Universidade do Minho
Escola de Medicina

Ana Rita Ferreira Fernandes

Development of a cellular ataxin-3 protein-protein interaction (PPI) assay for high-throughput screening of PPI modifiers

outubro de 2022

Development of a cellular ataxin-3 protein-protein interaction (PPI) assay for high throughput screening of PPI modifiers
Ana Rita Ferreira Fernandes
UMinho | 2022

Universidade do Minho
Escola de Medicina

Ana Rita Ferreira Fernandes

Development of a cellular ataxin-3 protein-protein interaction (PPI) assay for high-throughput screening of PPI modifiers

Dissertação de Mestrado
Mestrado em Ciências da Saúde

Trabalho efetuado sob a orientação do
Doutor Bruno Almeida
e da
Professora Doutora Patrícia Maciel

outubro de 2022

AGRADECIMENTOS/ACKNOWLEDGEMENTS

Em primeiro lugar quero agradecer à Professora Patrícia Maciel por me ter dado a oportunidade de integrar o seu grupo de investigação, bem como a toda a sua equipa que me proporcionou momentos de partilha e aprendizagem científica, nomeadamente à Andreia Castro, Sara Silva, Bruno Almeida, Liliana Costa, Daniela Garcia, Bruna Lomba, Jorge Fernandes, Stéphanie Oliveira, Joana Sousa, Cármen Vieira, Daniela Campos, Marta Costa, Sara Guerreiro, Daniela Monteiro, Joana Correia, Isabella França, João Barbosa, Jorge Silva, e ainda à Catarina Ferreira, Andreia Carvalho, Margarida Sousa, Beatriz Ferreira, André Mira, e Oleg Chertkov. Um agradecimento especial ao Bruno Almeida, por ter aceitado ser meu orientador, guiar-me neste projeto, e passar-me muito do seu conhecimento nesta área. À Liliana Costa, que tornou esta dissertação possível, foi a minha "mãe científica", passou-me os seus conhecimentos da melhor forma, e deu-me liberdade para errar e aprender. Além disso, tornou os dias no laboratório mais divertidos. Muito obrigada por tudo!

Obrigada também ao Professor Fernando Rodrigues, e ainda ao Jonas Campos, à Catarina Matos, e ao João do Canto, pela disponibilidade para me ajudarem quando precisei, e pela partilha do vosso alargado conhecimento. Não posso deixar de agradecer aos meus colegas de mestrado, em especial à Carolina e à Sara, por fazerem este percurso comigo, pelos desabafos, e pelos almoços e lanches juntos. I would also like to thank Dr. Stéphanie Cabantous from the Cancer Research Center of Toulouse that kindly sent us material that was crucial for the development of my work.

À AAUMinho, por ter proporcionado a maior aventura da minha vida, ter desafiado as minhas capacidades a todos os níveis, enriquecido as minhas soft skills, e pelas memórias e amigos para a vida.

Termino agradecendo aos meus pais, irmão, amigos e ao Leonardo, que me apoiam incondicionalmente, e me dão vida fora do laboratório.

FUNDING

The work presented in this thesis was performed in the Life and Health Sciences Research Institute (ICVS), Minho University. Financial support was provided by National funds, through the Foundation for Science and Technology (FCT) - project 02/SAICT/2017, supported by Norte Portugal Regional Operational Programme (NORTE 2020), under the PORTUGAL 2020 Partnership Agreement, through the European Regional Development Fund (ERDF).

STATEMENT OF INTEGRITY

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration.

I further declare that I have fully acknowledged the Code of Ethical Conduct of the University of Minho.

Resumo

Desenvolvimento de um sistema celular para deteção de interações proteína-proteína (IPPs) com a ataxina-3 e descoberta de moduladores destas IPPs

A doença de Machado-Joseph (DMJ) é uma doença neurodegenerativa provocada pela expansão de um trato de poliglutamina na proteína ataxina-3 (ATXN3). Evidências suportam a presença nuclear da ATXN3 como crucial para a patogénese da doença. No entanto, os mecanismos pelos quais provoca toxicidade no núcleo das células neuronais são desconhecidos. Uma vez que estudos recentes sugerem que a proteína desempenha papéis fisiológicos importantes no núcleo de células neuronais ao interagir com diferentes proteínas, surge a hipótese de que a ATXN3 expandida interage anormalmente com diferentes proteínas nesse compartimento, e de que essas interações interrompem vias celulares, contribuindo significativamente para a toxicidade neuronal. Tendo em conta esta informação, com este trabalho pretendeu-se desenvolver e otimizar o sistema da GFP (Proteína Verde Fluorescente, do inglês *Green Fluorescent Protein*) tripartida *in vivo*, que foi utilizado para validar uma interação proteína-proteína (IPP) nuclear relevante com a ATXN3, nomeadamente com o fator de *splicing* 9G8, uma vez que se observou que o seu padrão de ubiquitilação é alterado quando a expressão da ATXN3 é silenciada. A clonagem dos vetores necessários para o desenvolvimento do sistema foi baseada na técnica de clonagem de Gibson, sendo a ATXN3 e os interatores selecionados construídos em fusão com dois fragmentos da GFP, GFP10 e GFP11. Após interação, esses fragmentos aproximam-se e ligam de forma espontânea ao fragmento GFP1-9, expresso numa linha celular de mamífero previamente gerada, e é emitida fluorescência. O sistema foi validado com controlos positivos e negativos adequados por citometria de fluxo, imunofluorescência, e microscopia de fluorescência, seguida de coloração nuclear com DAPI (4',6'-diamino-2-fenil-indol). As mesmas técnicas validaram a interação ATXN3-9G8, reforçando o interesse no estudo de 9G8 na patogénese da DMJ e como potencial candidato para o rastreio de compostos terapêuticos. Uma mutação no sinal de localização nuclear da ATXN3 não impediu a sua interação nuclear com a 9G8, sugerindo que a ATXN3 consegue ser transportada independentemente deste sinal. Este sistema constitui uma ferramenta valiosa para detetar IPPs, estudar a sua localização subcelular, e identificar novos alvos terapêuticos, assim como moduladores promissores de IPPs relevantes por rastreio, que podem contribuir para o desenvolvimento de uma futura terapia para a DMJ. Além disso, este sistema pode ser uma ferramenta promissora para outros estudos, nomeadamente o estudo *in vivo* da oligomerização, fibrilação, acumulação, e formação de agregados proteicos de ATXN3.

Palavras-chave: Ataxina-3; Doença de Machado-Joseph; Interações proteína-proteína; Sistema da GFP tripartida

Abstract

Development of a cellular ataxin-3 protein-protein interaction (PPI) assay for high-throughput screening of PPI modifiers

Machado-Joseph disease (MJD) is a neurodegenerative disorder caused by expansion of a polyglutamine tract within the protein ataxin-3 (ATXN3). Evidence supports the nuclear presence of ATXN3 as key for pathogenesis. However, the mechanisms through which it promotes toxicity in neuronal cell nucleus remain mostly unknown. Since recent data suggests that the protein may play important physiological roles in the nucleus of neuronal cells through the interaction with several partners, we hypothesized that mutant ATXN3 interacts abnormally with different proteins in that compartment, and that those interactions disrupt cellular pathways, contributing significantly to their neuronal toxicity. Therefore, with this work we aimed to develop and optimize the tripartite split-GFP (triSFP) system *in vivo*, that was used to validate a relevant nuclear ATXN3 protein-protein interaction (PPI), namely with the splicing factor 9G8, as it was observed that its ubiquitylating pattern is altered when the expression of ATXN3 is silenced. Vector design and cloning necessary for the development of the triSFP system were based on Gibson Cloning technique, and ATXN3 and the selected interactors were successfully built in frame with GFP10 and GFP11 subunits. Upon interaction, these subunits get tethered and spontaneously assemble with the GFP1-9 subunits, expressed in a previously obtained mammalian cell line, to emit fluorescence. The triSFP system was validated with adequate positive and negative controls by flow cytometry, immunofluorescence, and fluorescence microscopy, followed by DAPI nuclear staining. The same techniques validated the ATXN3-9G8 candidate interaction, reinforcing the interest in the study of 9G8 in MJD pathogenesis and as a potential candidate for drug screening. A mutation at Nuclear Localization Signal (NLS) of ATXN3 did not avoid nuclear interaction between ATXN3 and 9G8, suggesting that ATXN3 is able to enter the nucleus in the absence of NLS recognition. This system constitutes a valuable tool to detect PPIs, to study their subcellular localization, and to identify novel therapeutic targets and promising modulators of relevant PPIs by high throughput screening assays, that might constitute lead molecules for future MJD therapy. Furthermore, this system can be a promising tool to other studies, namely the *in vivo* study of protein oligomerization, fibrilization, accumulation and ATXN3 aggregate formation.

Keywords: Ataxin-3; Machado-Joseph disease; Protein-protein interactions; Tripartite split-GFP system

INDEX

ABBREVIATION LIST

ADMET – Absorption, Distribution, Metabolism, Elimination and Toxicity
ALPHAScreen – Amplified Luminescent Proximity Homogeneous Assay Screen
ALS – Amyotrophic Lateral Sclerosis
ATXN3 – Ataxin-3
BiFC – Bimolecular Fluorescence Complementation
CAG – Cytosine-Adenine-Guanidine
CBP – CREB-Binding Protein
cDNA – Complementary DNA
CGG – Cytosine
Chk1 – Checkpoint kinase 1
CREB – CAMP-Response Element Binding
CTG – Cytosine-Thymine-Guanidine
DAPI - 4′,6-Diamidino-2-Phenylindole
Dsbs – Double Strand-Breaks
DUB – Deubiquitinating (Enzyme)
ELISA – Enzyme-Linked Immunosorbent Assay
ERAD – Endoplasmic Reticulum Associated Protein Degradation
F2H – Fluorescence Two-Hybrid
FACS – Flow Activated Cell Sorting
FC – Flow Cytometry
FITC – Fluorescein Isothiocyanate
FOXO – Forkhead Box O
FP – Fluorescence Polarization
FRDA – Friedreich Ataxia
FRET/BRET – Fluorescence/Bioluminescence Resonance Energy Transfer
FXTAS – Fragile X-Associated Tremor Ataxia Syndrome
GAA – Guanidine-Adenine-Adenine
GCG – Guanidine-Cytosine-Guanidine
GCN4 – General Control Nondepressible 4
GFP – Green Fluorescence Protein
HD – Huntington Disease
HDAC3 – Histone Deacetylase 3
HEK – Human Embryonic Kidney
ITC – Isothermal Titration Calorimetry
JD – Josephin Domain
KPNA3 – Karyopherin Subunit Alpha 3
LB – Luria Bertani
MAPT – Microtubule Associated Protein Tau
MD1 – Myotonic Dystrophy Type 1
MDC1 – Mediator of DNA Damage Checkpoint protein 1

MJD – Machado-Joseph Disease
MMP-2 – Matrix-Metalloproteinase 2
mRNA – Messenger Ribonucleic Acid
MST – Microscale Thermophoresis
Nbs – Nuclear Bodies
NER – Nucleotide Excision Repair
NES – Nuclear Export Signal
Niis – Neuronal Intranuclear Inclusions
NLS – Nuclear Localization Signal
NPC – Nuclear Pore Complex
OD – Optical Density
OPMD – Oculopharyngealmusculardystrophy
PCA – Protein-Fragment Complementation Assay
PCAF – P300/CREBBP Associated Factor
PCR – Polymerase Chain Reaction
PLA – Proximity Ligation Assay
PML – Promyelocytic Leukemia
PNKP – Polynucleotide Kinase 3'-Phosphatase
PolyQ – Polyglutamine
PPI – Protein-Protein Interaction
RNAPII - RNA Polymerase II
RNF4 – Ring Finger Protein 4
RNF8 – Ring Finger Protein 8
RT – Room Temperature
RT-PCR – Reverse Transcription Polymerase Chain Reaction
SBMA – Spinobulbar Muscular Atrophy
SCA3 – Spinocerebellar Ataxia Type 3
SCAS – Spinocerebellar Ataxias
SOC – Super Optimal broth with Catabolite repression
SOD2 – Superoxide Dismutase 2
SPR – Surface Plasmon Resonance
SRSF7 – Serine-Rich Splicing Factor 7
SSBs – Single Strand-Breaks
TC-NHEJ – Transcribed-Coupled Nonhomologous End-Joining
TR-FRET – Time-Resolved Forster Resonance Energy Transfer
Trisfp – Tripartite-split GFP
TRITC – Tetramethylrhodamine
UBDs – Ubiquitin Binding Domains
UPS – Ubiquitin-Proteasome System
VCP – Valosin-Containing Protein
WT – Wild-Type

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

Trinucleotide repeats, polyglutamine diseases and protein aggregation

Trinucleotide repeats are unstable regions within the genome found in several neurodevelopmental genes, being essential for brain function. Within protein coding regions they generate patterns of repeated amino acids which play a crucial role on the establishment of protein-protein interactions (PPIs), and contribute for evolution, by providing genetic variability. On the other hand, a variety of neurodegenerative disorders are characterized by the expansion of those unstable homopeptide sites. These are very dynamic regions that can form different secondary structures, that may interfere with basic cellular mechanisms. When gene expression levels and RNA or protein function or folding are compromised, serious disorders start to appear. Some examples of those conditions are: Spinobulbar Muscular Atrophy (SBMA), Huntington Disease (HD), and Spinocerebellar Ataxias (SCAS), characterized by cytosine-adenine-guanidine (CAG) expansion, Oculopharyngealmusculardystrophy (OPMD), characterized by guanidine-cytosine-guanidine (GCG) expansion, Myotonic Dystrophy Type 1 (MD1), characterized by cytosine-thymine-guanidine (CTG) expansion, Fragile X-Associated Tremor Ataxia Syndrome (FXTAS), characterized by cytosine-guanine-guanidine (CGG) expansion, and Friedreich's Ataxia (FRDA), characterized by guanidine-adenine-adenine (GAA expansion) [1].

Polyglutamine (polyQ) diseases are a group of nine human neurodegenerative inherited disorders that affect motor function, cognition, and mood, as they lead to progressive neurological impairment. They are characterized by the expansion of the CAG trinucleotide repeat within the respective disease causative gene, leading to the formation of an abnormal polyQ tract within the encoded protein. Each disease has different hallmarks depending on the affected protein's subcellular localization, abundance, structure, biological function, and the way the expansion affects its role in the cell [2]. However, one common characteristic of these disorders is the fact that mutant proteins are more prone to aggregate, and ultimately to form intranuclear and cytoplasmic inclusions. These aggregates are believed to contribute for neurodegeneration, but their mechanisms of pathogenicity are not fully clarified.

Even though disease-linked mutations are the main and general cause of polyQ diseases, other phenomena may play a role in the development of these diseases by contributing for protein misfolding and aggregate formation, such as post-translational modifications, prion-like behaviours, protein supersaturation, deficiencies in protein clearance pathways, increase in oxidative stress as well as endoplasmic reticulum stress, mitochondrial dysfunction and alteration of cytoplasmic membrane

permeability [3]. Altogether, endogenous and exogenous stress conditions affect proteostasis, and result in global misfolding and aggregation of the endogenous proteome.

Aggregation relies on a multistep process that begins with protein misfolding and formation of misfolded monomers, that quickly evolve to misfolded soluble oligomers, and then transit to insoluble forms. Soluble oligomers are considered toxic species, being key players in alterations of cell physiology and pathology. As so, it is hypothesized that this gain of toxic function, together with a partial loss of function due to capture of the protein in aggregates, ultimately lead to neurodegeneration [4], [5].

Machado-Joseph disease

Machado-Joseph disease (MJD), also known as spinocerebellar ataxia type 3 (SCA3), is a polyglutamine disease that results from the expansion of the trinucleotide CAG in the *ATXN3* gene, mapped to the chromosome 14q32.1 region. The human *ATXN3* gene has a length of approximately 48 Kb, comprises 11 exons, and the CAG repeats are located at exon 10 [6]–[8]. This expansion results in an abnormal polyQ tract near the C-terminus of the codified wild-type (WT) protein ataxin-3 (ATXN3). While in healthy individuals CAG repeats range from 12 to 44, in MJD patients the tract contains between 60 and 87 repeat units, and becomes unstable during transmission [9], [10]. The discovery of this dynamic expansion of the CAG repeats helped to clarify some features of MJD, namely the fact that the size of the expanded CAG repeat number is inversely correlated with the age of onset of the disease and is directly correlated with the severity of symptoms, reinforcing the central role of the expanded polyQ tract itself in the pathogenesis [11]. This explains why the clinical spectrum of the disease is so varied, even within the same family. Actually, descendants tend to have slightly higher number of CAG repeats, and consequently more severe forms of the disease, with faster progression [11]–[13].

MJD was first reported in individuals of Portuguese-Azorean ancestry in the 1970s, but quickly migrated to families of other geographic locations and ethnic backgrounds. It has a higher relative frequency in Portugal, Brazil, and Japan, but a significant number of cases has also been described in Canada, United States of America (USA), and Australia. Overall, MJD has an average prevalence of 1-5/100.000, being the most common dominant ataxia worldwide [14]–[16].

Between 1972 and 1976, members of the families William Machado, Thomas and Joseph were thought to suffer from different genetic conditions as they presented distinct phenotypes, as well as different ages of onset of the symptoms. In 1977, the Portuguese neurologist Paula Coutinho proposed for the first time the unification of the disease, and in 1980, the disease was officially called Machado-Joseph disease. The research performed by Corino Andrade and Paula Coutinho made a key contribution

to the definition of the disease, as they could delineate for the first-time a classification of the disease subtypes and define clinical criteria to diagnose the patients with MJD [17].

The main clinical hallmark of MJD is progressive ataxia, which is a motor coordination dysfunction that affects gait, balance, speech, and vision. Common symptoms related to the disfunction of the pyramidal tract or other motor clinical manifestations of the disease are muscle weakness, overresponsive reflexes, rigidity, slowed movement, tremor, involuntary eye movement and muscle contraction (**Figure 1**). However, nonmotor symptoms can also occur affecting sleep, cognition, and psychiatric normality of the patients [7], [9], [18].

As different patients present distinct clinical symptoms, it was possible to define a clinical classification with subtypes of the disease. The most severe ("type Joseph") is characterized by cerebellar ataxia and dominance of pyramidal and extrapyramidal anomalies. The symptoms appear early in life (10-30 years) and the disease progresses more quickly. In the second and most common type ("type Thomas"), symptoms include cerebellar ataxia, progressive external ophthalmoplegia, and with or without pyramidal signs. The symptoms manifest at an intermediate age (20-50 years). The third type ("type Machado") is marked by a late-onset (40-75 years) and slow progression of the disease. Symptoms include cerebellar deficits and peripheral alterations, with or without pyramidal signs. Curiously, different subtypes can be present in the same family [7]. Regarding life expectancy, MJD patients can expect a mean age of survival after disease onset of 21.18 years [19].

Regarding neuropathological features, MJD patients' brains have substantially less weight, and important areas like the deep cerebellar nuclei, pons, substantia nigra, medulla oblongata and some cranial nerves, are compromised [7].

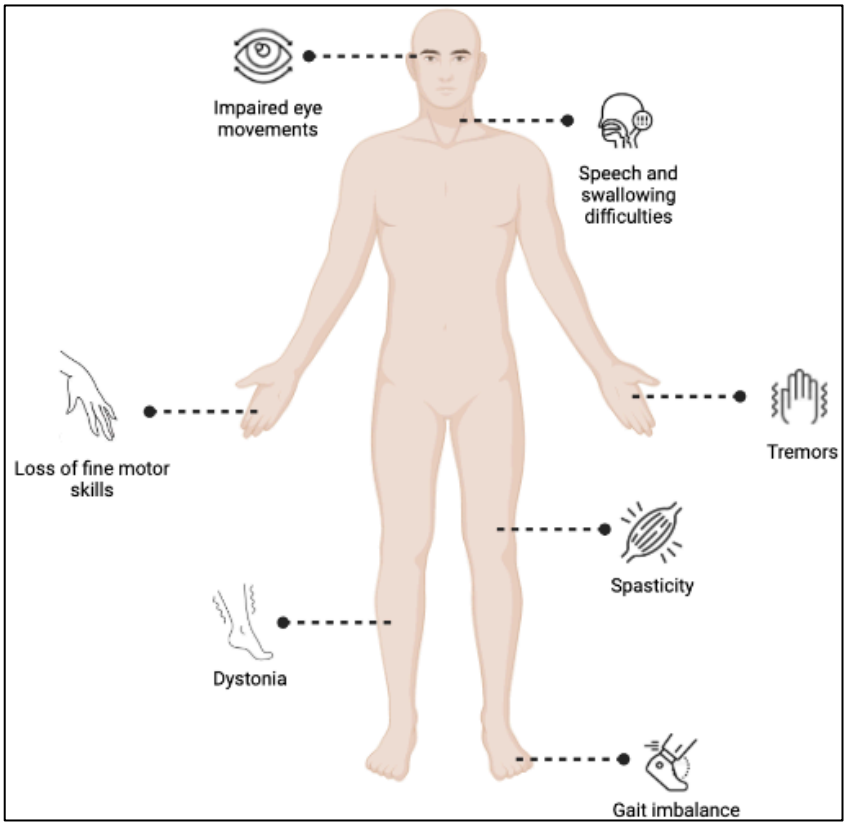


Figure 1 - Main clinical features of Machado-Joseph disease. Machado-Joseph disease presents heterogenous manifestations in several regions, that include impaired eye movements, speech and swallowing difficulties, deterioration of fine motor skills, dystonia, spasticity, tremors, and abnormalities in motor coordination, such as gait imbalance. Image was constructed with BioRender.com.

The protein ATXN3

ATXN3 is a small protein of 42 kilodaltons that contains a structured and well-conserved N-terminus with a catalytic Josephin domain (JD), and a flexible C-terminus, which is highly divergent throughout all known sequences. A wide range of phylogenetic groups such as protists, plants, invertebrates and vertebrates, contain JD-containing proteins that can be categorised in either ataxins or Josephins [20]. Furthermore, the interaction of ATXN3 with the same proteins in different species through its Josephin domain, such as NEDD8 in *Caenorhabditis elegans,* comproves that some interactions are evolutionarily conserved and have functional relevance [21]. The C-terminus contains a polyQ tract and two or three ubiquitin interaction motifs (UIMs), depending on the isoform [22], [23] (**Figure 2**). Even though several ATXN3 isoforms were identified, only two have been studied in more detail, namely the 2IUM isoform containing UIM1 and UIM2, and the 3UIM isoform containing UIM1, UIM2 and UIM3 [24], [25].

Concerning its expression, ATXN3 is ubiquitously expressed in neuronal and non-neuronal tissues. Regarding subcellular localization, ATXN3 is predominantly present in the cytoplasm but can also be found in the nucleus and mitochondria. This shuttling is in part mediated by two nuclear export signals (NES) and a nuclear localization signal (NLS), which are present in ATXN3 sequence [26]–[28].

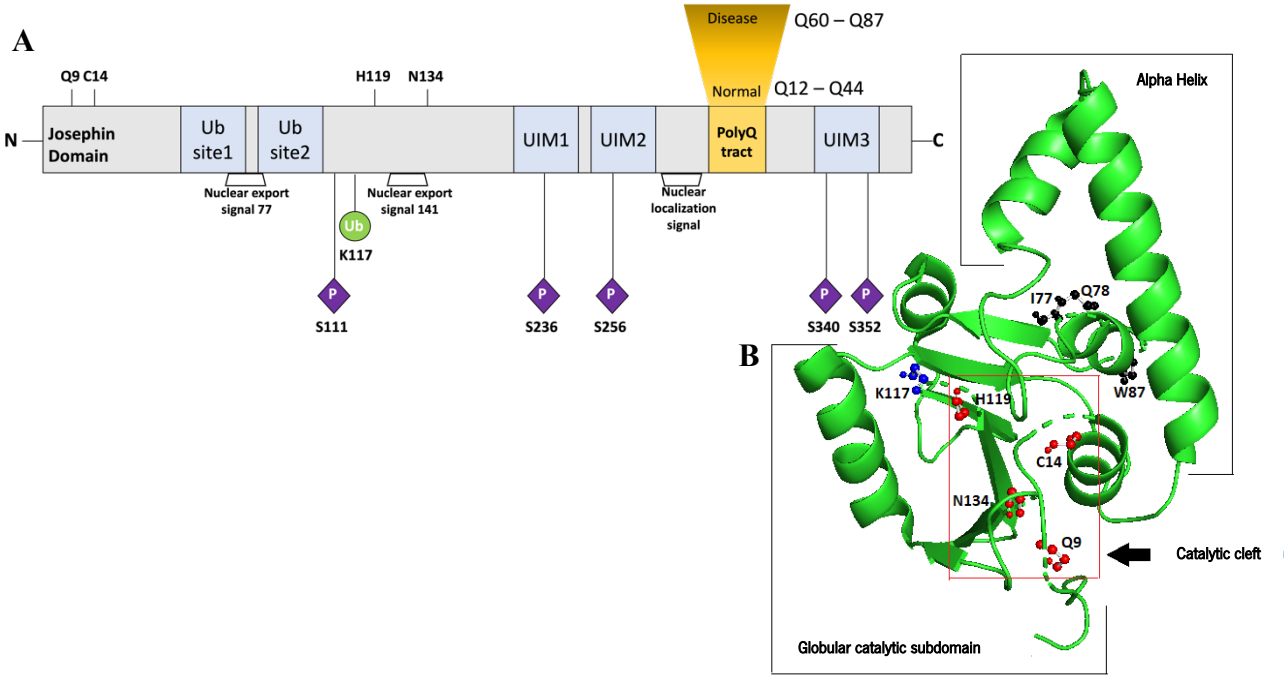


Figure 2 - Schematic representation and structure of ATXN3 protein. A - ATXN3 contains a catalytic domain at the N-terminus (Josephin domain), a polyQ tract at the C-terminus, and 3 UIMs (isoform 2, more prevalent in the brain). ATXN3 is monoubiquitylated at K117 residue and phosphorylated at residues S111, S236, S256, S340 and S352. It presents one nuclear localization signal, two nuclear export signals, and two ubiquitin binding sites. Image adapted from [29]. B – Structure of ATXN3 with the catalytic residues Q9, C14, H119 and N134 represented in red, in the catalytic fence, and residue K117 in blue near the fence, in the globular catalytic subdomain. The residues that establish ubiquitin binding domains (UBDs) are represented in black. UBD 1 (I77 e Q78) is localized near the catalytic fence, while UBD 2 (W87) is in the opposite side, separated by an alpha helix. Image obtained from the software PyMOL [30].

Evidence suggests that ATXN3 is involved in diverse cellular processes, particularly in protein quality control through the ubiquitin-proteasome system (UPS), a key mechanism for intracellular protein degradation, and which dysfunction has been associated with various neurodegenerative diseases. Furthermore, the discovery of ATXN3 deubiquitylating (DUB) activity *in vitro* was a major step towards understanding its biological function [31], [32]. Ubiquitylation is the covalent attachment of a ubiquitin (Ub) molecule or a polyUb chain to one or several lysine residues of a target protein. This post-

transcriptional modification regulates protein degradation by the proteasome, as well as DNA repair, chromatin remodelling, cell cycle progression, subcellular localization and signalling pathways [33]–[35]. Contrarily, deubiquitylation is the cleavage of ubiquitin or ubiquitin-like proteins from target proteins, affecting the same cellular mechanisms and controlling homeostasis [36].

ATXN3 is among 99 human DUBs, classified in 7 families, binds and cleaves polyUb chains with at least four Ub units, and according to some reports prefers lysine-63 (K63)-linked and mixed linkage polyUb chains over K48-linked chains [37]. The UIMs regulate the specificity of ATXN3 activity, and the protease activity resides in the catalytic domain present in the active site cysteine-14 (C14) in the JD [31], [38]. The DUB activity of ATXN3 is enhanced when the protein itself is ubiquitylated, and evidence shows that this modification is enhanced upon induction of cellular response to certain stressors, such as proteasome inhibition and unfolded protein response [39], [40]. *In vivo*, ATXN3 DUB activity is supported by two main pieces of evidence: (i) mutating the catalytic site (C14), and (ii) the depletion of ATXN3 in a mouse model, both lead to an increase of polyubiquitylated proteins, in cultured cells and tissues, respectively [41], [42]. In addition, ATXN3 was also shown to have deneddylase activity, as it can cleave the neural precursor cell expressed developmentally downregulated gene 8 (NEDD8) from its substrates. Neddylation regulates diverse cellular processes including protein degradation mediated by UPS, protein transcription and cell signalling [21].

Nevertheless, ATXN3 physiological relevance, as well as its substrates, have not been thoroughly elucidated. Previous studies reported that ATXN3 interacts with proteins involved in several cell mechanisms including cell quality control and protein folding [43], regulation of DNA transcription and repair [44], [45], cell structure and motility [46], mitochondrial respiration and immunometabolism [47], and even regulation of cancer cell stemness [48], which may indicate that ATXN3 has a physiological role in these pathways, and everything suggests that many more remain to be unravelled.

Some studies indicate that ATXN3 is a nonessential protein, and hence not necessary for normal cellular functioning. Accordingly, the depletion of ATXN3 in a *Caenorhabditis elegans* model did not alter the worm's lifespan and resulted in an increased stress resistance [46], [49]. Similarly, ATXN3 knock-down in rat and mouse models caused no signs of toxicity, and no deleterious effects, respectively [42], [50]. Since ATXN3 depletion does not affect an obvious network, it is difficult to gain insight into the molecular mechanisms in which the protein is involved. Therefore, the enlightenment of the molecular partners of ATXN3 is crucial not only to understand its biological functions, but also its involvement in disease processes.

ATXN3 in the nucleus and nuclear interactors

As previously mentioned, ATXN3 can be found in the cytoplasm, nucleus and mitochondria. However, in SCA3 patients and disease models, the expanded protein forms aggregates mainly in the nucleus of neuronal cells, resulting in neuronal intranuclear inclusions (NIIs), an important hallmark of MJD [51]. In fact, for some polyQ diseases, studies in cellular and animal models have shown that favoring the presence of the mutant protein in the nucleus by adding a strong nuclear localization signal increases its aggregation and toxicity, and by contrast, keeping the protein out of the nucleus by adding a nuclear export signal, reduces toxicity [52], [53]. Therefore, addressing the nuclear interactors of ATXN3 seems to be determinant given the relevance of ATXN3 nuclear localization for disease progression (**Figure 3**). Regarding the expanded form of ATXN3, however, there is yet no evidence to date showing that it interacts differentially with any known nuclear interactor.

As referred above, ATXN3 possesses two nuclear signals, NLS and NES, which enables its shuttling between the nucleus and the cytoplasm. NLS sequences are recognized by a group of proteins that are part of the nuclear pore complex (NPC), known as karyopherins. Karyopherin subunit alpha 3 (KPNA3) was hypothesized to interact with ATXN3 and it was shown that its overexpression increased ATXN3 nuclear presence, facilitating its transport to that compartment. Additionally, reducing KPNA3 levels alleviated the neurological phenotype induced by expanded ATXN3, characterized by nuclear toxicity [54]. Furthermore, one known ATXN3 interactor is described as the main regulator of the nuclear localization of ATXN3. Casein kinase 2 (CK2) interacts with and phosphorylates ATXN3, and increases its levels in the nucleus, suggesting a possible role of CK2 in the occurrence of NIIs. The phosphorylation-dependent cellular localization of ATXN3 also interferes with its nuclear function, since phosphorylation increases ATXN3's transcriptional repressor activity, as it happens in the matrix-metalloproteinase 2 (MMP-2) gene, where ATXN3 was proposed to act by binding specific chromatin regions [55].

The first ATXN3 interactors to be reported were hHR23A and hHR23B, the human homologs of the yeast DNA repair protein RAD23, which are engaged in the nucleotide excision repair (NER) pathway as well as in the endoplasmic reticulum associated protein degradation (ERAD) pathway, and interact with both normal and mutant ATXN3. However, the role that this DUB plays in those pathways remains to be explored. ATXN3 may be involved in the NER pathway by interacting with the ubiquitin-like domain of hHR23, which in turn is responsible for the interaction of this protein with some proteasome subunits. Indeed, another proposed and unaddressed hypothesis is the fact that the interaction between mutant ATXN3 and hHR23 affects the proteolytic pathway in some neurons, promoting their degeneration [56].

Polynucleotide kinase 3'-phosphatase or PNKP, a DNA end-processing enzyme involved in both single and double strand-break (SSBs and DSBs, respectively) repair, was also reported to interact with ATXN3. WT ATXN3 stimulates PNKP and DNA repair, while expanded ATXN3 diminishes this activity by trapping PNKP in the polyQ aggregates [44], [57]. A recent study explored this interaction and demonstrated that ATXN3 is part of a transcription-coupled nonhomologous end-joining (TC-NHEJ) DNA repair complex, together with PNKP, RNA polymerase II (RNAPII) and DNA Ligase IV [58], [59]. They presented two roles for ATXN3: (i) it activates PNKP to initiate DNA repair; and (ii) it deubiquitinates RNAPII after the repair, to resume transcription. Interestingly, absence of ATXN3 was shown to lead to damage accumulation particularly in the most transcribed genes within each cell type [59].

Another DNA repair protein associated with ATXN3 is Checkpoint kinase 1 (Chk1). The activation of this protein after DNA damage arrests the progression of the cell cycle, which either facilitates DNA repair or promotes cell death. ATXN3 was described to deubiquitinate Chk1, inhibiting its proteasomal degradation, and thus promoting its stabilization. The polyQ expansion does not seem to have an impact on the protease activity of ATXN3 towards Chk1 [60]. In the same context, another interactor of ATXN3 was reported**,** namely the mediator of DNA damage checkpoint protein 1 (MDC1), a known Ring finger protein 4 (RNF4) substrate. ATXN3 is recruited to DNA lesion sites by DNA damage-induced SUMOylation of chromatin-associated proteins, where it interacts with SUMO1 - Small Ubiquitin Like Modifier 1 - and counteracts with the E3 ubiquitin ligase RNF4, to prevent an early removal of MDC1 from DSBs. This way, DNA repair and DNA damage signaling can be initiated [61].

Evidence also suggests that ATXN3 interacts with Ring finger protein 8 (RNF8) in a complex, together with AAA + ATPase Valosin-Containing Protein (VCP) or p97. ATXN3 plays a key role in RFN8 homeostasis by deubiquitinating it and consequently slowing down its degradation rate. In the absence of DNA damage, RNF8 ubiquitinates itself, signaling for proteasome degradation. In response to DNA damage, the ATXN3-p97 complex prevents accumulation of RNF8 at the damaged site, which could delay the NHEJ-mediated DSB repair, an essential pathway for cell survival [62]. All the reported evidence suggests that ATXN3 is crucial for genome integrity, since it promotes the maintenance of the activity of several proteins involved in DNA repair.

Compelling evidence also supports a role for ATXN3 as a transcription regulator. For instance, ATXN3 was shown to inhibit CREB-mediated transcription in the 293T cell line through interaction with the cAMP-response element binding (CREB)-binding protein (CBP), p300, and p300/CBP associated factor (PCAF). Both the N and the C terminus of ATXN3 repress transcription, but while the C-terminus

binds coactivators and represses their transcription activity, the N-terminus of both WT and expanded ATXN3 binds H3 and H4 histones and blocks access to sites of acetylation [63].

ATXN3 also promotes histone deacetylation by interacting with histone deacetylase 3 (HDAC3) and the nuclear receptor co-repressor 1 (NCOR1). It is reported that for this repressive transcriptional role of normal ATXN3, UIMs are essential to identify ubiquitylated substrates and to monitor the activity of the N-terminus ubiquitin protease. This way, it is hypothesized that ATXN3 role in transcriptional regulation is coupled to its DUB activity, by stabilizing certain proteins required for transcriptional repression of specific genes [64].

Another role of ATXN3 in transcription regulation in response to oxidative stress was proposed. Both WT and expanded ATXN3 can interact with forkhead box O (FOXO) transcription factor FOXO4, but only WT ATXN3 promotes gene expression regulated by this transcription factor. In the presence of oxidative stress, ATXN3 and FOXO4 translocate to the nucleus, bind to the same region of the Superoxide Dismutase 2 (SOD2) gene promoter, and increase SOD2 expression, an antioxidant enzyme. This demonstrates a cytoprotective role of ATXN3 that might be impaired in expanded ATXN3, contributing to neurodegeneration [65].

The hypothesis of the regulatory role of ATXN3 coupled with its DUB activity seems to be true also for the interaction of ATXN3 and the serine-rich splicing factor 7 (SRSF7 or 9G8). 9G8 is an RNA-binding protein involved in the regulation of alternative splicing and polyadenylation, and also plays a role in the nuclear export of spliced RNA [66], [67]. 9G8 is part of the spliceosome machinery controlling alternative splicing of microtubule associated protein tau (MAPT) exon 10, the protein involved in the stabilization of the internal microtubules and associated with several neurodegenerative diseases. In fact, spliced Tau isoform levels are usually tightly controlled in mature human neurons, but imbalanced expression of 3R and 4R-Tau were observed in the brain of a mouse model of Down syndrome, which shows age-related neurodegeneration, and in human Pick's, Alzheimer and Parkinson disease patients, confirming the possibility that faulty tau exon 10 splicing contributes to the pathogenesis of those disorders [68]–[71]. In a neuronal cell line lacking ATXN3, the total and nuclear levels of this splicing factor were seen to be reduced when compared with control cells, and increased after proteasome inhibition, supporting the mentioned theory [45]. However, the physiological role of ATXN3 in this interaction remains unknown.

As mentioned before, NIIs are a common feature of polyQ diseases and might be a useful tool to understand their pathogenesis. Promyelocytic leukemia protein (PML) and CBP are two transcriptional

coactivators that colocalize with ATXN3 in MJD NIIs. PML nuclear bodies (NBs) are multiprotein complexes involved in a variety of important events such as apoptosis, transcription, and viral defense. The PML protein is its main component and one of its isoforms, PML-II, was shown to interact with ATXN3. Studies showed that PML-II recruits both WT and mutant ATXN3 to these nuclear bodies and abolishes its DUB activity, which may affect aggresome formation. However, the precise mechanisms through which this negative regulation happens remains to be unraveled [45], [72].

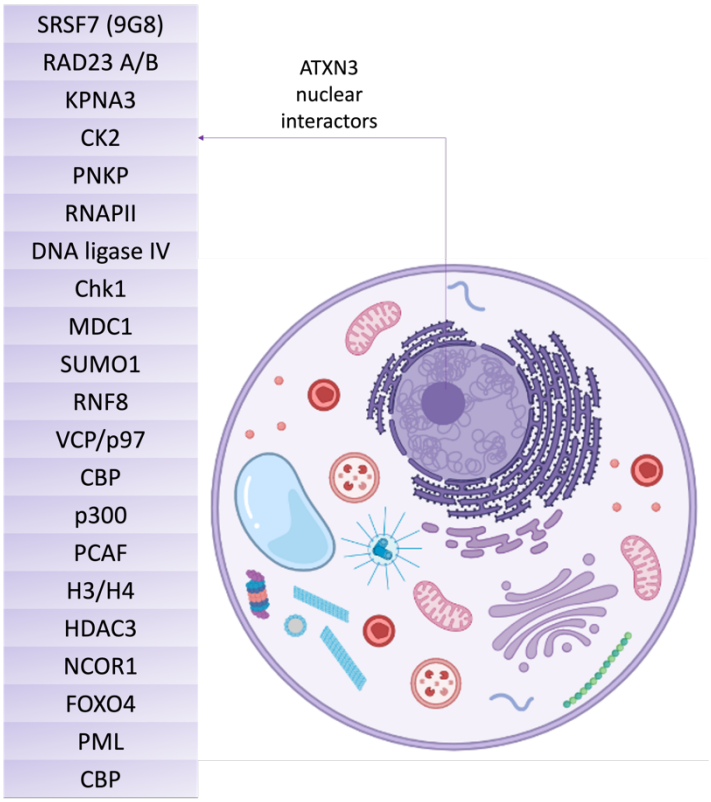


Figure 3 - Reported ATXN3 nuclear interactors. The represented nuclear interactors of ATXN3 are involved in several nuclear cell functions, such as DNA repair and genome integrity, transcription, RNA metabolism, and post-translational modifications, which suggests a role of ATXN3 in those mechanisms. Image was constructed with BioRender.com.

Protein-protein interaction (PPI) assays for drug screening

Proteins are the driving machines of living organisms, serving crucial functions by interacting with one another or with other biological macromolecules. PPIs are involved in the regulation of most biological processes, including gene expression, cell proliferation, signal transduction, transcription and apoptosis [73]. The dysregulation of interactions between specific pairs or groups of proteins can lead to various diseases. This can occur either through the loss of an essential interaction or through the formation of a

protein complex at an inappropriate time or location [74]. Literature and several databases provide already a vast knowledge regarding disease associated PPIs, exploring protein functions in networks and the molecular mechanisms involved in a certain disease [75].

Since numerous disorders are associated with abnormal PPIs, targeting those interactions is an attractive drug development strategy for disease treatment, either by disrupting or stabilizing them. Besides, targeting PPIs instead of single proteins increases drug specificity and efficacy. The interface of a PPI is formed by the combination of interacting domains of two particular proteins, which provides a higher level of uniqueness in comparison to one single protein. In contrast, a drug that targets the catalytic pocket of an enzyme will probably have an effect in a whole enzyme class with well conserved regions [76].

The modulation of disease-relevant PPIs used to be considered difficult and they were often referred as "undruggable", mainly due to PPIs' large and amorphous interfaces involved in their binding motifs. Over the years, this therapeutic strategy has become common mainly due to an increase in basic knowledge of PPI structure and energetics. An important discovery was the fact that there are "hotspots", rather than an evenly distribution of PPIs interacting energy across the surface area (Reviewed in [77]). Few key amino acids can be responsible for PPIs, which provides evidence that they can be modulated by small molecules, peptides or other scaffolds (Reviewed in [78]). Another major finding was the perception of plasticity of many protein interactions surfaces, allowing molecules to take advantage of their conformational changes. PPI-modulating compounds can inhibit or enhance a PPI by changing its conformation when they bind to a neighboring site on one of the interacting partners (allosteric PPI inhibitors or stabilizers) (Reviewed in [79]). Similarly, to disease-associated PPIs, compounds likely to modulate those interactions are available in the literature and in several databases. They allow the selection of chemical libraries with adequate ADMET (Absorption, Distribution, Metabolism, Elimination, Toxicity) profiles, likely to become useful therapeutic molecules [80][81]. PPIs are potential drug targets for a broad range of therapeutic areas, such as oncology, cancer immunotherapy, infectious diseases, neurological diseases, heart failure and inflammation and oxidative stress (Reviewed in [77]).

A variety of systems have been established to analyze PPIs. PPIs assays suitable for drug screening can be categorized as i) low-throughput target-based biochemical assays, such as isothermal titration calorimetry (ITC), surface plasmon resonance (SPR), microscale thermophoresis (MST), enzyme-linked immunosorbent assay (ELISA), pull-down and structure biology approaches (X-ray crystallography, nuclear magnetic resonance and fragment based lead discovery); ii) high-throughput target-based

biochemical assays, namely amplified luminescent proximity homogeneous assay screen (ALPHAScreen), time-resolved Forster resonance energy transfer (TR-FRET) and fluorescence polarization (FP); and iii) cell-based assays, suitable for high-throughput screening, such as co-immunoprecipitation, *in situ* proximity ligation (PLA), fluorescence two-hybrid (F2H), fluorescence/bioluminescence resonance energy transfer (FRET/BRET), bimolecular fluorescence complementation (BiFC), and tripartite-split GFP (green fluorescence protein) (triSFP) [82].

Cell-based assays

Cell-based assays are increasingly used to detect PPIs and screen PPI-modifying compounds, to obtain more biologically relevant results in a cellular context [83]. A mammalian system is likewise more favorable for drug discovery as it mimics proteins' post-translational modifications that occur in humans [84]. Moreover, those assays should allow a direct localization of PPIs, and an accurate detection of their modulation without complicated experimental setups. That important characteristic can be obtained with fluorescent proteins due to their intrinsic fluorescence, and with microscopy-based techniques [78][85]. Among many, we can highlight five prominent cell-based PPI detection methods: PLA, FRET, BRET, BiFC and triSFP (described next). It is important to point out that the activity of any compound identified with a specific assay needs to be confirmed in as many secondary assays as possible to make sure that they are not assay format-specific artefacts, and to avoid false positives (Reviewed in [78]).

In situ Proximity Ligation Assay (PLA)

The *in situ* PLA makes use of two antibodies directed against each protein of interest, generated in different species. Secondary antibodies that contain oligonucleotide-coupled probes are then added and bind conservative regions of the primary antibodies. When proteins interact, the oligonucleotide-coupled antibodies come into proximity and ligate into a circular DNA. The functional distance to allow hybridization and ligation is usually between 5 and 10 nm. This DNA is further amplified with a DNA polymerase and complementary fluorescent-labeled oligonucleotides are used to bind and detect DNA. This method provides a stable and high signal-to-noise ratio and allows the visualization and analysis of endogenous proteins, as they do not need to be tagged with a fluorescent marker. A limitation of this method is associated with the use of antibodies and their potential cross-reactivity (Reviewed in [86]).

Fluorescence and Bioluminescence Resonance Energy Transfer (FRET and BRET)

FRET is a non-radiative energy transfer process, from an excited donor fluorescent molecule, to another acceptor fluorescent molecule. This photophysical effect requires that i) the emission spectrum of the donor fluorophore overlaps the absorption spectrum of the acceptor fluorophore, ii) the donor and acceptor transition dipole orientations are approximately parallel, and iii) the donor and acceptor molecules are in close proximity (less than 10 nm apart). When coupled to two interacting proteins, the fluorescent molecules are brought to close proximity and induce a FRET signal. The most common pairs of donor and acceptor fluorophores are GFP and its derivatives [87]. BRET is very similar to FRET, but in this method occurs the transfer of energy between a luminescence donor and a fluorescence acceptor. As the BRET donor does not require an external light source, it prevents some issues associated with FRET such as autofluorescence, light scattering or photobleaching, and therefore represents an interesting alternative to that method [88]. Both techniques are very useful to monitor spatiotemporal changes in PPIs through microscopy and spectroscopy. However, they require a careful optimization of sensor FRET and BRET pairs, as well as significant image processing work, that can limit their use for high-content assays [82][89].

Protein-fragment Complementation Assays (PCAs)

Bimolecular Fluorescence Complementation (BiFC)

Protein-fragment complementation assays (PCAs) are a family of split systems where a reporter protein with enzymatic or fluorescent properties is split into non-active or non-fluorescent fragments, which are fused to interacting proteins of interest. When interaction occurs, the two fragments are brought into close proximity and reassemble spontaneously into a functional biosensor. BiFC is a type of PCA assay based on the reconstitution of a fluorescent protein, such as GFP. The BiFC assay also allows a direct visualization of spatial and temporal interaction between two interacting proteins *in vivo*, and the fluorescence signal provides a sensitive readout for detecting protein-protein interactions, even at a low expression level. Another advantage is the fact that the intensity of the fluorescence signal is proportional to the strength of the PPI [90]. A major limitation of this method is the spontaneous assembly of the fluorescent protein fragments prior to interaction, due to their large, bulky, and consequently aggregation-prone profile, which leads to enhanced background signals. Using smaller tags can reduce poor folding and self-assembly seen with split GFP [91]. This two-part split GFP is not well suited for bait and prey that

are expressed at different times or locations in a cell, because the fluorescent protein fragments irreversibly aggregate prior to interaction [85].

Tripartite Split-GFP (TriSFP)

The triSFP complementation assay, introduced by Cabantous *et al.* [91] in 2013, is a PCA-based very promising tool to study PPIs, *in vitro* and in living cells, that overcomes the poor folding and the background fluorescence arising from self-assembly observed in BiFC. This assay is based on the association of three fragments of the GFP protein: two short peptides, GFP10 and GFP11 (around 20 amino acids long), each tagged to one of the interacting partners ("bait" and "prey" proteins), and a third complementary large GFP1-9 detector fragment. When proteins interact, GFP10 and GFP11 assemble, allowing spontaneous self-association with GFP1–9 fragment to form a full-length GFP, and fluorescence is emitted. If proteins A and B do not interact, GFP10 and GFP11 are not tethered and entropy is too high to allow complementation with GFP1–9 [91] (**Figure 4**).

The main advantage of this system is the use of small fusion tags that reduce folding interference. This characteristic allows the identification and characterization of stable and soluble protein complexes. Moreover, the spontaneous assembly of the three-body system is unlikely to occur without the interaction between the proteins fused to GFP10 and GFP11 [85]. The third split-GFP detector fragment (GFP1-9) confers a strong specificity with low background signals in living cells [91].

Due to the irreversible assembly of the 3-body GFP fragments, the stability of the triSFP enables integration, accumulation, and subsequent detection even of transient interactions and low affinity complexes. The triSFP system can also be exploited to turn on the detection of PPIs by inducing the expression of the GFP1-9 fragment [91]. This assay is straightforward and flexible, and not only provides a convenient approach to detect and visualize PPIs in living cells in a robust way, as it can also serve as an ideal tool to monitor the dynamics and modulation of those interactions. Therefore, although advances are continuously being made with a lot of the assay technologies previously referred, a growing interest can be anticipated in the application of PCA-based high-throughput drug screenings, mainly of triSFP assay, with the goal of identifying compounds that selectively modulate relevant PPIs in living cells. Indeed, PPI modulation projects are likely to increase within the drug discovery field as they offer higher selectivity and as high-content methods continue to be improved to target PPIs involved in several diseases.

Therapy development for MJD

MJD and other polyQ diseases have a huge impact on the lives of patients and their families, as they slowly lead to a state of incapacity and ultimately death. Unfortunately, even though research worldwide is progressively leading to a better understanding of neurological diseases' mechanisms, there is still no disease-modifying treatment available for these disorders. Currently, the treatment approaches are mainly based on the management of the symptoms, increasing patient's quality of life. Patients can receive genetic counselling as well as to do speech therapy, physical exercise and physiotherapy [10], [92], [93]. In terms of pharmacological approaches, those include specific drugs to treat parkinsonian signs, restless legs syndrome and other motor-related symptoms [94]–[96].

Having in mind the goal of finding a disease-modifying treatment, it is necessary to understand the pathological mechanisms underlying the disease, overcome research challenges, and explore possible therapeutic strategies. PPIs might be the key to unravel this complex challenge.

AIMS

The role of ATXN3 PPIs in MJD pathogenesis is still not clear. Since recent data suggests that the protein may play important physiological roles in the nucleus of neuronal cells through the interaction with several partners, we hypothesize that mutant ATXN3 interacts abnormally with different proteins in that compartment and those interactions disrupt cellular pathways and contribute significantly to their neuronal toxicity.

Therefore, with this work we aimed to develop an assay based on the detection of relevant ATXN3 PPIs, that can be used in the future to screen compounds that modulate those interactions, either by stabilizing or disrupting them. For that, we aimed to build and to characterize a triSFP-based system in a mammalian cell line, and further validate the system by different techniques and with appropriate controls.

The splicing factor 9G8 is a previously identified relevant nuclear interactor of ATXN3, as it was observed that its ubiquitylating pattern is altered when the expression of ATXN3 is silenced [45]. As previously described, 9G8 has an evident role in the pathogenesis of several neurodegenerative diseases as it contributes for faulty splicing of the protein tau. This led us to hypothesize a role for 9G8 in MJD by interacting with ATXN3, which in turn motivated us to explore that interaction with the triSFP system and validate their nuclear encounter.

We expect that this work constitutes a starting point to future work aiming to identify novel therapeutic targets and promising modulators of PPIs, to test those compounds in MJD animal models, and to further translate them to the clinical context, with the aim of preventing or delaying the characteristic neurodegeneration of this disease, that remains fatal and untreatable, even though its genetic cause is known for decades.

Additionally, the successful generation and optimization of the triSFP-based ATXN3 PPI detection system should allow the production of a major tool for our laboratory to be used in this and other previously described applications, namely with the main goal of understanding the molecular mechanisms that drive ATXN3 aggregation and toxicity.

MATERIALS AND METHODS

Tripartite-split GFP system

As previously explained, in the triSFP system, the β-strands 10 and 11 of GFP (GFP10 and GFP11) are fused to "bait" and "prey" proteins and expressed in a mammalian cell line that co-expresses the "detector" fragment β-strand 1-9 (GFP1-9). When protein interaction occurs, GFP10 and GFP11 are in proximity, causing them to assemble spontaneously and with the GFP1-9 fragment to form a full-length GFP, and fluorescence is emitted. If proteins A and B are not interactors, GFP10 and GFP11 do not bind, and the high entropy does not allow complementation with GFP1-9 **(Figure 4)** [91].

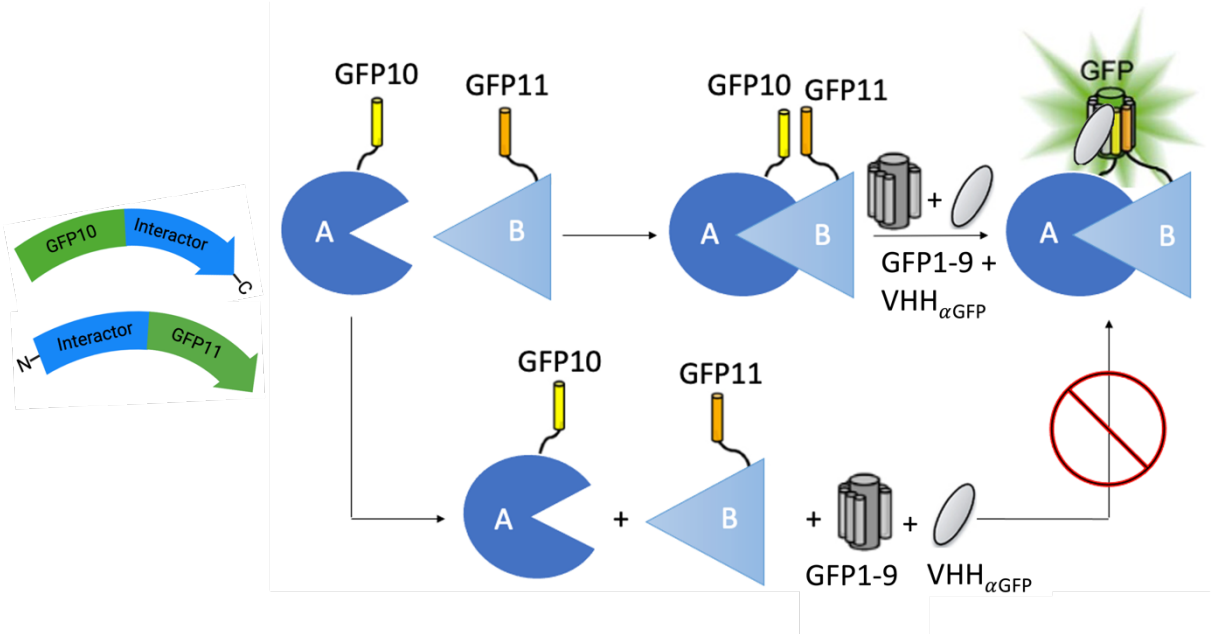


Figure 4 - Tripartite split-GFP complementation assay. β-strand 10 (GFP10) and β-strand 11 (GFP11) are fused to bait (A) and prey (B) proteins, respectively, and the detector fragment GFP1-9 (β-strand 1-9) and an anti-GFP VHH enhancer are added or expressed separately. When proteins A and B interact with each other, GFP10 and GFP11 assemble and then spontaneously associate with GFP1-9 and the anti-GFP VHH fragment to form a full-length GFP. If proteins A and B do not interact, GFP10 and GFP11 are not tethered, and entropy is too high to allow complementation with GFP1-9. GFP10 and GFP11 were cloned in the N and C-terminus, respectively, of each interactor. Image adapted from [91].

The known interactors of ATXN3 selected for this study were tubulin, the major constituent of microtubules, engaged in several cytoskeletal functions [97]; the human homolog of Saccharomyces cerevisiae Rad23 hHR23A, a protein associated with the proteasome and DNA repair, and with known activity in the endoplasmic reticulum associated protein degradation (ERAD) [98], [99]. The candidate interactor in study was the serine/arginine-rich splicing factor 9G8, involved in alternative splicing of Tau

[68]. Tubulin, hHR23A, and ATXN3 itself are confirmed interactors of ATXN3 that were used as positive controls in the triSFP system, and 9G8 is a previously identified relevant nuclear interactor of ATXN3, whose more detailed study might unravel relevant mechanisms for MJD. GCN4, a leucine zipper and transcriptional activator of amino acid biosynthesis in yeast that stabilizes and specifies protein dimer formation, and that does not have evidence of interacting with ATXN3, was also tested in the system as an ATXN3 interactor [100] (**Figure 5**). This zipper was already present in the original vectors with GFP10 and GFP11, kindly provided by Dr. Stéphanie Cabantous from the Cancer Research Center of Toulouse [91].

ATXN3 (WT – 14Q, and mutant form – 78Q) and the selected interactors were cloned in frame with GFP, creating two constructs for each protein, one with GFP10 attached to the N-terminus, and other with GFP11 attached to C-terminus. This provides more robustness to the system, as the site of interaction between ATXN3 and 9G8 is unknown. The use of both forms of ATXN3, WT and mutant form, can constitute a valuable characteristic of this system for future comparisons of their interactions. Every construct is detailed at **Table 2** (Supplementary Information).

