messenger RNA
ND	Neurodegenerative Disorder
NDAL	Neurodegeneration Research Laboratory
ng	Nanogram
NIPA1	Non-imprinted in Prader willi/Angelman Syndrome Region Protein 1
NLS	Nuclear Localization Signal
OD	Optical Density
OPTN	Optineurin
OR	Odds Ratio
p	Probability Value
PCR	Polymerase Chain Reaction
PD	Psrkinson’s Disease
PFN1	Profilin 1
pmole	Picomole
PSP	Progressive Supranuclear Palsy

PXX	Proline Rich Domain
R^2	Co-segregation Patterns
RNA	Ribonucleic Acid
RNase PH1	Ribonuclease P RNA Component H1
RRM	RNA Recognition Motif
RTK	Receptor Tirosine Domain
sALS	Sporadic Amyotrophic Lateral Sclerosis
SCA	Spinocerebellar Ataxia
SETX	Senataxin
SH	Src Homolgy Domain
SH2B3	SH2B Adaptor Protein
SIGMAR1	σ Non-opioid Receptor 1
SMN1	Survival Motor Neuron Protein1
SNP	Single Nucleotide Polymorphism
SOD1	Superoxide Dismutase1
SPG11	Spastic Paraplegia 11
T	Timine
TARDBP	Tar DNA-Binding Protein
TBE	Tris Borate EDTA
TDP-43	Tar DNA Binding Protein 43
TR	Turkey
txt	Text
u	Unit
UBQLN2	Ubiquilin 2
UMN	Upper Motor Neuron
USA	United States of America
UV	Ultra Violet
V	Volt
VAPB	Vesicle-associated Membrane Protein-associated Protein B
VCP	Valosin-containing Protein
VIC	VIC Fluorescent dye

1. INTRODUCTION

Neurodegenerative disorders (NDs) are worldwide the fourth common cause of death. Especially in the developed countries, NDs have significantly increasing mortality and morbidity rates (Bettens *et al.*, 2013, Hebert *et al.*, 2003). The term "neurodegeneration" is a combination of two words: "neuro-"corresponds to nerve cells, and "denegeneration" refers to a process of losing structure and/or function. According to the National Institute of Neurological Disorders and Stroke, there are hundreds of NDs worldwide, thus, the term has a very broad range of use. Therefore, the classification of NDs is made according to the type and the location of the neurons degenerated. For example, in Alzheimer's Disease (AD), neurons in the cerebral cortex; in Parkinson's Disease (PD), dopaminergic neurons of the substantia nigra; in spinocerebellar ataxias (SCA), the cerebellar neurons and in amyotrophic lateral sclerosis (ALS), the upper motor neurons (UMNs) in the motor cortex and the lower motor neurons (LMNs) in the brainstem and spinal cord specifically degenerate (Tsuji, 2010).

NDs have several common features: (i) they are late-onset, (ii) progressive, (iii) they have both familial and sporadic forms, (iv) they have accumulated proteinaceous aggregates (Gandhi and Wood, 2010, Tsuji, 2010). In the sporadic form of the NDs, which is much more abundant, there is a staggeringly complex interaction among many genetic, environmental and chance factors. However, it is the familial forms, which help to unravel the etiologies and the molecular pathology behind the degeneration processes. The common molecular pathogeneses of NDs are accumulation of misfolded proteins and aggregation of these proteins in neuronal and non-neuronal cells. For instance, amyloid-beta-positive inclusions exist in AD, alfa synuclein-positive inclusions are detectable in PD and Cu/Zn Superoxide Dismutase 1- (SOD1), trans active response DNA Binding protein 43- (TARDBP or TDP-43) positive inclusions are present in ALS (Al-Chalabi *et al.*, 2012, Polymenidou and Cleveland, 2011).

ALS, also known as Lou Gehrig's disease, is the most frequent adult-onset motor neuron disorder with no cure yet. Until it was first described and reported as a non-hereditary disorder by J. M. Charcot (Charcot, 1869), ALS was known as a progressive

muscular atrophy with family history (Aran, 1848). The disease is characterized by the selective loss and dysfunction of both UMNs of motor cortex and LMNs extending through brainstem and spinal cord to skeletal muscle (Figure 1.1). The clinical presentations of ALS vary majorly depending on the regions of the first neurons degenerated. While slow speech, clonic jaw reflexes and spasticity are common in UMN degeneration, weakness, atrophy and fasciculations are the representatives of LMN degeneration. In addition, a considerable proportion of cases also develop Frontotemporal Dementia (FTD) due to the varying degrees of degeneration of the neurons in prefrontal and temporal cortex. All of these degeneration events progressively result in paralysis and lead to death due to respiratory failure within three to five years (Andersen and Al-Chalabi, 2011, Kiernan *et al.*, 2011, Robberecht and Philips, 2013).

ALS has an incidence of 2-3 cases per 100,000 people and a prevalence of 4-6 per 100,000 cases annually with a fairly uniform distribution throughout the world. The disease is slightly more frequent in males (male to female ratio; 1.2-1.5:1). Higher incidences seen in the Island of Guam, the Kii peninsula in Japan and in the Gulf War veterans, which indicate contribution of environmental factors in addition to genetics (Abhinav *et al.*, 2007, Pratt *et al.*, 2012). Although ALS is described as adult-onset with a peak at 55-60 years, a juvenile form with an onset of younger than 25 years, is also present (Kiernan *et al.*, 2011, Orlacchio *et al.*, 2010).

In retrospective studies, the majority of ALS cases (86.5-99%) are reported without a family history and referred as sporadic (sALS) and the remaining cases are inherited and referred to as familial (fALS). Autosomal dominant, with complete or incomplete penetrance, is the most common inheritance pattern, whereas autosomal recessive and X-linked dominant modes of inheritance are also reported in fALS pedigrees (Andersen and Al-Chalabi, 2011, Robberecht and Philips, 2013).

Sporadic ALS and familial ALS share several clinical, pathological and genetic similarities. Heterogeneity in symptoms, variability in disease duration and progression are some of the common clinical features of sALS and fALS which make these two forms almost indistinguishable on clinical grounds. Heterogeneity in the ages of onset is another common feature among the two forms of the disease, however, the mean age of onset is

around one decade earlier in fALS (Cozzolino *et al.*, 2012, Robberecht and Philips, 2013). Existence of SOD1- or TDP-43- positive inclusions in the affected neurons of both sALS and fALS cases is one major neuropathological similarity. Presence of fALS-causative mutations in a minority of apparently sALS cases is an indicator of genetic similarity in addition to the common clinical and pathological features (Al-Chalabi *et al.*, 2012, Bosco *et al.*, 2010, Forsberg *et al.*, 2010).

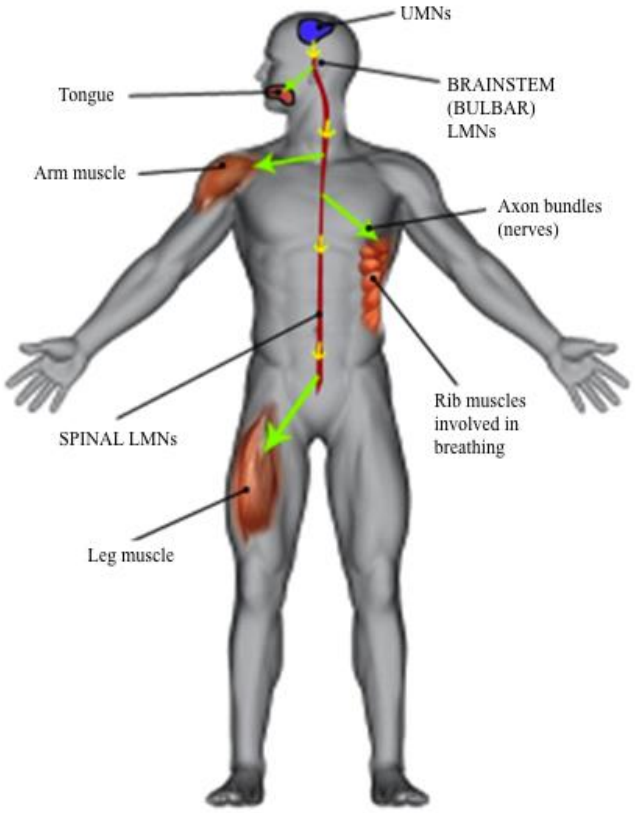


Figure 1.1. The affected neurons of the central nervous system in ALS.

The discovery of new causative genes and genetic risk factors contributing to ALS will definitely help us in: (i) improving the diagnosis strategy and classification of cases in order to better understand the clinical heterogeneity of the disease, (ii) developing new models to unravel the molecular mechanisms behind the pathology, (iii) assessing risk factors for the cases to prevent disease or its progression and (iv) designing effective medicines and innovative therapies. Thus, the genetic basis of ALS should be investigated more to dissect this devastating disease.

1.1. Genetics of ALS

The identification of SOD1 missense mutations in several fALS pedigrees in 1993 (Rosen, 1993) dramatically promoted the genetic research in ALS field and to date, more than 16 genes and several loci have been identified (Robberecht and Philips, 2013). The major fALS genes are listed in Table 1.1. These genes are responsible for approximately 50% of all fALS cases and the remaining part is still under investigation.

Table 1.1. Genes associated with familial ALS.

ALS Symbol/ Gene Name	**Gene Locus**	**Type of Inheritance/Onset**	**Estimated percentage of fALS**	**Reference**
ALS1/SOD1	21q22.1	AD, AR, De Novo /Adult	20	(Rosen, 1993)
ALS2/ALS2	2q33.2	AR/Juvenile	< 1	(Hadano *et al.*, 2001)
ALS3/Unknown	18q21	AD/Adult	Unknown	(Hand *et al.*, 2002)
ALS4/SETX	9q34	AD/Juvenile	Unknown	(Chen *et al.*, 2004)
ALS5/SPG11	15q21.1	AR/Juvenile	Unknown	(Orlacchio *et al.*, 2010)
ALS6/FUS	16q11.2	AD, AR, De Novo /Adult	1-5	(Kwiatkowski *et al.*, 2009)
ALS7/Unknown	20p13	AD/Adult	Unknown	(Sapp *et al.*, 2003)
ALS8/VAPB	20q13.3	AD/Adult	< 1	(Nishimura *et al.*, 2004)
ALS9/ANG	14q11.2	AD/Adult	< 1	(Greenway *et al.*, 2006)
ALS10/TDP-43	1p36.2	AD,AR/Adult	1-5	(Gitcho *et al.*, 2008, Sreedharan *et al.*, 2008)
ALS10/TDP-43	1p36.2	AD,AR/Adult	1-5	(Gitcho *et al.*, 2008, Sreedharan *et al.*, 2008)
ALS11/FIG4	6q21	AD/Adult	Unknown	(Chow *et al.*, 2009)

Table 1.1. Genes associated with familial ALS (continued).

ALS12/OPTN	10p15-p14	AD,AR/Adult	< 1	(Maruyama *et al.*, 2010)
ALS13/ATXN2	12q24	AD/Adult	< 1	(Elden *et al.*, 2010)
ALS14/VCP	9p13-p12	AD/Adult	< 1	(Johnson *et al.*, 2010)
ALS15/UBQLN2	Xp11	X-dominant/Adult	1-2	(Deng *et al.*, 2011)
ALS16/SIGMAR1	9p13.3	AD/Adult, AR/Juvenile	Unknown	(Al-Saif *et al.*, 2011, Luty *et al.*, 2010)
ALS18/PFN1	17p13.2	AD/Adult	Unknown	(Wu *et al.*, 2012)
ALS-FTD2/C9ORF72	9p21.2	AD/Adult	30-50	(DeJesus-Hernandez *et al.*, 2011, Renton *et al.*, 2011)

AD: Autosomal Dominant

AR: Autosomal Recessive

1.1.1. The Most Frequent ALS Genes

SOD1 (ALS1) is a small gene, encoding 153 evolutionarily well-conserved amino acids in its five exons. It is expressed in all cells, mainly located in cytosol, and it is known as a cellular antioxidant enzyme converting the superoxide radical (O_2^-), a product of oxidative phosphorylation events in mitochondria, to less harmful species (H_2O_2) (Redler and Dokholyan, 2012). Although reports vary in different populations, mutations in SOD1 account for 20% of fALS and 2% of sALS cases. Today, more than 169 different mutations (http://alsod.iop.kcl.ac.uk/), distributed in all of the five exons have been detected. The vast majority of these mutations are missense, reported majorly with a dominant inheritance pattern, except D90A, which is also recessively inherited (Andersen and Al-Chalabi, 2011). The role of SOD1 mutations has been widely investigated *in vivo* and *in vitro* over a decade. Novel evidence shows that, mutations in SOD1 may change the folding and the stability of the protein or lead to inefficient binding property through altering its active metal-binding sites (Shaw and Valentine, 2007). However, functional studies revealed that SOD1 still retains its wild-type or near wild-type enzyme activity when mutated, indicating that the SOD1-dependent ALS phenotype stems from a toxic

gain-of-function event rather than a loss-of-function (Bruijn *et al.*, 1998). Moreover, co-expression of wild-type SOD1 in mutant models cannot compensate or decrease the toxicity, further indicating that mutations in SOD1 may cause a novel toxic property; this, however, still needs to be clarified (Fischer and Glass, 2007, Redler and Dokholyan, 2012).

TDP-43 (ALS10) encodes for a 414 amino acid long protein in its six exons. It is ubiquitously expressed in all tissues including muscle and brain. It is a DNA/RNA binding protein and has two evolutionarily conserved RNA recognition motifs RRM1 and RRM2 and a nuclear localization signal (NLS) in the N-terminal region, which are crucial for its activity (Cohen *et al.*, 2011). It is involved in transcriptional regulation, RNA splicing, micro RNA biogenesis, cell death and maintenance of nuclear shape (Buratti and Baralle, 2009, Lagier-Tourenne and Cleveland, 2009). TDP-43 mutations were previously identified in cases with Frontotemporal Dementia (FTD), which sometimes shares similar clinical characteristics with ALS. Studies using affected neurons of FTD and ALS cases showed that TPD-43, which is generally located in the nucleus, is detectable in cytoplasmic inclusions. These inclusions are present in both sALS and fALS cases without SOD1 mutations (Al-Chalabi *et al.*, 2012). This holds the possibility that in a large proportion of ALS cases, mutations and other unknown factors alter TDP-43's cellular localization and may result in a new toxic gain-of-function of this protein, like in the case of SOD1. Genetic studies identified 50 different mutations, which are mostly missense and generally reside in exon 6. Mutations in this gene are almost always dominantly inherited and responsible for 1-5% of fALS and 0-2% of SALS cases (Robberecht and Philips, 2013).

The discovery of TDP-43 triggered the screening of homologous genes on the linkage regions and led to the identification of Fused Sarcoma (FUS/ALS6) mutations in dominant ALS pedigrees (Kwiatkowski *et al.*, 2009, Vance *et al.*, 2009). Subsequent studies also reported autosomal recessive inheritance pattern in several FUS pedigrees. Up to now, 42 mutations were documented on this gene, with about 5% of fALS and rare sALS patients having different types of FUS mutations, such as missense, deletions and insertions (Lagier-Tourenne and Cleveland, 2009). Although some of these mutations occur in the exons 3, 5 and 6, they generally occur in the exons 14-15, which encode for

the NLS (Dormann *et al.*, 2010). FUS is a ubiquitously expressed RNA binding protein and predominantly located in the nucleus. This also raised the question, whether the location of the protein changes when mutated or not, like in TDP-43. Although neuropathological studies detected FUS inclusions in rare cases, it is still under investigation whether ALS-causing mutations create a loss- or a gain-of-function in FUS patients (Al-Chalabi *et al.*, 2012).

An important breakthrough in ALS genetics was achieved with the recent identification of C9ORF72 (ALS-FTD2). The association of 9p21.2 with ALS was first identified in 2006 via linkage analyses (Morita *et al.*, 2006, Vance *et al.*, 2006). Following studies, enhanced by recent genomic technologies such as genome-wide association studies (GWAS), made the circle smaller and three genes, MOBKL2B, IFNK and C9ORF72, were considered as candidates (Laaksovirta *et al.*, 2010, Shatunov *et al.*, 2010). Finally, via resequencing analyses, a hexanucleotide repeat expansion in the intron one of C9ORF72 was detected in cases with ALS. In healthy individuals, the number of GGGGCC repeats are less than 23, however in patients the expansion of this repeat can reach up to 1600 (DeJesus-Hernandez *et al.*, 2011, Renton *et al.*, 2011). This discovery promoted the screening of the region in different populations. With large-scale studies, it was shown that surprisingly the GGGGCC expansion in C9ORF72 is the most frequent cause of both fALS (30-50%) and sALS (approximately 6%) (Majounie *et al.*, 2012, van der Zee *et al.*, 2013).

Ubiquilin 2 (UBQLN2 or ALS 15) is another recent ALS gene. It has one exon encoding for 624 amino acids. The protein product is a member of the ubiquitin-like protein family. It modulates the delivery of ubiquinated proteins to the proteasome complexes. Mutations in this gene were, for the first time, identified in a five-generation ALS family with an X-linked dominant inheritance pattern. Further screenings revealed that the mutations in the unique PXX domain of UBQLN2 gene exist in four additional Caucasian American families (Deng *et al.*, 2011). Following reports showed other mutations, some of which reside also outside of the PXX domain. Although the frequency of UBQLN2-ALS cases are variable in different populations, it is now estimated that around 1-2% of fALS cases have UBQLN2 mutations (Millecamps et al., 2012, Synofzik et al., 2012, van Doormaal et al., 2012, Gellera et al., 2013).

1.2. Proposed ALS Mechanisms

The clinical heterogeneity and the involvement of various genetic factors, which have roles in many different cellular processes, are indicators of the challenge in ALS research. However, with the help of intensive genetic and functional studies, a series of mechanisms underlying ALS are proposed. These mechanisms are mainly categorized as oxidative stress, mitochondrial dysfunction, axonal transport deficits, protein aggregation and abnormal RNA processing. In addition to these intracellular mechanisms, several extracellular events arising from the surrounding cells, such as glutamate excitotoxicity and neuro-inflammation, were also shown to have a role in ALS development (Ferraiuolo *et al.*, 2011) (Figure 1.2).

A more recent mechanism proposed in ALS are the abnormalities in RNA processing, which may lead to both the familial and sporadic forms of the disease. This mechanism was strongly stated due to the identification of two similar RNA binding proteins, TDP-43 and FUS (Ferraiuolo *et al.*, 2011, Robberecht and Philips, 2013), which have important RNA binding domains. These proteins target thousands of molecules, including the RNA product of the ALS gene, VCP, and have roles in diverse mechanisms from central nervous system development to synaptic functioning. Thus, mutations in TDP-43 and FUS may cause ALS through many ways, such as disturbing the processing and the axonal transport of the target molecules. Additively, the involvement of C9ORF72 enhanced the studies in RNA processing mechanism. The exact function of C9ORF72 is not determined yet, however, it is hypothesized that the expansion in the gene may abort the production of its protein and disturbs the balance of other proteins involved in transcription and splicing. Although it is yet impossible to fully understand how the RNA processing mechanism influences ALS pathology, increasing evidences indicate an important contribution of this mechanism to ALS development (Lagier-Tourenne *et al.*, 2012, Robberecht and Philips, 2013).

