CHARLES UNIVERSITY

Faculty of Pharmacy in Hradec Králové
Department of Pharmacology and Toxicology

Analysis of genomic regions bound and regulated by Ataxin-3

Diploma Thesis

Supervisor: PharmDr. Martina Čečková, Ph.D.

Specialized supervisor: PD Dr. Bernd Evert

Hradec Králové 2017 **Martin Svoreň**

I hereby confirm that this diploma thesis is my original copyrighted work. All the literature and other sources from which I drew during the processing are listed in the bibliography and properly cited. This work was not used to obtain equal or different degree.

Martin Svoreň

Acknowledgments

I would like to thank PD Dr. Bernd Evert for giving me the opportunity to work in his laboratory, for his enthusiasm and outstanding leadership. My thank belongs also to PharmDr. Martina Čečková, Ph.D. for help and comments on the diploma thesis.

Abstract

Charles University

Faculty of Pharmacy in Hradec Králové

Department of Pharmacology and Toxicology

Student: Martin Svoreň

Supervisor: PharmDr. Martina Čečková, Ph.D.

Specialized supervisor: PD Dr. Bernd Evert

Title of diploma thesis: Analysis of genomic regions bound and regulated by Ataxin-3

Spinocerebellar ataxia type 3 (SCA3), also known as Machado-Joseph disease, is a dominantly inherited neurodegenerative disease. In SCA3, the disease protein ataxin-3 (ATXN3) contains an abnormally long polyglutamine (polyQ) tract encoded by CAG repeat expansion. ATXN3 binds DNA and interacts with transcriptional regulators pointing toward a direct role of ATXN3 in transcription. It is conceivable that mutant ATXN3 triggers multiple, interconnected pathogenic cascades leading to neurotoxicity, however, the principal molecular pathomechanism remains elusive. Here, PCR analyses of 16 ATXN3-bound genomic regions recently identified by next generation sequencing of immunoprecipitated ATXN3-bound chromatin fragments confirmed enriched binding of ATXN3 to 5 genomic regions next to genes encoding CCAAT/enhancer binding protein delta (CEBPD), period circadian clock-2 (PER2), phosphatase and tensin homolog (PTEN), serine protease inhibitor family F2 (SERPINF2) and thrombospondin-1 (THBS1). To investigate putative regulatory effects of ATXN3, the ATXN3-bound genomic regions were subcloned in, luciferase reporter constructs. Subsequently, wild type (WT) and hemizygous ATXN3-knockout human neuroblastoma cell line (SH-SY5Y) were transfected with the constructs and analyzed for ATXN3-modulated changes in luciferase activity. In ATXN3 knockout cells, a repressive ATXN3-dependent change in luciferase activity was found for CEBPD and THBS1 suggesting that ATXN3 binds to these regions to enhance or maintain transcriptional repression. In cells overexpressing normal or mutant ATXN3, an ATXN3 isoform dependent, regulatory effect was also found for CEBPD and THBS1 suggesting that ATXN3

may be principally involved in the transcriptional regulation of these genes. The identified and analyzed genomic regions may contribute to pathogenetic processes involved in SCA3 disease.

Abstrakt

Karlova Univerzita

Farmaceutická fakulta v Hradci Králové

Katedra farmakológie a toxikológie

Študent: Martin Svoreň

Školiteľ: PharmDr. Martina Čečková, Ph.D.

Školiteľ-špecialista: Doc. Bernd Evert, Ph.D.

Názov diplomovej práce: Analýza oblastí genómu viazaných a regulovaných ataxínom 3

Spinocerebelárna ataxia typu 3 (SCA3), známa tiež ako Machado-Josephova choroba, je dominantne dedičné neurodegeneratívne ochorenie. Je spôsobené mutantným proteínom ataxínom 3 (ATXN3), ktorý obsahuje abnormálne dlhý polyglutamínový segment kódovaný opakujúcim sa CAG kodónom. ATXN3 je schopný viazať DNA a interaguje s celým radom transkripčných faktorov. Tieto fakty poukazujú na priamu rolu ataxínu 3 v transkripcii. Je pravdepodobné, že mutantný ATXN3 spúšťa mnohé, vzájomne prepojené patogénne kaskády, vedúce k neuronálnej toxicite, no hlavný patomechanizmus zodpovedný za rozvoj tejto neurodegenratívnej choroby na molekulárnej úrovni ostáva neznámy. Nedávno sa sekvenovaním imunoprecipitovaných chromatínových fragmentov viazaných ataxínom 3 podarilo identifikovať 16 genómových oblastí, ktoré boli za účelom potvrdenia prítomnosti väzby ataxínu 3 na DNA ďalej analyzované pomocou PCR. Zvýšenú väzbovosť špecifickú pre ATXN3 vykázalo 5 oblastí v blízkosti týchto génov: CCAAT/enhancer viažúci proteín delta (CEBPD), preiodický cirkadiálny clock-2 (PER2), fosfatázový a tenzínový homológ (PTEN), inhibítor serínových proteáz rodiny F2 (SERPINF2) a trombospondín-1 (THBS1). Tieto genómové oblasti boli následne vklonované do luciferázových reporterov za účelom preskúmania možného regulačného efektu sprostredkovaného ataxínom 3. Následne bol pripravenými reportermi transfekovaný divoký typ (wild type) a hemizigotný ATXN3-knockoutovaný podtyp bunkovej línie ľudského neuroblastómu (SH-SY5Y). V ďalšom korku bola analyzovaná ataxínom 3 modulovaná luciferázová aktivita. V knockoutovanej bunkovej línie bol pozorovaný pokles luciferázovej

aktivity sprostredkovaný ataxínom 3 v prípade genómových oblastí spojených s CEBPD a THBS1, naznačujúc, že ATXN3 sa viaže na tieto oblasti genómu za účelom vystupňovania alebo udržania transkripčnej represie. V bunkách, ktoré nadmerne exprimovali normálny alebo mutantný ATXN3 bol taktiež pozorovaný regulačný efekt mediovaný jednotlivými izoformami ataxínu 3 v prípade CEBPD a THBS1, čo poukazuje na možný vplyv ataxínu 3 na transkripčú reguláciu spomínaných génov. Identifikované a analyzované oblsasti ľudského genómu by sa mohli podieľať na patogenetických procesoch vedúcich k SCA3 manifestácii.

List of abbreviations

ATXN3	Ataxin 3
CBP	cAMP-response element binding protein
CEBPD	CCAAT/enhancer binding protein delta
ChIP	Chromatin immunoprecipitation
Ctr	control
dATP	deoxyadenosine triphosphate
dCTP	deoxycytidine triphosphate
dGTP	deoxyguanosine triphosphate
dTTP	deoxythymidine triphosphate
EtBr	ethidium bromide
FBS	fetal bovine serum
GFP	green fluorescent protein
HD	Huntington disease
HDL2	Huntington disease-like 2
iPSC	induced pluripotent stem cell
IPTG	isopropyl β-D-1-thiogalactopyranoside
MJD	Machado-Joseph disease
NIs	nuclear inclusions
PCR	polymerase chain reaction

PER2	period circadian clock-2
polyQ	polyglutamine
PTEN	phosphatase and tensin homolog
SCA3	Spinocerebellar ataxia type 3
SERPINF2	serine protease inhibitor family F2
SNP	single nucleotide polymorphism
TAE	tris(hydroxymethyl)aminomethane-acetate-ethylenediaminetetraacetic acid
TBP	TATA-binding protein
THBS1	thrombospondin-1
UPS	ubiquitin proteasome system
WT	wild type
x-gal	5-bromo-4-chloro-indolyl-β-D-galactopyranoside
β-gal	β-galactosidase

Contents

1. Introduction

With increasing life expectancy, thanks to medical advancements, the society is now facing the problem with ageing of population, which is connected to another huge issue, neurodegenerative diseases (YoungSoo Kim et al., 2012). It was estimated that these disorders are found in about 5% of patients suffering from brain disorders. The total cost of neurodegenerative diseases in 28 countries in Europe was evaluated at about 72 billion Euros per year (Nieoullon, 2011). Pathology of these diseases is defined by continual neuronal loss usually associated with protein aggregation (Barnham et al., 2004). In general, characteristic features of neurodegenerative disorders are late onset, sensory-motoric and cognitive abilities impairment resulting from neuronal loss deteriorating with time. So far, many of the neurodegenerative diseases are not curable and the exact, usually complex pathomechanism, is not described, as they are varying in their origin what divides them in several groups (Nieoullon, 2011).

Dominantly inherited neurodegenerative disorders, caused by polyglutamine repeats belong to the group of so called polyQ diseases. PolyQ diseases include spinocerebellar ataxias (SCA 1,2,6,7,17), and Huntington disease-like 2 (HDL2). Most commonly known polyQ diseases are Huntington disease (HD) and Spinocerebellar ataxia type 3 (SCA3).

In SCA3 disease, the mutation resulting in production of disease protein ATXN3 is well explained, nevertheless, the molecular mechanism resulting in neuronal death was not yet described. The main complication regarding ATXN3 is its involvement in numerous cellular processes and it is hard to define which one of them influence and trigger the onset of the disease (Costa & Paulson, 2012).

Consequently, this work is focused on identification of genes potentially involved in SCA3 pathogenesis, which could help improve the whole understanding of this severe, potentially lethal disease.

2. Theoretical part

2.1 Neurodegenerative diseases

Neurodegenerative diseases are defined as progressive neurological disorders characterized with selective neuronal loss in distinct parts of brain, resulting in various clinical presentations. Despite the broad etiological variations (e.g. hereditary disorders like Huntington's disease caused by genetic mutation), the common feature of these disorders is formation of inter- or intraneuronal aggregates and inclusions built up from proteins with altered physicochemical properties and other proteasomal components. The most prevalent and intensively studied neurodegenerative diseases are Alzheimer's disease, Parkinson's disease and amyotrophic lateral sclerosis (Kovacs, 2014; YoungSoo Kim et al., 2012).

2.2 Trinucleotide repeat disorders

Inherited neurological disorders characterized by expansions of trinucleotide repeats belong to the group of so called trinucleotide repeat disorders. They form the largest group of inherited neurodegenerative diseases with so far 16 described disorders. It was shown, that the trinucleotide expansions are unstable and the number of trinucleotide repeats might be higher in successive generations (Fu et al., 1991). These types of mutations are therefore called dynamic mutation. The severity of symptoms and the age of onset vary within these diseases and they are dependent on the number of instable trinucleotide repeats (Orr & Zoghbi, 2007). Trinucleotide repeat disorders differ also in exact expanded trinucleotide sequence and in their pathogenic mechanism.

2.3 Polyglutamine repeat disorders

Disorders, where expansions of trinucleotide repeat CAG is present in the coding regions of various genes, form a subgroup of trinucleotide repeat disorders, which are called polyglutamine (polyQ) diseases. CAG triplet codes for the amino acid glutamine, therefore the characteristic feature of these diseases is expanded

glutamine repeats in the respective mutant proteins. Each of these diseases have a different mutant protein (La Spada, et al., 1994; Zoghbi & Orr, 2000).

A characteristic feature of all these disorders is the formation of intraneuronal inclusions, which mainly include the expanded polyglutamine proteins and several other proteins of neuronal proteasome. In every polyglutamine disease, different types of neurons are affected, this result in a pattern of atrophy unique for each polyglutamine disease and consequently explains different symptoms of each disorder (Davies et al., 1998; Ross, 1997; Rubinsztein, et al., 1999).

A number mechanisms such as mitochondrial dysfunction, proteolytic cleavage of mutant protein, failure to clear the mutant protein, shuttling of the mutant protein to the nucleus and its subsequent aggregation, these all has shown to contribute to the observed neuronal loss, but the exact mechanism of neuronal death caused by the presence of the mutant protein is not known (Weber, et al., 2014). So far, there are nine polyglutamine disease known (see Introduction).

One of the most prevalent and best studied dominantly inherited ataxias is the Spinocerebellar ataxia type 3 (SCA3), which is also the main topic of this project.

2.4 Spinocerebellar ataxia type 3

Spinocerebellar ataxia type 3 (SCA3), also known as Machado-Joseph disease (MJD), is a neurodegenerative disease caused by an unstable CAG repeat expansion in the SCA3 gene leading to an expansion of polyglutamines (polyQ) in the corresponding protein, ataxin-3 (ATXN3) (Kawaguchi et al., 1994). The repeat expansion ranges from 12 to 44 triplets in healthy individuals and from 60 to 87 in SCA3 patients. In affected individuals, a slowly progressive ataxia syndrome appears, typically beginning between the ages of 20 and 50 years, characterized primarily by cerebellar ataxia and pyramidal signs. Presenting features include gait problems, speech difficulties, clumsiness, and often visual blurring and diplopia. Symptoms get worse over time leading to the need for assistive devices (including wheelchair) 10 to 15 years following onset. Currently, there is no cure for SCA3 but

treatments for symptoms can improve quality of life significantly (Costa & Paulson, 2012).

2.4.1 SCA3 pathology

Neuropathologically, SCA3 is characterized by neuronal loss in the pons, cerebellum, thalamus and spinal cord. The cerebellum typically shows atrophy while the cerebral cortex is spared in disease. As in most polyQ diseases, affected neurons in SCA3 show characteristic formation of nuclear inclusions (NIs) containing polyQ-expanded ATXN3, components of the ubiquitin proteasome system (UPS), chaperones, autophagy markers, transcription factors and other polyQ-containing proteins.

2.4.2 The disease protein – ataxin-3

ATXN3 is mainly localized in the cytoplasm but also shuttles between the cytoplasm and nucleus. It is a ubiquitously expressed deubiquitinase that controls protein quality and degradation of misfolded proteins via the UPS. ATXN3 has neuroprotective functions in stress response and regulation of aging-related signaling pathways. Moreover, ATXN3 binds DNA and interacts with transcription regulators pointing toward a direct role for ATXN3 in transcription (Costa & Paulson, 2012).

2.4.3 Neurodegeneration mediated by ataxin-3

The pathogenic process of neurodegeneration in SCA3 involves nuclear translocation and aggregation of ATXN3. Transgenic mice expressing mutant ATXN3 exclusively in the nucleus generate a severe neurodegenerative phenotype whereas expression of mutant ATXN3 directed to the cytoplasm causes only a mild phenotype (Bichelmeier et al., 2007). Recently, it was shown that nuclear import of ATXN3 is controlled by casein kinase 2-mediated phosphorylation of three serine residues in the C-terminus of ATXN3 (Mueller et al., 2009) whereas translocation to the cytoplasm is mediated by nuclear export signals (NES) in the N-terminus of ATXN3 (Antony et al., 2009; Macedo-Ribeiro et al., 2009). Aggregation of mutant

ATXN3 is enhanced by proteolysis that generates C-terminal polyQ-containing fragments which act as seeds for aggregation (Teixeira-Castro et al., 2011). Although mutant and normal ATXN3 undergo the same types of proteolytic cleavage, C-terminal fragments of ATXN3 are only found in brain homogenates from SCA3 patients and SCA3 transgenic mice (Colomer Gould et al., 2007; Goti, 2004). Nuclear accumulation of these fragments is caused most likely by the loss of NES signals after proteolytic release of the N-terminal part (Antony et al., 2009) and the lack of chaperone-mediated clearance by the UPS in the nucleus (Breuer, 2010). Proteolytic cleavage of ATXN3 can be mediated by calcium-dependent calpains. It is known that L-glutamate-induced excitation of SCA3 patient-specific, induced pluripotent stem cell (iPSC)-derived neurons initiates Ca^{2+}-dependent, calpain-mediated proteolysis of ATXN3 and formation of SDS-insoluble aggregates (Koch et al., 2011). Consistently, selective inhibition of calpain activity in transgenic SCA3 mice reduced proteolysis and nuclear aggregation of mutant ATXN3 and prevented polyQ-induced neurodegeneration (Simões et al., 2012).