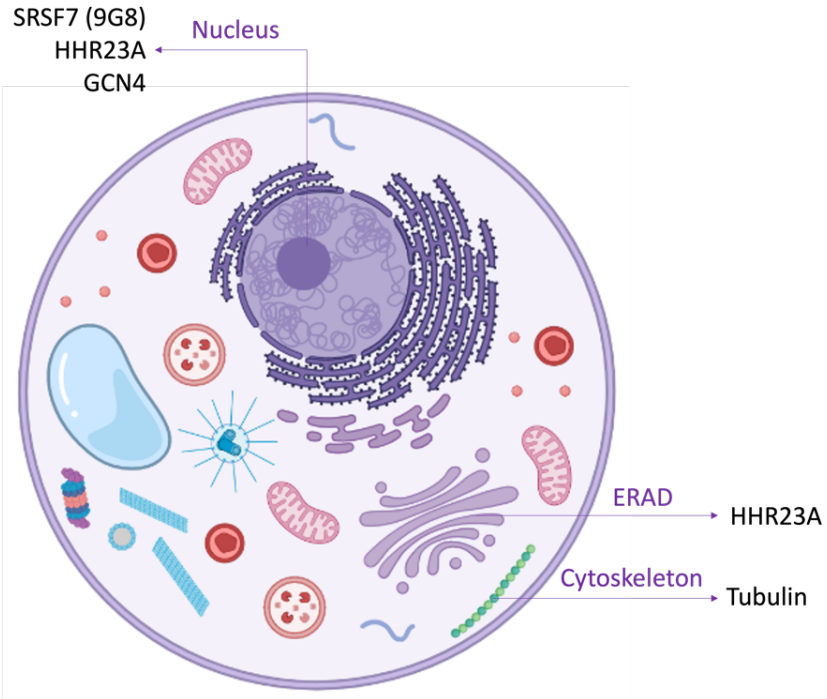


Figure 5 - Proteins selected for the study of its interaction with ATXN3 in the tripartite split-GFP system. 9G8 is a relevant nuclear ATXN3 interactor with a possible role in MJD neuropathogenesis, and tubulin and hHR23A were selected for positive controls as they are known ATXN3 interactors in the cytoplasm regarding cytoskeleton organization, as for tubulin, and, concerning hHR23A, in the endoplasmic-reticulum-associated protein degradation (ERAD) pathway at the cytoplasm, as well as at the nucleus for DNA repair functions. Although evidence of GCN4 interacting with ATXN3 does not exist, we decided to test it as an interactor of ATXN3. Image was constructed with BioRender.com.

Cloning strategy and experimental procedure

Vector design and cloning of genes of interest in the vectors in frame with GFP10 and GFP11 tags was performed based on Gibson Cloning, a robust exonuclease-based method that allows the assembly of any desired recombinant DNA fragment. The reaction involves a 5' exonuclease that generates overhangs, a polymerase that fills the gaps of the annealed single strand regions, and a DNA ligase that seals the nicks of the annealed and filled-in gaps (**Figure 6**) [101].

Firstly, the original vectors were modified to produce plasmids for general use, that would easily allow us to express any protein in frame with GFP10 and GFP11 (**Figure 7**). For that, pcDNA GFP10-zipper vector was cut with Fast Digest Kpn2I (BspEI) (FD0534) and XbaI (FD0684) enzymes, and the annealed oligonucleotides (5' – CCGGCGGCGGTGGAT – 3' and 5' – CTAGATCCACCGCCG – 3') were inserted. Oligonucleotides were annealed by mixing equal volumes in equimolar concentrations (2 ug each in a total volume of 50 ul), incubating the mix at 95°C for 5 minutes, and slowly cooling to room temperature. Annealed oligos were inserted through a 1:4 ligation, with incubation with T4 DNA ligase enzyme (1 Weiss Units/ul) for 10 minutes at 22°C. The reaction with the vector was further cut with Kpn2I enzyme and transformed. As the inserted oligonucleotides eliminate this site, only vectors with the successful insert of the linkers will provide *E. coli* resistance colonies after transformation. pcDNA zipper-GFP11 vector was cut with Fast Digest EcoRI (FD0274) and ClaI (Bsu15I) (FD0143) enzymes, annealed oligonucleotides (5' – AATTGCCACCATGGCTAT – 3' and 5' – CGATAGCCATGGTGGC – 3') were inserted, and the vector was further cut with EcoRI. As the linkers eliminate this site, only vectors with the successful insertion of the linkers will provide *E. coli* resistance after transformation. Furthermore, those linkers added the Kozak sequence, which functions as the protein translation initiation site.

Heat-shock transformation was performed to produce and isolate the modified pcDNA GFP10 and pcDNA GFP11 vectors. After thawing cells on ice for 20 minutes, 5 ul of vector was added to 50 ul of competent *E. coli* Dh5α cells, incubated 30 minutes on ice, and heat-shocked at 42°C for 45 seconds. By exposing cells to the heat shock, a pressure difference between the outside and the inside of the cell is created, which induces the formation of pores and allows the plasmid DNA to enter. Cells were then immediately put on ice for another 30 minutes to retain the plasmids inside the bacteria. Next, Luria Bertani (LB) medium was added, and bacteria was incubated at 37°C with agitation for at least 1 hour, to let them recover from heat shock and allow the production of the antibiotic resistance gene. After the incubation, 100 ul of cell suspension was spread onto agar medium containing ampicillin and grown overnight for about 12-18h at 37°C. In the following day, grown colonies were selected with a sterile

pipette tip and incubated overnight with agitation in a 50 ml falcon with 5-10 ml of LB liquid medium with ampicillin at 100 ug/ml (from stock of 100 mg/ml).

Plasmid isolation was performed the next day using the NucleoSpin® Plasmid QuickPure kit (Macherey-Nagel, 740588.250) following manufacturer instructions. In brief, pelleted bacteria from the 5-10 ml cell suspension were ressuspended in 500 ul of Buffer A1 and plasmid DNA was liberated from the *E. coli* host cells using 500 ul of Buffer A2, through SDS/alkaline cell lysis. 600 ul of buffer A3 was added to neutralize this lysate and allow DNA to bind the silica membrane of the columns. After a 10 minute centrifugation at 11.000 x g, supernatant was pipetted onto the NucleoSpin® Plasmid / Plasmid (NoLid) Column, centrifuged for 1 minute at 11.000 x g, the supernatant was discarded, and the column was washed with 600 ul of ethanolic Buffer A4, that removes salts, metabolites, and other soluble macromolecular cellular components. After centrifugation, ethanol in that buffer was removed with another centrifugation before eluting the DNA, as it might inhibit enzymatic reactions if not completely removed. Pure plasmid DNA was eluted to a new microcentrifuge tube with 50 ul of Buffer AE, and quantified using the Spectrophotometer NanoDrop™ (ThermoFisher Scientific). DNA quality was verified by the A260/A280 and A260/239 ratios.

As pcDNA GFP10 vector had its XbaI recognition site blocked by dam methylation, it had to be transformed again in dam- *E. coli* cells, namely the GM1519 F- (lambda cI857 S7)GM119 (dam-3 dcm-6 metB1 lacY1 galK2 galT22 tonA31 tsx-78 supE44 mtl-1 (thi-1)) *E. coli* K-12 strain. For this transformation the technique of electroporation was used. To remove salts and other components that might interfere with electroporation, cells were first grown up to 0.5 Optical Density (OD) at 600 nm, and then submitted to a sequence of centrifugations and washes with ice cold distilled water and glycerol at 10%. 25 ul of electrocompetent and thawed cells were mixed with approximately 100 pg of the DNA solution, and the mix was transferred into a chilled cuvette. Electroporation was performed with the following conditions: Voltage 1.8 kV, shunt resistor 200 Ω, and capacitor 25 μM. The typical time constant resulting from this setting is 4 milliseconds. 975 ul of preheated SOC (Super Optimal broth with Catabolite repression) medium was immediately added, gently mixed, and the content was transferred to a 1.5 ml microcentrifuge tube. The tubes were incubated at 37°C for approximately 2 hours, and 100 ul of cell suspension was spread onto agar medium containing ampicillin. In the following day, grown colonies were selected with a sterile pipette tip and incubated overnight with agitation in a 50 ml Falcon tube with 5-10 ml of LB liquid medium and 5-10 ul of ampicillin at 100 ug/ml. DNA purification was performed as previously described, quantified, and sequenced with the indicated primers (**Table 2,** Supplementary Information).

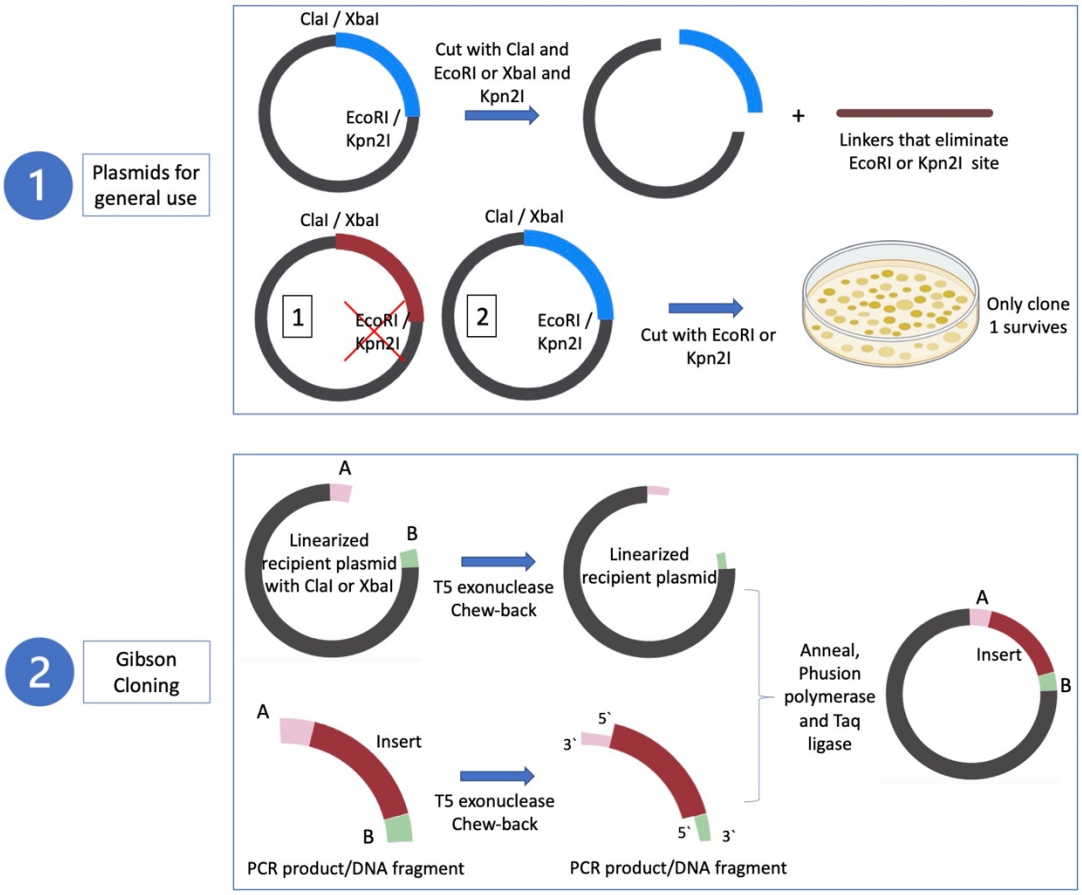


Figure 6 - Cloning strategy for the construction of vectors for the tripartite split-GFP system. 1 – Vectors pcDNA GFP10-zipper and pcDNA zipper-GFP11, sent by Dr. Stéphanie Cabantous, were modified to produce plasmids for general use, that will easily allow us to express any protein in frame with GFP10 and GFP11. 2 - Gibson Cloning for the insertion of genes of interest amplified by PCR, in the linearized produced plasmids. This technique employs three enzymatic activities in a single-tube master mix reaction: 5´ exonuclease, the 3´ extension activity of a DNA polymerase and DNA ligase activity. The 5´ exonuclease activity chews back the 5´ end sequences and exposes the complementary sequence for annealing. The polymerase activity then fills in the gaps on the annealed regions. A DNA ligase then seals the nick and covalently links the DNA fragments together. Image was constructed with BioRender.com.

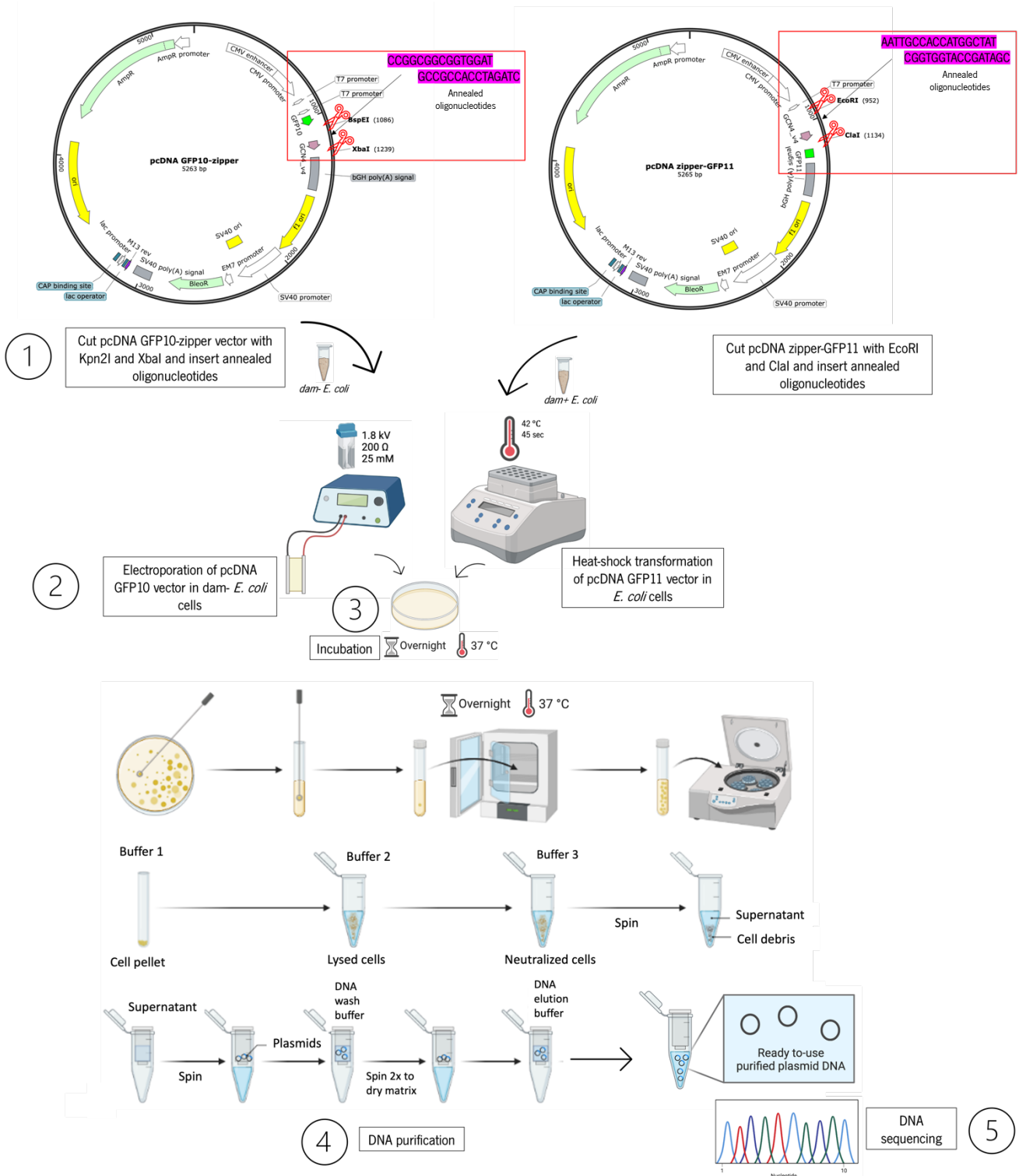


Figure 7 - Experimental procedure used to produce the vectors pcDNA GFP10 and pcDNA GFP11. GCN4 zipper from pcDNA GFP10-zipper and pcDNA zipper-GFP11 vectors was removed and the corresponding oligonucleotides were inserted (1). Electroporation of pcDNA GFP10 vector in *dam- E. coli* cells, and heat-shock transformation of pcDNA GFP11 vector in *dam+ E. coli* cells was perfomed to produce the constructed vectors (2). The bacteria that incorporated the vectors were incubated in an LB plate with ampicillin (3), and DNA was purified after growth of selected colonies (4). DNA sequencing allowed us to confirm the construction of the desired vectors (5). Image was created with SnapGene® software (from Insightful Science; available at snapgene.com) and BioRender.com.

After preparation of the vectors, we proceeded with cloning of the genes of interest upon PCR amplification with the indicated primers (**Table 1**) (**Figure 8**). PCR reactions were performed using Platinum SuperFi II PCR Master Mix (Thermo Fisher, 12368010), or with Phusion®High Fidelity PCR Kit (New England Biolabs (NEB), E0553S/L), with which ideal annealing temperature was obtained through gradient PCR (**Figure 36 to 39**, Supplementary Information).

PCR cycling conditions for Platinum enzyme were the following: 98°C for 30 seconds as initial denaturation, followed by 35 cycles of denaturing at 98°C for 10 seconds, annealing and extension at 72°C for 1 minute (2-step protocol), and a final extension at 72°C for 5 minutes. PCR cycling conditions for Phusion enzyme were as follows: 98°C for 30 seconds as initial denaturation, followed by 35 cycles of denaturing at 98°C for 10 seconds, annealing for 30 seconds with the temperature defined by gradient PCR, extension at 72°C for 1 minute, and a final extension at 72°C for 10 minutes.

For Gibson Cloning, the PCR primers used were especially long, as they should be up to 65 nucleotides in length. 20-40 nucleotides provide the requisite homology at the 5´end for the adjacent DNA fragment (vector), and 18-25 nucleotides are specific to the DNA element. PCR amplification was previously tested *in silico* using SnapGene® software (Insightful Science; available at snapgene.com).

Table 1 - Origin of each gene of interest, primers for their PCR amplification, and their product size.

Gene	Gene Origin	Forward Primer (5' to 3')	Reverse Primer (5' to 3')	Product size (bp)
WT ATXN3 for GFP10 vector	Vector pBRIT TAP-ATXN3 14Q Isoform 2 (3 UIMs)	CCTCCGGCGGCGGTGG ATCTAGAATGGAGTCCA TCTTCCACGAG	TCAGCGGGTTTAAACGGGC CCTCTATTTTTTTCCTTCTG TTTTCAAATC	1131
WT ATXN3 for GFP11 vector	Vector pBRIT TAP-ATXN3 14Q Isoform 2 (3 UIMs)	GGTGGAATTGCCACCAT GGCTATGGAGTCCATCT TCCACGA	GACCCACCACCTCCAGAG CCACCGCCACCATCTTTTT TTCCTTCTGTTTTCAAATC	1136
9G8 for GFP10 vector	Vector pIC111-9G8	GGTCCTCCGGCGGCGG TGGATCTATGTCGCGTT ACGGGCGGTACG	GGGTTTAAACGGGCCCTCT AGTCAGTCCATTCTTTCAG GACT	761
9G8 for GFP11 vector	Vector pIC111-9G8	GGAATTGCCACCATGG CTATCATGTCGCGTTAC GGGCGGTACGG	CCTCCAGAGCCACCGCCA CCATCGTCCATTCTTTCAG GACTTGCAC	758
Tubulin for GFP10 vector	cDNA of human SH-SY5Y cells	CCTCCGGCGGCGGTGG ATCTAGAATGCGTGAGT GCATCTCCATCCAC	TCAGCGGGTTTAAACGGGC CCTATGTATTCCTCTCCTT CTTCCTC	1400
Tubulin for GFP11 vector	cDNA of human SH-SY5Y cells	GGAATTGCCACCATGG CTATCATGCGTGAGTGC ATCTCCATCCA	ACCACCTCCAGAGCCACC GCCACCATCGTATTCCTCT CCTTCTTCCTC	1401
hHR23A for GFP10 vector	cDNA of human SH-SY5Y cells	GGTCCTCCGGCGGCGG TGGATCTATGGCCGTCA CCATCACGCTC	GGGTTTAAACGGGCCCTCT AGTCACTCGTCATCAAAGT TCTG	1136
hHR23A for GFP11 vector	cDNA of human SH-SY5Y cells	GGAATTGCCACCATGG CTATCATGGCCGTCAC CATCACGCTC	CCTCCAGAGCCACCGCCA CCATCCTCGTCATCAAAGT TCTGACTC	1133
Mutant ATXN3 for GFP10 vector	Vector pBRIT TAP-ATXN3 78Q Isoform 2 (3 UIMs)	CCTCCGGCGGCGGTGG ATCTAGAATGGAGTCCA TCTTCCACGAG	TCAGCGGGTTTAAACGGGC CCTCTATTTTTTGCCTTCGG TTTTCAGGTC	1323
Mutant ATXN3 for GFP11 vector	Vector pBRIT TAP-ATXN3 78Q Isoform 2 (3 UIMs)	GGTGGAATTGCCACCAT GGCTATGGAGTCCATCT TCCACGA	GACCCACCACCTCCAGAG CCACCGCCACCATCTTTTT TGCCTTCGGTTTTCAGGTC	1328

After PCR reaction, every insert was subjected to digestion with the enzyme DpnI (ER1701) for 1 hour at 37°C, to digest the methylated (parental) DNA template and select for the newly synthesized DNA. Amplification was always confirmed using electrophoresis through a 1% (wt/v) agarose gel stained with GreenSafe and photographed. Solutions were purified using the GRS PCR & Gel Band Purification Kit (100 preps, GK01.0100), following the manufacturer's instructions. Briefly, 100 ul of the PCR reaction solutions were first mixed with 5 volumes of Gel Solubilization Solution to ensure optimal pH, as DNA binding is facilitated by pH <7.5. The sample mixture was then transferred to a DNA fragment mini spin column, centrifuged at 14.000 x g for 30 seconds, and the flow-through was discarded. 600 ul of wash buffer 2 was added to remove contaminations like salts and soluble macromolecular components, centrifuged at 14.000 x g for 30 seconds to discard the flow-through, and centrifuged again to dry the

matrix of the column. To elute DNA into a new tube, 50 ul of elution buffer was added directly to the centre of the spin column, incubated for 2 minutes, and centrifuged another 2 minutes at 14.000 x g. The elution buffer provides low ionic strength conditions that lead to the release of DNA from silica of the column. DNA was further quantified, and purity was analysed as previously described.

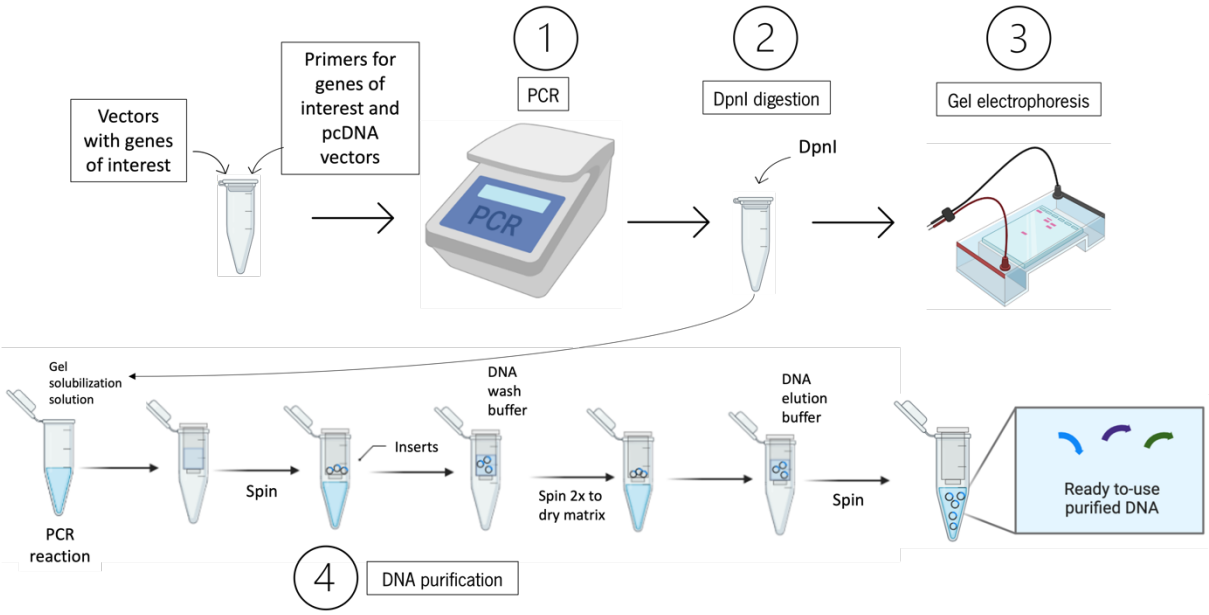


Figure 8 - Experimental procedure to isolate the genes of interest, to be inserted in the vectors pcDNA GFP10 and pcDNA GFP11. PCR reaction (1) was first performed to amplify the genes of interest, to be inserted in the final vectors pcDNA GFP10 and pcDNA GFP11, with the corresponding primers. The reaction was then submitted to DpnI digestion (2) to eliminate the parental methylated DNA, gel electrophoresis (3) allowed us to confirm the isolation of the gene of interest, and DNA was finally purified (4). Image was created with BioRender.com.

Next, the newly constructed pcDNA GFP10 and pcDNA GFP11 vectors were linearized to insert the genes of interest. pcDNA GFP10 vector was cut with XbaI enzyme, purified with the GRS PCR & Gel Band Purification Kit (GRiSP, 100 preps, GK01.0100), and mixed with the corresponding PCR product of the gene of interest, and GeneArt Gibson Assembly HiFi Master Mix (A46628). pcDNA GFP11 vector was cut with ClaI enzyme, purified with the same GRS PCR & Gel Band Purification Kit, and mixed with the corresponding PCR product of the gene of interest, as well as the same HiFi Master Mix. Next, commercial or laboratory-prepared competent bacteria were transformed with the cloned plasmids as described previously. *E. coli* cells were prepared in the laboratory by increasing the cell membrane permeability with calcium chloride ($CaCl_2$) and other supplements. DNA purification, as well as quantification, was performed as described previously (**Figure 9**).

Every construction was further confirmed, first by colony PCR, then by digestion with restriction enzymes, to analyze the expected fragment sizes, and finally by sequencing at STAB VIDA and further alignment using SnapGene® software (from Insightful Science; available at snapgene.com) with the appropriate primers (**Table 2**, Supplementary Information). For colony PCR, colonies were selected with a sterile pipette tip, and put in contact with 5 ul of pure DNase-free distilled water. 2 ul were used for PCR, and the other 3 ul for incubation with LB and the antibiotic ampicillin overnight for bacteria growth. PCR was performed with the primers used for gene isolation (**Table 1)** and a Taq DNA polymerase (Thermo Fisher, K0171). The cycling conditions were the following: 95°C for 3 minutes, to release the plasmid DNA from the bacteria to serve as PCR template, followed by 35 cycles of denaturing at 95°C for 30 seconds, annealing at 60°C for another 30 seconds, extension at 72°C for 1 minute and a final extension at 72°C for 10 minutes. Every cloning strategy was previously tested and performed *in silico* using the SnapGene® software (from Insightful Science; available at snapgene.com).

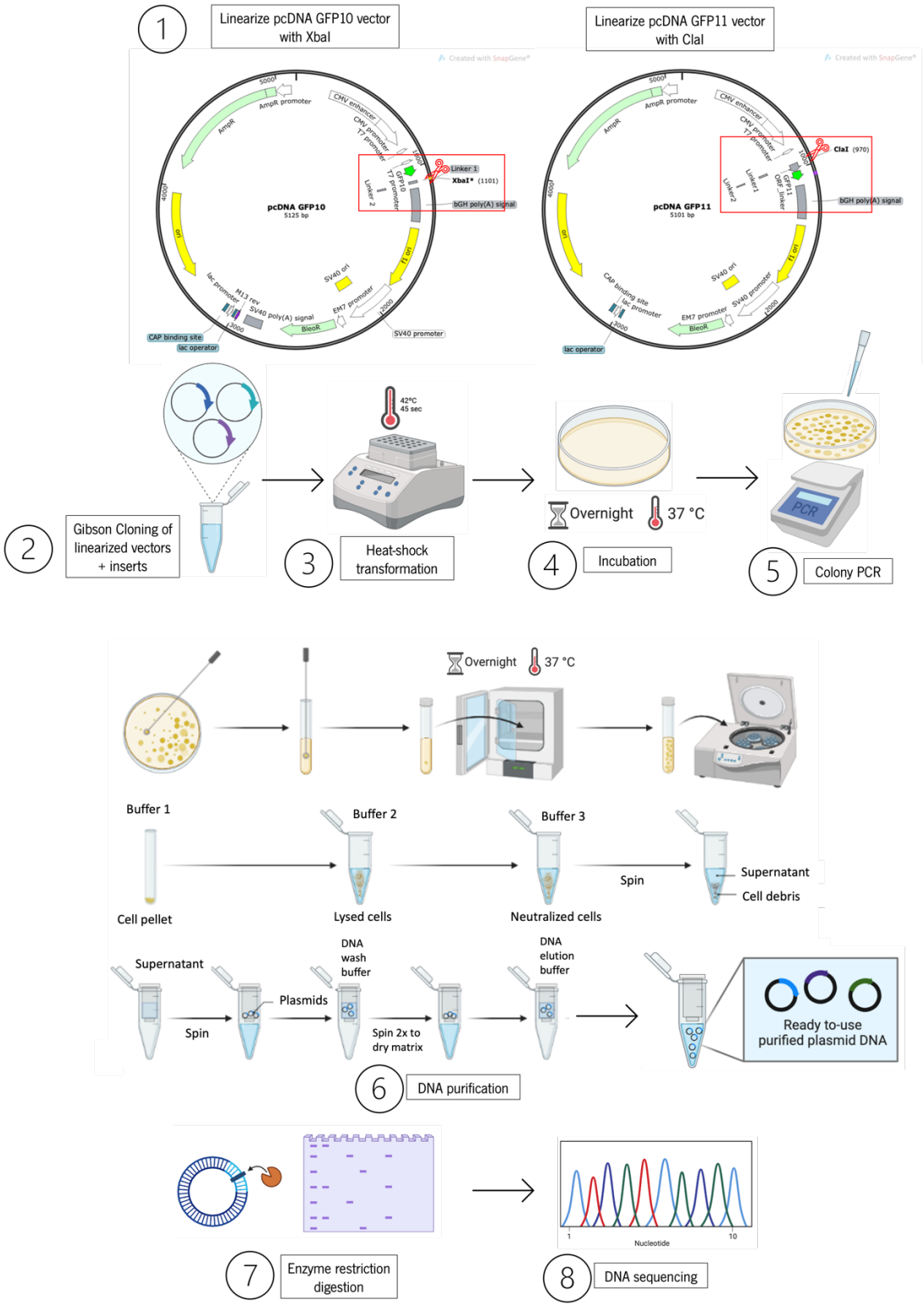


Figure 9 - Experimental procedure to produce the final vectors expressing GFP tags and the proteins of interest. pcDNA GFP10 and pcDNA GFP11 vectors were linearized (1) to insert the genes of interest by Gibson Cloning (2). After heat-shock transformation (3) and overnight incubation (4), colony PCR (5) was performed to confirm the insertion of the gene of interest before DNA purification (6). Restriction enzyme digestion (7) was performed to further confirm the successful vector construction before sequencing (8). Image was created with SnapGene® software (from Insightful Science; available at snapgene.com) and BioRender.com.

Transient Transfection

Transient transfections allow us to perform short-term studies of protein function in eukaryotic cells without viral methods. In this work we used a chemical method to transfect the vectors expressing our proteins of interest, that is based on creating an overall positive charge on the DNA molecules that we want to transfect, so that it can more easily cross the negatively charged cell membrane.

MRC5-SV (immortalized normal pulmonary human fibroblasts) cells expressing GFP1-9 and one single-domain antibody based on camelid heavy-chain antibodies (VHH or nanobody), engineered to boost GFP fluorescence by modulating the spectral properties of wild-type GFP, were kindly provided by Dr. Stéphanie Cabantous from the Cancer Research Center of Toulouse [82] (**Figure 66,** Supplementary Information). Cells were cultured in Opti-MEM™ I Reduced Serum Medium (Gibco, 31985-047), completed with 10% (v/v) fetal bovine serum (FBS) (Gibco,1050 0064), and 1% of Penicillin-Streptomycin (Gibco, 15140122), and routinely checked for contamination. The cell medium was supplemented with GlutaMAX (Gibco, 35050-061). Transfection of plasmids was performed using Fugene (FUGENE HD Transfection Reagent, Promega, E2312) according to the manufacturer's instructions.

For transient expression of interacting proteins, 2×10^5 MRC5-SV_1-9 cells per ml were seeded in ninety-six-well plates and transfected at a ratio 1:1 of the GFP10 and GFP11 fusion vectors (100 ng). ATXN3, either WT and mutant form, was transfected together with 9G8, the positive controls tubulin, hHR23A, ATXN3 itself, or the zipper GCN4. Proteins were all transfected attached to both GFP10 and GFP11, so blocking of a determined terminus would not influence interaction. pIC111+ATXN3 WT vector that express GFP was used as a transfection control, and the vector pcDNA GFP10-zipper-GFP11, also provided by Dr. Stéphanie Cabantous, controlled for GFP subunits assembly. For negative controls, ATXN3 (WT and mutant form) was transfected with an empty pcDNA GFP vector that does not express any fused protein, both empty pcDNA GFP vectors were transfected together, and pcDNA GFP vectors expressing WT ATXN3 were transfected alone (**Figure 10)**. Fluorescence was analyzed 24h after transfection using an Olympus IX81 Inverted Fluorescence Microscope.

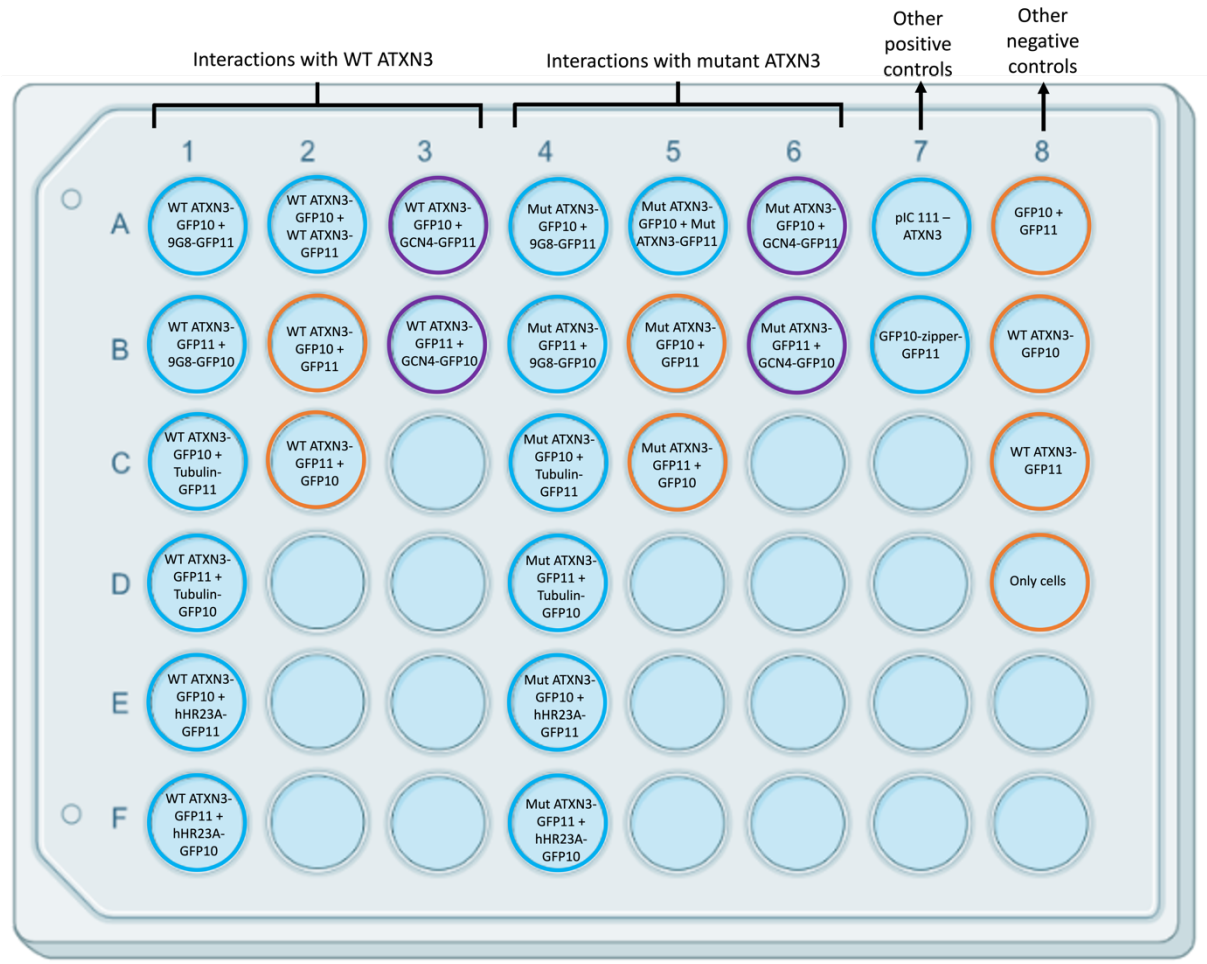


Figure 10 - Transfection scheme of the constructed vectors to characterize and optimize the triSFP system by analysing the different levels of fluorescence by flow cytometry and microscopy. Columns 1, 2 and 3 – Interactions with WT ATXN3. Blue wells: Transfection of WT ATXN3 with 9G8, tubulin, hHR23A, or ATXN3 itself, where fluorescence was expected to be seen. Orange wells: Transfection of WT ATXN3 with empty vectors that only express GFP10 or GFP11, where fluorescence must not be seen. Purple wells: Transfection of WT ATXN3 with GCN4, whose interaction was unknown. Columns 4, 5 and 6 – Interactions with mutant ATXN3. Blue wells: Transfection of mutant ATXN3 with 9G8, tubulin, hHR23A, and ATXN3 itself, where fluorescence was expected to be seen. Orange wells: Transfection of mutant ATXN3 with empty vectors that only express GFP10 or GFP11, where fluorescence must not be seen. Purple wells: Transfection of mutant ATXN3 with GCN4, whose interaction was unknown. Column 7 (blue wells)– Other positive controls. Transfection of PIC111-ATXN3, the vector that controls for the successful transfection, and GFP10-Zipper-GFP11, that controls for the successful assembly of GFP parts. Column 8 (orange wells) – Other negative controls. Transfection of empty vectors together that only express GFP10 or GFP11, transfection of each vector with WT ATXN3 alone to confirm that fluorescence is not detected without the remaining part of GFP, and cells only, without transfection or other condition, where fluorescence must not be seen.

Flow Cytometry of fluorescent cells

Flow Cytometry (FC) was used to identify the percentage of GFP positive cells in transfected MRC-SV GFP1-9 cells with the previously described constructs. After culture in a 96-well plate, plasmids were transfected, fluorescence was analysed by microscopy after 24 hours, and 48 hours later cells were analysed by FC. They were first detached from the plate with 30 ul of trypsin, and 100 ul of Optimem medium with FBS was then added, followed by up and down movements with the pippete. The suspension was then placed in microcentrifuge tubes, centrifuged at 900 g for 3 minutes, the medium was removed, and the pellet was resuspended with 150 ul of PBS 1x. Data was acquired in a LSRII flow cytometer (BD Bioscience) using FACS DIVA (BD Biosciences, San Jose, California, USA). The data was analysed using FlowJo Software (Tree Star, Ashland, OR, USA) as outlined in **Figure 11**.

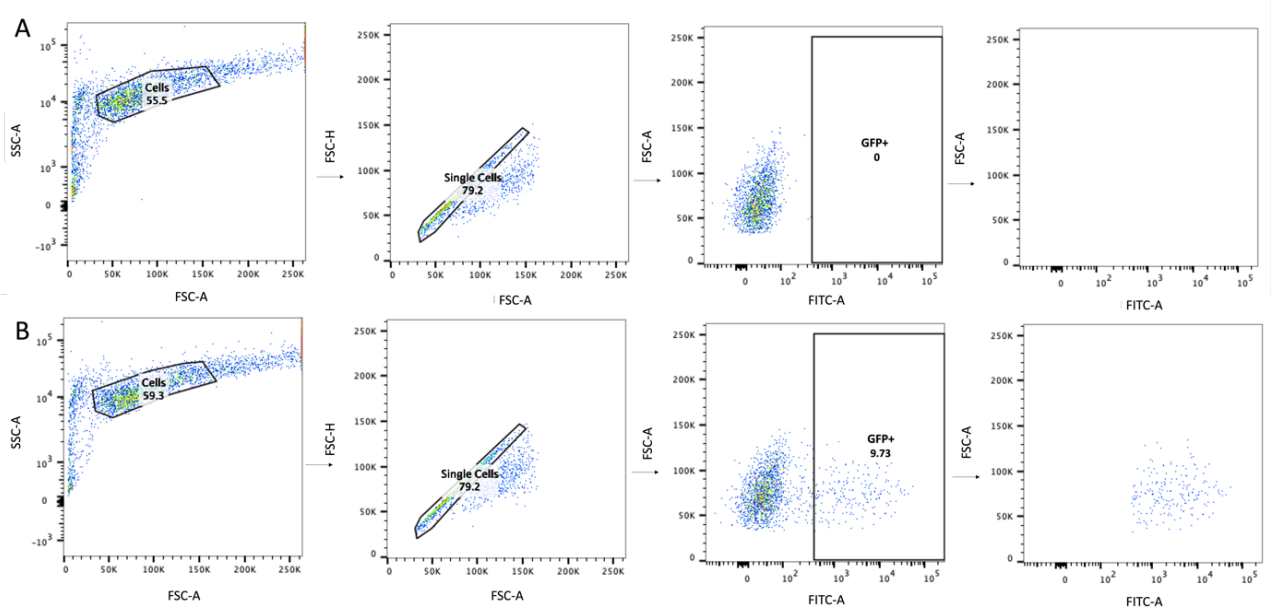


Figure 11 - Gating strategy to identify GFP positive cells using FlowJo software. A – Only cells. B – WT ATXN3-GFP10 + 9G8-GFP11. Gating for exclusion of debris and non-single-cell events were performed using forward vs side scatter (FSC vs SSC) area (A), and FSC-A vs FSC height (H) plots, respectively. FSC-A vs Fluorescein isothiocyanate (FITC)-A gate was defined based on the sample with only cells that express GFP1-9 (without transfection or other condition). The same strategy and gate were applied for every sample.

Immunofluorescence to validate transfection

Immunofluorescence staining of GFP10 was performed in cells expressing WT ATXN3–GFP11 and GFP10-9G8, and cells expressing GFP10-WT ATXN3 and GFP11 only, as well as WT ATXN3–GFP11 and GFP10 only, to validate the efficacy of transfection and expression of this tag.

After incubation of MRC-SV GFP1-9 expressing cells and transfection of the vectors in the following day, half of the medium was removed from the cultured cells and the same volume of 4% PFA in PBS was added for 10 minutes at RT. Subsequently, aspiration and incubation with 1000 ul of 4% PFA in PBS for 20 minutes at RT was performed, and fixed cells were rinsed twice with PBS 1x. To permeabilize cells, they were incubated with 0.5% Triton X-100 in PBS for 5 minutes at RT and rinsed twice in PBS 1x. Next, to block cells, they were incubated with 10% FBS in PBS for 60 minutes at RT and rinsed twice in PBS 1x. The cells were then incubated overnight at 4°C with the primary antibody anti-GFP10 from rabbit (diluted in 10% FBS in PBS at 1:2000). The day after incubation with primary antibody, cells were rinsed twice in 0.5% FBS diluted in PBS and incubated in secondary antibody (Goat anti-Rabbit IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor 594, A11012), diluted in 0.5% FBS in PBS at 1:1000, for 60 minutes at RT. Finally, cells were rinsed with 0.5% FBS in PBS. To analyze them, coverslips were mounted with mounting medium in a slide, and after dried, cells were analyzed under Olympus Widefield Inverted Microscope IX81.

DAPI nuclear staining

With the goal of confirming the nuclear location of the interaction between ATXN3 and 9G8, DAPI (4′,6-diamidino-2-phenylindole) nuclear staining was performed after transfection of the vectors that express those two proteins in fusion with GFP subunits. Cells were cultured in 24-well plates with sterilized coated coverslips. Coverslips were coated with 0.1% Gelatin from porcine skin (Sigma.Aldrich, G1890-100G) for 30 minutes in the incubator at 37°C before plating the cells. After incubation of cells and transfection of the vectors in the following day, half of the medium was removed from the cultured cells and the same volume of 4% PFA in PBS was added for 10 minutes at room temperature (RT) to fix cells. Subsequently, aspiration and incubation with 1000 ul of 4% PFA in PBS for 20 minutes at RT was performed, and fixed cells were rinsed twice with PBS 1x. To permeabilize the cells, they were incubated with 0.5% Triton X-100 in PBS for 5 minutes at RT and rinsed twice in PBS 1x. The cells were finally incubated with the DAPI counterstain (stock solution of 10 ug/mL) with a dilution of 1:2000, for 5 minutes at RT. To analyze cells, coverslips were mounted with mounting medium (Epredi Lab Vision PermaFluor Aqueous Mounting Medium, TA-030-FM) in a slide, and after dried, analyzed under Olympus LPS Confocal FV3000 microscope.

Site-directed mutagenesis to ATXN3 Nuclear Localization Signal

In order to further confirm and validate the nuclear interaction between ATXN3 and 9G8, a site directed mutagenesis was performed, to eliminate the nuclear localization signal (NLS) of ATXN3, both WT and mutant form. The NLS is a short peptide that facilitates the transport of ATXN3 from the cytoplasm into the nucleus of the cells. A previously described mutation of NLS sequence motif 282 Arg(R)-Lys(K)-Arg(R)-Arg (R) 285, shown to disrupt ATXN3 ability to enter the nucleus, was introduced at the conserved basic arginine (R) residue within this NLS, into a neutral threonine (T) residue (ATXN3 R282T) [28]. The mutation was inserted using the following designed primers for WT ATXN3: 5'-GTAGGCTTCTCGTCTCTTCGTAAGCTCTTCTGAAGTAAGA-3' and 5'-TCTTACTTCAGAAGAGCTTACGAAGAGACGAGAAGCCTAC-3', and the following for mutant ATXN3: 5'-TATGCTTCGCGACGTTTGGTCAGTTCTTCAGACGTCAG-3' and 5'-CTGACGTCTGAAGAACTGACCAAACGTCGCGAAGCATA-3'. Mutagenesis was performed in a PCR reaction with PfuTurbo DNA Polymerase (Agilent, 600250), DNA at 5 ug/ul, primers at 10 pm/ul, 10x cloned Pfu reaction buffer, and dNTPs (40mM), to a final volume of 50 ul. The PCR cycling conditions used were the following: denaturation at 95°C for 30 seconds, 16 cycles at 95°C for 30 seconds, annealing at 55°C for 1 minute, plus an extension at 68°C for 10 minutes, and a final extension at 68°C for 10 minutes. Following temperature cycling, the product vectors were treated with dpnI enzyme to digest methylated parental DNA, that does not include the desired mutation. The vectors were then transformed by heat-shock in *E. coli* cells, purified, quantified, and sequenced as previously described.

RESULTS

Vectors for triSFP system were constructed by Gibson Cloning

The previously described cloning strategy allowed us to successfully isolate every gene of interest for the triSFP system by PCR, namely ATXN3, 9G8 and the appropriate controls, and insert them in the corresponding vector, expressing GFP10 and GFP11 subunits. Every construct was confirmed by enzyme restriction digestion, in which we obtained the correct band sizes, as depicted in the following pictures **(Figure 12 to 24)**. Colony PCR to confirm insertion of gene in the respective vectors was performed before enzyme restriction digestion, with the results depicted in **Figures 40 to 49,** Supplementary Information. Sequencing of the newly inserted fragments confirmed the successful generation of the vectors (**Figures 60 to 61**, Supplementary Information).

pcDNA GFP10 vector

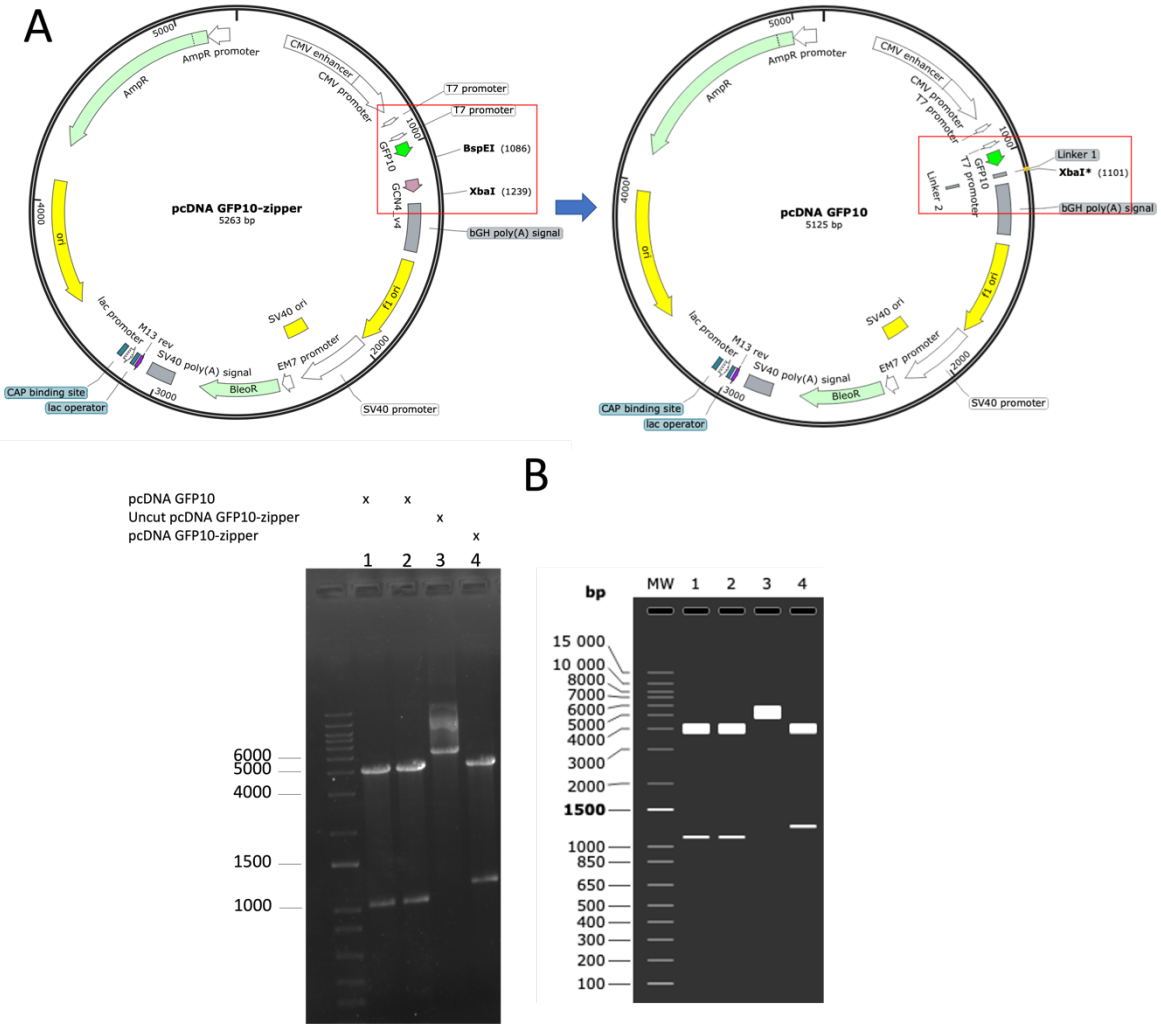


Figure 12 - Construction of the pcDNA GFP10 vector. A - pcDNA GFP10-zipper vector sent by Dr. Stéphanie Cabantous from the Cancer Research Center of Toulouse, with the backbone pcDNA 3.1; and pcDNA GFP10 vector after removal of GCN4 protein from the previous vector, with further addition of linkers that facilitate selection and future cloning. B – Enzyme restriction digestion to confirm the successful construction of the vector pcDNA GFP10. 1 and 2 – pcDNA GFP10 vector cut with BglII and ApaI enzymes (4026 bp + 1099 bp). 3 – Uncut pcDNA GFP10-zipper vector (5263 bp). 4 – pcDNA GFP10-zipper vector cut with BglII and ApaI enzymes (4026 bp + 1237 bp). Vectors and agarose gel with *in silico* confirmation of the expected bands after enzyme restriction digestion were obtained with SnapGene® software (from Insightful Science; available at snapgene.com).

pcDNA GFP11 vector

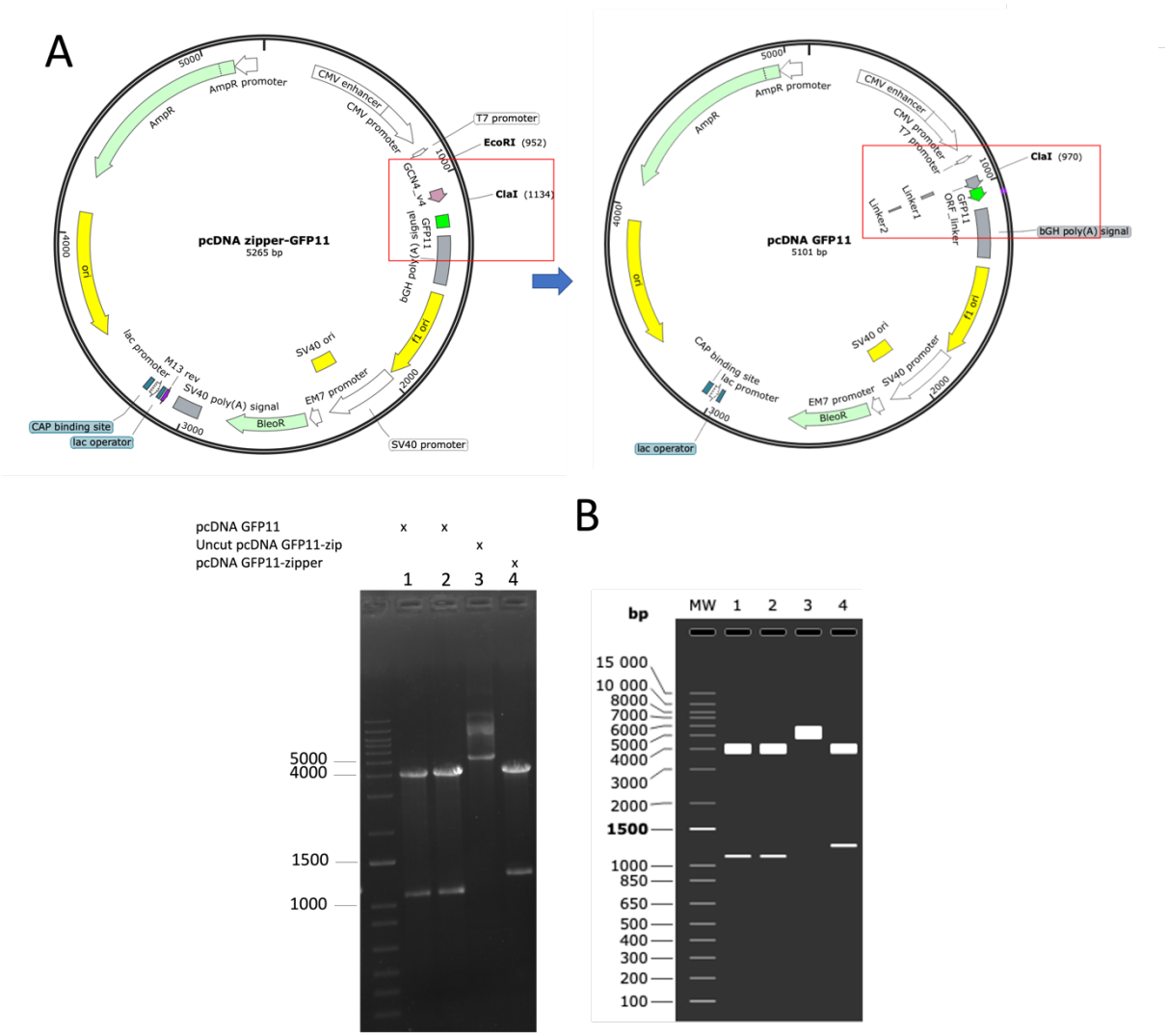


Figure 13 - Construction of the pcDNA GFP11 vector. A - pcDNA GFP11-zipper vector sent by Dr. Stéphanie Cabantous from the Cancer Research Center of Toulouse, with the backbone pcDNA 3.1; and pcDNA GFP11 vector after removal of GCN4 protein from the previous vector, with further addition of linkers that facilitate selection and future cloning. B – Enzyme restriction digestion to confirm the successful construction of the vector pcDNA GFP11; 1 and 2 – pcDNA GFP11 vector cut with BglII and ApaI enzymes (4026 bp + 1075 bp); 3 – Uncut pcDNA GFP11-zipper vector (5265 bp); 4 – pcDNA GFP11-zipper vector cut with BglII and ApaI enzymes (4026bp + 1239 bp). Vectors and agarose gel with *in silico* confirmation of the expected bands after enzyme restriction digestion were obtained with SnapGene® software (from Insightful Science; available at snapgene.com).