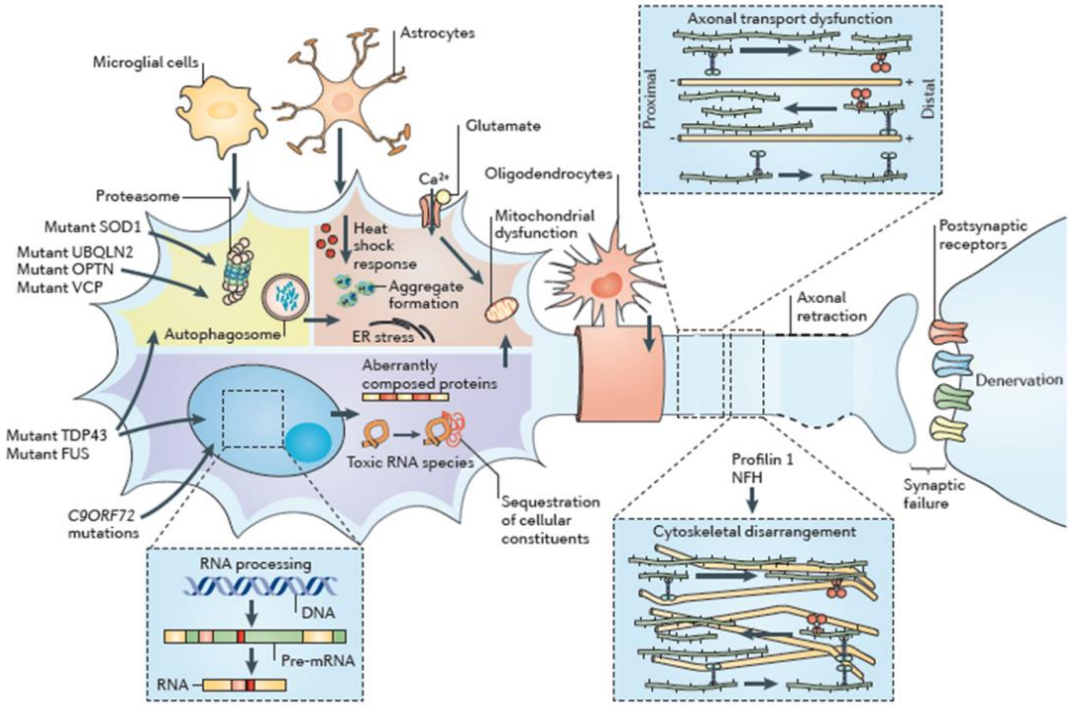


Figure 1.2. Molecular mechanisms in ALS (Robberecht and Philips, 2013).

The proposed ALS mechanisms indicate separate compartments in the cells, thus they may alone result in ALS, however, they can also together affect motor neuron death via multifactorial interactions. Thus, in addition to considering each mechanism individually, the interplay among these events should be considered. For example, the oxidative stress mechanism may result in apoptosis through the activation of microglial cells with secretion of reactive oxygen species or affecting mitochondrial functions that may directly induce apoptosis. In addition, dysfunctions in mitochondria may affect Ca^{+} levels in the surrounding organelles such as endoplasmic reticulum (ER). This might result in ER stress through losing control of the ER membrane pores, which may consequently cause defects in the protein quality control systems and lead to the production of misfolded proteins (Ferraiuolo *et al.*, 2011, Redler and Dokholyan, 2012).

The excess numbers of mechanisms proposed for ALS along with lack of information on how these mechanisms lead to selective motor neuron degeneration, and the association of several genes in diverse, partly overlapping pathways emphasizes the complex nature of the disease.

1.3. A Special Focus on the ATXN2 Locus

ATXN2, located on 12q.24.1, encompasses approximately 130 kb DNA with its 25 exons. It encodes a 1312 amino acid protein which is expressed in many tissues, such as brain, muscles, heart and kidneys. Although its exact function is unknown, several studies proposed possible roles for ataxin 2, such as regulating actin polymerization and endocytosis (Lastres-Becker *et al.*, 2008). On the other hand, the presence of two domains in the protein, Lsm (like SM) and LsmAD (Lsm-associated), common in several RNA processing molecules, suggests a likely role in RNA mechanisms as RNA modification, mRNA degradation and decapping (Chowdhury *et al.*, 2012).

In addition to Lsm and LsmAD, the N-terminal region of ataxin 2 includes a poly glutamine (PolyQ) repeat encoded by the triplet-repeat reads, CAG and CAA. Although it varies among individuals, the nucleotide composition of this triplet-repeat read is generally $(CAG)_8$-CAA-$(CAG)_4$-CAA-$(CAG)_8$ in most healthy individuals. However, the expansion of the PolyQ repeats to more than 33 results in manifestation of spinocerebellar ataxia type 2 (SCA2) or Parkinsonism depending on the predominantly affected brain site. While ataxin 2 aggregation in cerebellum or brainstem leads to SCA2, Parkinsonism manifests if the aggregation predominantly occurs in the basal ganglia (Figure 1.3) (Lagier-Tourenne and Cleveland, 2010, Lastres-Becker *et al.*, 2008).

In 2010, a yeast model, conducted for screening the modifiers of TDP-43 toxicity, indicated that the yeast orthologue of ataxin 2 is one of the modulators of TDP-43-dependent toxicity. The following drosophila study confirmed that the ataxin 2-modified TDP-43 toxicity was dose-sensitive. Additionally, flies that up-regulated both ataxin 2 and TDP-43 levels showed shorter lifespan than TDP-43 up-regulation alone. The same study also revealed a transient ataxin 2 and TDP-43 interaction through RNA binding, suggesting that the two proteins have roles in similar RNA processing mechanisms. Moreover, immunostaining experiments in spinal cord neurons of ALS cases showed alterations in the localization of ataxin 2. The physical and functional interactions between these two proteins and the perturbed localization of ataxin 2 in ALS pathology, together have raised the possibility that there could also be several genetic alterations in ATXN2 in cases with the disease. Consequently, genetic screening in 915 North American patients

revealed that intermediate PolyQ expansion (27-33Q) in ATXN2 confers risk for ALS development. Interestingly, there also were cases, whose affected neurons had TDP-43 accumulations modified by ataxin 2 with normal PolyQ lengths (Figure 1.3). The same study also suggested that the expansion significantly lowers the age of onset in their small cohort (n=65) (Elden *et al.*, 2010, Lagier-Tourenne and Cleveland, 2010).

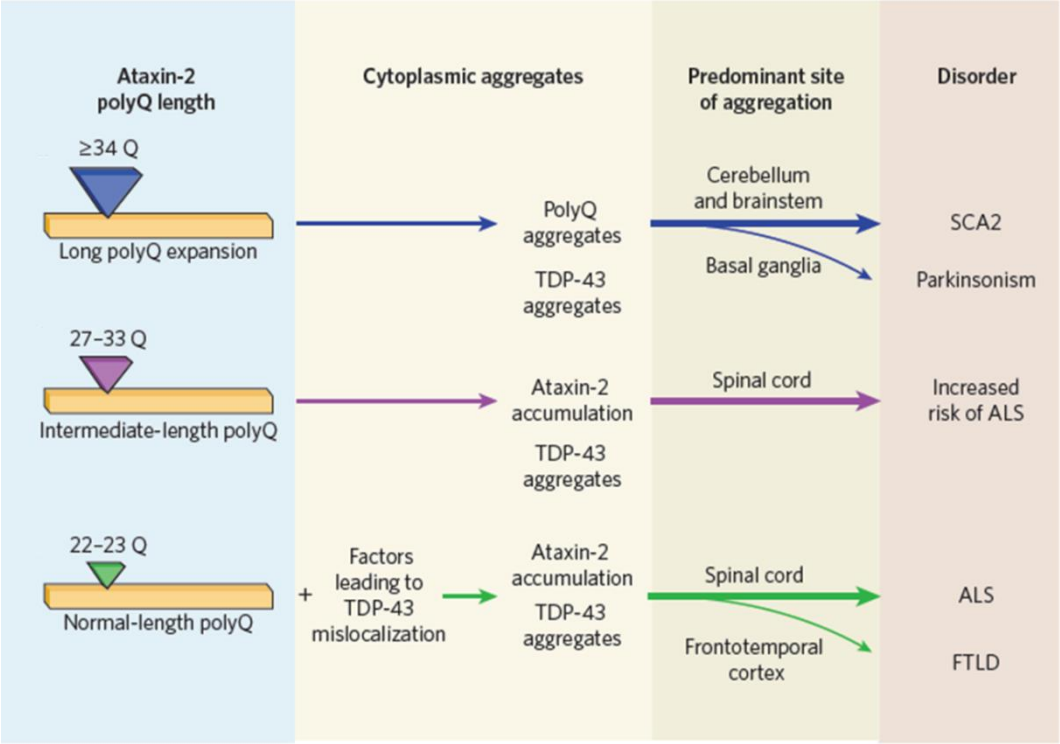


Figure 1.3. Ataxin-2 PolyQ length and related neurodegenerative disorders (Lagier-Tourenne and Cleveland, 2010).

The association of intermediate ATXN2 PolyQ expansions with ALS, triggered many screening studies in different ethnic populations. All follow-up studies in the European, Belgian, Caucasian, Italian, French-Canadian, South-West Chinese and Turkish populations confirmed this association (Table 1.2) (Chen *et al.*, 2011, Corrado *et al.*, 2011, Daoud *et al.*, 2011, Lahut *et al.*, 2012, Lee *et al.*, 2011b, Ross *et al.*, 2011, Van Damme *et al.*, 2011). However, as opposed to the first report, these studies did not find any clinical difference among ALS cases with PolyQ-expanded versus normal allele. Subsequently, a meta-analysis of all the previous reports, aiming to refine the length of risk alleles, revealed that the number of ALS patients carrying 30-33Q in ATXN2 is significantly higher than the controls (Gellera *et al.*, 2012).

Table 1.2. Major ATXN2 and ALS association studies.

Population	**# of cases**	**# of controls**	**Associated PolyQ expansion, (# of cases with expansion)**	**Reference**
North American	915 ALS	980	27-33, (43)	(Elden *et al.*, 2010)
European	894 sALS and 400 fALS	697	>30, (13)	(Lee *et al.*, 2011b)
Belgian and Dutch	1845 sALS and 91 fALS	2002	29-39, (28)	(Van Damme *et al.*, 2011)
Caucasian	532 ALS	4877	27-36, (33)	(Ross *et al.*, 2011)
Italian	219 sALS and 13 fALS	395	31-37, (7)	(Corrado *et al.*, 2011)
French-Canadian	461 sALS and 13 fALS	471	29-37, (25)	(Daoud *et al.*, 2011)
Southwest Chinese	345 sALS	350	24-31, (15)	(Chen *et al.*, 2011)
Turkish	158 sALS and 78 fALS	420	31 and 32, (4)	(Lahut *et al.*, 2012)

Identification of the association between ATXN2 and ALS brings up several questions. One important consideration is whether the intermediate length PolyQ expansion in ATXN2 is ALS-specific or not. One study reported that PolyQ repeat expansion (30-35Q) is significantly associated with ALS but not PD, using 559 ALS and 1490 PD cases from central Europe (Gispert *et al.*, 2012). Another study with a large cohort including ALS (n=532), AD (n=1530), FTD (n=641), PD (n=702) and progressive supranuclear palsy (PSP) (n=514) cases proved the association of ATXN2 and ALS. Interestingly, significant ATXN2 expansions (>30Q) were also found in PSP which is a taupathy whose cases usually do not have TDP-43 positive inclusions in their affected neurons (Ross *et al.*, 2011). This discovery raised the possibility that the ataxin 2 activity in ALS may also be independent of TDP-43 toxicity. Supportively, a subsequent report showed that the PolyQ expansion (31Q) also modifies the translocation of ALS-causing FUS mutants, such as R521C and R521H, and induces early and late apoptotic events as well (Farg *et al.*, 2013). A second consideration about the ATXN2-ALS association is that

whether PolyQ disease genes, other than ATXN2, are specific to ALS or not. In one study, intermediate length expansions in ATXN3 gene were evaluated and no association was found with ALS (Gispert *et al.*, 2012). Moreover, in a separate study, the genes responsible for several other spinocerebellar ataxia types, including SCA1, SCA3, SCA6, SCA7 and SCA17, in addition to the dentatorubral-pallidoluysian atrophy gene, and huntingtin which is the causative gene of Huntington disease, were screened in variable number of sALS cases. Among those genes, only ATXN2 PolyQ expansion (27-33Q) was found to be significantly associated with ALS (Lee *et al.*, 2011a). Last consideration is about the association of other genetic variations, such as interruptions in the PolyQ repeat sequence, single nucleotide polymorphisms (SNP), haplotypes and copy number variations (CNV), in/around ATXN2 with ALS. The PolyQ repeat region generally includes two CAA interruptions in the CAG repeats. Mutational loss of these interruptions was reported to increase the length instability of the PolyQ repeat region which may lead to the related diseases through expansion (Pujana *et al.*, 1999). A recent study, screening 40 ALS cases with ataxin 2 expansions (range 27-32Q), reported that all of the cases had 1-3 CAA interruption in the PolyQ region as opposed to SCA2 cases who generally carry pure CAG repeats. In the same study, genotyping of the two previously reported SNPs, rs695871 (major/minor allele: C/G) and rs695872 (major/minor allele: C/T), which are in the same exon with the PolyQ region were performed. Interestingly, while 21 patients bearing 3CAA interruptions had the GT haplotype, 18 of the remaining 19 patients with < 3CAA interruption had the CC haplotype. Furthermore, ALS cases with 3CAA interruptions were also shown to have earlier disease onset as compared to the < 3CAA interruption carriers (Choudhry *et al.*, 2001, Yu *et al.*, 2011). In addition to the previous reports, in the framework of this study, four ALS cases with intermediate PolyQ expansion (31-32) were reported in the Turkish ALS cohort (Lahut *et al.*, 2012).

To sum up, recent advances in ALS genetics revealed also the contribution of ATXN2 to disease pathology; this, however, raises a series of questions: (i) the ALS specific cut-off for ATXN2 PolyQ in other populations, (ii) existence of the ATXN2 PolyQ expansion and CAA interruptions in larger cohorts, (iii) most importantly, the identification and validation of ALS-associated genetic variations in the ATXN2 locus, such as SNPs and CNVs.

1.4. Genome-wide Association Studies in ALS

SNPs are single nucleotide changes and the most frequent form of DNA variations. They may affect several biological processes through changing amino acid sequence, mRNA stability and altering transcription factor binding affinity (Bush and Moore, 2012). GWAS is a chip-based array method that investigates SNPs in the entire genome in a large case-control cohort, in order to associate a genetic locus to the particular disease/phenotype.

Small-scale and time-consuming candidate gene approaches were replaced by more powerful, practical and applicable GWAS after three important achievements in the genetics field. The completion of the Human Genome Project revealed 99% of the genome sequence and identified thousands of SNPs. The International HapMap Project identified representatives of groups of SNPs as tagged-SNPs. These along with the with the advancements in chip-based array technologies, which decreased the costs drastically, made the simultaneous genotyping of millions of SNPs possible (Ikegawa, 2012).

GWAS relies on the "common disease/common variant" hypothesis, which suggests that the common diseases are influenced by common variants with low penetrance (effect). In this perspective, the establishment of a large case-control cohort is followed by genotyping the SNPs in the array chips. Later, the disease-associated SNPs, haplotypes or groups of SNPs segregating together are identified with intensive statistical methods (Figure 1.4). This information however, is not enough to associate a locus to the disease of interest. It needs to be validated in independent cohorts (Bush and Moore, 2012).

To date around 1,220 GWAS have been published in more than 230 traits (http://www.genome.gov). Although the majority of them failed to be validated in replication studies, the identification of novel loci and genes for age-related macular degeneration, Crohn's disease, type 2 diabetes and heart disease are several successful examples of GWAS (Bush and Moore, 2012). In addition to these, there are also a few achievements in GWAS in neurodegenerative diseases; among them, the most important one is the identification of ApoE-ε4, the major genetic risk factor for AD (Lambert *et al.*, 2009).

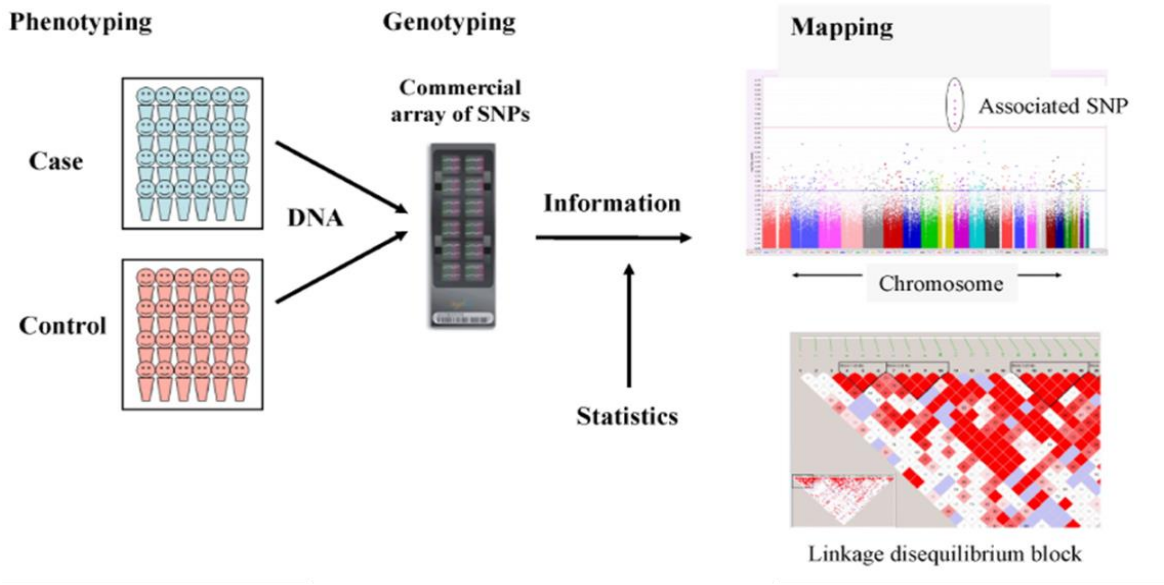


Figure 1.4. Schematic work-flow of a genome-wide association study (Ikegawa, 2012).

The situation in ALS is more complicated and the contribution of genetics to sALS is not well-defined yet. While some argue that genetic factors have almost no influence on sALS manifestation, others propose that the existence of mutant fALS genes in sALS cases, twin studies showing heritability of sALS and previously identified sALS variants are evidences for the contribution of genetics to sALS (Ahmeti *et al.*, 2013). GWAS, thought to shed light to this discussion, performed with large number of ALS cases from different populations including, Ireland, USA, UK, Finland, Germany, Sweden, Netherlands and France (http://www.genome.gov). These studies reported several candidate loci, however, replication studies using different samples consistently validated only the association of 9p21 locus with ALS (Cronin *et al.*, 2008, van Es *et al.*, 2009). Following studies on this region revealed the presence of the hexanucleotide repeat expansion in C9ORF72 in both sALS and fALS cases as discussed above.

Recently, a meta-analysis, with all previously obtained ALS-GWAS data, was performed by the International Consortium on Amyotrophic Lateral Sclerosis Genetics in order to identify new ALS susceptible loci and factors affecting age of onset. Among 4252 ALS cases versus 5123 controls, eight genomic loci including 9p21, were found to be significantly associated with the disease. On the other hand, the previously presented

susceptible variants were not found to be associated with risk of ALS. The most significantly associated locus in this study was 1p34.1 that bears a glycosyltransferase gene, beta-galactoside alpha-2,3-sialyltransferase 3. Among 4252 sALS cases, the minor allele of a SNP in this locus, rs3011225, was found to lower age at onset for two years. Moreover, the effect of the SNP was consistent among all the populations included in the study (Ahmeti *et al.*, 2013).

Despite the fact that the discovery of many genes and risk factors for several genes were achieved via GWAS, in many common complex diseases such as cancer and NDs, the "identified" genetic factors affect the pathogenesis slightly. It means that "common disease/common variant" hypothesis fails for such diseases (Ikegawa, 2012). In other words, the notion "significantly associated" does not mean that it is always biologically significant, too. For example, the additive effect of a SNP, previously shown to be significantly associated with height, was found to be less than half a centimeter (Sanna *et al.*, 2008). Thus, in order to explain the molecular basis of complex diseases such as ALS, it is crucial to take in addition to common variants, the rare variants into account which the GWAS strategy does not consider. For this reason, more recent techniques are required such as Whole Exome/Genome Sequencing which captures all variants in the genome, from rare SNPs to large InDels. However, the enormous amount of data obtained from GWAS may still provide insights into many genetic studies, such as genotype information of SNPs and genome-wide detection of CNVs.

1.4.1. Copy Number Variation Studies in ALS

In contrast to SNPs, CNVs are rarer in the genome (<5%); they are changes in the number of copies of DNA fragments (eg. deletions or duplications), ranging from kilobases to megabases. Most importantly, changes in the copy numbers of DNA fragments may have more dramatic functional consequences because they can severely influence transcription and function of a protein (Blauw *et al.*, 2008, Bush and Moore, 2012). Thus, disease-susceptible CNVs may explain a larger proportion of disease pathology.