2.4.4 Potential cause of Spinocerebellar ataxia type 3

The current understanding on the pathogenesis of polyQ diseases is that cellular toxicity and dysfunction is caused by soluble polyQ-containing oligomers of the expanded polyQ proteins (Takahashi et al., 2010). PolyQ oligomers form soluble aggregation intermediates with a large number of other proteins (e.g. transcription factors) and interfere with the protein homeostasis, mitochondrial function and transcriptional regulation (Costa & Paulson, 2012). Several transcriptional changes have been reported early in disease progression in various models of polyQ diseases suggesting that transcriptional dysregulation underlies disease pathogenesis in SCA3 and other polyQ disorders (Riley & Orr, 2006). In cell models of SCA3 several differentially expressed genes encoding transcription factors, inflammatory cytokines and cell surface proteins prior to the formation of NIs and cell death were identified (Evert et al., 2001; Evert et al., 1999; Evert et al., 2003; Jeub et al., 2006). Also, in cerebella of transgenic SCA3 mice, altered expression of genes involved in glutamatergic signaling and signal transduction occurs before the onset of

neurological symptoms (Chou et al., 2010). On the one hand, transcriptional dysregulation in polyQ disorders is caused by depletion of transcriptional components such as TBP (TATA binding protein) and CBP (cAMP-response element binding protein) into NIs. On the other hand, several of the polyQ proteins have normal functions in transcriptional regulation that are altered by polyQ expansion (Riley & Orr, 2006). It has been shown that mutant ATXN3 has an altered DNA binding reducing its ability to form histone deacetylating repressor complexes and to activate ATXN3-regulated genes (Araujo et al., 2011; Evert et al., 2006).

2.4.5 Genes potentially regulated by ataxin-3

Recently, using chromatin immunoprecipitation (ChIP) and next generation sequencing, Dr. Evert´s group identified several genomic regions significantly enriched by ataxin-3 in induced pluripotent stem (iPS) cells derived neurons from healthy controls and SCA3 patients. Moreover, differentially expressed mRNAs were previously found by RNA sequencing of iPS-derived neurons from controls and SCA3 patients. To identify ATXN3-regulated genes, the mRNA changes were correlated to differentially ATXN3-bound genomic regions resulting in a list of potential candidate genes regulated by ATXN3. Thus, using this set of genes, the aim of the thesis was (i) to validate enriched binding of ATXN3 to the genomic regions found by ChIP sequencing and (ii) to analyze the potential regulatory activity of the confirmed ATXN3-bound genomic regions by reporter assays.

3. Aims

The aim of the thesis was (i) to validate enriched binding of ATXN3 to genomic regions found by ChIP sequencing and (ii) to analyze the potential regulatory activity of confirmed ATXN3-bound genomic regions by reporter assays.

4. Experimental part

4.1 Materials

Sterile endotoxin-free water B. Braun Melsungen AG, Germany

dATP Solution (100 mM) 25µl, Thermo Scientific, USA

dCTP Solution (100 mM) 25µl, Thermo Scientific, USA

dGTP Solution (100 mM) 25µl, Thermo Scientific, USA

dTTP Solution (100 mM) 25µl, Thermo Scientific, USA

BioTherm™ Taq DNA Polymerase (5 U/µl), GeneCraft®, Germany

10x Taq DNA Polymerase Buffer A+, Segenetic, Germany

10x Taq DNA Polymerase Buffer without $MgCl_2$ GeneCraft®, Germany

$MgCl_2$ Solution for PCR (50 mM), GeneCraft®, Germany

Agarose low EEO, Applichem, Germany

GelRed™ Nucleic Acid gel stain, Biotium, USA

Opti-MEM® I Reduced Serum Medium, no phenol red, Gibco, UK

DMEM/F-12, GlutaMAX, Gibco, UK

10% active FBS, Biochrom GmbH, Germany

Nonessential amino acids 100mM, MEM NEAA 100x, Gibco, UK

Penicilin (10,000 U/ml), PenStrep, Gibco, UK

Streptomycin (10,000 µg/ml), PenStrep, Gibco, UK

Sodium pyruvate (1%), Gibco, UK

T4 DNA Ligase (5 U/µL), Thermo Scientific, USA

T4 DNA Ligase Buffer (10x), Thermo Scientific, USA

FastDigest Buffer (10X), Thermo Scientific, USA

BglII FastDigest, Thermo Scientific, USA

EcoRI FastDigest, Thermo Scientific, USA

HindIII FastDigest, Thermo Scientific, USA

PstI FastDigest, Thermo Scientific, USA

SacI FastDigest, Thermo Scientific, USA

XbaI FastDigest, Thermo Scientific, USA

XhoI FastDigest, Thermo Scientific, USA

LB-medium, Carl Roth, Germany

Select agar, Invitrogen, Spain

4.2 Methods

4.2.1 Chromatin preparation and immunoprecipitation (ChIP)

ChIP assays were performed using the Magna ChIP A/G kit (Millipore) following the instructions of the manufacturer. In brief, iPS-derived neural stem cells from controls and patients were seeded on 6-cm plates and differentiated to neurons for six weeks. Neuronal cells were fixed by adding formaldehyde (1% final concentration) and cross-linked adducts were resuspended and sonicated, resulting in an average chromatin fragment size of 400 bp. For ChIP, a mouse monoclonal antibody was used against ATXN3 (1H9, Millipore) and an unspecific mouse control IgG (Millipore). Protein-bound, immunoprecipitated DNA was reverse cross-linked and purified by using DNA purification mini-columns according to the instructions of the manufacturer.

4.2.2 Design and characteristics of used primers

The primers were designed using the software program Oligo (Visual Basic 6.0 SP5) which calculates melting and optimal annealing temperatures, considers the formation of hairpin loops and primer dimers. Designed primers were synthesized by Thermo Fisher Scientific (Germany).

Two different types of primers were used, (i) primers for the amplification and analysis of ATXN3-enriched genomic regions by ChIP-PCR and (ii) primers containing restriction sites for Sac I and Xho I (cleavage motives 5'- GAG CTC -3' and 5'- CTC GAG -3', respectively) for the subcloning of confirmed ATXN3-bound genomic regions into plasmids. Details about primers used for enrichment screening are listed in **Table 1** and for cloning are shown in **Table 2**.

Table 1 Primers used for analysis and verification of ATXN3-enriched genomic regions

Primer sequence (5' to 3')	Gene region	Melting temperature [°C]	Annealing temperature [°C]	Suggested temperatures [°C]	Length of gene region [bp]
ACAGCTGAGGCAGGCCCTGG	RRBP1(1)_For	59.5	59.3	58 / 60 /62	185
AGCTCCATTGCCCTTAAAGC	RRBP1(1)_Rev	55.5			
ACCTCATCCTGTCACCCCTG	SERPINF2(1)_For	52.4	57.1 - 60.8	52 / 54 / 56	185
CCTAGGAATCTGCATTTTTACC	SERPINF2(1)_Rev	53.4			
TCTTGAATTCTCCCAACAATCATC	ELAVL2_For	55.0	57.9 - 61.6	52 / 54 / 56	141
GGACACCTGAGCATCTTGCCTC	ELAVL2_Rev	56.2			
TCTTTCCACTAGCCTTGCTC	SEMA6A_For	50.7	51.6	50 / 52 /54	133
GTCAAGTTAAGAAGGGCAGG	SEMA6A_Rev	50.7			
ACAAACATCAAATACCAGAAAAGC	THBS1(1)_For	54.2	53.7	52 / 54 /56	231
CAAGAGTTTTGCTGTGGGATGG	THBS1(1)_Rev	55.8			
CCTAAAAGAGACTGGAGACCTGG	CEBPD-I(1)_For	54.9	57.0	56 / 58 / 60	239
GCACTGTGTGCTTGCATCTGG	CEBPD-I(1)_Rev	53.5			
ATCTATGGATATGAATAGGGCTG	DOCK10_For	52.8	56.9 - 60.6	56 / 58 / 60	163
CATTGAGTTCACCTTCCCAATC	DOCK10_Rev	53.2			
GTGGCTGAGGGCCCAGTTCC	COL6A3_For	59.1	57.7 - 61.4	58 / 60 / 62	162
CTGCCTTGTAGCTGGGAGAAG	COL6A3_Rev	54.3			
GTACATGTTTTCTTCCATTTTCC	TOX3_For	52.3	57.1 - 60.8	56 / 58 / 60	133
CTGTGGCATAGGTACTCATGGC	TOX3_Rev	53.8			
GAGTATGTGCCAGTTTTAAAAGAG	ARRDC3_For	52.6	57.2 - 60.9	56 / 58 / 60	132
CTTTTGCCCAGAAAAATGGCG	ARRDC3_Rev	60.3			
GTTGCCCTAGGGGAGATTCC	FOSL2_For	55.9	59.1	58 / 60 / 62	181
TGAGGAGTCCTCCTCCCACC	FOSL2_Rev	54.1			
TAAACAGATGTTGCAGAGGG	TLR4_For	48.9	55.8 - 59.5	56 / 58 / 60	137
CTTGTTCTCTTAATTTACTTGG	TLR4_Rev	47.9			
GGTCCAATAAACATCCCTGC	KLF10_For	52.5	54.4	52 / 54 /56	164
GAGTGGCTGAACCCATGTCC	KLF10_Rev	53.3			
CACAGAGTCCAGCAGGCAGC	PER2_For	53.3	57.6	56 / 58 / 60	141
CAGTCCCCTCCCCCTGGTGG	PER2_Rev	60.7			
GCCCCACAACAAATTCTCACC	THBS1(2)_For	55.5	57.6 - 61.3	58 / 60 / 62	146
GAGTTTTGCTGTGGGATGGAAG	THBS1(2)_Rev	55.4			
TGTGAGTCAGACTGCAGCTCTGG	CEBPD-I(2)_For	53.9	55.5	54 / 56 / 58	124
GCACTGTGTGCTTGCATCTGG	CEBPD-I(2)_Rev	53.5			
TGGACTCCAGGCCCAGCTC	RRBP1(2)_For	55.6	58.0 - 61.8	58 / 60 / 62	150
AGCTCCATTGCCCTTAAAGC	RRBP1_Rev	55.5			
ACCCCTGAGTGTGGCCCTGG	SERPINF2(2)_For	58.7	57.5 - 61.1	58 / 60 / 62	172
CCTAGGAATCTGCATTTTTACC	SERPINF2(1)_Rev	53.4			
ACCTCATCCTGTCACCCCTG	SERPINF2(1)_For	52.4	56.6 - 60.3	56 / 58 / 60	153
GTGACCCTGGTGGAGGAAC (1x SNP)	SERPINF2(2)_Rev	50.5			
ACCCCTGAGTGTGGCCCTGG	SERPINF2(2)_For	58.7	56.6 - 60.3	56 / 58 / 60	140
GTGACCCTGGTGGAGGAAC (1x SNP)	SERPINF2(2)_Rev	50.5			
TGCTTTTTTTCCGCCTTTCC	PTEN_For	58.5	52.4	52 / 54 / 56	117
GAAACATTCACCATGGCTGC	PTEN_Rev	52.2			
GGCTTCCTTCAGGGTCCAGC	CEBPD-III_For	57.4	55.3	54 / 56 / 58	148
GCAGAAATTTAACCACAGCCAC	CEBPD-III_Rev	54.5			

Table 2 Primers used for amplification and subcloning of confirmed ATXN3-bound genomic regions

Primer sequence (5' to 3')	Gene region	Melting temperatur e [°C]	Annealing temperature [°C]	Suggested temperatures [°C]	Length of gene region [bp]
gagctcGGTCCCACCTCTTTTATGG	THBS_SacI_For	50.7	60.6 - 64.3	58 / 60 / 62	968
ctcgagTTTCTTTCAGAGTTTGGGC	THBS_XhoI_Rev	49.3			
gagctCATCTCTGCTGGGGCAGG	PTEN_SacI_For	52.7	57.2 - 61.0	58 / 60 / 62	858
ctcgaGAACAACTGTGAGACTTGCAGG	PTEN_XhoI_Rev	50.7			
gagctcGGAAAAGTGAAATGGAATGG	CEBPD_SacI_For	51.0	56.2	54 / 56 / 58	901
ctcgagAATACCAGTGCCAGAGCAGC	CEBPD_XhoI_Rev	52.2			
gagctCAACCTCATCCTGTCACCCCTG	SERPINF2_SacI_For	56.2	58.1 - 61.8	58 / 60 / 62	940
ctcgaGTGAGACCCTCCAGCACAGTGG	SERPINF2_XhoI_Rev	55.5			
ctcgaGTTTTCAAGGTGATCCCTGC	PER2_XhoI_For	51.9	58.9	58 / 60 / 62	840
gagctCACCCCCTTCTGCGTTCAGC	PER2_SacI_Rev	58.1			

4.2.3 Polymerase chain reaction

The Polymerase Chain Reactions (PCR) were conducted in 25µl reaction volumes containing 0.2 µM of forward and 0.2 µM reverse oligonucleotide primers, 200 µM deoxynucleotide solution containing dATP, dTTP, dCTP und dGTP (Qiagen, Germany), 2.5 µl of the 10x PCR-Buffer containing 15 mM $MgCl_2$ (GC-002-006, BioThermTM, Genecraft; Germany) and 0.2 µl of the TaqPolymerase 5U/µl (GC-002-1000, BioThermTM, Genecraft; Germany). In some cases, 2.5 µl of DMSO and/or 2.5µl of Betaine were added to enhance specificity of PCR reaction. In other cases, Mg^{2+} concentration was increased up to 2.5 mM by using 2.5 µl of 10x Buffer without $MgCl_2$ and adding 1.25µl of 50mM $MgCl_2$ to increase PCR efficiency. The DNA templates were always preincubated at 45°C for 5 minutes and placed on ice. At least 1 to 2 µl (200 -400 ng) of the template DNA was added and filled up with nuclease-free water to its final volume. For analysis of ATXN3-enriched genomic regions, chromatin DNA immunoprecipitated with an ATXN3-specific antibody derived from iPSCs neurons of 3 SCA3 patients and 2 control subjects was used. For amplification and subcloning of confirmed ATXN3-bound regions, genomic DNA isolated from four different human cell lines was used as template. For each PCR, a negative control lacking template DNA was included, serving as a control for unspecific contamination.

The PCR reactions started with an initial DNA denaturation step at 94°C for 2 minutes followed by 32 cycles consisting of denaturation at 94°C for 45 seconds, annealing at 50-64°C for 45 seconds and elongation at 72°C for 45 seconds. After completion of the PCR cycles, a final step at 72°C for 5 minutes was included. The reactions were conducted in a Tprofessional Thermocycler (Biometra; Germany).

4.2.4 Agarose-TAE gel electrophoresis

4.2.4.1 Gels used for analysis and validation of ATXN3-enriched regions

The ChIP-PCR products were separated on 1,5% TAE agarose gel stained with Gel-red (2,5µl/100ml). Each agarose gel well was loaded with a mix of 10µl PCR product and 2µl 10x Xylene-Loading dye (Qiagen; Germany). For size discrimination, 4µl of a 100 bp DNA ladder and a 1 kb DNA ladder were used (500 ng/µl, Gene Craft; Germany). Gel electrophoresis was performed at 180V for 30-90 minutes (Electrophoresis Power Supply E831, Consort; Belgium) (41-2025-R, PeqLab Biotechnologie GmBH; Germany).

4.2.4.2 Gels used for preparation of DNA fragments and characterization of clones

DNA fragments obtained by restriction digests were separated on a 1% TAE agarose gel. Each agarose gel well was loaded with a mix of 15µl restriction digest and 2µl 10 Xylene-Loading dye (Qiagen; Germany). In addition, 10µl of a 1 kb DNA ladder were used (500 ng/µ, Gene Craft; Germany).

Gel electrophoresis was performed at 180V for 30-80 minutes (Electrophoresis Power Supply E831, Consort; Belgium) (41-2025-R, PeqLab Biotechnologie GmBH; Germany). Subsequently the agarose gel was taken out from electrophoresis chamber and stained in the dark in a 1x TAE solution containing ethidium bromide (0.5 µg/ml) (EtBr) for 10 minutes. Followed by two washes in H_2O in the dark for 10 minutes to remove unbound EtBr. The agarose gel was then placed on a ultraviolet illuminator to visualize EtBr-stained DNA. To enhance the

resolution and size discrimination of DNA fragments, a second electrophoresis was very often conducted.