pcDNA GFP10 – WT ATXN3 vector

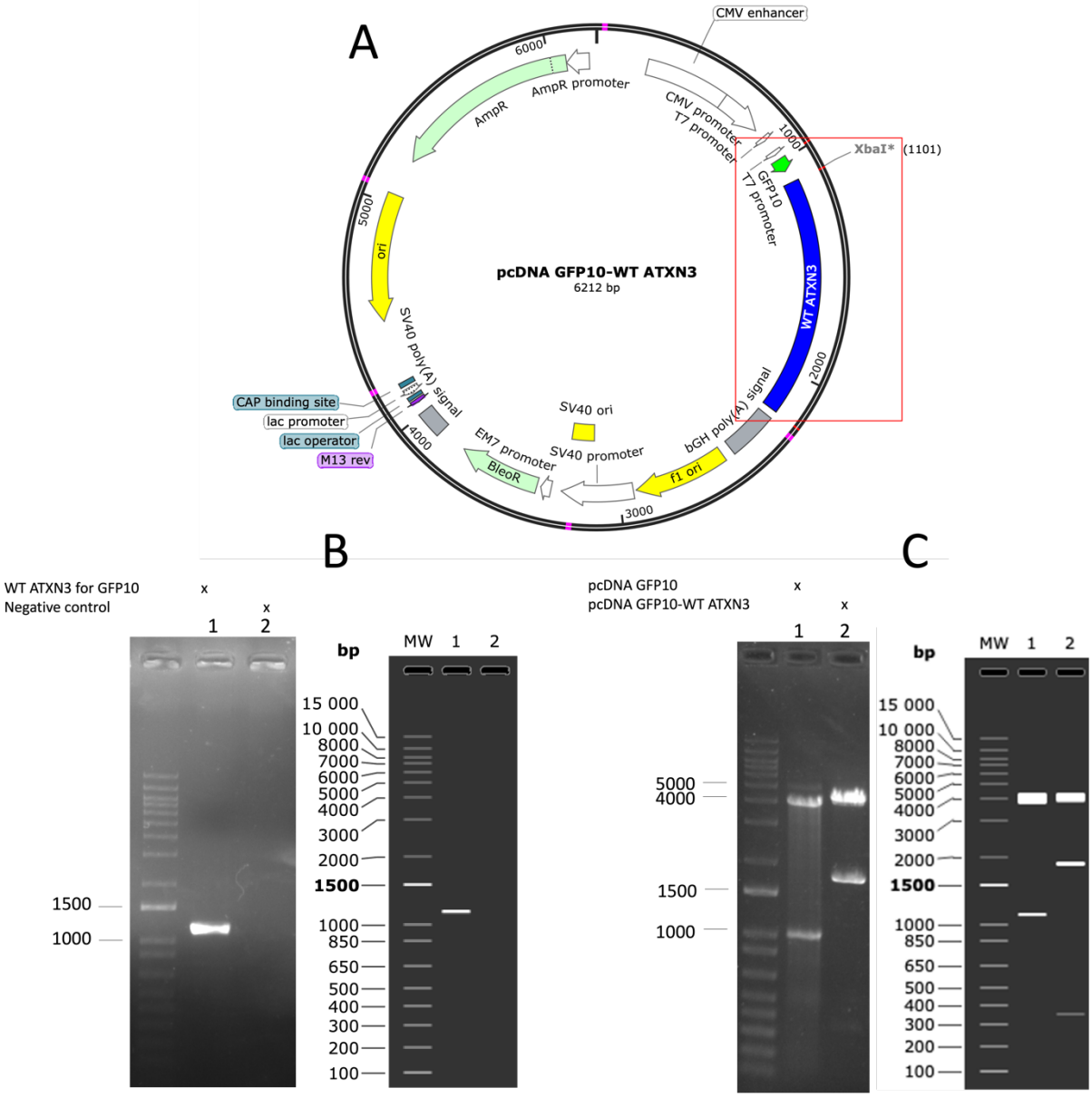


Figure 14 - Construction of the pcDNA GFP10-WT ATXN3 vector. A - pcDNA GFP10-WT ATXN3 vector built from pcDNA GFP10 vector after insertion of the WT ATXN3 cDNA (NM_004993.6). B – Gel electrophoresis from PCR amplification of WT ATXN3 for fusion with the GFP10 subunit, from the vector pBRIT TAP-ATXN3 WT; WT ATXN3 for GFP10 size is 1131 bp; 1 – WT ATXN3 for fusion with GFP10 amplification; 2 – Negative Control. C - Enzyme restriction digestion to confirm the successful insertion of WT ATXN3 in the pcDNA GFP10 vector before sequencing; 1 – pcDNA GFP10 vector cut with BglII and ApaI enzymes (4026 bp + 1099 bp); 2 - pcDNA GFP10-WT ATXN3 vector (6212 bp) cut with BglII and ApaI enzymes (4026 bp + 1697 bp + 458 bp). Vectors and agarose gel with *in silico* confirmation of the expected bands after enzyme restriction digestion were obtained with SnapGene® software (from Insightful Science; available at snapgene.com).

pcDNA WT ATXN3-GFP11 vector

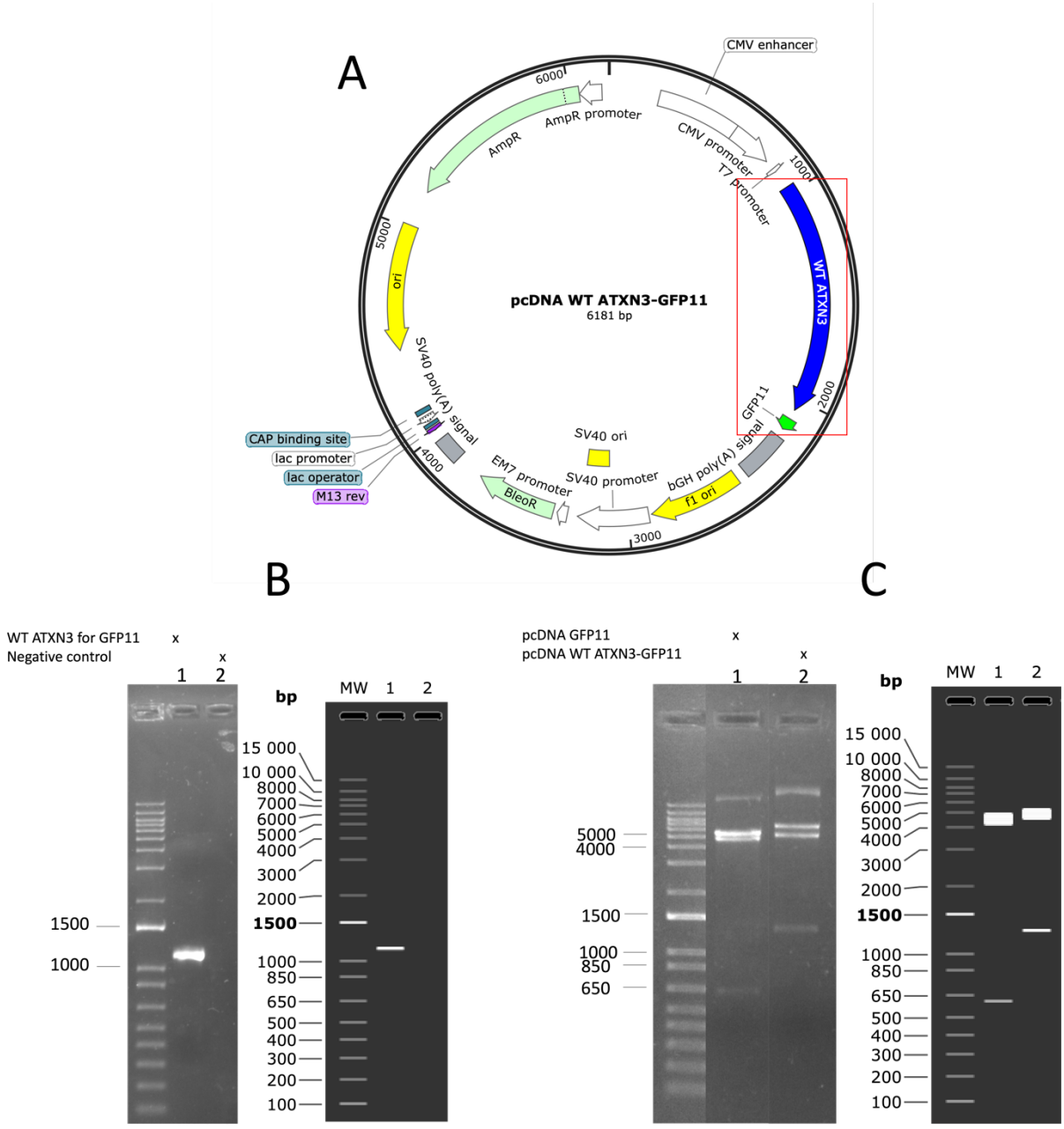


Figure 15 - Construction of the pcDNA WT ATXN3-GFP11 vector. A - pcDNA WT ATXN3-GFP11 vector built from pcDNA GFP11 vector after insertion of the WT ATXN3 cDNA (NM_004993.6). B – PCR amplification of WT ATXN3 for fusion with the GFP11 subunit from the vector pBRIT TAP-ATXN3 WT; WT ATXN3 for GFP11 size is 1136 bp; 1 – WT ATXN3 for fusion with GFP11 amplification; 2 – Negative Control. C - Enzyme restriction digestion to confirm the successful insertion of WT ATXN3 in the pcDNA GFP11 vector before sequencing; 1 – pcDNA GFP11 vector cut with ApaI and NdeI enzymes (4498 bp + 603 bp); 2 - pcDNA WT ATXN3-GFP11 vector (6181 bp) cut with ApaI and HindIII enzymes (4925 bp + 1256 bp). Vectors and agarose gel with *in silico* confirmation of the expected bands after enzyme restriction digestion were obtained with SnapGene® software (from Insightful Science; available at snapgene.com).

pcDNA GFP10 – 9G8 vector

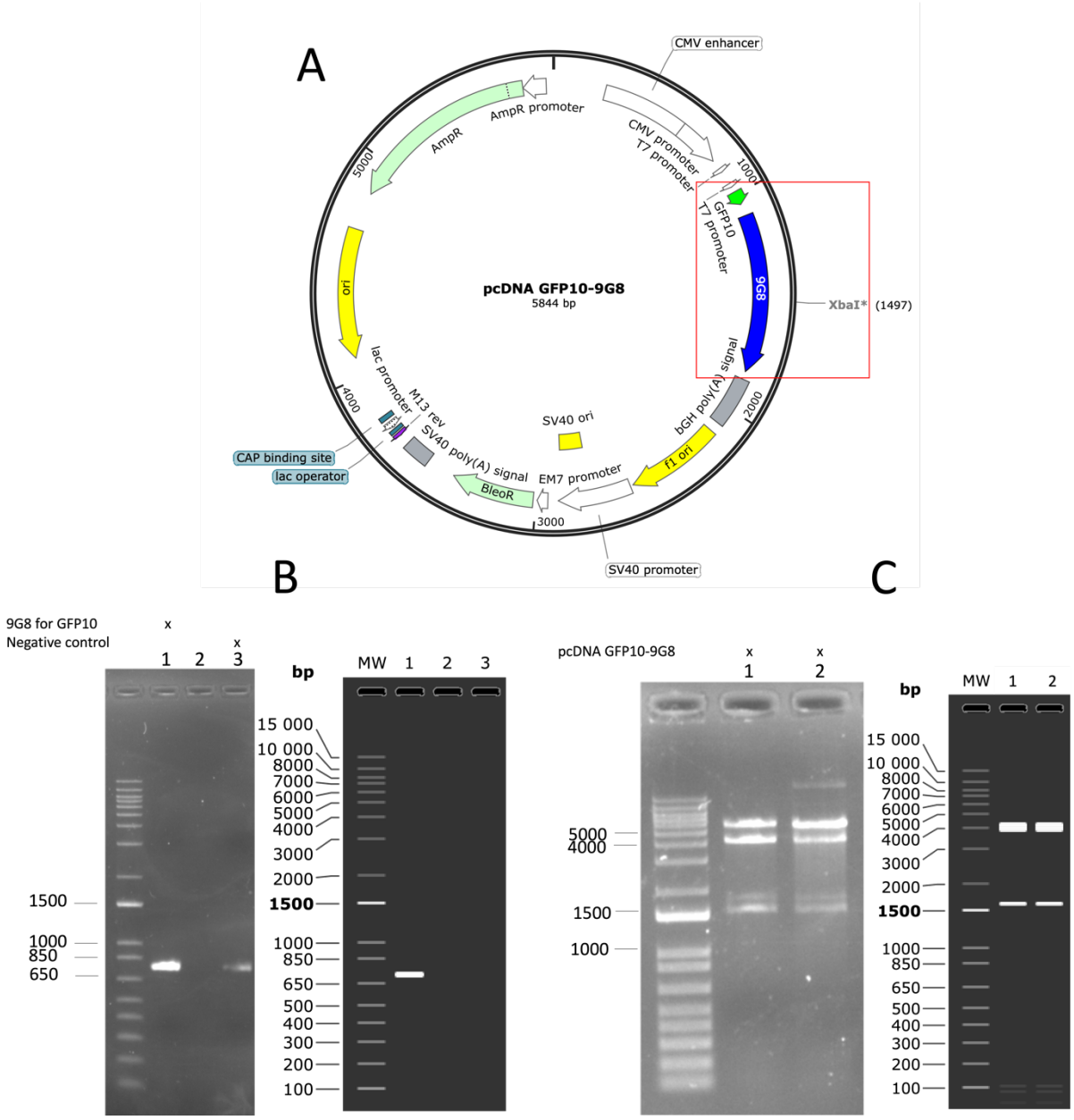


Figure 16 - Construction of the pcDNA GFP10-9G8 vector. A - pcDNA GFP10-9G8 vector built from pcDNA GFP10 vector after insertion of the 9G8 cDNA (NM_001031684.3). B – PCR amplification of 9G8 for fusion with the GFP10 subunit from the vector pIC111-9G8; 9G8 for GFP10 size is 761 bp; 1 – 9G8 for fusion with GFP10 amplification; 3 – Negative Control. C - Enzyme restriction digestion to confirm the successful insertion of 9G8 in the pcDNA GFP10 before sequencing; 1 and 2 – pcDNA GFP10-9G8 (5844 bp) cut with BglII and ApaI enzymes (4026 bp + 1581 bp + 108 bp + 87 bp + 42 bp). Vectors and agarose gel with *in silico* confirmation of the expected bands after enzyme restriction digestion were obtained with SnapGene® software (from Insightful Science; available at snapgene.com).

pcDNA 9G8 – GFP11 vector

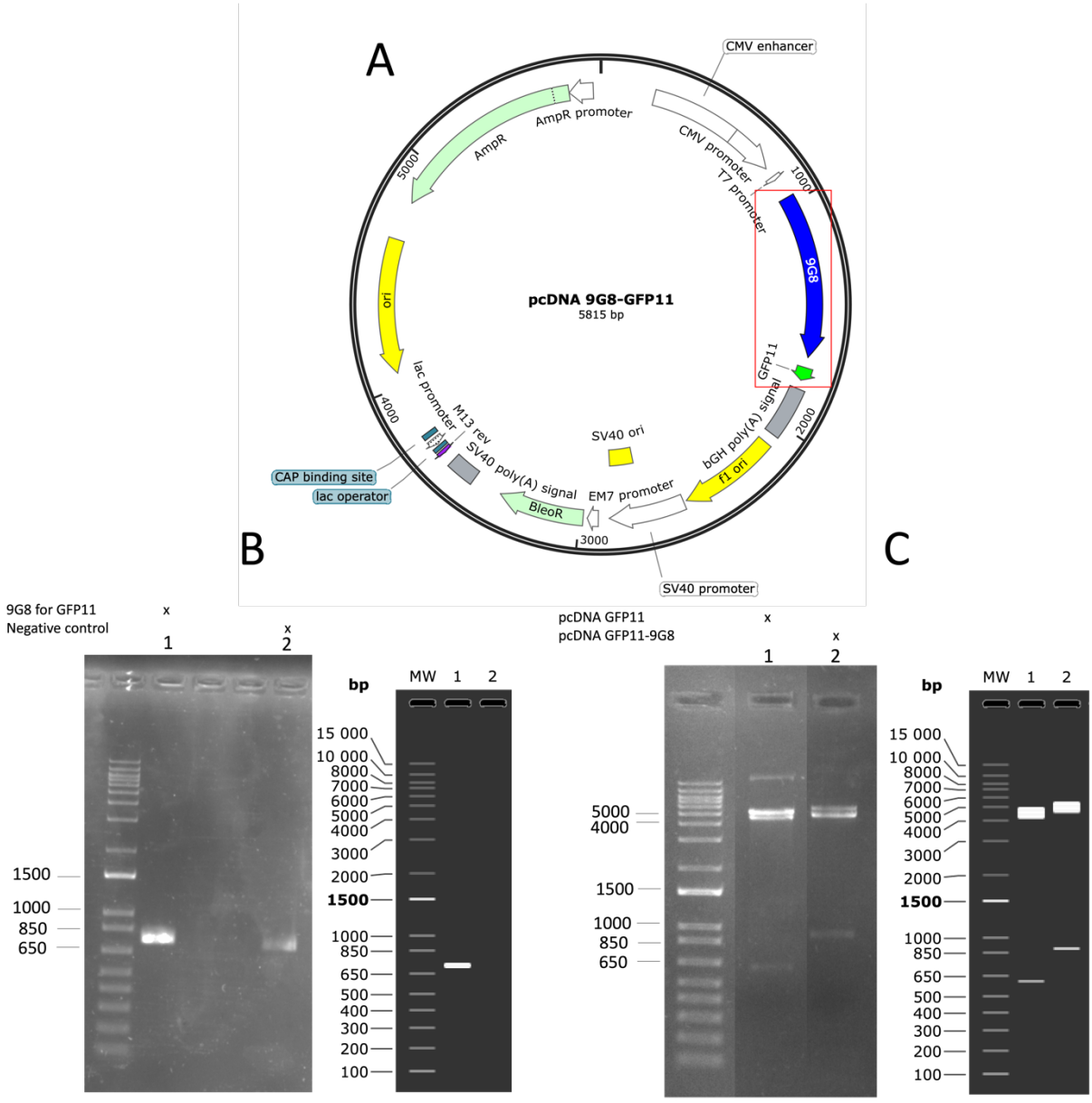


Figure 17 - Construction of the pcDNA 9G8-GFP11 vector. A - pcDNA 9G8-GFP11 vector built from pcDNA GFP11 vector after insertion of the 9G8 cDNA (NM_001031684.3). B – PCR amplification of 9G8 for fusion with the GFP11 subunit from the vector pIC111-9G8; 9G8 for GFP11 size is 758 bp; 1 – 9G8 for fusion with GFP11 amplification; 2 – Negative Control. C - Enzyme restriction digestion to confirm the successful insertion of 9G8 in the pcDNA GFP11 vector before sequencing; 1 – pcDNA GFP11 vector cut with ApaI and NdeI enzymes (4498 bp + 603 bp); 2 - pcDNA 9G8-GFP11 vector (5815 bp) cut with ApaI and HindIII enzymes (4925 bp + 890 bp). Vectors and agarose gel with *in silico* confirmation of the expected bands after enzyme restriction digestion were obtained with SnapGene® software (from Insightful Science; available at snapgene.com).

pcDNA GFP10-Tubulin vector

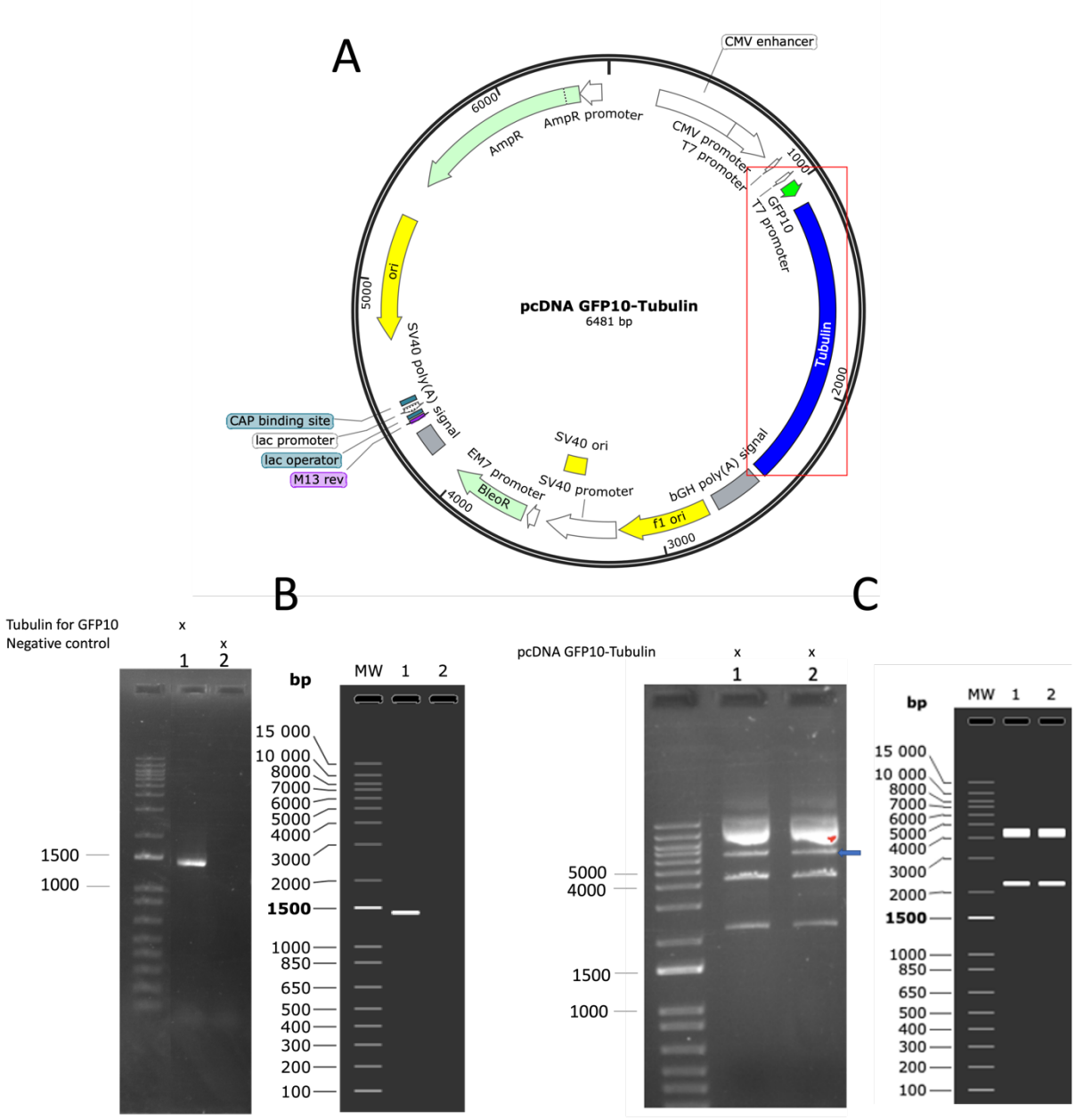


Figure 18 – Construction of the pcDNA GFP10-tubulin vector. A - pcDNA GFP10-tubulin vector built from pcDNA GFP10 vector after insertion of the tubulin cDNA (NM_006082.3). B – PCR amplification of tubulin for fusion with the GFP10 subunit from cDNA of human SH-SY5Y cells; Tubulin for GFP10 size is 1400 bp; 1 and 2 – Tubulin for fusion with GFP10 amplification; 3 – Negative Control. C - Enzyme restriction digestion to confirm the successful insertion of tubulin in the pcDNA GFP10 vector before sequencing; 1 and 2 – pcDNA GFP10-tubulin (6481 bp) cut with SalI enzyme (4259 bp + 2188 bp + 34 bp); Enzyme SalI cuts in 2 different sites, but as digestion was not totally effective, it cut some DNA only in one site, and the expected band is at 6481 bp (marked with a blue arrow). Vectors and agarose gel with *in silico* confirmation of the expected bands after enzyme restriction digestion were obtained with SnapGene® software (from Insightful Science; available at snapgene.com).

pcDNA Tubulin-GFP11 vector

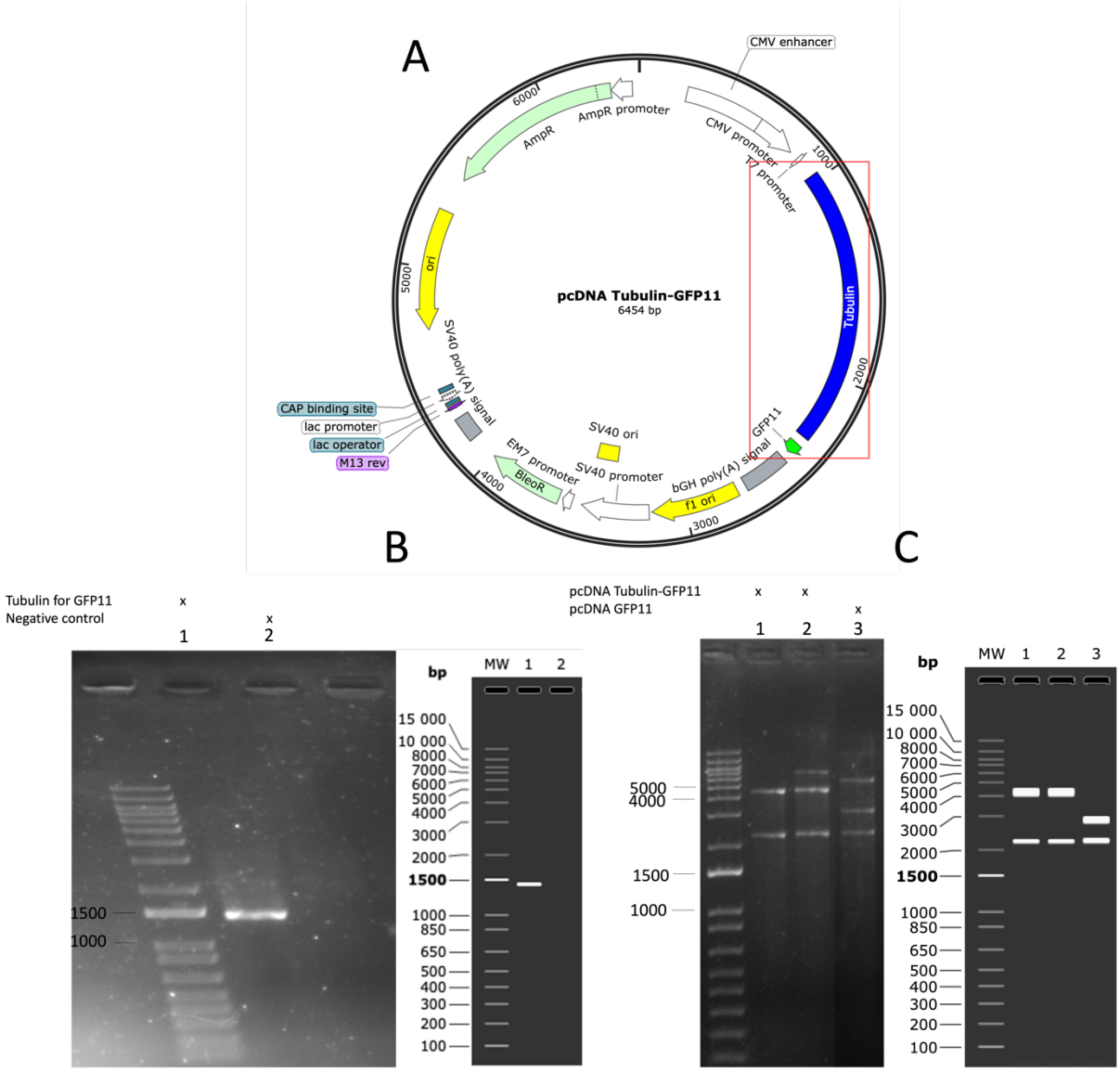


Figure 19 – Construction of the pcDNA tubulin-GFP11 vector. A - pcDNA tubulin-GFP11 vector built from pcDNA GFP11 vector after insertion of the tubulin cDNA (NM_006082.3). B – PCR amplification of tubulin for fusion with the GFP11 subunit from cDNA of human SH-SY5Y cells; Tubulin for GFP11 size is 1401 bp; 1 – Tubulin for fusion with GFP11 amplification; 2 – Negative Control. C - Enzyme restriction digestion to confirm the successful insertion of tubulin in the pcDNA GFP11 vector before sequencing; 1 and 2 - pcDNA tubulin-GFP11 (6454 bp) cut with SalI enzyme (4232 bp + 2188 bp + 34 bp); 3 – pcDNA GFP11 vector cut with SalI enzyme (2879 bp + 2188 bp + 34 bp). Vectors and agarose gel with *in silico* confirmation of the expected bands after enzyme restriction digestion were obtained with SnapGene® software (from Insightful Science; available at snapgene.com).

pcDNA GFP10-hHR23A vector

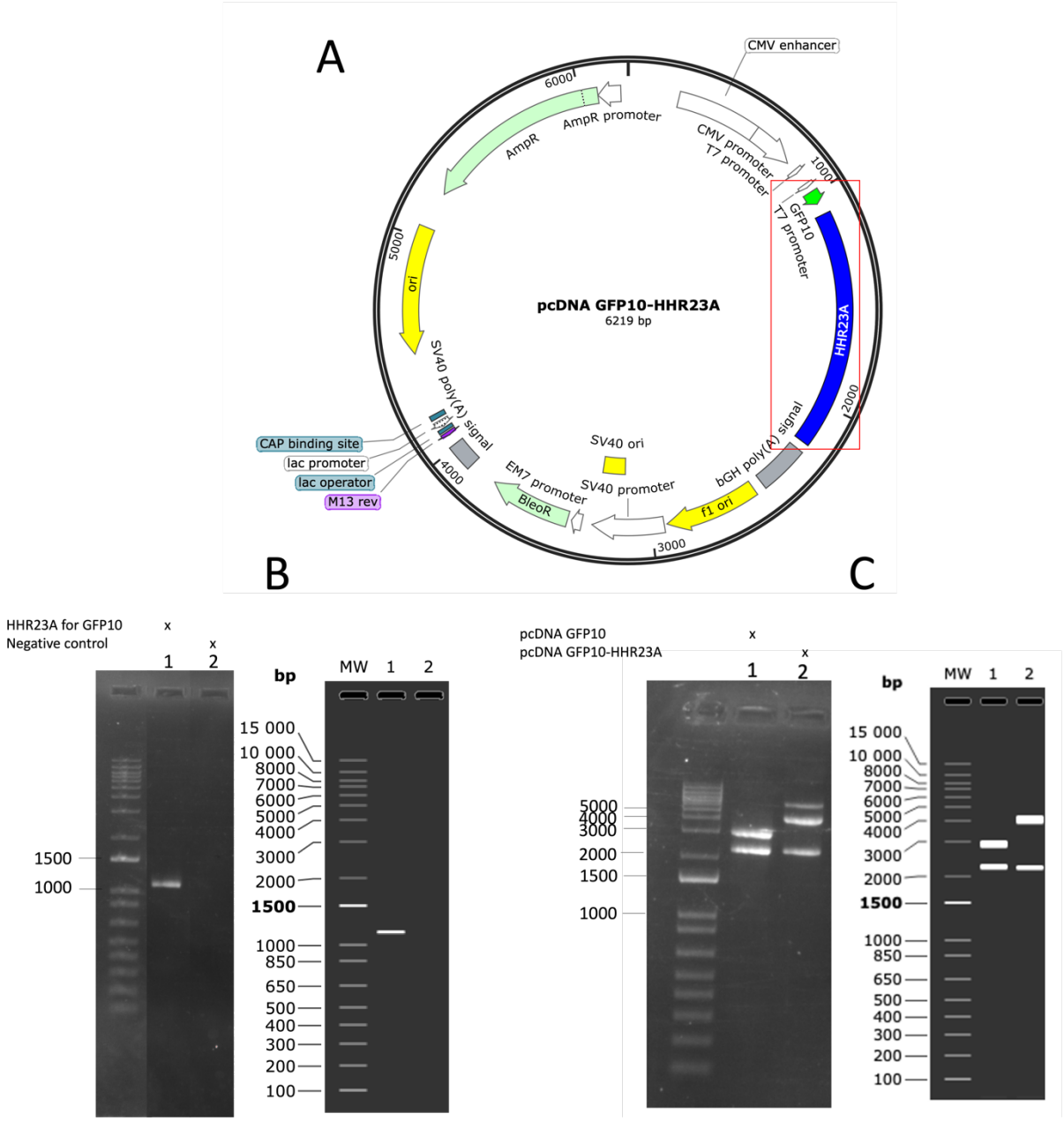
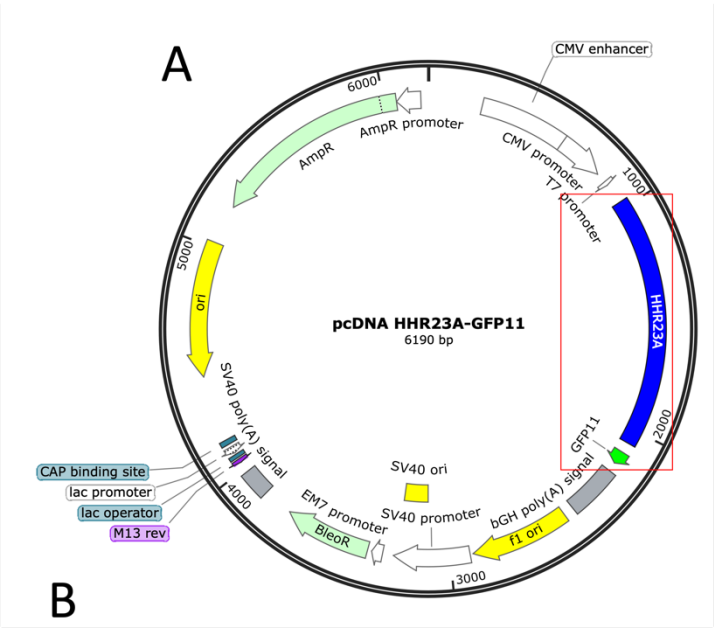


Figure 20 – Construction of the pcDNA GFP10-hHR23A vector. A – pcDNA GFP10-hHR23A vector built from pcDNA GFP10 vector after insertion of the HR23A cDNA (NM_005053.4). B – PCR amplification of hHR23A for fusion with the GFP10 subunit from cDNA of human SH-SY5Y cells; hHR23A for GFP10 size is 1136 bp; 1 and 2 – hHR23A for fusion with GFP10 amplification; 3 – Negative Control. C - Enzyme restriction digestion to confirm the successful insertion of hHR23A in pcDNA GFP10 vector before sequencing; 1 – pcDNA GFP10 vector cut with SalI enzyme (3041 bp + 2188 bp + 34 bp); 2 – pcDNA GFP10-hHR23A vector (6219 bp) cut with SalI enzyme (3997 bp + 2188 bp). Vectors and agarose gel with *in silico* confirmation of the expected bands after enzyme restriction digestion were obtained with SnapGene® software (from Insightful Science; available at snapgene.com).

pcDNA hHR23A-GFP11 vector



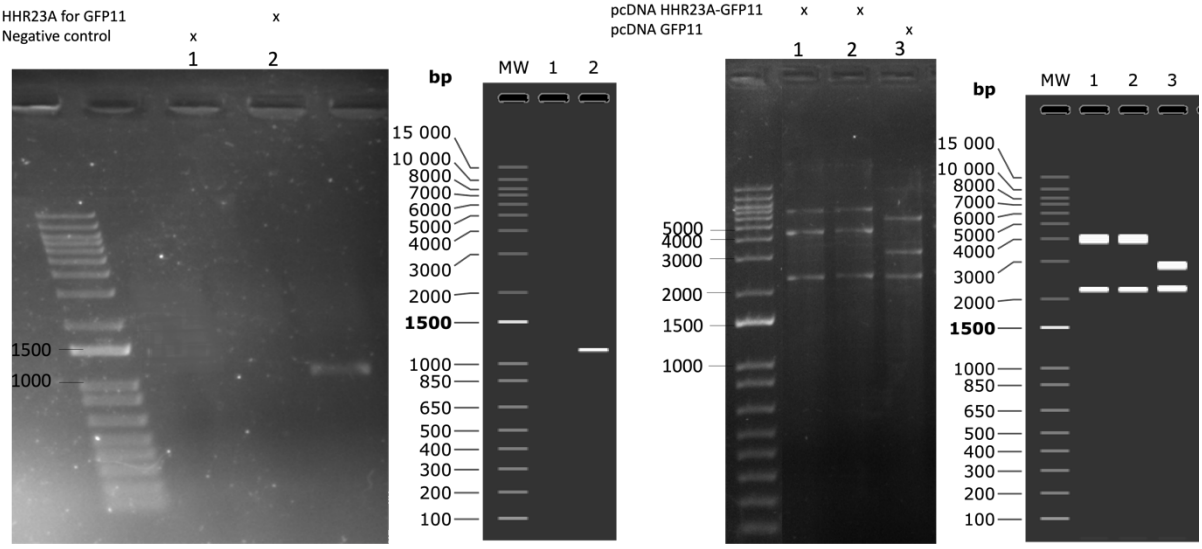


Figure 21 – Construction of the pcDNA hHR23A-GFP11 vector. A – pcDNA hHR23A-GFP11 vector built from pcDNA GFP11 vector after insertion of the hHR23A cDNA (NM_005053.4). B – PCR amplification of hHR23A for fusion with the GFP11 subunit from cDNA of human SH-SY5Y cells; hHR23A for GFP11 size is 1133 bp; 1 – Negative Control; 2 – hHR23A for fusion with GFP11 amplification. C - Enzyme restriction digestion to confirm the successful insertion of hHR23A in pcDNA GFP11 vector before sequencing; 1 and 2 - pcDNA hHR23A-GFP11 (6190 bp) cut with SalI enzyme (3968 bp + 2188 bp + 34 bp); 3 – pcDNA GFP11 vector cut with SalI enzyme (2879 bp + 2188 bp + 34 bp). Vectors and agarose gel with *in silico* confirmation of the expected bands after enzyme restriction digestion were obtained with SnapGene® software (from Insightful Science; available at snapgene.com).

pcDNA GFP10-Mutant ATXN3 vector

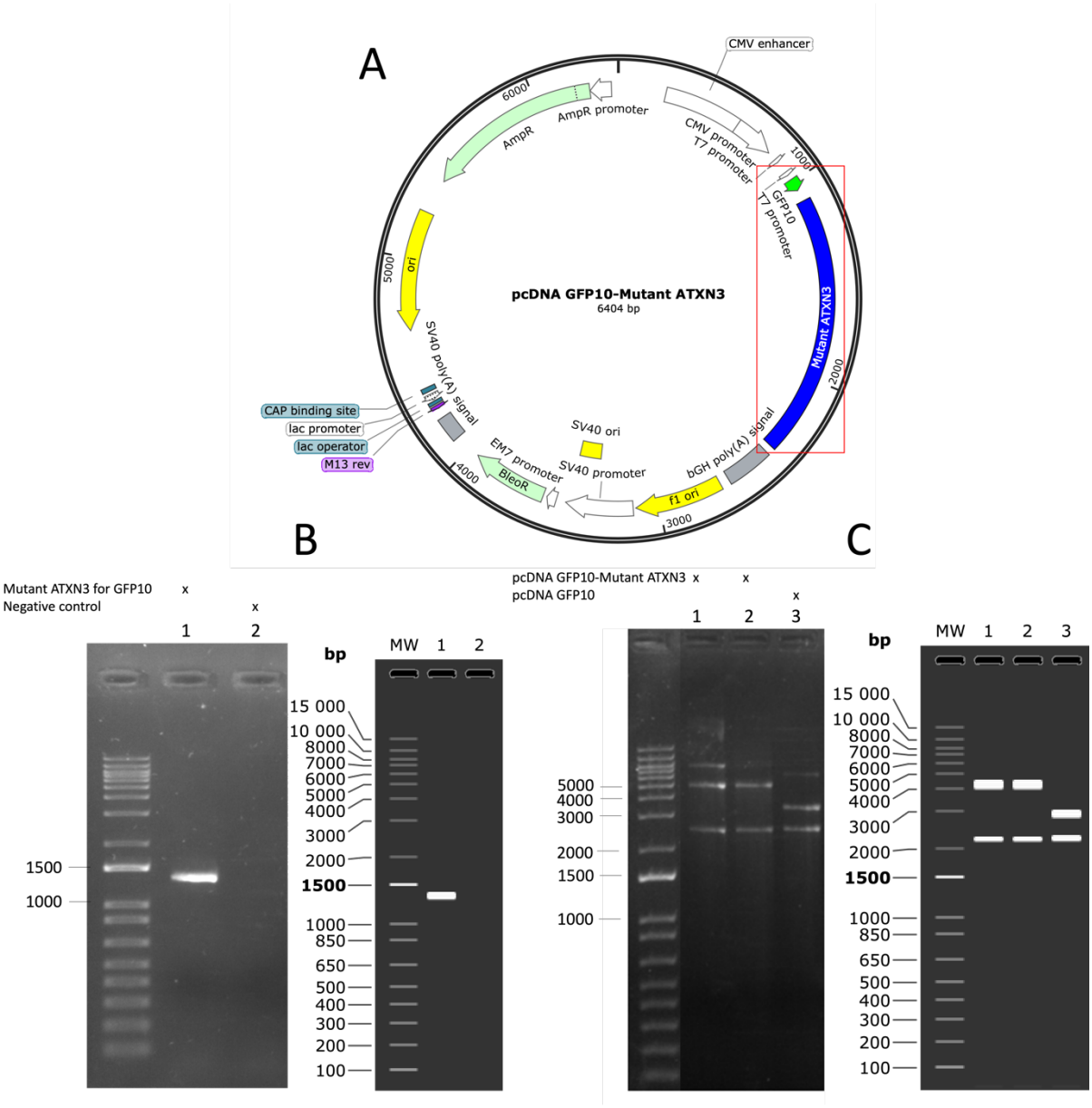


Figure 22 – Construction of the pcDNA GFP10-Mutant ATXN3 vector. A – pcDNA GFP10-Mutant ATXN3 vector built from pcDNA GFP10 vector after insertion of the mutant ATXN3 cDNA (NM_004993.6, expanded form with 78Q). B – PCR amplification of mutant ATXN3 for fusion with the GFP10 subunit from the vector pBRIT TAP-ATXN3 78Q; Mutant ATXN3 size for GFP10 is 1323 bp; 1 – Mutant ATXN3 for fusion with GFP10 amplification; 2 – Negative Control. C - Enzyme restriction digestion to confirm the successful insertion of mutant ATXN3 in the pcDNA GFP10 vector before sequencing; 1 and 2 – pcDNA GFP10-Mutant ATXN3 vector (6404 bp) cut with SalI enzyme (4182 bp + 2188 bp + 34 bp); 3 - pcDNA GF10 vector cut with SalI enzyme (3041 bp + 2188 bp + 34 bp). Vectors and agarose gel with *in silico* confirmation of the expected bands after enzyme restriction digestion were obtained with SnapGene® software (from Insightful Science; available at snapgene.com).

pcDNA Mutant ATXN3-GFP11 vector

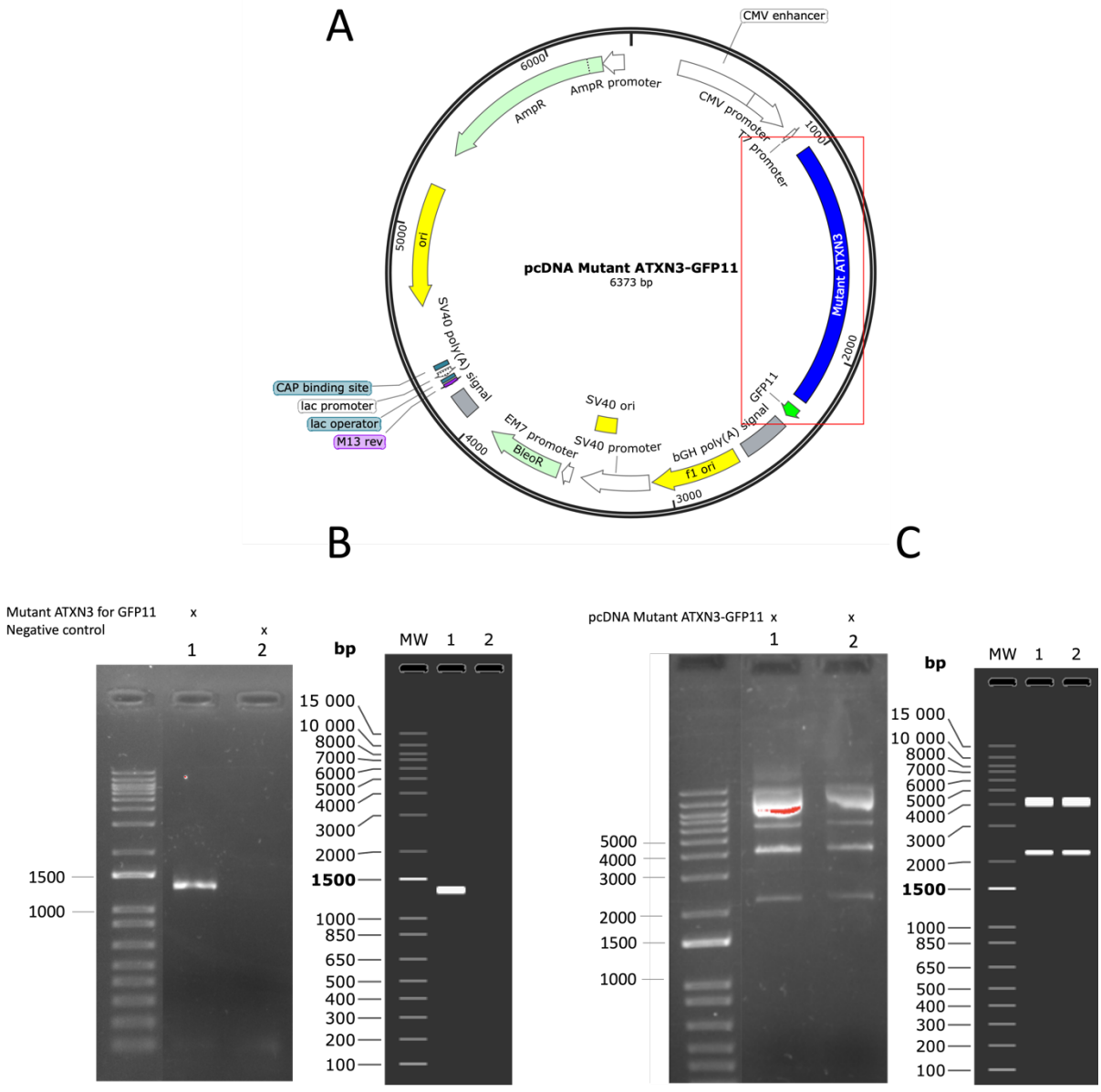


Figure 23 – Construction of the pcDNA Mutant ATXN3-GFP11 vector. A – pcDNA Mutant ATXN3-GFP11 vector built from pcDNA GFP11 vector after insertion of the mutant ATXN3 cDNA (NM_004993.6, expanded form with 78Q). B – PCR amplification of mutant ATXN3 for fusion with the GFP11 subunit from the vector pBRIT TAP-ATXN3 78Q; Mutant ATXN3 size for GFP11 is 1328 bp; 1 – Mutant ATXN3 for fusion with GFP11 amplification; 2 – Negative Control. C - Enzyme restriction digestion to confirm the successful insertion of mutant ATXN3 in the pcDNA GFP11 vector before sequencing; 1 and 2 – pcDNA mutant ATXN3-GFP11 vector (6373 bp) cut with SalI enzyme (4151 bp + 2188 bp + 34 bp). Vectors and agarose gel with *in silico* confirmation of the expected bands after enzyme restriction digestion were obtained with SnapGene® software (from Insightful Science; available at snapgene.com).

TriSFP system was validated by flow cytometry, immunofluorescence, and fluorescence microscopy

The constructs for the triSFP system expressing GFP10 and GFP11 subunits were further transfected in previously obtained MRC5-SV cells expressing the GFP1-9 subunit and the VHH enhancer, to test the system *in vivo*. 48 hours after transfection, flow cytometry allowed us to perform an initial detection and quantification of GFP fluorescence in transfections of ATXN3 (WT and mutant form) with the positive controls hHR23A, tubulin, and ATXN3 itself, whose interactions are reported in literature [56], [102], [103] (**Figure 24 and 26**). Regarding negative controls, as seen in **Figure 26**, no fluorescence was detected with ATXN3 fused to one of GFP subunits transfected with the other subunit alone (GFP10-ATXN3 + GFP11; ATXN3-GFP11 + GFP10), neither with isolated GFP subunits expressed together (GFP10 + GFP11), or even just ATXN3 fused to one subunit (GFP10-ATXN3; ATXN3-GFP11). Other positive controls (GFP10-zipper-GFP11 construct for control of GFP assembly, and pIC111-ATXN3 for control of transfection) showed expression of the respective genes and significant amounts of fluorescence were detected (**Figure 26**). Interestingly, some fluorescence was observed with cells expressing ATXN3 and the GCN4 leucine zipper, a transcriptional activator previously studied by the authors that developed the triSFP assay [82], but for which we were not expecting an interaction with ATXN3. Importantly, significant amounts of fluorescence were detected with this system with transfection of ATXN3 and the candidate interactor splicing factor 9G8 (**Figure 24 and 26**). The percentage of cells displaying fluorescence is detailed in **Table 3**, Supplementary Information.

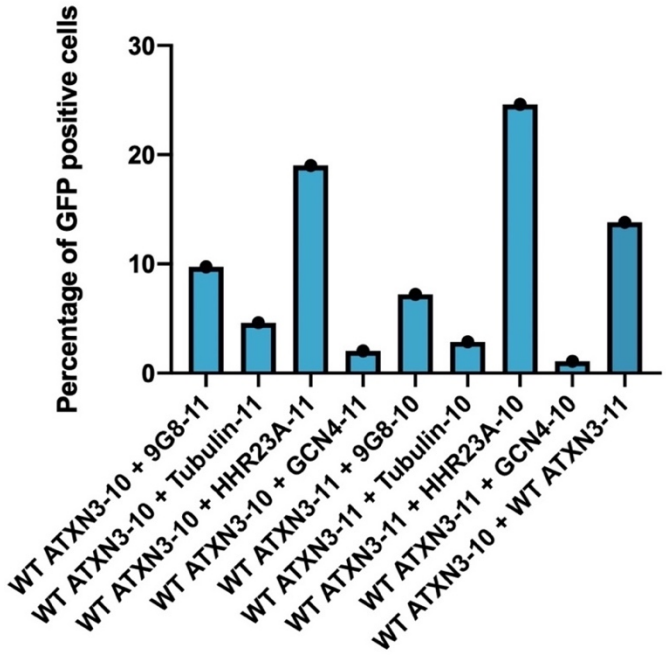


Figure 24 – Flow cytometry analysis of percentage of GFP positive cells in interactions between WT ATXN3 attached to one subunit of GFP, and 9G8, tubulin, hHR23A, GCN4 and ATXN3 itself attached to the other subunit of GFP. 10 is short for GFP10 subunit, and 11 is short for GFP11 subunit. Column graphic was obtained with GraphPad Prism version 9.0.0 for macOS, GraphPad Software, San Diego, California USA, www.graphpad.com.

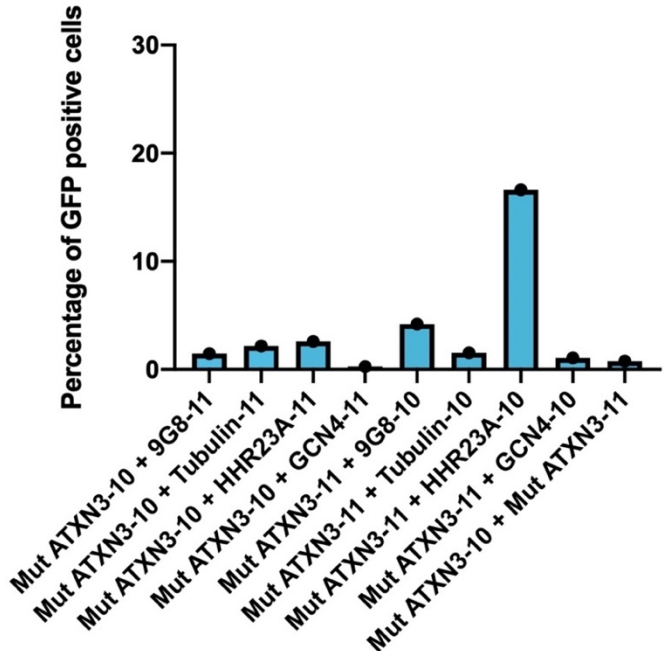


Figure 25 – Flow cytometry analysis of percentage of GFP positive cells in interactions between mutant ATXN3 attached to one subunit of GFP, and 9G8, tubulin, hHR23A, GCN4 and ATXN3 itself attached to the other subunit of GFP. 10 is short for GFP10 subunit, and 11 is short for GFP11 subunit. Column graphic was obtained with GraphPad Prism version 9.0.0 for macOS, GraphPad Software, San Diego, California USA, www.graphpad.com.

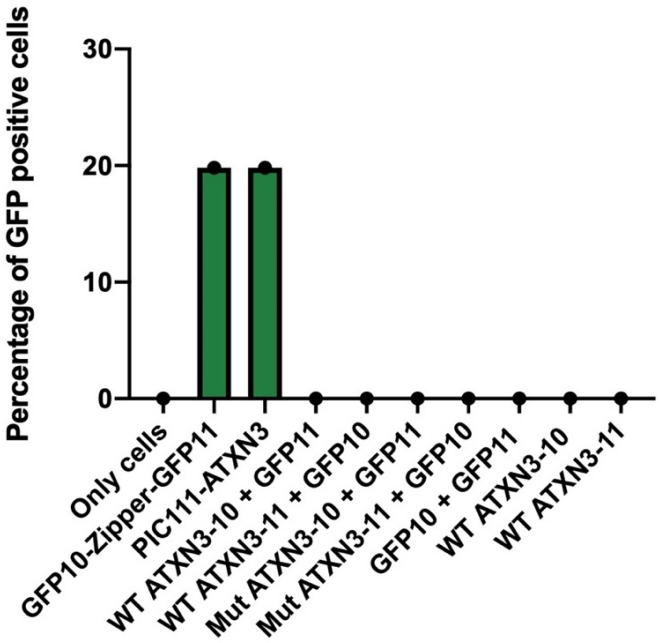


Figure 26 – Flow cytometry analysis of percentage of GFP positive cells in the following controls: only cells (negative control), GFP10-Zipper-GFP11 and PIC111-ATXN3 (positive controls), and ATXN3 (WT and mutant form) attached with one of GFP subunit interacting with the other subunit alone, without any protein, GFP subunits without any protein attached, and WT ATXN3 attached to one of GFP subunits alone (negative controls). 10 is short for GFP10 subunit, and 11 is short for GFP11 subunit. Column graphic was obtained with GraphPad Prism version 9.0.0 for macOS, GraphPad Software, San Diego, California USA, www.graphpad.com.

After flow cytometry analysis, we wanted to validate that the observed fluorescence was indeed due to proteins encountering and GFP tags assembling, and that the absence of it was not due to transfection or expression problems. For that, immunofluorescence against GFP10 was performed to cells expressing WT ATXN3-GFP11 and GFP10-9G8, and GFP10-WT ATXN3 and GFP11 only, as well as WT ATXN3-GFP11 and GFP10 only. Imaging with TRITC filter confirms expression of GFP10 in all conditions, but FITC filter is only visible when both GFP tags are fused to interactors, validating the triSFP system with emission of fluorescence only upon protein interaction **(Figure 27).**

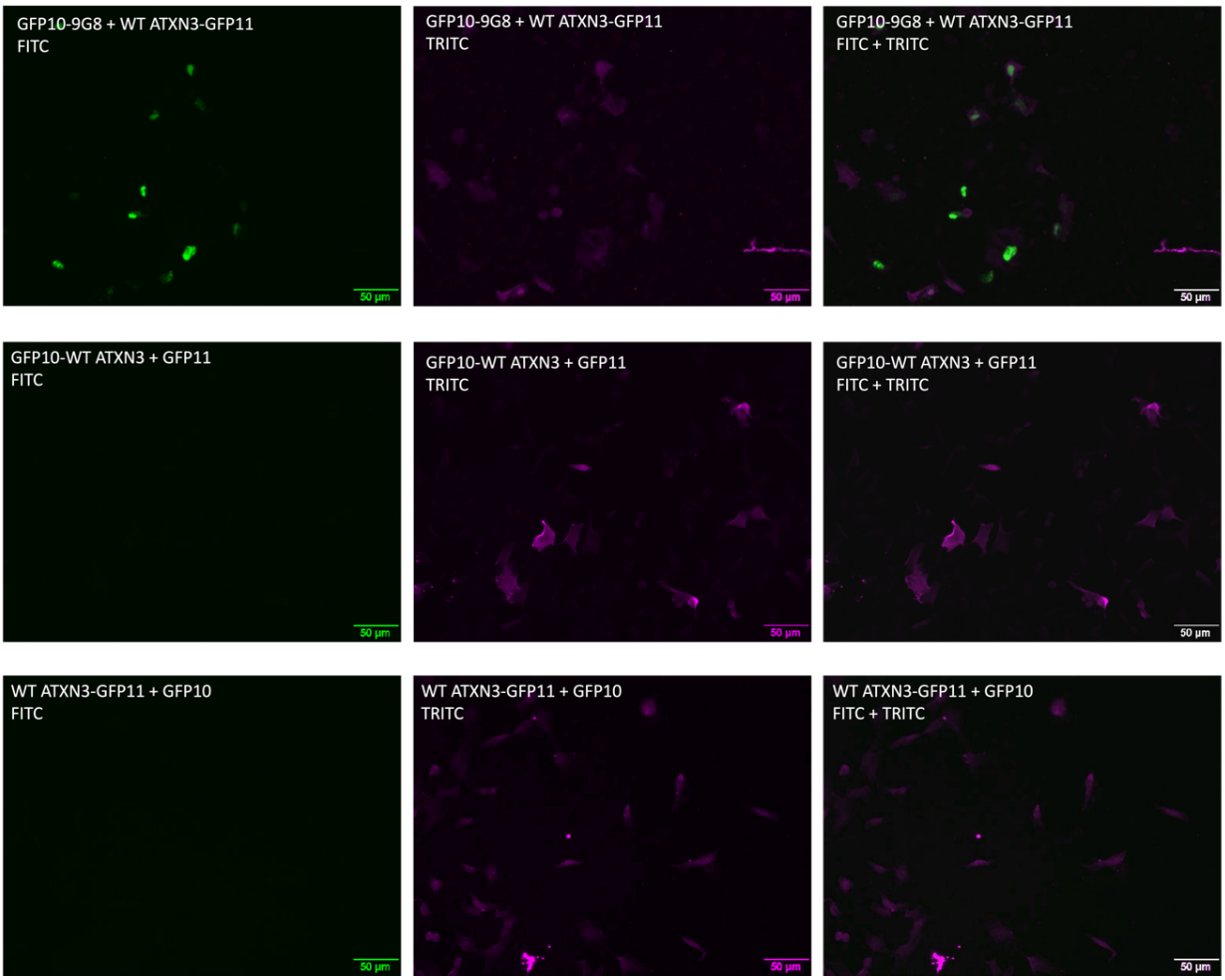


Figure 27 - Immunofluorescence against GFP10 for interaction between WT ATXN3-GFP11 and GFP10-9G8, between GFP10-WT ATXN3 and GFP11 only, as well as WT ATXN3-GFP11 and GFP10 only. Images were acquired using the Olympus IX81 Inverted Fluorescence Microscope. Scale bars: 50 um.