Recently, copy number variations in the SMN1, the causative gene for spinal muscular atrophy, were found to be associated with ALS. While duplication of this gene was reported to be a risk factor for sALS, homozygous deletions of the gene was shown to be a protective factor in the Swedish population (Blauw *et al.*, 2012, Corcia *et al.*, 2012). In addition to these, there are five genome-wide CNV studies in ALS, all depending on capturing SNP intensity as representatives of CNVs. Although no significant association has been reported so far, there are several possible candidates such as DPP6 and NIPA1 loci which need to be validated (Blauw *et al.*, 2010). Moreover, data from the first GWAS study in the Turkish ALS cohort revealed three different, possibly ALS-associated CNVs (Uyan *et. al.*, 2013 submitted to PLOSONE). These are the duplication of MAP4K3 which encodes for a key protein involved in many cellular signaling events, the total deletion of HLA-B which has a role in antigen presentation, and the partial deletion of EPHA3 whose protein product is implicated in axon guidance (Lam *et al.*, 2010, Mendez *et al.*, 2009, Vaidya *et al.*, 2003). Although these CNVs were found to be promising in the study cohort, they need to be further validated by another genotyping technique and confirmed in other and larger ALS cohorts.

2. PURPOSE

The recent acceleration in ALS genetics led to the identification of many genes and several associated loci, however the genetic basis of the majority of ALS cases is still unknown. This information may only be achieved by investigating possible disease-causing and -modifying genetic variations, such as mutations, SNPs and CNVs in the affected individuals. A recent comprehensive study combining yeast screen, fly/cell biology and human genetics led to the identification of the association of intermediate length polyQ expansions in the ATXN2 gene and ALS. Subsequent genetic screening studies in several different ethnic populations also showed this association. However, this association still needs to be investigated in other cohorts to confirm it worldwide. In addition to PolyQ expansions, a few studies investigated SNPs in this locus, reporting two possible disease-modifying SNPs in the ATXN2 gene.

The immense complexity of ALS pushes the scientists forward to investigate the recently discovered CNVs' impact on the disease which may have drastic biological effects. Thus far, genome-wide CNV studies on ALS cases revealed many candidates, however these need to be confirmed in independent replication studies. A study performed in our lab revealed for the first time, variations in MAP4K3, EPHA3 and HLA-B genes as ALS risk factors in the Turkish cohort (Uyan *et al.*, 2013, submitted to PLOSONE). To further confirm the association of these genes with ALS risk, CNV genotypes converted from GWAS intensity data had to be validated by additional/and more sensitive CNV genotyping techniques.

In this respect, this study has two major aims:

- To understand the contribution of the ATXN2 locus structure to ALS by
  (i) screening the state of ATXN2 PolyQ repeat length in a large Turkish ALS cohort and matched control population
  (ii) searching for ALS-associated SNPs or haplotypes in the ATXN2 locus

- To validate the results of a recent study performed at NDAL

(i) by genotyping the CNVs in a subset of the Turkish cohort understudy, including both patients and healthy controls, with specifically designed probes

3. MATERIALS

3.1. DNA Samples

In the framework of this thesis, a total of 256 ALS patients (158 sALS and 78 fALS), referred to our laboratory from several hospitals throughout Turkey and 420 anonymously collected, neurologically healthy Turkish controls from Haydarpaşa State Hospital were investigated. Written informed consent was obtained from all patients. The Ethics Committee of Boğaziçi University approved the use of patient samples. All cases and controls included into the ATXN2 polyQ repeat expansion analysis. SNP genotyping was performed in 135 ALS patients and 50 controls, CNV detection in 24 ALS patients and 24 controls.

3.2. Polymerase Chain Reaction (PCR) Amplification of ATXN2 PolyQ Repeat Region

3.2.1. Primers

The primers for the amplification of ATXN2 PolyQ repeat region were purchased from Integrated DNA Technologies, Inc. USA (Table 3.1).

Table 3.1. Primers used for PCR amplification.

ATXN2	**Forward :** 5'- GGGCCCCTCACCATGTCG -39
	Reverse-1 : 5'- /56-FAM/CGG GCT TGC GGA CAT TGG -3'
	Reverse-2 : 5'- CGG GCT TGC GGA CAT TGG -3'

3.2.2. Enzymes and Chemicals

Enzymes and chemicals used in the amplification of ATXN2 PolyQ repeat region are shown in Table 3.2.

Table 3.2. Components of PCR amplification.

FastStart Taq DNA Polymerase	5U/ µl, Roche, USA
$MgCl_2$	25mM in dH_2O Takara, Japan
10X PCR Reaction Buffer without $MgCl_2$	500 mM Tris/HCl (pH 8.0) 100 mM KCl 50 mM $(NH_4)_2SO_4$ (pH 8.3) Roche, USA
5X GC-RICH solution	Roche, USA
Deoxyribonucleotides (dNTP)	25 mM, Takara, Japan

3.2.3. Software for Evaluating PCR Results

GeneScan and sequencing data were analyzed using the software in Table 3.3.

Table. 3.3. Software for GeneScan and sequencing analyses.

Peak Scanner	Peak Scanner™ Software v1.0, Life Technologies Corporation, USA
FinchTV v.1.4.0	Geospiza, USA

3.3. Agarose Gel Electrophoresis

All the necessary chemicals for agarose gel electrophoresis are compiled in Table 3.4.

Table 3.4. Materials used in agarose gel electrophoresis.

Agarose	Prona, EU
10X TBE Buffer	0.89 M Tris-Base 0.89 M Boric acid 20 mM Na_2EDTA (pH 8.3)
Ethidium Bromide	MP Biomedicals, France
6X Loading Dye	10mM Tris-HCl (pH 7.6) 0.03% Bromophenol Blue 0.03% Xylene Cyanol FF 60% glycerol 60 mM EDTA Fermentas, Lithuania
DNA Ladder	1000 bp (Invitrogen, USA) 50 bp (MBI Fermentas, Lithuania)

3.4. SNP Dataset from Turkish ALS-GWAS

The study setup of the previously performed Turkish ALS-GWAS is shown in Table 3.5.

Table 3.5. Properties of the Turkish ALS-GWAS.

sALS # vs control #	116 sALS vs 109 Turkish controls
Genotyping Platform	Illumina HumanOmniExpress SNP Array, Illumina, San Diego, CA, USA
Genotyped SNP #	733,202

3.5. SNP Genotyping

3.5.1. SNP Genotyping Assays, Solutions and Software

2X TaqMan Universal PCR Master Mix, No AmpErase UNG and all assays for SNP genotyping were purchased from Applied Biosystems, Inc., USA (Table 3.6).

Table 3.6. TaqMan SNP genotyping assays.

Assay ID	Targeted SNP	Located Gene
C_31831338_10	rs10849949	ATXN2
C_2539513_20	rs2073950	
C_2539510_20	rs2301621	
C_2981072_10	rs3184504	SH2B3
C_2539516_20	rs739496	
Custom Assay	rs2239194	

Stepone Software v.2.2.2, Applied Biosystems Inc., USA was used for SNP genotyping.

3.6. CNV Detection

3.6.1. Copy Number Assays

All TaqMan assays specific to the CNV of interest and buffers and solutions were purchased from Applied Biosystems (Table 3.7, Table 3.8).

Table 3.7. TaqMan CNV assays.

Assay ID	**Targeted Gene**	
Hs02852136_cn	MAP4K3	Applied Biosytems, Inc., USA
Hs03605931_cn	HLA-B	
Hs03458738_cn	EPHA3	

Table 3.8. Buffers and solutions for CNV detection.

2X TaqMan Genotyping Master Mix	Applied Biosystems, Inc., USA
TaqMan Copy Number Reference Assay, 20X	

3.6.2. Software

Installed software for CNV detection are shown in Table 3.9.

Table 3.9. Software used for CNV detection.

Stepone Software v.2.2.2	Applied Biosystems Inc., USA
Copy Caller v2.0	

3.7. Kits

QIAquick Gel Extraction Kit (Qiagen, USA) was used for DNA extraction from agarose gel.

3.8. Equipment

All the lab equipment used in this thesis is compiled in Table 3.10.

Table 3.10. Equipment used in this thesis.

96-well plate	MicroAmp Fast, Applied Biosystems, USA
96-well plate sealer	Applied Biosystems, USA
Autoclave	Model ASB260T, Astell, UK
Balances	Model VA124, Gec Avery, UK Model CC081, Gec Avery, UK TE612, Sartorious, Germany
Centrifuges	Allegra X22-R, Beckman Coulter, USA Centrifuge 1-15, Sigma, USA Centrifuge 2-16K, Sigma, USA
DNA Isolation Insturement	MagNA Pure Compact, Roche
Electrophoretic Equipments	Wide Mini-Sub Gel GT, BIO-RAD, USA Mini-Sub Gel GT, BIO-RAD, USA
Eppendorf Tubes	1,5 ml Boil-Proof Microtubes, Axygen, USA 0,2 ml and 0,5 ml Thin Wall Flat Cap, Axygen, USA
Gel Documentation System	GelDoc Documentation System, BIO-RAD, USA
Laptop Computer	Latitude, Dell, USA
Magnetic Stirrer	Chiltern Hotplate, HS31, UK Hotplates MR3001, Heidolph, Germany

Table 3.10. Equipment used in this thesis (continued).

Ovens	MD 554, Microvave Oven, Arçelik Turkey BD53, Binder, Germany
pH Meter	PB-11, Sartorious, Germany
Power Supply	EC250-90, Thermo Scientific, USA
Thermalcyclers	Techgene, Techne, TC-312, Bibby Scientific, UK
Real-Time PCR System	StepOnePlus, Applied Biosystems Inc., USA
Refrigerator	4250T, Arçelik, Turkey
Shaker	Duomax Platform Shaker, Heidolph, Germany
Spectrophotometer	NanoDrop ND-2000c, Thermo, USA
Tips	Universal Fit Filter Tips, Axygen, USA
Vortex	Fisons WhirliMixer, UK Reax Top, Heidolph, Germany
Water Bath	Gemo DT104, TEST Laboratuvar Cihazları, TR
Water Purification	Millipore, USA

4. METHODS

4.1. DNA Extraction

DNA extraction was performed according to the instructions of MagNA Pure Compact Instrument (Roche), a robotic system, designed for extraction of nucleic acids (DNA or RNA) from tissue or blood samples.

4.2. Quality Control of Genomic DNA

4.2.1. Qualitative Analysis by Agarose Gel Electrophoresis

DNA quality was tested by running the samples in 1% agarose gel. One gram of agarose was dissolved in 100 ml 0.5X TBE buffer. The DNA-intercalating agent EtBr which enables visualization of DNA under UV light was added into the solution with a final concentration of 0.5 µg/ml. The solution was gently mixed and poured into an electrophoresis plate and left to polymerize at room temperature. After polymerization, the gel was put into an electrophoresis tank with 0.5X TBE. Five µl PCR product was mixed with 6X loading dye, loaded and run at 120 Volt for 20-25 minutes. Finally, the gels were put under UV light, the images captured and documented.

4.2.2. Quantitative Analysis by Spectrometric Measurement of DNA

Measurement of DNA concentrations was done by a NanoDrop spectrometer. At 260 nm optical density (OD 260), the elution buffer of MagNA Pure Compact Instrument was used as blank. One µl from each sample was put on the optical tip of the instrument and concentrations were documented.

4.3. ATXN2 PolyQ Repeat Determination

4.3.1. PCR Amplification of ATXN2 PolyQ Region

The amplification of the PolyQ region was performed by the specific Forward and Reverse-1 primers (Table 3.1). 30-40 ng genomic DNA was used in each PCR reaction with the reagents listed in Table 4.1. After pipetting all components, the volume of each PCR-mix was completed to 25 µl with dH_2O. The cycling conditions are shown in Table 4.2.

Table 4.1. PolyQ repeat region PCR amplification components.

Reagent	**[Stock]**	**Volume (µl)**	**[Final]**
$MgCl_2$	25 mM	1.5	3.5 mM
Buffer	10X	2.5	X
GC Rich	5X	5.0	X
dNTP	5 mM	1.0	0.8 mM
Forward Primer	12.5 pmole/µl	1.0	0.5 mM
Reverse Primer 1	12.5 pmole/µl	1.0	0.5 mM
Taq Polymerase	5 u/ µl	0.23	1 u/µl

Table 4.2. Cycling conditions for PCR.

Initial Denaturation	94 ^oC	5 minutes	-
Denaturation:	95 ^oC	60 seconds	30 cycles
Annealing:	55.7 ^oC	60 seconds	
Extension:	72 ^oC	90 seconds	
Final Extension	72 ^oC	5 minutes	-

4.3.2. Determination of ATXN2 PolyQ Region by GeneScan Analysis

For PolyQ repeat size determination GeneScan analysis was used. GeneScan analysis, based on capillary electrophoresis, compares the emission of fluorescently-labeled primers in the PCR product with the size standards. Subsequent analysis in compatible software allows determination of the sizes of DNA fragments. In this study, 10 µl PCR product from all individuals together with positive and negative controls, produced by 5´ FAM-labeled primer (Reverse-1), were subjected to each set in GeneScan Analysis (Macrogen, Korea). The results were analyzed by PeakScanner Software by using 500 LIZ size standards.

4.4. ATXN2 Sequencing

To test whether the intermediate length polyQ repeats contain any CAA sequences, ALS cases with an intermediate length PolyQ expansion were subjected to DNA sequencing. For this, Forward and Reverse-2 primers were used according to the protocol above.

After running (120 V, 35 minutes) the PCR products in a 2% gel, the short and long bands were cut from the gel under UV light and separately labeled. Prior to sequencing, QIAquick Gel Extraction Kit was used to purify PCR products from gel, following manufacturer's instructions. Sequencing with 500 ng/10 µl PCR-amplified DNA and 5 pmole/µl primers was performed at RefGen Company (Ankara, Turkey). Results were analyzed with compatible software (FinchTV v.1.4.0).

4.5. SNP and Haplotype Analysis from Turkish ALS-GWAS

The results of the Turkish ALS-GWAS data, previously performed in our lab in a collaborative study, was used for single marker and haplotype analyses in the ATXN2 region. Twenty eight SNPs in a 250 kb region on chromosome 12q (110,300,000–110,550,000), including the ATXN2 gene, was investigated. Hardy-Weinberg disequilibrium (HWE) and Minor Allele Frequency (MAF) scores of these SNPs were calculated using the PLINK toolset (http://pngu.mgh.harvard.edu/purcell/plink/). To SNPs

which pass the threshold for the HWE ($p>0.05$) and MAF ($f<0.01$) scores, single marker analysis was applied. The significance of single marker analysis was calculated. Haplotype analysis was performed for investigating ALS-association of haplotypes on this region using the Haplowiev 4.2 program (http://www.broad.mit.edu/haploview). After selecting the candidate haplotypes, 1000 permutation test (multiple testing) was applied to eliminate false positive results. Linkage Disequilibrium (LD) plots and co-segregation patterns (R^2) of neighbouring SNPs were calculated to select candidates for SNP Genotyping by TaqMan assays.

4.6. SNP Genotyping by TaqMan Assays

Six SNPs from ALS-GWAS data were subsequently genotyped by TaqMan SNP Genotyping Assays in 135 ALS cases and 50 Turkish healthy controls that were not present in the GWAS data.

The SNP genotyping reaction requires three components: (i) genomic DNA from individuals to be tested, (ii) TaqMan Universal PCR Master Mix, No AmpErase UNG which contains AmpliTaq Gold® DNA polymerase and (iii) SNP Genotyping Assay which contains sequence-specific forward and reverse primers and two TaqMan® MGB probes, labeled with VIC or FAM dyes. The method principally depends on PCR amplification of the desired genomic region and detection of the intensity, emitted by VIC or FAM labeled probes which specifically bind to two different alleles of the relevant SNP.

Prior to genotyping, genomic DNA from all samples was diluted to a final concentration of 2 ng/µl (± 0.3) with distilled water. After sample preperation, the real-time PCR instrument was set for SNP genotyping. First, the genotyping preference was selected and adjusted to the 96 well-plate option for standard run (Table 4.3). The ID of SNP to be genotyped was typed to the assay name box. In the next step, to design the plate, all samples and the target SNPs were assigned to each well of the plate. One negative control (no-template-control) was required in each assay. After setting the instrument, SNP Genotyping Assays were thawed, briefly vortexed, spinned down and put on ice before preparing the reaction mix (Table 4.4). The 96-well plate was placed on ice (in its case). Five ng of genomic DNA from each sample was pipetted into the wells then, the reaction

mix was distributed to the samples. After adding all components, the 96-well plate was sealed with MicroAmp™ Optical Adhesive Film and spinned down to remove the air bubbles.

Table 4.3. SNP genotyping cycling conditions.

Stage	**Temperature**	**Time**
AmpliTaq Gold Enzyme Activation	95 °C	10 minutes
Cycle (50 cycles)	92 °C	15 seconds
	60 °C	60 seconds

Table 4.4. SNP genotyping reaction mix.

Component	**µl per well**
TaqMan Universal PCR Master Mix (2X), No AmpErase UNG	10.0
20X SNP Genotyping Assay	0.5
DNase-free water	7.0
Total volume per well	17.5

After PCR cycling, allelic discrimination analysis screen in the StepOne Software was used to analyze the genotyping data. The data sheet of each SNP assay provides the information for the genotype-dye correlation. Determination of the genotypes of each allele was performed according to the allelic discrimination plot in the results screen. Finally, SNP genotypes from the GWAS data and the TaqMan assays, were combined for analyses and the odds ratios (OR) were calculated.

4.7. CNV Detection by TaqMan Assays

The candidate CNVs previously identified by the ALS-GWAS in the MAP4K3, HLA-B and EPHA3 genes were further investigated using the TaqMan Copy Number Assay protocol, which is a simple method designed for detecting and measuring CNVs in mouse and human genomes. The method principally depends on two main steps. The first step is a duplex real-time PCR reaction which amplifies the CNV of interest and a reference sequence (RNase P H1 RNA gene), known to exist in two copies in the human genome. Next, the C_T values are determined by measurement of the intensities of the FAM™ dye-labeled MGB probe which binds to target, and the VIC® dye-labeled TAMRA™ probe, binding to the reference sequence (Figure 4.1). Finally, comparative C_T ($\Delta\Delta C_T$) is determined by a software or manually.

Each DNA sample was first diluted to 5 ng/µl (± 0.4) with distilled water and the MagNA Pure Elution Buffer was used as blank. Then, the real-time PCR instrument was set: First, the 96-well plate Quantitation-Standard Curve option was selected to observe C_T values of each assay. Then, the names and dye labels (VIC or FAM) of target assay and reference were entered into the instrument (as specified in data sheet of each assay). Later, all required wells and the plate layout were designed and the cycling conditions were selected (Table 4.5). When the instrument was ready, 4 µl (20 ng) of all DNA samples were pipetted in to the bottom of each well as triplicates; all components of the CNV detection reaction mix were gently vortexed and spinned down. The reaction mixture was prepared in an Eppendorf tube after calculating the required total volumes of each component (Table 4.6). The mixture was vortexed and spinned down, and 16 µl of the reaction mix was pipetted to each well. The plate was sealed with MicroAmp™ Optical Adhesive Film and spinned down to remove the air bubbles before it was placed into the PCR instrument.

