1 2 9 0

UNIVERSIDADE D
COIMBRA

André Filipe Vieira da Conceição

CRISPR/CAS9 AS A TOOL FOR GENE THERAPY IN MACHADO-JOSEPH DISEASE: SILENCING *ATXN3* AND CAG EXPANSION CORRECTION

Dissertação de Mestrado em Biologia Celular e Molecular com especialização em Neurobiologia, orientada pelo Doutor Carlos Adriano de Matos (Centro de Investigação em Biomedicina, Universidade do Algarve) e pela Professora Doutora Ana Luísa Carvalho (Departamento de Ciências da Vida, Faculdade de Ciências e Tecnologia, Universidade de Coimbra), apresentada ao Departamento de Ciências da Vida da Universidade de Coimbra.

Setembro de 2019

André Filipe Vieira da Conceição

CRISPR/Cas9 as a tool for gene therapy in Machado-Joseph disease: silencing *ATXN3* and CAG expansion correction

Dissertação de Mestrado em Biologia Celular e Molecular com especialização em Neurobiologia, orientada pelo Doutor Carlos Adriano de Matos (Centro de Investigação em Biomedicina, Universidade do Algarve) e pela Professora Doutora Ana Luísa Carvalho (Departamento de Ciências da Vida, Faculdade de Ciências e Tecnologia, Universidade de Coimbra), apresentada ao Departamento de Ciências da Vida da Universidade de Coimbra.

Setembro 2019

O trabalho aqui apresentado foi realizado no grupo de investigação liderado pelo Prof. Doutor Luís Pereira de Almeida, no Grupo de Vectores, Terapia Génica e Celular do CNC - Centro de Neurociências e Biologia Celular e do CIBB - Centro de Inovação em Biomedicina e Biotecnologia, Universidade de Coimbra, Portugal.

Este trabalho foi financiado pelo FEDER através do Programa Operacional Regional Centro 2020, do Programa Operacional de Fatores de Competitividade (COMPETE 2020) e por Fundos Nacionais através da FCT (Fundação para a Ciência e a Tecnologia) – projetos BrainHealth2020 (CENTRO-01-0145-FEDER-000008), ViraVector (CENTRO-01-0145-FEDER- 022095), CortaCAGs (POCI-01-0145-FEDER-016719), SpreadSilencing (POCI-01-0145-FEDER-029716 e POCI-01-0145-FEDER-007440), bem como pelos projetos EXPL/NEU-NMC/0331/2012 (FCT), POCI-01-0145-FEDER-032309 e pelos SynSpread, ESMI e ModelPolyQ no âmbito do EU Joint Programme - Neurodegenerative Disease Research (JPND), os dois últimos cofinanciados pelo programa H2020 da União Europeia, GA No. 643417; ainda pela "National Ataxia Foundation" (USA), pelo American Portuguese Biomedical Research Fund (APBRF) e pelo Richard Chin and Lily Lock Machado Joseph Disease Research Fund.

The work presented here was carried out at the research group led by Prof. Luís Pereira de Almeida, Group of Vectors, Gene and Cell Therapy of CNC - Center for Neuroscience and Cell Biology and of CIBB- Center for Innovative Biomedicine and Biotechnology, University of Coimbra, Portugal.

This work was funded by the ERDF through the Regional Operational Program Center 2020, Competitiveness Factors Operational Program (COMPETE 2020) and National Funds through FCT (Foundation for Science and Technology) - BrainHealth2020 projects (CENTRO-01-0145-FEDER-000008), ViraVector (CENTRO-01-0145-FEDER-022095), CortaCAGs (PTDC/NEU-NMC/0084/2014|POCI-01-0145-FEDER-016719), SpreadSilencing (POCI-01-0145-FEDER-029716, POCI-01-0145-FEDER-032309) as well as SynSpread, ESMI and ModelPolyQ under the EU Joint Program - Neurodegenerative Disease Research (JPND), the last two co-funded by the European Union H2020 program, GA No.643417; by the National Ataxia Foundation (USA), the American Portuguese Biomedical Research Fund (APBRF) and the Richard Chin and Lily Lock Machado-Joseph Disease Research Fund.