4.2.5 Ligation of PCR amplified products to pCR2.1 vector

PCR products synthesized by Taq DNA polymerase contain 3′ deoxyadenosine overhangs and therefore are suitable to be ligated into the linearized pCR2.1 Topo TA vector containing 3′ thymidine overhangs (Thermo Fisher Scientific Inc.). The multiple cloning site (MCS) of the pCR2.1 vector contains several restriction sites for endonucleases (for instance EcoR I, Xho I and Sac I) and is located inside the E. coli lacZα gene, which encodes for the α-peptide of β-galactosidase (β-gal). The successful ligation of a PCR product within the MCS of the pCR2.1 vector results in disruption of the α-peptide and plasmid recircularization. In addition, the pCR2.1 vector carries resistance genes for ampicillin and kanamycin.

The ligation reaction consisting of 1 µl of Salt Solution, 1µl of the Topo TA pCR2.1 Vector, 1µl of fresh PCR product and 3µl of distillated water was conducted in 50µl reaction tube. The ligation reaction was incubated for 10 minutes at room temperature and immediately used for the transformation of competent bacterial cells.

4.2.6 Transformation of Competent Cells with pCR2.1-Topo Vector

Chemically competent E. coli (Top10F') cells, susceptible to ampicillin, were transformed in volumes of 52 µl consisting of 2 µl ligation product and 50 µl competent cells kept on ice constantly. The suspension was mixed gently and incubated on ice for 10 to 30 min, followed by a heat shock at 42°C in a water bath (MultiTempIII, Amersham Pharmacia Biotech AB, Sweden) for exactly 30 seconds. The reaction was then removed and placed on ice immediately and incubated for another 2 minutes. After adding 250 µl of LB medium without ampicillin, the transformed bacterial cells were incubated at 37 °C for 60 minutes by 225 rpm on an

Innova®40 thermoshaker (New Brunswick Scientific; USA) and then plated on LB-agar plates containing ampicillin.

4.2.7 Blue/White colony screening of recombinant pCR2.1 clones

The competent E. coli cells (Top10F') carry a mutation in their LacZ gene (LacZΔM15). The LacZ gene encodes β-galactosidase (β-gal), which enables the bacteria to hydrolyze lactose into glucose and galactose. The deletion of the amino acids 11-41 of β-gal (LacZΔM15 mutation) disables the protein to form its active form, a tetramer, which renders it nonfunctional. The function can be complemented by plasmids expressing amino acids 1-59 (α-peptide) of ß-gal.

LB-agar plates containing 75µl/ml of ampicillin were gently dried under the constant sterile-air flow. Then 40µl of a 5-bromo-4-chloro-indolyl-β-D-galactopyranoside (X-Gal) stock solution (20mg/ml) and 40 µl isopropyl β-D-1-thiogalactopyranoside (IPTG) (10 mM) were evenly spread on the plates. X-Gal is a colorless analogue of lactose and can be hydrolyzed by β-gal into galactose and 5-bromo-4-chloro-3-hydroxyindole. 5-bromo-4-chloro-3-hydroxyindole forms dimers which are oxidized into 5,5'-dibromo-4,4'-dichloro-indigo, an intensively blue substance. The plates were exposed to sterile-air flow for 30 minutes to dry. The expression of β-gal in bacteria was enhanced by the addition of IPTG.

Two X-Gal/IPTG plates were used for each transformation reaction. 50 µl of the transformation reaction were spread on the first plate, the remaining 250 µl were spread on the second plate. The plates were incubated over night at 37 °C. Successfully transformed bacterial cells carrying the pCR2.1 vector are resistant against ampicillin and form bacterial colonies on the plate. White colored colonies indicate that the region encoding for the α-peptide of the vector is disrupted by the subcloned insert and that the β-gal of the E. coli cells remains non-functional.

Three to six white colonies were picked using the tip of a pipette and transferred into a 15ml falcon tube containing 5 ml of LB medium containing

ampicillin (75µl/ml). The cells were incubated at 37 °C with 225 rpm on an Innova®40 thermoshaker (New Brunswick Scientific; USA) over night.

4.2.8 Plasmid DNA Mini isolation of pCR2.1 and pGL4.23 construct

Isolation of plasmid DNA was performed using the ZR Plasmid Miniprep™ - Classic Kit (Zymo Research Corp.; USA). For cryo conservation and reinoculation, 500 µl from the transformed E. coli suspension were transferred into a 1,5 ml Eppendorf tube and stored at 4°C. The remaining bacterial suspension was centrifugated at 4000 rpm for 30 minutes at 4°C (Centrifuge 5810R, Eppendorf; Germany).

The supernatant was discarded and the pellet re-suspended entirely in 200 µl P1 buffer. The prepared solution was transferred into a clear 1.5 ml Eppendorf tube. Consequently 200 µl of the P2 buffer was added and mixed by inverting the tube 3 times. Then 400 µl of P3 buffer was added and the obtained solution was gently mixed by inverting the tube 5 times. The mixture was incubated for 2 minutes at room temperature. Then, the sample was centrifuged (Biofuge Pico, Heraeus; Germany) for 2 minutes at 13 000 rpm.

A Zymo-Spin™ IIN column was placed in a collection tube and the supernatant was transferred into the column. The Zymo-Spin™ IIN/Collection tube assembly was centrifuged for 30 seconds at 13 000 rpm. The flow-through was discarded and 200 µl of the Endo-Wash buffer was added to the column followed by centrifugation for 1 minute at 13 000 rpm. Than 400 µl of the Plasmid Wash Buffer was added to the column and centrifuged for 30 seconds at 13 000 rpm. The Zymo-Spin™ IIN column was then placed into a clean 1.5 ml safety-lock Eppendorf tube and the plasmid DNA was eluted with 30 µl of water by centrifugation for 30 seconds at 13 000 rpm.

4.2.9 Restriction digests of pCR2.1 and pGL4.23 vector constructs

To confirm the presence of subcloned inserts in pCR2.1 and pGL4.23 vectors, the isolated plasmids were digested with various restriction endonucleases to discriminate the size, orientation and number of the inserted sequences.

The digest reactions were conducted in 15 µl volumes. Each reaction consisted of 2 µl plasmid DNA, 0.5 µl of restriction enzyme, 1.5 µl 10x restriction buffer and was filled to its end volume with water. The mixtures were incubated for 15 minutes at 37°C and the length of produced fragments was estimated via gel electrophoresis in a 1% TAE gel.

4.2.10 Preparative digests of pGL4.23 luciferase reporter and pCR2.1 vectors containing subcloned inserts

The digest reactions were performed in 40 µl volumes. 4 µl to 27 µl of the plasmid DNA (2-4 µg), were digested with 3 µl Sac I and then with 3 µl Xho I in the presence of 4 µl 10x Buffer. Each reaction was filled up to a total volume of 40 µl with water. The mixtures were incubated for 15 minutes at 37°C. The digested plasmids were then separated on a 1% TAE agarose gel by gel electrophoresis.

4.2.11 Isolation of subcloned DNA fragments

The separated vector and insert fragments were excised from the agarose gel with a scalpel and placed into a clean 1.5 ml Eppendorf tube. Excised agarose slices containing the separated vector/inserts were weighted and 3 times the volume of ADB buffer solution was added (Zymoclean™ Gel DNA Recovery Kit). The tubes were placed in a water bath at 50°C and incubated for 5-10 minutes, until the gel slice was completely dissolved.

The agarose gel solution was transferred to a Zymo-Spin™ Column in a collection tube and centrifuged at 13 000 rpm for 60 seconds. The obtained flow-through was discarded. Then, 200 µl of DNA wash buffer were added to the columns and centrifuged for 30 seconds again. The washing step was repeated once. The

column was transferred into a 1.5 ml tube and 6µl of water was added. The DNA was eluted by centrifugation for 60 seconds.

4.2.12 Ligation of isolated DNA fragments to pGL4.23 luciferase reporter vector

The ligations with pGL4.23 were performed in 10 µl reaction volumes. Approximately 30-40 ng of the insert were ligated to 60 ng of linearized pGL4.23 vector. Then, 1 µl ligase, 1 µl 10x ligase buffer and 1-7 µl water were added to the mixture.

First, the vector, insert and water were mixed together and incubated for 5 minutes at 45 °C. The solution was then transferred on ice. Finally, 10x ligase buffer and ligase were added. The mixture was incubated at 14 °C overnight.

4.2.13 Transformation of competent E. coli cells with ligation products of pGL4.23 vector

The transformations were conducted in 52 µl volumes consisting of 2 µl of the ligation product and 50 µl competent E. coli cells. The ligation product was added to 50 µl of competent cells, incubated on ice for 20 min and transformed by a heat shock in a water bath at 42 °C for exactly 30 seconds. The reaction was then placed on ice immediately and incubated for 2 minutes. Then, 250 µl LB medium were added to the suspension. The mixture was incubated at 37 °C for 1 hour by 225 rpm.

4.2.14 Plasmid DNA Maxi isolation of pGL4.23 constructs

The EndoFree Plasmid Maxi Kit (Qiagen) was used for the endotoxin-free isolation of the recombinant pGL4.23 reporter plasmids. For this, 500 µl of the respective E. coli cultures, confirmed by restriction analysis of the isolated plasmids described above were used to inoculate a 100 ml LB medium containing 75 µl/ml of ampicillin in an Erlenmeyer flask. The suspension was incubated for 12-17 hours at

37°C and 225 rpm. The cell suspension was split into two 50 ml falcon flask and the bacterial cells were centrifugated at 4000 rpm at 4°C for 30 minutes.

The supernatant was discarded, the bacteria pellets were thoroughly resuspended in 5 ml P1 buffer and transferred to a 50 ml falcon flask. Then, 10 ml of P2 buffer was added and the solution was mixed gently by inverting 5 times and vortexing once for few seconds. Then, the mixture was incubated at room temperature for 5 minutes. Subsequently, 10 ml of P3 buffer were added and the bacterial lysate was mixed immediately by inverting 5 times and vortexed for 5 seconds once more. The lysate was centrifuged for 30 minutes at 4000 rpm.

The solution was then transferred to a clean 50 ml falcon tube without disturbing the thick layer located on its surface. After addition of 2.5 ml endotoxin removal (ER) buffer, the lysate was mixed by inverting and kept on ice for 30 minutes on a rotating platform.

A Qiagen-tip 500 was equilibrated with 10 ml QBT buffer and the column was emptied by gravity flow. Then, the lysate was placed to the Qiagen-tip 500 and allowed to flow through the column by gravity flow. After washing the Qiagen-tip two times with 30 ml QC buffer, the DNA was eluted from the Qiagen-tip in an ultracentrifuge tube using 15 ml QN buffer.

The DNA was precipitated by adding 11 ml room-temperature isopropanol. The solution was immediately mixed and centrifuged at 15.000 x g for 30 minutes at 4°C. The supernatant was carefully decanted without disturbing the DNA pellet.

The DNA pellet was washed with 5 ml room-temperatured 70% ethanol and centrifuged at 15,000 x g for 5-10 minutes. The supernatant was carefully removed. The DNA pellet was airdried for 5-8 minutes and dissolved in 200-400 µl TE buffer.

4.2.15 Cultivation of Wild Type and ATXN3-knockout Human Neuroblastoma Cells

The human neuroblastoma cells (SH-SY5Y) were cultured in DMEM/F-12, GlutaMAX (Gibco; UK) supplemented with 10% active FBS (Biochrom GmbH; Germany), 1% nonessential amino acids (100mM, MEM NEAA 100x, Gibco; UK), 10,000 U/ml of penicillin and 10,000 µg/ml streptomycin (PenStrep, Gibco; UK) and 1% sodium pyruvate (Gibco;UK). The mediums were stored at 4°C.

The cell lines were cultivated in cell culture flasks with a culture area of 150 cm^2 (Cellstar®Cellculture Flask, Greiner Bio-One GmbH; Germany) in an incubator (Hera Cell, Heraeus Instruments; Germany) at 37 °C and 5% CO2 concentration. For splitting of the cells, the medium was removed and the cells were washed once with 10 to 15 ml of PBS (Dulbecco, Biochrom GmbH; Germany).

The cells were incubated with 1 ml of 5% Trypsin-EDTA (10x) (Gibco; UK) for 3-5 minutes at 37°C to dissociate the cells from the culture flask surface. The unattached cells were resuspended in 9 ml medium. The wildtype SH-SY5Y cells were split at a dilution of 1:5 to 1:10 depending on the cell concentration two to three times a week. The ATXN3-KO cells previously generated by CRISPR-Cas9 mediated gene knockout of ATXN3 in the parental wildtype SH-SY5Y cells were subcultured at a dilution of 1:2. The cell splitting was done under sterile conditions in a laminar flow (Hera safe, Heraeus Instruments; Germany).

4.2.16 Transfection of Wild Type Neuroblastoma Cells (SH-WT)

The SH-WT cells were seeded in 24-well tissue culture plates (Cellstar, Greiner Bio-One; Germany) at a density of 400,000 cells/well 24 hours prior to transfection. The cell counting was performed with a Neubauer chamber.

For transfection, two separate solutions were prepared. Solution A contained 6 µl of the transfection reagent Lipofectamine 2000 (Invitrogen; USA) and 200 µl of Opti-Mem (Gibco; UK). The solution was mixed and incubated for at least 5 minutes at room temperature.

Solution B contained 200 µl Opti-Mem, 2µl of the respective pGL4 luciferase reporter construct (1 µg/µl), 2 µl of the actin gene-promoter-driven co-reporter *Renilla* luciferase (pRL-Actin; 10 ng/µl) and 1 µl of GFP-expression construct (1µg/µl) or 1 µl GFP-ATXN3-expression construct (1µg/µl). Solution A was added to Solution B and incubated for 20 minutes at room temperature.

The medium was removed from the tissue culture plates and approximately 100µl of the transfection mixture was added into each well. Each transfection was carried out four times. After two hours incubation at 37°C, 500 µl medium was added to each well and the plates were incubated for additional 24 hours.

4.2.17 Transfection of ATXN3-KO Neuroblastoma Cells (SH-KO)

The SH-KO cells were seeded in the first three columns of 24-well tissue culture plates (Cellstar, Greiner Bio-One; Germany) at a density of 200,000 cells/well and SH-WT at the same density were seeded at remaining three columns of the plate, 24 hours prior to transfection. The cell counting was performed with a Neubauer Chamber.

On the day of transfection two separate solutions were prepared. Solution A contained 9 µl of the transfection reagent Lipofectamine 2000 (Invitrogen; USA) and 300 µl of Opti-Mem (Gibco; UK). The solution was mixed and incubated for at least 5 minutes at room temperature.

Solution B contained 300 µl Opti-Mem, 4.8 µl of the respective pGL4 luciferase reporter construct (1 µg/µl) and 3 µl of the co-reporter pRL-Actin (10 ng/µl). Solution A was added to Solution B and incubated for 20 minutes at room temperature.

The medium was removed from the tissue culture plates and approximately 100 µl of the transfection mixture was added into each well. Each transfection was carried out three times. After two hours incubation, 500 µl medium was added to each well and the plates were incubated for another 24 hours.

4.2.18 Lysis of transfected cells

After removal of the medium, cells were washed once with 500µl PBS per well and lysed by adding 50 µl passive lysis buffer (Promega, Madison, USA) to each well. The plates were incubated for 20 minutes at room temperature on a shaker (KS260 Basic; IKA®; Germany). The cell lysates were placed on ice and transferred on 96-well tissue culture plates (Sarstetedt Inc.; USA) and either used directly for luciferase assay or stored at -20 °C.

4.2.19 Luciferase reporter assay

For measurement of luciferase activity, white 96-well plates were loaded with 10 µl of each lysate sample per well using luciferin as substrate. In addition, for normalization of transfection efficiency, 10 µl of each lysate sample in a separate well was also measured for the *Renilla* luciferase activity using coelenterazine as substrate (co-reporter pRL-Actin under the control of a human ß-actin promoter was always co-transfected). The luciferase activities were measured in a microplate luminometer (CentroLB960, Berthold Technologies; Germany). The reader was programmed (MicroWin2000) to inject each 40 µl of *Renilla*/ F*irefly* substrates in each well and measure the luminescence signal for 10 seconds. Signal values were exported to and analyzed in MS Excel and GraphPad Prism.