To further characterize and validate the detected fluorescence, transfections were analysed by fluorescence microscopy. In accordance with what is described in literature [56], [102], [103], fluorescence of transfection of ATXN3 with its positive controls was observed in the following cell compartments: throughout the whole cells for interaction of ATXN3 with hHR23A, which was expected as this protein is known to be involved in both DNA repair at the nucleus and in the endoplasmic reticulum associated protein degradation (ERAD) pathway in the cytoplasm; exclusively in the cytoplasm for the ATXN3-tubulin interaction, to be expected as this protein is the major constituent of microtubules; and throughout the whole cell for interaction of ATXN3 with itself, as ATXN3 can be found in cytoplasm and nucleus, and it interacts with itself, regulating its own levels, ubiquitination pattern, and subcellular localization [103]. The cytosolic interaction between ATXN3 and tubulin, and both nuclear and cytosolic interaction between ATXN3 and hHR23A, were confirmed by DAPI nuclear staining. Once again, no

fluorescence was detected with the negative control of ATXN3 transfected with the other GFP tag without an interactor attached. Altogether, this data validates the triSFP system.

The interaction of ATXN3 with 9G8 was again detected, and importantly, the two proteins were seen to interact in the nucleus of the cells, as confirmed by DAPI nuclear staining, which reinforces the interest of the study of this interaction. The system also showed mainly a nuclear interaction between ATXN3 and the transcriptional activator GCN4 leucine zipper (**Figure 28 to 34**).

GFP10-WT ATXN3 + Interactor-GFP11

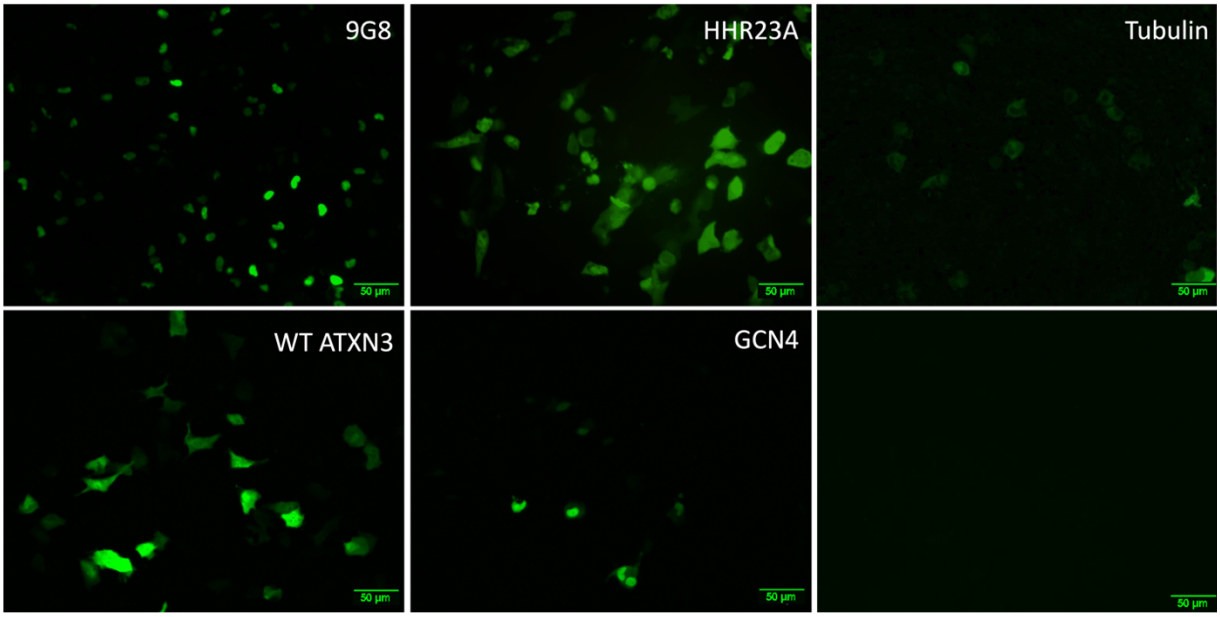


Figure 28 - Wild-type ATXN3 fused with GFP10, interacting with 9G8, hHR23A, tubulin, ATXN3 itself and GCN4 fused with GFP11. The last image (below right) shows transfection of WT ATXN3 fused to GFP10 with GFP11 only, without an interactor, which constitutes a negative control. Images were acquired using the Olympus IX81 Inverted Fluorescence Microscope, 24 hours after transfection. Scale bars: 50 um.

WT ATXN3-GFP11 + GFP10-Interactor

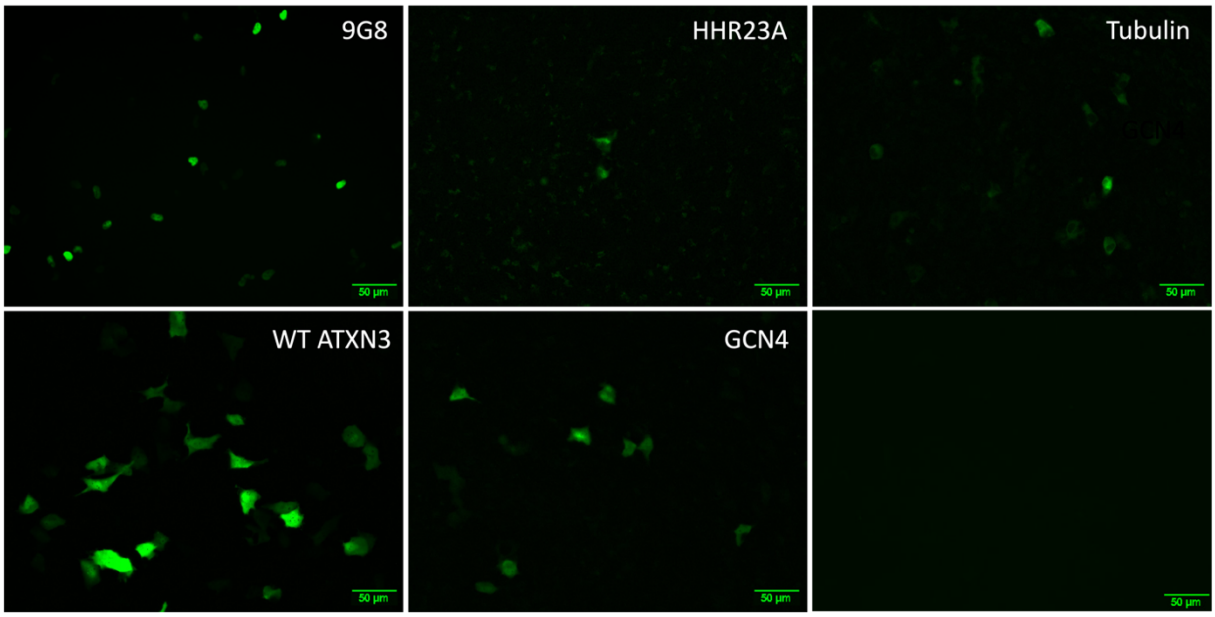


Figure 29 - Wild-type ATXN3 fused with GFP11, interacting with 9G8, hHR23A, tubulin, ATXN3 itself and GCN4 fused with GFP10. The last image (below right) shows transfection of WT ATXN3 fused to GFP11 with GFP10 only, without an interactor, which constitutes a negative control. Images were acquired using the Olympus IX81 Inverted Fluorescence Microscope, 24 hours after transfection. Scale bars: 50 um.

GFP10-Mutant ATXN3 + Interactor-GFP11

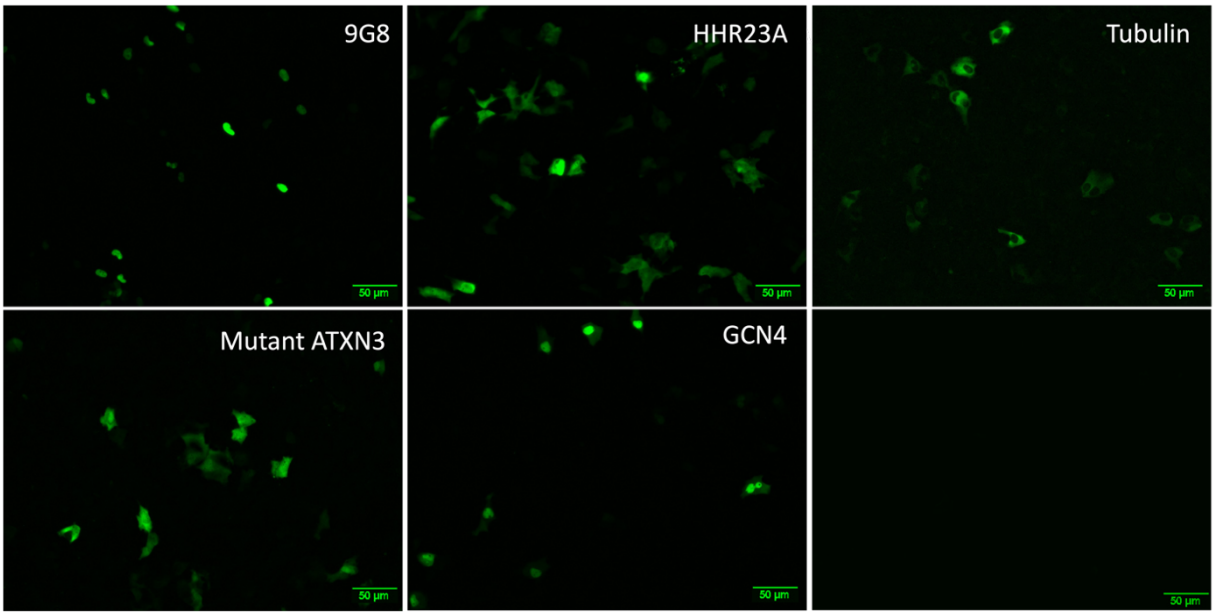


Figure 30 - Mutant ATXN3 fused with GFP10, interacting with 9G8, hHR23A, tubulin, ATXN3 itself and GCN4 fused with GFP11. The last image (below right) shows transfection of mutant ATXN3 fused to GFP10 with GFP11 only, without an interactor, which constitutes a negative control. Images were acquired using the Olympus IX81 Inverted Fluorescence Microscope, 24 hours after transfection. Scale bars: 50 um.

Mutant ATXN3-GFP11 + GFP10-Interactor

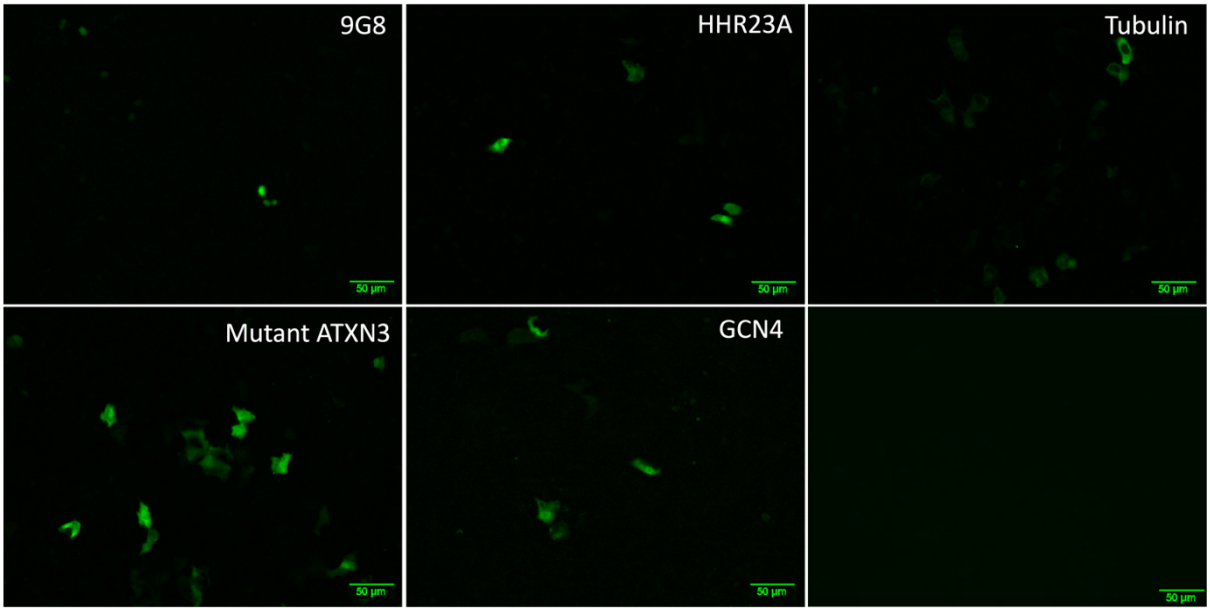


Figure 31 - Mutant ATXN3 fused with GFP11, interacting with 9G8, hHR23A, tubulin, ATXN3 itself and GCN4 fused with GFP10. The last image (below right) shows transfection of mutant ATXN3 fused to GFP11 with GFP10 only, without an interactor, which constitutes a negative control. Images were acquired using the Olympus IX81 Inverted Fluorescence Microscope, 24 hours after transfection. Scale bars: 50 um.

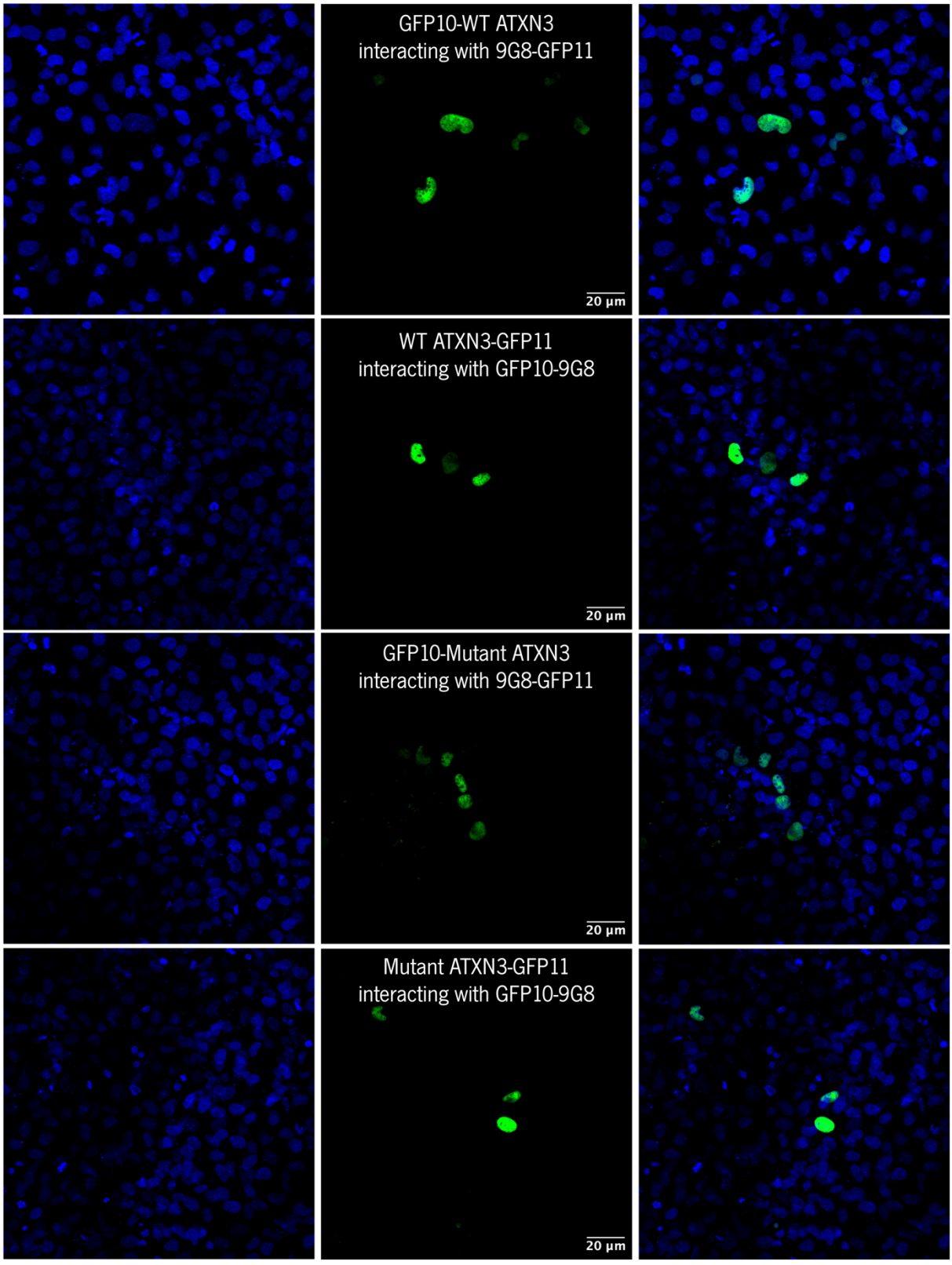


Figure 32 – DAPI nuclear staining of interactions between ATXN3 and 9G8. GFP10-WT ATXN3 interacting with 9G8-GFP11, WT ATXN3-GFP11 interacting with GFP10-9G8, GFP10-Mutant ATXN3 interacting with 9G8-GFP11, and Mutant ATXN3-GFP11 interacting with GFP10-9G8. Images were acquired using the Olympus LPS Confocal FV3000 microscope. Scale bars: 20 um.

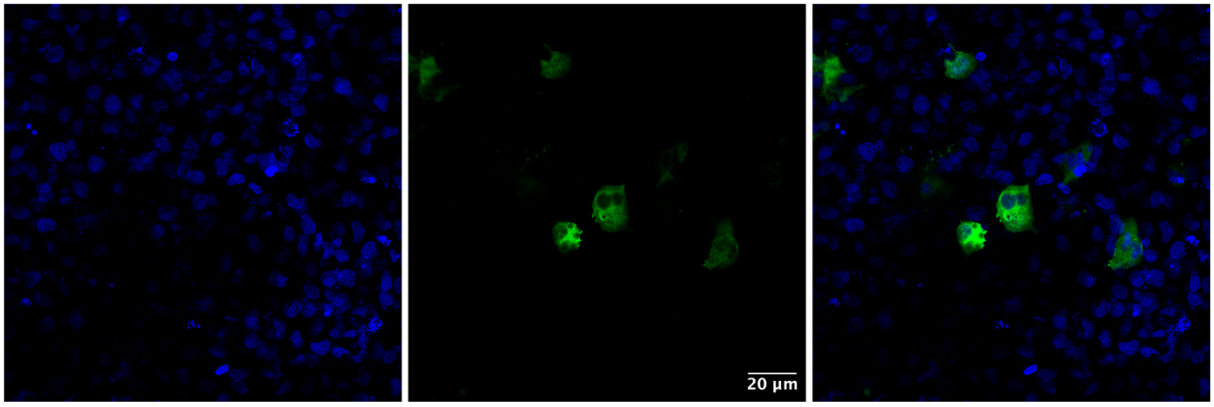


Figure 33 – DAPI nuclear staining of interaction between tubulin and ATXN3. Image was acquired using the Olympus LPS Confocal FV3000 microscope. Scale bar: 20 um.

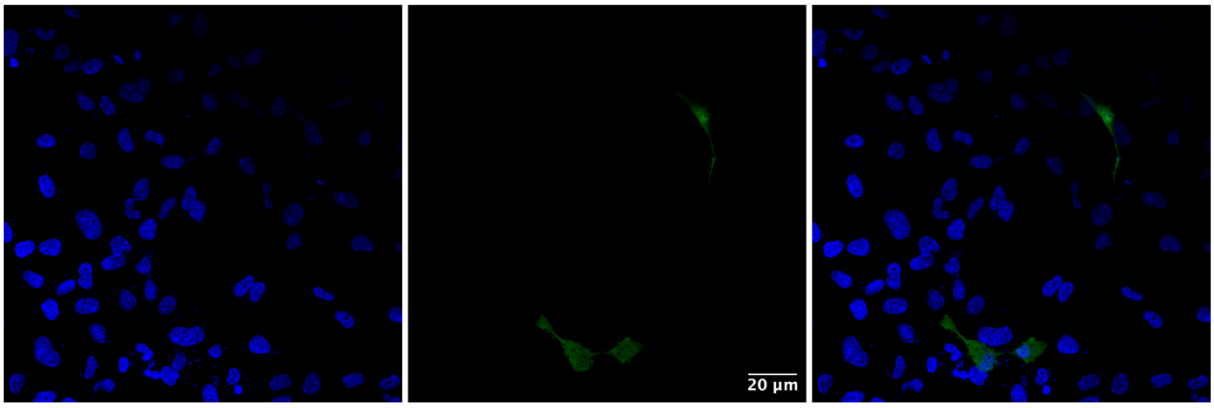


Figure 34 – DAPI nuclear staining of interaction between hHR23A and ATXN3. Image was acquired using the Olympus LPS Confocal FV3000 microscope. Scale bar: 20 um.

Mutation at Nuclear Localization Signal did not avoid interaction of ATXN3 with 9G8

Our previous results validate the triSFP system, which in turn allowed us to detect a nuclear interaction between ATXN3 and 9G8. Considering that, we decided to perform a site directed mutagenesis to the nuclear localization signal of ATXN3, both WT and mutant form, to evaluate the impact of this mutation in that interaction. Sequencing analysis confirmed the presence of the NLS mutation at pcDNA WT ATXN3 - GFP11 (WT ATXN3 R282T NLS-GFP11), pcDNA GFP10 - Mutant ATXN3 (GFP10 - Mutant ATXN3 R282T NLS) and pcDNA Mutant ATXN3 - GFP11 (Mutant ATXN3 R282T NLS-GFP11) vectors, but was not successfully inserted at the pcDNA GFP10 - WT ATXN3 (GFP10 - WT ATXN3 R282T NLS) vector (**Figures 62 to 65**, Supplementary Information). Then, we analyzed by fluorescence microscopy if the mutation would interrupt the nuclear interaction between 9G8 and ATXN3, as previous studies demonstrated that mutating this site avoids ATXN3 translocation to the nucleus [28]. Contrarily to what we hypothesized, the mutation did not disrupt ATXN3 ability to enter the nucleus, as fluorescence was

still observed in that compartment when NLS mutated ATXN3 was transfected with 9G8, as depicted in the following images (**Figure 35).**

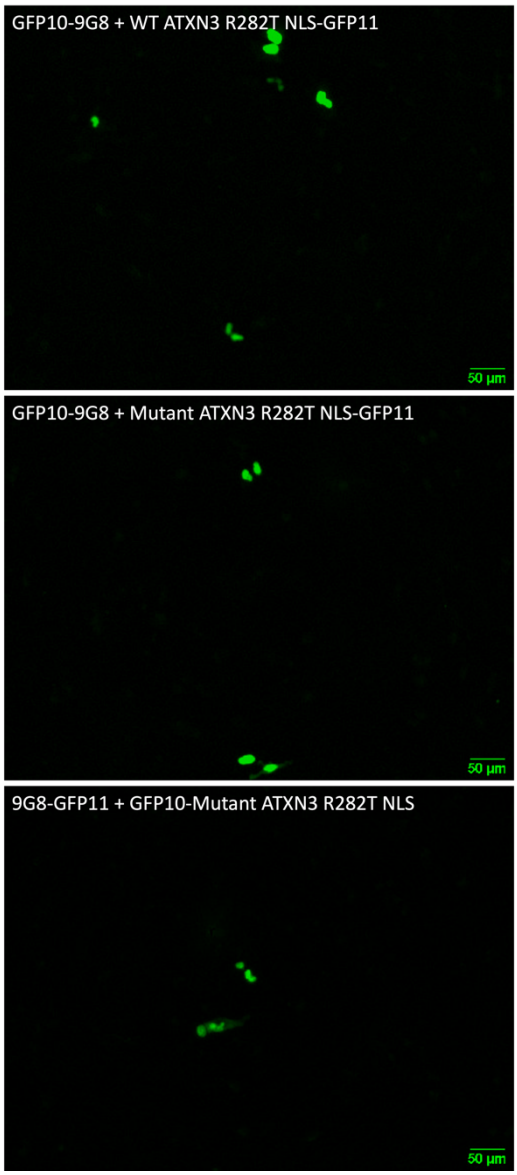


Figure 35 – WT ATXN3 R282T NLS-GFP11 and mutant ATXN3 R282T NLS-GFP11 interacting with GFP10-9G8, and GFP10-mutant ATXN3 R282T NLS interacting with 9G8-GFP11. Images were acquired using the Olympus IX81 Inverted Fluorescence Microscope, 24 hours after transfection. Scale bars: 50 **u**m.

DISCUSSION

Faced with the need of a robust cell-based high-throughput system to detect disease relevant ATXN3 PPIs *in vivo* and further modulate them for MJD therapy, the tripartite-split-GFP system was successfully constructed with adequate positive and negative controls, and with a relevant ATXN3 interaction for the study of MJD pathogenesis, namely with the splicing factor 9G8.

Flow cytometry allowed us to test the development of the triSFP system and perform an initial quantification of GFP fluorescence in cells with different transfection conditions. Results strongly validated positive controls, namely ATXN3 interacting with hHR23A, tubulin, and ATXN3 itself, and GFP10-Zipper-GFP11 and pIC111-ATXN3. Negative controls were also validated, namely ATXN3 fused to one of GFP subunits transfected with the other subunit alone, isolated GFP subunits expressed together, and just ATXN3 fused to one subunit, as fluorescence was not detected (**Figure 24 to 26**). This confirms the validity of the system and provides further support to the ATXN3-9G8 interaction, confirmed by the same technique. Additionally, our results are in agreement with those obtained by Cabantous *et al.* at [91] and Koraïchi *et al.* at [82], that developed the same system and obtained similar fluorescence levels with their constructs. Comparisons of levels of fluorescence between different conditions obtained by flow cytometry can give as a hint regarding protein interactions affinity, but it is important to keep in mind that transfection efficiency severely affects the levels of GFP fluorescence. As so, comparisons of interactions between WT and mutant ATXN3 with the different partners, with special interest in 9G8, as well as of interactions through different termini, do not provide reliable information in the current system and strong conclusions cannot be taken. More flow cytometry data could be collected from different transfections with the same constructs, and, most importantly, from stably transfected cell lines, to perform that analysis more robustly.

Immunofluorescence against GFP10 results validated transfection and expression of this GFP tag, reinforcing the successful construction of the triSFP system (**Figure 27**). The fact that imaging with TRITC filter confirms expression of GFP10 in all conditions, but the FITC filter is only visible when both GFP tags are fused to interactors, confirms that, on the one hand, protein interaction is necessary to see fluorescence, and on the other hand, that the absence of fluorescence when one of the GFP tags is expressed alone, is not due to transfection problems. Ideally, we would also perform immunofluorescence for the GFP11 and GFP1-9 subunits, but good quality antibodies for those tags were not available. Alternatively, and complementing these results, RT-PCR (reverse transcription-polymerase chain reaction) can be performed to confirm the expression of all GFP tags at the mRNA level, after RNA extraction.

Additionally, immunofluorescence or RT-PCR for ATXN3, 9G8 and the other interactors can be performed to confirm their expression.

Fluorescence microscopy reinforced the robustness of the system by providing evidence of the subcellular localization of the positive control interactors (hHR23A, tubulin and ATXN3 itself), which was in accordance with what is described in literature [56], [102], [103]. Verification of the interaction of ATXN3 with 9G8 by flow cytometry, and their nuclear encounter confirmed by fluorescence microscopy and DAPI nuclear staining, is essential to consider 9G8 as a potential candidate for drug screening, as the nuclear presence of ATXN3 seems to contribute for protein aggregation and neurodegeneration in MJD. This validation is in line with published [45], and unpublished observations from our team, that confirm the interaction of 9G8 with ATXN3 in cells and brain tissues of a MJD mouse model [104], and even imply some impact of this interaction on ATXN3 aggregation.

Interestingly, this system also allowed the detection of a non-described and unexpected interaction between ATXN3 and the GCN4 leucine zipper, a transcriptional activator of amino acid biosynthesis in yeast. The system showed mainly a nuclear interaction with ATXN3, which is in line with the role of GCN4 in transcription [100], and the presence of ATXN3 in that compartment, where it exerts different functions, among which transcription regulation. As previously described, ATXN3 appears to have a role as a transcription regulator by inhibiting CREB-mediated transcription through interaction with CBP, p300, and PCAF; promotes histone deacetylation by interacting with HDAC3 and NCOR1; and regulates transcription in response to oxidative stress by interacting with the transcription factor FOXO4 [63]–[65]. Further experiments could be performed to validate this interaction, such as a pull-down assay followed by co-immunoprecipitation or SPR. Indeed, the triSFP system can constitute a valuable tool to identify new interactors. Being sensitive to distance between GFP domains, we estimate that the assay only detects close or direct interactions. However, further *in vitro* assays are necessary to validate those interactions, as the fact that an external protein to the complex may be mediating the interaction in cells cannot be excluded.

After having the system validated, site-directed mutagenesis to NLS of ATXN3, which was reported to avoid its import to the nucleus in yeast cells [28], was performed to evaluate its impact in ATXN3 and 9G8 interaction. Nuclear fluorescence was still detected with the mutation, which means that, most likely, the mutation at NLS is not stopping translocation of ATXN3 to the nucleus (**Figure 35**). Interestingly, another mutation at the same motif, namely from Arg-Lys-Arg-Arg to His-Asn-His-His, as well as its deletion, was described to have no effect on subcellular localization of ATXN3 at Human embryonic

kidney (HEK) 293 cells, suggesting that this NLS is not required for the import of ATXN3 to the nucleus [55], [105]. Further analysis can be performed to understand what is happening, such as performing other described mutations at the NLS, and DAPI nuclear staining can further help to validate the correct location of the detected nuclear interactions.

Regarding the interacting termini of proteins, the literature indicates that hHR23A interacts with ATXN3, both normal and mutant form, through its N-terminus, namely the Josephin domain. SPR measurements identify in ATXN3 three separate tubulin interacting regions, one also within the Josephin Domain and the two others within the unstructured region [106]. Furthermore, ATXN3 interacts with itself through its Josephin domain. There is yet no information available regarding interaction of ATXN3 with 9G8 or GCN4. Even though the flow cytometry assay that we performed in transiently transfected cells has the disadvantage of not providing strong conclusions regarding comparisons between interactions through different termini, the detected fluorescence with all interactors attached to both GFP tags (GFP10 blocks N-terminus, and GFP10 blocks C-terminus) suggests that the tags did not interfere with any interaction.

The described work was performed using a pulmonary fibroblast cell line. Even though it is not a neuronal cell line, the fact that the studied interactions were previously validated in neuronal cell lines, opens up the possibility of using this cellular system as a first approach to study these interactions, namely in the screening of compounds to modulate them. However, we aim to further generate similar but stable cell lines in the SH-SY5Y background by lentiviral transduction.

CONCLUSIONS AND FUTURE PERSPECTIVES

This work allowed us to develop and optimize a relevant transient system to detect PPIs and study their subcellular localization *in vivo*. This system avoids protein interference or aggregation, due to the small sizes of GFP10 and GFP11 tagging peptides, and provides specificity of the triSFP complementation, so that false signals are not detected with unspecific binding.

Having the triSFP system fully optimized and characterized for ATXN3 PPIs detection, the next step would be to sub-clone ATXN3 and 9G8 into a lentiviral vector, which would be employed to obtain a similar but constitutive expression system. Ideally, a bicistronic vector could be used to express both partners each attached to both GFP tags. On the other hand, having only ATXN3 subcloned into a constitutive expression system would allow us to transiently add any other protein of interest and analyze its interaction with ATXN3 in this system. After developing and characterizing the constitutive system, it

can be used for the screening with different compounds by high-throughput approaches. By having a stable line constitutively expressing the interactors, we make sure expression levels are the same in all used compounds, and differences in fluorescence levels are not due to differences in initial expression levels. Those compounds can be selected from drug libraries enriched in structures similar no known PPI inhibitors, or from peptide libraries enriched in small molecules that may affect the interaction. If a positive hit is not found on the screening, focus can be redirected to structural information about the hotspots of interaction to design peptides that are able to disrupt ATXN3-9G8 interaction surfaces. A non-toxic concentration should then be defined for every compound, and a kinetic analysis of cells fluorescence levels can be performed in different timepoints. Flow cytometry or flow activated cell sorting (FACS) is essential for the screening process, as it spots slight differences of fluorescence that would be difficult to detect through microscopy. That way, it is possible to create a precise calibration system of fluorescence intensity by analyzing different levels of fluorescence when adding drugs to the system.

Besides its clear importance for studying PPIs, the triSFP system can also constitute an important and promising tool in our laboratory for other studies. A valuable characteristic of this system is that it can be used to visualize aggregates *in vivo* and understand the molecular mechanisms that drive protein oligomerization and accumulation. Indeed, previous studies used a similar system to investigate ALS-linked proteins, α-synuclein, and tau aggregation molecular mechanisms, revealing promising results [3], [107], [108]. With this system, those observations can be compared between WT and mutant form of ATXN3, with particular interest in the study of the polyQ expanded ATXN3 interaction with itself. Other applications previously described include visualization of neuronal synapses in living organisms, morphogen gradients and protein interactions in developmental biology pathways, among others [109]–[111], [3].

REFERENCES

[1] B. Almeida, S. Fernandes, I. A. Abreu, e S. Macedo-Ribeiro, «Trinucleotide repeats: a structural perspective», *Front. Neurol.*, vol. 4, p. 76, 2013, doi: 10.3389/fneur.2013.00076.

[2] C. A. Matos, S. de Macedo-Ribeiro, e A. L. Carvalho, «Polyglutamine diseases: The special case of ataxin-3 and Machado–Joseph disease», *Prog. Neurobiol.*, vol. 95, n.º 1, pp. 26–48, set. 2011, doi: 10.1016/j.pneurobio.2011.06.007.

[3] E. K. Don *et al.*, «In vivo Validation of Bimolecular Fluorescence Complementation (BiFC) to Investigate Aggregate Formation in Amyotrophic Lateral Sclerosis (ALS)», *Mol. Neurobiol.*, vol. 58, n.º 5, pp. 2061–2074, 2021, doi: 10.1007/s12035-020-02238-0.

[4] C. Haass e D. J. Selkoe, «Soluble protein oligomers in neurodegeneration: lessons from the Alzheimer's amyloid β-peptide», *Nat. Rev. Mol. Cell Biol.*, vol. 8, n.º 2, pp. 101–112, fev. 2007, doi: 10.1038/nrm2101.

[5] S. Ye, C.-H. Hsiung, Y. Tang, e X. Zhang, «Visualizing the Multistep Process of Protein Aggregation in Live Cells», *Acc. Chem. Res.*, vol. 55, n.º 3, pp. 381–390, fev. 2022, doi: 10.1021/acs.accounts.1c00648.

[6] Y. Ichikawa *et al.*, «The genomic structure and expression of MJD, the Machado-Joseph disease gene», *J. Hum. Genet.*, vol. 46, n.º 7, pp. 413–422, 2001, doi: 10.1007/s100380170060.

[7] C. Bettencourt e M. Lima, «Machado-Joseph Disease: from first descriptions to new perspectives», *Orphanet J. Rare Dis.*, vol. 6, p. 35, jun. 2011, doi: 10.1186/1750-1172-6-35.

[8] Y. Kawaguchi *et al.*, «CAG expansions in a novel gene for Machado-Joseph disease at chromosome 14q32.1», *Nat. Genet.*, vol. 8, n.º 3, pp. 221–228, nov. 1994, doi: 10.1038/ng1194-221.

[9] J. D. Da Silva, A. Teixeira-Castro, e P. Maciel, «From Pathogenesis to Novel Therapeutics for Spinocerebellar Ataxia Type 3: Evading Potholes on the Way to Translation», *Neurother. J. Am. Soc. Exp. Neurother.*, vol. 16, n.º 4, pp. 1009–1031, out. 2019, doi: 10.1007/s13311-019-00798-1.

[10] P. Maciel *et al.*, «Improvement in the molecular diagnosis of Machado-Joseph disease», *Arch. Neurol.*, vol. 58, n.º 11, pp. 1821–1827, nov. 2001, doi: 10.1001/archneur.58.11.1821.

[11] P. Maciel *et al.*, «Correlation between CAG repeat length and clinical features in Machado-Joseph disease», *Am. J. Hum. Genet.*, vol. 57, n.º 1, pp. 54–61, jul. 1995.

[12] G. Cancel *et al.*, «Marked phenotypic heterogeneity associated with expansion of a CAG repeat sequence at the spinocerebellar ataxia 3/Machado-Joseph disease locus», *Am. J. Hum. Genet.*, vol. 57, n.º 4, pp. 809–816, out. 1995.

[13] D. R. Carvalho, A. La Rocque-Ferreira, I. M. Rizzo, E. U. Imamura, e C. E. Speck-Martins, «Homozygosity enhances severity in spinocerebellar ataxia type 3», *Pediatr. Neurol.*, vol. 38, n.º 4, pp. 296–299, abr. 2008, doi: 10.1016/j.pediatrneurol.2007.12.006.

[14] T. Li *et al.*, «Is the High Frequency of Machado-Joseph Disease in China Due to New Mutational Origins?», *Front. Genet.*, vol. 9, p. 740, fev. 2019, doi: 10.3389/fgene.2018.00740.

[15] C. A. Matos, L. P. de Almeida, e C. Nóbrega, «Machado-Joseph disease/spinocerebellar ataxia type 3: lessons from disease pathogenesis and clues into therapy», *J. Neurochem.*, vol. 148, n.º 1, pp. 8–28, jan. 2019, doi: 10.1111/jnc.14541.

[16] A. T. Meira, J. L. Pedroso, F. Boller, G. L. Franklin, O. G. P. Barsottini, e H. A. G. Teive, «Reconstructing the History of Machado-Joseph Disease», *Eur. Neurol.*, vol. 83, n.º 1, pp. 99–104, 2020, doi: 10.1159/000507191.

[17] B. C. M. Garcia, F. M. B. Germiniani, P. Marques, J. Sequeiros, e H. A. G. Teive, «Paula Coutinho's outstanding contribution to the definition of Machado-Joseph disease», *Arq. Neuropsiquiatr.*, vol. 75, n.º 10, pp. 748–750, out. 2017, doi: 10.1590/0004-282x20170127.

[18] N. Mendonça, M. C. França, A. F. Gonçalves, e C. Januário, «Clinical Features of Machado-Joseph Disease», *Adv. Exp. Med. Biol.*, vol. 1049, pp. 255–273, 2018, doi: 10.1007/978-3-319-71779-1_13.

[19] C. Kieling, P. R. Prestes, M. L. Saraiva-Pereira, e L. B. Jardim, «Survival estimates for patients with Machado-Joseph disease (SCA3)», *Clin. Genet.*, vol. 72, n.º 6, pp. 543–545, dez. 2007, doi: 10.1111/j.1399-0004.2007.00910.x.

[20] N. Tzvetkov e P. Breuer, «Josephin domain-containing proteins from a variety of species are active de-ubiquitination enzymes», *Biol. Chem.*, vol. 388, n.º 9, pp. 973–978, set. 2007, doi: 10.1515/BC.2007.107.

[21] A. Ferro *et al.*, «NEDD8: A new ataxin-3 interactor», *Biochim. Biophys. Acta BBA - Mol. Cell Res.*, vol. 1773, n.º 11, pp. 1619–1627, nov. 2007, doi: 10.1016/j.bbamcr.2007.07.012.

[22] L. Masino *et al.*, «Domain architecture of the polyglutamine protein ataxin-3: a globular domain followed by a flexible tail», *FEBS Lett.*, vol. 549, n.º 1, pp. 21–25, ago. 2003, doi: 10.1016/S0014-5793(03)00748-8.

[23] G. Nicastro, R. P. Menon, L. Masino, P. P. Knowles, N. Q. McDonald, e A. Pastore, «The solution structure of the Josephin domain of ataxin-3: structural determinants for molecular recognition», *Proc Natl Acad Sci U A*, vol. 102, n.º 30, pp. 10493–8, 2005, doi: 10.1073/pnas.0501732102.

[24] G. M. Harris, K. Dodelzon, L. Gong, P. Gonzalez-Alegre, e H. L. Paulson, «Splice Isoforms of the Polyglutamine Disease Protein Ataxin-3 Exhibit Similar Enzymatic yet Different Aggregation Properties», *PLoS ONE*, vol. 5, n.º 10, out. 2010, doi: 10.1371/journal.pone.0013695.

[25] D. Weishäupl *et al.*, «Physiological and pathophysiological characteristics of ataxin-3 isoforms», *J Biol Chem*, vol. 294, n.º 2, pp. 644–661, 2019, doi: 10.1074/jbc.RA118.005801.

[26] C. Pozzi *et al.*, «Study of subcellular localization and proteolysis of ataxin-3», *Neurobiol. Dis.*, vol. 30, n.º 2, pp. 190–200, mai. 2008, doi: 10.1016/j.nbd.2008.01.011.

[27] P. M. A. Antony *et al.*, «Identification and functional dissection of localization signals within ataxin-3», *Neurobiol. Dis.*, vol. 36, n.º 2, pp. 280–292, nov. 2009, doi: 10.1016/j.nbd.2009.07.020.

[28] S. Macedo-Ribeiro, L. Cortes, P. Maciel, e A. L. Carvalho, «Nucleocytoplasmic Shuttling Activity of Ataxin-3», *PLOS ONE*, vol. 4, n.º 6, p. e5834, jun. 2009, doi: 10.1371/journal.pone.0005834.

[29] M. do C. Costa e H. L. Paulson, «Toward understanding Machado–Joseph disease», *Prog. Neurobiol.*, vol. 97, n.º 2, pp. 239–257, mai. 2012, doi: 10.1016/j.pneurobio.2011.11.006.

[30] W. L. DeLano, «The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC». 2010.

[31] B. Burnett, «The polyglutamine neurodegenerative protein ataxin-3 binds polyubiquitylated proteins and has ubiquitin protease activity», *Hum. Mol. Genet.*, vol. 12, n.º 23, pp. 3195–3205, set. 2003, doi: 10.1093/hmg/ddg344.

[32] B. J. Winborn *et al.*, «The Deubiquitinating Enzyme Ataxin-3, a Polyglutamine Disease Protein, Edits Lys ^{63} Linkages in Mixed Linkage Ubiquitin Chains», *J. Biol. Chem.*, vol. 283, n.º 39, pp. 26436–26443, set. 2008, doi: 10.1074/jbc.M803692200.

[33] A. Hershko e A. Ciechanover, «The Ubiquitin System», *Annu. Rev. Biochem.*, vol. 67, n.º 1, pp. 425–479, 1998, doi: 10.1146/annurev.biochem.67.1.425.

[34] E. Oh, D. Akopian, e M. Rape, «Principles of Ubiquitin-Dependent Signaling», *Annu. Rev. Cell Dev. Biol.*, vol. 34, n.º 1, pp. 137–162, 2018, doi: 10.1146/annurev-cellbio-100617-062802.

[35] K. K. Deol, S. Lorenz, e E. R. Strieter, «Enzymatic Logic of Ubiquitin Chain Assembly», *Front. Physiol.*, vol. 10, 2019, doi: 10.3389/fphys.2019.00835.

[36] F. E. R. Turcu, K. H. Ventii, e K. D. Wilkinson, «Regulation and Cellular Roles of Ubiquitin-specific Deubiquitinating Enzymes», *Annu. Rev. Biochem.*, vol. 78, pp. 363–397, 2009, doi: 10.1146/annurev.biochem.78.082307.091526.

[37] B. J. Winborn *et al.*, «The Deubiquitinating Enzyme Ataxin-3, a Polyglutamine Disease Protein, Edits Lys63 Linkages in Mixed Linkage Ubiquitin Chains», *J. Biol. Chem.*, vol. 283, n.º 39, pp. 26436–26443, set. 2008, doi: 10.1074/jbc.M803692200.

[38] N. Elu *et al.*, «Identification of substrates for human deubiquitinating enzymes (DUBs): An up-to-date review and a case study for neurodevelopmental disorders», *Semin. Cell Dev. Biol.*, p. S1084952122000039, jan. 2022, doi: 10.1016/j.semcdb.2022.01.001.

[39] S. V. Todi, B. J. Winborn, K. M. Scaglione, J. R. Blount, S. M. Travis, e H. L. Paulson, «Ubiquitination directly enhances activity of the deubiquitinating enzyme ataxin-3», *EMBO J.*, vol. 28, n.º 4, pp. 372–382, fev. 2009, doi: 10.1038/emboj.2008.289.

[40] S. V. Todi *et al.*, «Activity and Cellular Functions of the Deubiquitinating Enzyme and Polyglutamine Disease Protein Ataxin-3 Are Regulated by Ubiquitination at Lysine 117», *J. Biol. Chem.*, vol. 285, n.º 50, pp. 39303–39313, dez. 2010, doi: 10.1074/jbc.M110.181610.

[41] S. J. S. Berke, Y. Chai, G. L. Marrs, H. Wen, e H. L. Paulson, «Defining the Role of Ubiquitin-interacting Motifs in the Polyglutamine Disease Protein, Ataxin-3», *J. Biol. Chem.*, vol. 280, n.º 36, pp. 32026–32034, set. 2005, doi: 10.1074/jbc.M506084200.

[42] I. Schmitt *et al.*, «Inactivation of the mouse Atxn3 (ataxin-3) gene increases protein ubiquitination», *Biochem. Biophys. Res. Commun.*, vol. 362, n.º 3, pp. 734–739, out. 2007, doi: 10.1016/j.bbrc.2007.08.062.

[43] K. M. Scaglione *et al.*, «Ube2w and Ataxin-3 Coordinately Regulate the Ubiquitin Ligase CHIP», *Mol. Cell*, vol. 43, n.º 4, pp. 599–612, ago. 2011, doi: 10.1016/j.molcel.2011.05.036.

[44] R. Gao *et al.*, «Inactivation of PNKP by Mutant ATXN3 Triggers Apoptosis by Activating the DNA Damage-Response Pathway in SCA3», *PLOS Genet.*, vol. 11, n.º 1, p. e1004834, jan. 2015, doi: 10.1371/journal.pgen.1004834.

[45] A. Neves-Carvalho *et al.*, «Regulation of neuronal mRNA splicing and Tau isoform ratio by ATXN3 through deubiquitylation of splicing factors», *bioRxiv*, p. 711424, jul. 2019, doi: 10.1101/711424.

[46] A.-J. Rodrigues *et al.*, «Absence of ataxin-3 leads to cytoskeletal disorganization and increased cell death», *Biochim. Biophys. Acta BBA - Mol. Cell Res.*, vol. 1803, n.º 10, pp. 1154–1163, out. 2010, doi: 10.1016/j.bbamcr.2010.07.004.

[47] T. P. Chapman *et al.*, «Ataxin-3 Links NOD2 and TLR2 Mediated Innate Immune Sensing and Metabolism in Myeloid Cells», *Front. Immunol.*, vol. 10, p. 1495, jul. 2019, doi: 10.3389/fimmu.2019.01495.

[48] H. Liu *et al.*, «The Machado–Joseph Disease Deubiquitinase Ataxin-3 Regulates the Stability and Apoptotic Function of p53», *PLOS Biol.*, vol. 14, n.º 11, p. e2000733, nov. 2016, doi: 10.1371/journal.pbio.2000733.

[49] A.-J. Rodrigues *et al.*, «Functional genomics and biochemical characterization of the C. elegans orthologue of the Machado-Joseph disease protein ataxin-3», *FASEB J.*, vol. 21, n.º 4, pp. 1126–1136, jan. 2007, doi: 10.1096/fj.06-7002com.

[50] S. Alves *et al.*, «Silencing ataxin-3 mitigates degeneration in a rat model of Machado–Joseph disease: no role for wild-type ataxin-3?», *Hum. Mol. Genet.*, vol. 19, n.º 12, pp. 2380–2394, jun. 2010, doi: 10.1093/hmg/ddq111.

[51] H. L. Paulson *et al.*, «Intranuclear Inclusions of Expanded Polyglutamine Protein in Spinocerebellar Ataxia Type 3», *Neuron*, vol. 19, n.º 2, pp. 333–344, ago. 1997, doi: 10.1016/S0896-6273(00)80943-5.

[52] I. A. Klement *et al.*, «Ataxin-1 Nuclear Localization and Aggregation: Role in Polyglutamine-Induced Disease in SCA1 Transgenic Mice», *Cell*, vol. 95, n.º 1, pp. 41–53, out. 1998, doi: 10.1016/S0092-8674(00)81781-X.

[53] U. Bichelmeier *et al.*, «Nuclear localization of ataxin-3 is required for the manifestation of symptoms in SCA3: in vivo evidence», *J. Neurosci. Off. J. Soc. Neurosci.*, vol. 27, n.º 28, pp. 7418–7428, jul. 2007, doi: 10.1523/JNEUROSCI.4540-06.2007.

[54] A. S. Sowa *et al.*, «Karyopherin α-3 is a key protein in the pathogenesis of spinocerebellar ataxia type 3 controlling the nuclear localization of ataxin-3», *Proc. Natl. Acad. Sci. U. S. A.*, vol. 115, n.º 11, pp. E2624–E2633, mar. 2018, doi: 10.1073/pnas.1716071115.

[55] T. Mueller, P. Breuer, I. Schmitt, J. Walter, B. O. Evert, e U. Wüllner, «CK2-dependent phosphorylation determines cellular localization and stability of ataxin-3», *Hum. Mol. Genet.*, vol. 18, n.º 17, pp. 3334–3343, set. 2009, doi: 10.1093/hmg/ddp274.

[56] G. Wang, N. Sawai, S. Kotliarova, I. Kanazawa, e N. Nukina, «Ataxin-3, the MJD1 gene product, interacts with the two human homologs of yeast DNA repair protein RAD23, hHR23A and HHR23B», *Hum. Mol. Genet.*, vol. 9, n.º 12, pp. 1795–1803, jul. 2000, doi: 10.1093/hmg/9.12.1795.

[57] A. Chatterjee *et al.*, «The Role of the Mammalian DNA End-processing Enzyme Polynucleotide Kinase 3'-Phosphatase in Spinocerebellar Ataxia Type 3 Pathogenesis», *PLOS Genet.*, vol. 11, n.º 1, p. e1004749, jan. 2015, doi: 10.1371/journal.pgen.1004749.

[58] A. Chakraborty *et al.*, «Classical non-homologous end-joining pathway utilizes nascent RNA for error-free double-strand break repair of transcribed genes», *Nat. Commun.*, vol. 7, p. 13049, 05 2016, doi: 10.1038/ncomms13049.

[59] A. Chakraborty *et al.*, «Deficiency in classical nonhomologous end-joining–mediated repair of transcribed genes is linked to SCA3 pathogenesis», *Proc. Natl. Acad. Sci.*, vol. 117, n.º 14, pp. 8154–8165, abr. 2020, doi: 10.1073/pnas.1917280117.

[60] Y. Tu *et al.*, «Ataxin-3 promotes genome integrity by stabilizing Chk1», *Nucleic Acids Res.*, vol. 45, n.º 8, pp. 4532–4549, mai. 2017, doi: 10.1093/nar/gkx095.

[61] A. Pfeiffer *et al.*, «Ataxin-3 consolidates the MDC1-dependent DNA double-strand break response by counteracting the SUMO-targeted ubiquitin ligase RNF4», *EMBO J.*, vol. 36, n.º 8, pp. 1066–1083, abr. 2017, doi: 10.15252/embj.201695151.

[62] A. N. Singh *et al.*, «The p97–Ataxin 3 complex regulates homeostasis of the DNA damage response E3 ubiquitin ligase RNF8», *EMBO J.*, vol. 38, n.º 21, p. e102361, nov. 2019, doi: 10.15252/embj.2019102361.

[63] F. Li, T. Macfarlan, R. N. Pittman, e D. Chakravarti, «Ataxin-3 is a histone-binding protein with two independent transcriptional corepressor activities», *J Biol Chem*, vol. 277, n.º 47, pp. 45004–12, 2002, doi: 10.1074/jbc.M205259200.

[64] B. O. Evert *et al.*, «Ataxin-3 represses transcription via chromatin binding, interaction with histone deacetylase 3, and histone deacetylation», *J Neurosci*, vol. 26, n.º 44, pp. 11474–86, 2006, doi: 10.1523/jneurosci.2053-06.2006.

[65] J. Araujo *et al.*, «FOXO4-dependent upregulation of superoxide dismutase-2 in response to oxidative stress is impaired in spinocerebellar ataxia type 3», *Hum. Mol. Genet.*, vol. 20, n.º 15, pp. 2928–2941, ago. 2011, doi: 10.1093/hmg/ddr197.

[66] Y. Huang, T. A. Yario, e J. A. Steitz, «A molecular link between SR protein dephosphorylation and mRNA export», *Proc. Natl. Acad. Sci. U. S. A.*, vol. 101, n.º 26, pp. 9666–9670, jun. 2004, doi: 10.1073/pnas.0403533101.

[67] M. Müller-McNicoll *et al.*, «SR proteins are NXF1 adaptors that link alternative RNA processing to mRNA export», *Genes Dev.*, vol. 30, n.º 5, pp. 553–566, mar. 2016, doi: 10.1101/gad.276477.115.

[68] L. Gao, J. Wang, Y. Wang, e A. Andreadis, «SR protein 9G8 modulates splicing of tau exon 10 via its proximal downstream intron, a clustering region for frontotemporal dementia mutations», *Mol. Cell. Neurosci.*, vol. 34, n.º 1, pp. 48–58, jan. 2007, doi: 10.1016/j.mcn.2006.10.004.

[69] C. Luk *et al.*, «Development and assessment of sensitive immuno-PCR assays for the quantification of cerebrospinal fluid three- and four-repeat tau isoforms in tauopathies», *J. Neurochem.*, vol. 123, n.º 3, pp. 396–405, nov. 2012, doi: 10.1111/j.1471-4159.2012.07911.x.

[70] N. K. Magdalinou *et al.*, «Identification of candidate cerebrospinal fluid biomarkers in parkinsonism using quantitative proteomics», *Parkinsonism Relat. Disord.*, vol. 37, pp. 65–71, abr. 2017, doi: 10.1016/j.parkreldis.2017.01.016.

[71] X. Yin *et al.*, «Dyrk1A overexpression leads to increase of 3R-tau expression and cognitive deficits in Ts65Dn Down syndrome mice», *Sci. Rep.*, vol. 7, n.º 1, p. 619, abr. 2017, doi: 10.1038/s41598-017-00682-y.

[72] Y. Chai, L. Wu, J. D. Griffin, e H. L. Paulson, «The role of protein composition in specifying nuclear inclusion formation in polyglutamine disease», *J Biol Chem*, vol. 276, n.º 48, pp. 44889–97, 2001, doi: 10.1074/jbc.M106575200.

[73] K.-J. Wu, P.-M. Lei, H. Liu, C. Wu, C.-H. Leung, e D.-L. Ma, «Mimicking Strategy for Protein–Protein Interaction Inhibitor Discovery by Virtual Screening», *Molecules*, vol. 24, n.º 24, dez. 2019, doi: 10.3390/molecules24244428.

[74] T. F. Massoud, R. Paulmurugan, A. De, P. Ray, e S. S. Gambhir, «Reporter gene imaging of protein–protein interactions in living subjects», *Curr. Opin. Biotechnol.*, vol. 18, n.º 1, pp. 31–37, fev. 2007, doi: 10.1016/j.copbio.2007.01.007.

[75] F. Zhu, F. Li, X. Ling, Q. Liu, e B. Shen, «Disease associated protein-protein interaction network reconstruction based on comprehensive influence analysis», p. 29, 2019, doi: : https://doi.org/10.1101/2019.12.18.880997.

[76] C. Corbel *et al.*, «Screening for Protein-Protein Interaction Inhibitors Using a Bioluminescence Resonance Energy Transfer (BRET)–Based Assay in Yeast», *SLAS Discov. Adv. Sci. Drug Discov.*, vol. 22, n.º 6, pp. 751–759, jul. 2017, doi: 10.1177/2472555216689530.

[77] X. Ran e J. E. Gestwicki, «Inhibitors of Protein-Protein Interactions (PPIs): An Analysis of Scaffold Choices and Buried Surface Area», *Curr. Opin. Chem. Biol.*, vol. 44, pp. 75–86, jun. 2018, doi: 10.1016/j.cbpa.2018.06.004.

[78] S. Gul e K. Hadian, «Protein–protein interaction modulator drug discovery: past efforts and future opportunities using a rich source of low- and high-throughput screening assays», *Expert Opin. Drug Discov.*, vol. 9, n.º 12, pp. 1393–1404, dez. 2014, doi: 10.1517/17460441.2014.954544.

[79] I. Petta, S. Lievens, C. Libert, J. Tavernier, e K. De Bosscher, «Modulation of Protein–Protein Interactions for the Development of Novel Therapeutics», *Mol. Ther.*, vol. 24, n.º 4, pp. 707–718, abr. 2016, doi: 10.1038/mt.2015.214.

[80] D. Fayne, «De-peptidising protein–protein interactions – big jobs for small molecules», *Drug Discov. Today Technol.*, vol. 10, n.º 4, pp. e467–e474, dez. 2013, doi: 10.1016/j.ddtec.2013.08.002.

[81] M. A. Kuenemann, L. M. L. Bourbon, C. M. Labbé, B. O. Villoutreix, e O. Sperandio, «Which Three-Dimensional Characteristics Make Efficient Inhibitors of Protein–Protein Interactions?», *J. Chem. Inf. Model.*, vol. 54, n.º 11, pp. 3067–3079, nov. 2014, doi: 10.1021/ci500487q.

[82] F. Koraïchi *et al.*, «High-content tripartite split-GFP cell-based assays to screen for modulators of small GTPase activation», *J. Cell Sci.*, vol. 131, n.º 1, jan. 2018, doi: 10.1242/jcs.210419.