À minha Tia Amélia

The world was young, the mountains green,
No stain yet on the Moon was seen,
No words were laid on stream or stone
When Durin woke and walked alone.
He named the nameless hills and dells;
He drank from yet untasted wells;
He stooped and looked in Mirrormere,
And saw a crown of stars appear,
As gems upon a silver thread,
Above the shadow of his head.

The world was fair, the mountains tall,
In Elder Days before the fall
(…)

A king he was on carven throne
In many-pillared halls of stone
With golden roof and silver floor,
And runes of power upon the door.
The light of sun and star and moon
In shining lamps of crystal hewn
Undimmed by cloud or shade of night
There shone for ever fair and bright.

There hammer on the anvil smote,
There chisel clove, and graver wrote;
There forged was blade, and bound was hilt;
The delver mined, the mason built.
There beryl, pearl, and opal pale,
And metal wrought like fishes' mail,
Buckler and corslet, axe and sword,
And shining spears were laid in hoard.
Unwearied then were Durin's folk;
Beneath the mountains music woke:
The harpers harped, the minstrels sang,
And at the gates the trumpets rang.

The world is grey, the mountains old,
The forge's fire is ashen-cold;
No harp is wrung, no hammer falls:
The darkness dwells in Durin's halls;
The shadow lies upon his tomb
(…)
But still the sunken stars appear
In dark and windless Mirrormere;
There lies his crown in water deep,
Till Durin wakes again from sleep.

— J.R.R. Tolkien

"If you only do what you can do,
you will never be more than you are now."

— Master Shifu

Agradecimentos

Aos ombros de gigantes – é assim que se progride e se caminha em direção ao futuro. Se é verdade que trabalhei muito e me esforcei para completar esta etapa, também é verdade que sozinho nunca o conseguiria. Assim sendo, é para mim uma honra agradecer a todos aqueles que, de uma forma ou outra, contribuíram para a realização deste trabalho.

Professor Doutor Luís Almeida, as primeiras palavras são para si. Obrigado pela oportunidade que me deu de trabalhar consigo e com o seu incrível grupo. Ao acolher-me no seu laboratório permitiu que eu começasse a caminhar no mundo da ciência. Foi a minha primeira experiência num laboratório e fico-lhe muito agradecido pelo voto que confiança.

Doutor Carlos! Ora bem, como vou eu agradecer-te? Um simples obrigado nunca será suficiente para demonstrar o quão grato estou por me teres acompanhado nesta façanha. Agradeço-te em primeiro lugar, por aquilo que de mais valor me poderias ter dado: conhecimento. Sem dúvida, a partilha do conhecimento tem sido desde o início da humanidade, aquilo que nos fez evoluir. E tu, com toda a certeza, tens vindo também a deixar essa marca. Ter-te como orientador foi do mais enriquecedor e gratificante que já vivi. Não falo apenas por me teres ensinado quase tudo o que sei fazer no laboratório, mas também pelo facto de ser incrível ter como orientador um “poço de conhecimento”, especialmente para uma pessoa como eu que nunca saiu da idade dos “porquês”. 9999 perguntas num dia, era um dia tranquilo para ti! Contigo é possível discutir ciência, história, filosofias a vida, enfim, tudo. Não posso também deixar de te agradecer teres-me contado, como que em jeito aprovação, que irias partir para uma nova aventura e de me teres dado a tua palavra que continuarias a orientar-me independentemente da distância de 440 Km. Pois bem, quero todos saibam que o Doutor Carlos cumpriu honrosamente com a sua palavra! Continuaste a aconselhar-me, a guiar-me, a ajudar-me e acima de tudo, a dar-me força quando errava. Não estares fisicamente presente no laboratório ensinou-me a pensar, a aprender, a planear e a superar dificuldades “sozinho”. Enfim, ajudou-me a crescer e obrigado por isso. Não posso também deixar de referenciar o pato assado, moelas, Tropical, Largo do Carmo, DnD, MTG entre muitos outros – porque nem só de ciência vive o homem. Fizeste um bom trabalho e espero que estejas orgulhoso de mim. Obrigado.

Professora Ana Luísa, antes de mais quero agradecer-lhe por ter aceite ser minha orientadora, espero ter estado à altura, sei que não é fácil. Sinto-me um afortunado por ter tido a oportunidade de assistir às suas incríveis aulas. Foi sem dúvida uma das melhores professoras que tive e por isso também lhe fico agradecido. Ser seu aluno fez crescer o meu gosto pela ciência e sobretudo fez-me sentir que estava no lugar certo. Tenho também de agradecer a disponibilidade que sempre mostrou para comigo dentro e fora da sala de aula. Para mim, a Professora Ana Luísa é exemplo a seguir! Obrigado.