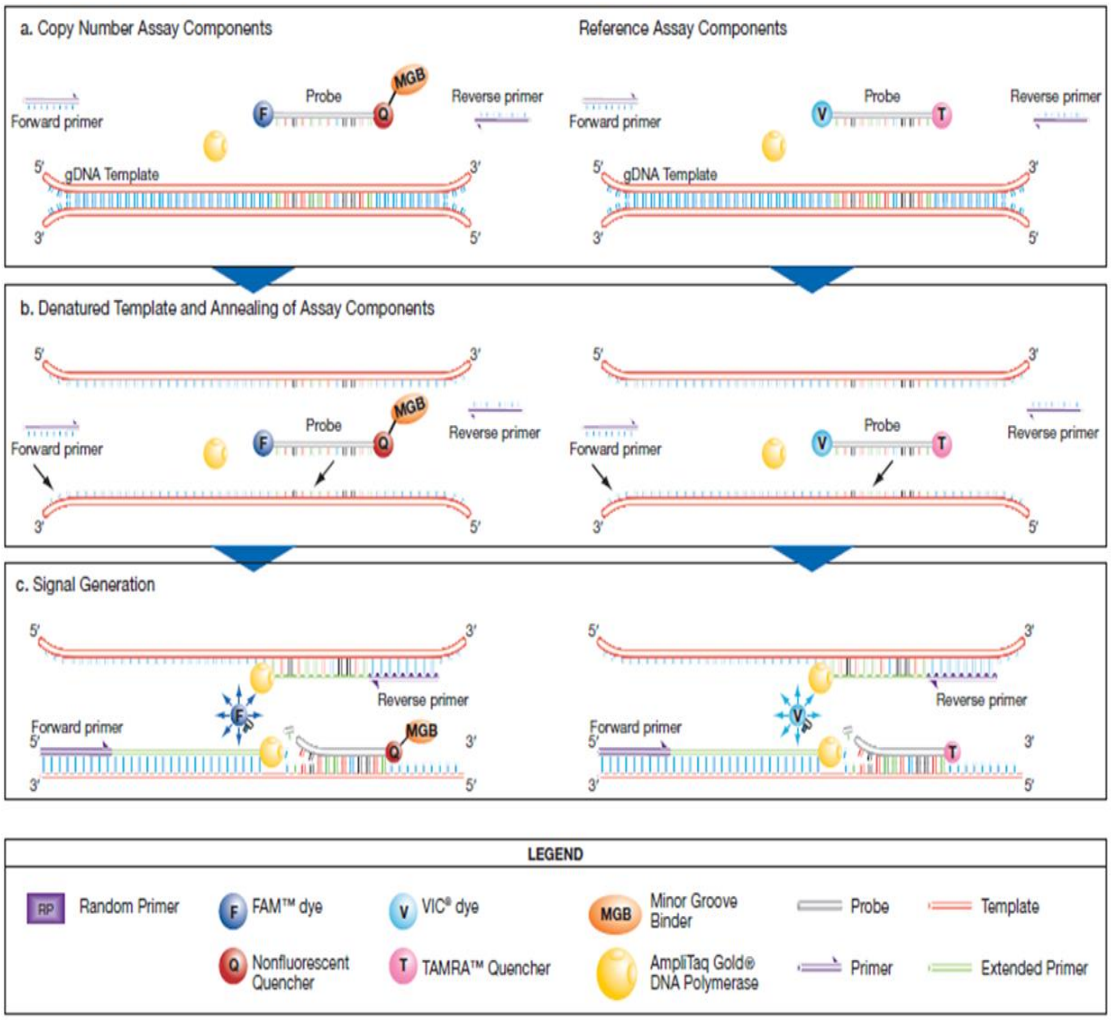


Figure 4.1. Detection of target and reference DNA sequences in a duplex real-time PCR.

Table 4.5. CNV detection cycling conditions.

Stage	Temperature	Time
Hold	95 °C	10 minutes
Cycles (40)	95 °C	15 seconds
	60 °C	60 seconds

Table 4.6. CNV detection reaction mix.

Component	**µl per well**
2X TaqMan Genotyping Master Mix	10.0
TaqMan Copy Number Assay, 20X	1.0
TaqMan Copy Number Reference Assay, 20X	1.0
Nuclease-free water	4.0
Total volume	**16.0**

The C_T data obtained from StepOne software were imported to CopyCaller program as txt file after the PCR was finished. Data were analyzed by Without Calibrator option which asks the user to enter the expected (most probable) copy number for the CNV of interest to the program. For the probes designed for the detection of CNVs in MAP4K3, HLA-B and EPHA3; three, zero and one copies were typed to the program, respectively. Finally, samples were ready to be analyzed and the output was documented.

5. RESULTS

5.1. ATXN2 PolyQ Expansion Analyses

The ATXN2 PolyQ repeat sizes of 236 ALS cases and 420 healthy Turkish controls were investigated after the PCR amplification of the region. The PCR products were first run on agarose gels to test the quality of the amplifications. Although the expanded alleles were visible on the agarose gel, the exact number of PolyQ repeat was indiscernible in this method (Figure 5.1).

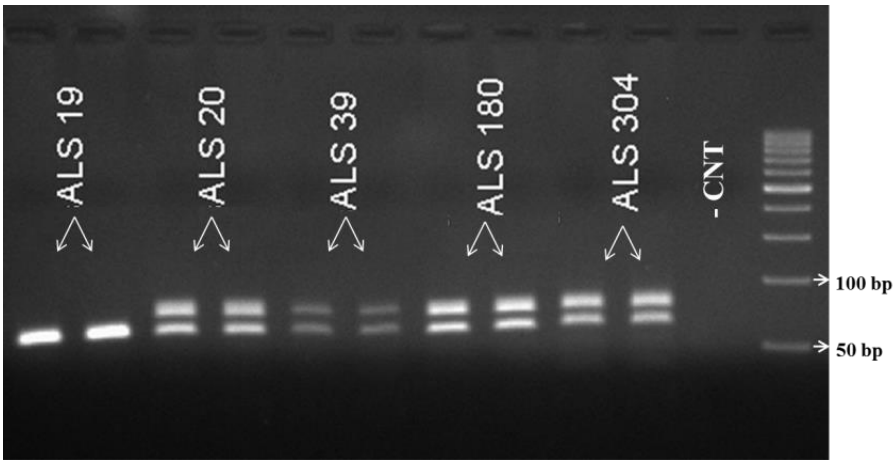


Figure 5.1. A representative gel image of the PCR products. ALS 19 is homozygous and the remaining four individuals are heterozygous for the ATXN2 PolyQ repeat.

Subsequently, GeneScan analysis was applied to all samples to size each allele. GeneScan analyses confirmed the results obtained from gel electrophoresis, giving precise information about the PolyQ repeat size of each allele. To validate the GeneScan results, 10% of the samples were reanalyzed and consistent results were obtained. Representative examples, showing a homozygous sample with normal (22/22) and a heterozygous individual with intermediate length PolyQ repeat number (22/32), are shown in Figure 5.2.

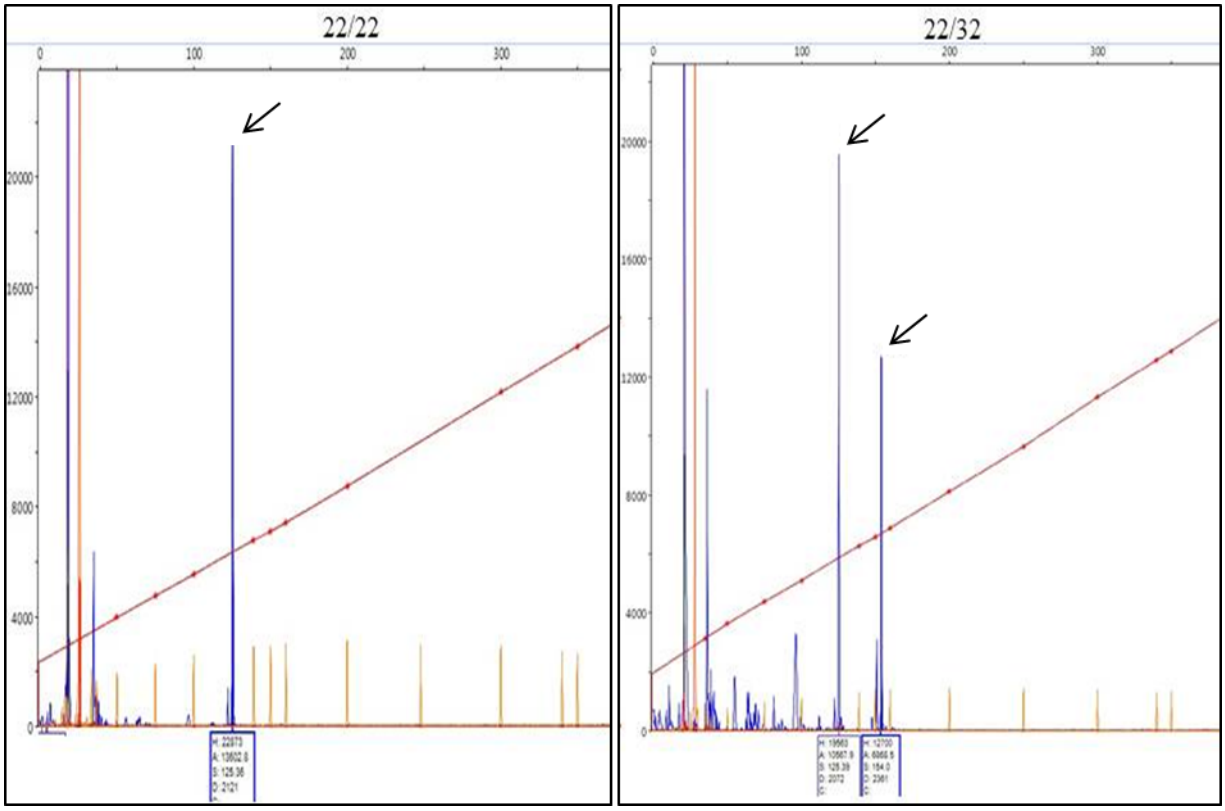


Figure 5.2. Representative examples of GeneScan results. Arrows indicate the allele peaks. Left panel: a homozygous case with normal repeats (22/22); right panel: a heterozygous case with 22/32 PolyQ repeats.

GeneScan analyses revealed that the lengths of PolyQ repeats varied from 13 to 32 in ALS cases and 15 to 29 in controls. The 22 repeat PolyQ was the most abundant allele accounting for approximately 92% of cases and controls (Figure 5.3). Twenty eight ALS cases (18 fALS and 10 sALS) carrying mutations in different ALS genes were also included into this study in order to investigate the possible influence of the ATXN2 PolyQ expansion on ages at disease onset. None of the 28 cases was found to harbour an expansion in the ATXN2 PolyQ region (Table 5.1).

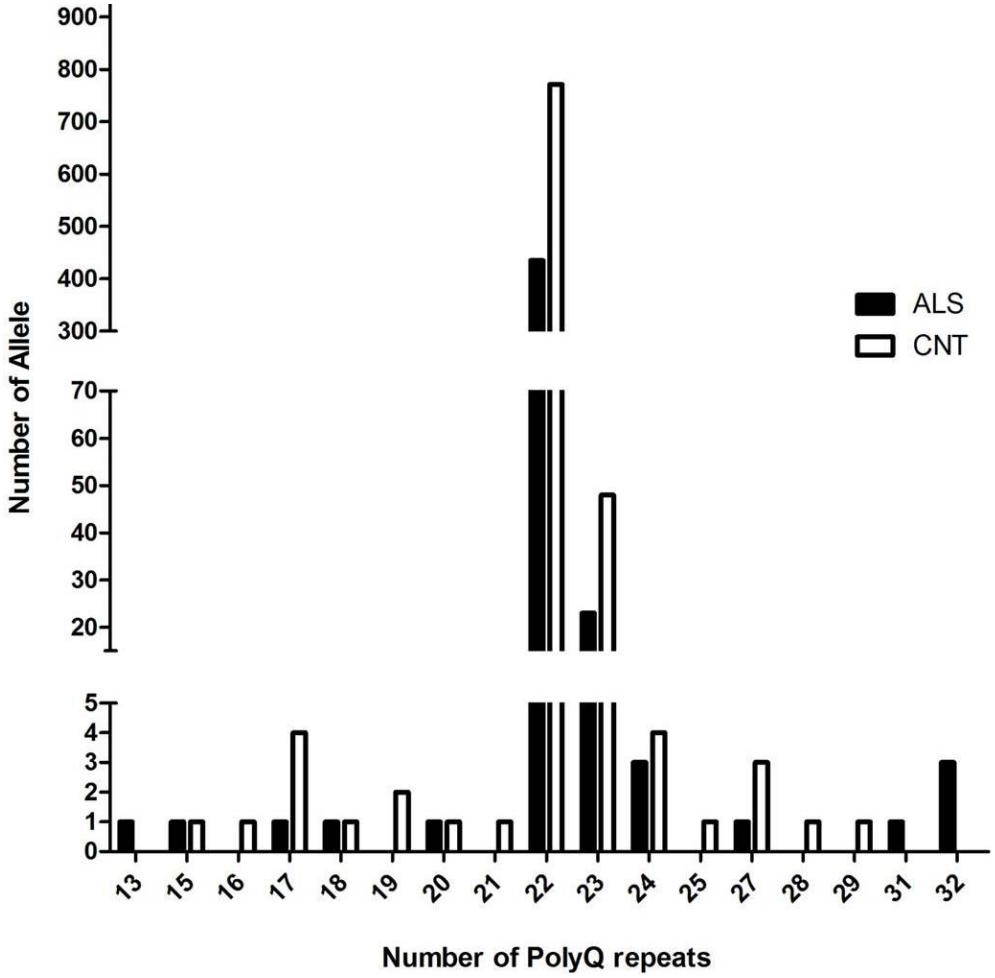


Figure 5.3. Distribution of the ATXN2 PolyQ repeat sizes.

For statistics, two different values were applied. The commonly used cutoff (27-33) in several previous studies investigating ATXN2-ALS association, was not statistically significant in this analysis (p=0.168). The 31 or 32 repeat alleles were present in heterozygous state in four ALS cases (one 22/31, three 22/32), however there was no control with >30 PolyQ repeat. Fisher's exact test showed that a cutoff of >30 PolyQ appeared optimal to discriminate ALS samples from controls in this study (p= 0.01721) (Table 5.2).

Table 5.1. The ATXN2 PolyQ repeat sizes of the Turkish cases with ALS mutations.

Patient No	Gender	AO	Mutation	ATXN2 PolyQ repeat
fALS 61	female	53	SOD1: L144F	22/22
fALS 221	male	20	SOD1: A4S	22/22
fALS 226	male	19	SOD1: H71Y	22/22
fALS 147	male	50	SOD1: D90A	17/22
fALS 310	male	55	SOD1: D90A	22/22
fALS 191	male	28	SOD1: N86S	22/22
fALS 264	male	16	FUS: Δ143-148	22/22
sALS 131	female	16	UBQLN2: M392I	22/22
fALS 175	male	22	UBQLN2: P525S	22/22
fALS 268	male	26	UBQLN2: P560S	22/22
sALS 350	male	14	UBQLN2: M392I	22/22
fALS 132	male	14	SPG11: F2265L	22/22
fALS 167	male	13	PLEKHG5: P630H	22/22
fALS 256	male	42	OPTN: Δ 2 bp in 359 aa	22/22
sALS 19	male	32	C9ORF72: Expansion	22/22
sALS 37	female	69	C9ORF72: Expansion	22/22
fALS 56	female	46	C9ORF72: Expansion	22/22
sALS 70	male	48	C9ORF72: Expansion	22/23
sALS 83	male	49	C9ORF72: Expansion	22/22
sALS 120	female	48	C9ORF72: Expansion	22/23
fALS 184	male	46	C9ORF72: Expansion	22/22
sALS 201	female	55	C9ORF72: Expansion	22/22
fALS 202	male	56	C9ORF72: Expansion	22/22
fALS 298	male	80	C9ORF72: Expansion	22/23
fALS 321	male	63	C9ORF72: Expansion	22/22
fALS 336	female	37	C9ORF72: Expansion	22/23
sALS 368	male	42	C9ORF72: Expansion	22/22
sALS 376	female	52	C9ORF72: Expansion	22/22

AO: Age of onset, Δ: deletion, aa: amino acid.

Table 5.2. The frequencies of the ATXN2 expansion sizes among cases and controls.

	≤30 repeats	**> 30 repeats**	**p-value[a]**
ALS (n=236)	232 (98.3%)	4 (1.69%)	0.01721
Controls (n=420)	420 (100%)	0	

[a] Fisher's exact test

The last part of the study, DNA sequencing was performed to ALS cases with an expanded ATXN2 allele to further test the PolyQ sizes determined by GeneScan analysis and to test whether these cases contain any CAA interruption in their repeat regions. Among four cases with intermediate length PolyQ, one (sALS 39) had no CAA, however the remaining three (sALS 20, sALS 180 and fALS 304) had one CAA interruption in the PolyQ region. Additionally, the similar interruption patterns of the CAA triplets in three cases were found in these cases. Clinical information and sequence data are shown in Table 5.3.

Table 5.3. Clinical characteristics of the Turkish ALS patients with PolyQ expansion.

ALS No	**Gender**	**Birth**	**AO**	**AD**	**SO**	**Genotype**	**Sequence Composition**
sALS 20*	female	1950	52	57	LE	22/31	$(CAG)_{21}CAA(CAG)_9$
sALS 39	male	1962	39	44	LE	22/32	$(CAG)_{32}$
sALS 180*	female	1929	77	-	Bulbar	22/32	$(CAG)_{22}CAA(CAG)_8$
fALS 304*	female	1982	8	alive	Bulbar, LE	23/32	$(CAG)_{23}CAA(CAG)_8$

AO: age of onset, AD: age at death, SO: site of onset, LE: lower extremity, * case with the GGGAA haplotype.

5.2. SNP and Haplotype Association Analyses in the ATXN2 Locus

In addition to screening of the ATXN2 PolyQ repeat size distribution, SNP and haplotype associations in this region were investigated. In this extension, two datasets were combined: (i) the previously performed Turkish ALS-GWAS data of NDAL (116 ALS

cases and 109 controls), and (ii) genotype information of the selected five SNPs in additional 185 Turkish individuals (135 ALS cases and 50 controls). The combination of these two datasets will be referred to in this thesis as ‘combined-data’.

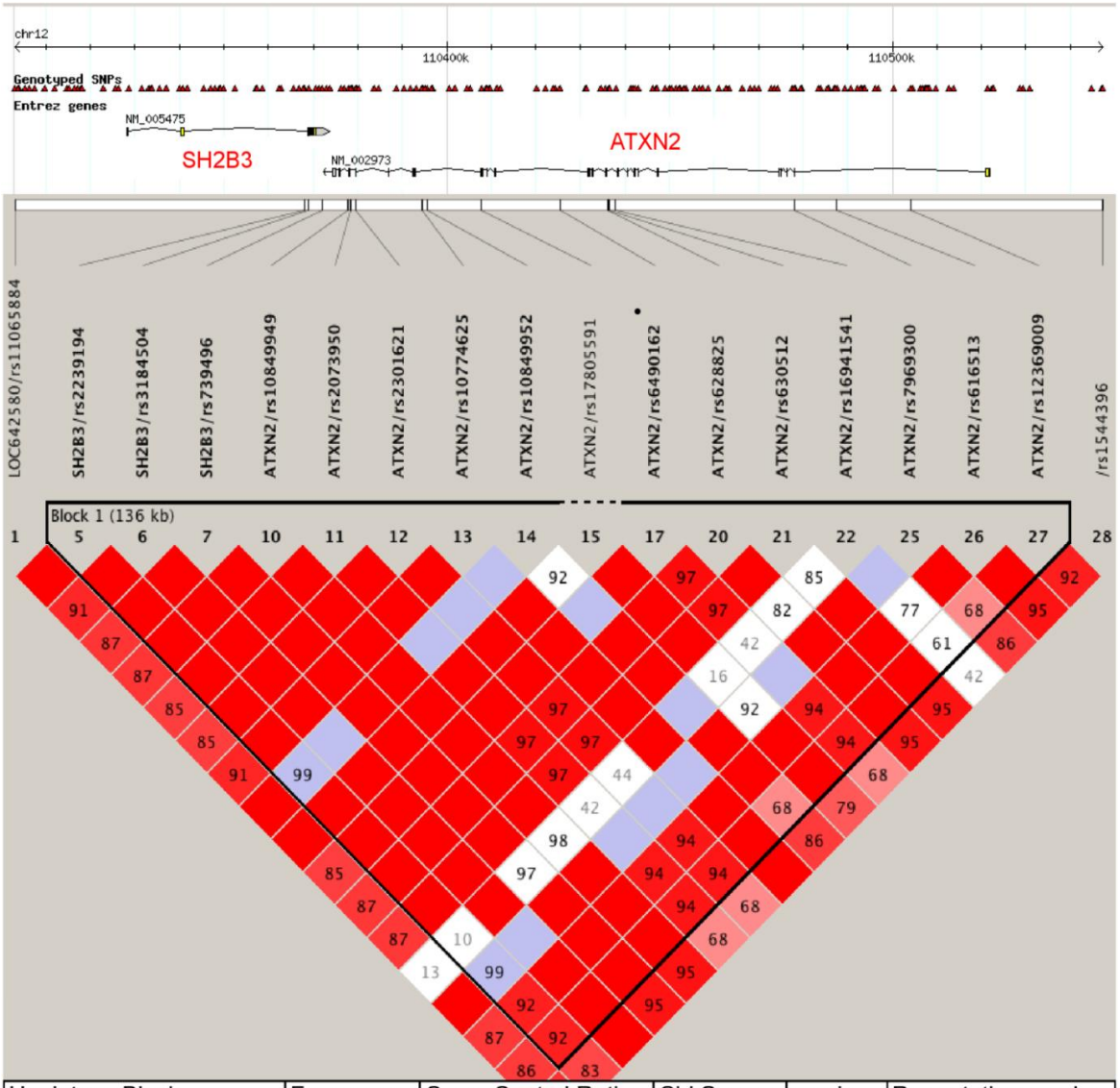
5.2.1. SNP and Haplotype Association Analyses of the ALS-GWAS Data

The SNP data of the Turkish GWAS study, performed with 116 ALS patients and 109 healthy controls, was first subjected to single marker analysis using Haploview Software. A 250 kb region (110,300,000-110,550,000), including 28 SNPs in the ATXN2 locus was investigated. Among 28 SNPs, ten were excluded due to low HWE and MAF scores and the remaining 18 were subjected to single marker analysis. In the allelic test, none of the SNPs showed significant association, however, one SNP (rs2239194) within the SH2B3 gene which is the downstream neighbor of ATXN2 showed a trend towards association (p=0.063) (Table 5.4).