4.2.20 Statistical analysis

For statistical analysis was used GraphPad Prism 7.0 software (GraphPad Software, Inc., San Diego, California, USA). Statistical significance was analyzed using paired t-test or one-way ANOVA. A P value < 0.05 was considered statistically significant. Data are presented as mean ± standard deviation (SD).

4.3 Results

4.3.1 ChIP-PCR optimization

For PCR amplification of specific ATXN3-enriched genomic regions previously identified by ChIP sequencing, the reaction conditions for each primer pair used were optimized by using various annealing temperatures varying Mg^{2+} concentrations and occasionally by the addition of betaine and/or DMSO as is illustrated on **Figure 1**. Optimized conditions are shown in **Table 3**.

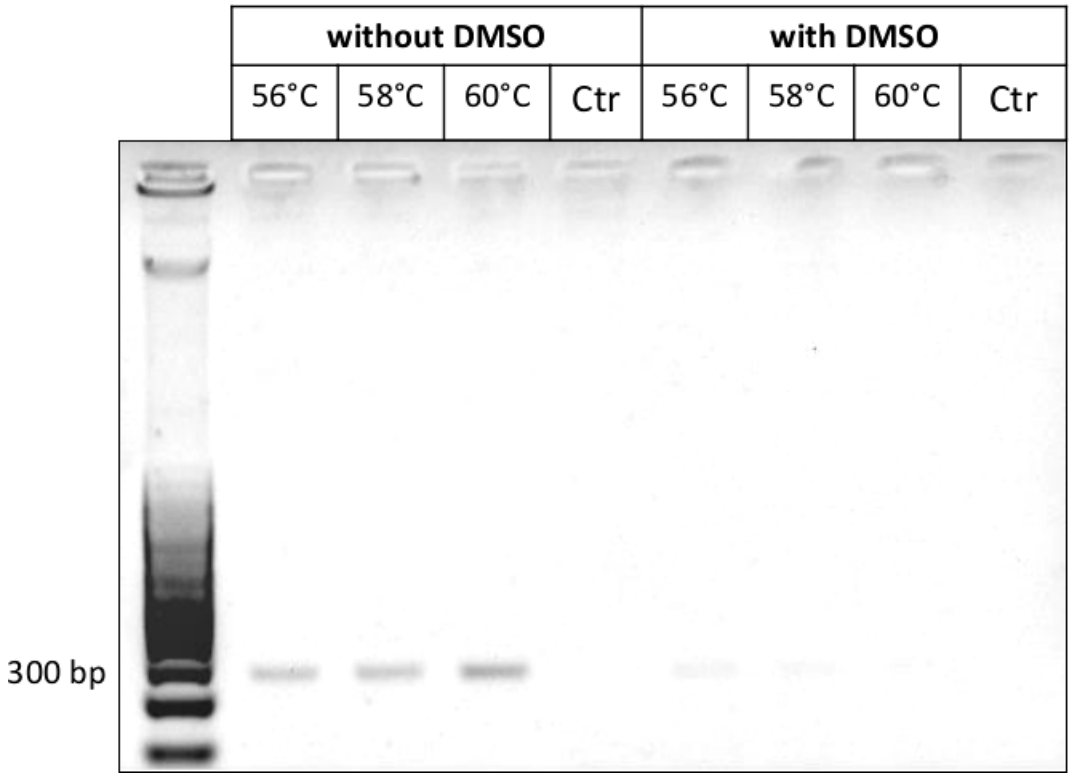

Figure 1 Analysis of PCR products containing 239 bp of the CEBPD-I(1) gene region by agarose gel electrophoresis. *Optimization was performed using a temperature gradient (56°C, 58°C, 60°C) with and without DMSO (10%). 1,5% TAE agarose gel, 180V for 30 minutes, Gel-Red, Template: chromatin DNA of SCA3 patient FS11; Ctr without DNA, 10 µl PCR product+ 2µl Loading buffer, 4 µl marker*

Table 3 Optimized conditions identified for the amplification of specific genomic regions

Gene region	Optimized temperature [°C]	Concentration of Mg2+ [mM]	Betaine [%]	Length of gene region [bp]
RRBP1(1)	62	1,5	-	185
SERPINF2(1)	52	1,5	-	185
ELAVL2	52	1,5	-	141
SEMA6A	54	1,5	-	133
THBS1(1)	52	1,5	-	231
CEBPD-I(1)	60	1,5	-	239
DOCK10	58	1,5	-	163
COL6A3	62	1,5	-	162
TOX3	56	1,5	10	133
ARRDC3	60	1,5	-	132
FOSL2	62	1,5	-	181
TLR4	54	2,5	-	137
KLF10	56	1,5	-	164
PER2	60	1,5	-	141
THBS1(2)	62	1,5	-	146
CEBPD-I(2)	58	2,5	-	124
RRBP1(2)_For RRBP1(1)_Rev	62	1,5	-	150
SERPINF2(2)_For SERPINF2(1)_Rev	60	1,5	-	172
SERPINF2(1)_For SERPINF2(2)_Rev	60	1,5	-	153
SERPINF2(2)	58	1,5	-	140
PTEN	58	1,5	-	117
CEBPD-III	62	1,5	10	148

4.3.2 Validation of ATXN3 enriched genomic regions by ChIP-PCR

For ChIP-PCR, chromatin DNA prepared from differentiated iPS-derived neurons of 2 SCA3 patients and 2 control subjects was immunoprecipitated using an ATXN3-specific antibody. The purified ATXN3-bound chromatin fragments were then used as DNA templates for the amplification of the 22 genomic regions (previously identified by ChIP sequencing) with each primer pair under optimized conditions (see Table 1). For some genomic regions, new primers and smaller-sized amplicons were chosen for specific amplification by PCR (see **Table 1**). Enrichment of ATXN3 was confirmed for genomic regions located within or closed to genes encoding: phosphatase and tensin homolog (PTEN), thrombospondin 1 (THBS1), serine protease inhibitor family F2 (SERPINF2), CCAAT/enhancer binding protein delta (CEBPD-I) and period circadian clock 2 (PER2) (**Figure 2**). Interestingly, the enrichment of ATXN3 to genomic regions of both CEBPD-I and THBS1 was significantly lower in SCA3 patient (#2) compared to control (#2) suggesting that mutant ATXN3 has probably altered or reduced binding properties in neurons of SCA3 patients.

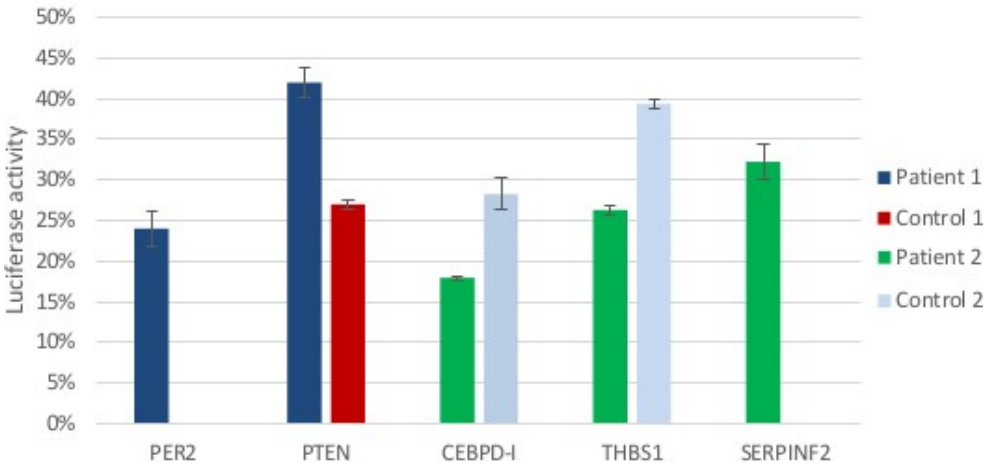

***Figure 2 Verified ATXN3-enriched genomic regions**. Quantification of ATXN3 enriched genomic regiones located within the CEBPD-I , PER2, PTEN, THBS1 and SERPINF2 gene analyzed by ChIP-PCR. The amount of immunoprecipitated chromatin DNA was normalized to the amount of input chromatin DNA determined by densitometric quantification of the generated PCR products. Values are expressed in percentage of the respective input signal (100%). The results are presented as the mean ± SD of three independent experiments (n=3)*

4.3.3 PCR optimization and amplification of selected genomic regions

For the confirmed ATXN3-bound regions, new primers including restriction sites (Sac I and Xho I) for subcloning in the pGL4 luciferase vector were designed and used to amplify genomic regions containing the respective ATXN3-enriched sequence together with 5' and 3' flanking regions. Optimization of PCR conditions was performed similarly as for the ChIP-PCRs, and is exemplarily shown in **Figure 3**. Genomic DNAs isolated from 3 different human cell lines, a neuronal stem cell line (NSC), tumor cell line (HeLa) and lymphoblastoid cell line (Lym). were used as templates. After determination of the optimal conditions (**Table 4**), the confirmed ATXN3-bound genomic regions were specifically amplified and used for cloning into pCR2.1 vector.

Table 4 Optimized conditions identified for the amplification of specific genomic regions used for subcloning into pCR2.1 and pGL4.23 vectors

Gene region	Optimized temperature [°C]	Length of genomic region [bp]
CEBPD-I	60	901
PER2	58	840
PTEN	62	858
SERPINF2	58	940
THBS1	62	968

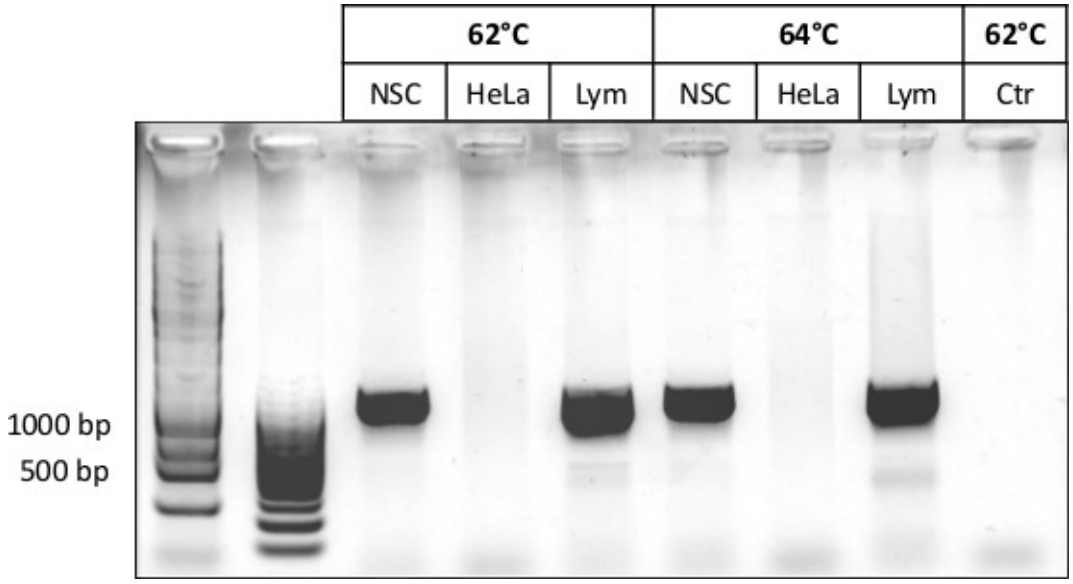

Figure 3 Analysis of PCR products containing 858 bp of the PTEN gene region by agarose gel electrophoresis. *Optimization was performed using different temperatures (62°C and 64°C) 1,0% TAE agarose gel, 180V for 30 minutes, GelRed, Template: Genomic DNA from a neuronal stem cell line (NSC), tumor cell line (HeLa) and lymphoblastoid cell line (Lym); Ctr without DNA, 10 µl PCR product+ 2µl Loading buffer, 4 µl Marker of each 100 bp and 1 kb ladder.*

4.3.4 Subcloning of PCR products into pCR2.1 vector and characterization of recombinant clones

The amplificated PCR products of each genomic region were inserted and ligated to the pCR 2.1 vector (see Methods). After transformation of competent E. coli cells with the pCR2.1 ligation products, ampicillin-resistant clones were grown on agar plates and recombinant clones were identified by Blue/White screening. Individual clones were picked and grown in small scale culture medium containing ampicillin and used for isolation of plasmid DNA.

To verify the presence, size and orientation of the subcloned inserts, plasmid DNA of several recombinant clones was isolated and digested with restriction enzymes Eco RI and SacI/XhoI. Single EcoRI digests of pCR 2.1 clones were used to discriminate the size of the subcloned inserts whereas double digests with SacI and XhoI were used to release the inserts for further subsequent forced cloning into pGL4 reporter vector also restricted with SacI and XhoI.

4.3.5 Cloning of genomic inserts into pGL 4.23 reporter vector and characterization of recombinant clones

The purified SacI/XhoI inserts of the respective pCR2.1 clones (4.3.4) were cloned into a SacI/XhoI-restricted pGL 4.23 reporter vector in front of the luciferase gene. After transformation of competent E. coli cells with the pGL4.23 ligation products, ampicillin-resistant clones were selected on agar plates. Individual clones were picked and propagated for the isolation of plasmid DNA. To confirm the presence and size of the cloned inserts, plasmid DNA isolated from several recombinant clones was digested with the restriction enzymes SacI and XhoI (**Figure 4**). Moreover, to assure that the subcloned inserts are present as single copies, additional restriction digests were performed, as listed in **Table 5**. The generated pGL4.23 constructs were verified by sequencing.

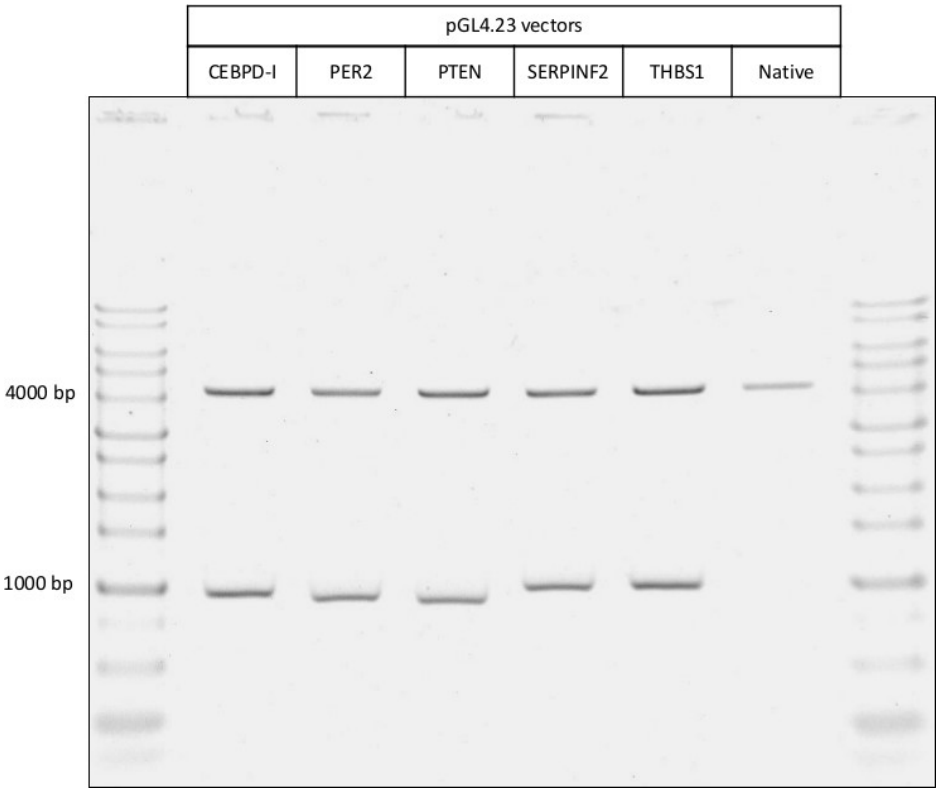

Figure 4 Restriction digests of generated plasmids using enzymes XhoI and SacI resulting in linearized pGL4 vector backbones and release of the subcloned inserts, *1,0% TAE agarose gel, 180V for 45 minutes, EtBr, 2µl Loading buffer, 4 µl Marker of 1 kb ladder.*

Table 5 Confirmation digest of cloned plasmids

Reporter construct	Restriction enzyme	Approximate length of expected fragments [bp]
pGL4.23-CEBPD-I	Pst I	3631 1485 and 62
pGL4.23-PER2	Apa I	4696 419
pGL4.23-PTEN	Hind III	4588 543
pGL4.23-SERPINF2	Bgl II	4785 430
pGL4.23-THBS1	Xba I	2656 2590

4.3.6 Regulatory activity of an ATXN3-bound, proximal region upstream of the CCAAT/enhancer binding protein delta (CEBPD) gene

In **Figure 5** the chromosomal localization of the ATXN3-enriched genomic region upstream of the CEBPD gene and the amplificated region cloned into the reporter vector pGL4.23 in front of a minimal promoter and a *Firefly* luciferase (luc2) is shown.