Professor Carlos Duarte, não, não me esqueci! Muito do que sei de Neurobiologia devo-o a si. Obrigado por responder às minhas 1001 perguntas, que com toda a certeza 1000 delas eram descabidas. Quero também salientar o subtil bom humor com que dá as aulas, fazendo com fossem ainda mais interessantes. Não posso deixar passar também a disponibilidade que demonstrou para apoiar o meu futuro.

Quero deixar o meu sincero apreço àqueles que me deram a mão após o Carlos ter embarcado numa nova aventura. Dina, foste realmente um grande pilar para mim dentro do laboratório. Se as experiências do cresil estão aqui a ti o devo. Se hoje sei purificar proteína e RNA de um cerebelo a ti o devo. Obrigado por teres perdido um ou outro fim de semana a orientar-me. Sem a tua ajuda este trabalho teria sido bem mais pobre. Muito obrigado! Agradeço ao David pelas horas que "perdeu" a tentar desvendar aquele quebra-cabeças que foi o *piggyBac*. Agradeço também teres-me ajudado e ensinado sempre que precisava, especialmente nos qPCR. Tenho indiscutivelmente de agradecer à Sara por todo o trabalho que já desenvolveu e tem vido a desenvolver que ajudou em muito a realização deste trabalho. Um obrigado muito especial à D. Rosário pela sua disponibilidade e simpatia que sempre demonstrou comigo. Perdeu muitas horas a ouvir-me, a aconselhar-me e a ajudar-me, sobretudo nos meus PCR e eletroforeses. Os seus anos de experiência foram uma mais valia à realização deste trabalho.

Agradeço também aos meus companheiros de "guerra", Fred, Valada, Mário e Passos pelos bons momentos dentro e fora do laboratório. UNO after lunch, é algo que não tem preço. Fred e Mário, obrigado por me ajudarem a agilizar as minhas experiências sempre que vos pedia. Foram muitas as gargalhadas e os bons momento que vocês todos me proporcionaram.

Uma palavra de gratidão ao professor Clévio por me ter aconselhado este mestrado e por ter apoiado a minha candidatura. Sem ele muito provavelmente não me encontraria aqui hoje. Espero não o ter desiludido e ter cumprido com as espectativas. Obrigado!

Um agradecimento muito especial ao Dr. Gordo, que sempre foi muito melhor amigo do que aquilo que eu merecia. É incrível pensar como é que já nos conhecemos há 20 anos e ainda é possível contar contigo para tudo. Espero que daqui a mais 20 anos ainda estejas presente na minha vida. Gosto muito de ti, para mim és como um irmão. (Gostas mais de pão ou de manteiga? (…) e se tivesses fome?).

Um obrigado muito especial ao Toino e à D. Fátima que há quase mais de 5 anos me têm tratado como um filho. Com a ajuda deles todo este percurso tem vindo a ser muito mais fácil. Não tenho como agradecer o que têm feito por mim. Na casa deles sempre me fizeram sentir como se estivesse na minha. Estou-lhes profundamente agradecido.

Por fim, mas não menos especial – à semelhança do que acontece com a parte mais importante de uma boa casa, os seus alicerces, que ficam na base sem serem vistos – agradeço à Catarina por toda a dedicação, apoio e carinho que meu deu e dá. Foi fundamental para cumprir mais esta etapa. Espero que o futuro seja risonho e próspero para nós. Tens sido das pessoas mais importantes da minha vida. Agradeço à minha irmã, pela paciência que tem comigo especialmente porque quase nunca lhe atender o telefone à primeira. Agradeço também teres sempre a porta aberta para me receber. E claro, o agradecimento que para mim é mais importante vai para a minha Tia, que sem ela nunca seria o que sou hoje. Tem feito um esforço hercúleo para eu estar aqui hoje. Devo tudo à minha Tia, e nunca vou ser capaz de lhe retribuir o que tem feito por mim. Sempre me apoiou e sempre me deu a oportunidade de seguir os meus sonhos e de fazer as minhas escolhas. Para mim é uma Mãe.

VA FAILL!