After single marker analysis, SNPs located in the region were further subjected to haplotype analysis by Haploview Software. Six different haplotype blocks with various frequencies in the cohort, ranging from 39.5% to 4.5% , were detected. Remarkably, a 15-SNP haplotype block (GGGGAAGAGAAGGAC, MAF=0.149) in a 136 kb region was found to be significantly associated with ALS (p=0.0057). After permutation test the haplotype retained its significance (p= 0.02). Interestingly, three out of the 15 SNPs were in SH2B3 and the remaining 12 SNPs were in the ATXN2 gene. This haplotype was represented in 19.4% of cases and 10.1% of controls (Figure 5.4). Two among four ALS cases with expanded ATXN2-PolyQ (sALS 20 and sALS 180) were found to harbour this haplotype heterozygously, however, the remaining two cases were not present in the ALS-GWAS data, so they could not be evaluated.

Table 5.4. Single marker association analysis of the ALS-GWAS data.

SNP ID	Associated Allele	Case, Control Ratios	Chi Square	p-value
LOC642580/rs11065884	G	0.348, 0.317	0.494	0.482
SH2B3/rs2239194	G	0.921, 0.867	3.457	0.063
SH2B3/rs3184504	A	0.396, 0.394	0.001	0.98
SH2B3/rs739496	G	0.329, 0.298	0.49	0.4838
ATXN2/rs10849949	G	0.330, 0.298	0.54	0.4623
ATXN2/rs2073950	A	0.281, 0.239	1.03	0.3102
ATXN2/rs2301621	A	0.283, 0.239	1.127	0.2884
ATXN2/rs10774625	A	0.409, 0.399	0.043	0.8358
ATXN2/rs10849952	A	0.952, 0.940	0.308	0.5791
ATXN2/rs17805591	G	0.969, 0.959	0.361	0.548
ATXN2/rs6490162	G	0.281, 0.239	1.03	0.3102
ATXN2/rs628825	A	0.319, 0.294	0.327	0.5677
ATXN2/rs630512	A	0.322, 0.294	0.416	0.5187
ATXN2/rs16941541	G	0.948, 0.945	0.018	0.8927
ATXN2/rs7969300	G	0.952, 0.940	0.308	0.5791
ATXN2/rs616513	A	0.322, 0.284	0.738	0.3903
ATXN2/rs12369009	C	0.307, 0.280	0.397	0.5284
/rs1544396	A	0.778, 0.743	0.76	0.3832

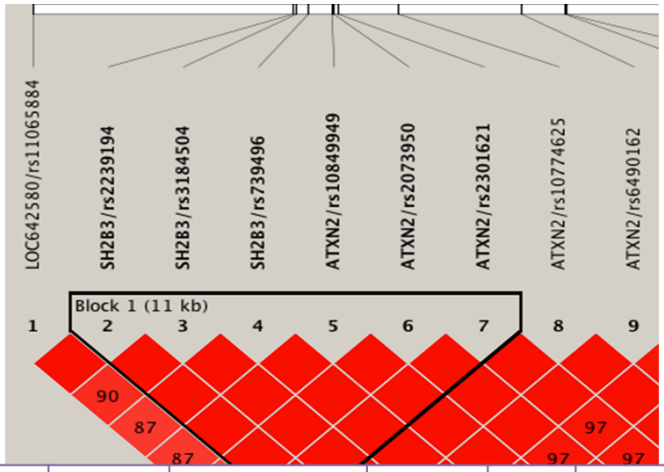


Haplotype Block	Frequency	Case, Control Ratios	Chi Square	p-value	Permutation p-value
GAAAGGAAAGGGGCA	0.395	0.396, 0.394	0.001	0.9800	1.000
GGAAGGGAAGGGGCA	0.228	0.209, 0.248	0.969	0.3250	0.927
GGGGAAGAGAAGGAC	0.149	0.194, 0.101	7.644	0.0057	0.020
AGGGAAGAGAAGGAC	0.101	0.080, 0.124	2.361	0.1244	0.567
GGAAGGGAAGGAGCA	0.054	0.052, 0.055	0.018	0.8927	1.000
GGGGGGGGAAAGAAC	0.045	0.035, 0.055	1.077	0.2993	0.918

Figure 5.4. Haplotype block analysis in the ATXN2 locus.The disease-associated block is highlighted in red.

5.2.2. SNP and Haplotype Association Analyses of the Combined-data

The identification of the ALS-associated haplotype, including SNPs in both ATXN2 and SH2B3 genes, initiated further experiments. To test the strength of this genetic association, the number of individuals included into the study was increased by genotyping with TaqMan SNP genotyping assays. Six SNPs (rs2239194, rs3184504, rs739496, rs10849949, rs2073950 and rs2301621) were selected as the representatives of the 15 SNP haplotype (Figure 5.5). Haploview detected five different haplotype blocks from these six SNPs in the ALS-GWAS data, and one of them was found to be statistically significant after permutation test (p= 0.02), confirming the representation of the six SNPs.

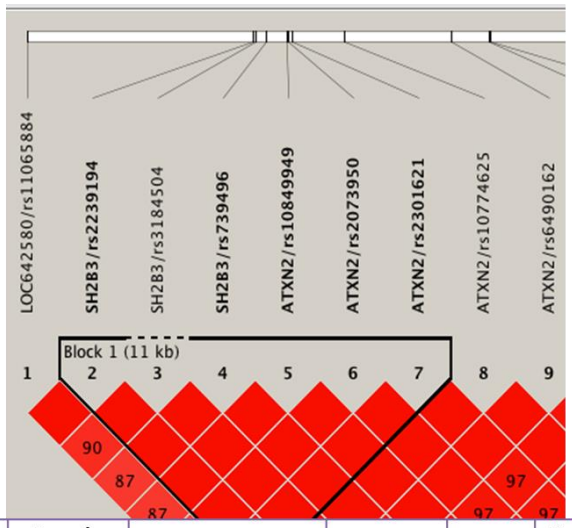


Haplotype	Freq. in population	Case, control ratios	Chi square	p-value	Permutation p-value
GAAAGG	0.400	0.402, 0.398	0.006	0.9379	1.000
GGAAGG	0.291	0.274, 0.307	0.765	0.4438	0.875
GGGGAA	0.153	0.199, 0.106	8.13	0.0062	0.020
AGGGAA	0.106	0.080, 0.133	3.349	0.0676	0.285
GGGGGG	0.054	0.049, 0.060	0.262	0.6096	0.982

Figure 5.5. Haplotype analysis of the six selected SNPs in the GWAS data. The disease-associated block is highlighted in red. The major parameters of the haplotype (case, control ratio and p-value) are similar to the 15-SNP haplotype.

Among the six TaqMan assays for the representative SNPs, one assay (targeting rs3184504) did not function properly (in trial experiments). Thus, the Haploview analysis using the ALS-GWAS data was repeated for the remaining five SNPs. As a result, the new 5-SNP haplotype (GGGAA) had similar values with the ALS-associated 15-SNP haplotype

in terms of its frequency in the population, the case and control ratios, and the p-values (Figure 5.4, Figure 5.6).



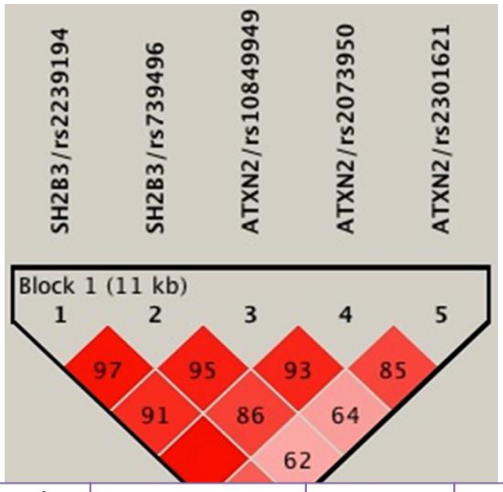
Haplotype	Freq. in population	Case, control ratios	Chi square	p-value	Permutation p-value
GAAGG	0.682	0.665, 0.699	0.58	0.4454	0.855
GGGAA	0.157	0.205, 0.106	8.13	0.0044	0.019
AGGAA	0.107	0.080, 0.134	3.34	0.0673	0.233
GGGGG	0.055	0.049, 0.060	0.26	0.609	0.976

Figure 5.6. Haplotype analysis of the five selected SNPs using the ALS-GWAS data. The disease-associated block is highlighted in red.

The five SNPs were genotyped in an additional 135 Turkish cases and 50 controls. The genotype results were combined with the ALS-GWAS data. Single marker analysis of the combined-data showed no significant association, although one SNP (rs2073950) had borderline significance (p=0.057) (Table 5.5).

Table 5.5. Single marker analysis of the 5 SNPs of the combined-data.

SNP ID	Associated Allele	Case, Control Ratios	Chi Square	p-value
SH2B3/rs2239194	G	0.909, 0.885	1.266	0.26
SH2B3/rs739496	A	0.712, 0.711	0.0	0.98
ATXN2/rs10849949	G	0.308, 0.289	0.358	0.54
ATXN2/rs2073950	A	0.288, 0.228	3.169	0.057
ATXN2/rs2301621	G	0.597, 0.578	0.295	0.58



Haplotype	Freq. in population	Case, control ratios	Chi square	p-value	Permutation p-value
GAAGG	0.528	0.531, 0.524	0.041	0.939	1.00
GAAGA	0.158	0.138, 0.187	3.532	0.060	0.319
GGGAA	0.125	0.140, 0.103	2.438	0.118	0.512
AGGAA	0.085	0.067, 0.113	5.22	0.022	0.128
GGGGG	0.038	0.034, 0.043	0.456	0.499	1.00
GAGAA	0.02	0.032, 0.000	10.594	0.0011	0.007

Figure 5.7. Haplotype analysis from the combined-data. The haplotype, GGGAA, representing the 15-SNP haplotype was not significant; a new ALS-associated haplotype, GAGAA, was identified.

Haplotype analysis of the combined-data revealed that the GGGAA haplotype was not significantly associated with ALS (p= 0.118) (Figure 5.7). Compared to haplotype results of the five SNPs using ALS-GWAS data, the frequency of the haplotype in cases was decreased and controls were approximately the same (Figure 5.6, Figure 5.7). Remarkably, the three ALS cases with both expanded PolyQ repeats and CAA repeat interruptions were found to harbor the GGGAA haplotype heterozygously (Table 5.3).

Interestingly, a new haplotype, GAGAA, was found in 15 ALS cases with normal length ATXN2 PolyQ repeat and none of the controls. After the permutation test, the GAGAA haplotype remained significant (p=0.007).

5.3. CNV Detection Analyses

In the last part of this thesis, the CNVs that were previously found to be associated with ALS were validated in 48 individuals (24 cases and 24 controls). The CNVs which were previously found to be ALS-associated were identified using PennCNV tool, which detects CNVs via evaluating the SNP intensities of the ALS-GWAS data. On the other hand, for validation studies, CNVs were detected after the real-time PCR amplification experiments, using CopyCaller Software.

5.3.1. CNV Detection in the MAP4K3 Gene

In a previous study performed in our lab (Uyan *et al.*, 2013), PennCNV analysis has shown a novel heterozygous duplication at the coding region of the MAP4K3 gene to confer risk for ALS (in 14/117 cases and 2/109 controls). However, validation studies of all the individuals, predicted to have this duplication, plus seven additional control samples, predicted to be normal (two copies), failed to confirm the PennCNV results. According to the CopyCaller analysis, all of the 14 ALS cases and eight of the nine controls were found to have two copies of the CNV at the MAP4K3 gene and one control had heterozygous deletion of the CNV (Figure 5.8). Comparison of the results from both the PennCNV and CopyCaller analyses, is represented in Figure 5.8.

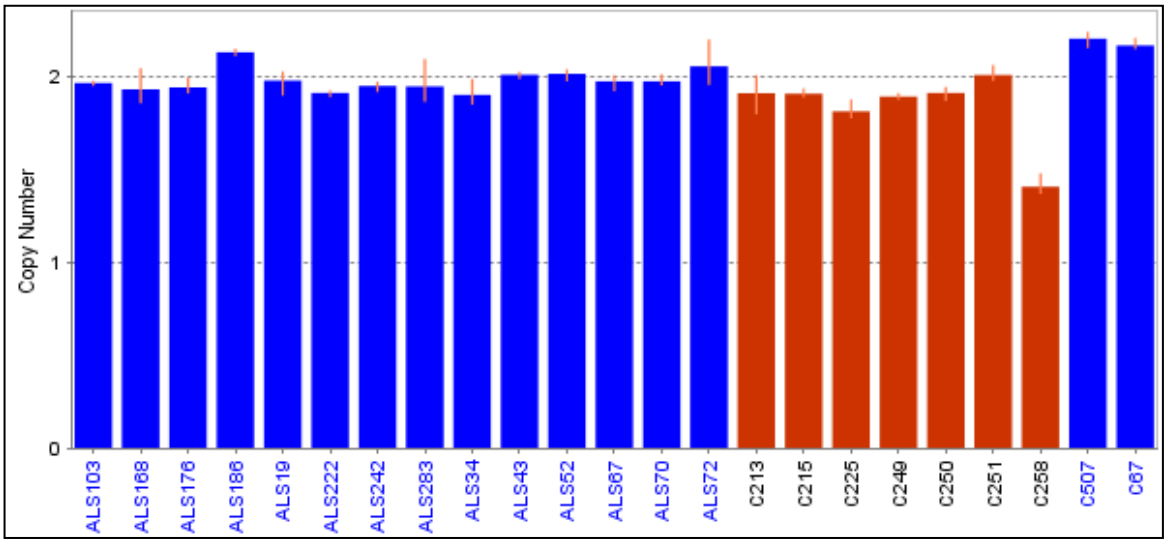


Figure 5.8. CopyCaller CNV detection of the MAP4K3-CNV. All individuals except C258 were found to have two copies of the CNV. Only C258 had one copy of the gene. Individuals shown in blue bars were predicted to have three copies, and individuals in red were predicted to have two copies, according to the previous PennCNV results (Uyan *et al*., 2013).

5.3.2. CNV Detection in the HLA-B Gene

In Uyan's study, the PennCNV tool has associated a total deletion at the coding region of the HLA-B gene to ALS. However, CopyCaller analysis confuted this association by showing that the individuals predicted to have the deletion, have actually two copies. The comparison of the results of PennCNV and the CopyCaller analyses of HLA-B is shown in Figure 5.9.

5.3.3. CNV Detection in the EPHA3 Gene

In Uyan's study, PennCNV has shown a heterozygous deletion at the EPHA3 locus as a protective factor against ALS. Remarkably, CopyCaller Software confirmed the EPHA3-CNV results of the individuals detected by PennCNV. However, two controls, that had been shown to have one copy each, were found to have a total deletion (Figure 5.10).

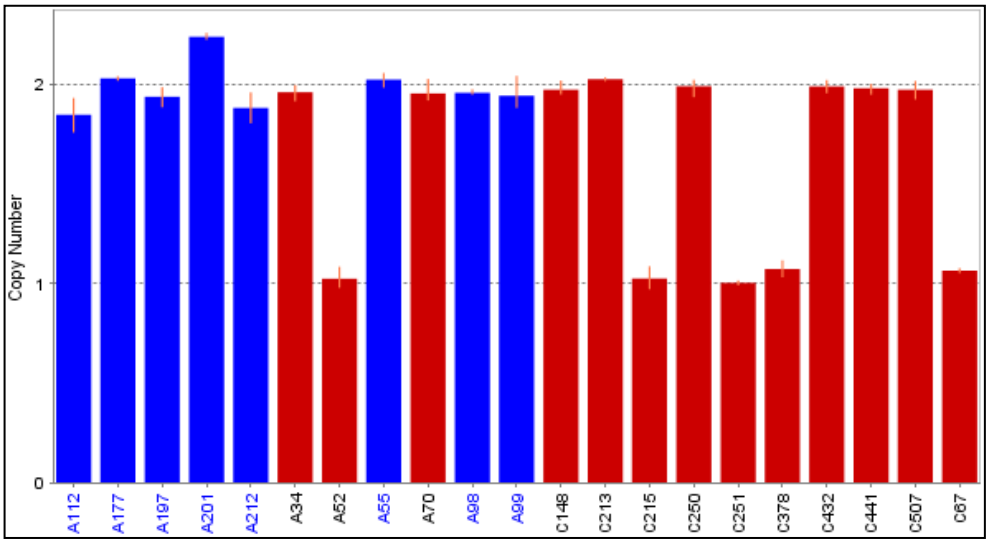


Figure 5.9. CopyCaller CNV detection of the HLA-B-CNV. A total of 16 individuals (10 cases and eight controls) have two copies and the remaining has one copy. According to the previous PennCNV analysis, individuals shown with blue bars were predicted to have a total deletion, and red ones were shown to have two copies of the CNV.

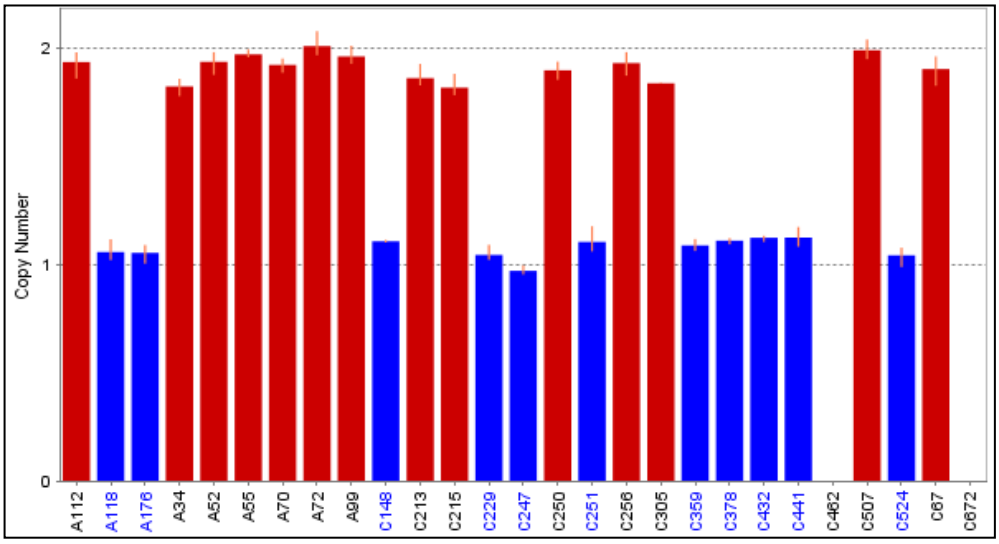


Figure 5.10. CopyCaller CNV detection of the EPHA3-CNV. CopyCaller detected 14 individuals (red bars) with two copies, 11 individuals (blue bars) with one copy each and two individuals with one copy of the CNV. According to the PennCNV results, individuals shown with blue bars and two controls (C462 and C672) were shown to have one copy of the CNV.

6. DISCUSSION

ALS research aiming to unravel the etiology of the disease has acquired exceptional growth in the recent 20 years. In this period, many different approaches, such as genetic, functional and clinical studies have been conducted. One of the pioneering disciplines which dissects the disease and provides invaluable information to the ALS field are genetic studies. Conventional genetic methods along with the recent developments in the genomic field, such as GWAS and next-generation sequencing technologies, led to the recent identification of more than 25 genes and several loci. However, the information obtained so far, describes only 60% of fALS and a very small portion of sALS world-wide (Robberecht and Philips, 2013). In this respect, identification of new genes and risk and/or protective factors of ALS are very crucial for the development of effective treatments and innovative therapies.