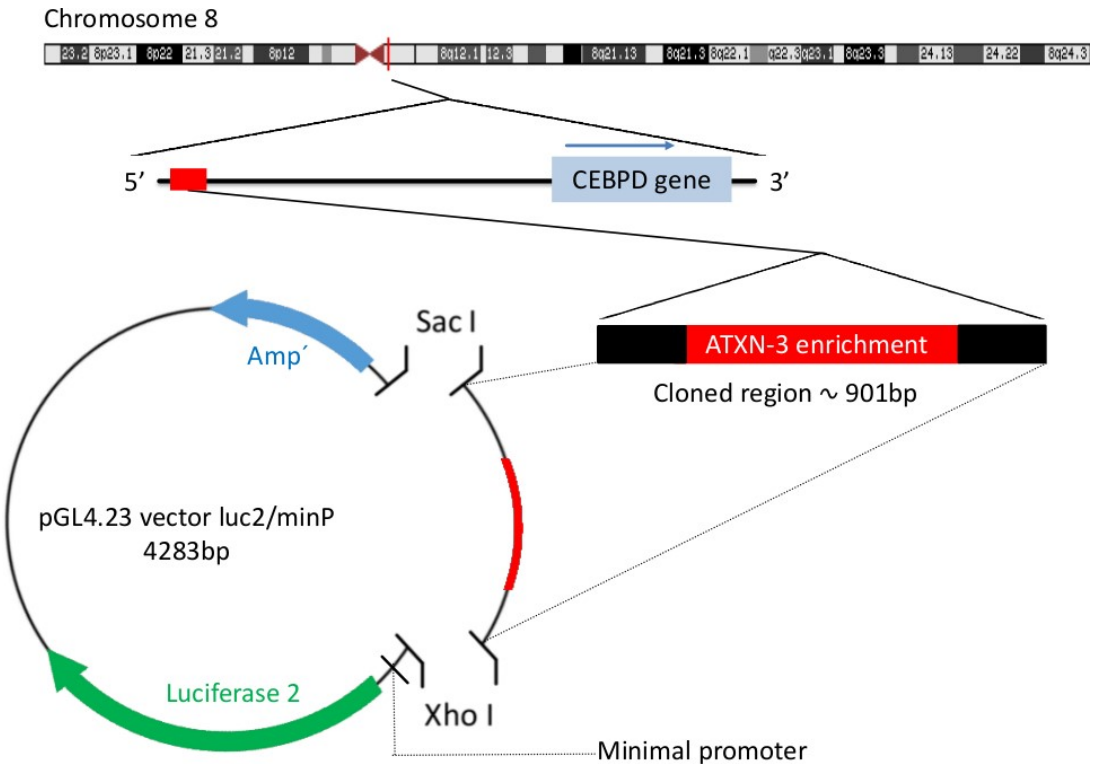

Figure 5 Chromosomal localization of the cloned CEBPD genomic region and construction of the pGL4-CEBPD-I reporter vector

4.3.6.1 Reporter activity of the pGL4-CEBPD-I construct in human neuroblastoma SH-SY5Y cells overexpressing normal and mutant ATXN3

To study the functional effects of ATXN3 on CEBPD gene transcription, SH-SY5Y cells were transfected with the pGL4 luciferase reporter containing the 901 bp genomic region upstream of the CEBPD gene and expression plasmids encoding green fluorescent protein (GFP) and GFP fusion proteins of human full length

normal (Q13) and mutant (Q77) ATXN3. Co-expression of normal ATXN3Q13 increased reporter activity by approximately 40% compared to controls transfected with a GFP expression plasmid only (**Figure 6**). A similar although lower increase in reporter activity (~25%) was observed when mutant ATXN3Q77 was co-expressed, pointing towards an enhancing effect of normal ATXN3 and a partially reduced activating capacity of mutant ATXN3 on this genomic region.

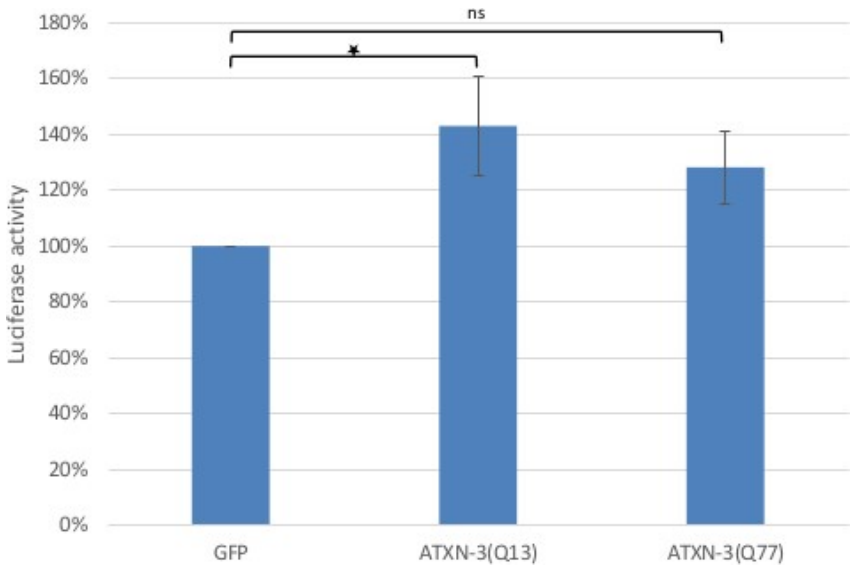

Figure 6 Mean Luciferase Activity (%) of pGL4-CEBPD-I construct in Human Neuroblastoma Cells. *The results are presented as the mean ± SD of four independent experiments (n=4). Data were statistically analysed using One-way ANOVA, P < 0.05*

4.3.6.2 Reporter activity of the pGL4-CEBPD-I construct in hemizygous ATXN3 Knockout human neuroblastoma SH-SY5Y cells

To further analyze potential regulatory effects mediated by ATXN3 on the subcloned genomic CEBPD-I region, a previously generated human neuroblastoma SH-SY5Y cell line exhibiting only one functional allele for the expression of ATXN3 (hemizygous knockout) were transfected with the pGL4-CEBPD-I luciferase reporter construct and compared to the luciferase activities obtained in wildtype SH-SY5Y cells.

In both wildtype and ATXN3 knockout SH-SY5Y cells, the pGL4-CEBPD-I reporter activity was strongly repressed to one-fifth of the activities obtained in cells transfected with the empty control pGL4 vector. Interestingly, ATXN3-KO cells (~20%) showed a less strong repression of the CEBPD-mediated reporter activity compared to wildtype cells (~10%) as shown in **Figure 7**, suggesting that the ATXN3-mediated repression is partially reduced in ATXN3-KO cells through the hemizygous knockout of endogenous ATXN3. The partial loss of repression in KO cells, however, indicates that endogenous ATXN3 suppresses rather than enhances the CEBPD-mediated activity found under ATXN3 overexpression conditions (4.3.6.1).

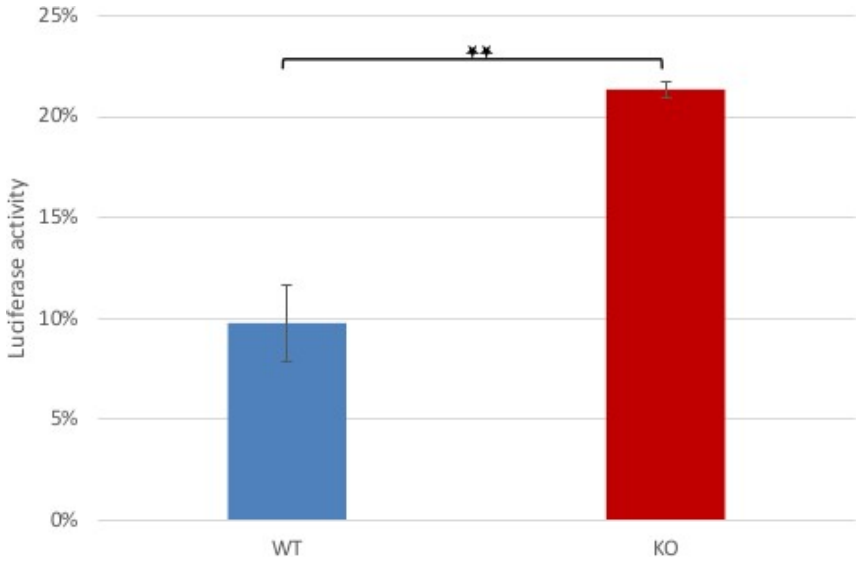

Figure 7 Mean Luciferase Activity (% of native pGL4 reporter activity) of pGL4-CEBPD-I construct in WT and hemizygous ATXN3 Knockout Human Neuroblastoma Cells. *The results are presented as the mean ± SD of three independent experiments (n=3). Data were statistically analysed using paired t-test, $P < 0.05$*

4.3.7 Regulatory activity of an ATXN3-bound, proximal region upstream of the period circadian clock-2 (PER2) gene

In **Figure 8** the chromosomal localization of the ATXN3-enriched genomic region upstream of the PER2 gene and the amplificated region cloned into the reporter vector pGL4.23 in front of a minimal promoter and a *Firefly* luciferase (luc2) is shown.

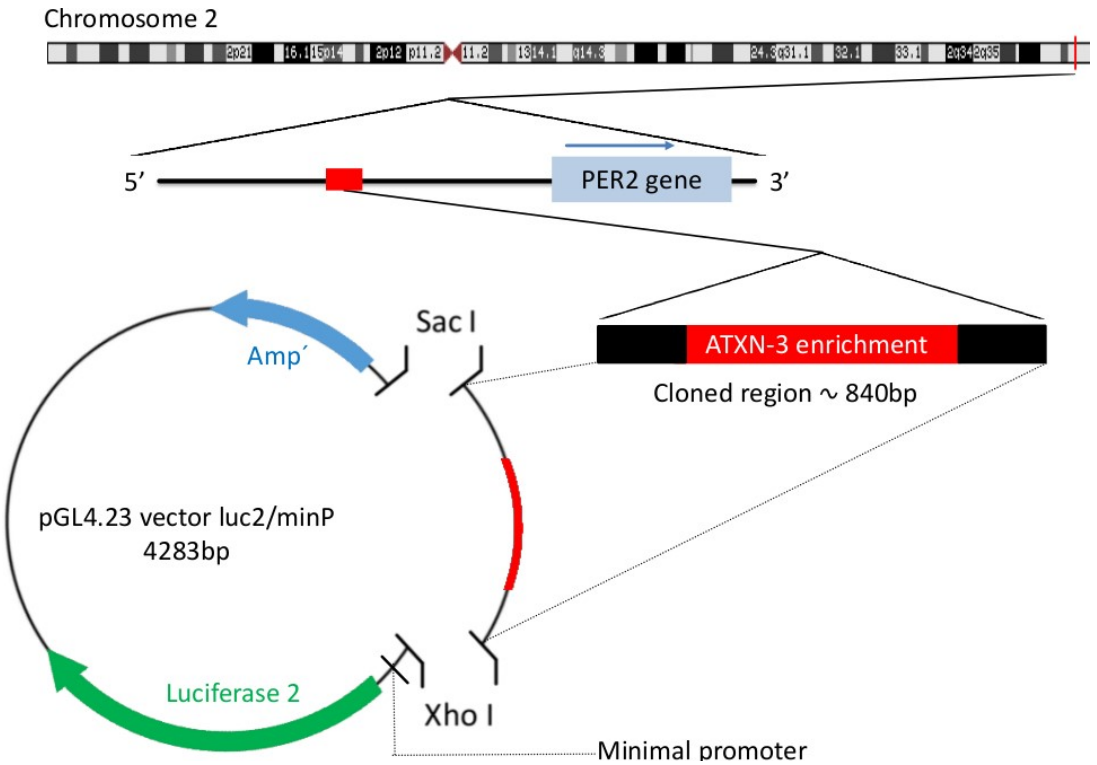

Figure 8 Chromosomal localization of the cloned PER2 genomic region and construction of the pGL4-PER2 reporter vector

4.3.7.1 Reporter activity of pGL4-PER2 construct in human neuroblastoma SH-SY5Y cells overexpressing normal and mutant ATXN3

To further analyze a possible regulatory effect mediated by ATXN3 on the subcloned PER2 genomic region in vitro, the pGL4-PER2 luciferase reporter construct was co-transfected with expression plasmids encoding green fluorescent protein (GFP) and GFP fusion proteins of human full length normal (Q13) and mutant (Q77) ATXN3 into the human neuroblastoma cell line SH-SY5Y.

The genomic region linked to PER2 slightly increased luciferase activity in response to co-expression of normal ATXN-3 whereas the activity mildly decreased when mutant ATXN-3 was co-expressed (**Figure 9**). The small changes in luciferase activity found for the PER2-associated region gene are unlikely to account for a substantial ATXN3-mediated regulatory effect under the conditions analyzed.

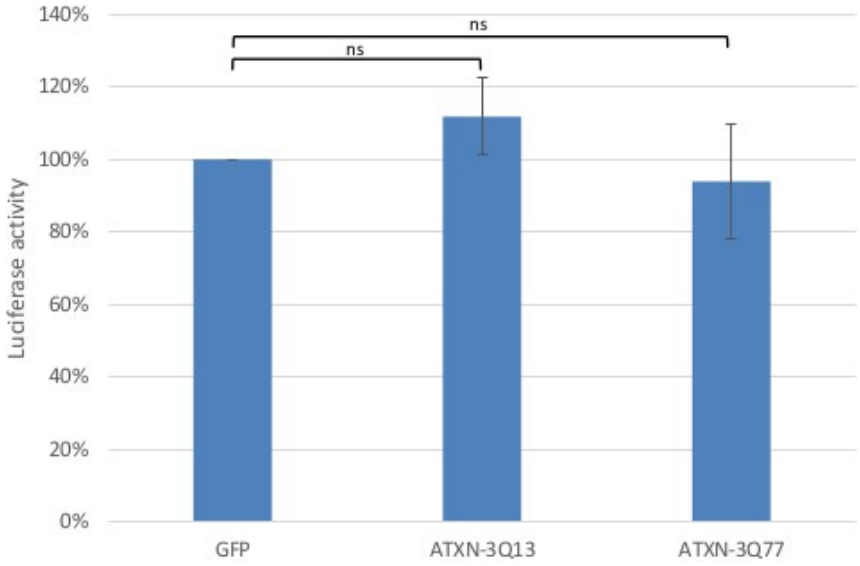

Figure 9 Mean Luciferase Activity (%) of pGL4-PER2 construct in Human Neuroblastoma Cells. *The results are presented as the mean ± SD of four independent experiments (n=4). Data were statistically analysed using One-way ANOVA, $P < 0.05$*

4.3.7.2 Reporter activity of pGL4-PER2 construct in hemizygous ATXN3 KO human neuroblastoma cells

To study a putative regulatory effect of ATXN3 on the PER2 subcloned region., ATXN3 knock-out and wildtype SH-SY5Y cells were transfected with pGL4-PER2 luciferase reporter construct. In both wildtype and ATXN3 knockout SH-SY5Y cells, the pGL4-PER2 reporter activity was strongly repressed to approximately 15% of the activities obtained in cells transfected with the empty control pGL4 vector (**Figure 10**).

However, the PER2 gene-linked genomic region showed no differences in luciferase activity between ATXN3 KO and WT cells, suggesting that repression of the subcloned region is not ATXN3-dependent under the conditions analyzed.