Table of contents

Resumo 1

Abstract 3

List of abbreviation 5

List of images 9

List of tables 11

Chapter 1 – Introduction 13

1.1 Machado-Joseph disease 15

  1.1.1 Disease description and clinical features 15

  1.1.2 Epidemiology 16

  1.1.3 Disease aetiology 16

  1.1.4 Atxn-3 protein 18

  1.1.5 Mutant atxn-3 protein and MJD pathogenesis 20

1.2 Models of Machado-Joseph disease 21

  1.2.1 Importance of human disease models 21

  1.2.2 Cell cultures as disease models 22

  1.2.3 Animal models of Machado-Joseph disease 24

  1.2.4 *Caenorhabditis elegans* models 25

  1.2.5 *Drosophila melanogaster* models 27

  1.2.6 Zebrafish model 28

  1.2.7 Marmoset – A non-human primate model 29

  1.2.8 Lentiviral rodent models 30

  1.2.9 Transgenic mouse models expressing truncated human ataxin-3 31

  1.2.10 Knock-in mouse models 32

  1.2.11 Transgenic mouse models expressing full-length human ataxin-3 .. 33

1.3 Gene-based therapies for Machado-Joseph disease 37

1.3.1 Current perspectives 37

1.3.2 Machado-Joseph gene therapies based on RNA-targeting 38

1.3.4 Toxicity of expanded CAG sequence-bearing mRNA molecules 41

1.4 CRISPR/Cas9-based gene therapy tools 41

1.4.1 The biology of CRISPR/Cas9 41

1.4.2 CRISPR/Cas9 as a tool for gene editing 45

1.4.3 CRISPR/Cas9 as a tool for gene expression regulation 48

1.4.4 Expanding CRISPR/Cas9 potential 49

1.4.5 Generation of mouse models using CRISPR/Cas9 52

1.4.6 CRISPR/Cas9 towards Machado-Joseph disease 53

1.5 Objectives 55

Chapter 2 – Materials and methods 59

Plasmids 61

Plasmid amplification 61

Single guided RNAs (sgRNA) design 61

sgRNA cloning 62

Design and generation of donor templates 63

Cell culture and transfection 65

Preparation of cell lysates 65

Antibodies 66

HEK cells genotyping 66

Lentiviral production 68

Animal experimentation 68

Lentiviral stereotaxic injections 69

Behavioural tests 69

Brain tissue collection 70

Histological processing and cresyl violet staining 70

Cerebellar RNA and protein extraction 71

Western blot 71

qPCR 72

Microscopy 72

Statistical analysis 73

Chapter 3 – Results and discussion 75

3.1 Using CRISPR/Cas9 to overexpand and retract the *ATXN3* CAG repetition 77

3.1.1 Overexpanding CAG trinucleotide repeats in vitro 77

3.1.2 Correcting the ATXN3 gene in vitro 82

3.2 dCas9-KRAB as a tool for in vivo silencing of the *ATXN3* gene 94

Chapter 4 – Conclusion 103

Conclusion and future perspectives 105

References 109

Resumo

A doença de Machado-Joseph (DMJ), também conhecida por ataxia espinocerebelosa tipo 3, é uma doença neurodegenerativa e a forma mais comum de ataxia hereditária dominante no mundo. É uma doença rara, porém tem uma prevalência significativa em algumas regiões de Portugal, nomeadamente no arquipélago dos Açores. Os sintomas de DMJ incluem perda de coordenação motora bem como outros sinais neurológicos, sendo o cerebelo a região do cérebro mais afetada. A DMJ surge de uma expansão anormal de trinucleótidos CAG no exão 10 do gene *ATXN3*, que codifica para a proteína ataxina-3 (atxn-3). No contexto da doença, a atxn-3 contém uma sequência de glutaminas (poliQ) anormalmente longa. As funções biológicas da proteína atxn-3 não são totalmente conhecidas, contudo tem sido descrito que a atxn-3 expandida adquire uma função tóxica que desregula o normal funcionamento de diversos sistemas celulares, levando à morte neuronal. Recentemente, tem sido sugerido que tratos CAG anormais presentes em transcritos de mRNA podem também ser tóxicos. Tal como em outras doenças de poliglutaminas, um maior número repetições CAG tem sido associado a sintomas mais severos, com um aparecimento mais precoce.