SOD1-based ALS has been the main focus in ALS field for 20 years, however, the discovery of TDP-43 pointed to a new mechanism, based on RNA biology, into disease pathogenesis and soon after, led to changing paradigms in ALS research. TDP-43 accounts for around 1-5% of fALS and 0-2% of sALS cases. The mislocalization of the protein is detected in ubiquitinated cytoplasmic inclusions of neurons of almost all nonSOD1-ALS cases, indicating a remarkable role in disease pathology. TDP-43 is involved in several important RNA pathways, and although the exact mechanism (gain- or loss-of-function) is unknown, mutations in this gene result in abnormalities in RNA processing. Following this, the identification of TDP-43 as an ALS gene impelled researchers to investigate other RNA/DNA binding and/or processing proteins and by this means, other important ALS genes, such as FUS and ATXN2 were identified (Al-Chalabi *et al.*, 2012).

It is still uncertain how the potential gain-of-function activity of TDP-43 may contribute to ALS pathogenesis. In 2010, a comprehensive study investigating possible modifiers of TDP-43 toxicity identified 27 genes that enhance the TDP-43-dependent toxicity. Among those genes, ATXN2, the causative gene for SCA2, was further shown to increase the TDP-43 toxicity *in vitro* and in flies (Elden *et al.*, 2010). Subsequently, genetic screening in a large ALS cohort and follow-up studies in many populations,

revealed that ATXN2 intermediate length PolyQ is associated with ALS (Chen *et al.*, 2011, Corrado *et al.*, 2011, Daoud *et al.*, 2011, Elden *et al.*, 2010, Gispert *et al.*, 2012, Lee *et al.*, 2011a, Lee *et al.*, 2011b, Van Damme *et al.*, 2011).

The developments in the genomic field have dramatically accelerated especially studies investigating sALS genetics. GWAS is one of the powerful genomic approaches used to discover the candidate loci involved in ALS. Up to now, more than 10 large-scale ALS-GWAS, leading to the identification of several disease-associated loci have been published (http://www.genome.gov). Although next-generation sequencing technologies have widely replaced GWAS in recent years, the large amount of information obtained from GWAS is still useful for investigating the disease association of rarer variants, such as CNVs.

In the framework of this thesis, the contribution of the ATXN2 locus to ALS in a Turkish cohort was interrogated, and other candidate disease-associated CNVs in the particular regions of the genome were analyzed.

6.1. ATXN2 PolyQ Expansion Analyses

In this study, among 236 ALS cases, four were found to carry >30 PolyQ (1.69%), however, none of the controls had larger repeat sizes than 29. This difference was statistically significant. In studies investigating ATXN2 PolyQ repeats in different populations, the ALS-specific cutoff of ATXN2 PolyQ size was found to be variable (Table 1.2). Although Turkey is a large country with a great ethnic heterogeneity, the cutoff PolyQ size of the Turkish population was expected to be more similar to the large and mixed European cohort rather than the Chinese, French-Canadian and Italian cohorts in whom similar analyses were performed. As anticipated, in the Turkish cohort, the ALS-specific cutoff was shown to be >30, in accordance with the study investigating a European cohort, comprising individuals with different genetic backgrounds.

This study also included Turkish fALS samples, with a defined ALS mutation, in order to investigate the oligogenic status of ALS and the possible effects of ATXN2 PolyQ repeat expansion on disease phenotype, e.g. ages at disease onset; however, none of the

ALS cases with defined ALS mutations was found to carry the expansion. Interestingly, one juvenile fALS case was identified (without any known causative gene defect) who had an expansion. Although follow-up studies confuted the results of the Elden study, which reported a decreased age at onset in the ALS cases with PolyQ expansion, it is still important to analyze clinical features of the ALS cases with PolyQ expansion in other and larger cohorts (Daoud *et al.*, 2011, Elden *et al.*, 2010, Gellera *et al.*, 2012). Unfortunately, the number of our ALS cases with a PolyQ expansion was too limited, to systematically investigate the genotype-phenotype correlations of the PolyQ expansion.

6.1.1. Possible Roles of ATXN2 in ALS Pathology

Neurodegenerative diseases share common features. When the clinical features of SCA2 and ALS are compared, the association between ATXN2 and ALS is not very surprising, since the degeneration of motor neurons, seen in ALS, was also shown in SCA2 (Nanetti *et al.*, 2009). This clinical overlap indicates shared molecular mechanisms between these two diseases; however, the exact contribution of the PolyQ expansion of ATXN2 on ALS pathology has not been thoroughly understood yet.

A few studies, investigating ATXN2 and TDP-43 interactions, highlighted several partly overlapping mechanisms. The transient interaction of these proteins through RNA binding and their presence in ubiquitinated cytoplasmic inclusions, together pointed to abnormalities in RNA processing, which is one of the commonly proposed ALS mechanisms of pathogenesis (Elden *et al.*, 2010). As previously reported in several studies, phosphorylation of TDP-43 increases its aggregation propensity and toxicity. Additionally, the C-terminal cleavage of TDP-43 is known to increase the mislocalization, phosphorylation and ubiquitination of this protein which together result in enhanced TDP-43 aggregation in both neuronal and glial cells. The study, reporting ATXN2-dependent biochemical modifications of TDP-43 revealed that the PolyQ expansion both enhances the stress-induced C-terminal cleavage of TDP-43 and results in caspase activation which leads to cell death in patient-derived lymphoblast cells (Hart and Gitler, 2012). Moreover, one study investigating the interaction of ATXN2 and FUS showed that the co-expression of PolyQ-expanded ATXN2 with several ALS-causing mutants of FUS, results in increased ER stress and Golgi fragmentation, which are frequent in ALS and trigger

apoptosis (Farg *et al.*, 2013). The current knowledge highlights an important role of ATXN2 in ALS pathology, however it remains to be dissected.

Recently, a study investigating the sequence composition of ALS cases with PolyQ expansion reported that all cases had one to three CAA interruptions (Yu *et al.*, 2011). Sequencing of the expanded ATXN2 PolyQ samples in the Turkish ALS cohort revealed that one out of four cases with the PolyQ expansion harbored no CAA interruption whereas the remaining three cases had only one CAA interruption. The interruption of PolyQ-encoding CAG repeat sequence in the ATXN2 gene by a CAA is known to increase the length stability (Lastres-Becker *et al.*, 2008). It was shown that this interruption changes the secondary structure of the RNA product rather than the protein sequence since both codons encode for glutamine (Sobczak and Krzyzosiak, 2005). Thus, it can be speculated that RNA-dependent ATXN2 toxicity may contribute to ALS in addition to protein-based effects.

6.2. SNP and Haplotype Association Analyses in the ATXN2 Locus

In order to further investigate the genetic association of the ATXN2 locus with ALS, SNP and haplotype analyses were performed. For this purpose, the previously completed SNP data of the ALS-GWAS was used (Uyan, M.Sc. Thesis, Bogazici University, 2012). This study contains 117 out of 236 ALS samples included into the present study. There was no significant ALS-associated SNP at this locus, which may be due to the small sample size or heterogeneity of the Turkish population. However, a large 15-SNP-haplotype increasing ALS risk was shown using haplotype analysis. Interestingly, three of the 15 SNPs were found to be located in the neighboring gene of ATXN2, which is SH2B3.

Subsequently, additional Turkish ALS cases and controls were genotyped by the TaqMan SNP assay for the five SNPs which represent the 15-SNP-haplotype block. Although SNP genotyping with TaqMan probes is not comparable to GWAS, it is a very accurate and fast method to genotype the SNPs of interest. In the combined-data obtained, a new haplotype block, GAGAA, was found to be associated with ALS. On the other hand, the GGGAA haplotype, previously shown to be significant in the ALS-GWAS, lost its

significance when the cohort size was increased from 117 to 236. Furthermore, the cases with CAA interruption were found to be arising from a similar genetic origin, because they all harbored the GGGAA haplotype. This was in accordance with the recent study investigating the CAA interruption in ALS cases (Yu *et al.*, 2011).

6.2.1. SH2B3 in ALS and ATXN2-SH2B3 Interactions

The identification of a genetic association between the 5-SNP-haplotype, which includes SNPs of the SH2B3 gene and ALS, suggests a possible role of SH2B3 in ALS pathology. SH2B3 is a member of an adaptor protein family, SH2B (1-3), which contains pleckstrin motifs, proline-rich regions and Src Homology 2 (SH2) domains. These domains provide the proteins to interact with the members of receptor tyrosine kinase pathway (RTK), which regulates very crucial cellular events, such as translation, proliferation and survival (Devalliere and Charreau, 2011). In addition, rs3184504, located in the exon 3 of SH2B3, which is in the ALS-associated 15-SNP-haplotype block, was previously found to be associated with multiple sclerosis (Alcina *et al.*, 2010). These molecular and genetic evidences support that, SH2B3 may have a role in ALS pathology.

Investigating the molecular interactions between SH2B3 and ataxin 2 may also provide valuable insights to understand the contribution of these proteins to ALS. A recent study reported that ataxin 2 with normal length PolyQ binds to several Src Homology 3 (SH3) domains *in vitro*, which can also form a regulatory apparatus with SH2 domain. Further *in vivo* evidences showed that ataxin 2 with long length PolyQ (SCA2 range) result in decreased protein levels of one important member of the RTK pathway, the endogenous growth factor receptor-bound protein 2, which contains both SH2 and SH3 domains (Drost *et al.*, 2013). These results indicate that ATXN2 PolyQ expansion diminishes the compensatory mechanisms following cellular damage and degeneration. However, in the case of ALS, it should be further investigated whether the intermediate length PolyQ expansion modulates RTK activity, if yes, to what extent, or not.

6.3. Validation of the Candidate ALS-Associated CNVs

The advancements in genomic technologies in recent years have tremendously accelerated gene discovery studies. One of these technologies, which provides important information to understand complex traits as also emphasized frequently in this thesis is GWAS. Many genes and trait-associated loci were discovered by using this method. However, among more than 10 GWAS performed in ALS, only the disease association of chromosome 9p21 locus was discovered, whereas most of the other candidates failed to be replicated in other cohorts (Laaksovirta *et al.*, 2010, Shatunov *et al.*, 2010, van Es *et al.*, 2009).

The "common disease/common variant" hypothesis on which GWAS relies, is losing its popularity due to the failure in replication studies and the mild effects of common variants (SNPs) indicating that rare variants (possibly even individual-specific) may explain the most important part of complex diseases. Meanwhile, CNVs which are rarer than SNPs but may have a higher penetrance were found to be associated with several complex diseases, such as autism and epilepsy (de Kovel *et al.*, 2010, Marshall *et al.*, 2008).

Few studies investigating genome-wide CNVs in ALS cases found several candidate loci so far. All these studies benefited from the SNP intensity from GWAS to genotype CNVs, which is an indirect way, thus leading to many false positive results. Recently, we investigated genome-wide CNVs in the Turkish ALS cohort (116 cases and 109 controls) and reported CNVs in three gene regions using the PennCNV tool. Among them, a heterozygous deletion in MAP4K3 and a homozygous deletion in HLA-B genes were found to be risk factors for ALS; on the other hand a heterozygous deletion in EPHA3 was shown to be protective (Uyan *et al*., 2013). However, considering the small size and the large heterogeneity of the cohort, we concluded that it would be more informative to genotype the individuals with an accurate genotyping method to eliminate false positives. For validation, the samples in which CNVs were identified by PennCNV were further genotyped by TaqMan CNV Assay. CNVs in MAP4K3 and HLA-B were excluded as false positives and only the CNV in EPHA3 could be validated.

Ephrin (Eph) receptors are a subfamily of RTK and consist of 14 members. The known functions of Ephs are regulation of cell-cell interaction and neuronal cell migration, axon guidance and vasculogenesis during embryonic development (Lisle *et al.*, 2013). In 2012, EPHA4, one member of Eph A subfamily, was shown to modify ALS. Both loss-of-function and knock-down of this gene *in vivo* was shown to rescue disease phenotype and increase survival (Van Hoecke *et al.*, 2012). In the framework of this thesis, it was shown that a heterozygous deletion in the EPHA3 gene was significantly higher in control individuals than in ALS cases. To further confirm these results and draw a firmer conclusion on the role of this CNV, it should be genotyped in larger cohorts and the molecular consequences of the deletion in EPHA3 should be investigated. Our preliminary data suggest that, like EPHA4, EPHA3 may also modulate disease pathology and that ephs may be possible targets for therapeutic intervention.

6.4. Future Goals

The first aim of this thesis was to investigate the genetic contribution of the ATXN2 locus on ALS. Our results revealed the presence of the intermediate length ATXN2 PolyQ expansions in Turkish ALS patients and additionally an ALS-associated haplotype including SNPs, located in the neighbouring SH2B3 gene. The second aim of this thesis was to validate candidate CNV loci which were shown to be associated with ALS in the framework of a previous thesis in this lab. One CNV in the EPHA3 locus which may be a protective factor for ALS was validated (Uyan, M.Sc. Thesis, Bogazici University, 2012). In the light of these observations, several future experiments should be designed to:

- evaluate the ATXN2 ALS association: The number of ALS cases and controls included into both ATXN2 PolyQ size screening and SNP and haplotype analyses should be increased.

- understand genotype-phenotype correlations: ALS cases with PolyQ expansions should be investigated in a lager cohort.

- question the segregation of the SH2B3 gene mutations with the disease: ALS cases should be sequenced.

- analyze the possible contribution of ataxin 2 and SH2B3 proteins to ALS: molecular interactions between these two proteins should be investigated *in vitro* and *in vivo*.

- gain insights into the association of EPHA3 locus with ALS: a larger case and control cohort, including the individuals genotyped by GWAS, should be genotyped.

EPHA3 is one of the members of receptor tyrosine kinase (RTK) family and ATXN2 and SH2B3 encode proteins that both interact with growth receptor tyrosine kinases. Our novel observation suggests that proteins of the RTK pathway and their interaction partners may be promising targets for therapeutic interventions.

REFERENCES

Abhinav, K., B. Stanton, C. Johnston, J. Hardstaff, R.W. Orrell, R. Howard, J. Clarke, M. Sakel, M.A. Ampong, C.E. Shaw, P.N. Leigh and A. Al-Chalabi, 2007, "Amyotrophic Lateral Sclerosis in South-East England: A Population-Based Study. The South-East England Register for Amyotrophic Lateral Sclerosis (Seals Registry)", *Neuroepidemiology*, Vol. 29, No. 1-2, pp. 44-48.

Ahmeti, K.B., S. Ajroud-Driss, A. Al-Chalabi, P.M. Andersen, J. Armstrong, A. Birve, H.M. Blauw, R.H. Brown, L. Bruijn, W. Chen, A. Chio, M.C. Comeau, S. Cronin, F.P. Diekstra, A. Soraya Gkazi, J.D. Glass, J.D. Grab, E.J. Groen, J.L. Haines, O. Hardiman, S. Heller, J. Huang, W.Y. Hung, J.M. Jaworski, A. Jones, H. Khan, J.E. Landers, C.D. Langefeld, P.N. Leigh, M.C. Marion, R.L. McLaughlin, V. Meininger, J. Melki, J.W. Miller, G. Mora, M.A. Pericak-Vance, E. Rampersaud, W. Robberecht, L.P. Russell, F. Salachas, C.G. Saris, A. Shatunov, C.E. Shaw, N. Siddique, T. Siddique, B.N. Smith, R. Sufit, S. Topp, B.J. Traynor, C. Vance, P. van Damme, L.H. van den Berg, M.A. van Es, P.W. van Vught, J.H. Veldink, Y. Yang and J.G. Zheng, 2013, "Age of Onset of Amyotrophic Lateral Sclerosis Is Modulated by a Locus on 1p34.1", *Neurobiology of Aging*, Vol. 34, No. 1, pp. 357 e357-319.

Al-Chalabi, A., A. Jones, C. Troakes, A. King, S. Al-Sarraj and L.H. van den Berg, 2012, "The Genetics and Neuropathology of Amyotrophic Lateral Sclerosis", *Acta Neuropathologica*, Vol. 124, No. 3, pp. 339-352.

Al-Saif, A., F. Al-Mohanna and S. Bohlega, 2011, "A Mutation in Sigma-1 Receptor Causes Juvenile Amyotrophic Lateral Sclerosis", *Annals of Neurology*, Vol. 70, No. 6, pp. 913-919.

Alcina, A., K. Vandenbroeck, D. Otaegui, A. Saiz, J.R. Gonzalez, O. Fernandez, M.L. Cavanillas, M.C. Cenit, R. Arroyo, I. Alloza, M. Garcia-Barcina, A. Antiguedad, L. Leyva, G. Izquierdo, M. Lucas, M. Fedetz, M.J. Pinto-Medel, J. Olascoaga, Y.

Blanco, M. Comabella, X. Montalban, E. Urcelay and F. Matesanz, 2010, "The Autoimmune Disease-Associated Kif5a, Cd226 and Sh2b3 Gene Variants Confer Susceptibility for Multiple Sclerosis", *Genes and Immunity*, Vol. 11, No. 5, pp. 439-445.

Andersen, P.M. and A. Al-Chalabi, 2011, "Clinical Genetics of Amyotrophic Lateral Sclerosis: What Do We Really Know?", *Nature Reviews Neurology*, Vol. 7, No. 11, pp. 603-615.

Bettens, K., K. Sleegers and C. Van Broeckhoven, 2013, "Genetic Insights in Alzheimer's Disease", *Lancet Neurology*, Vol. 12, No. 1, pp. 92-104.

Blauw, H.M., A. Al-Chalabi, P.M. Andersen, P.W. van Vught, F.P. Diekstra, M.A. van Es, C.G. Saris, E.J. Groen, W. van Rheenen, M. Koppers, R. Van't Slot, E. Strengman, K. Estrada, F. Rivadeneira, A. Hofman, A.G. Uitterlinden, L.A. Kiemeney, S.H. Vermeulen, A. Birve, S. Waibel, T. Meyer, S. Cronin, R.L. McLaughlin, O. Hardiman, P.C. Sapp, M.D. Tobin, L.V. Wain, B. Tomik, A. Slowik, R. Lemmens, D. Rujescu, C. Schulte, T. Gasser, R.H. Brown, Jr., J.E. Landers, W. Robberecht, A.C. Ludolph, R.A. Ophoff, J.H. Veldink and L.H. van den Berg, 2010, "A Large Genome Scan for Rare Cnvs in Amyotrophic Lateral Sclerosis", *Human Molecular Genetics*, Vol. 19, No. 20, pp. 4091-4099.

Blauw, H.M., C.P. Barnes, P.W. van Vught, W. van Rheenen, M. Verheul, E. Cuppen, J.H. Veldink and L.H. van den Berg, 2012, "Smn1 Gene Duplications Are Associated with Sporadic Als", *Neurology*, Vol. 78, No. 11, pp. 776-780.

Blauw, H.M., J.H. Veldink, M.A. van Es, P.W. van Vught, C.G. Saris, B. van der Zwaag, L. Franke, J.P. Burbach, J.H. Wokke, R.A. Ophoff and L.H. van den Berg, 2008, "Copy-Number Variation in Sporadic Amyotrophic Lateral Sclerosis: A Genome-Wide Screen", *Lancet Neurology*, Vol. 7, No. 4, pp. 319-326.

Bosco, D.A., G. Morfini, N.M. Karabacak, Y. Song, F. Gros-Louis, P. Pasinelli, H. Goolsby, B.A. Fontaine, N. Lemay, D. McKenna-Yasek, M.P. Frosch, J.N. Agar, J.P.

Julien, S.T. Brady and R.H. Brown, Jr., 2010, "Wild-Type and Mutant Sod1 Share an Aberrant Conformation and a Common Pathogenic Pathway in Als", *Nature Neuroscience*, Vol. 13, No. 11, pp. 1396-1403.

Bruijn, L.I., M.K. Houseweart, S. Kato, K.L. Anderson, S.D. Anderson, E. Ohama, A.G. Reaume, R.W. Scott and D.W. Cleveland, 1998, "Aggregation and Motor Neuron Toxicity of an Als-Linked Sod1 Mutant Independent from Wild-Type Sod1", *Science*, Vol. 281, No. 5384, pp. 1851-1854.

Buratti, E. and F.E. Baralle, 2009, "The Molecular Links between Tdp-43 Dysfunction and Neurodegeneration", *Advances in Genetics*, Vol. 66, pp. 1-34.

Bush, W.S. and J.H. Moore, 2012, "Chapter 11: Genome-Wide Association Studies", *PLoS Computational Biology*, Vol. 8, No. 12, pp. e1002822.

Chen, Y., R. Huang, Y. Yang, K. Chen, W. Song, P. Pan, J. Li and H.F. Shang, 2011, "Ataxin-2 Intermediate-Length Polyglutamine: A Possible Risk Factor for Chinese Patients with Amyotrophic Lateral Sclerosis", *Neurobiology of Aging*, Vol. 32, No. 10, pp. 1925 e1921-1925.