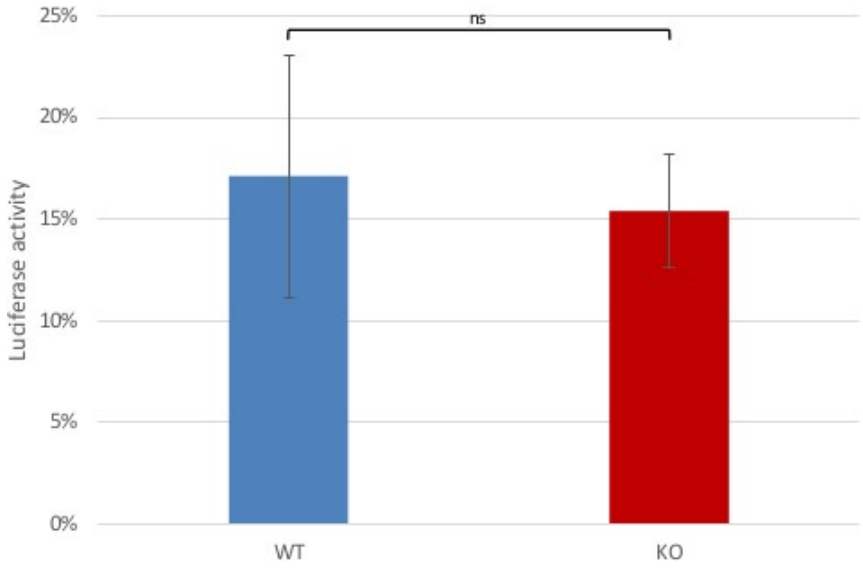

Figure 10 Mean Luciferase Activity (% of native pGL4 vector activity) of pGL4-PER2 construct in WT and heterozygous ATXN3-knockout Human Neuroblastoma Cells. *The results are presented as the mean ± SD of three independent experiments (n=3). Data were statistically analysed using paired t-test, $P < 0.05$*

4.3.8 Regulatory activity of an ATXN3-bound, intronic region within the phosphatase and tensin homolog (PTEN) gene

To study the regulatory activity of the ATXN3-enriched proximal PTEN region, a reporter construct was generated by PCR amplification of a 858 bp genomic fragment from intron 2 of the PTEN gene (with total 9 exons) and insertion into the reporter vector pGL4.23 upstream of a minimal promoter and a *Firefly* luciferase (luc2) gene (**Figure 11**).

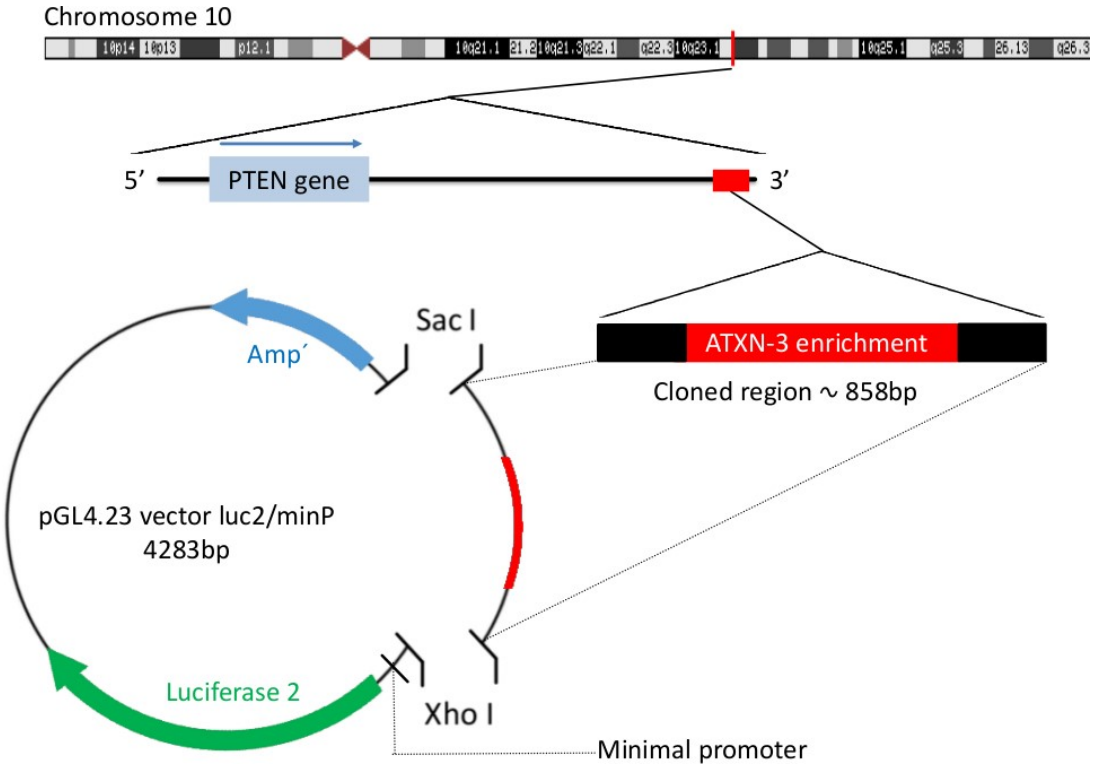

Figure 11 Chromosomal localization of the cloned PTEN genomic region and construction of the pGL4-PTEN reporter vector

4.3.8.1 Reporter activity of pGL4-PTEN construct in human neuroblastoma SH-SY5Y cells overexpressing normal and mutant ATXN3

To further analyze a possible regulatory effect mediated by ATXN3 on the subcloned PTEN genomic region in vitro, the pGL4-PTEN luciferase reporter construct was co-transfected with expression plasmids encoding green fluorescent protein (GFP) and GFP fusion proteins of human full length normal (Q13) and mutant (Q77) ATXN3 into the human neuroblastoma cell line SH-SY5Y.

The gene region linked to PTEN showed a slightly increased luciferase activity in response to co-expression of normal ATXN-3. Similarly. a small increase in luciferase activity was found when mutant ATXN-3 was co-expressed (**Figure 12**). Thus, a regulatory effect mediated by ATXN3 on this genomic region under the conditions analyzed is unlikely.

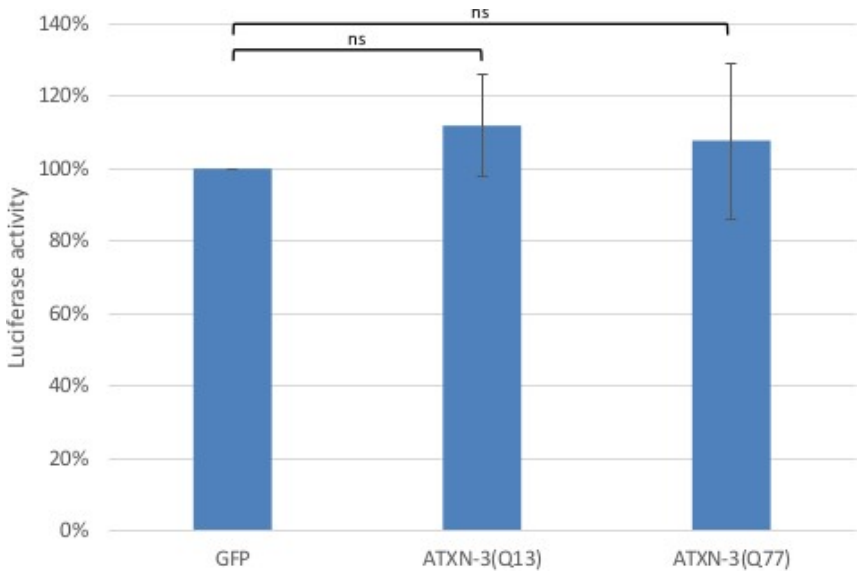

Figure 12 Mean Luciferase Activity (%) of pGL4-PTEN construct in Human Neuroblastoma Cells. *The results are presented as the mean ± SD of four independent experiments (n=4). Data were statistically analysed using One-way ANOVA, P < 0.05*

4.3.8.2 Reporter activity of pGL4-PTEN construct in hemizygous ATXN3 KO human neuroblastoma cells

To investigate a putative regulatory effect of ATXN3 on the PTEN subcloned region, ATXN3 knock-out human neuroblastoma SH-SY5Y and wildtype SH-SY5Y cells were transfected with the pGL4-PTEN luciferase reporter construct. In both wildtype and ATXN3 knockout SH-SY5Y cells, the pGL4-PTEN reporter activity was slightly repressed to ~70% in WT and 80% in KO cells compared to the activities obtained in cells transfected with the empty control pGL4 vector (**Figure 13**).

However, no significant differences in luciferase activity of the PTEN-linked region were found in ATXN3-KO and WT cells (**Figure 13**), suggesting no ATXN3-dependent regulatory activity of the subcloned genomic region.

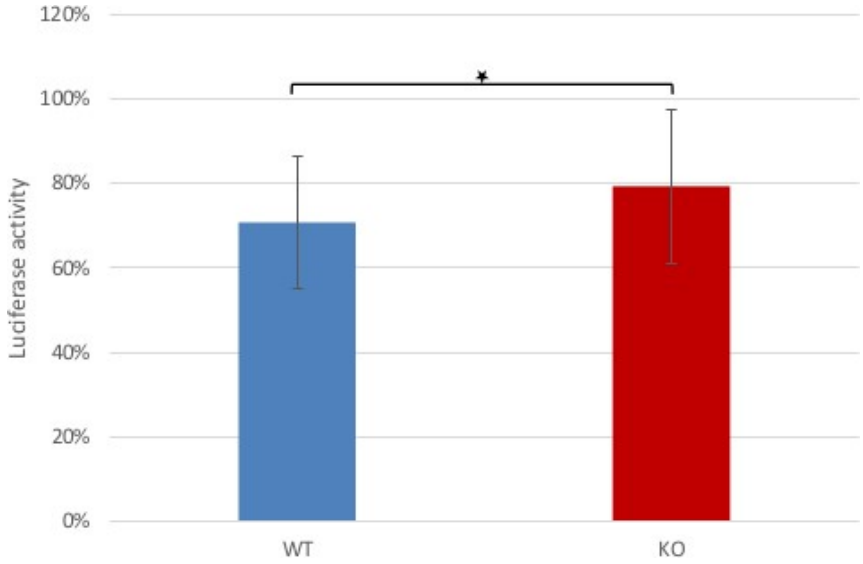

Figure 13 Mean Luciferase Activity (% of native pGL4 vector activity) of pGL4-PTEN construct in WT and heterozygous ATXN3-knockout Human Neuroblastoma Cells. *The results are presented as the mean ± SD of three independent experiments (n=3). Data were statistically analysed using paired t-test, P < 0.05*

4.3.9 Regulatory activity of an ATXN3-bound, intronic region within the serine protease inhibitor family F2 (SERPINF2) gene

To study the regulatory activity of the ATXN3-enriched proximal SERPINF2 region, a reporter construct was generated by PCR amplification of a 940 bp genomic fragment from intron 5 (with total 10 exons) and insertion into the reporter vector pGL4.23 upstream of a minimal promoter and a *Firefly* luciferase (luc2) gene (**Figure 14**)

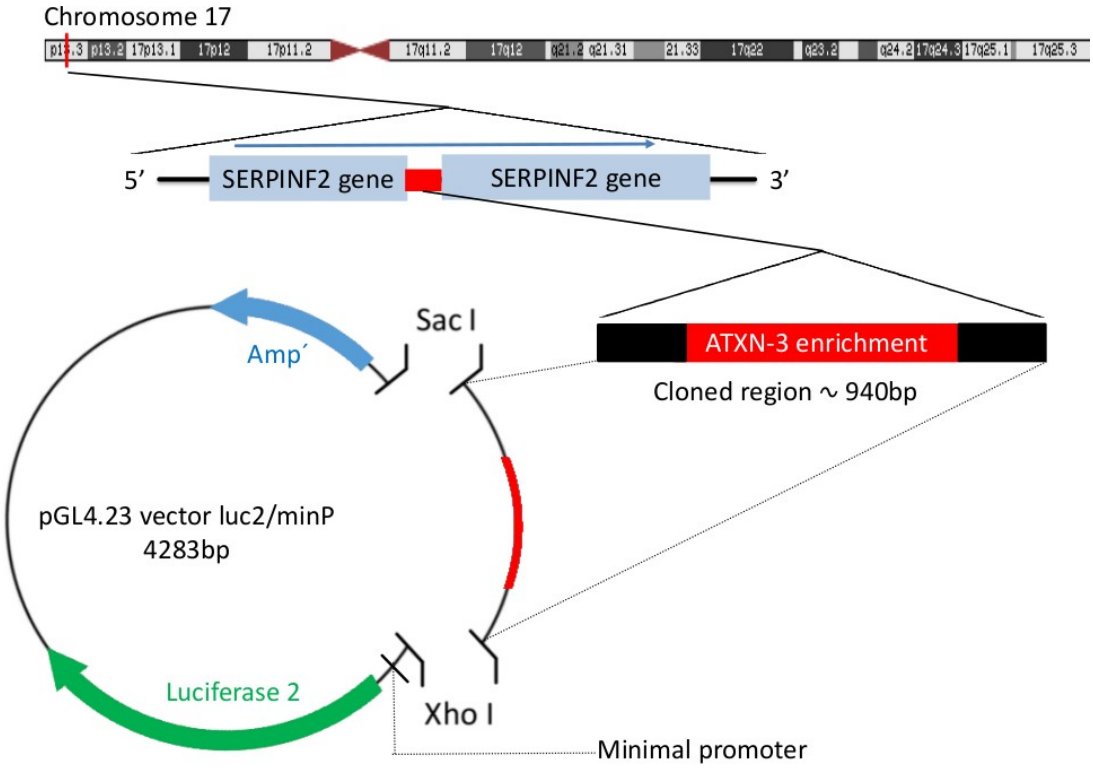

Figure 14 Chromosomal localization of the cloned SERPINF2 genomic region and construction of the pGL4-SERPINF2 reporter vector

4.3.9.1 Reporter activity of pGL4-SERPINF2 construct in human neuroblastoma SH-SY5Y cells overexpressing normal and mutant ATXN3

To analyze a possible regulatory effect mediated by ATXN3 on the subcloned SERPINF2 genomic region in vitro, the pGL4-SERPINF2 luciferase reporter construct was co-transfected with expression plasmids encoding green fluorescent protein (GFP) and GFP fusion proteins of human full length normal (Q13) and mutant (Q77) ATXN3 into the human neuroblastoma cell line SH-SY5Y.

The gene region linked to SERPINF2 slightly decreased activity in response to co-expression of normal ATXN-3. This tendency increased even more when mutant ATXN-3 was co-expressed (**Figure 15**) pointing toward a very weak ATXN3-dependent repressive effect of the subcloned genomic region under the conditions analyzed.

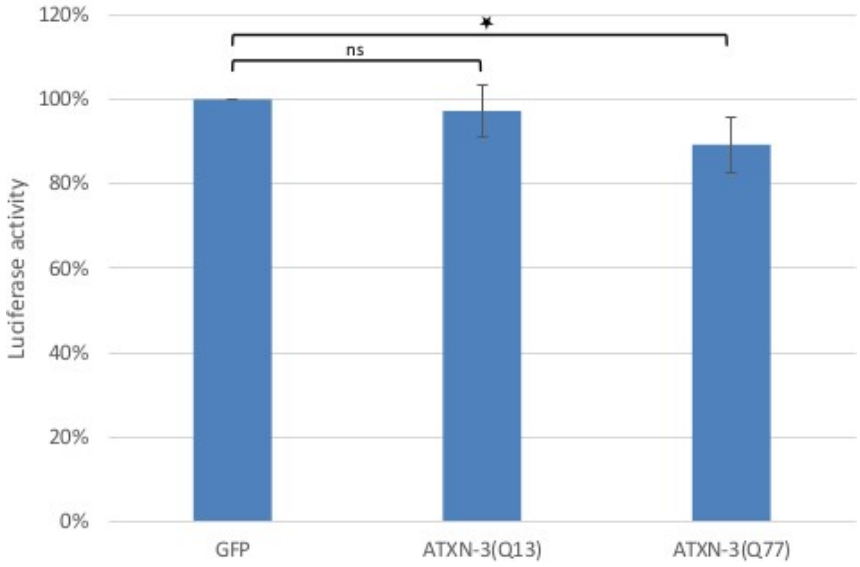

Figure 15 Mean Luciferase Activity (%) of pGL4-SERPINF2 construct in Human Neuroblastoma Cells. *The results are presented as the mean ± SD of four independent experiments (n=4). Data were statistically analysed using One-way ANOVA, P < 0.05*

4.3.9.2 Reporter activity of pGL4-SERPINF2 construct in hemizygous ATXN3 KO human neuroblastoma cells

A repressive effect was also found in ATXN3 knock-out human neuroblastoma wildtype SH-SY5Y cells transfected with pGL4-SERPINF2 luciferase reporter construct. In both wildtype and ATXN3 knockout SH-SY5Y cells, the pGL4-SERPINF2 reporter activity was strongly repressed to one third of the activities obtained in cells transfected with the empty control pGL4 vector (**Figure 16**).