[83] S. Dwane e P. A. Kiely, «Tools used to study how protein complexes are assembled in signaling cascades», *Bioeng. Bugs*, vol. 2, n.º 5, pp. 247–259, 2011, doi: 10.4161/bbug.2.5.17844.

[84] L. Yurlova *et al.*, «The Fluorescent Two-Hybrid Assay to Screen for Protein–Protein Interaction Inhibitors in Live Cells: Targeting the Interaction of p53 with Mdm2 and Mdm4», *J. Biomol. Screen.*, vol. 19, n.º 4, pp. 516–525, abr. 2014, doi: 10.1177/1087057113518067.

[85] J.-D. Pedelacq, G. S. Waldo, e S. Cabantous, «High-Throughput Protein–Protein Interaction Assays Using Tripartite Split-GFP Complementation», em *High-Throughput Protein Production and*

Purification, vol. 2025, R. Vincentelli, Ed. New York, NY: Springer New York, 2019, pp. 423–437. doi: 10.1007/978-1-4939-9624-7_20.

[86] I. Weibrecht *et al.*, «Proximity ligation assays: a recent addition to the proteomics toolbox», *Expert Rev. Proteomics*, vol. 7, n.º 3, pp. 401–409, jun. 2010, doi: 10.1586/epr.10.10.

[87] J. M. Walker e R. Rapley, Eds., *Molecular Biomethods Handbook*. Totowa, NJ: Humana Press, 2008. doi: 10.1007/978-1-60327-375-6.

[88] H. Kobayashi, L.-P. Picard, A.-M. Schönegge, e M. Bouvier, «Bioluminescence resonance energy transfer–based imaging of protein–protein interactions in living cells», *Nat. Protoc.*, vol. 14, n.º 4, pp. 1084–1107, abril 2019, doi: 10.1038/s41596-019-0129-7.

[89] A. Pietraszewska-Bogiel e T. W. J. Gadella, «FRET microscopy: from principle to routine technology in cell biology: FRET MICROSCOPY», *J. Microsc.*, vol. 241, n.º 2, pp. 111–118, fev. 2011, doi: 10.1111/j.1365-2818.2010.03437.x.

[90] H.-T. Lai e C.-M. Chiang, «Bimolecular Fluorescence Complementation (BiFC) Assay for Direct Visualization of Protein-Protein Interaction in vivo», *Bio-Protoc.*, vol. 3, n.º 20, 2013, doi: 10.21769/bioprotoc.935.

[91] S. Cabantous *et al.*, «A New Protein-Protein Interaction Sensor Based on Tripartite Split-GFP Association», *Sci. Rep.*, vol. 3, n.º 1, p. 2854, dez. 2013, doi: 10.1038/srep02854.

[92] I. Miyai *et al.*, «Cerebellar Ataxia Rehabilitation Trial in Degenerative Cerebellar Diseases», *Neurorehabil. Neural Repair*, vol. 26, n.º 5, pp. 515–522, jun. 2012, doi: 10.1177/1545968311425918.

[93] M. Svensson, J. Lexell, e T. Deierborg, «Effects of Physical Exercise on Neuroinflammation, Neuroplasticity, Neurodegeneration, and Behavior: What We Can Learn From Animal Models in Clinical Settings», *Neurorehabil. Neural Repair*, vol. 29, n.º 6, pp. 577–589, jul. 2015, doi: 10.1177/1545968314562108.

[94] P. J. Tuite, E. A. Rogaeva, P. H. St George-Hyslop, e A. E. Lang, «Dopa-responsive parkinsonism phenotype of Machado-Joseph disease: confirmation of 14q CAG expansion», *Ann. Neurol.*, vol. 38, n.º 4, pp. 684–687, out. 1995, doi: 10.1002/ana.410380422.

[95] C. Buhmann, A. Bussopulos, e M. Oechsner, «Dopaminergic response in Parkinsonian phenotype of Machado-Joseph disease», *Mov. Disord. Off. J. Mov. Disord. Soc.*, vol. 18, n.º 2, pp. 219–221, fev. 2003, doi: 10.1002/mds.10322.

[96] R. Nandagopal e S. G. K. Moorthy, «Dramatic levodopa responsiveness of dystonia in a sporadic case of spinocerebellar ataxia type 3», *Postgrad. Med. J.*, vol. 80, n.º 944, pp. 363–365, jun. 2004, doi: 10.1136/pgmj.2003.015297.

[97] P. Binarová e J. Tuszynski, «Tubulin: Structure, Functions and Roles in Disease», *Cells*, vol. 8, n.º 10, p. 1294, out. 2019, doi: 10.3390/cells8101294.

[98] H.-C. Hsieh *et al.*, «hHR23A, a human homolog of Saccharomyces cerevisiae Rad23, regulates xeroderma pigmentosum C protein and is required for nucleotide excision repair», *Biochem. Biophys. Res. Commun.*, vol. 335, n.º 1, pp. 181–187, set. 2005, doi: 10.1016/j.bbrc.2005.07.067.

[99] B. Medicherla, Z. Kostova, A. Schaefer, e D. H. Wolf, «A genomic screen identifies Dsk2p and Rad23p as essential components of ER-associated degradation», *EMBO Rep.*, vol. 5, n.º 7, pp. 692–697, jul. 2004, doi: 10.1038/sj.embor.7400164.

[100] E. K. O'Shea, J. D. Klemm, P. S. Kim, e T. Alber, «X-Ray Structure of the GCN4 Leucine Zipper, a Two-Stranded, Parallel Coiled Coil», *Science*, vol. 254, n.º 5031, pp. 539–544, out. 1991, doi: 10.1126/science.1948029.

[101] D. G. Gibson, L. Young, R.-Y. Chuang, J. C. Venter, C. A. Hutchison, e H. O. Smith, «Enzymatic assembly of DNA molecules up to several hundred kilobases», *Nat. Methods*, vol. 6, n.º 5, pp. 343–345, mai. 2009, doi: 10.1038/nmeth.1318.

[102] S. Mazzucchelli *et al.*, «Proteomic and biochemical analyses unveil tight interaction of ataxin-3 with tubulin», *Int. J. Biochem. Cell Biol.*, vol. 41, n.º 12, pp. 2485–2492, dez. 2009, doi: 10.1016/j.biocel.2009.08.003.

[103] S. V. Todi, M. N. Laco, B. J. Winborn, S. M. Travis, H. M. Wen, e H. L. Paulson, «Cellular Turnover of the Polyglutamine Disease Protein Ataxin-3 Is Regulated by Its Catalytic Activity», *J. Biol. Chem.*, vol. 282, n.º 40, pp. 29348–29358, out. 2007, doi: 10.1074/jbc.M704126200.

[104] A. Silva-Fernandes *et al.*, «Chronic Treatment with 17-DMAG Improves Balance and Coordination in A New Mouse Model of Machado-Joseph Disease», *Neurotherapeutics*, vol. 11, n.º 2, pp. 433–449, abr. 2014, doi: 10.1007/s13311-013-0255-9.

[105] P. Breuer, A. Haacke, B. O. Evert, e U. Wüllner, «Nuclear Aggregation of Polyglutamine-expanded Ataxin-3», *J. Biol. Chem.*, vol. 285, n.º 9, pp. 6532–6537, fev. 2010, doi: 10.1074/jbc.M109.036335.

[106] M. Bonanomi *et al.*, «Interactions of ataxin-3 with its molecular partners in the protein machinery that sorts protein aggregates to the aggresome», *Int. J. Biochem. Cell Biol.*, vol. 51, pp. 58–64, jun. 2014, doi: 10.1016/j.biocel.2014.03.015.

[107] M. Masuda-Suzukake *et al.*, «Pathological alpha-synuclein propagates through neural networks», *Acta Neuropathol. Commun.*, vol. 2, p. 88, ago. 2014, doi: 10.1186/s40478-014-0088-8.

[108] W. Chun, G. S. Waldo, e G. V. W. Johnson, «Split GFP complementation assay: a novel approach to quantitatively measure aggregation of tau in situ: effects of GSK3beta activation and caspase 3 cleavage», *J. Neurochem.*, vol. 103, n.º 6, pp. 2529–2539, dez. 2007, doi: 10.1111/j.1471-4159.2007.04941.x.

[109] M. A. Basson, «Signaling in Cell Differentiation and Morphogenesis», *Cold Spring Harb. Perspect. Biol.*, vol. 4, n.º 6, p. a008151, jun. 2012, doi: 10.1101/cshperspect.a008151.

[110] S. A. Harvey e J. C. Smith, «Visualisation and quantification of morphogen gradient formation in the zebrafish», *PLoS Biol.*, vol. 7, n.º 5, p. e1000101, mai. 2009, doi: 10.1371/journal.pbio.1000101.

[111] S. Feng *et al.*, «Bright split red fluorescent proteins for the visualization of endogenous proteins and synapses», *Commun. Biol.*, vol. 2, p. 344, set. 2019, doi: 10.1038/s42003-019-0589-x.

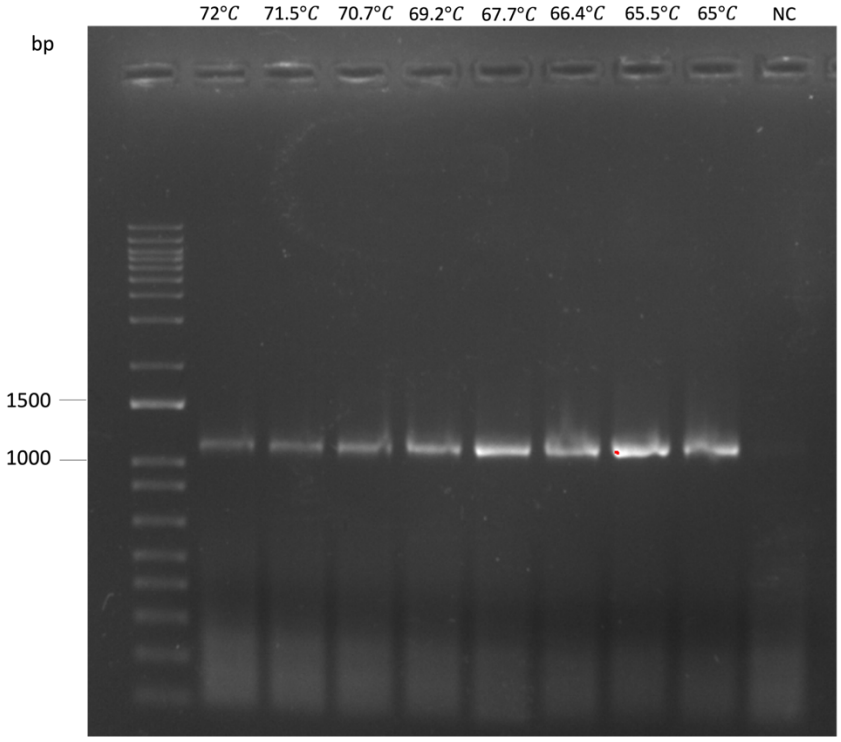


Figure 36 - Gradient PCR of WT ATXN3 gene amplification for GFP11 vector. The selected temperature for gene isolation was 67.7°C. NC – Negative control.

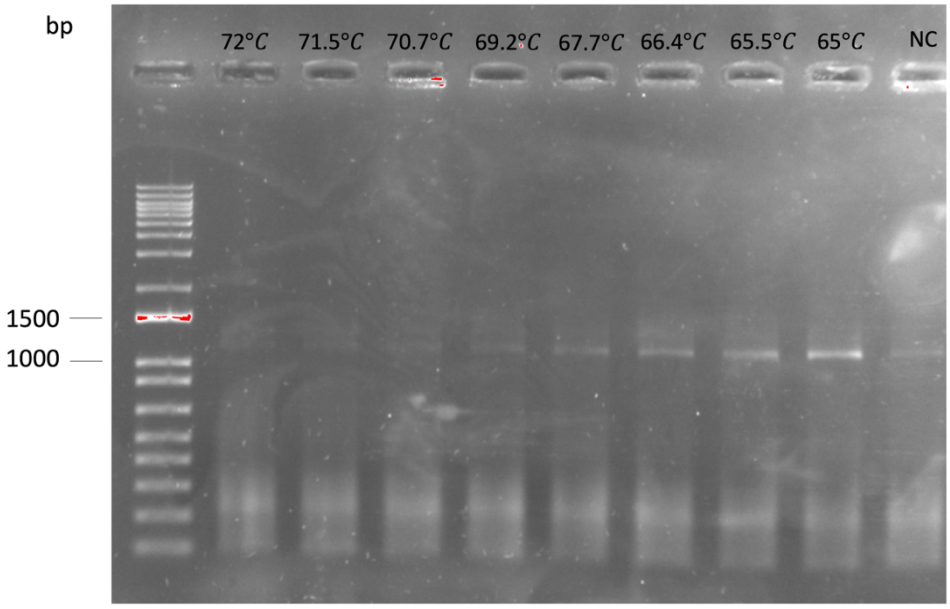


Figure 37 - Gradient PCR of mutant ATXN3 gene amplification for GFP11 vector. The selected temperature for gene isolation was 65°C. NC – Negative control.

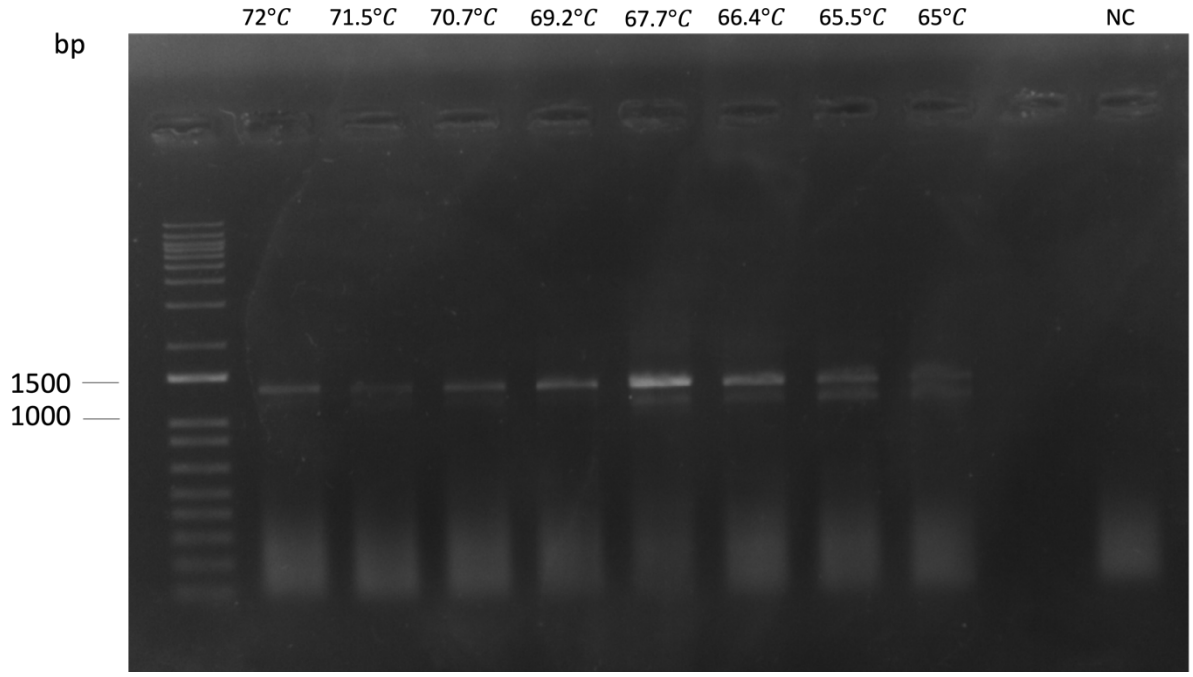


Figure 38 - Gradient PCR of mutant ATXN3 gene amplification for GFP10 vector. The selected temperature for gene isolation was 69.2°$\boldsymbol{C}$. NC – Negative control.

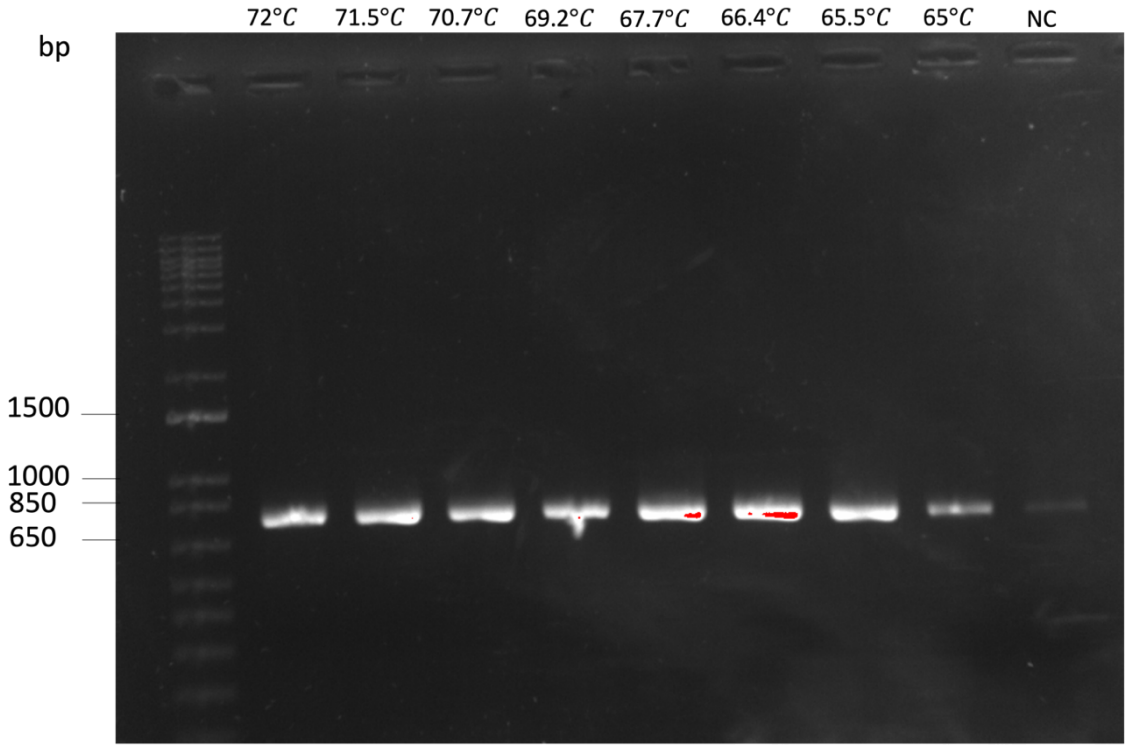


Figure 39 - Gradient PCR of 9G8 gene amplification for GFP10 vector. The selected temperature for gene isolation was 67.7°$\boldsymbol{C}$. NC – Negative control.

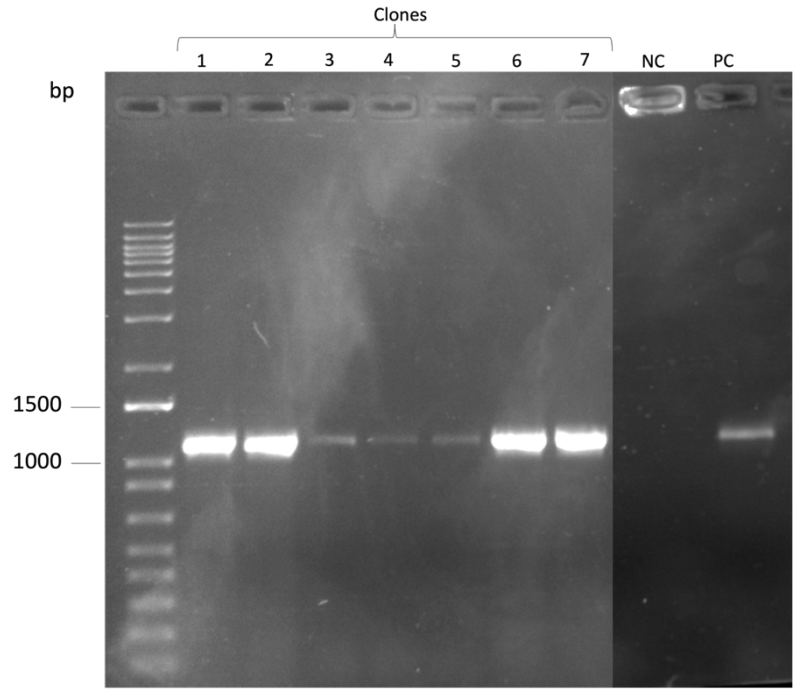


Figure 40 – Colony PCR of the pcDNA GFP10-WT ATXN3 vector before purification, to confirm insertion of WT ATXN3 in pcDNA GFP10 vector. WT ATXN3 for GFP10 size is 1131 bp. Clones 2 and 6 were selected. NC – Negative control. PC – Positive control.

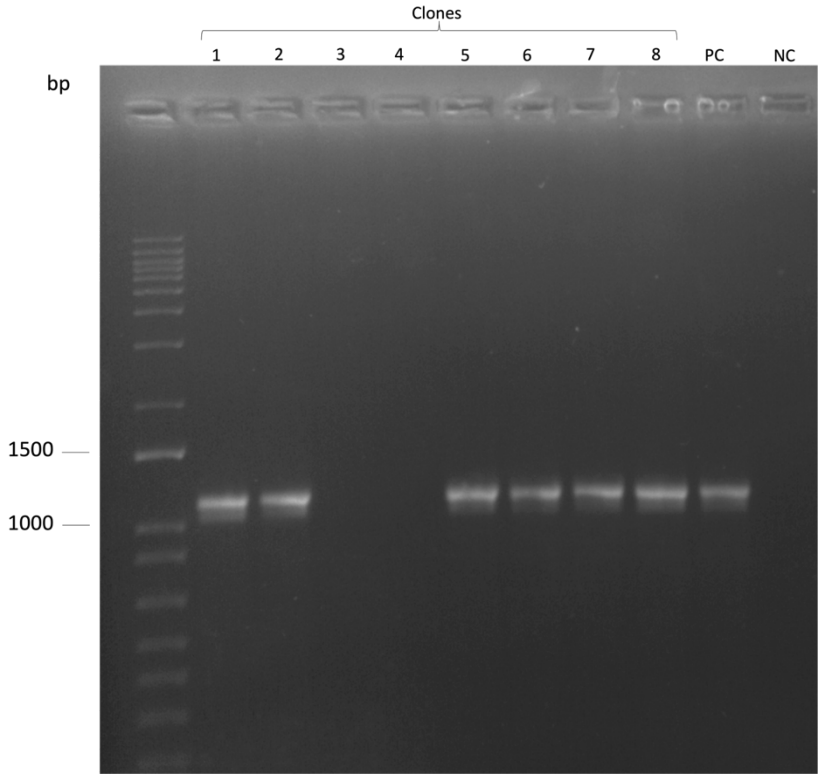


Figure 41 – Colony PCR of the pcDNA WT ATXN3-GFP11 vector before purification, to confirm insertion of WT ATXN3 in pcDNA GFP11 vector. WT ATXN3 for GFP11 size is 1136 bp. Clone 2 was selected. NC – Negative control. PC – Positive control.

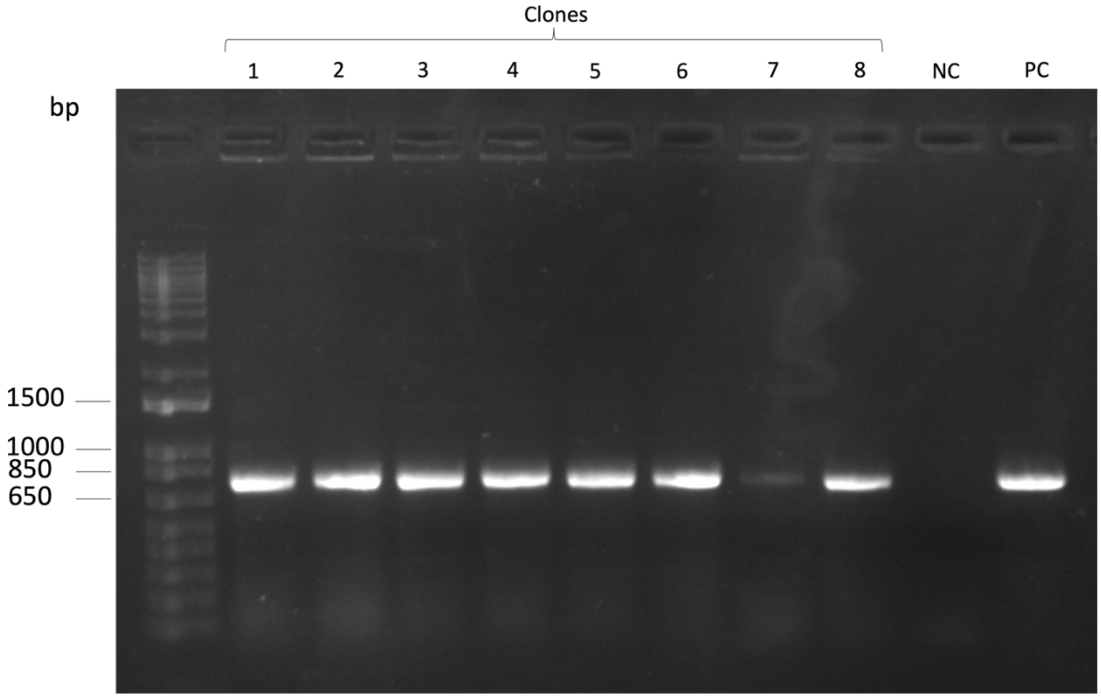


Figure 42 – Colony PCR of the pcDNA GFP10-9G8 vector before purification, to confirm insertion of 9G8 in pcDNA GFP10 vector. 9G8 for GFP10 size is 761 bp. Clones 1 and 2 were selected. NC – Negative control. PC – Positive control.

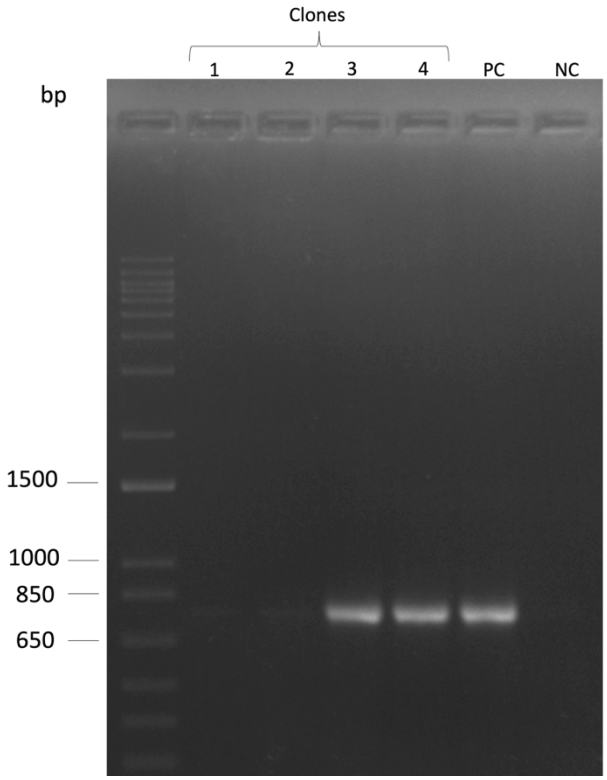


Figure 43 – Colony PCR of the pcDNA 9G8-GFP11 vector before purification, to confirm insertion of 9G8 in pcDNA GFP11 vector. 9G8 for GFP11 size is 758 bp. Clone 4 was selected. NC – Negative control. PC – Positive control.

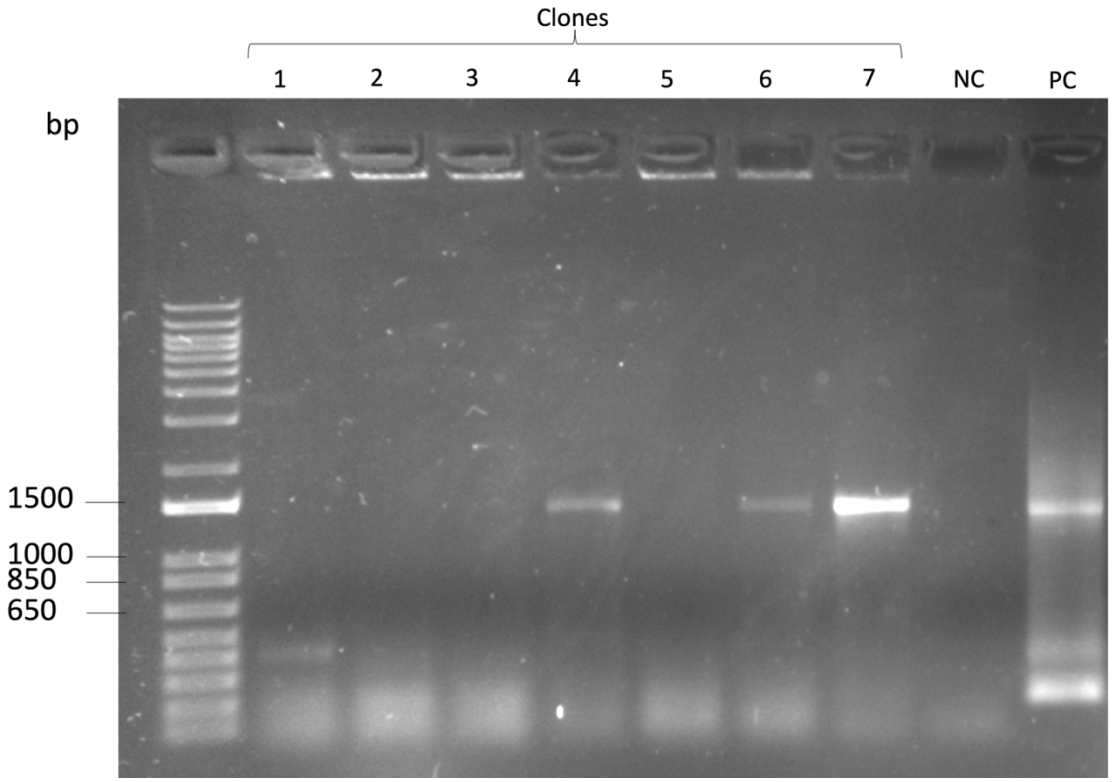


Figure 44 – Colony PCR of the pcDNA GFP10-Tubulin vector before purification, to confirm insertion of tubulin in pcDNA GFP10 vector. Tubulin size for GFP10 is 1400 bp. Clones 4 and 7 were selected. NC – Negative control. PC – Positive control.

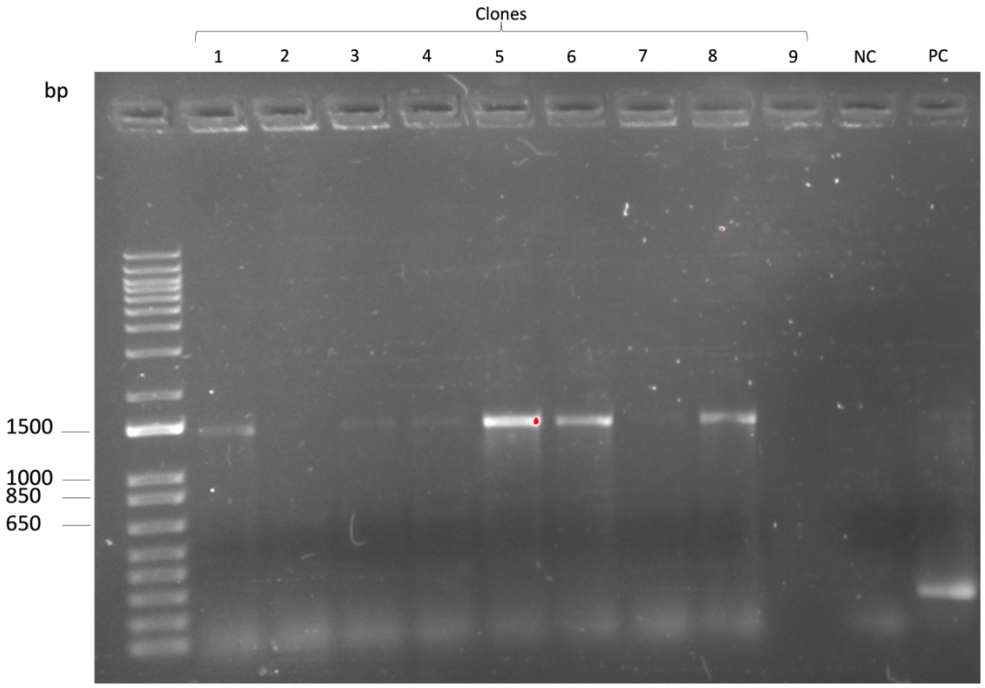


Figure 45 – Colony PCR of the pcDNA Tubulin-GFP11 vector before purification, to confirm insertion of tubulin in pcDNA GFP11 vector. Tubulin size for GFP11 is 1401 bp. Clones 5 and 6 were selected. NC – Negative control. PC – Positive control.

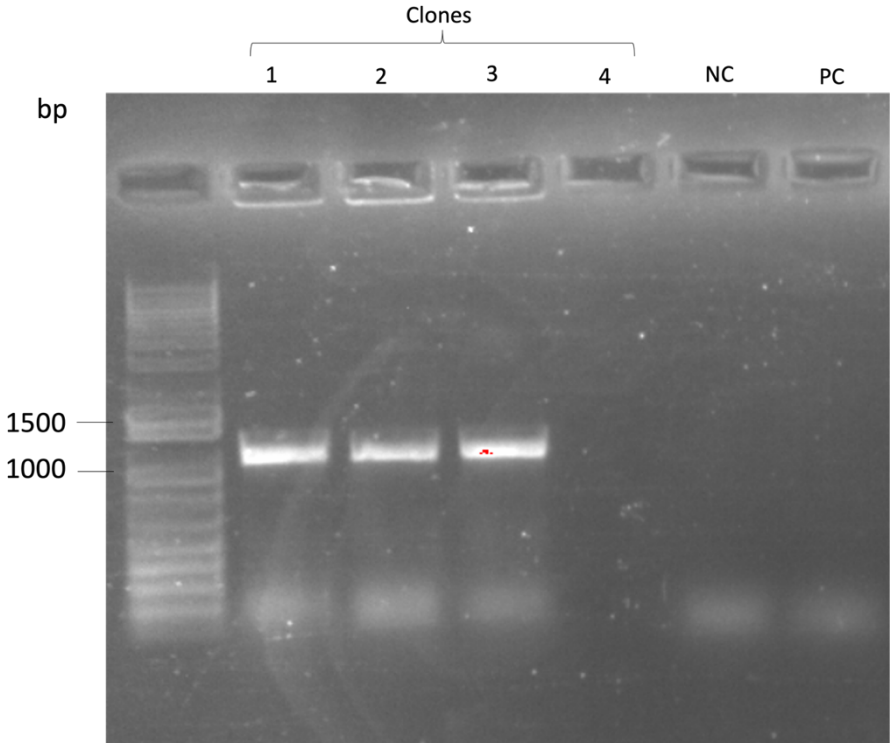


Figure 46 – Colony PCR of the pcDNA GFP10-hHR23A vector before purification, to confirm insertion of hHR23A in pcDNA GFP10 vector. hHR23A for GFP10 size is 1136 bp. Clone 3 was selected. NC – Negative control. PC – Positive control.

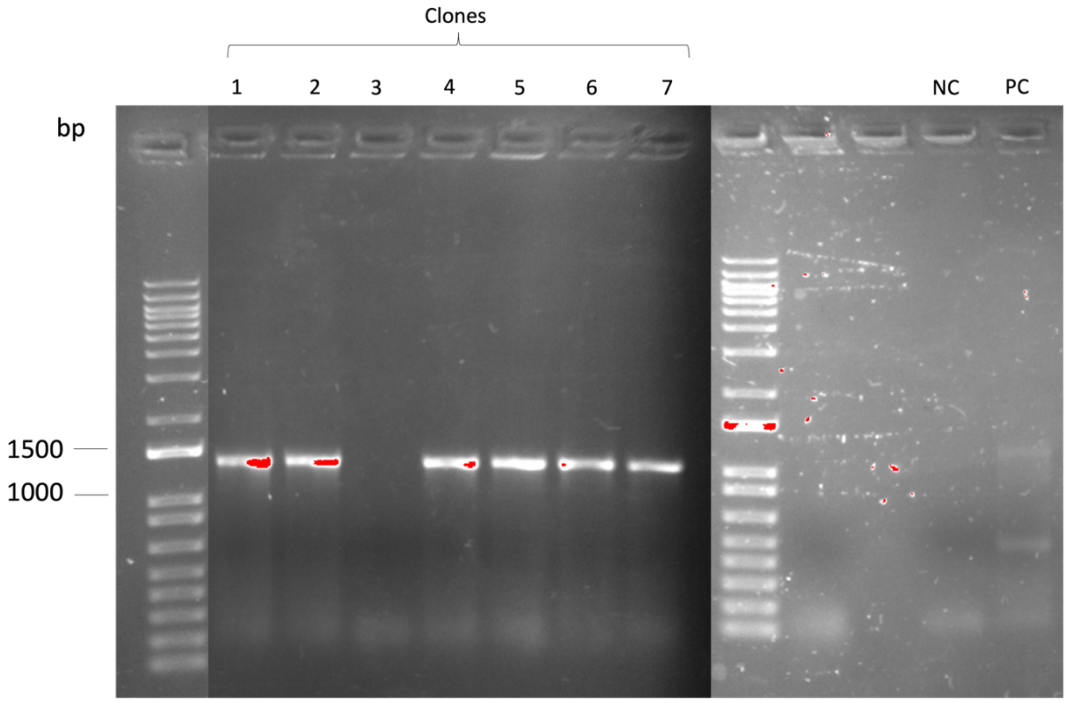


Figure 47 – Colony PCR of the pcDNA hHR23A-GFP11 vector before purification, to confirm insertion of hHR23A in pcDNA GFP11 vector. hHR23A for GFP11 size is 1133 bp. Clone 2 was selected. NC – Negative control. PC – Positive control.

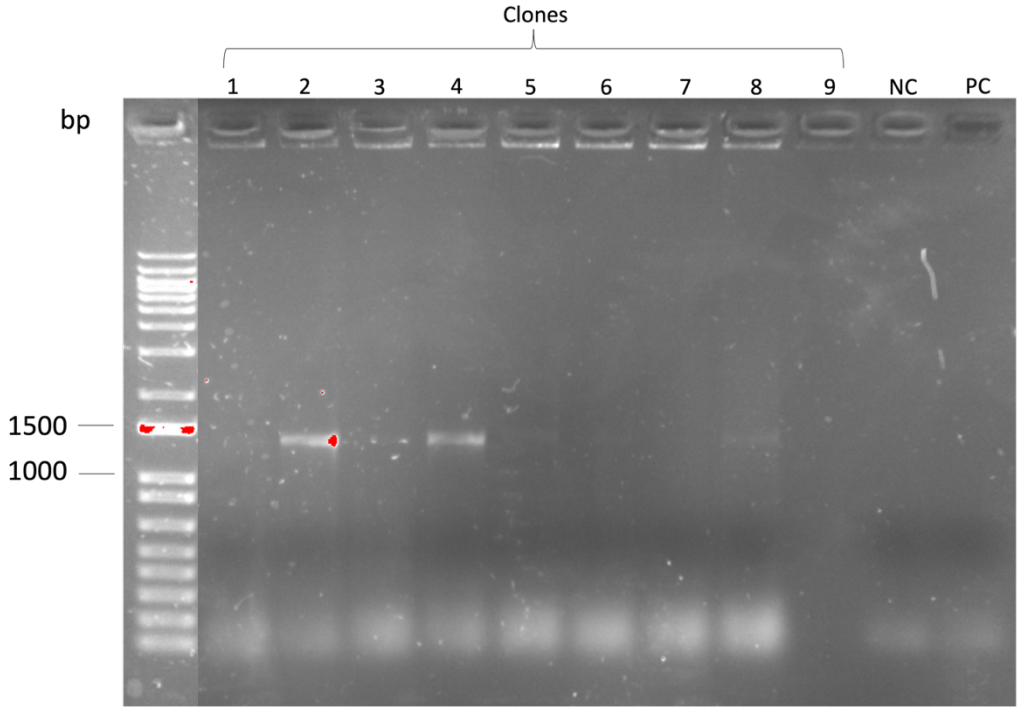


Figure 48 – Colony PCR of the pcDNA GFP10-Mutant ATXN3 vector before purification, to confirm insertion of Mutant ATXN3 in pcDNA GFP10 vector. Mutant ATXN3 for GFP10 size is 1323 bp. Clones 2 and 4 were selected. NC – Negative control. PC – Positive control.

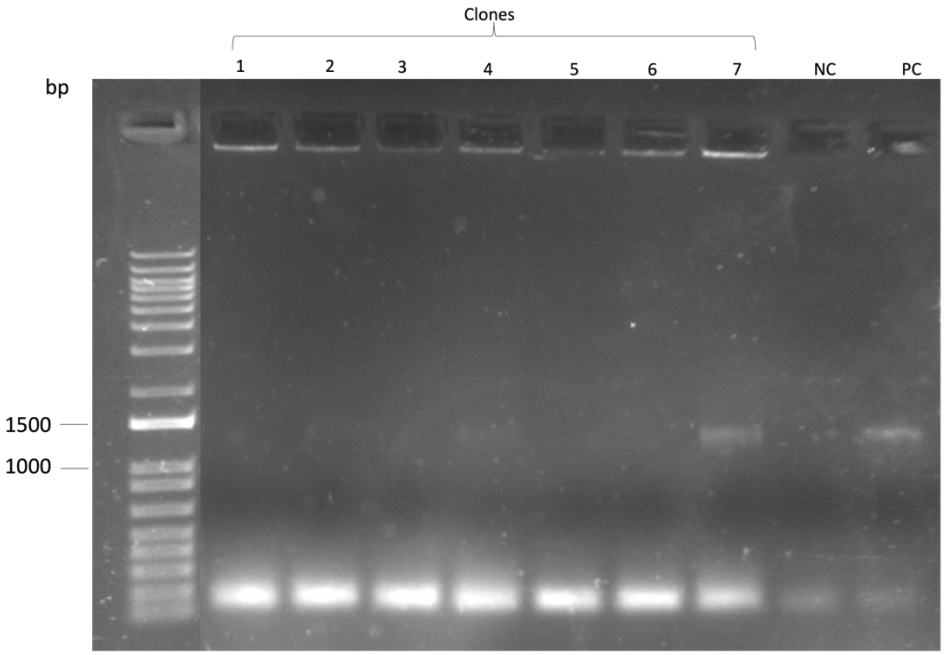


Figure 49 – Colony PCR of the pcDNA Mutant ATXN3-GFP11 vector before purification, to confirm insertion of Mutant ATXN3 in pcDNA GFP11 vector. Mutant ATXN3 for GFP11 size is 1328 bp. Clones 4 and 7 were selected. NC – Negative control. PC – Positive control.

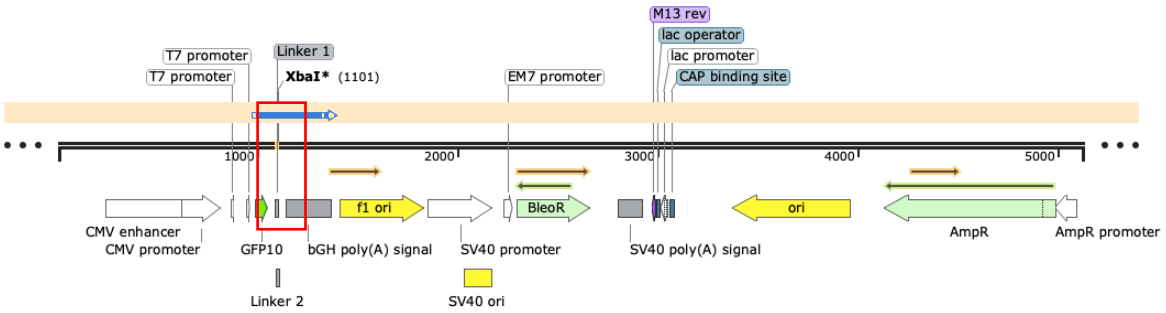
Table 2 – Constructs used for triSFP system, enzymes used for confirmation of insert of genes of interest by restriction digestion, primers used for vectors sequencing, selected clones for sequencing, and construct sizes (bp). All vectors have the backbone pcDNA 3.1.

Constructs	Enzyme restriction for confirmation	Sequencing primers (5' to 3')	Sequenced clones**	Construct size (bp)
pcDNA GFP10-zipper*				5263
pcDNA zipper-GFP11*				5265
pcDNA GFP10	ApaI and BglII	Primer Forward (Fw) T7 promoter: TAATACGACTCACTATAGGG	**Clone 1** Clone 3	5125
pcDNA GFP11	ApaI and BglII	Primer Fw T7 promoter - TAATACGACTCACTATAGGG	**Clone 2** Clone 4	5101
pcDNA GFP10-WT ATXN3	ApaI and BglII	Primer Fw T7 promoter - TAATACGACTCACTATAGGG Primer Reverse (Rv) - GGGGCAAACAACAGATGGCTG	**Clone 2** Clone 6	6212
pcDNA WT ATXN3-GFP11	ApaI and HindIII	Primer Fw T7 promoter - TAATACGACTCACTATAGGG Primer Rv - GGCGTCGGTGATGCCGGCGGC	**Clone 2**	6181
pcDNA GFP10-9G8	ApaI and BglII	Primer Fw T7 promoter - TAATACGACTCACTATAGGG Primer Rv - GGGGCAAACAACAGATGGCTG	**Clone 1** Clone 2	5844
pcDNA 9G8-GFP11	ApaI and HindIII	Primer Fw T7 promoter - TAATACGACTCACTATAGGG Primer Rv - GGCGTCGGTGATGCCGGCGGC	**Clone 4**	5815
pcDNA GFP10-Tubulin	SalI	Primer Fw T7 promoter - TAATACGACTCACTATAGGG Primer Rv - GGGGCAAACAACAGATGGCTG	**Clone 4** Clone 7	6481
pcDNA Tubulin-GFP11	SalI	Primer Fw T7 promoter - TAATACGACTCACTATAGGG Primer Rv - GGCGTCGGTGATGCCGGCGGC	Clone 5 **Clone 6**	6454
pcDNA GFP10-hHR23A	SalI	Primer Fw T7 promoter - TAATACGACTCACTATAGGG Primer Rv - GGGGCAAACAACAGATGGCTG	**Clone 3**	6219
pcDNA hHR23A-GFP11	SalI	Primer Fw T7 promoter - TAATACGACTCACTATAGGG Primer Rv - GGCGTCGGTGATGCCGGCGGC	**Clone 2**	6190
pcDNA GFP10-Mutant ATXN3	SalI	Primer Fw T7 promoter - TAATACGACTCACTATAGGG Primer Rv - GGGGCAAACAACAGATGGCTG	Clone 2 **Clone 4**	6404

pcDNA Mutant ATXN3-GFP11	SalI	Primer Fw T7 promoter - TAATACGACTCACTATAGGG Primer Rv - GGCGTCGGTGATGCCGGCGGC	**Clone 4** Clone 7	6373
pcDNA WT ATXN3 R282T NLS – GFP11		Primer Fw T7 promoter - TAATACGACTCACTATAGGG	**Clone 1**	6181
pcDNA GFP10 - Mutant ATXN3 R282T NLS		Primer Fw T7 promoter - TAATACGACTCACTATAGGG Primer Rv - GGGGCAAACAACAGATGGCTG	**Clone 1**	6404
pcDNA Mutant ATXN3 R282T NLS – GFP11		Primer Fw T7 promoter - TAATACGACTCACTATAGGG Primer Rv - ACTAGAAGGCACAGTCGAGGC	**Clone 1**	6373

* Constructs sent by Dr. Stéphanie Cabantous from the Cancer Research Center of Toulouse.
** Clones in bold were the ones selected for transfection.

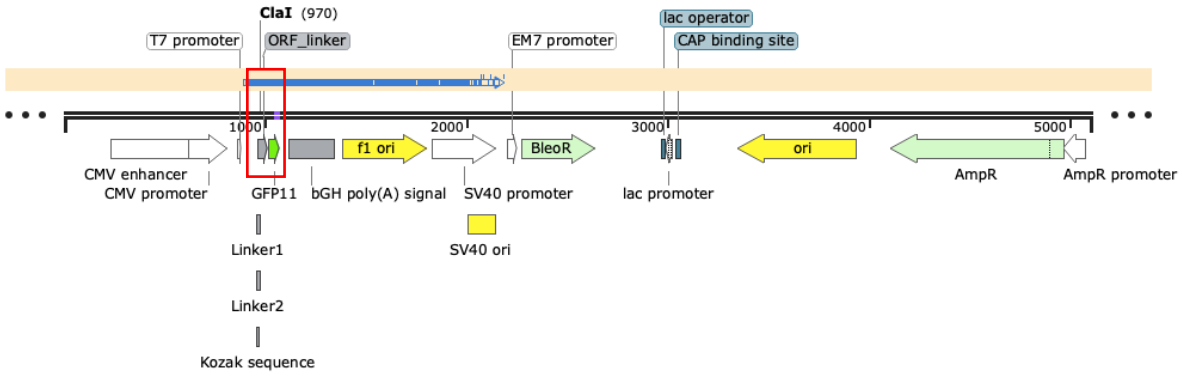
pcDNA GFP10 vector alignment



pcDNA GFP10_cloning	987	ATGGGCGACCTGCCCGACGACCACTACCTGTCCACCCAGACCATCC
Primer Forward pcDNA GFP10	16	NTGGGCGACCTGCCCGACGACCACTACCTGTCCACCCAGACCATCC
pcDNA GFP10_cloning	1033	TGAGCAAGGACCTGAACATCGATGGTGGCGGTGGCTCTGGAGGTGGTGGG
Primer Forward pcDNA GFP10	62	TGAGCAAGGACCTGAACATCGATGGTGGCGGTGGCTCTGGAGGTGGTGGG
pcDNA GFP10_cloning	1083	TCCTCCGGCGGCGGTGGATCTAGAGGGCCCGTTTAAACCCGCTGATCAGC
Primer Forward pcDNA GFP10	112	TCCTCCGGCGGCGGTGGATCTAGAGGGCCCGTTTAAACCCGCTGATCAGC
pcDNA GFP10_cloning	1133	CTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCG
Primer Forward pcDNA GFP10	162	CTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCG

Figure 50 - Sequence alignment of the pcDNA GFP10 vector with primer detailed in table 2, obtained from SnapGene® software (from Insightful Science; available at snapgene.com). Red – Unconfirmed sequenced nucleotide.