Chen, Y.Z., C.L. Bennett, H.M. Huynh, I.P. Blair, I. Puls, J. Irobi, I. Dierick, A. Abel, M.L. Kennerson, B.A. Rabin, G.A. Nicholson, M. Auer-Grumbach, K. Wagner, P. De Jonghe, J.W. Griffin, K.H. Fischbeck, V. Timmerman, D.R. Cornblath and P.F. Chance, 2004, "DNA/Rna Helicase Gene Mutations in a Form of Juvenile Amyotrophic Lateral Sclerosis (Als4)", *American Journal of Human Genetics*, Vol. 74, No. 6, pp 1128-1135.

Choudhry, S., M. Mukerji, A.K. Srivastava, S. Jain and S.K. Brahmachari, 2001, "Cag Repeat Instability at Sca2 Locus: Anchoring Caa Interruptions and Linked Single Nucleotide Polymorphisms", *Human Molecular Genetics*, Vol. 10, No. 21, pp. 2437-2446.

Chow, C.Y., J.E. Landers, S.K. Bergren, P.C. Sapp, A.E. Grant, J.M. Jones, L. Everett, G.M. Lenk, D.M. McKenna-Yasek, L.S. Weisman, D. Figlewicz, R.H. Brown and M.H. Meisler, 2009, "Deleterious Variants of Fig4, a Phosphoinositide Phosphatase, in Patients with Als", *American Journal of Human Genetics*, Vol. 84, No. 1, pp. 85-88.

Chowdhury, A., K.K. Raju, S. Kalurupalle and S. Tharun, 2012, "Both Sm-Domain and C-Terminal Extension of Lsm1 Are Important for the Rna-Binding Activity of the Lsm1-7-Pat1 Complex", *RNA*, Vol. 18, No. 5, pp. 936-944.

Cohen, T.J., V.M. Lee and J.Q. Trojanowski, 2011, "Tdp-43 Functions and Pathogenic Mechanisms Implicated in Tdp-43 Proteinopathies", *Trends in Molecular Medicine*, Vol. 17, No. 11, pp. 659-667.

Corcia, P., C. Ingre, H. Blasco, R. Press, J. Praline, C. Antar, C. Veyrat-Durebex, Y.O. Guettard, W. Camu, P.M. Andersen, P. Vourc'h and C.R. Andres, 2012, "Homozygous Smn2 Deletion Is a Protective Factor in the Swedish Als Population", *European Journal of Human Genetics*, Vol. 20, No. 5, pp. 588-591.

Corrado, L., L. Mazzini, G.D. Oggioni, B. Luciano, M. Godi, A. Brusco and S. D'Alfonso, 2011, "Atxn-2 Cag Repeat Expansions Are Interrupted in Als Patients", *Human Genetics*, Vol. 130, No. 4, pp. 575-580.

Cozzolino, M., M.G. Pesaresi, V. Gerbino, J. Grosskreutz and M.T. Carri, 2012, "Amyotrophic Lateral Sclerosis: New Insights into Underlying Molecular Mechanisms and Opportunities for Therapeutic Intervention", *Antioxidants & Redox Signaling*, Vol. 17, No. 9, pp. 1277-1330.

Cronin, S., S. Berger, J. Ding, J.C. Schymick, N. Washecka, D.G. Hernandez, M.J. Greenway, D.G. Bradley, B.J. Traynor and O. Hardiman, 2008, "A Genome-Wide Association Study of Sporadic Als in a Homogenous Irish Population", *Human Molecular Genetics*, Vol. 17, No. 5, pp. 768-774.

Daoud, H., V. Belzil, S. Martins, M. Sabbagh, P. Provencher, L. Lacomblez, V. Meininger, W. Camu, N. Dupre, P.A. Dion and G.A. Rouleau, 2011, "Association of Long Atxn2 Cag Repeat Sizes with Increased Risk of Amyotrophic Lateral Sclerosis", *Archives of Neurology*, Vol. 68, No. 6, pp. 739-742.

de Kovel, C.G., H. Trucks, I. Helbig, H.C. Mefford, C. Baker, C. Leu, C. Kluck, H. Muhle, S. von Spiczak, P. Ostertag, T. Obermeier, A.A. Kleefuss-Lie, K. Hallmann, M. Steffens, V. Gaus, K.M. Klein, H.M. Hamer, F. Rosenow, E.H. Brilstra, D.K. Trenite, M.E. Swinkels, Y.G. Weber, I. Unterberger, F. Zimprich, L. Urak, M. Feucht, K. Fuchs, R.S. Moller, H. Hjalgrim, P. De Jonghe, A. Suls, I.M. Ruckert, H.E. Wichmann, A. Franke, S. Schreiber, P. Nurnberg, C.E. Elger, H. Lerche, U. Stephani, B.P. Koeleman, D. Lindhout, E.E. Eichler and T. Sander, 2010, "Recurrent Microdeletions at 15q11.2 and 16p13.11 Predispose to Idiopathic Generalized Epilepsies", *Brain*, Vol. 133, No. Pt 1, pp. 23-32.

DeJesus-Hernandez, M., I.R. Mackenzie, B.F. Boeve, A.L. Boxer, M. Baker, N.J. Rutherford, A.M. Nicholson, N.A. Finch, H. Flynn, J. Adamson, N. Kouri, A. Wojtas, P. Sengdy, G.Y. Hsiung, A. Karydas, W.W. Seeley, K.A. Josephs, G. Coppola, D.H. Geschwind, Z.K. Wszolek, H. Feldman, D.S. Knopman, R.C. Petersen, B.L. Miller, D.W. Dickson, K.B. Boylan, N.R. Graff-Radford and R. Rademakers, 2011, "Expanded Ggggcc Hexanucleotide Repeat in Noncoding Region of C9orf72 Causes Chromosome 9p-Linked Ftd and Als", *Neuron*, Vol. 72, No. 2, pp. 245-256.

Deng, H.X., W. Chen, S.T. Hong, K.M. Boycott, G.H. Gorrie, N. Siddique, Y. Yang, F. Fecto, Y. Shi, H. Zhai, H. Jiang, M. Hirano, E. Rampersaud, G.H. Jansen, S. Donkervoort, E.H. Bigio, B.R. Brooks, K. Ajroud, R.L. Sufit, J.L. Haines, E. Mugnaini, M.A. Pericak-Vance and T. Siddique, 2011, "Mutations in Ubqln2 Cause Dominant X-Linked Juvenile and Adult-Onset Als and Als/Dementia", *Nature*, Vol. 477, No. 7363, pp. 211-215.

Devalliere, J. and B. Charreau, 2011, "The Adaptor Lnk (Sh2b3): An Emerging Regulator in Vascular Cells and a Link between Immune and Inflammatory Signaling", *Biochemical Pharmacology*, Vol. 82, No. 10, pp. 1391-1402.

Dormann, D., R. Rodde, D. Edbauer, E. Bentmann, I. Fischer, A. Hruscha, M.E. Than, I.R. Mackenzie, A. Capell, B. Schmid, M. Neumann and C. Haass, 2010, "Als-Associated Fused in Sarcoma (Fus) Mutations Disrupt Transportin-Mediated Nuclear Import", *The EMBO Journal*, Vol. 29, No. 16, pp. 2841-2857.

Drost, J., D. Nonis, F. Eich, O. Leske, E. Damrath, E.R. Brunt, I. Lastres-Becker, R. Heumann, J. Nowock and G. Auburger, 2013, "Ataxin-2 Modulates the Levels of Grb2 and Src but Not Ras Signaling", *Journal of Molecular Neuroscience*, pp. 659-667

Elden, A.C., H.J. Kim, M.P. Hart, A.S. Chen-Plotkin, B.S. Johnson, X. Fang, M. Armakola, F. Geser, R. Greene, M.M. Lu, A. Padmanabhan, D. Clay-Falcone, L. McCluskey, L. Elman, D. Juhr, P.J. Gruber, U. Rub, G. Auburger, J.Q. Trojanowski, V.M. Lee, V.M. Van Deerlin, N.M. Bonini and A.D. Gitler, 2010, "Ataxin-2 Intermediate-Length Polyglutamine Expansions Are Associated with Increased Risk for Als", *Nature*, Vol. 466, No. 7310, pp. 1069-1075.

Farg, M.A., K.Y. Soo, S.T. Warraich, V. Sundaramoorthy, I.P. Blair and J.D. Atkin, 2013, "Ataxin-2 Interacts with Fus and Intermediate-Length Polyglutamine Expansions Enhance Fus-Related Pathology in Amyotrophic Lateral Sclerosis", *Human Molecular Genetics*, Vol. 22, No. 4, pp. 717-728.

Ferraiuolo, L., J. Kirby, A.J. Grierson, M. Sendtner and P.J. Shaw, 2011, "Molecular Pathways of Motor Neuron Injury in Amyotrophic Lateral Sclerosis", *Nature Reviews Neurology*, Vol. 7, No. 11, pp. 616-630.

Fischer, L.R. and J.D. Glass, 2007, "Axonal Degeneration in Motor Neuron Disease", *Neurodegenerative Diseases*, Vol. 4, No. 6, pp. 431-442.

Forsberg, K., P.A. Jonsson, P.M. Andersen, D. Bergemalm, K.S. Graffmo, M. Hultdin, J. Jacobsson, R. Rosquist, S.L. Marklund and T. Brannstrom, 2010, "Novel Antibodies Reveal Inclusions Containing Non-Native Sod1 in Sporadic Als Patients", *PLoS One*, Vol. 5, No. 7, pp. e11552.

Gandhi, S. and N.W. Wood, 2010, "Genome-Wide Association Studies: The Key to Unlocking Neurodegeneration?", *Nature Neuroscience*, Vol. 13, No. 7, pp. 789-794.

Gellera, C., N. Ticozzi, V. Pensato, L. Nanetti, A. Castucci, B. Castellotti, G. Lauria, F. Taroni, V. Silani and C. Mariotti, 2012, "Ataxin2 Cag-Repeat Length in Italian Patients with Amyotrophic Lateral Sclerosis: Risk Factor or Variant Phenotype? Implication for Genetic Testing and Counseling", *Neurobiology of Aging*, Vol. 33, No. 8, pp. 1847 e1815-1821.

Gispert, S., A. Kurz, S. Waibel, P. Bauer, I. Liepelt, C. Geisen, A.D. Gitler, T. Becker, M. Weber, D. Berg, P.M. Andersen, R. Kruger, O. Riess, A.C. Ludolph and G. Auburger, 2012, "The Modulation of Amyotrophic Lateral Sclerosis Risk by Ataxin-2 Intermediate Polyglutamine Expansions Is a Specific Effect", *Neurobiology of Disease*, Vol. 45, No. 1, pp. 356-361.

Gitcho, M.A., R.H. Baloh, S. Chakraverty, K. Mayo, J.B. Norton, D. Levitch, K.J. Hatanpaa, C.L. White, 3rd, E.H. Bigio, R. Caselli, M. Baker, M.T. Al-Lozi, J.C. Morris, A. Pestronk, R. Rademakers, A.M. Goate and N.J. Cairns, 2008, "Tdp-43 A315t Mutation in Familial Motor Neuron Disease", *Annals of Neurology*, Vol. 63, No. 4, pp. 535-538.

Greenway, M.J., P.M. Andersen, C. Russ, S. Ennis, S. Cashman, C. Donaghy, V. Patterson, R. Swingler, D. Kieran, J. Prehn, K.E. Morrison, A. Green, K.R. Acharya, R.H. Brown, Jr. and O. Hardiman, 2006, "Ang Mutations Segregate with Familial and 'Sporadic' Amyotrophic Lateral Sclerosis", *Nature Genetics*, Vol. 38, No. 4, pp. 411-413.

Hadano, S., C.K. Hand, H. Osuga, Y. Yanagisawa, A. Otomo, R.S. Devon, N. Miyamoto, J. Showguchi-Miyata, Y. Okada, R. Singaraja, D.A. Figlewicz, T. Kwiatkowski, B.A. Hosler, T. Sagie, J. Skaug, J. Nasir, R.H. Brown, Jr., S.W. Scherer, G.A. Rouleau, M.R. Hayden and J.E. Ikeda, 2001, "A Gene Encoding a Putative Gtpase Regulator Is Mutated in Familial Amyotrophic Lateral Sclerosis 2", *Nature Genetics*, Vol. 29, No. 2, pp. 166-173.

Hand, C.K., J. Khoris, F. Salachas, F. Gros-Louis, A.A. Lopes, V. Mayeux-Portas, C.G. Brewer, R.H. Brown, Jr., V. Meininger, W. Camu and G.A. Rouleau, 2002, "A Novel Locus for Familial Amyotrophic Lateral Sclerosis, on Chromosome 18q", *American Journal of Human Genetics*, Vol. 70, No. 1, pp. 251-256.

Hart, M.P. and A.D. Gitler, 2012, "Als-Associated Ataxin 2 Polyq Expansions Enhance Stress-Induced Caspase 3 Activation and Increase Tdp-43 Pathological Modifications", *The Journal of Neuroscience*, Vol. 32, No. 27, pp. 9133-9142.

Hebert, L.E., P.A. Scherr, J.L. Bienias, D.A. Bennett and D.A. Evans, 2003, "Alzheimer Disease in the Us Population: Prevalence Estimates Using the 2000 Census", *Archives of Neurology*, Vol. 60, No. 8, pp. 1119-1122.

Ikegawa, S., 2012, "A Short History of the Genome-Wide Association Study: Where We Were and Where We Are Going", *Genomics & Informatics*, Vol. 10, No. 4, pp. 220-225.

Johnson, J.O., J. Mandrioli, M. Benatar, Y. Abramzon, V.M. Van Deerlin, J.Q. Trojanowski, J.R. Gibbs, M. Brunetti, S. Gronka, J. Wuu, J. Ding, L. McCluskey, M. Martinez-Lage, D. Falcone, D.G. Hernandez, S. Arepalli, S. Chong, J.C. Schymick, J. Rothstein, F. Landi, Y.D. Wang, A. Calvo, G. Mora, M. Sabatelli, M.R. Monsurro, S. Battistini, F. Salvi, R. Spataro, P. Sola, G. Borghero, G. Galassi, S.W. Scholz, J.P. Taylor, G. Restagno, A. Chio and B.J. Traynor, 2010, "Exome Sequencing Reveals Vcp Mutations as a Cause of Familial Als", *Neuron*, Vol. 68, No. 5, pp. 857-864.

Kiernan, M.C., S. Vucic, B.C. Cheah, M.R. Turner, A. Eisen, O. Hardiman, J.R. Burrell and M.C. Zoing, 2011, "Amyotrophic Lateral Sclerosis", *Lancet*, Vol. 377, No. 9769, pp. 942-955.

Kwiatkowski, T.J., Jr., D.A. Bosco, A.L. Leclerc, E. Tamrazian, C.R. Vanderburg, C. Russ, A. Davis, J. Gilchrist, E.J. Kasarskis, T. Munsat, P. Valdmanis, G.A. Rouleau, B.A. Hosler, P. Cortelli, P.J. de Jong, Y. Yoshinaga, J.L. Haines, M.A. Pericak-Vance, J. Yan, N. Ticozzi, T. Siddique, D. McKenna-Yasek, P.C. Sapp, H.R. Horvitz, J.E. Landers and R.H. Brown, Jr., 2009, "Mutations in the Fus/Tls Gene on Chromosome 16 Cause Familial Amyotrophic Lateral Sclerosis", *Science*, Vol. 323, No. 5918, pp. 1205-1208.

Laaksovirta, H., T. Peuralinna, J.C. Schymick, S.W. Scholz, S.L. Lai, L. Myllykangas, R. Sulkava, L. Jansson, D.G. Hernandez, J.R. Gibbs, M.A. Nalls, D. Heckerman, P.J. Tienari and B.J. Traynor, 2010, "Chromosome 9p21 in Amyotrophic Lateral Sclerosis in Finland: A Genome-Wide Association Study", *Lancet Neurology*, Vol. 9, No. 10, pp. 978-985.

Lagier-Tourenne, C. and D.W. Cleveland, 2009, "Rethinking Als: The Fus About Tdp-43", *Cell*, Vol. 136, No. 6, pp. 1001-1004.

Lagier-Tourenne, C. and D.W. Cleveland, 2010, "Neurodegeneration: An Expansion in Als Genetics", *Nature*, Vol. 466, No. 7310, pp. 1052-1053.

Lagier-Tourenne, C., M. Polymenidou, K.R. Hutt, A.Q. Vu, M. Baughn, S.C. Huelga, K.M. Clutario, S.C. Ling, T.Y. Liang, C. Mazur, E. Wancewicz, A.S. Kim, A. Watt, S. Freier, G.G. Hicks, J.P. Donohue, L. Shiue, C.F. Bennett, J. Ravits, D.W. Cleveland and G.W. Yeo, 2012, "Divergent Roles of Als-Linked Proteins Fus/Tls and Tdp-43 Intersect in Processing Long Pre-Mrnas", *Nature Neuroscience*, Vol. 15, No. 11, pp. 1488-1497.

Lahut, S., O. Omur, O. Uyan, Z.S. Agim, A. Ozoguz, Y. Parman, F. Deymeer, P. Oflazer, F. Koc, H. Ozcelik, G. Auburger and A.N. Basak, 2012, "Atxn2 and Its

Neighbouring Gene Sh2b3 Are Associated with Increased Als Risk in the Turkish Population", *PLoS One*, Vol. 7, No. 8, pp. e42956.

Lam, D., S. Shah, I.P. de Castro, S.H. Loh and L.M. Martins, 2010, "Drosophila Happyhour Modulates Jnk-Dependent Apoptosis", *Cell Death & Disease*, Vol. 1, pp. e66.

Lambert, J.C., S. Heath, G. Even, D. Campion, K. Sleegers, M. Hiltunen, O. Combarros, D. Zelenika, M.J. Bullido, B. Tavernier, L. Letenneur, K. Bettens, C. Berr, F. Pasquier, N. Fievet, P. Barberger-Gateau, S. Engelborghs, P. De Deyn, I. Mateo, A. Franck, S. Helisalmi, E. Porcellini, O. Hanon, M.M. de Pancorbo, C. Lendon, C. Dufouil, C. Jaillard, T. Leveillard, V. Alvarez, P. Bosco, M. Mancuso, F. Panza, B. Nacmias, P. Bossu, P. Piccardi, G. Annoni, D. Seripa, D. Galimberti, D. Hannequin, F. Licastro, H. Soininen, K. Ritchie, H. Blanche, J.F. Dartigues, C. Tzourio, I. Gut, C. Van Broeckhoven, A. Alperovitch, M. Lathrop and P. Amouyel, 2009, "Genome-Wide Association Study Identifies Variants at Clu and Cr1 Associated with Alzheimer's Disease", *Nature Genetics*, Vol. 41, No. 10, pp. 1094-1099.

Lastres-Becker, I., U. Rub and G. Auburger, 2008, "Spinocerebellar Ataxia 2 (Sca2)", *Cerebellum*, Vol. 7, No. 2, pp. 115-124.

Lee, T., Y.R. Li, A. Chesi, M.P. Hart, D. Ramos, N. Jethava, D. Hosangadi, J. Epstein, B. Hodges, N.M. Bonini and A.D. Gitler, 2011a, "Evaluating the Prevalence of Polyglutamine Repeat Expansions in Amyotrophic Lateral Sclerosis", *Neurology*, Vol. 76, No. 24, pp. 2062-2065.

Lee, T., Y.R. Li, C. Ingre, M. Weber, T. Grehl, O. Gredal, M. de Carvalho, T. Meyer, O.B. Tysnes, G. Auburger, S. Gispert, N.M. Bonini, P.M. Andersen and A.D. Gitler, 2011b, "Ataxin-2 Intermediate-Length Polyglutamine Expansions in European Als Patients", *Human Molecular Genetics*, Vol. 20, No. 9, pp. 1697-1700.

Lisle, J.E., I. Mertens-Walker, R. Rutkowski, A.C. Herington and S.A. Stephenson, 2013, "Eph Receptors and Their Ligands: Promising Molecular Biomarkers and

Therapeutic Targets in Prostate Cancer", *Biochim Biophys Acta*, Vol. 1835, No. 2, pp. 243-257.