No significant ATXN3-dependent differences for the SERPINF2-linked region were found between ATXN3 KO and WT cells except for a small reduction of the repressive effect in KO cells (**Figure 16**).

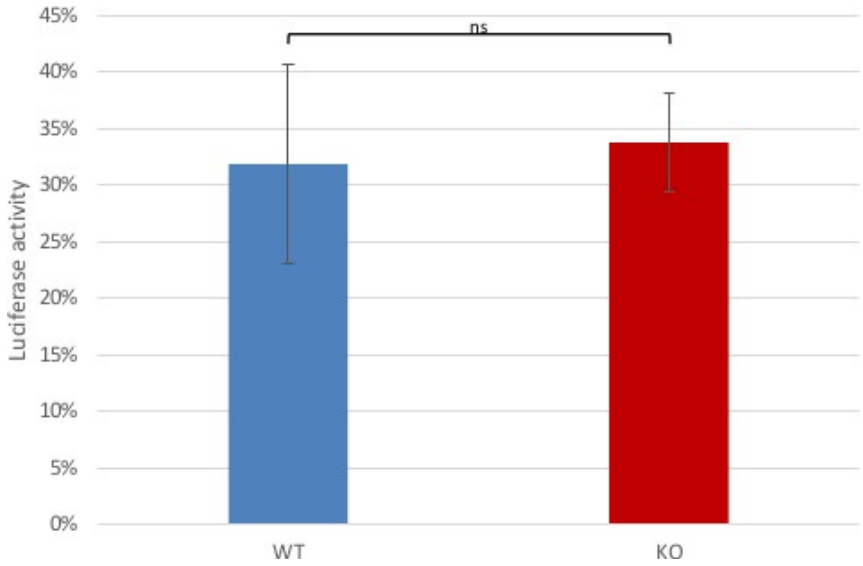

Figure 16 Mean Luciferase Activity (% of native pGL4 vector activity) of pGL4-SERPINF2 construct in WT and heterozygous ATXN3-knockout Human. *The results are presented as the mean ± SD of three independent experiments (n=3). Data were statistically analysed using paired t-test, $P < 0.05$*

4.3.10 Regulatory activity of an ATXN3-bound, proximal region downstream of the thrombospondin-1 (THBS1) gene

In **Figure 17** the chromosomal localization of the ATXN3-enriched genomic region upstream of the THBS1 gene and the amplificated region cloned into the reporter vector pGL4.23 in front of a minimal promoter and a *Firefly* luciferase (luc2) is shown.

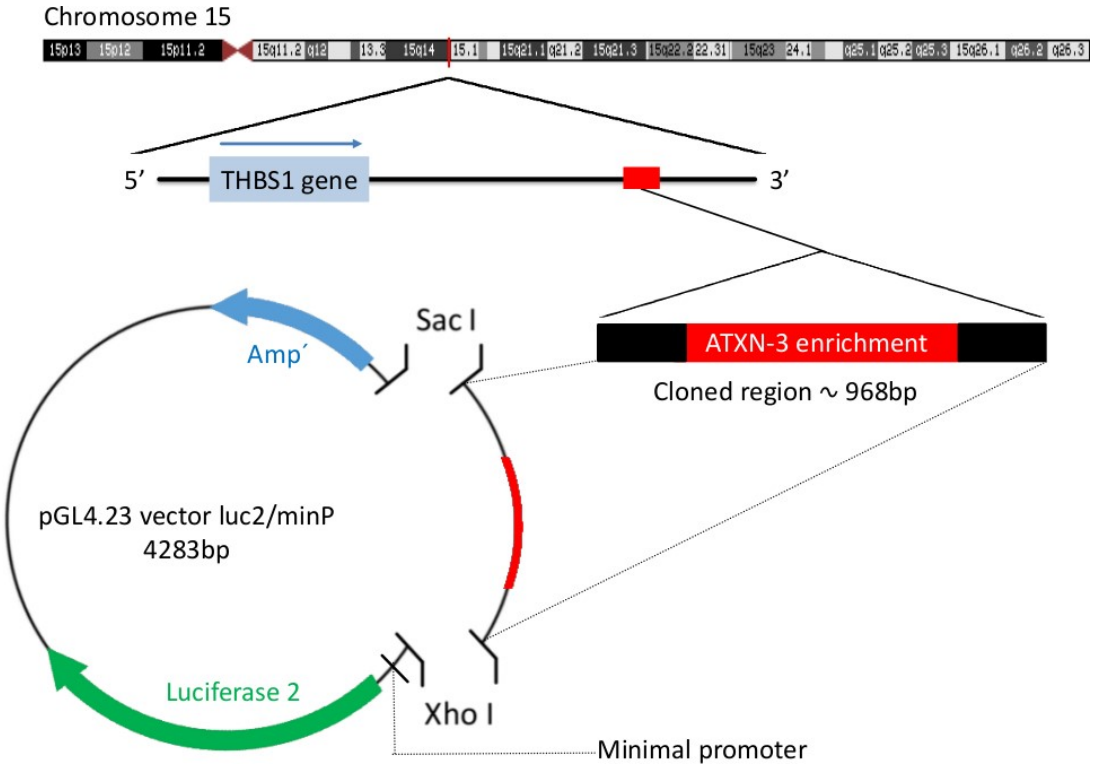

Figure 17 Chromosomal localization of the cloned THBS1 genomic region and construction of the pGL4-THBS1 reporter vector

4.3.10.1 Reporter activity of pGL4-THBS1 construct in human neuroblastoma SH-SY5Y cells overexpressing normal and mutant ATXN3

To analyze a possible regulatory effect mediated by ATXN3 on the subcloned THBS1 genomic region in vitro, the pGL4-THBS1 luciferase reporter construct was co-transfected with expression plasmids encoding green fluorescent protein (GFP) and GFP fusion proteins of human full length normal (Q13) and mutant (Q77) ATXN3 into the human neuroblastoma cell line SH-SY5Y.

The luciferase activity of the subcloned region linked to the THBS1 gene showed an increase of circa 20% in response to the co-expression of normal ATXN3. This tendency increased by another 20% when mutant ATXN3 was co-expressed (**Figure 18**), suggesting an ATXN3-dependent activation of the THBS1 genomic region that might be relevant for SCA3 pathogenesis.

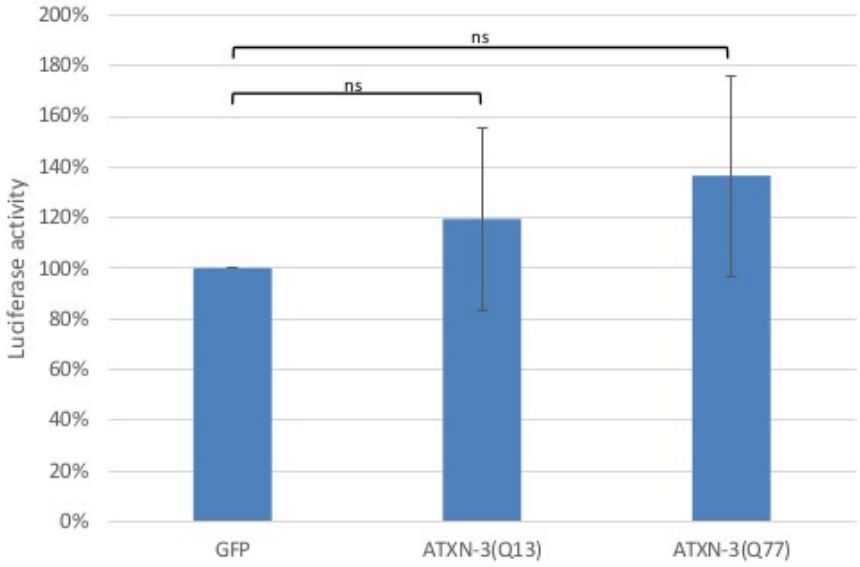

Figure 18 Mean Luciferase Activity (%) of pGL4-THBS1 construct in Human Neuroblastoma Cells. *The results are presented as the mean ± SD of four independent experiments (n=4). Data were statistically analysed using One-way ANOVA, P < 0.05*

4.3.10.2 Reporter activity of pGL4-THBS1 construct in ATXN3 KO human neuroblastoma cell line

The possible modulating effect of ATXN3 on the regulatory activity of the THBS1 subcloned genomic region was also analyzed in ATXN3 knock-out and wildtype human neuroblastoma SH-SY5Y cells line transfected with pGL4-THBS1 luciferase reporter construct. In both wildtype and ATXN3 knockout SH-SY5Y cells, the pGL4-THBS1 reporter activity was strongly repressed to approximately one-fifth of the activities obtained in cells transfected with the empty control pGL4 vector (**Figure 19**). Repression of luciferase activity was stronger in WT (10%) compared to KO (20%) cells suggesting that endogenous ATXN3 present in SH-SY5Y cells probably acts as a transcriptional repressor on the subcloned genomic element (**Figure 19**).

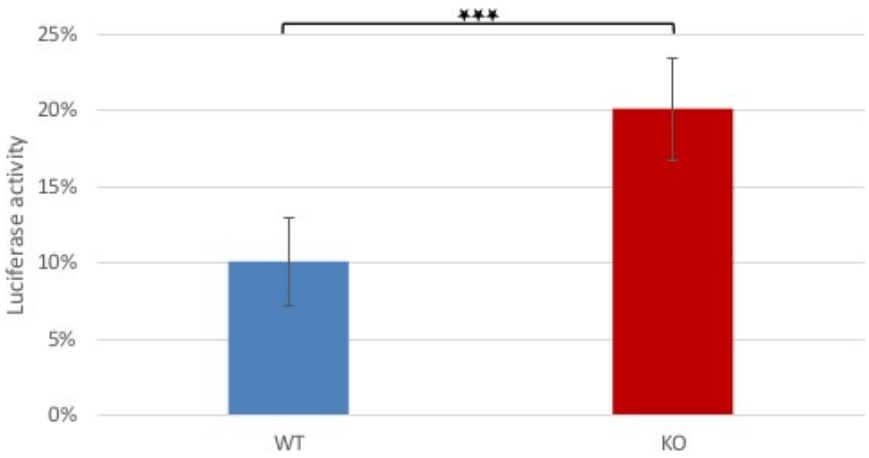

Figure 19 Mean Luciferase Activity (% of native pGL4 vector activity) of pGL4-THBS1 construct in WT and heterozygous ATXN3-knockout Human Neuroblastoma Cells. *The results are presented as the mean ± SD of three independent experiments (n=3). Data were statistically analysed using paired t-test, $P < 0.05$*

5. Discussion and conclusions

5.1 Discussion

Transcriptional dysregulation represents a possible pathogenetic mechanism in polyQ diseases and has been shown in various polyQ disease models suggesting that differential expression of genes contributes to disease pathogenesis in SCA3 and other polyQ disorders (Riley & Orr, 2006). ATXN3 is known to interact with several transcription factors, binds to specific genomic regions and regulates expression of distinct genes (Costa & Paulson, 2012; Dueñas et al., 2006). In this thesis, 16 ATXN3-bound genomic regions, previously identified by ChIP sequencing, correlating to differentially expressed mRNAs in human iPS-derived neurons were verified by ChIP-PCR analyses using ATXN3-immunoprecipitated chromatin from different human iPS-derived neurons. ChIP-PCR analyses confirmed 5 ATXN3-enriched genomic regions which were used for the generation of reporter constructs containing the identified ATXN3-bound regions in front of a luciferase gene. The reporter constructs were used to study the potential capacity of ATXN3 to modulate transcription of the identified genomic regions in human neuroblastoma SH-SY5Y cells overexpressing different ATXN3 isoforms and SH-SY5Y-derived ATXN3 knockout cells expressing only one functional allele of *ATXN3*.

In both wildtype and ATXN3 knockout SH-SY5Y cells, the subcloned ATXN3-bound genomic regions of CEBPD-I, PER2, SERPINF2 and THBS1 strongly repressed the transcription of the reporter luciferase whereas the subcloned genomic region of PTEN did not significantly alter the luciferase activity compared to the activity of the empty reporter construct. In contrast to yet known ATXN3-regulated genes (Araujo et al., 2011; Evert et al., 2006; Reina et al., 2012; Sacco et al., 2014), the genomic regions analyzed here are not localized within core promoter regions of the genes but in intronic or genomic regions up/downstream of the associated genes. Thus, it is likely that the ATXN3-bound genomic regions studied here, except PTEN, function as repressing elements to silence transcription of the associated genes.

However, in regard to ATXN3-modulated regulation, only the genomic ATXN3-bound regions associated with the CEBPD-I and THBS1 gene exerted ATXN3-dependent changes in luciferase activity (**Figure 7** and **Figure 19**, respectively). Compared to wildtype cells, the repressing capacity of CEBPD-I and THBS1 in ATXN3 knockout cells was reduced by half suggesting that ATXN3 binds to these regions to enhance or maintain transcriptional repression. The mechanism of repression is not clear yet but could principally involve formation of repressor complexes and histone deacetylation as recently shown for the ATXN3-mediated repression of the matrix metalloproteinase-2 gene transcription (Evert et al., 2006). Moreover, ATXN3 is known to interact with the TATA-box binding protein (TBP) and to inhibit histone acetyltransferase activity (Li, et al., 2002). Since transcription of the reporter luciferase in the generated constructs is controlled by a minimal promoter containing a TATA box it is possible that ATXN3 interferes with the assembly and formation of the transcription initiation complex at the minimal promoter. Future studies, using the reporter constructs without minimal promoter or deleted TATA box are required to show if ATXN3 confers its repressive capacity via TBP-mediated transcription.

Similarly, in wildtype cells overexpressing GFP fusion proteins of normal and mutant ATXN3, ATXN3-dependent changes were found only for the reporter constructs containing the genomic regions associated with the CEBPD-I and THBS1 genes. For CEBPD-I, co-expression of normal and mutant human full-length ATXN3 clearly increased the luciferase activity compared to the control co-expressing GFP protein only (**Figure 6**). Interestingly, the increase in luciferase activity mediated by mutant ATXN3 was lower than by normal ATXN3 possibly indicating a partial loss of function of the mutant ATXN3 protein. Among other functions, CEBPD is a crucial transcription factor involved in inflammatory processes in the human brain (Ramji & Foka, 2002) and has been shown to be significantly upregulated in rat mesencephalic cells overexpressing mutant ATXN3 and brain sections of SCA3 patients (Evert et al., 2003). Thus, ATXN3 normally may be involved in the regulation of the CEBPD gene expression and altered CEBPD expression may be

associated with SCA3 disease. However, future studies are required to investigate the implication of CEBPD in SCA3 pathogenesis in more detail.

For THBS1, normal and mutant human full-length ATXN3 also increased the luciferase activity with an upward trend for mutant ATXN3 suggesting that mutant ATXN3 exerts a gain of function mediating aberrant transcriptional activation of the THBS1 genomic region (**Figure 18**). THBS1 is involved in proliferation and differentiation of neural progenitor cells and critically supports neuronal migration in brain tissue (Lu & Kipnis, 2010). Moreover, THBS1 was found to be upregulated in mesencephalic cells overexpressing mutant ATXN3 (Evert et al., 2003). These findings suggest that ATXN3 could be involved in transcriptional regulation of the THBS1 gene expression. Recently, it has been reported, that induction of CEBPD contributes to the repression of THBS1 transcription (Ko et al., 2015). Further studies are needed to clarify whether ATXN3 participates in the regulation of CEBPD and THBS1 and, if transcriptional dysregulation of these targets contributes to SCA3 pathogenesis.