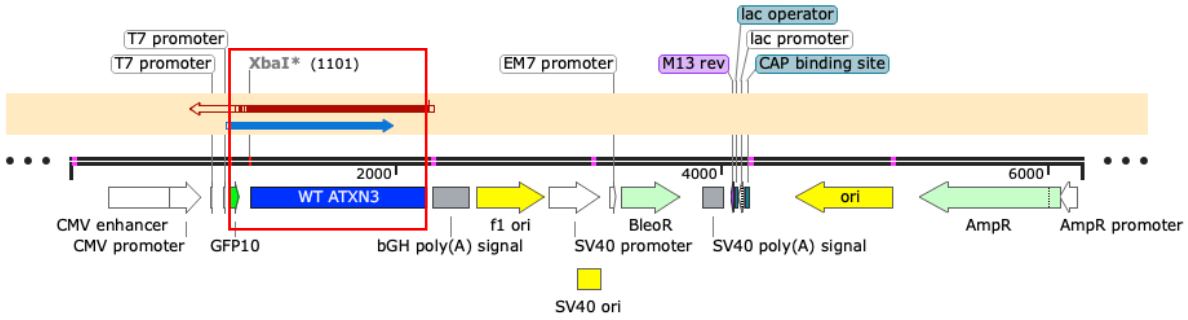
pcDNA GFP11 vector alignment



```
pcDNA GFP11                  907   ACTTAAGCTTGGTACCGAGCTCGGATCCACTAGTCCAGTGTGGTGGAATT
Primer Forward pcDNA GFP11   16    ACTTAAGCTTGGTACCGAGCTCGGATCCACTAGTCCAGTGTGGTGGAATT

pcDNA GFP11                  957   GCCACCATGGCTATCGATGGTGGCGGTGGCTCTGGAGGTGGTGGGTCCTC
Primer Forward pcDNA GFP11   66    GCCACCATGGCTATCGATGGTGGCGGTGGCTCTGGAGGTGGTGGGTCCTC

pcDNA GFP11                  1007  CGGAGAGAAGCGCGACCACATGGTGCTGCTGGAGTACGTGACCGCCGCCG
Primer Forward pcDNA GFP11   116   CGGAGAGAAGCGCGACCACATGGTGCTGCTGGAGTACGTGACCGCCGCCG

pcDNA GFP11                  1057  GCATCACCGACGCCTCCTAATCTAGAGGGCCCGTTTAAACCCGCTGATCA
Primer Forward pcDNA GFP11   166   GCATCACCGACGCCTCCTAATCTAGAGGGCCCGTTTAAACCCGCTGATCA

pcDNA GFP11                  1107  GCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCC
Primer Forward pcDNA GFP11   216   GCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCC
```

Figure 51 - Sequence alignment of the pcDNA GFP11 vector with primer detailed in table 2, obtained from SnapGene® software (from Insightful Science; available at snapgene.com).

pcDNA GFP10 - WT ATXN3 vector alignment



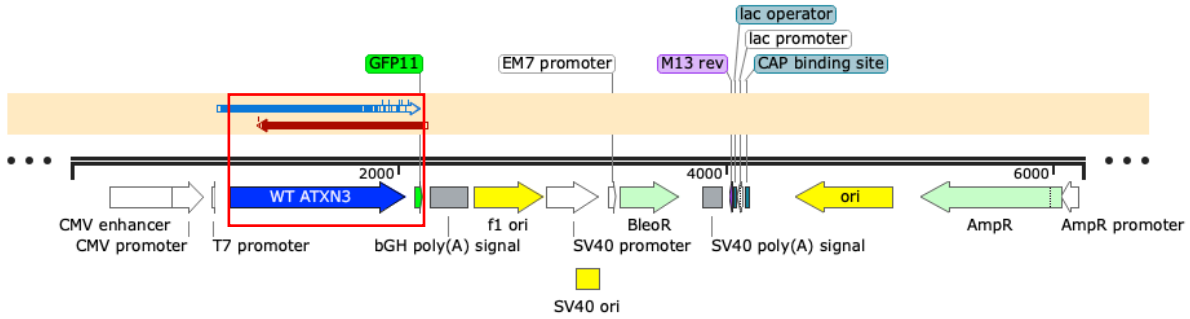

```
pcDNA GFP10-WT ATXN3                 987   ATGGGCGACCTGCC
Primer Forward pcDNA GFP10-WT ATXN3  22    NTGGGCGACCTGCC
Primer Reverse pcDNA GFP10-WT ATXN3  212   AT-GGCGACTTGCC

pcDNA GFP10-WT ATXN3                 1001  CGACGACCACTACCTGTCCACCCAGACCATCCTGAGCAAGGACCTGAACA
Primer Forward pcDNA GFP10-WT ATXN3  36    CGACGACCACTACCTGTCCACCCAGACCATCCTGAGCAAGGACCTGAACA
Primer Reverse pcDNA GFP10-WT ATXN3  260   CGAGGACCCATACCTGTCCACCCAGACCNT-CTGAGCAAGGACCTGAACA

pcDNA GFP10-WT ATXN3                 1051  TCGATGGTGGCGGTGGCTCTGGAGGTGGTGGGTCCTCCGGCGGCGGTGGA
Primer Forward pcDNA GFP10-WT ATXN3  86    TCGATGGTGGCGGTGGCTCTGGAGGTGGTGGGTCCTCCGGCGGCGGTGGA
Primer Reverse pcDNA GFP10-WT ATXN3  309   TCGATGGTGGC-GTGGCTCTGGAGGTGGT-GGTCCTCCGGCGGCGGTGGA
```

```
pcDNA GFP10-WT ATXN3                1101 TCTAGAATGGAGTCCATCTTCCACGAGAAACAAGAAGGCTCACTTTGTGC
Primer Forward pcDNA GFP10-WT ATXN3 136  TCTAGAATGGAGTCCATCTTCCACGAGAAACAAGAAGGCTCACTTTGTGC
Primer Reverse pcDNA GFP10-WT ATXN3 357  TCTAGAATGGAGTCCATCTTCCACGAGAAACAAGAAGGCTCACTTTGTGC

pcDNA GFP10-WT ATXN3                1151 TCAACATTGCCTGAATAACTTATTGCAAGGAGAATATTTTAGCCCTGTGG
Primer Forward pcDNA GFP10-WT ATXN3 186  TCAACATTGCCTGAATAACTTATTGCAAGGAGAATATTTTAGCCCTGTGG
Primer Reverse pcDNA GFP10-WT ATXN3 407  TCAACATTGCCTGAATAACTTATTGCAAGGAGAATATTTTAGCCCTGTGG

pcDNA GFP10-WT ATXN3                1201 AATTATCCTCAATTGCACATCAGCTGGATGAGGAGGAGAGGATGAGAATG
Primer Forward pcDNA GFP10-WT ATXN3 236  AATTATCCTCAATTGCACATCAGCTGGATGAGGAGGAGAGGATGAGAATG
Primer Reverse pcDNA GFP10-WT ATXN3 457  AATTATCCTCAATTGCACATCAGCTGGATGAGGAGGAGAGGATGAGAATG

pcDNA GFP10-WT ATXN3                1251 GCAGAAGGAGGAGTTACTAGTGAAGATTATCGCACGTTTTTACAGCAGCC
Primer Forward pcDNA GFP10-WT ATXN3 286  GCAGAAGGAGGAGTTACTAGTGAAGATTATCGCACGTTTTTACAGCAGCC
Primer Reverse pcDNA GFP10-WT ATXN3 507  GCAGAAGGAGGAGTTACTAGTGAAGATTATCGCACGTTTTTACAGCAGCC

pcDNA GFP10-WT ATXN3                1301 TTCTGGAAATATGGATGACAGTGGTTTTTTCTCTATTCAGGTTATAAGCA
Primer Forward pcDNA GFP10-WT ATXN3 336  TTCTGGAAATATGGATGACAGTGGTTTTTTCTCTATTCAGGTTATAAGCA
Primer Reverse pcDNA GFP10-WT ATXN3 557  TTCTGGAAATATGGATGACAGTGGTTTTTTCTCTATTCAGGTTATAAGCA

pcDNA GFP10-WT ATXN3                1351 ATGCCTTGAAAGTTTGGGGTTTAGAACTAATCCTGTTCAACAGTCCAGAG
Primer Forward pcDNA GFP10-WT ATXN3 386  ATGCCTTGAAAGTTTGGGGTTTAGAACTAATCCTGTTCAACAGTCCAGAG
Primer Reverse pcDNA GFP10-WT ATXN3 607  ATGCCTTGAAAGTTTGGGGTTTAGAACTAATCCTGTTCAACAGTCCAGAG

pcDNA GFP10-WT ATXN3                1401 TATCAGAGGCTCAGGATCGATCCTATAAATGAAAGATCATTTATATGCAA
Primer Forward pcDNA GFP10-WT ATXN3 436  TATCAGAGGCTCAGGATCGATCCTATAAATGAAAGATCATTTATATGCAA
Primer Reverse pcDNA GFP10-WT ATXN3 657  TATCAGAGGCTCAGGATCGATCCTATAAATGAAAGATCATTTATATGCAA

pcDNA GFP10-WT ATXN3                1451 TTATAAGGAACACTGGTTTACAGTTAGAAAATTAGGAAAACAGTGGTTTA
Primer Forward pcDNA GFP10-WT ATXN3 486  TTATAAGGAACACTGGTTTACAGTTAGAAAATTAGGAAAACAGTGGTTTA
Primer Reverse pcDNA GFP10-WT ATXN3 707  TTATAAGGAACACTGGTTTACAGTTAGAAAATTAGGAAAACAGTGGTTTA

pcDNA GFP10-WT ATXN3                1501 ACTTGAATTCTCTCTTGACGGGTCCAGAATTAATATCAGATACATATCTT
Primer Forward pcDNA GFP10-WT ATXN3 536  ACTTGAATTCTCTCTTGACGGGTCCAGAATTAATATCAGATACATATCTT
Primer Reverse pcDNA GFP10-WT ATXN3 757  ACTTGAATTCTCTCTTGACGGGTCCAGAATTAATATCAGATACATATCTT

pcDNA GFP10-WT ATXN3                1551 GCACTTTTCTTGGCTCAATTACAACAGGAAGGTTATTCTATATTTGTCGT
Primer Forward pcDNA GFP10-WT ATXN3 586  GCACTTTTCTTGGCTCAATTACAACAGGAAGGTTATTCTATATTTGTCGT
Primer Reverse pcDNA GFP10-WT ATXN3 807  GCACTTTTCTTGGCTCAATTACAACAGGAAGGTTATTCTATATTTGTCGT

pcDNA GFP10-WT ATXN3                1601 TAAGGGTGATCTGCCAGATTGCGAAGCTGACCAACTCCTACAGATGATTA
Primer Forward pcDNA GFP10-WT ATXN3 636  TAAGGGTGATCTGCCAGATTGCGAAGCTGACCAACTCCTACAGATGATTA
Primer Reverse pcDNA GFP10-WT ATXN3 857  TAAGGGTGATCTGCCAGATTGCGAAGCTGACCAACTCCTACAGATGATTA
```

```
pcDNA GFP10-WT ATXN3                1651 GGGTCCAACAGATGCATCGACCAAAACTTATTGGAGAAGAATTAGCACAA
Primer Forward pcDNA GFP10-WT ATXN3 686  GGGTCCAACAGATGCATCGACCAAAACTTATTGGAGAAGAATTAGCACAA
Primer Reverse pcDNA GFP10-WT ATXN3 907  GGGTCCAACAGATGCATCGACCAAAACTTATTGGAGAAGAATTAGCACAA

pcDNA GFP10-WT ATXN3                1701 CTAAAAGAGCAAAGAGTCCATAAAACAGACCTGGAACGAGTGTTAGAAGC
Primer Forward pcDNA GFP10-WT ATXN3 736  CTAAAAGAGCAAAGAGTCCATAAAACAGACCTGGAACGAGTGTTAGAAGC
Primer Reverse pcDNA GFP10-WT ATXN3 957  CTAAAAGAGCAAAGAGTCCATAAAACAGACCTGGAACGAGTGTTAGAAGC

pcDNA GFP10-WT ATXN3                1751 AAATGATGGCTCAGGAATGTTAGACGAAGATGAGGAGGATTTGCAGAGGG
Primer Forward pcDNA GFP10-WT ATXN3 786  AAATGATGGCTCAGGAATGTTAGACGAAGATGAGGAGGATTTGCAGAGGG
Primer Reverse pcDNA GFP10-WT ATXN3 1007 AAATGATGGCTCAGGAATGTTAGACGAAGATGAGGAGGATTTGCAGAGGG

pcDNA GFP10-WT ATXN3                1801 CTCTGGCACTAAGTCGCCAAGAAATTGACATGGAAGATGAGGAAGCAGAT
Primer Forward pcDNA GFP10-WT ATXN3 836  CTCTGGCACTAAGTCGCCAAGAAATTGACATGGAAGATGAGGAAGCAGAT
Primer Reverse pcDNA GFP10-WT ATXN3 1057 CTCTGGCACTAAGTCGCCAAGAAATTGACATGGAAGATGAGGAAGCAGAT

pcDNA GFP10-WT ATXN3                1851 CTCCGCAGGGCTATTCAGCTAAGTATGCAAGGTAGTTCCAGAAACATATC
Primer Forward pcDNA GFP10-WT ATXN3 886  CTCCGCAGGGCTATTCAGCTAAGTATGCAAGGTAGTTCCAGAAACATATC
Primer Reverse pcDNA GFP10-WT ATXN3 1107 CTCCGCAGGGCTATTCAGCTAAGTATGCAAGGTAGTTCCAGAAACATATC

pcDNA GFP10-WT ATXN3                1901 TCAAGATATGACACAGACATCAGGTACAAATCTTACTTCAGAAGAGCTTC
Primer Forward pcDNA GFP10-WT ATXN3 936  TCAAGATATGACACAGACATCAGGTACAAATCTTACTTCAGAAGAGCTTC
Primer Reverse pcDNA GFP10-WT ATXN3 1157 TCAAGATATGACACAGACATCAGGTACAAATCTTACTTCAGAAGAGCTTC

pcDNA GFP10-WT ATXN3                1951 GGAAGAGACGAGAAGCCTACTTTGAAAAACAGCAGCAAAAGCAGCAACAG
Primer Forward pcDNA GFP10-WT ATXN3 986  GGAAGAGACGAGAAGCCTACTTTGAAAAACAGCAGCAAA-----------
Primer Reverse pcDNA GFP10-WT ATXN3 1207 GGAAGAGACGAGAAGCCTACTTTGAAAAACAGCAGCAAAAGCAGCAACAG

pcDNA GFP10-WT ATXN3                2001 CAGCAGCAGCAGCAGCAGGGGGACCTATCAGGACAGAGTTCACATCC
Primer Forward pcDNA GFP10-WT ATXN3 1024 ---------------------------------------------------
Primer Reverse pcDNA GFP10-WT ATXN3 1257 CAGCAGCAGCAGCAGCAGGGGGACCTATCAGGACAGAGTTCACATCC

pcDNA GFP10-WT ATXN3                2051 ATGTGAAAGGCCAGCCACCAGTTCAGGAGCACTTGGGAGTGATCTAGGTG
Primer Forward pcDNA GFP10-WT ATXN3 1024 ---------------------------------------------------
Primer Reverse pcDNA GFP10-WT ATXN3 1307 ATGTGAAAGGCCAGCCACCAGTTCAGGAGCACTTGGGAGTGATCTAGGTG

pcDNA GFP10-WT ATXN3                2101 ATGCTATGAGTGAAGAAGACATGCTTCAGGCAGCTGTGACCATGTCTTTA
Primer Forward pcDNA GFP10-WT ATXN3 1024 ---------------------------------------------------
Primer Reverse pcDNA GFP10-WT ATXN3 1357 ATGCTATGAGTGAAGAAGACATGCTTCAGGCAGCTGTGACCATGTCTTTA

pcDNA GFP10-WT ATXN3                2151 GAAACTGTCAGAAATGATTTGAAAACAGAAGGAAAAAAA
Primer Forward pcDNA GFP10-WT ATXN3 1024 ------------------------------------
Primer Reverse pcDNA GFP10-WT ATXN3 1407 GAAACTGTCAGAAATGATTTGAAAACAGAAGGAAAAAAA
```

Figure 52 - Sequence alignment of the pcDNA GFP10-WT ATXN3 vector with primers detailed in table 2, obtained from SnapGene® software (from Insightful Science; available at snapgene.com).

pcDNA WT ATXN3-GFP11 vector alignment



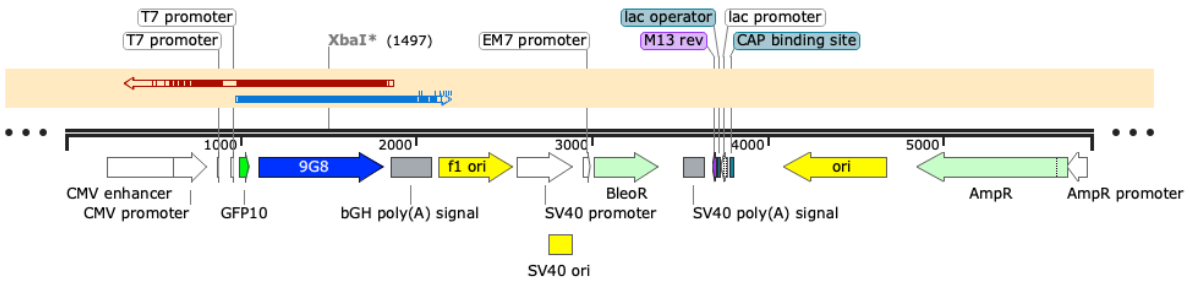

```
pcDNA WT ATXN3-GFP11          969  ATGGAGTCCATCTTCCACGAGAAACAAGAAGG
Primer Forward WT ATXN3-GFP11 78   ATGGAGTCCATCTTCCACGAGAAACAAGAAGG
Primer Reverse WT ATXN3-GFP11 0    --------------------------------

pcDNA WT ATXN3-GFP11          1001 CTCACTTTGTGCTCAACATTGCCTGAATAACTTATTGCAAGGAGAATATT
Primer Forward WT ATXN3-GFP11 110  CTCACTTTGTGCTCAACATTGCCTGAATAACTTATTGCAAGGAGAATATT
Primer Reverse WT ATXN3-GFP11 0    -------------------------------------------------

pcDNA WT ATXN3-GFP11          1051 TTAGCCCTGTGGAATTATCCTCAATTGCACATCAGCTGGATGAGGAGGAG
Primer Forward WT ATXN3-GFP11 160  TTAGCCCTGTGGAATTATCCTCAATTGCACATCAGCTGGATGAGGAGGAG
Primer Reverse WT ATXN3-GFP11 0    -------------------------------------------------

pcDNA WT ATXN3-GFP11          1101 AGGATGAGAATGGCAGAAGGAGGAGTTACTAGTGAAGATTATCGCACGTT
Primer Forward WT ATXN3-GFP11 210  AGGATGAGAATGGCAGAAGGAGGAGTTACTAGTGAAGATTATCGCACGTT
Primer Reverse WT ATXN3-GFP11 1    -----------------------------------ATTATCGGCAGGT

pcDNA WT ATXN3-GFP11          1151 TTTACAGCAGCCTTCTGGAAATATGGATGACAGTGGTTTTTTCTCTATTC
Primer Forward WT ATXN3-GFP11 260  TTTACAGCAGCCTTCTGGAAATATGGATGACAGTGGTTTTTTCTCTATTC
Primer Reverse WT ATXN3-GFP11 14   TTTACAGCAGCCTTCCGGAAATATGGATGACAGTGGTTTTTTCTCTATTC

pcDNA WT ATXN3-GFP11          1201 AGGTTATAAGCAATGCCTTGAAAGTTTGGGGTTTAGAACTAATCCTGTTC
Primer Forward WT ATXN3-GFP11 310  AGGTTATAAGCAATGCCTTGAAAGTTTGGGGTTTAGAACTAATCCTGTTC
Primer Reverse WT ATXN3-GFP11 64   AGGTTATAAGCAATGCCTTGAAAGTTTGGGGTTTAGAACTAATCCTGTTC

pcDNA WT ATXN3-GFP11          1251 AACAGTCCAGAGTATCAGAGGCTCAGGATCGATCCTATAAATGAAAGATC
Primer Forward WT ATXN3-GFP11 360  AACAGTCCAGAGTATCAGAGGCTCAGGATCGATCCTATAAATGAAAGATC
Primer Reverse WT ATXN3-GFP11 114  AACAGTCCAGAGTATCAGAGGCTCAGGATCGATCCTATAAATGAAAGATC
```

```
pcDNA WT ATXN3-GFP11          1301 ATTTATATGCAATTATAAGGAACACTGGTTTACAGTTAGAAAATTAGGAA
Primer Forward WT ATXN3-GFP11 410  ATTTATATGCAATTATAAGGAACACTGGTTTACAGTTAGAAAATTAGGAA
Primer Reverse WT ATXN3-GFP11 164  ATTTATATGCAATTATAAGGAACACTGGTTTACAGTTAGAAAATTAGGAA

pcDNA WT ATXN3-GFP11          1351 AACAGTGGTTTAACTTGAATTCTCTCTTGACGGGTCCAGAATTAATATCA
Primer Forward WT ATXN3-GFP11 460  AACAGTGGTTTAACTTGAATTCTCTCTTGACGGGTCCAGAATTAATATCA
Primer Reverse WT ATXN3-GFP11 214  AACAGTGGTTTAACTTGAATTCTCTCTTGACGGGTCCAGAATTAATATCA

pcDNA WT ATXN3-GFP11          1401 GATACATATCTTGCACTTTTCTTGGCTCAATTACAACAGGAAGGTTATTC
Primer Forward WT ATXN3-GFP11 510  GATACATATCTTGCACTTTTCTTGGCTCAATTACAACAGGAAGGTTATTC
Primer Reverse WT ATXN3-GFP11 264  GATACATATCTTGCACTTTTCTTGGCTCAATTACAACAGGAAGGTTATTC

pcDNA WT ATXN3-GFP11          1451 TATATTTGTCGTTAAGGGTGATCTGCCAGATTGCGAAGCTGACCAACTCC
Primer Forward WT ATXN3-GFP11 560  TATATTTGTCGTTAAGGGTGATCTGCCAGATTGCGAAGCTGACCAACTCC
Primer Reverse WT ATXN3-GFP11 314  TATATTTGTCGTTAAGGGTGATCTGCCAGATTGCGAAGCTGACCAACTCC

pcDNA WT ATXN3-GFP11          1501 TACAGATGATTAGGGTCCAACAGATGCATCGACCAAAACTTATTGGAGAA
Primer Forward WT ATXN3-GFP11 610  TACAGATGATTAGGGTCCAACAGATGCATCGACCAAAACTTATTGGAGAA
Primer Reverse WT ATXN3-GFP11 364  TACAGATGATTAGGGTCCAACAGATGCATCGACCAAAACTTATTGGAGAA

pcDNA WT ATXN3-GFP11          1551 GAATTAGCACAACTAAAAGAGCAAAGAGTCCATAAAACAGACCTGGAACG
Primer Forward WT ATXN3-GFP11 660  GAATTAGCACAACTAAAAGAGCAAAGAGTCCATAAAACAGACCTGGAACG
Primer Reverse WT ATXN3-GFP11 414  GAATTAGCACAACTAAAAGAGCAAAGAGTCCATAAAACAGACCTGGAACG

pcDNA WT ATXN3-GFP11          1601 AGTGTTAGAAGCAAATGATGGCTCAGGAATGTTAGACGAAGATGAGGAGG
Primer Forward WT ATXN3-GFP11 710  AGTGTTAGAAGCAAATGATGGCTCAGGAATGTTAGACGAAGATGAGGAGG
Primer Reverse WT ATXN3-GFP11 464  AGTGTTAGAAGCAAATGATGGCTCAGGAATGTTAGACGAAGATGAGGAGG

pcDNA WT ATXN3-GFP11          1651 ATTTGCAGAGGGCTCTGGCACTAAGTCGCCAAGAAATTGACATGGAAGAT
Primer Forward WT ATXN3-GFP11 760  ATTTGCAGAGGGCTCTGGCACTAAGTCGCCAAGAAATTGACATGGAAGAT
Primer Reverse WT ATXN3-GFP11 514  ATTTGCAGAGGGCTCTGGCACTAAGTCGCCAAGAAATTGACATGGAAGAT

pcDNA WT ATXN3-GFP11          1701 GAGGAAGCAGATCTCCGCAGGGCTATTCAGCTAAGTATGCAAGGTAGTTC
Primer Forward WT ATXN3-GFP11 810  GAGGAAGCAGATCTCCGCAGGGCTATTCAGCTAAGTATGCAAGGTAGTTC
Primer Reverse WT ATXN3-GFP11 564  GAGGAAGCAGATCTCCGCAGGGCTATTCAGCTAAGTATGCAAGGTAGTTC

pcDNA WT ATXN3-GFP11          1751 CAGAAACATATCTCAAGATATGACACAGACATCAGGTACAAATCTTACTT
Primer Forward WT ATXN3-GFP11 860  CAGAAACATATCTCAAGATATGACACAGACATC-NNNACAAATCTTACTT
Primer Reverse WT ATXN3-GFP11 614  CAGAAACATATCTCAAGATATGACACAGACATCAGGTACAAATCTTACTT

pcDNA WT ATXN3-GFP11          1801 CAGAAGAGCTTCGGAAGAGACGAGAAGCCTACTTTGAAAAACAGCAGCAA
Primer Forward WT ATXN3-GFP11 909  CAGAAGAGCTTCGGAAGAGACGAGAAGCCTACTTTGAAAAACAGCAGC-A
Primer Reverse WT ATXN3-GFP11 664  CAGAAGAGCTTCGGAAGAGACGAGAAGCCTACTTTGAAAAACAGCAGCAA
```

```
pcDNA WT ATXN3-GFP11           1851 AAGCAGCAACAGCAGCAGCAGCAGCAGCAGGGGGGACCTATCAGGACA
Primer Forward WT ATXN3-GFP11 958  AAGCAGCAACAGCAGCAGCAGCAGCAGC-NGGGGACCTATCAGGACN
Primer Reverse WT ATXN3-GFP11 714  AAGCAGCAACAGCAGCAGCAGCAGCAGCAGGGGGGACCTATCAGGACA

pcDNA WT ATXN3-GFP11           1901 GAGTT-CACATCCATGTGAAAGGCCAGCCACCAGTTCAGGAG-CACTTGG
Primer Forward WT ATXN3-GFP11 1007 GAGTTNNNCATCCATGTGAAAGGCCAGCCNCCNGTTCAGNAGCCACTTGG
Primer Reverse WT ATXN3-GFP11 764  GAGTT-CACATCCATGTGAAAGGCCAGCCACCAGTTCAGGAG-CACTTGG

pcDNA WT ATXN3-GFP11           1949 GAGTGATCTAGGTGATGCTATGAGTGAAGAAGACATGCTTCAGGCAGCTG
Primer Forward WT ATXN3-GFP11 1057 GANTGATCTAGGTGATGCTATNANTGAAGAAGANNNGCTTCAGGCAGCTG
Primer Reverse WT ATXN3-GFP11 812  GAGTGATCTAGGTGATGCTATGAGTGAAGAAGACATGCTTCAGGCAGCTG

pcDNA WT ATXN3-GFP11           1999 TGACCATG-TCTTTAGAAACTG-TCAGAAATGATTTGAAAACAGAAGGAA
Primer Forward WT ATXN3-GFP11 1107 TGACCATGTTCTTTNNNAACTGNNCAGAAATGATTNGAAANNNAAGGAAA
Primer Reverse WT ATXN3-GFP11 862  TGACCATG-TCTTTAGAAACTG-TCAGAAATGATTTGAAAACAGAAGGAA

pcDNA WT ATXN3-GFP11           2047 AAAAAGATGGT-GGCGGTGGCTCTGGAGGTGGTGGGTCCTCCG-GAGAGA
Primer Forward WT ATXN3-GFP11 1157 AAAANAANGGTGGGCGGNGGCNNNNGNNNNGNGGNCCCTNCGNNANANN
Primer Reverse WT ATXN3-GFP11 910  AAAAAGATGGT-GGCGGTGGCTCTGGAGGTGGTGGGTCCTCCG-GAGAGA

pcDNA WT ATXN3-GFP11           2095 AGCGCGACCACATG
Primer Forward WT ATXN3-GFP11 1207 NNNNGNACNCNNTG
Primer Reverse WT ATXN3-GFP11 958  AGCGCGACCACATG
```

Figure 53 - Sequence alignment of the pcDNA WT ATXN3-GFP11 vector with primers detailed in table 2, obtained from SnapGene® software (from Insightful Science; available at snapgene.com).

pcDNA GFP10 – 9G8 vector alignment



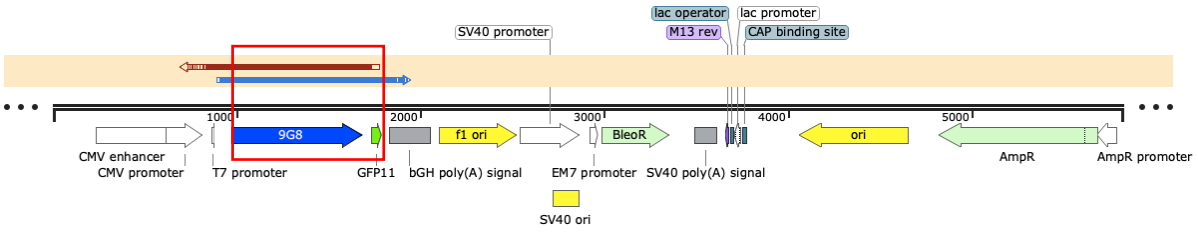

```
pcDNA GFP10-9G8           987  ATGGGCGACCTGCC
Primer Forward GFP10-9G8 18   NTGGGCGACCTGCC
Primer Reverse GFP10-9G8 560  ATGGGCGACCTGCC
```

```
pcDNA GFP10-9G8          1001 CGACGACCACTACCTGTCCACCCAGACCATCCTGAGCAAGGACCTGAACA
Primer Forward GFP10-9G8 32   CGACGACCACTACCTGTCCACCCAGACCATCCTGAGCAAGGACCTGAACA
Primer Reverse GFP10-9G8 574  CGACGACCACTACCTGTCCACCCAGACCATCCTGAGCAAGGACCTGAACA

pcDNA GFP10-9G8          1051 TCGATGGTGGCGGTGGCTCTGGAGGTGGTGGGTCCTCCGGCGGCGGTGGA
Primer Forward GFP10-9G8 82   TCGATGGTGGCGGTGGCTCTGGAGGTGGTGGGTCCTCCGGCGGCGGTGGA
Primer Reverse GFP10-9G8 624  TCGATGGTGGCGGTGGCTCTGGAGGTGGTGGGTCCTCCGGCGGCGGTGGA

pcDNA GFP10-9G8          1101 TCTATGTCGCGTTACGGGCGGTACGGAGGAGAAACCAAGGTGTATGTTGG
Primer Forward GFP10-9G8 132  TCTATGTCGCGTTACGGGCGGTACGGAGGAGAAACCAAGGTGTATGTTGG
Primer Reverse GFP10-9G8 674  TCTATGTCGCGTTACGGGCGGTACGGAGGAGAAACCAAGGTGTATGTTGG

pcDNA GFP10-9G8          1151 TAACCTGGGAACTGGCGCTGGCAAAGGAGAGTTAGAAAGGGCTTTCAGTT
Primer Forward GFP10-9G8 182  TAACCTGGGAACTGGCGCTGGCAAAGGAGAGTTAGAAAGGGCTTTCAGTT
Primer Reverse GFP10-9G8 724  TAACCTGGGAACTGGCGCTGGCAAAGGAGAGTTAGAAAGGGCTTTCAGTT

pcDNA GFP10-9G8          1201 ATTATGGTCCTTTAAGAACTGTATGGATTGCGAGAAATCCTCCAGGATTT
Primer Forward GFP10-9G8 232  ATTATGGTCCTTTAAGAACTGTATGGATTGCGAGAAATCCTCCAGGATTT
Primer Reverse GFP10-9G8 774  ATTATGGTCCTTTAAGAACTGTATGGATTGCGAGAAATCCTCCAGGATTT

pcDNA GFP10-9G8          1251 GCCTTTGTGGAATTCGAAGATCCTAGAGATGCAGAAGATGCAGTACGAGG
Primer Forward GFP10-9G8 282  GCCTTTGTGGAATTCGAAGATCCTAGAGATGCAGAAGATGCAGTACGAGG
Primer Reverse GFP10-9G8 824  GCCTTTGTGGAATTCGAAGATCCTAGAGATGCAGAAGATGCAGTACGAGG

pcDNA GFP10-9G8          1301 ACTGGATGGAAAGGTGATTTGTGGCTCCCGAGTGAGGGTTGAACTATCGA
Primer Forward GFP10-9G8 332  ACTGGATGGAAAGGTGATTTGTGGCTCCCGAGTGAGGGTTGAACTATCGA
Primer Reverse GFP10-9G8 874  ACTGGATGGAAAGGTGATTTGTGGCTCCCGAGTGAGGGTTGAACTATCGA

pcDNA GFP10-9G8          1351 CAGGCATGCCTCGGAGATCACGTTTTGATAGACCACCTGCCCGACGTCCC
Primer Forward GFP10-9G8 382  CAGGCATGCCTCGGAGATCACGTTTTGATAGACCACCTGCCCGACGTCCC
Primer Reverse GFP10-9G8 924  CAGGCATGCCTCGGAGATCACGTTTTGATAGACCACCTGCCCGACGTCCC

pcDNA GFP10-9G8          1401 TTTGATCCAAATGATAGATGCTATGAGTGTGGCGAAAAGGGACATTATGC
Primer Forward GFP10-9G8 432  TTTGATCCAAATGATAGATGCTATGAGTGTGGCGAAAAGGGACATTATGC
Primer Reverse GFP10-9G8 974  TTTGATCCAAATGATAGATGCTATGAGTGTGGCGAAAAGGGACATTATGC

pcDNA GFP10-9G8          1451 TTATGATTGTCATCGTTACAGCCGGCGAAGAAGAAGCAGGTCACGGTCTA
Primer Forward GFP10-9G8 482  TTATGATTGTCATCGTTACAGCCGGCGAAGAAGAAGCAGGTCACGGTCTA
Primer Reverse GFP10-9G8 1024 TTATGATTGTCATCGTTACAGCCGGCGAAGAAGAAGCAGGTCACGGTCTA

pcDNA GFP10-9G8          1501 GATCACATTCTCGATCCAGAGGAAGGCGATACTCTCGCTCACGCAGCAGG
Primer Forward GFP10-9G8 532  GATCACATTCTCGATCCAGAGGAAGGCGATACTCTCGCTCACGCAGCAGG
Primer Reverse GFP10-9G8 1074 GATCACATTCTCGATCCAGAGGAAGGCGATACTCTCGCTCACGCAGCAGG
```

```
pcDNA GFP10-9G8          1551 AGCAGGGGACGAAGGTCAAGGTCAGCATCTCCTCGACGATCAAGATCTAT
Primer Forward GFP10-9G8 582  AGCAGGGGACGAAGGTCAAGGTCAGCATCTCCTCGACGATCAAGATCTAT
Primer Reverse GFP10-9G8 1124 AGCAGGGGACGAAGGTCAAGGTCAGCATCTCCTCGACGATCAAGATCTAT

pcDNA GFP10-9G8          1601 CTCTCTTCGTAGATCAAGATCAGCTTCACTCAGAAGATCTAGGTCTGGTT
Primer Forward GFP10-9G8 632  CTCTCTTCGTAGATCAAGATCAGCTTCACTCAGAAGATCTAGGTCTGGTT
Primer Reverse GFP10-9G8 1174 CTCTCTTCGTAGATCAAGATCAGCTTCACTCAGAAGATCTAGGTCTGGTT

pcDNA GFP10-9G8          1651 CTATAAAAGGATCGAGGTATTTCCAATCCCCGTCGAGGTCAAGATCAAGA
Primer Forward GFP10-9G8 682  CTATAAAAGGATCGAGGTATTTCCAATCCCCGTCGAGGTCAAGATCAAGA
Primer Reverse GFP10-9G8 1224 CTATAAAAGGATCGAGGTATTTCCAATCCCCGTCGAGGTCAAGATCAAGA

pcDNA GFP10-9G8          1701 TCCAGGTCTATTTCACGACCAAGAAGCAGCCGATCAAAGTCCAGATCTCC
Primer Forward GFP10-9G8 732  TCCAGGTCTATTTCACGACCAAGAAGCAGCCGATCAAAGTCCAGATCTCC
Primer Reverse GFP10-9G8 1274 TCCAGGTCTATTTCACGACCAAGAAGCAGCCGATCAAAGTCCAGATCTCC

pcDNA GFP10-9G8          1751 ATCTCCAAAAAGAAGTCGTTCCCCATCAGGAAGTCCTCGCAGAAGTGCAA
Primer Forward GFP10-9G8 782  ATCTCCAAAAAGAAGTCGTTCCCCATCAGGAAGTCCTCGCAGAAGTGCAA
Primer Reverse GFP10-9G8 1324 ATCTCCAAAAAGAAGTCGTTCCCCATCAGGAAGTCCTCGCAGAAGTGCAA

pcDNA GFP10-9G8          1801 GTCCTGAAAGAATGGACTGA
Primer Forward GFP10-9G8 832  GTCCTGAAAGAATGGACTGA
Primer Reverse GFP10-9G8 1374 GTCCTGAAAGAATGGACTGA
```

Figure 54 - Sequence alignment of the pcDNA GFP10-9G8 vector with primers detailed in table 2, obtained from SnapGene® software (from Insightful Science; available at snapgene.com).

pcDNA 9G8 – GFP11 vector alignment



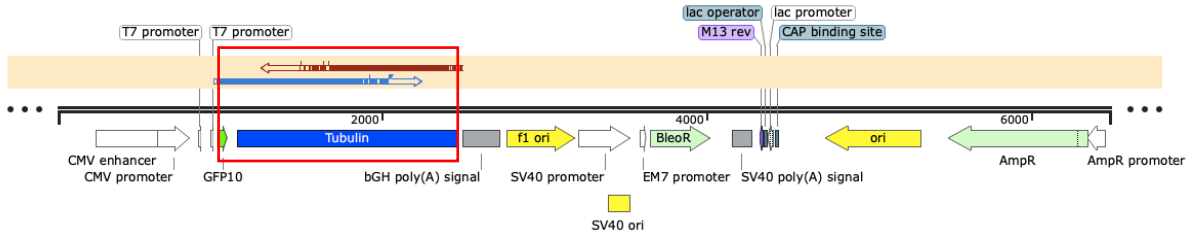

```
pcDNA 9G8-GFP11          972  ATGTCGCGTTACGGGCGGTACGGAGGAGA
Primer Forward 9G8-GFP11 81   ATGTCGCGTTACGGGCGGTACGGAGGAGA
Primer Reverse 9G8-GFP11 266  ATGTCGCGTTACGGGCGGTACGGAGGAGA

pcDNA 9G8-GFP11          1001 AACCAAGGTGTATGTTGGTAACCTGGGAACTGGCGCTGGCAAAGGAGAGT
Primer Forward 9G8-GFP11 110  AACCAAGGTGTATGTTGGTAACCTGGGAACTGGCGCTGGCAAAGGAGAGT
Primer Reverse 9G8-GFP11 295  AACCAAGGTGTATGTTGGTAACCTGGGAACTGGCGCTGGCAAAGGAGAGT
```

```
pcDNA 9G8-GFP11          1051 TAGAAAGGGCTTTCAGTTATTATGGTCCTTTAAGAACTGTATGGATTGCG
Primer Forward 9G8-GFP11 160  TAGAAAGGGCTTTCAGTTATTATGGTCCTTTAAGAACTGTATGGATTGCG
Primer Reverse 9G8-GFP11 345  TAGAAAGGGCTTTCAGTTATTATGGTCCTTTAAGAACTGTATGGATTGCG

pcDNA 9G8-GFP11          1101 AGAAATCCTCCAGGATTTGCCTTTGTGGAATTCGAAGATCCTAGAGATGC
Primer Forward 9G8-GFP11 210  AGAAATCCTCCAGGATTTGCCTTTGTGGAATTCGAAGATCCTAGAGATGC
Primer Reverse 9G8-GFP11 395  AGAAATCCTCCAGGATTTGCCTTTGTGGAATTCGAAGATCCTAGAGATGC

pcDNA 9G8-GFP11          1151 AGAAGATGCAGTACGAGGACTGGATGGAAAGGTGATTTGTGGCTCCCGAG
Primer Forward 9G8-GFP11 260  AGAAGATGCAGTACGAGGACTGGATGGAAAGGTGATTTGTGGCTCCCGAG
Primer Reverse 9G8-GFP11 445  AGAAGATGCAGTACGAGGACTGGATGGAAAGGTGATTTGTGGCTCCCGAG

pcDNA 9G8-GFP11          1201 TGAGGGTTGAACTATCGACAGGCATGCCTCGGAGATCACGTTTTGATAGA
Primer Forward 9G8-GFP11 310  TGAGGGTTGAACTATCGACAGGCATGCCTCGGAGATCACGTTTTGATAGA
Primer Reverse 9G8-GFP11 495  TGAGGGTTGAACTATCGACAGGCATGCCTCGGAGATCACGTTTTGATAGA

pcDNA 9G8-GFP11          1251 CCACCTGCCCGACGTCCCTTTGATCCAAATGATAGATGCTATGAGTGTGG
Primer Forward 9G8-GFP11 360  CCACCTGCCCGACGTCCCTTTGATCCAAATGATAGATGCTATGAGTGTGG
Primer Reverse 9G8-GFP11 545  CCACCTGCCCGACGTCCCTTTGATCCAAATGATAGATGCTATGAGTGTGG

pcDNA 9G8-GFP11          1301 CGAAAAGGGACATTATGCTTATGATTGTCATCGTTACAGCCGGCGAAGAA
Primer Forward 9G8-GFP11 410  CGAAAAGGGACATTATGCTTATGATTGTCATCGTTACAGCCGGCGAAGAA
Primer Reverse 9G8-GFP11 595  CGAAAAGGGACATTATGCTTATGATTGTCATCGTTACAGCCGGCGAAGAA

pcDNA 9G8-GFP11          1351 GAAGCAGGTCACGGTCTAGATCACATTCTCGATCCAGAGGAAGGCGATAC
Primer Forward 9G8-GFP11 460  GAAGCAGGTCACGGTCTAGATCACATTCTCGATCCAGAGGAAGGCGATAC
Primer Reverse 9G8-GFP11 645  GAAGCAGGTCACGGTCTAGATCACATTCTCGATCCAGAGGAAGGCGATAC

pcDNA 9G8-GFP11          1401 TCTCGCTCACGCAGCAGGAGCAGGGGACGAAGGTCAAGGTCAGCATCTCC
Primer Forward 9G8-GFP11 510  TCTCGCTCACGCAGCAGGAGCAGGGGACGAAGGTCAAGGTCAGCATCTCC
Primer Reverse 9G8-GFP11 695  TCTCGCTCACGCAGCAGGAGCAGGGGACGAAGGTCAAGGTCAGCATCTCC

pcDNA 9G8-GFP11          1451 TCGACGATCAAGATCTATCTCTCTTCGTAGATCAAGATCAGCTTCACTCA
Primer Forward 9G8-GFP11 560  TCGACGATCAAGATCTATCTCTCTTCGTAGATCAAGATCAGCTTCACTCA
Primer Reverse 9G8-GFP11 745  TCGACGATCAAGATCTATCTCTCTTCGTAGATCAAGATCAGCTTCACTCA

pcDNA 9G8-GFP11          1501 GAAGATCTAGGTCTGGTTCTATAAAAGGATCGAGGTATTTCCAATCCCCG
Primer Forward 9G8-GFP11 610  GAAGATCTAGGTCTGGTTCTATAAAAGGATCGAGGTATTTCCAATCCCCG
Primer Reverse 9G8-GFP11 795  GAAGATCTAGGTCTGGTTCTATAAAAGGATCGAGGTATTTCCAATCCCCG

pcDNA 9G8-GFP11          1551 TCGAGGTCAAGATCAAGATCCAGGTCTATTTCACGACCAAGAAGCAGCCG
Primer Forward 9G8-GFP11 660  TCGAGGTCAAGATCAAGATCCAGGTCTATTTCACGACCAAGAAGCAGCCG
Primer Reverse 9G8-GFP11 845  TCGAGGTCAAGATCAAGATCCAGGTCTATTTCACGACCAAGAAGCAGCCG
```

```
pcDNA 9G8-GFP11          1601 ATCAAAGTCCAGATCTCCATCTCCAAAAAGAAGTCGTTCCCCATCAGGAA
Primer Forward 9G8-GFP11 710  ATCAAAGTCCAGATCTCCATCTCCAAAAAGAAGTCGTTCCCCATCAGGAA
Primer Reverse 9G8-GFP11 895  ATCAAAGTCCAGATCTCCATCTCCAAAAAGAAGTCGTTCCCCATCAGGAA

pcDNA 9G8-GFP11          1651 GTCCTCGCAGAAGTGCAAGTCCTGAAAGAATGGACGATGGTGGCGGTGGC
Primer Forward 9G8-GFP11 760  GTCCTCGCAGAAGTGCAAGTCCTGAAAGAATGGACGATGGTGGCGGTGGC
Primer Reverse 9G8-GFP11 945  GTCCTCGCAGAAGTGCAAGTCCTGAAAGAATGGACGATGGTGGCGGTGGC

pcDNA 9G8-GFP11          1701 TCTGGAGGTGGTGGGTCCTCCGGAGAGAAGCGCGACCACATGGTGCTGCT
Primer Forward 9G8-GFP11 810  TCTGGAGGTGGTGGGTCCTCCGGAGAGAAGCGCGACCACATGGTGCTGCT
Primer Reverse 9G8-GFP11 995  TCTGGAGGTGGTGGGTCCTCCGGAGAGAAGCGCGACCACA-NGTGNTGNN

pcDNA 9G8-GFP11          1751 GGAGTACGTGACCGCCGCCGGCATCACCGACGCCTCCTAA
Primer Forward 9G8-GFP11 860  GGAGTACGTGACCGCCGCCGGCATCACCGACGCCTCCTAA
Primer Reverse 9G8-GFP11 1044 NNNGNNNNN------------------------------
```

Figure 55 - Sequence alignment of the pcDNA 9G8-GFP11 vector with primers detailed in table 2, obtained from SnapGene® software (from Insightful Science; available at snapgene.com).

pcDNA GFP10 – Tubulin vector alignment




```
pcDNA GFP10-Tubulin          1107 ATGCGTGAGTGCATCTCCATCCACGTTGGCCAGGCTGGTGTCCA
Primer Forward GFP10-Tubulin 143  ATGCGTGAGTGCATCTCCATCCACGTTGGCCAGGCTGGTGTCCA
Primer Reverse GFP10-Tubulin 39   ----------------------------------------GTTAA

pcDNA GFP10-Tubulin          1151 GATTGGCAATGCCTGCTGGGAGCTCTACTGCCTGGAACA--CGGCATCCA
Primer Forward GFP10-Tubulin 187  GATTGGCAATGCCTGCTGGGAGCTCTACTGCCTGGAACA--CGGCATCCA
Primer Reverse GFP10-Tubulin 45   AAAAANNAANACCTGGGGGGGATNTTTTTTTNGGGAAAACCGGNNTTCC

pcDNA GFP10-Tubulin          1199 GCCC----GATGGCCAGATGCCAAGTGACAAGACCATT--------GGGG
Primer Forward GFP10-Tubulin 235  GCCC----GATGGCCAGATGCCAAGTGACAAGACCATT--------GGGG
Primer Reverse GFP10-Tubulin 95   CCCCNNNGGGGGCCNGAAGCCAAAAGGAAAAAANNCCCCGGGGGGGGGG
```

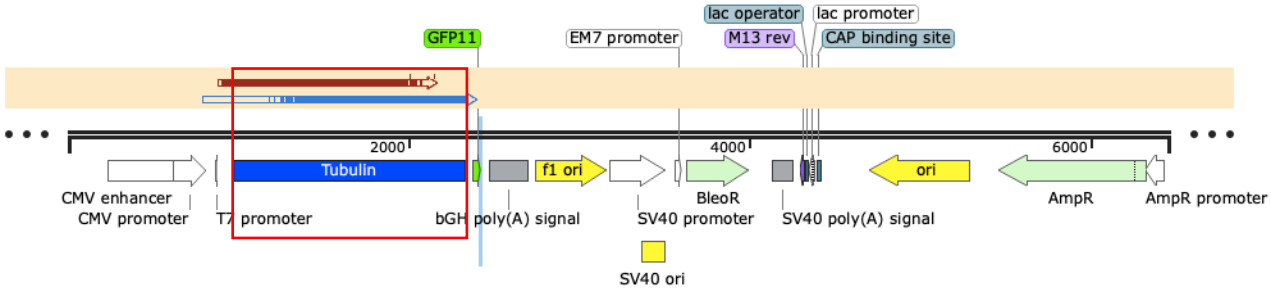
```
pcDNA GFP10-Tubulin          1237 GAGGAGATGACTCCTTCAACACCTTCTTCAGTGAGACGGGCGCTGG----
Primer Forward GFP10-Tubulin 273  GAGGAGATGACTCCTTCAACACCTTCTTCAGTGAGACGGGCGCTGG----
Primer Reverse GFP10-Tubulin 145  GGGGANATTTNTTTTTTAAACCNNTTTNTTTAAGGGGGGGGGGGGGGG

pcDNA GFP10-Tubulin          1282 --CAAGCACGTGCCCCGGGCTGTGTTTGT----AGACTTGGAACCC--AC
Primer Forward GFP10-Tubulin 318  --CAAGCACGTGCCCCGGGCTGTGTTTGT----AGACTTGGAACCC--AC
Primer Reverse GFP10-Tubulin 195  GAAAACCCCGGGCCCCCGGGGGNGTTTTTTTTAAGATTGGGAACCCCACC

pcDNA GFP10-Tubulin          1325 AGTC-ATTGATGAAG-TTCGCACTGGCACCTACC-GCCAG-CTCTTCCAC
Primer Forward GFP10-Tubulin 361  AGTC-ATTGATGAAG-TTCGCACTGGCACCTACC-GCCAG-CTCTTCCAC
Primer Reverse GFP10-Tubulin 245  ATTCAATTGAGGAAANTTCGCACTGGCCCTACCGGCCAGTTTTTTCCAC

pcDNA GFP10-Tubulin          1371 CCTG-AGCAGCTCATCACAGGCAAGGAAGATGCTG-CCAATAACTATGCC
Primer Forward GFP10-Tubulin 407  CCTG-AGCAGCTCATCACAGGCAAGGAAGATGCTG-CCAATAACTATGCC
Primer Reverse GFP10-Tubulin 295  CCTGAAGCAGTTCATCCCNGGCAAGGAAGATGCTGCCCAATAACTATGCC

pcDNA GFP10-Tubulin          1419 CGAGGGCACTACA-CCATT-GGCAAGGAGATCATTGACCTTGTGTTGGAC
Primer Forward GFP10-Tubulin 455  CGAGGGCACTACA-CCATT-GGCAAGGAGATCATTGACCTTGTGTTGGAC
Primer Reverse GFP10-Tubulin 345  CGAGGGCATTACACCCATTGGGCAAGGAGATCATTGACCTTGTGTTGGAC

pcDNA GFP10-Tubulin          1467 CGAATTCGCAAGCTGGCTGACCAGTGCACCGGTCTTCAGGGCTTCTTGGT
Primer Forward GFP10-Tubulin 503  CGAATTCGCAAGCTGGCTGACCAGTGCACCGGTCTTCAGGGCTTCTTGGT
Primer Reverse GFP10-Tubulin 395  CGAATTCGCAAGCTGGCTGACCAGTGCACCGGTCTTCAGGGCTTCTTGGT

pcDNA GFP10-Tubulin          1517 TTTCCACAGCTTTGGTGGGGGAACTGGTTCTGGGTTCACCTCCCT-GCTC
Primer Forward GFP10-Tubulin 553  TTTCCACAGCTTTGGTGGGGGAACTGGTTCTGGGTTCACCTCCCT-GCTC
Primer Reverse GFP10-Tubulin 445  TTTCCACAGCTTTGGTGGGGGAACTGGTTCTGGGTTCACCTCCCTGGTTC

pcDNA GFP10-Tubulin          1566 ATGGAACGTC-TCTCAGTTGATTATGGCAAGAAGT-CCAAGCTGGAGTTC
Primer Forward GFP10-Tubulin 602  ATGGAACGTC-TCTCAGTTGATTATGGCAAGAAGT-CCAAGCTGGAGTTC
Primer Reverse GFP10-Tubulin 495  ATGGAACGTCTTCTCAGTTGATTATGGCAAGAAGTCCCAAGCTGGAGTTC

pcDNA GFP10-Tubulin          1614 TCCATTTA-CCCAGCACCCC--AGGTTTCCACAGCTGTAGTTGAGCCCTA
Primer Forward GFP10-Tubulin 650  TCCATTTA-CCCAGCACCCC--AGGTTTCCACAGCTGTAGTTGAGCCCTA
Primer Reverse GFP10-Tubulin 545  TCCATTTACCCCAGCACCCCCAGGTTTTCCACAGCTGTAGTTGAGCCCTT

pcDNA GFP10-Tubulin          1661 CAA--CTCCATCCTCACCACCCACACCACCCTGGAGCACTCTGATTGTGC
Primer Forward GFP10-Tubulin 697  CAA--CTCCATCCTCACCACCCACACCACCCTGGAGCACTCTGATTGTGC
Primer Reverse GFP10-Tubulin 595  CAAATTTCCTTTCTCACCACCCACACCACCCTGGAGCACTCTGATTGTGC

pcDNA GFP10-Tubulin          1709 CTTCATGGTAGACAATGAGGCCATCTATGACATCTGTCGTAGAAACCTCG
Primer Forward GFP10-Tubulin 745  CTTCATGGTAGACAATGAGGCCATCTATGACATCTGTCGTAGAAACCTCG
Primer Reverse GFP10-Tubulin 645  CTTCATGGTAGACAATGAGGCCATCTATGACATCTGTCGTAGAAACCTCG
```

```
pcDNA GFP10-Tubulin          1759 ATATCGAGCGCCCAACCTACACTAACCTTAACCGCCTTATTAGCCAGATT
Primer Forward GFP10-Tubulin 795  ATATCGAGCGCCCAACCTACACTAACCTTAACCGCCTTATTAGCCAGATT
Primer Reverse GFP10-Tubulin 695  ATATCGAGCGCCCAACCTACACTAACCTTAACCGCCTTATTAGCCAGATT

pcDNA GFP10-Tubulin          1809 GTGTCCTCCATCACTGCTTCCCTGAGATTTGATGGAGCCCTGAATGTTGA
Primer Forward GFP10-Tubulin 845  GTGTCCTCCATCACTGCTTCCCTGAGATTTGATGGAGCCCTGAATGTTGA
Primer Reverse GFP10-Tubulin 745  GTGTCCTCCATCACTGCTTCCCTGAGATTTGATGGAGCCCTGAATGTTGA

pcDNA GFP10-Tubulin          1859 CCTGACAGAATTCCAGACCAACCTGGTGCCCTACCCCCGCATCCACTTCC
Primer Forward GFP10-Tubulin 895  CCTGACAGAATTCCAGACCAACCTGGTGCCCTACCCCCGCATCCACTTCC
Primer Reverse GFP10-Tubulin 795  CCTGACAGAATTCCAGACCAACCTGGTGCCCTACCCCCGCATCCACTTCC

pcDNA GFP10-Tubulin          1909 CTCTGGCCACATATGCCCCTGTCATCTCTGCTGAGAAAGCCTACCATGAA
Primer Forward GFP10-Tubulin 945  CTCTGGCCACATATGCCCCTGTCATCTCTGCTGAGAAAGCCTACCATGAA
Primer Reverse GFP10-Tubulin 845  CTCTGGCCACATATGCCCCTGTCATCTCTGCTGAGAAAGCCTACCATGAA

pcDNA GFP10-Tubulin          1959 CAGCTTTCTGTAGCAGAGATCACCAATGCTTGCTTTGAGCCAGCCAACCA
Primer Forward GFP10-Tubulin 995  CAGCTTTCTGTAGCAAAGATCACCAATGCTTGCTTTGAGCCAGCCAACCA
Primer Reverse GFP10-Tubulin 895  CAGCTTTCTGTAGCAGAGATCACCAATGCTTGCTTTGAGCCAGCCAACCA

pcDNA GFP10-Tubulin          2009 GATGGTGAAATGTGACCCTCGCCATGGTAAATACATGGCTTGCTGCCTGT
Primer Forward GFP10-Tubulin 1045 GATGGTGAAATGTGACCCTCGCCATGGTAAATACATGGCTTGCTGCCTGT
Primer Reverse GFP10-Tubulin 945  GATGGTGAAATGTGACCCTCGCCATGGTAAATACATGGCTTGCTGCCTGT

pcDNA GFP10-Tubulin          2059 TGTACC-GTGGTGACGTGGTTCCCAAAGATGTCAATGCTGCCATTGCCAC
Primer Forward GFP10-Tubulin 1095 TGTACCGGGGGTGACGTGGTTCCCAAAAATGTCAATGCTGCCATTGCCAC
Primer Reverse GFP10-Tubulin 995  TGTACC-GTGGTGACGTGGTTCCCAAAGATGTCAATGCTGCCATTGCCAC

pcDNA GFP10-Tubulin          2108 CATCAAAACCAAGCGCAGCATCCAGTTTGTGG-ATTGGTGCCCCACTGGC
Primer Forward GFP10-Tubulin 1145 CATCAAAACCAAGCGCAGCATCCAGTTTGTGGAATGGGGGCCCCCCTGGC
Primer Reverse GFP10-Tubulin 1044 CATCAAAACCAAGCGCAGCATCCAGTTTGTGG-ATTGGTGCCCCACTGGC

pcDNA GFP10-Tubulin          2157 TTCAA-GGTTGGCATCAACTACCAGCCT-CCCACTGTGGTGCCTGGTGG-
Primer Forward GFP10-Tubulin 1195 TTCAAGGGTGGGAATCAACTACCAACCTCCCCACTGGGGGCCTGGGGGA
Primer Reverse GFP10-Tubulin 1093 TTCAA-GGTTGGCATCAACTACCAGCCT-CCCACTGTGGTGCCTGGTGG-

pcDNA GFP10-Tubulin          2204 AGACCTGGCCAAGGTACAGAGAGCTGTGTGCATGCTGAGCAACACCACAG
Primer Forward GFP10-Tubulin 1245 AAACCTGGCCAGGGAAAAAAAANTTGGTGGTNTTNTTAAAAACCCNACC
Primer Reverse GFP10-Tubulin 1140 AGACCTGGCCAAGGTACAGAGAGCTGTGTGCATGCTGAGCAACACCACAG

pcDNA GFP10-Tubulin          2254 CCATTGCTGAGGC-CTGGGCTCGCCTGGACCACAAGTTTGACCTGATGTA
Primer Forward GFP10-Tubulin 1295 CCTTTTTTAAGGCGGGGGGCCCCCNNAACCAAAATTT------------
Primer Reverse GFP10-Tubulin 1190 CCATTGCTGAGGC-CTGGGCTCGCCTGGACCACAAGTTTGACCTGATGTA
```

```
pcDNA GFP10-Tubulin           2303 TGCCAAGCGTGCCTTTGTTCACTGGTACGTGGGTGAGGGGATGGAGGAAG
Primer Forward GFP10-Tubulin 1332 ------------------------------------------------
Primer Reverse GFP10-Tubulin 1239 TGCCAAGCGTGCCTTTGTTCACTGGTACGTGGGTGAGGGGATGGAGGAAG

pcDNA GFP10-Tubulin           2353 GCGAGTTTTCAGAGGCCCGTGAAGATATGGCTGCCCTTGAGAAGGATTAT
Primer Forward GFP10-Tubulin 1332 ------------------------------------------------
Primer Reverse GFP10-Tubulin 1289 GCGAGTTTTCAGAGGCCCGTGAAGATATGGCTGCCCTTGAGAAGGATTAT

pcDNA GFP10-Tubulin           2403 GAGGAGGTTGGTGTGGATTCTGTTGAAGGAGAGGGTGAGGAAGAAGGAGA
Primer Forward GFP10-Tubulin 1332 -------------------------------------------------
Primer Reverse GFP10-Tubulin 1339 GAGGAGGTTGGTGCGGATTCTGTCGAAGGAGAGGGTGAGGAAGAAGGAGA

pcDNA GFP10-Tubulin           2453 GGAATAC
Primer Forward GFP10-Tubulin 1332 -------
Primer Reverse GFP10-Tubulin 1389 GGAATNC
```

Figure 56 - Sequence alignment of the pcDNA GFP10-Tubulin vector with primers detailed in table 2, obtained from SnapGene® software (from Insightful Science; available at snapgene.com). Red – Unconfirmed sequenced nucleotides.

pcDNA Tubulin - GFP11 vector alignment




```
pcDNA Tubulin-GFP11           972  ATGCGTGAGTGCA
Primer Forward Tubulin-GFP11 89   ATGCGTGAGTGCA
Primer Reverse Tubulin-GFP11 170  ------GGGTGAG

pcDNA Tubulin-GFP11           985  TCTCCATCCACGTTGGCCAGGCTGGTGTCCAGATTGGCAATGCCTGCTGG
Primer Forward Tubulin-GFP11 102  TCTCCATCCACGTTGGCCAGGCTGGTGTCCAGATTGGCAATGCCTGCTGG
Primer Reverse Tubulin-GFP11 212  GTTTCNNNCCCGTGGGCAGG---GGGGTCCAAA-TGGCAA-GCCTGT--G

pcDNA Tubulin-GFP11           1035 GAGCTCTACTGCCTGGAACACGGCATCCAGCCCGATGGCCAGATGCCAAG
Primer Forward Tubulin-GFP11 152  GAGCTCTACTGCCTGGAACACGGCATCCAGCCCGATGGCCAGATGCCAAG
Primer Reverse Tubulin-GFP11 255  GAATTTTA-TGCT--GAAC-CNGC-TCCAGCC---------GANGCCAAT
```