Luty, A.A., J.B. Kwok, C. Dobson-Stone, C.T. Loy, K.G. Coupland, H. Karlstrom, T. Sobow, J. Tchorzewska, A. Maruszak, M. Barcikowska, P.K. Panegyres, C. Zekanowski, W.S. Brooks, K.L. Williams, I.P. Blair, K.A. Mather, P.S. Sachdev, G.M. Halliday and P.R. Schofield, 2010, "Sigma Nonopioid Intracellular Receptor 1 Mutations Cause Frontotemporal Lobar Degeneration-Motor Neuron Disease", *Annals of Neurology*, Vol. 68, No. 5, pp. 639-649.

Majounie, E., A.E. Renton, K. Mok, E.G. Dopper, A. Waite, S. Rollinson, A. Chio, G. Restagno, N. Nicolaou, J. Simon-Sanchez, J.C. van Swieten, Y. Abramzon, J.O. Johnson, M. Sendtner, R. Pamphlett, R.W. Orrell, S. Mead, K.C. Sidle, H. Houlden, J.D. Rohrer, K.E. Morrison, H. Pall, K. Talbot, O. Ansorge, D.G. Hernandez, S. Arepalli, M. Sabatelli, G. Mora, M. Corbo, F. Giannini, A. Calvo, E. Englund, G. Borghero, G.L. Floris, A.M. Remes, H. Laaksovirta, L. McCluskey, J.Q. Trojanowski, V.M. Van Deerlin, G.D. Schellenberg, M.A. Nalls, V.E. Drory, C.S. Lu, T.H. Yeh, H. Ishiura, Y. Takahashi, S. Tsuji, I. Le Ber, A. Brice, C. Drepper, N. Williams, J. Kirby, P. Shaw, J. Hardy, P.J. Tienari, P. Heutink, H.R. Morris, S. Pickering-Brown and B.J. Traynor, 2012, "Frequency of the C9orf72 Hexanucleotide Repeat Expansion in Patients with Amyotrophic Lateral Sclerosis and Frontotemporal Dementia: A Cross-Sectional Study", *Lancet Neurology*, Vol. 11, No. 4, pp. 323-330.

Marshall, C.R., A. Noor, J.B. Vincent, A.C. Lionel, L. Feuk, J. Skaug, M. Shago, R. Moessner, D. Pinto, Y. Ren, B. Thiruvahindrapduram, A. Fiebig, S. Schreiber, J. Friedman, C.E. Ketelaars, Y.J. Vos, C. Ficicioglu, S. Kirkpatrick, R. Nicolson, L. Sloman, A. Summers, C.A. Gibbons, A. Teebi, D. Chitayat, R. Weksberg, A. Thompson, C. Vardy, V. Crosbie, S. Luscombe, R. Baatjes, L. Zwaigenbaum, W. Roberts, B. Fernandez, P. Szatmari and S.W. Scherer, 2008, "Structural Variation of Chromosomes in Autism Spectrum Disorder", *American Journal of Human Genetics*, Vol. 82, No. 2, pp. 477-488.

Maruyama, H., H. Morino, H. Ito, Y. Izumi, H. Kato, Y. Watanabe, Y. Kinoshita, M. Kamada, H. Nodera, H. Suzuki, O. Komure, S. Matsuura, K. Kobatake, N. Morimoto, K. Abe, N. Suzuki, M. Aoki, A. Kawata, T. Hirai, T. Kato, K. Ogasawara, A. Hirano, T. Takumi, H. Kusaka, K. Hagiwara, R. Kaji and H. Kawakami, 2010, "Mutations of Optineurin in Amyotrophic Lateral Sclerosis", *Nature*, Vol. 465, No. 7295, pp. 223-226.

Mendez, R., N. Aptsiauri, A. Del Campo, I. Maleno, T. Cabrera, F. Ruiz-Cabello, F. Garrido and A. Garcia-Lora, 2009, "Hla and Melanoma: Multiple Alterations in Hla Class I and Ii Expression in Human Melanoma Cell Lines from Estdab Cell Bank", *Cancer Immunology, Immunotherapy*, Vol. 58, No. 9, pp. 1507-1515.

Morita, M., A. Al-Chalabi, P.M. Andersen, B. Hosler, P. Sapp, E. Englund, J.E. Mitchell, J.J. Habgood, J. de Belleroche, J. Xi, W. Jongjaroenprasert, H.R. Horvitz, L.G. Gunnarsson and R.H. Brown, Jr., 2006, "A Locus on Chromosome 9p Confers Susceptibility to Als and Frontotemporal Dementia", *Neurology*, Vol. 66, No. 6, pp. 839-844.

Nanetti, L., R. Fancellu, C. Tomasello, C. Gellera, D. Pareyson and C. Mariotti, 2009, "Rare Association of Motor Neuron Disease and Spinocerebellar Ataxia Type 2 (Sca2): A New Case and Review of the Literature", *Journal of Neurology*, Vol. 256, No. 11, pp. 1926-1928.

Nishimura, A.L., M. Mitne-Neto, H.C. Silva, A. Richieri-Costa, S. Middleton, D. Cascio, F. Kok, J.R. Oliveira, T. Gillingwater, J. Webb, P. Skehel and M. Zatz, 2004, "A Mutation in the Vesicle-Trafficking Protein Vapb Causes Late-Onset Spinal Muscular Atrophy and Amyotrophic Lateral Sclerosis", *American Journal of Human Genetics*, Vol. 75, No. 5, pp. 822-831.

Orlacchio, A., C. Babalini, A. Borreca, C. Patrono, R. Massa, S. Basaran, R.P. Munhoz, E.A. Rogaeva, P.H. St George-Hyslop, G. Bernardi and T. Kawarai, 2010, "Spatacsin Mutations Cause Autosomal Recessive Juvenile Amyotrophic Lateral Sclerosis", *Brain*, Vol. 133, No. Pt 2, pp. 591-598.

Polymenidou, M. and D.W. Cleveland, 2011, "The Seeds of Neurodegeneration: Prion-Like Spreading in Als", *Cell*, Vol. 147, No. 3, pp. 498-508.

Pratt, A.J., E.D. Getzoff and J.J. Perry, 2012, "Amyotrophic Lateral Sclerosis: Update and New Developments", *Degenerative Neurological and Neuromuscular Disease*, Vol. 2012, No. 2, pp. 1-14.

Pujana, M.A., J. Corral, M. Gratacos, O. Combarros, J. Berciano, D. Genis, I. Banchs, X. Estivill and V. Volpini, 1999, "Spinocerebellar Ataxias in Spanish Patients: Genetic Analysis of Familial and Sporadic Cases. The Ataxia Study Group", *Human Genetics*, Vol. 104, No. 6, pp. 516-522.

Redler, R.L. and N.V. Dokholyan, 2012, "The Complex Molecular Biology of Amyotrophic Lateral Sclerosis (Als)", *Progress in Molecular Biology and Translational Science*, Vol. 107, pp. 215-262.

Renton, A.E., E. Majounie, A. Waite, J. Simon-Sanchez, S. Rollinson, J.R. Gibbs, J.C. Schymick, H. Laaksovirta, J.C. van Swieten, L. Myllykangas, H. Kalimo, A. Paetau, Y. Abramzon, A.M. Remes, A. Kaganovich, S.W. Scholz, J. Duckworth, J. Ding, D.W. Harmer, D.G. Hernandez, J.O. Johnson, K. Mok, M. Ryten, D. Trabzuni, R.J. Guerreiro, R.W. Orrell, J. Neal, A. Murray, J. Pearson, I.E. Jansen, D. Sondervan, H. Seelaar, D. Blake, K. Young, N. Halliwell, J.B. Callister, G. Toulson, A. Richardson, A. Gerhard, J. Snowden, D. Mann, D. Neary, M.A. Nalls, T. Peuralinna, L. Jansson, V.M. Isoviita, A.L. Kaivorinne, M. Holtta-Vuori, E. Ikonen, R. Sulkava, M. Benatar, J. Wuu, A. Chio, G. Restagno, G. Borghero, M. Sabatelli, D. Heckerman, E. Rogaeva, L. Zinman, J.D. Rothstein, M. Sendtner, C. Drepper, E.E. Eichler, C. Alkan, Z. Abdullaev, S.D. Pack, A. Dutra, E. Pak, J. Hardy, A. Singleton, N.M. Williams, P. Heutink, S. Pickering-Brown, H.R. Morris, P.J. Tienari and B.J. Traynor, 2011, "A Hexanucleotide Repeat Expansion in C9orf72 Is the Cause of Chromosome 9p21-Linked Als-Ftd", *Neuron*, Vol. 72, No. 2, pp. 257-268.

Robberecht, W. and T. Philips, 2013, "The Changing Scene of Amyotrophic Lateral Sclerosis", *Nature Reviews Neuroscience*, pp. 717-728.

Rosen, D.R., 1993, "Mutations in Cu/Zn Superoxide Dismutase Gene Are Associated with Familial Amyotrophic Lateral Sclerosis", *Nature*, Vol. 364, No. 6435, pp. 362.

Ross, O.A., N.J. Rutherford, M. Baker, A.I. Soto-Ortolaza, M.M. Carrasquillo, M. DeJesus-Hernandez, J. Adamson, M. Li, K. Volkening, E. Finger, W.W. Seeley, K.J. Hatanpaa, C. Lomen-Hoerth, A. Kertesz, E.H. Bigio, C. Lippa, B.K. Woodruff, D.S. Knopman, C.L. White, 3rd, J.A. Van Gerpen, J.F. Meschia, I.R. Mackenzie, K. Boylan, B.F. Boeve, B.L. Miller, M.J. Strong, R.J. Uitti, S.G. Younkin, N.R. Graff-Radford, R.C. Petersen, Z.K. Wszolek, D.W. Dickson and R. Rademakers, 2011, "Ataxin-2 Repeat-Length Variation and Neurodegeneration", *Human Molecular Genetics*, Vol. 20, No. 16, pp. 3207-3212.

Sanna, S., A.U. Jackson, R. Nagaraja, C.J. Willer, W.M. Chen, L.L. Bonnycastle, H. Shen, N. Timpson, G. Lettre, G. Usala, P.S. Chines, H.M. Stringham, L.J. Scott, M. Dei, S. Lai, G. Albai, L. Crisponi, S. Naitza, K.F. Doheny, E.W. Pugh, Y. Ben-Shlomo, S. Ebrahim, D.A. Lawlor, R.N. Bergman, R.M. Watanabe, M. Uda, J. Tuomilehto, J. Coresh, J.N. Hirschhorn, A.R. Shuldiner, D. Schlessinger, F.S. Collins, G. Davey Smith, E. Boerwinkle, A. Cao, M. Boehnke, G.R. Abecasis and K.L. Mohlke, 2008, "Common Variants in the Gdf5-Uqcc Region Are Associated with Variation in Human Height", *Nature Genetics*, Vol. 40, No. 2, pp. 198-203.

Sapp, P.C., B.A. Hosler, D. McKenna-Yasek, W. Chin, A. Gann, H. Genise, J. Gorenstein, M. Huang, W. Sailer, M. Scheffler, M. Valesky, J.L. Haines, M. Pericak-Vance, T. Siddique, H.R. Horvitz and R.H. Brown, Jr., 2003, "Identification of Two Novel Loci for Dominantly Inherited Familial Amyotrophic Lateral Sclerosis", *American Journal of Human Genetics*, Vol. 73, No. 2, pp. 397-403.

Shatunov, A., K. Mok, S. Newhouse, M.E. Weale, B. Smith, C. Vance, L. Johnson, J.H. Veldink, M.A. van Es, L.H. van den Berg, W. Robberecht, P. Van Damme, O. Hardiman, A.E. Farmer, C.M. Lewis, A.W. Butler, O. Abel, P.M. Andersen, I. Fogh, V. Silani, A. Chio, B.J. Traynor, J. Melki, V. Meininger, J.E. Landers, P. McGuffin, J.D. Glass, H. Pall, P.N. Leigh, J. Hardy, R.H. Brown, Jr., J.F. Powell, R.W. Orrell, K.E. Morrison, P.J. Shaw, C.E. Shaw and A. Al-Chalabi, 2010, "Chromosome 9p21

in Sporadic Amyotrophic Lateral Sclerosis in the Uk and Seven Other Countries: A Genome-Wide Association Study", *Lancet Neurology*, Vol. 9, No. 10, pp. 986-994.

Shaw, B.F. and J.S. Valentine, 2007, "How Do Als-Associated Mutations in Superoxide Dismutase 1 Promote Aggregation of the Protein?", *Trends in Biochemical Sciences*, Vol. 32, No. 2, pp. 78-85.

Sobczak, K. and W.J. Krzyzosiak, 2005, "Cag Repeats Containing Caa Interruptions Form Branched Hairpin Structures in Spinocerebellar Ataxia Type 2 Transcripts", *The Journal of Biological Chemistry*, Vol. 280, No. 5, pp. 3898-3910.

Sreedharan, J., I.P. Blair, V.B. Tripathi, X. Hu, C. Vance, B. Rogelj, S. Ackerley, J.C. Durnall, K.L. Williams, E. Buratti, F. Baralle, J. de Belleroche, J.D. Mitchell, P.N. Leigh, A. Al-Chalabi, C.C. Miller, G. Nicholson and C.E. Shaw, 2008, "Tdp-43 Mutations in Familial and Sporadic Amyotrophic Lateral Sclerosis", *Science*, Vol. 319, No. 5870, pp. 1668-1672.

Tsuji, S., 2010, "Genetics of Neurodegenerative Diseases: Insights from High-Throughput Resequencing", *Human Molecular Genetics*, Vol. 19, No. R1, pp. R65-70.

Vaidya, A., A. Pniak, G. Lemke and A. Brown, 2003, "Epha3 Null Mutants Do Not Demonstrate Motor Axon Guidance Defects", *Molecular and Cellular Biology*, Vol. 23, No. 22, pp. 8092-8098.

Van Damme, P., J.H. Veldink, M. van Blitterswijk, A. Corveleyn, P.W. van Vught, V. Thijs, B. Dubois, G. Matthijs, L.H. van den Berg and W. Robberecht, 2011, "Expanded Atxn2 Cag Repeat Size in Als Identifies Genetic Overlap between Als and Sca2", *Neurology*, Vol. 76, No. 24, pp. 2066-2072.

Van der Zee, J., I. Gijselinck, L. Dillen, T. Van Langenhove, J. Theuns, S. Engelborghs, S. Philtjens, M. Vandenbulcke, K. Sleegers, A. Sieben, V. Baumer, G. Maes, E. Corsmit, B. Borroni, A. Padovani, S. Archetti, R. Perneczky, J. Diehl-Schmid, A. de Mendonca, G. Miltenberger-Miltenyi, S. Pereira, J. Pimentel, B. Nacmias, S. Bagnoli, S. Sorbi, C. Graff, H.H. Chiang, M. Westerlund, R. Sanchez-Valle, A.

Llado, E. Gelpi, I. Santana, M.R. Almeida, B. Santiago, G. Frisoni, O. Zanetti, C. Bonvicini, M. Synofzik, W. Maetzler, J.M. Vom Hagen, L. Schols, M.T. Heneka, F. Jessen, R. Matej, E. Parobkova, G.G. Kovacs, T. Strobel, S. Sarafov, I. Tournev, A. Jordanova, A. Danek, T. Arzberger, G.M. Fabrizi, S. Testi, E. Salmon, P. Santens, J.J. Martin, P. Cras, R. Vandenberghe, P.P. De Deyn, M. Cruts, C. Van Broeckhoven, J. Muller Vom Hagen, A. Ramirez, D. Kurzwelly, C. Sachtleben, W. Mairer, C. Firmo, A. Antonell, J. Molinuevo, A. Kinhult Stahlbom, H. Thonberg, I. Nennesmo, A. Borjesson-Hanson, V. Bessi, I. Piaceri, M. Helena Ribeiro, M. Rosario Almeida, C. Oliveira, J. Massano, C. Garret, P. Pires, A. Danel, G. Maria Fabrizi, S. Ferrari and T. Cavallaro, 2013, "A Pan-European Study of the C9orf72 Repeat Associated with Ftld: Geographic Prevalence, Genomic Instability, and Intermediate Repeats", *Human Mutation*, Vol. 34, No. 2, pp. 363-373.

Van Es, M.A., J.H. Veldink, C.G. Saris, H.M. Blauw, P.W. van Vught, A. Birve, R. Lemmens, H.J. Schelhaas, E.J. Groen, M.H. Huisman, A.J. van der Kooi, M. de Visser, C. Dahlberg, K. Estrada, F. Rivadeneira, A. Hofman, M.J. Zwarts, P.T. van Doormaal, D. Rujescu, E. Strengman, I. Giegling, P. Muglia, B. Tomik, A. Slowik, A.G. Uitterlinden, C. Hendrich, S. Waibel, T. Meyer, A.C. Ludolph, J.D. Glass, S. Purcell, S. Cichon, M.M. Nothen, H.E. Wichmann, S. Schreiber, S.H. Vermeulen, L.A. Kiemeney, J.H. Wokke, S. Cronin, R.L. McLaughlin, O. Hardiman, K. Fumoto, R.J. Pasterkamp, V. Meininger, J. Melki, P.N. Leigh, C.E. Shaw, J.E. Landers, A. Al-Chalabi, R.H. Brown, Jr., W. Robberecht, P.M. Andersen, R.A. Ophoff and L.H. van den Berg, 2009, "Genome-Wide Association Study Identifies 19p13.3 (Unc13a) and 9p21.2 as Susceptibility Loci for Sporadic Amyotrophic Lateral Sclerosis", *Nature Genetics*, Vol. 41, No. 10, pp. 1083-1087.

Van Hoecke, A., L. Schoonaert, R. Lemmens, M. Timmers, K.A. Staats, A.S. Laird, E. Peeters, T. Philips, A. Goris, B. Dubois, P.M. Andersen, A. Al-Chalabi, V. Thijs, A.M. Turnley, P.W. van Vught, J.H. Veldink, O. Hardiman, L. Van Den Bosch, P. Gonzalez-Perez, P. Van Damme, R.H. Brown, Jr., L.H. van den Berg and W. Robberecht, 2012, "Epha4 Is a Disease Modifier of Amyotrophic Lateral Sclerosis in Animal Models and in Humans", *Nat Med*, Vol. 18, No. 9, pp. 1418-1422.

Vance, C., A. Al-Chalabi, D. Ruddy, B.N. Smith, X. Hu, J. Sreedharan, T. Siddique, H.J. Schelhaas, B. Kusters, D. Troost, F. Baas, V. de Jong and C.E. Shaw, 2006, "Familial Amyotrophic Lateral Sclerosis with Frontotemporal Dementia Is Linked to a Locus on Chromosome 9p13.2-21.3", *Brain*, Vol. 129, pp. 868-876.

Vance, C., B. Rogelj, T. Hortobagyi, K.J. De Vos, A.L. Nishimura, J. Sreedharan, X. Hu, B. Smith, D. Ruddy, P. Wright, J. Ganesalingam, K.L. Williams, V. Tripathi, S. Al-Saraj, A. Al-Chalabi, P.N. Leigh, I.P. Blair, G. Nicholson, J. de Belleroche, J.M. Gallo, C.C. Miller and C.E. Shaw, 2009, "Mutations in Fus, an Rna Processing Protein, Cause Familial Amyotrophic Lateral Sclerosis Type 6", *Science*, Vol. 323, No. 5918, pp. 1208-1211.

Wu, C.H., C. Fallini, N. Ticozzi, P.J. Keagle, P.C. Sapp, K. Piotrowska, P. Lowe, M. Koppers, D. McKenna-Yasek, D.M. Baron, J.E. Kost, P. Gonzalez-Perez, A.D. Fox, J. Adams, F. Taroni, C. Tiloca, A.L. Leclerc, S.C. Chafe, D. Mangroo, M.J. Moore, J.A. Zitzewitz, Z.S. Xu, L.H. van den Berg, J.D. Glass, G. Siciliano, E.T. Cirulli, D.B. Goldstein, F. Salachas, V. Meininger, W. Rossoll, A. Ratti, C. Gellera, D.A. Bosco, G.J. Bassell, V. Silani, V.E. Drory, R.H. Brown, Jr. and J.E. Landers, 2012, "Mutations in the Profilin 1 Gene Cause Familial Amyotrophic Lateral Sclerosis", *Nature*, Vol. 488, No. 7412, pp. 499-503.

Yu, Z., Y. Zhu, A.S. Chen-Plotkin, D. Clay-Falcone, L. McCluskey, L. Elman, R.G. Kalb, J.Q. Trojanowski, V.M. Lee, V.M. Van Deerlin, A.D. Gitler and N.M. Bonini, 2011, "Polyq Repeat Expansions in Atxn2 Associated with Als Are Caa Interrupted Repeats", *PLoS One*, Vol. 6, No. 3, pp. e17951.