Atualmente, não há terapias capazes de tratar ou atrasar o curso da doença, e, portanto, a DMJ continua a ser fatal. Algumas estratégias terapêuticas que têm sido testadas atuam ao nível pós-transcricional e são por isso incapazes de prevenir os putativos efeitos tóxicos dos transcritos de mRNA que contêm uma cadeia de CAGs expandida. Assim sendo, novas estratégias que atuem a um nível pré-transcricional podem vir a ser vantajosas.

O estudo de doenças humanas apenas consegue ser robusto quando é usado um modelo de doença adequado. Porém, os modelos de murganho presentemente usados para estudar a DMJ ou não reproduzem de forma fiel o contexto genético da doença ou apresentam sintomas pouco acentuados e uma progressão lenta que limitam a sua utilização.

Na última década, a área da edição genética tem tido desenvolvimentos extraordinários. O uso da CRISPR/Cas9 como uma ferramenta para edição genética trouxe a possibilidade de manipular o genoma virtualmente em qualquer *locus*, de uma forma barata, fácil e flexível. Além do mais, variantes da CRISPR/Cas9, como é o caso da

Cas9 cataliticamente inativa (dCas9), têm mostrado resultados promissores no que diz respeito à regulação da transcrição genética.

Na primeira parte deste trabalho pretendemos estabelecer uma estratégia *in vitro* para aperfeiçoar um modelo de DMJ em murganho designado por yeast artificial chromosome (YAC) MJD-Q84.2, que é considerado como o modelo que melhor recapitula os sintomas e o contexto genético humano em condições de doença. Neste trabalho, utilizámos ferramentas de edição genética para sobreexpandir o trato CAG do exão 10 do gene *ATXN3*, de 84 para 141 CAGs; ao aplicar esta estratégia ao modelo de murganho seria expectável que se criasse um modelo robusto, com um fenótipo mais severo e em que os sintomas se manifestassem mais cedo em comparação com o já existente modelo YAC MJD-Q84.2. Contudo, os resultados *in vitro* descritos neste trabalho mostraram que a estratégia usada parece ser ineficiente na sobreexpansão das repetições CAG em células HEK 293T.

Paralelamente, usando também a CRISPR/Cas9, foi testada uma estratégia para corrigir o gene *ATXN3* mutante, que pudesse ser usada para criar linhas celulares isogénicas derivadas de pacientes de DMJ. O estudo aqui apresentado, usando células HEK 293T, mostrou ser possível integrar um fragmento de DNA contendo 14 repetições do trinucleótido CAG e uma cassete de seleção no exão 10 do gene *ATXN3*. Contudo, esta abordagem não se mostrou totalmente fiável e que foi detetada integração independente da atividade da Cas9.

Na segunda parte deste trabalho foram avaliados os efeitos da dCas9-KRAB como estratégia de silenciamento a nível pré-transcricional do gene humano da *ATXN3* num modelo animal de DMJ. Os resultados *in vivos* mostraram que esta abordagem tem potencial terapêutico, melhorando a performance motora num modelo DJM transgénico severamente afetado.

Abstract

Machado-Joseph disease (MJD), also known as spinocerebellar ataxia type 3, is a neurodegenerative disorder considered to be the most common form of autosomal dominantly-inherited ataxia in the world. It is a rare disorder, although it has a significative prevalence in some regions of Portugal, especially in the archipelago of Azores. MJD symptoms include a loss of motor coordination and other neurological signs, with the cerebellum being the most affected brain region. MJD arises from an abnormal CAG trinucleotide expansion within the exon 10 of the human *ATXN3* gene, which encodes for a protein named ataxin-3 (atxn-3), that bears an aberrant polyglutamine (polyQ) tract in the disease context. The exact biological function of atxn-3 is not fully understood, but it has been described that, when expanded atxn-3 undergoes a toxic gain-of-function that deregulates normal cellular pathways leading to neuronal loss. Recently, it has been suggested that abnormal CAG tracts in mRNA transcripts might also be toxic. As with other polyglutamine diseases, greater numbers of CAG repeats are associated with earlier ages of onset and more severe symptoms.

Currently, no therapies capable of delaying or treating the disease are available, and MJD remains a fatal disorder. Many of the therapeutic strategies that have been tested act at a post-transcriptional level, being unable to prevent the putative toxicity of CAG expanded mRNA transcripts. Therefore, new strategies that act at a pre-transcriptional level may be advantageous.