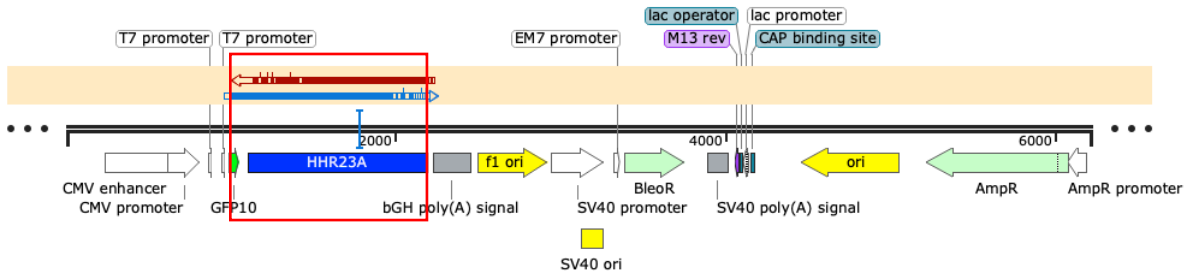
```
pcDNA Tubulin-GFP11          1085 T---GACAAGACCATTGGGGGAGGAGATGACTCCTTCAACACCTTCTTCA
Primer Forward Tubulin-GFP11 202  T---GACAAGACCATTGGGGGAGGAGATGACTCCTTCAACACCTTCTTCA
Primer Reverse Tubulin-GFP11 291  CCAAGNCAAGACCTT---GGGGAGNAATGA-TCCTTTANC-CCTTTTTCA

pcDNA Tubulin-GFP11          1132 GTGAGACGGGCGCTGGCAAGCACGTGCCCCGGGCTGTGTTTGTAGACTTG
Primer Forward Tubulin-GFP11 249  GTGAGACGGGCGCTGGCAAGCACGTGCCCCGGGCTGTGTTTGTAGACTTG
Primer Reverse Tubulin-GFP11 336  GT--ANCGGGCG-TGGCAAGCA-GTG-CCCGGGCTG-GTTTG-AAACTTG

pcDNA Tubulin-GFP11          1182 GAACCCACAGTCATTGATGAAGTTCGCACTGGCACCTACCGCCAGCTCTT
Primer Forward Tubulin-GFP11 299  GAACCCACAGTCATTGATGAAGTTCGCACTGGCACCTACCGCCAGCTCTT
Primer Reverse Tubulin-GFP11 379  GAACCCCCAGTCA-TGAT-AAGTTCGC-CTGGCACCTA-CGCCAGTTTTT

pcDNA Tubulin-GFP11          1232 CCACCCTGAGCAGCTCATCACAGGCAAGGAAGATGCTGCCAATAACTATG
Primer Forward Tubulin-GFP11 349  CCACCCTGAGCAGCTCATCACAGGCAAGGAAGATGCTGCCAATAACTATG
Primer Reverse Tubulin-GFP11 425  CCA-CCTGAGCAGTTCATCACAGGCAAGGAAGATGCTGCCAAT-ACTATG

pcDNA Tubulin-GFP11          1282 CCCGAGGGCACTACACCATTGGCAAGGAGATCATTGACCTTGTGTTGGAC
Primer Forward Tubulin-GFP11 399  CCCGAGGGCACTACACCATTGGCAAGGAGATCATTGACCTTGTGTTGGAC
Primer Reverse Tubulin-GFP11 473  CCCGAGGGCACTACACCATTGGCAAGGAGATCATTGACCTTGTG-TGGAC

pcDNA Tubulin-GFP11          1332 CGAATTCGCAAGCTGGCTGACCAGTGCACCGGTCTTCAGGGCTTCTTGGT
Primer Forward Tubulin-GFP11 449  CGAATTCGCAAGCTGGCTGACCAGTGCACCGGTCTTCAGGGCTTCTTGGT
Primer Reverse Tubulin-GFP11 522  CGAATTCGCAAGCTGGCTGACCAGTGCACCGGTCTTCAGGGCTTCTTGGT

pcDNA Tubulin-GFP11          1382 TTTCCACAGCTTTGGTGGGGGAACTGGTTCTGGGTTCACCTCCCTGCTCA
Primer Forward Tubulin-GFP11 499  TTTCCACAGCTTTGGTGGGGGAACTGGTTCTGGGTTCACCTCCCTGCTCA
Primer Reverse Tubulin-GFP11 572  TTTCCACAGCTTTGGTGGGGGAACTGGTTCTGGGTTCACCTCCCTGCTCA

pcDNA Tubulin-GFP11          1432 TGGAACGTCTCTCAGTTGATTATGGCAAGAAGTCCAAGCTGGAGTTCTCC
Primer Forward Tubulin-GFP11 549  TGGAACGTCTCTCAGTTGATTATGGCAAGAAGTCCAAGCTGGAGTTCTCC
Primer Reverse Tubulin-GFP11 622  TGGAACGTCTCTCAGTTGATTATGGCAAGAAGTCCAAGCTGGAGTTCTCC

pcDNA Tubulin-GFP11          1482 ATTTACCCAGCACCCCAGGTTTCCACAGCTGTAGTTGAGCCCTACAACTC
Primer Forward Tubulin-GFP11 599  ATTTACCCAGCACCCCAGGTTTCCACAGCTGTAGTTGAGCCCTACAACTC
Primer Reverse Tubulin-GFP11 672  ATTTACCCAGCACCCCAGGTTTCCACAGCTGTAGTTGAGCCCTACAACTC

pcDNA Tubulin-GFP11          1532 CATCCTCACCACCCACACCACCCTGGAGCACTCTGATTGTGCCTTCATGG
Primer Forward Tubulin-GFP11 649  CATCCTCACCACCCACACCACCCTGGAGCACTCTGATTGTGCCTTCATGG
Primer Reverse Tubulin-GFP11 722  CATCCTCACCACCCACACCACCCTGGAGCACTCTGATTGTGCCTTCATGG

pcDNA Tubulin-GFP11          1582 TAGACAATGAGGCCATCTATGACATCTGTCGTAGAAACCTCGATATCGAG
Primer Forward Tubulin-GFP11 699  TAGACAATGAGGCCATCTATGACATCTGTCGTAGAAACCTCGATATCGAG
Primer Reverse Tubulin-GFP11 772  TAGACAATGAGGCCATCTATGACATCTGTCGTAGAAACCTCGATATCGAG
```

```
pcDNA Tubulin-GFP11          1632 CGCCCAACCTACACTAACCTTAACCGCCTTATTAGCCAGATTGTGTCCTC
Primer Forward Tubulin-GFP11 749  CGCCCAACCTACACTAACCTTAACCGCCTTATTAGCCAGATTGTGTCCTC
Primer Reverse Tubulin-GFP11 822  CGCCCAACCTACACTAACCTTAACCGCCTTATTAGCCAGATTGTGTCCTC

pcDNA Tubulin-GFP11          1682 CATCACTGCTTCCCTGAGATTTGATGGAGCCCTGAATGTTGACCTGACAG
Primer Forward Tubulin-GFP11 799  CATCACTGCTTCCCTGAGATTTGATGGAGCCCTGAATGTTGACCTGACAG
Primer Reverse Tubulin-GFP11 872  CATCACTGCTTCCCTGAGATTTGATGGAGCCCTGAATGTTGACCTGACAG

pcDNA Tubulin-GFP11          1732 AATTCCAGACCAACCTGGTGCCCTACCCCCGCATCCACTTCCCTCTGGCC
Primer Forward Tubulin-GFP11 849  AATTCCAGACCAACCTGGTGCCCTACCCCCGCATCCACTTCCCTCTGGCC
Primer Reverse Tubulin-GFP11 922  AATTCCAGACCAACCTGGTGCCCTACCCCCGCATCCACTTCCCTCTGGCC

pcDNA Tubulin-GFP11          1782 ACATATGCCCCTGTCATCTCTGCTGAGAAAGCCTACCATGAACAGCTTTC
Primer Forward Tubulin-GFP11 899  ACATATGCCCCTGTCATCTCTGCTGAGAAAGCCTACCATGAACAGCTTTC
Primer Reverse Tubulin-GFP11 972  ACATATGCCCCTGTCATCTCTGCTGAGAAAGCCTACCATGAACAGCTTTC

pcDNA Tubulin-GFP11          1832 TGTAGCAGAGATCACCAATGCTTGCTTTGAGCCAGCCAACCAGATGGTGA
Primer Forward Tubulin-GFP11 949  TGTAGCAGAGATCACCAATGCTTGCTTTGAGCCAGCCAACCAGATGGTGA
Primer Reverse Tubulin-GFP11 1022 TGTAGCAGAGATCACCAATGCTTGCTTTGAGCCAGCCAACCAGATGGTGA

pcDNA Tubulin-GFP11          1882 AATGTGACCCTCGCCATGGTAAATACATGGCTTGCTGCCTGTTGTACCGT
Primer Forward Tubulin-GFP11 999  AATGTGACCCTCGCCATGGTAAATACATGGCTTGCTGCCTGTTGTACCGT
Primer Reverse Tubulin-GFP11 1072 AATGTGACCCTCGCCATGGTAAATACATGGCTTGCTGCCTGTTGTACCGT

pcDNA Tubulin-GFP11          1932 GGTGACGTGGTTCCCAAAGATGTCAATGCTGCCATTGCCACCATCAAAAC
Primer Forward Tubulin-GFP11 1049 GGTGACGTGGTTCCCAAAGATGTCAATGCTGCCATTGCCACCATCAAAAC
Primer Reverse Tubulin-GFP11 1122 GGTGACGTGGTTCCCAAAGATGTCAATGCTGCCATTGCCACCATCAAAAC

pcDNA Tubulin-GFP11          1982 CAAGCGCAGCATCCAGTTT-GTGGATTGGTGCCCCACTGGCTTCAAGGTT
Primer Forward Tubulin-GFP11 1099 CAAGCGCAGCATCCAGTTTGGGGGATTGGTGCCCCACTGGCTTCAAGGTT
Primer Reverse Tubulin-GFP11 1172 CAAGCGCAGCATCCAGTTT-GTGGATTGGTGCCCCACTGGCTTCAAGGTT

pcDNA Tubulin-GFP11          2031 GGCATCAACTACCAGCCTCCCACTGTGGTGCCTGGTGGAGACCTGGCCAA
Primer Forward Tubulin-GFP11 1149 GGCATCAACTACCAGCCT-CCACTGTGGGGCCTGGTGGAGACCTGGCCAA
Primer Reverse Tubulin-GFP11 1221 GGCATCAACTACCAGCCTCCCACTGTGGTGCCTGGTGGAGACCTGGCCAA

pcDNA Tubulin-GFP11          2081 GGTACAGAGAGCTGTGTGCATGCTGAGCAACACCACAGCCATTGCTG-AG
Primer Forward Tubulin-GFP11 1198 GGTACAAAAAACTTGGTGCATGCTGAACAACCCCA-AGCCTTTGCTGAAG
Primer Reverse Tubulin-GFP11 1271 GGTACAGAGAGCTGTGTGCATGCTGAGCAACACCACAGCCATTGCTG-AG

pcDNA Tubulin-GFP11          2130 GCCTGGGCTCGCCTGGACCACAAG--TTTGACCTGATGTATGCCAAGCGT
Primer Forward Tubulin-GFP11 1247 GCTGGGGTTNGCCTGGGACCCNAGATTTGACCTGATTTTTGCCAAGGG-
Primer Reverse Tubulin-GFP11 1320 GCCTGGGCTCGCCTGGACCACAAG--TTTGACCTGATGTATGCCAAGCGT
```

```
pcDNA Tubulin-GFP11           2178 GCCTTTGTTCACTGGTACGTGGGTGAGGGGATGGAGGAAGGCGAGTTTTC
Primer Forward Tubulin-GFP11 1295 ------------------------------------------------
Primer Reverse Tubulin-GFP11 1368 GCCTTTGTTCACTGGTACGTGGGTGAGGGGATGGAGGAAGGCGAGTTTTC

pcDNA Tubulin-GFP11           2228 AGAGGCCCGTGAAGATATGGCTGCCCTTGAGAAGGATTATGAGGAGGTTG
Primer Forward Tubulin-GFP11 1295 ------------------------------------------------
Primer Reverse Tubulin-GFP11 1418 AGAGGCCCGTGAAGATATGGCTGCCCTTGAGAAGGATTATGAGGAGGTTG

pcDNA Tubulin-GFP11           2278 GTGTGGATTCTGTTGAAGGAGAGGGTGAGGAAGAAGGAGAGGAATACGAT
Primer Forward Tubulin-GFP11 1295 ------------------------------------------------
Primer Reverse Tubulin-GFP11 1468 GTGTGGATTCTGTTGAAGGAGAGGGTGAGGAAGAAGGAGAGGAATACGAT

pcDNA Tubulin-GFP11           2328 GGTGGCGGTGGCTCTGGAGGTGGTGGGT
Primer Forward Tubulin-GFP11 1295 ---------------------------
Primer Reverse Tubulin-GFP11 1518 GGTGGCGGTGGCTCTGGAGGTGGTGGGT
```

Figure 57 - Sequence alignment of the pcDNA Tubulin-GFP11 vector with primers detailed in table 2, obtained from SnapGene® software (from Insightful Science; available at snapgene.com).

pcDNA GFP10 – hHR23A vector alignment




```
pcDNA GFP10-hHR23A           987  ATGGGCGACCTGCC
Primer Forward GFP10-hHR23A  25   NTGGGCGACCTGCC
Primer Reverse GFP10-hHR23A  0    -------------

pcDNA GFP10-hHR23A           1001 CGACGACCACTACCTGT--CCACCCAGACCATCCTGAGC--AAGGACCTG
Primer Forward GFP10-hHR23A  39   CGACGACCACTACCTGT--CCACCCAGACCATCCTGAGC--AAGGACCTG
Primer Reverse GFP10-hHR23A  1    ----------TCCCTTTTCCCCCCCAACCCCTTCTNGACAAAAGGCCCTT

pcDNA GFP10-hHR23A           1047 AACATCGATGGTGGCGGTGGCTCTGGAGGTGGTGGGTCCTCCGGCGGCGG
Primer Forward GFP10-hHR23A  85   AACATCGATGGTGGCGGTGGCTCTGGAGGTGGTGGGTCCTCCGGCGGCGG
Primer Reverse GFP10-hHR23A  41   AAAATTGGGGGGGG-GGGGGGTTTGGGGGGGGGGGGTCCCCCGGGGGGGG
```

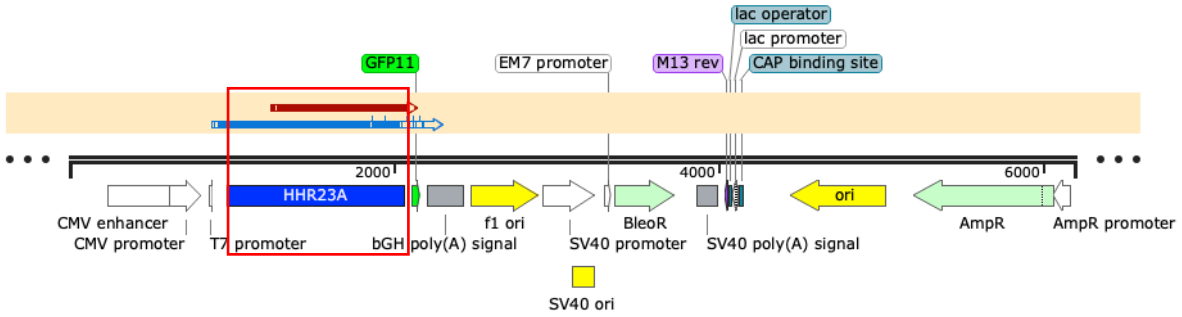
```
pcDNA GFP10-hHR23A          1097 TGGATCTATGGCC--GTCACCATCACGCTC--AAAACGCTGCAGCAGCAG
Primer Forward GFP10-hHR23A 135  TGGATCTATGGCC--GTCACCATCACGCTC--AAAACGCTGCAGCAGCAG
Primer Reverse GFP10-hHR23A 90   GGGTTTTTGGCCCGTCCACCATCCCCCCTCAAAAACGGTGCAGCAGCAG

pcDNA GFP10-hHR23A          1143 ACCTTC--AAGATCC-GCATGGAGCC-TGACGAGAC--GGTGAAGGTGCT
Primer Forward GFP10-hHR23A 181  ACCTTC--AAGATCC-GCATGGAGCC-TGACGAGAC--GGTGAAGGTGCT
Primer Reverse GFP10-hHR23A 140  ACCTTTCAAAGATCCGGCATGGAGCCTTGACGAAGCCGGGTAAAGGTGCT

pcDNA GFP10-hHR23A          1187 AAAGGAGAAGATAGAAGCTGAGAAGGG-TCGTGATGCC-TTCCCCGTGGC
Primer Forward GFP10-hHR23A 225  AAAGGAGAAGATAGAAGCTGAGAAGGG-TCGTGATGCC-TTCCCCGTGGC
Primer Reverse GFP10-hHR23A 190  AAAGGAGAAGATAGAAGCTGAGAAGGGTTCGTGATGCCTTTCCCCGGGGG

pcDNA GFP10-hHR23A          1235 TGGACAGAAACTCA-TCTATGCCGGCAAGATCTTGAGTGACGATGTCCCT
Primer Forward GFP10-hHR23A 273  TGGACAGAAACTCA-TCTATGCCGGCAAGATCTTGAGTGACGATGTCCCT
Primer Reverse GFP10-hHR23A 240  TGGACAGAAACTCATTCTATGCCGGCAAGATTTTGAGTGACGATGTCCCT

pcDNA GFP10-hHR23A          1284 ATCAGGGACTATCGCATCGATGAGAAGAACTTTGTGGTCGTCATGGTGAC
Primer Forward GFP10-hHR23A 322  ATCAGGGACTATCGCATCGATGAGAAGAACTTTGTGGTCGTCATGGTGAC
Primer Reverse GFP10-hHR23A 290  ATCAGGGACTATCGCATCGATGAGAAGAACTTTGTGGTCGTCATGGTGAC

pcDNA GFP10-hHR23A          1334 CAAGACCAAAGCCGGCCAGGGTACC-TCAGCACCCCCAGAGGCCTCACCC
Primer Forward GFP10-hHR23A 372  CAAGACCAAAGCCGGCCAGGGTACC-TCAGCACCCCCAGAGGCCTCACCC
Primer Reverse GFP10-hHR23A 340  CAAGACCAAAGCCGGCCAGGGTACCTTCAGCACCCCCAGAGGCCTCACCC

pcDNA GFP10-hHR23A          1383 ACAGCTGCCCCAGAGTCCTCTACATCCTTCCCGCCTGCCCCCACCTCAGG
Primer Forward GFP10-hHR23A 421  ACAGCTGCCCCAGAGTCCTCTACATCCTTCCCGCCTGCCCCCACCTCAGG
Primer Reverse GFP10-hHR23A 390  ACAGCTGCCCCAGAGTCCTCTACATCCTTCCCGCTTGCCCCCACTTCAGG

pcDNA GFP10-hHR23A          1433 CATGTCCCATCCCCCACCTGCCGCCAGAGAGGACAAGAGCCCATCAGAGG
Primer Forward GFP10-hHR23A 471  CATGTCCCATCCCCCACCTGCCGCCAGAGAGGACAAGAGCCCATCAGAGG
Primer Reverse GFP10-hHR23A 440  CATGTCCCATCCCCCACCTGCCGCCAGAGAGGACAAGAGCCCATCAGAGG

pcDNA GFP10-hHR23A          1483 AATCCGCCCCCACGACGTCCCCAGAGTCTGTGTCAGGCTCTGTTCCCTCT
Primer Forward GFP10-hHR23A 521  AATCCGCCCCCACGACGTCCCCAGAGTCTGTGTCAGGCTCTGTTCCCTCT
Primer Reverse GFP10-hHR23A 490  AATCCGCCCCCACGACGTCCCCAGAGTCTGTGTCAGGCTCTGTTCCCTCT

pcDNA GFP10-hHR23A          1533 TCAGGTAGCAGCGGGCGAGAGGAAGACGCGGCCTCCACGCTAGTGACGGG
Primer Forward GFP10-hHR23A 571  TCAGGTAGCAGCGGGCGAGAGGAAGACGCGGCCTCCACGCTAGTGACGGG
Primer Reverse GFP10-hHR23A 540  TCAGGTAGCAGCGGGCGAGAGGAAGACGCGGCCTCCACGCTAGTGACGGG

pcDNA GFP10-hHR23A          1583 CTCTGAGTATGAGACGATGCTGACGGAGATCATGTCCATGGGCTATGAGC
Primer Forward GFP10-hHR23A 621  CTCTGAGTATGAGACGATGCTGACGGAGATCATGTCCATGGGCTATGAGC
Primer Reverse GFP10-hHR23A 590  CTCTGAGTATGAGACGATGCTGACGGAGATCATGTCCATGGGCTATGAGC
```

```
pcDNA GFP10-hHR23A          1633 GAGAGCGGGTCGTGGCCGCCCTGAGAGCCAGCTACAACAACCCCCACCGA
Primer Forward GFP10-hHR23A 671  GAGAGCGGGTCGTGGCCGCCCTGAGAGCCAGCTACAACAACCCCCACCGA
Primer Reverse GFP10-hHR23A 640  GAGAGCGGGTCGTGGCCGCCCTGAGAGCCAGCTACAACAACCCCCACCGA

pcDNA GFP10-hHR23A          1683 GCCGTGGAGTATCTGCTCACGGGAATTCCTGGGAGCCCCGAGCCGGAACA
Primer Forward GFP10-hHR23A 721  GCCGTGGAGTATCTGCTCACGGGAATTCCTGGGAGCCCCGAGCCGGAACA
Primer Reverse GFP10-hHR23A 690  GCCGTGGAGTATCTGCTCACGGGAATTCCTGGGAGCCCCGAGCCGGAACA

pcDNA GFP10-hHR23A          1733 CGGTTCTGTCCAGGAGAGCCAGGTATCGGAGCAGCCGGCCACGGAAGCAG
Primer Forward GFP10-hHR23A 771  CGGTTCTGTCCAGGAGAGCCAGGTATCGGAGCAGCCGGCCACGGAAGCAG
Primer Reverse GFP10-hHR23A 740  CGGTTCTGTCCAGGAGAGCCAGGTATCGGAGCAGCCGGCCACGGAAGCAG

pcDNA GFP10-hHR23A          1783 GAGAGAACCCCCTGGAGTTCCTGCGGGACCAGCCCCAGTTCCAGAACATG
Primer Forward GFP10-hHR23A 821  GAGAGAACCCCCTGGAGTTCCTGCGGGACCAGCCCCAGTTCCAGAACATG
Primer Reverse GFP10-hHR23A 790  GAGAGAACCCCCTGGAGTTCCTGCGGGACCAGCCCCAGTTCCAGAACATG

pcDNA GFP10-hHR23A          1833 CGGCAGGTGATTCAGCAGAACCCTGCGCTGCTGCCCGCCCTGCTCCAGCA
Primer Forward GFP10-hHR23A 871  CGGCAGGTGATTCAGCAGAACCCTGCGCTGCTGCCCGCCCTGCTCCAGCA
Primer Reverse GFP10-hHR23A 840  CGGCAGGTGATTCAGCAGAACCCTGCGCTGCTGCCCGCCCTGCTCCAGCA

pcDNA GFP10-hHR23A          1883 GCTGGGCCAGGAGAACCCTCAGCTTTTACAGCAAATCAGCCGGCACCAGG
Primer Forward GFP10-hHR23A 921  GCTGGGCCAGGAGAACCCTCAGCTTTTACAGCAAATCAGCCGGCACCAGG
Primer Reverse GFP10-hHR23A 890  GCTGGGCCAGGAGAACCCTCAGCTTTTACAGCAAATCAGCCGGCACCAGG

pcDNA GFP10-hHR23A          1933 AGCAGTTCATCCAGATGCTGAACGAGCCCCCTGGGGAGCTGGCGGACATC
Primer Forward GFP10-hHR23A 971  AGCAGTTCATCCAGATGCTGAACGAGCCCCCTGGGGAGCTGGCGGACATC
Primer Reverse GFP10-hHR23A 940  AGCAGTTCATCCAGATGCTGAACGAGCCCCCTGGGGAGCTGGCGGACATC

pcDNA GFP10-hHR23A          1983 TCAGATGTGGAGGGGGAGGTGGGCGCCATAGGAGAGGAGGCCCCGCAGAT
Primer Forward GFP10-hHR23A 1021 TCAAATGTGAAGGGGGAAGTGGGCGCCATAGGAAGGAGGCCCCGCNGAT
Primer Reverse GFP10-hHR23A 990  TCAGATGTGGAGGGGGAGGTGGGCGCCATAGGAGAGGAGGCCCCGCAGAT

pcDNA GFP10-hHR23A          2033 GAACTACATCCAGGT-GACGCCGCAGGAGAAAGAAGCTATAGAGAGGTTG
Primer Forward GFP10-hHR23A 1071 GAACTACATCCAGGTGGACGCCGCAGGAAAAAGAAACTTTAAAANGGTTG
Primer Reverse GFP10-hHR23A 1040 GAACTACATCCAGGT-GACGCCGCAGGAGAAAGAAGCTATAGAGAGGTTG

pcDNA GFP10-hHR23A          2082 AAGGCCCTGGGCTTCCCAGAGAGCCTGGTCATCCAGGCCTATTTCGCGTG
Primer Forward GFP10-hHR23A 1121 AAGGCCCTGGGCTTCCCAGAGAGCCTGGTCTTCCAGGCCTATTTCGCGGG
Primer Reverse GFP10-hHR23A 1089 AAGGCCCTGGGCTTCCCAGAGAGCCTGGTCATCCAGGCCTATTTCGCGTG

pcDNA GFP10-hHR23A          2132 TGAAAAAAATGAGAACTTGGC-TGCCAACTTCCTCCTGAGTCAGAAC-TT
Primer Forward GFP10-hHR23A 1171 TGAAAAAAATGAAAACTTGGCTTGCCACTTTCCTCCTGAATCAAAACTTT
Primer Reverse GFP10-hHR23A 1139 TGAAAAAAATGAGAACTTGGC-TGCCAACTTCCTCCTGAGTCAGAAC-TT
```

```
pcDNA GFP10-hHR23A          2180 TGATGAC-GAGTGACTAGAGGGCCCGTTTAAACCCGC-TGATCAGCCTCG
Primer Forward GFP10-hHR23A 1221 TGATGACAAAGGGACTAAANGGCCGGTTTAAACCCCCTTGAACACCTCC
Primer Reverse GFP10-hHR23A 1187 TGATGAC-GAGTGACTAGAGGGCCCG-TTAAACCCGC-TGATCAGCCTCG
```

Figure 58 - Sequence alignment of the pcDNA GFP10-hHR23A vector with primers detailed in table 2, obtained from SnapGene® software (from Insightful Science; available at snapgene.com). Red – Unconfirmed sequenced nucleotide.

pcDNA hHR23A – GFP11 vector alignment



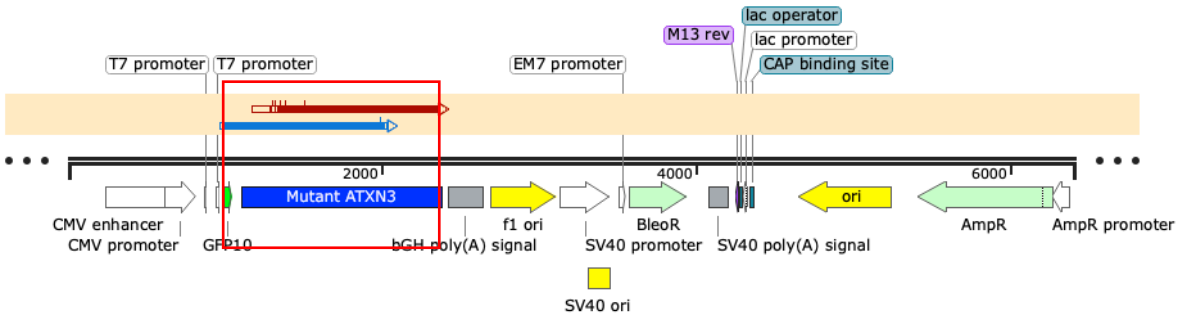

```
pcDNA hHR23A-GFP11          972  ATGGCTATCATGGCCGTCACCATCACGCTCAAAACGCT
Primer Forward hHR23A-GFP11 89   ATGGCTATCATGGCCGTCACCATCACGCTCAAAACGCT
Primer Reverse hHR23A-GFP11 0    -------------------------------------

pcDNA hHR23A-GFP11          1001 GCAGCAGCAGACCTTCAAGATCCGCATGGAGCCTGACGAGACGGTGAAGG
Primer Forward hHR23A-GFP11 118  GCAGCAGCAGACCTTCAAGATCCGCATGGAGCCTGACGAGACGGTGAAGG
Primer Reverse hHR23A-GFP11 0    -------------------------------------------------

pcDNA hHR23A-GFP11          1051 TGCTAAAGGAGAAGATAGAAGCTGAGAAGGGTCGTGATGCCTTCCCCGTG
Primer Forward hHR23A-GFP11 168  TGCTAAAGGAGAAGATAGAAGCTGAGAAGGGTCGTGATGCCTTCCCCGTG
Primer Reverse hHR23A-GFP11 0    -------------------------------------------------

pcDNA hHR23A-GFP11          1101 GCTGGACAGAAACTCATCTATGCCGGCAAGATCTTGAGTGACGATGTCCC
Primer Forward hHR23A-GFP11 218  GCTGGACAGAAACTCATCTATGCCGGCAAGATCTTGAGTGACGATGTCCC
Primer Reverse hHR23A-GFP11 0    -------------------------------------------------

pcDNA hHR23A-GFP11          1151 TATCAGGGACTATCGCATCGATGAGAAGAACTTTGTGGTCGTCATGGTGA
Primer Forward hHR23A-GFP11 268  TATCAGGGACTATCGCATCGATGAGAAGAACTTTGTGGTCGTCATGGTGA
Primer Reverse hHR23A-GFP11 0    -------------------------------------------------

pcDNA hHR23A-GFP11          1201 CCAAGACCAAAGCCGGCCAGGGTACCTCAGCACCCCCAGAGGCCTCACCC
Primer Forward hHR23A-GFP11 318  CCAAGACCAAAGCCGGCCAGGGTACCTCAGCACCCCCAGAGGCCTCACCC
Primer Reverse hHR23A-GFP11 1    ----------------------------------------GGCCTCACCC
```

```
pcDNA hHR23A-GFP11          1251 ACAGCTGCCCCAGAGTCCTCTACATCCTTCCCGCCTGCCCCCACCTCAGG
Primer Forward hHR23A-GFP11 368  ACAGCTGCCCCAGAGTCCTCTACATCCTTCCCGCCTGCCCCCACCTCAGG
Primer Reverse hHR23A-GFP11 11   ACAGCTGCCCCAGAGTCCTCTACATCCTTTCCGCCTGCCCCCACCTCAGG

pcDNA hHR23A-GFP11          1301 CATGTCCCATCCCCCACCTGCCGCCAGAGAGGACAAGAGCCCATCAGAGG
Primer Forward hHR23A-GFP11 418  CATGTCCCATCCCCCACCTGCCGCCAGAGAGGACAAGAGCCCATCAGAGG
Primer Reverse hHR23A-GFP11 61   CATGTCCCATCCCCCACCTGCCGCCAGAGAGGACAAGAGCCCATCAGAGG

pcDNA hHR23A-GFP11          1351 AATCCGCCCCCACGACGTCCCCAGAGTCTGTGTCAGGCTCTGTTCCCTCT
Primer Forward hHR23A-GFP11 468  AATCCGCCCCCACGACGTCCCCAGAGTCTGTGTCAGGCTCTGTTCCCTCT
Primer Reverse hHR23A-GFP11 111  AATCCGCCCCCACGACGTCCCCAGAGTCTGTGTCAGGCTCTGTTCCCTCT

pcDNA hHR23A-GFP11          1401 TCAGGTAGCAGCGGGCGAGAGGAAGACGCGGCCTCCACGCTAGTGACGGG
Primer Forward hHR23A-GFP11 518  TCAGGTAGCAGCGGGCGAGAGGAAGACGCGGCCTCCACGCTAGTGACGGG
Primer Reverse hHR23A-GFP11 161  TCAGGTAGCAGCGGGCGAGAGGAAGACGCGGCCTCCACGCTAGTGACGGG

pcDNA hHR23A-GFP11          1451 CTCTGAGTATGAGACGATGCTGACGGAGATCATGTCCATGGGCTATGAGC
Primer Forward hHR23A-GFP11 568  CTCTGAGTATGAGACGATGCTGACGGAGATCATGTCCATGGGCTATGAGC
Primer Reverse hHR23A-GFP11 211  CTCTGAGTATGAGACGATGCTGACGGAGATCATGTCCATGGGCTATGAGC

pcDNA hHR23A-GFP11          1501 GAGAGCGGGTCGTGGCCGCCCTGAGAGCCAGCTACAACAACCCCCACCGA
Primer Forward hHR23A-GFP11 618  GAGAGCGGGTCGTGGCCGCCCTGAGAGCCAGCTACAACAACCCCCACCGA
Primer Reverse hHR23A-GFP11 261  GAGAGCGGGTCGTGGCCGCCCTGAGAGCCAGCTACAACAACCCCCACCGA

pcDNA hHR23A-GFP11          1551 GCCGTGGAGTATCTGCTCACGGGAATTCCTGGGAGCCCCGAGCCGGAACA
Primer Forward hHR23A-GFP11 668  GCCGTGGAGTATCTGCTCACGGGAATTCCTGGGAGCCCCGAGCCGGAACA
Primer Reverse hHR23A-GFP11 311  GCCGTGGAGTATCTGCTCACGGGAATTCCTGGGAGCCCCGAGCCGGAACA

pcDNA hHR23A-GFP11          1601 CGGTTCTGTCCAGGAGAGCCAGGTATCGGAGCAGCCGGCCACGGAAGCAG
Primer Forward hHR23A-GFP11 718  CGGTTCTGTCCAGGAGAGCCAGGTATCGGAGCAGCCGGCCACGGAAGCAG
Primer Reverse hHR23A-GFP11 361  CGGTTCTGTCCAGGAGAGCCAGGTATCGGAGCAGCCGGCCACGGAAGCAG

pcDNA hHR23A-GFP11          1651 CAGGAGAGAACCCCCTGGAGTTCCTGCGGGACCAGCCCCAGTTCCAGAAC
Primer Forward hHR23A-GFP11 768  CAGGAGAGAACCCCCTGGAGTTCCTGCGGGACCAGCCCCAGTTCCAGAAC
Primer Reverse hHR23A-GFP11 411  CAGGAGAGAACCCCCTGGAGTTCCTGCGGGACCAGCCCCAGTTCCAGAAC

pcDNA hHR23A-GFP11          1701 ATGCGGCAGGTGATTCAGCAGAACCCTGCGCTGCTGCCCGCCCTGCTCCA
Primer Forward hHR23A-GFP11 818  ATGCGGCAGGTGATTCAGCAGAACCCTGCGCTGCTGCCCGCCCTGCTCCA
Primer Reverse hHR23A-GFP11 461  ATGCGGCAGGTGATTCAGCAGAACCCTGCGCTGCTGCCCGCCCTGCTCCA

pcDNA hHR23A-GFP11          1751 GCAGCTGGGCCAGGAGAACCCTCAGCTTTTACAGCAAATCAGCCGGCACC
Primer Forward hHR23A-GFP11 868  GCAGCTGGGCCAGGAGAACCCTCAGCTTTTACAGCAAATCAGCCGGCACC
Primer Reverse hHR23A-GFP11 511  GCAGCTGGGCCAGGAGAACCCTCAGCTTTTACAGCAAATCAGCCGGCACC
```

```
pcDNA hHR23A-GFP11          1801 AGGAGCAGTTCATCCAGATGCTGAACGAGCCCCCTGGGGAGCTGGCGGAC
Primer Forward hHR23A-GFP11 918  AGGAGCAGTTCATCCAGATGCTGAACGAGCCCCCTGGGGAGCTGGCGGAC
Primer Reverse hHR23A-GFP11 561  AGGAGCAGTTCATCCAGATGCTGAACGAGCCCCCTGGGGAGCTGGCGGAC

pcDNA hHR23A-GFP11          1851 ATCTCAGATGTG--GAGGGGGAGGTGGGCGCCATAGGAGAGGAGGCCCCG
Primer Forward hHR23A-GFP11 968  ATCTCAGATGTGAAGGGGGGGAGGTGGGCGCCATAGGAGAGGAGGCCCCG
Primer Reverse hHR23A-GFP11 611  ATCTCAGATGTG--GAGGGGGAGGTGGGCGCCATAGGAGAGGAGGCCCCG

pcDNA hHR23A-GFP11          1899 CAGATGAACTACATCCAGGTGACGCCGCAGGAGAAAGAAGCTATAGAG-A
Primer Forward hHR23A-GFP11 1018 CAGATGAACTACATCCAGGTGACGCCGCAGGAGAAAGAAGCTATAGAGAA
Primer Reverse hHR23A-GFP11 659  CAGATGAACTACATCCAGGTGACGCCGCAGGAGAAAGAAGCTATAGAG-A

pcDNA hHR23A-GFP11          1948 GGTTGAAGGCCCTGGGCTTCCCAGAGAGCCTGGTCATCCAGGCCTATTTC
Primer Forward hHR23A-GFP11 1068 GGTTGAAGGCCCTGGGCTTCCCAGAGAGCCTGGTCATCCAGGCCTATTTC
Primer Reverse hHR23A-GFP11 708  GGTTGAAGGCCCTGGGCTTCCCAGAGAGCCTGGTCATCCAGGCCTATTTC

pcDNA hHR23A-GFP11          1998 GCGTGTGAAAAAAATGAGAACTTGGCTGCCAACTTCCTCCTGAGTCAGAA
Primer Forward hHR23A-GFP11 1118 GCGTGTGAAAAAAATGAGAACTTGGCTGCCAACTTCCTCCTGAGTCAAAA
Primer Reverse hHR23A-GFP11 758  GCGTGTGAAAAAAATGAGAACTTGGCTGCCAACTTCCTCCTGAGTCAGAA

pcDNA hHR23A-GFP11          2048 CTTTGATGACGAGGATGGTGGCGGTGGCT-CTGGAGGTGGTGG
Primer Forward hHR23A-GFP11 1168 CTTTGATGACAAGAATGGTGCCGGTGGCTNNNGGAGGTGGTGG
Primer Reverse hHR23A-GFP11 808  CTTTGATGACGAGGATGGTGGCGGTGGCT-CTGGAGGTGGTGG
```

Figure 59 - Sequence alignment of the pcDNA hHR23A-GFP11 vector with primers detailed in table 2, obtained from SnapGene® software (from Insightful Science; available at snapgene.com).

pcDNA GFP10 – Mutant ATXN3 vector alignment




```
pcDNA GFP10-Mutant ATXN3          987  ATGGGCGACCTGCC
Primer Forward GFP10-Mutant ATXN3 23   ATGGGCGACCTGCC
Primer Reverse GFP10-Mutant ATXN3 0    --------------
```

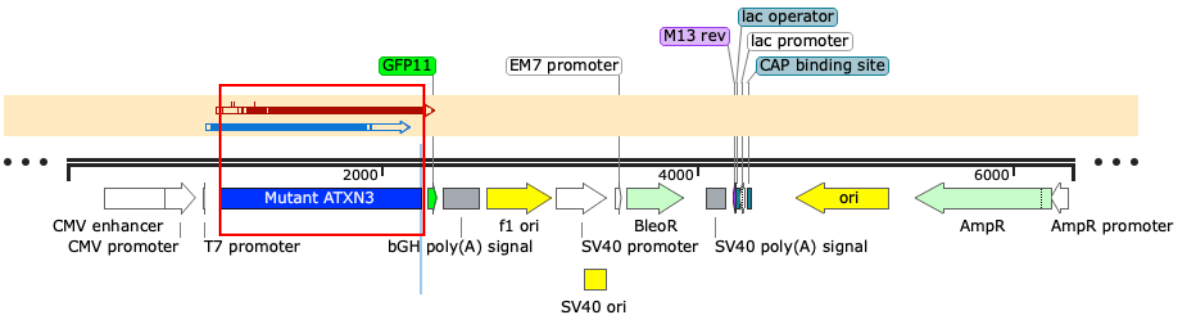
```
pcDNA GFP10-Mutant ATXN3          1001 CGACGACCACTACCTGTCCACCCAGACCATCCTGAGCAAGGACCTGAACA
Primer Forward GFP10-Mutant ATXN3 37  CGACGACCACTACCTGTCCACCCAGACCATCCTGAGCAAGGACCTGAACA
Primer Reverse GFP10-Mutant ATXN3 0   --------------------------------------------------

pcDNA GFP10-Mutant ATXN3          1051 TCGATGGTGGCGGTGGCTCTGGAGGTGGTGGGTCCTCCGGCGGCGGTGGA
Primer Forward GFP10-Mutant ATXN3 87  TCGATGGTGGCGGTGGCTCTGGAGGTGGTGGGTCCTCCGGCGGCGGTGGA
Primer Reverse GFP10-Mutant ATXN3 1   ------------------------------GGG-----------------

pcDNA GFP10-Mutant ATXN3          1101 TCTAGAATGGAGTCCATCTTCCACGAGAAACAAGAAGGCTCACTTTGTGC
Primer Forward GFP10-Mutant ATXN3 137 TCTAGAATGGAGTCCATCTTCCACGAGAAACAAGAAGGCTCACTTTGTGC
Primer Reverse GFP10-Mutant ATXN3 3   --------------------------------------------------

pcDNA GFP10-Mutant ATXN3          1151 TCAACATTGCCTGAATAACTTATTGCAAGGAGAATATTTTAGCCCTGTGG
Primer Forward GFP10-Mutant ATXN3 187 TCAACATTGCCTGAATAACTTATTGCAAGGAGAATATTTTAGCCCTGTGG
Primer Reverse GFP10-Mutant ATXN3 3   ----------------------------GGGAAATATTT---CCCT----

pcDNA GFP10-Mutant ATXN3          1201 AATTATCCTCAATTGCACATCAGCTGGATGAGGAGGAGA-GGATGAGAA-
Primer Forward GFP10-Mutant ATXN3 237 AATTATCCTCAATTGCACATCAGCTGGATGAGGAGGAGA-GGATGAGAA-
Primer Reverse GFP10-Mutant ATXN3 19  AAATTCCCCCACTTAAA---GGGGGGGAGGGGAGGGAGAGGGATGAAAAA

pcDNA GFP10-Mutant ATXN3          1248 -TGGCAGAAGGAGGAG---TTACTAGTGAAGATTATCGCAC----GTTTT
Primer Forward GFP10-Mutant ATXN3 284 -TGGCAGAAGGAGGAG---TTACTAGTGAAGATTATCGCAC----GTTTT
Primer Reverse GFP10-Mutant ATXN3 66  TTGCCAAAAGGGGGGAGTTTTTTAGGGAAGATTTTTTCCCCGGTTTTT

pcDNA GFP10-Mutant ATXN3          1291 TAC-AGCAGCCTTCTGG--AAATATGGATGA-CAGTGGTTTTTT-CTCTA
Primer Forward GFP10-Mutant ATXN3 327 TAC-AGCAGCCTTCTGG--AAATATGGATGA-CAGTGGTTTTTT-CTCTA
Primer Reverse GFP10-Mutant ATXN3 116 TACAAGCAGCCTTTTGGGAAAATATGGAAGACCAGTGGTTTTTTNTTCTA

pcDNA GFP10-Mutant ATXN3          1336 TTCAGGTTATAA-GCAATG-CCTTGAAAGTTTGGGGTTTAGAACTAATCC
Primer Forward GFP10-Mutant ATXN3 372 TTCAGGTTATAA-GCAATG-CCTTGAAAGTTTGGGGTTTAGAACTAATCC
Primer Reverse GFP10-Mutant ATXN3 166 TTCAGGTTATAANGCAATGCCCTTGAAAGTTTGGGGTTTAGAACTAATCC

pcDNA GFP10-Mutant ATXN3          1384 TGT-TCAACAGTCCAGAGTATCAGAGGCTCAGGATCGATCCTATAAATGA
Primer Forward GFP10-Mutant ATXN3 420 TGT-TCAACAGTCCAGAGTATCAGAGGCTCAGGATCGATCCTATAAATGA
Primer Reverse GFP10-Mutant ATXN3 216 TGTNTCAACAGTCCAGAGTATCAGAGGCTCAGGATCGATCCTATAAATGA

pcDNA GFP10-Mutant ATXN3          1433 AAGATCATTTATATGCAATTATAAGGAACACTGGTTTACAGTTAGAAAAT
Primer Forward GFP10-Mutant ATXN3 469 AAGATCATTTATATGCAATTATAAGGAACACTGGTTTACAGTTAGAAAAT
Primer Reverse GFP10-Mutant ATXN3 266 AAGATCATTTATATGCAATTATAAGGAACACTGGTTTACAGTTAGAAAAT

pcDNA GFP10-Mutant ATXN3          1483 TAGG-AAAACAGTGG-TTTAACTTGAA-TTCTCTCTTGACGGGTCCAGAA
Primer Forward GFP10-Mutant ATXN3 519 TAGG-AAAACAGTGG-TTTAACTTGAA-TTCTCTCTTGACGGGTCCAGAA
Primer Reverse GFP10-Mutant ATXN3 316 TAGGAAAAACAGTGGTTTAACTTGAATTCTCTCTTGACGGGTCCAGAA
```

```
pcDNA GFP10-Mutant ATXN3          1530 TTAATATCAGATACATATCTTGCACTTTTCTTGGCTCAATTACAACAGGA
Primer Forward GFP10-Mutant ATXN3 566  TTAATATCAGATACATATCTTGCACTTTTCTTGGCTCAATTACAACAGGA
Primer Reverse GFP10-Mutant ATXN3 366  TTAATATCAGATACATATCTTGCACTTTTCTTGGCTCAATTACAACAGGA

pcDNA GFP10-Mutant ATXN3          1580 AGGTTATTCTATATTTGTCGTTAAGGGTGATCTGCCAGATTGCGAAGCTG
Primer Forward GFP10-Mutant ATXN3 616  AGGTTATTCTATATTTGTCGTTAAGGGTGATCTGCCAGATTGCGAAGCTG
Primer Reverse GFP10-Mutant ATXN3 416  AGGTTATTCTATATTTGTCGTTAAGGGTGATCTGCCAGATTGCGAAGCTG

pcDNA GFP10-Mutant ATXN3          1630 ACCAACTCCTGCAGATGATTAGGGTCCAACAGATGCATCGACCAAAACTG
Primer Forward GFP10-Mutant ATXN3 666  ACCAACTCCTGCAGATGATTAGGGTCCAACAGATGCATCGACCAAAACTG
Primer Reverse GFP10-Mutant ATXN3 466  ACCAACTCCTGCAGATGATTAGGGTCCAACAGATGCATCGACCAAAACTG

pcDNA GFP10-Mutant ATXN3          1680 ATTGGCGAAGAACTGGCTCAACTGAAAGAACAGCGTGTGCATAAGACCGA
Primer Forward GFP10-Mutant ATXN3 716  ATTGGCGAAGAACTGGCTCAACTGAAAGAACAGCGTGTGCATAAGACCGA
Primer Reverse GFP10-Mutant ATXN3 516  ATTGGCGAAGAACTGGCTCAACTGAAAGAACAGCGTGTGCATAAGACCGA

pcDNA GFP10-Mutant ATXN3          1730 CCTGGAACGTGTCCTGGAAGCAAATGACGGCAGCGGCATGCTGGATGAAG
Primer Forward GFP10-Mutant ATXN3 766  CCTGGAACGTGTCCTGGAAGCAAATGACGGCAGCGGCATGCTGGATGAAG
Primer Reverse GFP10-Mutant ATXN3 566  CCTGGAACGTGTCCTGGAAGCAAATGACGGCAGCGGCATGCTGGATGAAG

pcDNA GFP10-Mutant ATXN3          1780 ACGAAGAAGATCTGCAGCGTGCCCTGGCACTGTCTCGTCAGGAAATTGAT
Primer Forward GFP10-Mutant ATXN3 816  ACGAAGAAGATCTGCAGCGTGCCCTGGCACTGTCTCGTCAGGAAATTGAT
Primer Reverse GFP10-Mutant ATXN3 616  ACGAAGAAGATCTGCAGCGTGCCCTGGCACTGTCTCGTCAGGAAATTGAT

pcDNA GFP10-Mutant ATXN3          1830 ATGGAAGACGAAGAAGCAGATCTGCGTCGCGCTATTCAGCTGTCAATGCA
Primer Forward GFP10-Mutant ATXN3 866  ATGGAAGACGAAGAAGCAGATCTGCGTCGCGCTATTCAGCTGTCAATGCA
Primer Reverse GFP10-Mutant ATXN3 666  ATGGAAGACGAAGAAGCAGATCTGCGTCGCGCTATTCAGCTGTCAATGCA

pcDNA GFP10-Mutant ATXN3          1880 GGGCAGCTCTCGTAACATCTCGCAGGACATGACCCAGACGAGCGGTACCA
Primer Forward GFP10-Mutant ATXN3 916  GGGCAGCTCTCGTAACATCTCGCAGGACATGACCCAGACGAGCGGTACCA
Primer Reverse GFP10-Mutant ATXN3 716  GGGCAGCTCTCGTAACATCTCGCAGGACATGACCCAGACGAGCGGTACCA

pcDNA GFP10-Mutant ATXN3          1930 ATCTGACGTCTGAAGAACTGCGCAAACGTCGCGAAGCATATTTTGAAAAA
Primer Forward GFP10-Mutant ATXN3 966  ATCTGACGTCTGAAGAACTGCGCAAACGTCGCGAAGCATATTTTGAAAAA
Primer Reverse GFP10-Mutant ATXN3 766  ATCTGACGTCTGAAGAACTGCGCAAACGTCGCGAAGCATATTTTGAAAAA

pcDNA GFP10-Mutant ATXN3          1980 CAG-CAACAG-AAGCAACAACAGCAGCAACAACAACAACAACAGCAGCAG
Primer Forward GFP10-Mutant ATXN3 1016 CAGCCAACAGAAAGCAACAACAGCAGCAACAACAACAACAACGGCGGCAG
Primer Reverse GFP10-Mutant ATXN3 816  CAG-CAACAG-AAGCAACAACAGCAGCAACAACAACAACAACAGCAGCAG

pcDNA GFP10-Mutant ATXN3          2028 CAGCAGCAACAACAACAGCAACAACAACAGCAACAGCAGCAACAGCAACA
Primer Forward GFP10-Mutant ATXN3 1066 CAGCAGCAACAACAACGGCAACAACAACGGCAACGGCGGCAACAGCAACG
Primer Reverse GFP10-Mutant ATXN3 864  CAGCAGCAACAACAACAGCAACAACAACAGCAACAGCAGCAACAGCAACA
```

```
pcDNA GFP10-Mutant ATXN3          2078 GCAACAGCAGCAACAACAGCAACAGCAACAACAGCAACAGCAACAACAGC
Primer Forward GFP10-Mutant ATXN3 1116 GCAACGGCAGCAACAA--------------------------------
Primer Reverse GFP10-Mutant ATXN3 914  GCAACAGCAGCAACAACAGCAACAGCAACAACAGCAACAGCAACAACAGC

pcDNA GFP10-Mutant ATXN3          2128 AACAACAGCAACAACAGCAACAACAGCAACAACAGCAACAACAGCAACAG
Primer Forward GFP10-Mutant ATXN3 1131 -------------------------------------------------
Primer Reverse GFP10-Mutant ATXN3 964  AACAACAGCAACAACAGCAACAACAGCAACAACAGCAACAACAGCAACAG

pcDNA GFP10-Mutant ATXN3          2178 CAACAACAGCAGCAGCAACAGCAGCAACAACAACAGAGAGATCTGTCAGG
Primer Forward GFP10-Mutant ATXN3 1131 -------------------------------------------------
Primer Reverse GFP10-Mutant ATXN3 1014 CAACAACAGCAGCAGCAACAGCAGCAACAACAACAGAGAGATCTGTCAGG

pcDNA GFP10-Mutant ATXN3          2228 CCAGAGTTCCCATCCGTGTGAACGTCCGGCCACCTCAAGCGGTGCACTGG
Primer Forward GFP10-Mutant ATXN3 1131 -------------------------------------------------
Primer Reverse GFP10-Mutant ATXN3 1064 CCAGAGTTCCCATCCGTGTGAACGTCCGGCCACCTCAAGCGGTGCACTGG

pcDNA GFP10-Mutant ATXN3          2278 GTAGTGATCTGGGTGACGCCATGTCCGAAGAAGACATGCTGCAGGCAGCA
Primer Forward GFP10-Mutant ATXN3 1131 -------------------------------------------------
Primer Reverse GFP10-Mutant ATXN3 1114 GTAGTGATCTGGGTGACGCCATGTCCGAAGAAGACATGCTGCAGGCAGCA

pcDNA GFP10-Mutant ATXN3          2328 GTGACGATGTCCCTGGAAACCGTGCGTAACGACCTGAAAACCGAAGGCAA
Primer Forward GFP10-Mutant ATXN3 1131 ---------------------------------CCGGAAACGGGA-----
Primer Reverse GFP10-Mutant ATXN3 1164 GTGACGATGTCCCTGGAAACCGTGCGTAACGACCTGAAAACCGAAGGCAA

pcDNA GFP10-Mutant ATXN3          2378 AAAA
Primer Forward GFP10-Mutant ATXN3 1143 ----
Primer Reverse GFP10-Mutant ATXN3 1214 AAAA
```

Figure 60 - Sequence alignment of the pcDNA GFP10-Mutant ATXN3 vector with primers detailed in table 2, obtained from SnapGene® software (from Insightful Science; available at snapgene.com).

pcDNA Mutant ATXN3 – GFP11 vector alignment



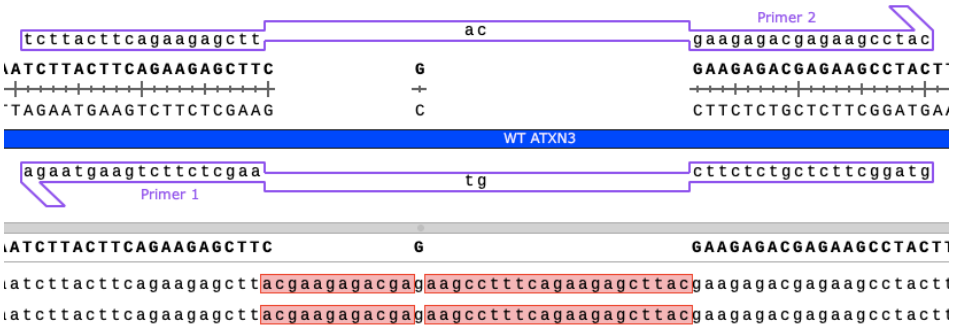
```
pcDNA Mutant ATXN3-GFP11          969  ATGGCTATGGAGTCCATCTTCCACGAGAAACAAGAAGG
Primer Forward Mutant ATXN3-GFP11 84   ATGGCTATGGAGTCCATCTTCCACGAGAAACAAGAAGG
Primer Reverse Mutant ATXN3-GFP11 13   GGGTTT---GNNTTTTTTTCCNGAAA---AAAAAAG

pcDNA Mutant ATXN3-GFP11          1001 CTCACTTTGTGCTCAACATTGCCTGAATAACTTATTGCAAGGAGAATATT
Primer Forward Mutant ATXN3-GFP11 116  CTCACTTTGTGCTCAACATTGCCTGAATAACTTATTGCAAGGAGAATATT
Primer Reverse Mutant ATXN3-GFP11 57   GTCCTTTG-GTTTAACCTT-CCCGAAAAATTTTTTGCAAGGGGAAATAT

pcDNA Mutant ATXN3-GFP11          1051 TTAGCCCT-GTGGAATTATCCTCAATTGCACATCAGCTGGATGAGGAGGA
Primer Forward Mutant ATXN3-GFP11 166  TTAGCCCT-GTGGAATTATCCTCAATTGCACATCAGCTGGATGAGGAGGA
Primer Reverse Mutant ATXN3-GFP11 105  TTAGCCCCGGGGGAATTTTCCTCAATTGCCCATCAGCTGGATGGGGAGG-

pcDNA Mutant ATXN3-GFP11          1100 GAGGATGAGAATGGCAGAAGGAGGAGTTACTAGTGAAGATTATCGCACGT
Primer Forward Mutant ATXN3-GFP11 215  GAGGATGAGAATGGCAGAAGGAGGAGTTACTAGTGAAGATTATCGCACGT
Primer Reverse Mutant ATXN3-GFP11 154  GAGGATGAGAATGGCAGAAGGAGGAGTTATTAGGGAAGATTATCGCACGT

pcDNA Mutant ATXN3-GFP11          1150 TTTTACAGCAGCCTTCTGGAAATATGGATGACAGTGGTTTTTTC-TCTAT
Primer Forward Mutant ATXN3-GFP11 265  TTTTACAGCAGCCTTCTGGAAATATGGATGACAGTGGTTTTTTC-TCTAT
Primer Reverse Mutant ATXN3-GFP11 204  TTTTACAGCAGCCTTCTGGAAATATGGATGACAGTGGTTTTTTCTTCTAT

pcDNA Mutant ATXN3-GFP11          1199 TCAGGTTATAAGCAATGCCTTGAAAGTTTGGGGTTTAGAACTAATCCTGT
Primer Forward Mutant ATXN3-GFP11 314  TCAGGTTATAAGCAATGCCTTGAAAGTTTGGGGTTTAGAACTAATCCTGT
Primer Reverse Mutant ATXN3-GFP11 254  TCAGGTTATAAGCAATGCCTTGAAAGTTTGGGGTTTAGAACTAATCCTGT

pcDNA Mutant ATXN3-GFP11          1249 TCAACAGTCCAGAGTATCAGAGGCTCAGGATCGATCCTATAAATGAAAGA
Primer Forward Mutant ATXN3-GFP11 364  TCAACAGTCCAGAGTATCAGAGGCTCAGGATCGATCCTATAAATGAAAGA
Primer Reverse Mutant ATXN3-GFP11 304  TCAACAGTCCAGAGTATCAGAGGTTCAGGATCGATCCTATAAATGAAAGA

pcDNA Mutant ATXN3-GFP11          1299 TCATTTATATGCAATTATAAGGAACACTGGTTTACAGTTAGAAAATTAGG
Primer Forward Mutant ATXN3-GFP11 414  TCATTTATATGCAATTATAAGGAACACTGGTTTACAGTTAGAAAATTAGG
Primer Reverse Mutant ATXN3-GFP11 354  TCATTTATATGCAATTATAAGGAACACTGGTTTACAGTTAGAAAATTAGG

pcDNA Mutant ATXN3-GFP11          1349 AAAACAGTGGTTTAACTTGAATTCTCTCTTGACGGGTCCAGAATTAATAT
Primer Forward Mutant ATXN3-GFP11 464  AAAACAGTGGTTTAACTTGAATTCTCTCTTGACGGGTCCAGAATTAATAT
Primer Reverse Mutant ATXN3-GFP11 404  AAAACAGTGGTTTAACTTGAATTCTCTCTTGACGGGTCCAGAATTAATAT

pcDNA Mutant ATXN3-GFP11          1399 CAGATACATATCTTGCACTTTTCTTGGCTCAATTACAACAGGAAGGTTAT
Primer Forward Mutant ATXN3-GFP11 514  CAGATACATATCTTGCACTTTTCTTGGCTCAATTACAACAGGAAGGTTAT
Primer Reverse Mutant ATXN3-GFP11 454  CAGATACATATCTTGCACTTTTCTTGGCTCAATTACAACAGGAAGGTTAT

pcDNA Mutant ATXN3-GFP11          1449 TCTATATTTGTCGTTAAGGGTGATCTGCCAGATTGCGAAGCTGACCAACT
Primer Forward Mutant ATXN3-GFP11 564  TCTATATTTGTCGTTAAGGGTGATCTGCCAGATTGCGAAGCTGACCAACT
Primer Reverse Mutant ATXN3-GFP11 504  TCTATATTTGTCGTTAAGGGTGATCTGCCAGATTGCGAAGCTGACCAACT
```