The ATXN3-mediated changes observed for the ATXN3-bound genomic regions associated with PER2, PTEN and SERPINF2 gene in ATXN3 overexpressing SH-SY5Y wildtype cells were only moderate indicating that regulatory effects mediated by ATXN3 on these regions under the conditions analyzed are unlikely (**Figure 9**, **Figure 12** and **Figure 15**, respectively). However, since ATXN3 is involved in the cellular stress response (Costa & Paulson, 2012), the reporter activity of the constructs should also be analyzed in cells exposed to various stress stimuli including oxidative stress and heat shock. In addition, future studies should include co-expression of ATXN3 without a GFP fusion protein as well as other ATXN3 isoforms (Costa & Paulson, 2012). ATXN3 is alternatively spliced to encode either a C-terminal hydrophobic stretch or a third ubiquitin interacting motif (Harris et al., 2010). Moreover, at least 20 different protein isoforms of ATXN3 from 56 identified alternative transcripts of the gene have been predicted (Bettencourt et al., 2010, 2013) demonstrating the high variability of ATXN3 gene

transcripts and the existence of cell type-specific isoforms for individual cellular functions.

The enormous diversity of ATXN3 splice variants may be one explanation for the opposite effects encountered for ATXN3-dependent discrepancies obtained in reporter assays for the analyzed genomic regions. In this work two different experimental approaches were used (i) reporter assays in human neuroblastoma SH-SY5Y cells overexpressing artificial GFP fusion proteins of ATXN3 and (ii) reporter assays in human neuroblastoma SH-SY5Y ATXN3 knockout cells expressing endogenous ATXN3 from one functional allele of *ATXN3*. On one side, the ATXN3-mediated increased luciferase activities found for CEBPD and THBS1 in SH-SY5Y wildtype cells may be simply caused by an altered property of the artificial ATXN3 variant fused to GFP. On the other side, the loss of repressing activity by the loss of one copy of endogenous ATXN3 found for CEBPD and THBS1 regions in ATXN3-knockout cells may be also caused by off-target effects resulting from genetic modification of the cell line. Moreover, since the ATXN3-bound genomic regions were inserted in front of minimal promoter in the reporter constructs, activating and repressing effects may have interfered with each other. Therefore, additional experiments should also include reporter constructs of the analyzed genomic regions without a minimal promoter. In addition, further experiments in SH-SY5Y ATXN3 knockout cells are required to analyze whether co-expression of ATXN3 is capable to rescue the loss of repression found in KO cells for CEBPD and THBS1.

In this thesis, enriched binding of ATXN3 to five genomic regions in iPS-derived neurons from controls and SCA3 patients was confirmed. The analysis of their potential gene regulatory activity revealed that two of the ATXN3-bound genomic regions exert altered gene regulatory activities in response to the loss of endogenous and co-expressed recombinant ATXN3. Correspondingly, the binding of ATXN3 to genomic regions of both CEBPD-I and THBS1 was significantly lower in SCA3 patient compared to control indicating that altered binding properties of the mutant ATXN3 protein may contribute to differential regulatory effects and altered expression of the associated genes in SCA3 (**Figure 2**). Thus, the identified genomic

regions are likely to be regulated by ATXN3 and might be implicated in pathological processes in SCA3 disease.

5.2 Conclusion

ATXN3 is widely expressed in numerous mammalian tissues and is involved in many diverse processes in mammalian cells. Besides its function in protein quality control, ATXN3 plays an important role in transcriptional regulation. We focused our efforts on the identification of genes regulated by ATXN3 to better understand the molecular mechanisms underlying the pathogenesis of the severe neurodegenerative SCA3 disease.

Our data confirmed enriched binding of ATXN3 to five genomic regions in iPS-derived neurons from controls and SCA3 patients. The analysis of their potential gene regulatory activity revealed that two of the ATXN3-bound genomic regions, CEBPD and THBS1, exert altered gene regulatory activities in response to the loss of endogenous and co-expressed recombinant ATXN3. The identified genomic regions are likely to be regulated by ATXN3 and might be implicated in pathological processes in SCA3 disease.

6. References

Antony, P. M. A., Mäntele, S., Mollenkopf, P., Boy, J., Kehlenbach, R. H., Riess, O., & Schmidt, T. (2009). Identification and functional dissection of localization signals within ataxin-3. *Neurobiology of Disease*, *36*(2), 280–292. ISSN 0969-9961

Araujo, J., Breuer, P., Dieringer, S., Krauss, S., Dorn, S., Zimmermann, K., … Evert, B. O. (2011). FOXO4-dependent upregulation of superoxide dismutase-2 in response to oxidative stress is impaired in spinocerebellar ataxia type 3. *Human Molecular Genetics*, *20*(15), 2928–2941. ISSN 0964-6906

Bettencourt, C., Raposo, M., Ros, R., Montiel, R., Bruges-Armas, J., & Lima, M. (2013). Transcript Diversity of Machado-Joseph Disease Gene (ATXN3) Is Not Directly Determined by SNPs in Exonic or Flanking Intronic Regions. *Journal of Molecular Neuroscience*, *49*(3), 539–543. ISSN 0895-8696

Bettencourt, C., Santos, C., Montiel, R., Do Carmo Costa, M., Cruz-Morales, P., Santos, L. R., … Lima, M. (2010). Increased transcript diversity: Novel splicing variants of Machado-Joseph Disease gene (ATXN3). *Neurogenetics*, *11*(2), 193–202. ISSN 1364-6745

Bichelmeier, U., Schmidt, T., Hübener, J., Boy, J., Rüttiger, L., Häbig, K., … Riess, O. (2007). Nuclear localization of ataxin-3 is required for the manifestation of symptoms in SCA3: in vivo evidence. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *27*(28), 7418–7428. ISSN 0270-6474

Breuer, P., Haacke, A., Evert, B. O., & W??llner, U. (2010). Nuclear aggregation of polyglutamine-expanded ataxin-3: fragments escape the cytoplasmic quality control. *Journal of Biological Chemistry*, *285*(9), 6532–6537. ISSN 0021-9258

Colomer Gould, V. F., Goti, D., Pearce, D., Gonzalez, G. A., Gao, H., Bermudez de Leon, M., … Brown, D. R. (2007). A Mutant ataxin-3 fragment results from processing at a site N-terminal to amino acid 190 in brain of Machado-Joseph disease-like transgenic mice. *Neurobiology of Disease*, *27*(3), 362–369. ISSN 0969-9961

Costa, M. do C., & Paulson, H. L. (2012). Toward understanding Machado-Joseph disease. *Progress in Neurobiology*, *97*(2), 239–257. ISSN 0301-0082

Davies, S. W., Beardsall, K., Turmaine, M., DiFiglia, M., Aronin, N., & Bates, G. P. (1998). Are neuronal intranuclear inclusions the common neuropathology of triplet-repeat disorders with polyglutamine-repeat expansions? *Lancet*, *351*(9096), 131–133. ISSN 0140-6736

Dueñas, A. M., Goold, R., & Giunti, P. (2006). Molecular pathogenesis of spinocerebellar ataxias. *Brain*, *129*(6), 1357–1370. ISSN 0006-8950

Evert, B. O., Araujo, J., Vieira-Saecker, A. M., de Vos, R. a I., Harendza, S., Klockgether, T., & Wüllner, U. (2006). Ataxin-3 represses transcription via chromatin binding, interaction with histone deacetylase 3, and histone deacetylation. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *26*(44), 11474–11486. ISSN 0270-6474

Evert, B. O., Vogt, I. R., Kindermann, C., Ozimek, L., de Vos, R. a, Brunt, E. R., … Wüllner, U. (2001). Inflammatory genes are upregulated in expanded ataxin-3-expressing cell lines and spinocerebellar ataxia type 3 brains. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *21*(15), 5389–5396. ISSN 1529-2401

Evert, B. O., Vogt, I. R., Vieira-Saecker, A. M., Ozimek, L., de Vos, R. A. I., Brunt, E. R. P., … Wüllner, U. (2003). Gene expression profiling in ataxin-3 expressing cell lines reveals distinct effects of normal and mutant ataxin-3. *J Neuropathol Exp Neurol*, *62*(10), 1006–18. ISSN 0022-3069

Evert, B. O., Wüllner, U., Schulz, J. B., Weller, M., Groscurth, P., Trottier, Y., … Klockgether, T. (1999). High level expression of expanded full-length ataxin-3 in vitro causes cell death and formation of intranuclear inclusions in neuronal cells. *Human Molecular Genetics*, *8*(7), 1169–1176. ISSN 0964-6906

Fu, Y. H., Kuhl, D. P., Pizzuti, A., Pieretti, M., Sutcliffe, J. S., Richards, S., … Warren, S. T. (1991). Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. *Cell*, *67*(6), 1047–58. ISSN 0092-8674

Goti, D. (2004). A Mutant Ataxin-3 Putative-Cleavage Fragment in Brains of Machado-Joseph Disease Patients and Transgenic Mice Is Cytotoxic above a Critical Concentration. *Journal of Neuroscience*, *24*(45), 10266–10279. ISSN 0270-6474

Harris, G. M., Dodelzon, K., Gong, L., Gonzalez-Alegre, P., & Paulson, H. L. (2010). Splice isoforms of the polyglutamine disease protein ataxin-3 exhibit similar enzymatic yet different aggregation properties. *PLoS ONE*, *5*(10). ISSN 1932-6203

Chou, A. H., Chen, C. Y., Chen, S. Y., Chen, W. J., Chen, Y. L., Weng, Y. S., & Wang, H. L. (2010). Polyglutamine-expanded ataxin-7 causes cerebellar dysfunction by inducing transcriptional dysregulation. *Neurochemistry International*, *56*(2), 329–339. ISSN 0197-0186

J., E., M., E., & F., B. (2004). Neurodegenerative diseases and oxidative stress. *Biomedicine and Pharmacotherapy*, *58*(1), 39–46. ISSN 0753-3322

Jeub, M., Herbst, M., Spauschus, A., Fleischer, H., Klockgether, T., Wuellner, U., & Evert, B. O. (2006). Potassium channel dysfunction and depolarized resting membrane potential in a cell model of SCA3. *Experimental Neurology*, *201*(1), 182–192. ISSN 0014-4886

Kawaguchi, Y., Okamoto, T., Taniwaki, M., Aizawa, M., Inoue, M., Katayama, S., … Akiguchi, I. (1994). CAG expansions in a novel gene for Machado-Joseph disease at chromosome 14q32.1. *Nature Genetics*, *8*(3), 221–228. ISSN 1061-4036 (Print)

Ko, C. Y., Chu, Y. Y., Narumiya, S., Chi, J. Y., Furuyashiki, T., Aoki, T., … Wang, J. M. (2015). The CCAAT/enhancer-binding protein delta/miR135a/thrombospondin 1 axis mediates PGE2-induced angiogenesis in Alzheimer's disease. *Neurobiology of Aging*, *36*(3), 1356–1368. ISSN 1558-1497

Koch, P., Breuer, P., Peitz, M., Jungverdorben, J., Kesavan, J., Poppe, D., … Brüstle, O. (2011). Excitation-induced ataxin-3 aggregation in neurons from patients with Machado–Joseph disease. *Nature*, *480*, 543–546. ISSN 0028-0836

Kovacs: *Neuropathology of Neurodegenerative Diseases*. 2012. Cambridge University Press, ISBN: 978-110758866-0;978-110767420-2, 1-7.

Li, F., Macfarlan, T., Pittman, R. N., & Chakravarti, D. (2002). Ataxin-3 is a histone-binding protein with two independent transcriptional corepressor activities. *Journal of Biological Chemistry*, *277*(47), 45004–45012. ISSN 0021-9258

Lu, Z., & Kipnis, J. (2010). Thrombospondin 1--a key astrocyte-derived neurogenic factor. *FASEB Journal : Official Publication of the Federation of American Societies for Experimental Biology*, *24*(6), 1925–34. ISSN 1530-6860

Macedo-Ribeiro, S., Cortes, L., Maciel, P., & Carvalho, A. L. (2009). Nucleocytoplasmic shuttling activity of ataxin-3. *PLoS ONE*, *4*(6). ISSN 1932-6203

Mueller, T., Breuer, P., Schmitt, I., Walter, J., Evert, B. O., & Wüllner, U. (2009). CK2-dependent phosphorylation determines cellular localization and stability of ataxin-3. *Human Molecular Genetics*, *18*(17), 3334–3343. ISSN 0964-6906

Nieoullon, A. (2011). Neurodegenerative diseases and neuroprotection: current views and prospects. *Journal of Applied Biomedicine*, *9*(4), 173–183. ISSN 1214-021X

Orr, H. T., & Zoghbi, H. Y. (2007). Trinucleotide Repeat Disorders - annurev.neuro.29.051605.113042. *Annual Review of Neuroscience*. ISSN 0147-006X

Ramji, D. P., & Foka, P. (2002). CCAAT/enhancer-binding proteins: structure, function and regulation. *The Biochemical Journal*, *365*(Pt 3), 561–75. ISSN 0264-6021

Reina, C. P., Nabet, B. Y., Young, P. D., & Pittman, R. N. (2012). Basal and stress-induced Hsp70 are modulated by ataxin-3. *Cell Stress and Chaperones*, *17*(6), 729–742. ISSN 1355-8145

Riley, B. E., & Orr, H. T. (2006). Polyglutamine neurodegenerative diseases and regulation of transcription: Assembling the puzzle. *Genes and Development*, *20*(16), 2183–2192. ISSN 0890-9369

Ross, C. A. (1997). Intranuclear neuronal inclusions: A common pathogenic mechanism for glutamine-repeat neurodegenerative diseases? *Neuron*, *19*(6), 1147–1150. ISSN 0896-6273

Rubinsztein, D. C., Wyttenbach, A., & Rankin, J. (1999). Intracellular inclusions, pathological markers in diseases caused by expanded polyglutamine tracts? *Journal of Medical Genetics*, *36*(4), 265–270. ISSN 0022-2593

Sacco, J. J., Yau, T. Y., Darling, S., Patel, V., Liu, H., Urbé, S., … Coulson, J. M. (2014). The deubiquitylase Ataxin-3 restricts PTEN transcription in lung cancer cells. *Oncogene*, *33*(August), 4265–4272. ISSN 1476-5594

Simões, A. T., Gonçalves, N., Koeppen, A., Déglon, N., Kügler, S., Duarte, C. B., & Pereira De Almeida, L. (2012). Calpastatin-mediated inhibition of calpains in the mouse brain prevents mutant ataxin 3 proteolysis, nuclear localization and aggregation, relieving Machado-Joseph disease. *Brain*, *135*(8), 2428–2439. ISSN 1460-2156

Takahashi, T., Katada, S., & Onodera, O. (2010). Polyglutamine diseases: Where does toxicity come from? What is toxicity? Where are we going? *Journal of Molecular Cell Biology*, *2*(4), 180–191. ISSN 1674-2788

Teixeira-Castro, A., Ailion, M., Jalles, A., Brignull, H. R., Vilacã, J. L., Dias, N., … Maciel, P. (2011). Neuron-specific proteotoxicity of mutant ataxin-3 in C. elegans: Rescue by the DAF-16 and HSF-1 pathways. *Human Molecular Genetics*, *20*(15), 2996–3009. ISSN 0964-6906

Warren, S. T., & Nelson, D. L. (1993). Trinucleotide repeat expansions in neurological disease. *Current Opinion in Neurobiology*, *3*(5), 752–759. ISSN 0959-4388

Weber, J. J., Sowa, A. S., Binder, T., & Hübener, J. (2014). From pathways to targets: Understanding the mechanisms behind polyglutamine disease. *BioMed Research International*, *2014*. ISSN 2314-6141

YoungSoo Kim, Yunkyung Kim, Onyou Hwang and Dong Jin Kim (2012). Pathology of Neurodegenerative Diseases, Brain Damage - Bridging Between Basic Research and Clinics, Dr. Alina Gonzalez-Quevedo (Ed.), ISBN: 978-953-51-0375-2, InTech

Zoghbi, H. Y., & Orr, H. T. (2000). Glutamine Repeats and Neurodegeneration. *Annual Review of Neuroscience*. ISSN 0147-006X