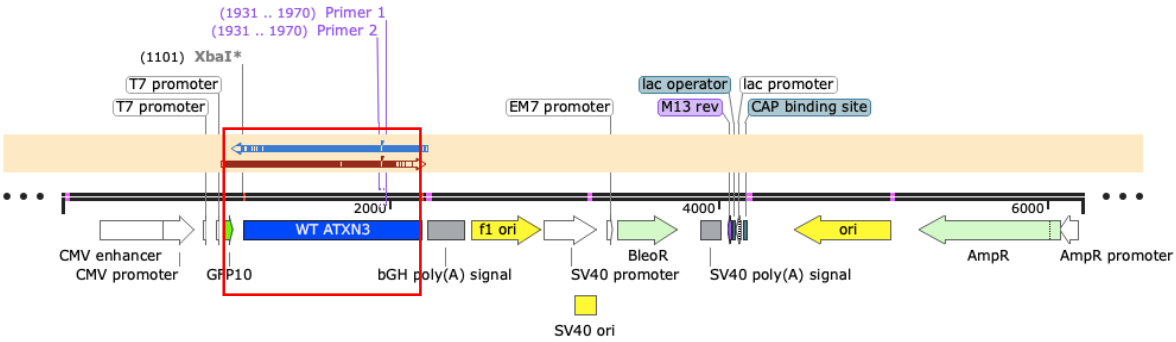
```
pcDNA Mutant ATXN3-GFP11          1499 CCTGCAGATGATTAGGGTCCAACAGATGCATCGACCAAAACTGATTGGCG
Primer Forward Mutant ATXN3-GFP11 614  CCTGCAGATGATTAGGGTCCAACAGATGCATCGACCAAAACTGATTGGCG
Primer Reverse Mutant ATXN3-GFP11 554  CCTGCAGATGATTAGGGTCCAACAGATGCATCGACCAAAACTGATTGGCG

pcDNA Mutant ATXN3-GFP11          1549 AAGAACTGGCTCAACTGAAAGAACAGCGTGTGCATAAGACCGACCTGGAA
Primer Forward Mutant ATXN3-GFP11 664  AAGAACTGGCTCAACTGAAAGAACAGCGTGTGCATAAGACCGACCTGGAA
Primer Reverse Mutant ATXN3-GFP11 604  AAGAACTGGCTCAACTGAAAGAACAGCGTGTGCATAAGACCGACCTGGAA

pcDNA Mutant ATXN3-GFP11          1599 CGTGTCCTGGAAGCAAATGACGGCAGCGGCATGCTGGATGAAGACGAAGA
Primer Forward Mutant ATXN3-GFP11 714  CGTGTCCTGGAAGCAAATGACGGCAGCGGCATGCTGGATGAAGACGAAGA
Primer Reverse Mutant ATXN3-GFP11 654  CGTGTCCTGGAAGCAAATGACGGCAGCGGCATGCTGGATGAAGACGAAGA

pcDNA Mutant ATXN3-GFP11          1649 AGATCTGCAGCGTGCCCTGGCACTGTCTCGTCAGGAAATTGATATGGAAG
Primer Forward Mutant ATXN3-GFP11 764  AGATCTGCAGCGTGCCCTGGCACTGTCTCGTCAGGAAATTGATATGGAAG
Primer Reverse Mutant ATXN3-GFP11 704  AGATCTGCAGCGTGCCCTGGCACTGTCTCGTCAGGAAATTGATATGGAAG

pcDNA Mutant ATXN3-GFP11          1699 ACGAAGAAGCAGATCTGCGTCGCGCTATTCAGCTGTCAATGCAGGGCAGC
Primer Forward Mutant ATXN3-GFP11 814  ACGAAGAAGCAGATCTGCGTCGCGCTATTCAGCTGTCAATGCAGGGCAGC
Primer Reverse Mutant ATXN3-GFP11 754  ACGAAGAAGCAGATCTGCGTCGCGCTATTCAGCTGTCAATGCAGGGCAGC

pcDNA Mutant ATXN3-GFP11          1749 TCTCGTAACATCTCGCAGGACATGACCCAGACGAGCGGTACCAATCTGAC
Primer Forward Mutant ATXN3-GFP11 864  TCTCGTAACATCTCGCAGGACATGACCCAGACGAGCGGTACCAATCTGAC
Primer Reverse Mutant ATXN3-GFP11 804  TCTCGTAACATCTCGCAGGACATGACCCAGACGAGCGGTACCAATCTGAC

pcDNA Mutant ATXN3-GFP11          1799 GTCTGAAGAACTGCGCAAACGTCGCGAAGCATATTTTGAAAAACAGCAAC
Primer Forward Mutant ATXN3-GFP11 914  GTCTGAAGAACTGCGCAAACGTCGCGAAGCATATTTTGAAAAACAGCAAC
Primer Reverse Mutant ATXN3-GFP11 854  GTCTGAAGAACTGCGCAAACGTCGCGAAGCATATTTTGAAAAACAGCAAC

pcDNA Mutant ATXN3-GFP11          1849 AGAAGCAACAACAGCAGCAACAACAACAACAACAGCAGCAGCAGCAGCAA
Primer Forward Mutant ATXN3-GFP11 964  AGAAGCAACAACAGCAGCAACAACAACAACAACAGCAGCAGCAGCAGCAA
Primer Reverse Mutant ATXN3-GFP11 904  AGAAGCAACAACAGCAGCAACAACAACAACAACAGCAGCAGCAGCAGCAA

pcDNA Mutant ATXN3-GFP11          1899 CAACAACAGCAACAACAACAGCAACAGCAGCAACAGCAACAGCAACAGCA
Primer Forward Mutant ATXN3-GFP11 1014 CAACAACGGCAACAACAACGGCAACAGCAGCAACAGCAACGGCAACGGCA
Primer Reverse Mutant ATXN3-GFP11 954  CAACAACAGCAACAACAACAGCAACAGCAGCAACAGCAACAGCAACAGCA

pcDNA Mutant ATXN3-GFP11          1949 GCAACAACAGCAACAGCAACAACAGCAACAGCAACAACAGCAACAACAGC
Primer Forward Mutant ATXN3-GFP11 1064 GCAACAACGGCAACGGCAACAACGGCAACNGCAACAANAGNAACAACGGC
Primer Reverse Mutant ATXN3-GFP11 1004 GCAACAACAGCAACAGCAACAACAGCAACAGCAACAACAGCAACAACAGC

pcDNA Mutant ATXN3-GFP11          1999 AACAACAGCAACAACAGCAACAACAGCAACAACAGCAACAGCAACAACAG
Primer Forward Mutant ATXN3-GFP11 1114 AACAACGGCAACAACGGCAACAACGGCAACAACGGCAACGGGAACAACGG
Primer Reverse Mutant ATXN3-GFP11 1054 AACAACAGCAACAACAGCAACAACAGCAACAACAGCAACAGCAACAACAG
```

```
pcDNA Mutant ATXN3-GFP11          2049 CAGCAGCAACAGCAGCAACAACAACAGAGAGATCTGTCAGGCCAGAGTTC
Primer Forward Mutant ATXN3-GFP11 1164 GGGCAGCAACGGC-GCCACCACCACGGAAAGATTTGTCAGGCCA-AATTC
Primer Reverse Mutant ATXN3-GFP11 1104 CAGCAGCAACAGCAGCAACAACAACAGAGAGATCTGTCAGGCCAGAGTTC

pcDNA Mutant ATXN3-GFP11          2099 CCATCCGTGTGAACGTCCGGCCACCTCAAGCGGTGCACTGGGTAGTGATC
Primer Forward Mutant ATXN3-GFP11 1212 CCATCCCGGTNAACCTCCGGCCCCNTCNGG-GGGGAANTGGNTAAGGATT
Primer Reverse Mutant ATXN3-GFP11 1154 CCATCCGTGTGAACGTCCGGCCACCTCAAGCGGTGCACTGGGTAGTGATC

pcDNA Mutant ATXN3-GFP11          2149 TGGGTGACGCCATGTCCGAAGAAGACATGCTGCAGGCAGCAGTGACGATG
Primer Forward Mutant ATXN3-GFP11 1261 NGGGGAAC-CCCTGNTCCAAAAAAAANTGTTTCAGG--------------
Primer Reverse Mutant ATXN3-GFP11 1204 TGGGTGACGCCATGTCCGAAGAAGACATGCTGCAGGCAGCAGTGACGATG

pcDNA Mutant ATXN3-GFP11          2199 TCCCTGGAAACCGTGCGTAACGACCTGAAAACCGAAGGCAAAAAAGATGG
Primer Forward Mutant ATXN3-GFP11 1295 ------------------------------------------------
Primer Reverse Mutant ATXN3-GFP11 1254 TCCCTGGAAACCGTGCGTAACGACCTGAAAACCGAAGGCAAAAAAGATGG

pcDNA Mutant ATXN3-GFP11          2249 TGGCGGTGGCTCTGGAGGTGGTGGGTCCTCCGGAGAGAAGCGCGACCACA
Primer Forward Mutant ATXN3-GFP11 1295 ------------------------------------------------
Primer Reverse Mutant ATXN3-GFP11 1304 TGGCGGTGGCTCTGGAGGTGGTGGGTCCTCCGGAGAGAAGCGCGACCACA
```

Figure 61 - Sequence alignment of the pcDNA Mutant ATXN3-GFP11 vector with primers detailed in table 2, obtained from SnapGene® software (from Insightful Science; available at snapgene.com).

pcDNA GFP10 – ATXN3 WT R282T NLS vector alignment





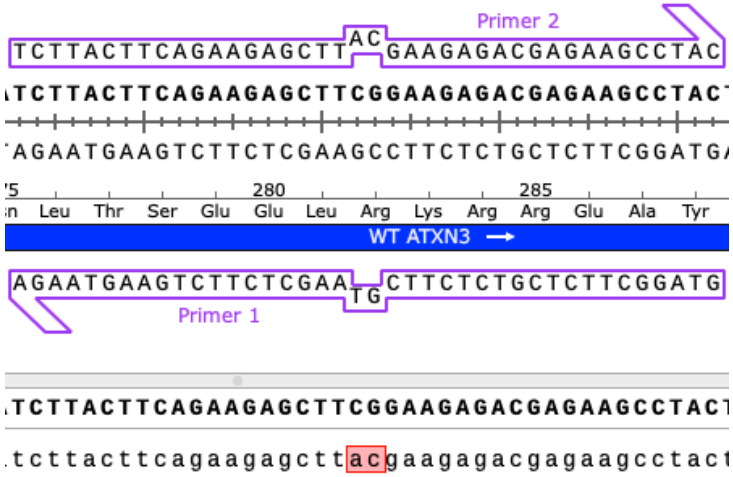

```
pcDNA GFP10-WT ATXN3              987  ATGGGCGACCTGCC
Primer Fw NLS Mut. GFP10-WT ATXN3 18   NTGGGCGACCTGCC
Primer Rv NLS Mut. GFP10-WT ATXN3 0    -------------

pcDNA GFP10-WT ATXN3              1001 CGACGACCACTACCTGTCCACCCAGACCATCCTGAGCAAGGACCTGAACA
Primer Fw NLS Mut. GFP10-WT ATXN3 32   CGACGACCACTACCTGTCCACCCAGACCATCCTGAGCAAGGACCTGAACA
Primer Rv NLS Mut. GFP10-WT ATXN3 1    ----------------------------------AGCAAGG-CCTGAA-A

pcDNA GFP10-WT ATXN3              1051 TCGATGGTGGCGGTGGCTCTGGAGGTGGTGGGTCCTCCGGCGGCGGTGGA
Primer Fw NLS Mut. GFP10-WT ATXN3 82   TCGATGGTGGCGGTGGCTCTGGAGGTGGTGGGTCCTCCGGCGGCGGTGGA
Primer Rv NLS Mut. GFP10-WT ATXN3 15   TCGATGG-GGCGGTGGCNC-GGAGGTNG-GGGTCCTCCGGGGGCGGTGGA

pcDNA GFP10-WT ATXN3              1101 TCTAGAATGGAGTCCATCTTCCACGAGAAACAAGAAGGCTCACTTTGTGC
Primer Fw NLS Mut. GFP10-WT ATXN3 132  TCTAGAATGGAGTCCATCTTCCACGAGAAACAAGAAGGCTCACTTTGTGC
Primer Rv NLS Mut. GFP10-WT ATXN3 62   TCTAGAATGGAGTCCTTTTTCC-CGAGAAACAAGAAGGCTCACTTTGTGC

pcDNA GFP10-WT ATXN3              1151 TCAACATTGCCTGAATAACTTATTGCAAGGAGAATATTTTAGCCCTGTGG
Primer Fw NLS Mut. GFP10-WT ATXN3 182  TCAACATTGCCTGAATAACTTATTGCAAGGAGAATATTTTAGCCCTGTGG
Primer Rv NLS Mut. GFP10-WT ATXN3 111  TCAACA-TGCCTGAATAACTTA-TGCAAGGAGAATA-TTTAGCCCTGTGG

pcDNA GFP10-WT ATXN3              1201 AATTATCCTCAATTGCACATCAGCTGGATGAGGAGGAGAGGATGAGAATG
Primer Fw NLS Mut. GFP10-WT ATXN3 232  AATTATCCTCAATTGCACATCAGCTGGATGAGGAGGAGAGGATGAGAATG
Primer Rv NLS Mut. GFP10-WT ATXN3 158  AATTATCCTCAATTGCACATCAGCNGGATGAGGAGGAGAGGATGAGAATG

pcDNA GFP10-WT ATXN3              1251 GCAGAAGGAGGAGTTACTAGTGAAGATTATCGCACGTTTTTACAGCAGCC
Primer Fw NLS Mut. GFP10-WT ATXN3 282  GCAGAAGGAGGAGTTACTAGTGAAGATTATCGCACGTTTTTACAGCAGCC
Primer Rv NLS Mut. GFP10-WT ATXN3 208  GCAGAAGGAGGAGTTACTAGTGAAGATTATCGCACGTTTTTACAGCAGCC

pcDNA GFP10-WT ATXN3              1301 TTCTGGAAATATGGATGACAGTGGTTTTTTCTCTATTCAGGTTATAAGCA
Primer Fw NLS Mut. GFP10-WT ATXN3 332  TTCTGGAAATATGGATGACAGTGGTTTTTTCTCTATTCAGGTTATAAGCA
Primer Rv NLS Mut. GFP10-WT ATXN3 258  TTCTGGAAATATGGATGACAGTGGTTTTTTCTCTATTCAGGTTATAAGCA

pcDNA GFP10-WT ATXN3              1351 ATGCCTTGAAAGTTTGGGGTTTAGAACTAATCCTGTTCAACAGTCCAGAG
Primer Fw NLS Mut. GFP10-WT ATXN3 382  ATGCCTTGAAAGTTTGGGGTTTAGAACTAATCCTGTTCAACAGTCCAGAG
Primer Rv NLS Mut. GFP10-WT ATXN3 308  ATGCCTTGAAAGTTTGGGGTTTAGAACTAATCCTGTTCAACAGTCCAGAG
```

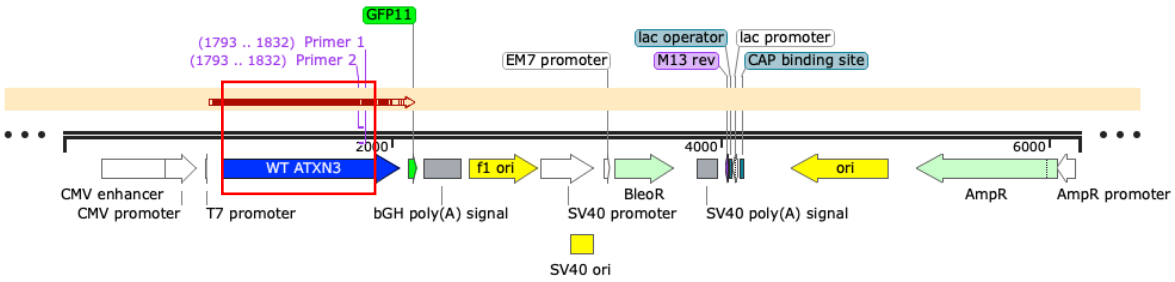
```
pcDNA GFP10-WT ATXN3              1401 TATCAGAGGCTCAGGATCGATCCTATAAATGAAAGATCATTTATATGCAA
Primer Fw NLS Mut. GFP10-WT ATXN3 432  TATCAGAGGCTCAGGATCGATCCTATAAATGAAAGATCATTTATATGCAA
Primer Rv NLS Mut. GFP10-WT ATXN3 358  TATCAGAGGCTCAGGATCGATCCTATAAATGAAAGATCATTTATATGCAA

pcDNA GFP10-WT ATXN3              1451 TTATAAGGAACACTGGTTTACAGTTAGAAAATTAGGAAAACAGTGGTTTA
Primer Fw NLS Mut. GFP10-WT ATXN3 482  TTATAAGGAACACTGGTTTACAGTTAGAAAATTAGGAAAACAGTGGTTTA
Primer Rv NLS Mut. GFP10-WT ATXN3 408  TTATAAGGAACACTGGTTTACAGTTAGAAAATTAGGAAAACAGTGGTTTA

pcDNA GFP10-WT ATXN3              1501 ACTTGAATTCTCTCTTGACGGGTCCAGAATTAATATCAGATACATATCTT
Primer Fw NLS Mut. GFP10-WT ATXN3 532  ACTTGAATTCTCTCTTGACGGGTCCAGAATTAATATCAGATACATATCTT
Primer Rv NLS Mut. GFP10-WT ATXN3 458  ACTTGAATTCTCTCTTGACGGGTCCAGAATTAATATCAGATACATATCTT

pcDNA GFP10-WT ATXN3              1551 GCACTTTTCTTGGCTCAATTACAACAGGAAGGTTATTCTATATTTGTCGT
Primer Fw NLS Mut. GFP10-WT ATXN3 582  GCACTTTTCTTGGCTCAATTACAACAGGAAGGTTATTCTATATTTGTCGT
Primer Rv NLS Mut. GFP10-WT ATXN3 508  GCACTTTTCTTGGCTCAATTACAACAGGAAGGTTATTCTATATTTGTCGT

pcDNA GFP10-WT ATXN3              1601 TAAGGGTGATCTGCCAGATTGCGAAGCTGACCAACTCCTACAGATGATTA
Primer Fw NLS Mut. GFP10-WT ATXN3 632  TAAGGGTGATCTGCCAGATTGCGAAGCTGACCAACTCCTACAGATGATTA
Primer Rv NLS Mut. GFP10-WT ATXN3 558  TAAGGGTGATCTGCCAGATTGCGAAGCTGACCAACTCCTACAGATGATTA

pcDNA GFP10-WT ATXN3              1651 GGGTCCAACAGATGCATCGACCAAAACTTATTGGAGAAGAATTAGCACAA
Primer Fw NLS Mut. GFP10-WT ATXN3 682  GGGTCCAACAGATGCATCGACCAAAACTTATTGGAGAAGAATTAGCACAA
Primer Rv NLS Mut. GFP10-WT ATXN3 608  GGGTCCAACAGATGCATCGACCAAAACTTATTGGAGAAGAATTAGCACAA

pcDNA GFP10-WT ATXN3              1701 CTAAAAGAGCAAAGAGTCCATAAAACAGACCTGGAACGAGTGTTAGAAGC
Primer Fw NLS Mut. GFP10-WT ATXN3 732  CTAAGAGAGCAAAGAGTCCATAAAACAGACCTGGAACGAGTGTTAGAAGC
Primer Rv NLS Mut. GFP10-WT ATXN3 658  CTAAAAGAGCAAAGAGTCCATAAAACAGACCTGGAACGAGTGTTAGAAGC

pcDNA GFP10-WT ATXN3              1751 AAATGATGGCTCAGGAATGTTAGACGAAGATGAGGAGGATTTGCAGAGGG
Primer Fw NLS Mut. GFP10-WT ATXN3 782  AAATGATGGCTCAGGAATGTTAGACGAAGATGAGGAGGATTTGCAGAGGG
Primer Rv NLS Mut. GFP10-WT ATXN3 708  AAATGATGGCTCAGGAATGTTAGACGAAGATGAGGAGGATTTGCAGAGGG

pcDNA GFP10-WT ATXN3              1801 CTCTGGCACTAAGTCGCCAAGAAATTGACATGGAAGATGAGGAAGCAGAT
Primer Fw NLS Mut. GFP10-WT ATXN3 832  CTCTGGCACTAAGTCGCCAAGAAATTGACATGGAAGATGAGGAAGCAGAT
Primer Rv NLS Mut. GFP10-WT ATXN3 758  CTCTGGCACTAAGTCGCCAAGAAATTGACATGGAAGATGAGGAAGCAGAT

pcDNA GFP10-WT ATXN3              1851 CTCCGCAGGGCTATTCAGCTAAGTATGCAAGGTAGTTCCAGAAACATATC
Primer Fw NLS Mut. GFP10-WT ATXN3 882  CTCCGCAGGGCTATTCAGCTAAGTATGCAAGGTAGTTCCAGAAACATATC
Primer Rv NLS Mut. GFP10-WT ATXN3 808  CTCCGCAGGGCTATTCAGCTAAGTATGCAAGGTAGTTCCAGAAACATATC

pcDNA GFP10-WT ATXN3              1901 TCAAGATATGACACAGACATCAGGTACAAATCTTACTTCAGAAGAGCTT-
Primer Fw NLS Mut. GFP10-WT ATXN3 932  TCAAGATATGACACAGACATCAGGTACAAATCTTACTTCAGAAGAGCTTA
Primer Rv NLS Mut. GFP10-WT ATXN3 858  TCAAGATATGACACAGACATCAGGTACAAATCTTACTTCAGAAGAGCTTA
```

```
pcDNA GFP10-WT ATXN3              1949 --------CG----------------------GAAGAGACGAGAAGCCT
Primer Fw NLS Mut. GFP10-WT ATXN3 982  CGAAGAGACGAGAAGCCTTTCAGAAGAGCTTACGAAGAGACGAGAAGCCT
Primer Rv NLS Mut. GFP10-WT ATXN3 908  CGAAGAGACGAGAAGCCTTTCAGAAGAGCTTACGAAGAGACGAGAAGCCT

pcDNA GFP10-WT ATXN3              1969 ACTTTGAAAAACAGCAGCAAAAGCAGCAACAGCAGCAGCAGCAGCAGCAG
Primer Fw NLS Mut. GFP10-WT ATXN3 1032 ACTTTGAAAAACAGCAGCAAAAGCAGCAACAGCAGCAGCAGCAGCAGCAG
Primer Rv NLS Mut. GFP10-WT ATXN3 958  ACTTTGAAAAACAGCAGCAAAAGCAGCAACAGCAGCAGCAGCAGCAGCAG

pcDNA GFP10-WT ATXN3              2019 CAGGGGGACCTATCAGGACAGAGTTCACATCCATGTGAAAGGCCAGCCAC
Primer Fw NLS Mut. GFP10-WT ATXN3 1082 CAGGGGGACCTATCAGGACGGAGTTCANTTCC-TGTGAAAGGCC-GCCAC
Primer Rv NLS Mut. GFP10-WT ATXN3 1008 CAGGGGGACCTATCAGGACAGAGTTCACATCCATGTGAAAGGCCAGCCAC

pcDNA GFP10-WT ATXN3              2069 CAGTTCAGGAGCACTTGGGAGTGATCTAGGTGATGCTATGAGTGAAGAAG
Primer Fw NLS Mut. GFP10-WT ATXN3 1130 CAGTTCAGGAGC-NTTGGGAGTGATCTAGGGGATGCTAT-AATGAAGAGA
Primer Rv NLS Mut. GFP10-WT ATXN3 1058 CAGTTCAGGAGCACTTGGGAGTGATCTAGGTGATGCTATGAGTGAAGAAG

pcDNA GFP10-WT ATXN3              2119 ACATGCTTCAGGCAGCTGTGACCATGTCTTTAGAAACTGTCAGAAATGAT
Primer Fw NLS Mut. GFP10-WT ATXN3 1177 --NTGCTTCAGG-AGCTGTGACC-TGTCTTTA-AAACTGTC-NAAATGAT
Primer Rv NLS Mut. GFP10-WT ATXN3 1108 ACATGCTTCAGGCAGCTGTGACCATGTCTTTAGAAACTGTCAGAAATGAT

pcDNA GFP10-WT ATXN3              2169 TTGAAAACAGAAGGAAAAAAA
Primer Fw NLS Mut. GFP10-WT ATXN3 1222 TTGAAACC---AAGGAAAAAA
Primer Rv NLS Mut. GFP10-WT ATXN3 1158 TTGAAAACAGAAGGAAAAAAA
```

Figure 62 - Sequence alignment of the pcDNA GFP10 – WT ATXN3 R282T NLS vector with primers detailed in table 2, obtained from SnapGene® software (from Insightful Science; available at snapgene.com). Pink: inserted primers with mutation. Red: incorrect insertion of nucleotides.

pcDNA ATXN3 WT R282T NLS – GFP11 vector alignment



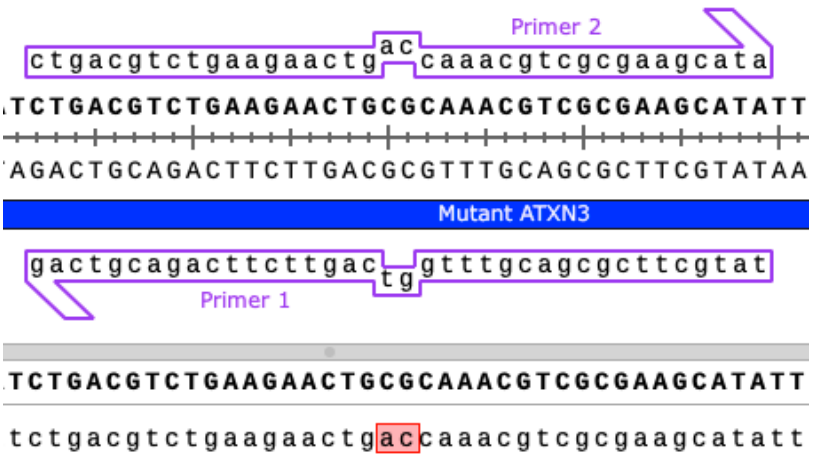


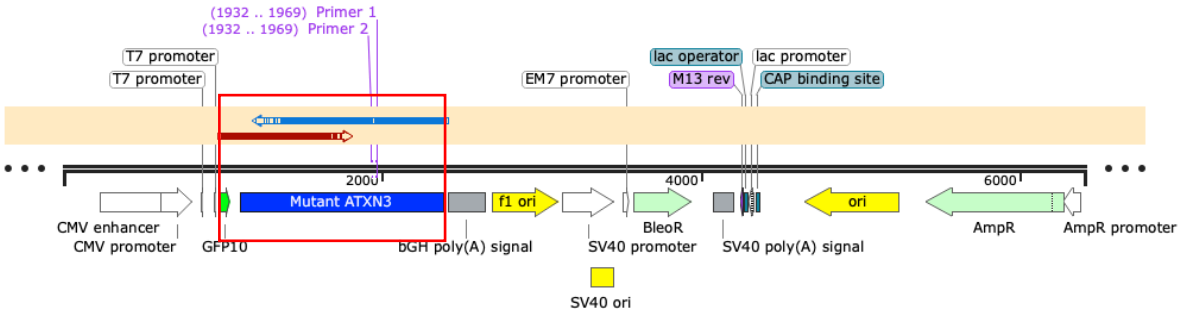

```
pcDNA WT ATXN3-GFP11           969  ATGGAGTCCATCTTCCACGAGAAACAAGAAGG
Primer Fw WT ATXN3-GFP11 R282T 80   ATGGAGTCCATCTTCCACGAGAAACAAGAAGG

pcDNA WT ATXN3-GFP11           1001 CTCACTTTGTGCTCAACATTGCCTGAATAACTTATTGCAAGGAGAATATT
Primer Fw WT ATXN3-GFP11 R282T 112  CTCACTTTGTGCTCAACATTGCCTGAATAACTTATTGCAAGGAGAATATT

pcDNA WT ATXN3-GFP11           1051 TTAGCCCTGTGGAATTATCCTCAATTGCACATCAGCTGGATGAGGAGGAG
Primer Fw WT ATXN3-GFP11 R282T 162  TTAGCCCTGTGGAATTATCCTCAATTGCACATCAGCTGGATGAGGAGGAG

pcDNA WT ATXN3-GFP11           1101 AGGATGAGAATGGCAGAAGGAGGAGTTACTAGTGAAGATTATCGCACGTT
Primer Fw WT ATXN3-GFP11 R282T 212  AGGATGAGAATGGCAGAAGGAGGAGTTACTAGTGAAGATTATCGCACGTT

pcDNA WT ATXN3-GFP11           1151 TTTACAGCAGCCTTCTGGAAATATGGATGACAGTGGTTTTTTCTCTATTC
Primer Fw WT ATXN3-GFP11 R282T 262  TTTACAGCAGCCTTCTGGAAATATGGATGACAGTGGTTTTTTCTCTATTC

pcDNA WT ATXN3-GFP11           1201 AGGTTATAAGCAATGCCTTGAAAGTTTGGGGTTTAGAACTAATCCTGTTC
Primer Fw WT ATXN3-GFP11 R282T 312  AGGTTATAAGCAATGCCTTGAAAGTTTGGGGTTTAGAACTAATCCTGTTC

pcDNA WT ATXN3-GFP11           1251 AACAGTCCAGAGTATCAGAGGCTCAGGATCGATCCTATAAATGAAAGATC
Primer Fw WT ATXN3-GFP11 R282T 362  AACAGTCCAGAGTATCAGAGGCTCAGGATCGATCCTATAAATGAAAGATC

pcDNA WT ATXN3-GFP11           1301 ATTTATATGCAATTATAAGGAACACTGGTTTACAGTTAGAAAATTAGGAA
Primer Fw WT ATXN3-GFP11 R282T 412  ATTTATATGCAATTATAAGGAACACTGGTTTACAGTTAGAAAATTAGGAA

pcDNA WT ATXN3-GFP11           1351 AACAGTGGTTTAACTTGAATTCTCTCTTGACGGGTCCAGAATTAATATCA
Primer Fw WT ATXN3-GFP11 R282T 462  AACAGTGGTTTAACTTGAATTCTCTCTTGACGGGTCCAGAATTAATATCA

pcDNA WT ATXN3-GFP11           1401 GATACATATCTTGCACTTTTCTTGGCTCAATTACAACAGGAAGGTTATTC
Primer Fw WT ATXN3-GFP11 R282T 512  GATACATATCTTGCACTTTTCTTGGCTCAATTACAACAGGAAGGTTATTC

pcDNA WT ATXN3-GFP11           1451 TATATTTGTCGTTAAGGGTGATCTGCCAGATTGCGAAGCTGACCAACTCC
Primer Fw WT ATXN3-GFP11 R282T 562  TATATTTGTCGTTAAGGGTGATCTGCCAGATTGCGAAGCTGACCAACTCC

pcDNA WT ATXN3-GFP11           1501 TACAGATGATTAGGGTCCAACAGATGCATCGACCAAAACTTATTGGAGAA
Primer Fw WT ATXN3-GFP11 R282T 612  TACAGATGATTAGGGTCCAACAGATGCATCGACCAAAACTTATTGGAGAA
```

```
pcDNA WT ATXN3-GFP11           1551 GAATTAGCACAACTAAAAGAGCAAAGAGTCCATAAAACAGACCTGGAACG
Primer Fw WT ATXN3-GFP11 R282T 662  GAATTAGCACAACTAAAAGAGCAAAGAGTCCATAAAACAGACCTGGAACG

pcDNA WT ATXN3-GFP11           1601 AGTGTTAGAAGCAAATGATGGCTCAGGAATGTTAGACGAAGATGAGGAGG
Primer Fw WT ATXN3-GFP11 R282T 712  AGTGTTAGAAGCAAATGATGGCTCAGGAATGTTAGACGAAGATGAGGAGG

pcDNA WT ATXN3-GFP11           1651 ATTTGCAGAGGGCTCTGGCACTAAGTCGCCAAGAAATTGACATGGAAGAT
Primer Fw WT ATXN3-GFP11 R282T 762  ATTTGCAGAGGGCTCTGGCACTAAGTCGCCAAGAAATTGACATGGAAGAT

pcDNA WT ATXN3-GFP11           1701 GAGGAAGCAGATCTCCGCAGGGCTATTCAGCTAAGTATGCAAGGTAGTTC
Primer Fw WT ATXN3-GFP11 R282T 812  GAGGAAGCAGATCTCCGCAGGGCTATTCAGCTAAGTATGCAAGGTAGTTC

pcDNA WT ATXN3-GFP11           1751 CAGAAACATATCTCAAGATATGACACAGACATCAGGTACAAATCTTACTT
Primer Fw WT ATXN3-GFP11 R282T 862  CAGAAACATATCTCAAGATATGACACAGACATCAGGTACAAATCTTACTT

pcDNA WT ATXN3-GFP11           1801 CAGAAGAGCTTCGGAAGAGACGAGAAGCCTACTTTGAAAAACAGCAGCAA
Primer Fw WT ATXN3-GFP11 R282T 912  CAGAAGAGCTTACGAAGAGACGAGAAGCCTACTTTGAAAAACAGCAGCAA

pcDNA WT ATXN3-GFP11           1851 AAGCAGCAACAGCAGCAGCAGCAGCAGCAGGGGGACCTATCAGGACA
Primer Fw WT ATXN3-GFP11 R282T 962  AAGCAGCAACAGCAGCAGCAGCAGCAGCAGGGGGACCTATCAGGACA
```

Figure 63 - Sequence alignment of the pcDNA WT ATXN3 R282T NLS-GFP11 vector with primers detailed in table 2, obtained from SnapGene. Pink: inserted primers with mutation. Blue: correct insertion of NLS mutation.

pcDNA GFP10 – Mutant ATXN3 R282T NLS vector alignment






```
pcDNA GFP10-Mutant ATXN3           987  ATGGGCGACCTGCC
Primer Fw GFP10-Mutant ATXN3 R282T 17   NTGGGCGACCTGCC
Primer Rv GFP10-Mutant ATXN3 R282T 0    -------------

pcDNA GFP10-Mutant ATXN3           1001 CGACGACCACTACCTGTCCACCCAGACCATCCTGAGCAAGGACCTGAACA
Primer Fw GFP10-Mutant ATXN3 R282T 31   CGACGACCACTACCTGTCCACCCAGACCATCCTGAGCAAGGACCTGAACA
Primer Rv GFP10-Mutant ATXN3 R282T 0    ------------------------------------------------

pcDNA GFP10-Mutant ATXN3           1051 TCGATGGTGGCGGTGGCTCTGGAGGTGGTGGGTCCTCCGGCGGCGGTGGA
Primer Fw GFP10-Mutant ATXN3 R282T 81   TCGATGGTGGCGGTGGCTCTGGAGGTGGTGGGTCCTCCGGCGGCGGTGGA
Primer Rv GFP10-Mutant ATXN3 R282T 0    ------------------------------------------------

pcDNA GFP10-Mutant ATXN3           1101 TCTAGAATGGAGTCCATCTTCCACGAGAAACAAGAAGGCTCACTTTGTGC
Primer Fw GFP10-Mutant ATXN3 R282T 131  TCTAGAATGGAGTCCATCTTCCACGAGAAACAAGAAGGCTCACTTTGTGC
Primer Rv GFP10-Mutant ATXN3 R282T 0    ------------------------------------------------

pcDNA GFP10-Mutant ATXN3           1151 TCAACATTGCCTGAATAACTTATTGCAAGGAGAATATTTTAGCCCTGTGG
Primer Fw GFP10-Mutant ATXN3 R282T 181  TCAACATTGCCTGAATAACTTATTGCAAGGAGAATATTTTAGCCCTGTGG
Primer Rv GFP10-Mutant ATXN3 R282T 1    -----------------------------------TTTAGCCCTG-GG

pcDNA GFP10-Mutant ATXN3           1201 AATTATCCTCAATTGCACATCAGCTGGATGAGGAGGAGAGGATGAGAATG
Primer Fw GFP10-Mutant ATXN3 R282T 231  AATTATCCTCAATTGCACATCAGCTGGATGAGGAGGAGAGGATGAGAATG
Primer Rv GFP10-Mutant ATXN3 R282T 13   AATTTTCCTCAATT-CCCATTAGC-GGGTGGGGGGGAG-GGATG-GAATG

pcDNA GFP10-Mutant ATXN3           1251 GCAGAAGGAGGAGTTACTAGTGAAGATTATCGCACGTTTTTACAGCAGCC
Primer Fw GFP10-Mutant ATXN3 R282T 281  GCAGAAGGAGGAGTTACTAGTGAAGATTATCGCACGTTTTTACAGCAGCC
Primer Rv GFP10-Mutant ATXN3 R282T 59   GCAG-AGGGGGAGTTACTAGGGAAGATTATCGCCCGTTTTTACAGCAGCT

pcDNA GFP10-Mutant ATXN3           1301 TTCTGGAAATATGGATGACAGTGGTTTTTTCTCTATTCAGGTTATAAGCA
Primer Fw GFP10-Mutant ATXN3 R282T 331  TTCTGGAAATATGGATGACAGTGGTTTTTTCTCTATTCAGGTTATAAGCA
Primer Rv GFP10-Mutant ATXN3 R282T 108  TTC-GGAAANAAGGATGACAGTGGTTTTTTCTCTATCC-GGTTATAAGCA
```

```
pcDNA GFP10-Mutant ATXN3           1351 ATGCCTTGAAAGTTTGGGGTTTAGAACTAATCCTGTTCAACAGTCCAGAG
Primer Fw GFP10-Mutant ATXN3 R282T 381  ATGCCTTGAAAGTTTGGGGTTTAGAACTAATCCTGTTCAACAGTCCAGAG
Primer Rv GFP10-Mutant ATXN3 R282T 156  A-GCCCTNAAAGTTTGGGGTTTAGAACTAATCCTGTTCAACAGTCCAGAG

pcDNA GFP10-Mutant ATXN3           1401 TATCAGAGGCTCAGGATCGATCCTATAAATGAAAGATCATTTATATGCAA
Primer Fw GFP10-Mutant ATXN3 R282T 431  TATCAGAGGCTCAGGATCGATCCTATAAATGAAAGATCATTTATATGCAA
Primer Rv GFP10-Mutant ATXN3 R282T 205  TATCAGAGGCTCAGGATCGATCCTATAAATGAAAGATCATTTATATGCAA

pcDNA GFP10-Mutant ATXN3           1451 TTATAAGGAACACTGGTTTACAGTTAGAAAATTAGGAAAACAGTGGTTTA
Primer Fw GFP10-Mutant ATXN3 R282T 481  TTATAAGGAACACTGGTTTACAGTTAGAAAATTAGGAAAACAGTGGTTTA
Primer Rv GFP10-Mutant ATXN3 R282T 255  TTATAAGGAACACTGGTTTACAGTTAGAAAATTAGGAAAACAGTGGTTTA

pcDNA GFP10-Mutant ATXN3           1501 ACTTGAATTCTCTCTTGACGGGTCCAGAATTAATATCAGATACATATCTT
Primer Fw GFP10-Mutant ATXN3 R282T 531  ACTTGAATTCTCTCTTGACGGGTCCAGAATTAATATCAGATACATATCTT
Primer Rv GFP10-Mutant ATXN3 R282T 305  ACTTGAATTCTCTCTTGACGGGTCCAGAATTAATATCAGATACATATCTT

pcDNA GFP10-Mutant ATXN3           1551 GCACTTTTCTTGGCTCAATTACAACAGGAAGGTTATTCTATATTTGTCGT
Primer Fw GFP10-Mutant ATXN3 R282T 581  GCACTTTTCTTGGCTCAATTACAACAGGAAGGTTATTCTATATTTGTCGT
Primer Rv GFP10-Mutant ATXN3 R282T 355  GCACTTTTCTTGGCTCAATTACAACAGGAAGGTTATTCTATATTTGTCGT

pcDNA GFP10-Mutant ATXN3           1601 TAAGGGTGATCTGCCAGATTGCGAAGCTGACCAACTCCTGCAGATGATTA
Primer Fw GFP10-Mutant ATXN3 R282T 631  TAAGGGTGATCTGCCAGATTGCGAAGCTGACCAACTCCTGCAGATGATTA
Primer Rv GFP10-Mutant ATXN3 R282T 405  TAAGGGTGATCTGCCAGATTGCGAAGCTGACCAACTCCTGCAGATGATTA

pcDNA GFP10-Mutant ATXN3           1651 GGGTCCAACAGATGCATCGACCAAAACTGATTGGCGAAGAACTGGCTCAA
Primer Fw GFP10-Mutant ATXN3 R282T 681  GGGTCCAACAGATGCATCGACCAAAACTGATTGGCGACTAACTGGCTCAA
Primer Rv GFP10-Mutant ATXN3 R282T 455  GGGTCCAACAGATGCATCGACCAAAACTGATTGGCGAAGAACTGGCTCAA

pcDNA GFP10-Mutant ATXN3           1701 CTGAAAGAACAGCGTGTGCATAAGACCGACCTGGAACGTGTCCTGGAAGC
Primer Fw GFP10-Mutant ATXN3 R282T 731  CTGAAAGAACAGCGTGTGCATAACACCGACCTGTAACGTGTCCTGGAAGC
Primer Rv GFP10-Mutant ATXN3 R282T 505  CTGAAAGAACAGCGTGTGCATAAGACCGACCTGGAACGTGTCCTGGAAGC

pcDNA GFP10-Mutant ATXN3           1751 AAATGACGGCAGCGGCATGCTGGATGAAGACGAAGAAGATCTGCAGCGTG
Primer Fw GFP10-Mutant ATXN3 R282T 781  AAATGATGGTAGCGCCATGCTGGATGAAGATGAACAATATCTGCTGTGTG
Primer Rv GFP10-Mutant ATXN3 R282T 555  AAATGACGGCAGCGGCATGCTGGATGAAGACGAAGAAGATCTGCAGCGTG

pcDNA GFP10-Mutant ATXN3           1801 CCCTGGCACTGTCTCGTCAGGAAATTGATATGGAAGACGAAGAAGCAGAT
Primer Fw GFP10-Mutant ATXN3 R282T 831  CCCTGCCACCGTTCCTCTGCAA--------------------------
Primer Rv GFP10-Mutant ATXN3 R282T 605  CCCTGGCACTGTCTCGTCAGGAAATTGATATGGAAGACGAAGAAGCAGAT

pcDNA GFP10-Mutant ATXN3           1851 CTGCGTCGCGCTATTCAGCTGTCAATGCAGGGCAGCTCTCGTAACATCTC
Primer Fw GFP10-Mutant ATXN3 R282T 853  ------------------------------------------------
Primer Rv GFP10-Mutant ATXN3 R282T 655  CTGCGTCGCGCTATTCAGCTGTCAATGCAGGGCAGCTCTCGTAACATCTC
```

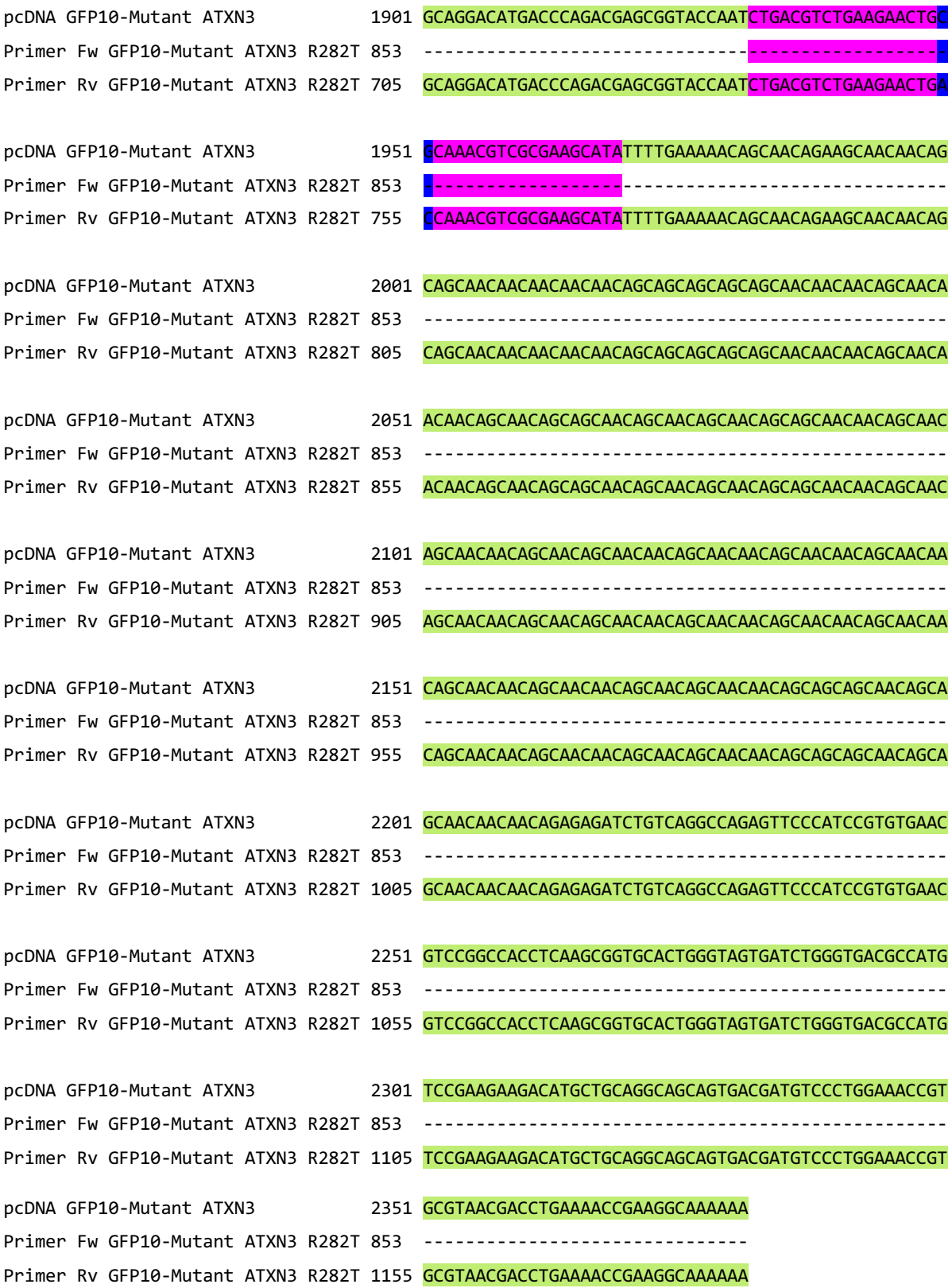
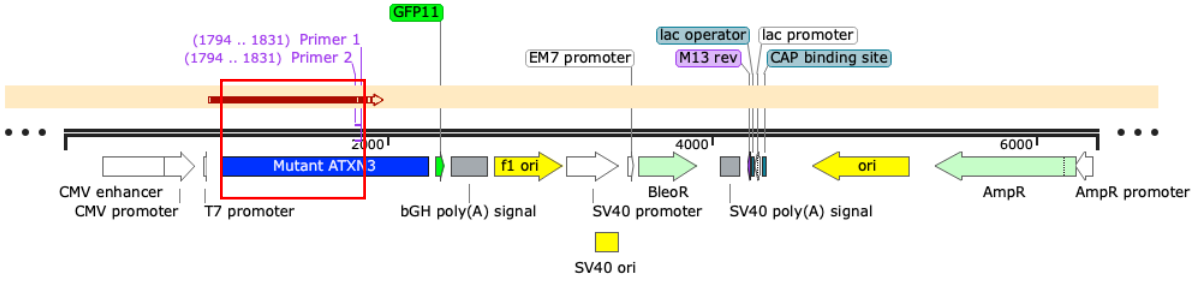


Figure 64 - Sequence alignment of the pcDNA GFP10 – Mutant ATXN3 R282T NLS vector with primers detailed in table 2, obtained from SnapGene® software (from Insightful Science; available at snapgene.com). Pink: inserted primers with mutation. Blue: correct insertion of NLS mutation. Red – Unconfirmed sequenced nucleotide.

pcDNA Mutant ATXN3 R282T NLS – GFP11 vector alignment




```
pcDNA Mutant ATXN3-GFP11         969  ATGGAGTCCATCTTCCACGAGAAACAAGAAGG
Primer Fw Mut ATXN3-GFP11 R282T  78   ATGGAGTCCATCTTCCACGAGAAACAAGAAGG

pcDNA Mutant ATXN3-GFP11         1001 CTCACTTTGTGCTCAACATTGCCTGAATAACTTATTGCAAGGAGAATATT
Primer Fw Mut ATXN3-GFP11 R282T  110  CTCACTTTGTGCTCAACATTGCCTGAATAACTTATTGCAAGGAGAATATT

pcDNA Mutant ATXN3-GFP11         1051 TTAGCCCTGTGGAATTATCCTCAATTGCACATCAGCTGGATGAGGAGGAG
Primer Fw Mut ATXN3-GFP11 R282T  160  TTAGCCCTGTGGAATTATCCTCAATTGCACATCAGCTGGATGAGGAGGAG

pcDNA Mutant ATXN3-GFP11         1101 AGGATGAGAATGGCAGAAGGAGGAGTTACTAGTGAAGATTATCGCACGTT
Primer Fw Mut ATXN3-GFP11 R282T  210  AGGATGAGAATGGCAGAAGGAGGAGTTACTAGTGAAGATTATCGCACGTT

pcDNA Mutant ATXN3-GFP11         1151 TTTACAGCAGCCTTCTGGAAATATGGATGACAGTGGTTTTTTCTCTATTC
Primer Fw Mut ATXN3-GFP11 R282T  260  TTTACAGCAGCCTTCTGGAAATATGGATGACAGTGGTTTTTTCTCTATTC
```

```
pcDNA Mutant ATXN3-GFP11        1201 AGGTTATAAGCAATGCCTTGAAAGTTTGGGGTTTAGAACTAATCCTGTTC
Primer Fw Mut ATXN3-GFP11 R282T 310  AGGTTATAAGCAATGCCTTGAAAGTTTGGGGTTTAGAACTAATCCTGTTC

pcDNA Mutant ATXN3-GFP11        1251 AACAGTCCAGAGTATCAGAGGCTCAGGATCGATCCTATAAATGAAAGATC
Primer Fw Mut ATXN3-GFP11 R282T 360  AACAGTCCAGAGTATCAGAGGCTCAGGATCGATCCTATAAATGAAAGATC

pcDNA Mutant ATXN3-GFP11        1301 ATTTATATGCAATTATAAGGAACACTGGTTTACAGTTAGAAAATTAGGAA
Primer Fw Mut ATXN3-GFP11 R282T 410  ATTTATATGCAATTATAAGGAACACTGGTTTACAGTTAGAAAATTAGGAA

pcDNA Mutant ATXN3-GFP11        1351 AACAGTGGTTTAACTTGAATTCTCTCTTGACGGGTCCAGAATTAATATCA
Primer Fw Mut ATXN3-GFP11 R282T 460  AACAGTGGTTTAACTTGAATTCTCTCTTGACGGGTCCAGAATTAATATCA

pcDNA Mutant ATXN3-GFP11        1401 GATACATATCTTGCACTTTTCTTGGCTCAATTACAACAGGAAGGTTATTC
Primer Fw Mut ATXN3-GFP11 R282T 510  GATACATATCTTGCACTTTTCTTGGCTCAATTACAACAGGAAGGTTATTC

pcDNA Mutant ATXN3-GFP11        1451 TATATTTGTCGTTAAGGGTGATCTGCCAGATTGCGAAGCTGACCAACTCC
Primer Fw Mut ATXN3-GFP11 R282T 560  TATATTTGTCGTTAAGGGTGATCTGCCAGATTGCGAAGCTGACCAACTCC

pcDNA Mutant ATXN3-GFP11        1501 TGCAGATGATTAGGGTCCAACAGATGCATCGACCAAAACTGATTGGCGAA
Primer Fw Mut ATXN3-GFP11 R282T 610  TGCAGATGATTAGGGTCCAACAGATGCATCGACCAAAACTGATTGGCGAA

pcDNA Mutant ATXN3-GFP11        1551 GAACTGGCTCAACTGAAAGAACAGCGTGTGCATAAGACCGACCTGGAACG
Primer Fw Mut ATXN3-GFP11 R282T 660  GAACTGGCTCAACTGAAAGAACAGCGTGTGCATAAGACCGACCTGGAACG

pcDNA Mutant ATXN3-GFP11        1601 TGTCCTGGAAGCAAATGACGGCAGCGGCATGCTGGATGAAGACGAAGAAG
Primer Fw Mut ATXN3-GFP11 R282T 710  TGTCCTGGAAGCAAATGACGGCAGCGGCATGCTGGATGAAGACGAAGAAG

pcDNA Mutant ATXN3-GFP11        1651 ATCTGCAGCGTGCCCTGGCACTGTCTCGTCAGGAAATTGATATGGAAGAC
Primer Fw Mut ATXN3-GFP11 R282T 760  ATCTGCAGCGTGCCCTGGCACTGTCTCGTCAGGAAATTGATATGGAAGAC

pcDNA Mutant ATXN3-GFP11        1701 GAAGAAGCAGATCTGCGTCGCGCTATTCAGCTGTCAATGCAGGGCAGCTC
Primer Fw Mut ATXN3-GFP11 R282T 810  GAAGAAGCAGATCTGCGTCGCGCTATTCAGCTGTCAATGCAGGGCAGCTC

pcDNA Mutant ATXN3-GFP11        1751 TCGTAACATCTCGCAGGACATGACCCAGACGAGCGGTACCAATCTGACGT
Primer Fw Mut ATXN3-GFP11 R282T 860  TCGTAACATCTCGCAGGACATGACCCAGACGAGCGGTACCAATCTGACGT

pcDNA Mutant ATXN3-GFP11        1801 CTGAAGAACTGCGCAAACGTCGCGAAGCATATTTTGAAAAACAGCAACAG
Primer Fw Mut ATXN3-GFP11 R282T 910  CTGAAGAACTGACCAAACGTCGCGAAGCATATTTTGAAAAACAGCAACAG
```

Figure 65 - Sequence alignment of the pcDNA Mutant ATXN3 R282T NLS – GFP11 vector with primers detailed in table 2, obtained from SnapGene® software (from Insightful Science; available at snapgene.com). Pink: inserted primers with mutation. Blue: correct insertion of NLS mutation.

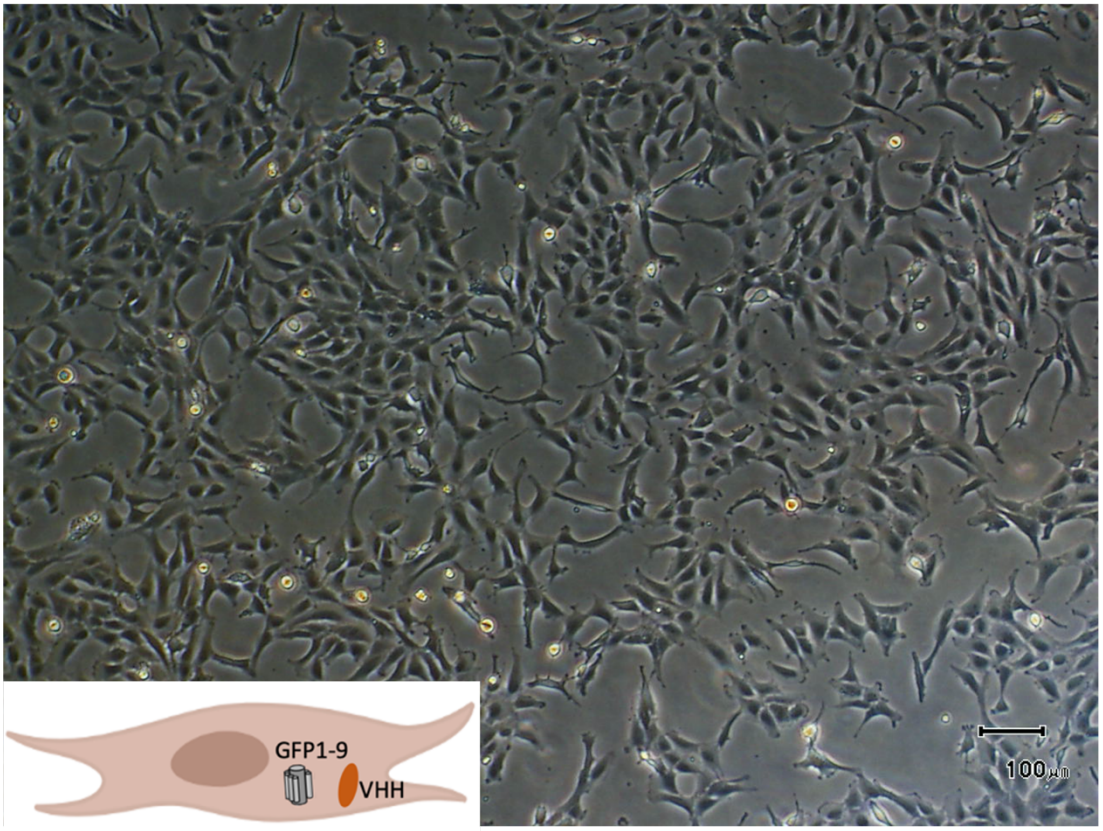


Figure 66 – MRC5-SV (immortalized normal pulmonary human fibroblasts) cells expressing GFP1-9 and one single-domain antibody based on camelid heavy-chain antibodies (VHH or nanobody).

Table 3 - Percentage of GFP positive cells obtained by flow cytometry.

Transfection	Percentage of GFP positive cells (%)
Only cells	0
GFP10-Zipper-GFP11	19.8
PIC111-ATXN3	19.9
WT ATXN3-GFP10	0
WT ATXN3-GFP11	0
GFP10 + GFP11	0
WT ATXN3-GFP10 + GFP11	0
WT ATXN3-GFP11 + GFP10	0
Mutant ATXN3-GFP10 + GFP11	0
Mutant ATXN3-GFP11 + GFP10	0
ATXN3-GFP10 + 9G8-GFP11	9.73
ATXN3-GFP10 + Tubulin-GFP11	4.62
ATXN3-GFP10 + hHR23A-GFP11	19.0
ATXN3-GFP10 + GCN4-GFP11	1.85
ATXN3-GFP11 + 9G8-GFP10	7.22
ATXN3-GFP11 + Tubulin-GFP10	2.85
ATXN3-GFP11 + hHR23A-GFP10	24.6
ATXN3-GFP11 + GCN4-GFP10	1.08
ATXN3-GFP10 + ATXN3-GFP11	13.8
Mutant ATXN3-GFP10 + 9G8-GFP11	1.18
Mutant ATXN3-GFP10 + Tubulin-GFP11	1.97
Mutant ATXN3-GFP10 + hHR23A-GFP11	1.15
Mutant ATXN3-GFP10 + GCN4-GFP11	0.25
Mutant ATXN3-GFP11 + 9G8-GFP10	4.18
Mutant ATXN3-GFP11 + Tubulin-GFP10	1.26
Mutant ATXN3-GFP11 + hHR23A-GFP10	17.2
Mutant ATXN3-GFP11 + GCN4-GFP10	1.05
Mutant ATXN3-GFP10 + Mutant ATXN3-GFP11	0.45