However, disease studies can only be robust and feasible when the adequate disease model is used. Yet, the available mouse models used to study MJD either fail to properly mimic the actual genetic context of the disease or are of limited use due to slow and mild symptom progression.

In the last decade, the field of gene editing has been thriving. The use of the CRISPR/Cas9 system as a tool for gene editing brought the possibility of performing low-cost, flexible and easy genomic manipulation virtually at any *loci*. Moreover, CRISPR/Cas9 variants, such as the catalytic inactivated Cas9 (dCas9) have shown promising results regarding gene transcription regulation.

In the first part of this work, we planned to establish an *in vitro* strategy to refine the yeast artificial chromosome (YAC) MJD-Q84.2 mouse, which is the MJD model that best recapitulates the symptoms and the genetic context of the human condition. In this work we proposed to use gene editing tools to overexpand the *ATXN3* CAG tract in the exon 10, from 84 to 141 CAG; applying this strategy to the model would be expected to generate a robust mouse model that displayed a more severe phenotype and an earlier disease onset, comparing to the YAC MJD-Q84.2 model. However, the *in vitro* results here presented, using HEK 293T cells, showed that the strategy used was inefficient at overexpanding the CAG repeats.

At the same time, a CRISPR/Cas9-based strategy for correcting the pathogenic *ATXN3* gene that could be used to generate isogenic MJD patient-derived cell lines was also tested. The study here presented, using HEK 293T cells, showed that it is possible to integrate a DNA fragment containing 14 CAG repeats plus a selection cassette in the exon 10 of the *ATXN3* gene. However, this strategy was not completely infallible, and integration seemed to be independent from Cas9 activity.

In the second part of this work, we assessed the effects of dCas9-KRAB as a strategy to pre-transcriptionally silence mutant *ATXN3* in a MJD mouse model. *In vivo* results showed that this approach has a therapeutic potential to improve motor performance in a severely-impaired transgenic mouse model of MJD.

Keywords: Machado-Joseph disease (MJD), SCA3, atxn-3, CRISPR/Cas9, gene editing, homology-directed repair, YAC MJD-Q84.2, KRAB, gene silencing.

List of abbreviation

4-HT	4-hydroxytamoxifen
AAVS1	adeno-associated virus integration site 1
AID	cytosine deaminase
ASOs	antisense oligonucleotides
ATCC	American Type Culture Collection
ATXN3 / atnx-3	antaxin-3
AVV	adeno-associated virus
BCA	bicinchoninic acid
bp	base pair
CAA	cytosine-adenine-adenine
CAG	cytosine-adenine-guanine
cAMP	cyclic adenosine monophosphate
Cas2	CRISPR associated protein 2
Cas9	CRISPR associated protein 9
CD71	transferrin receptor I (known as cluster of differentiation 71)
cDNA	coding DNA
Chk1	checkpoint kinase 1
C-Myc	cellular myelocytomatosis factor
CopGFP	copepod GFP
CpG	cytosine-guanine
CREBBP	cAMP response element-binding-binding protein
CRISPR	clustered regularly interspaced short palindromic repeats
CRISPRi	CRISPR interference
crRNA	CRISPR RNA
Csn2	CRISPR-associated protein Csn2
CXCR4	C-X-C chemokine receptor type 4
dCas9	inactive / dead Cas9
DHFR	dihydropholate reductase
DMJ	doença de Machado-Joseph
DNMT3	DNA methyltransferase 3
EMBL-EBI	European Bioinformatics Institute - European Bioinformatics Institute

EMX1	empty spiracles homeobox 1
ER	endoplasmic reticulum
ERAD	reticulum-associated degradation
ES	embryonic stem cells
FDA	food and drug administration
FKBP	FK506 Binding Protein 12
FokI	*Flavobacterium okeanokoites* domain I
FOXO4	forkhead box O4
FRB	FKBP-rapamycin binding
Gapdh	glyceraldehyde-3-phosphate dehydrogenase
GFP	green fluorescent protein
HA tag	Human influenza hemagglutinin epitope tag
HD	Huntington disease
HDAC3	histone deacetylase 3
HDAC6	histone deacetylase 6
HDR	Homology-directed repair
HEK 293T	human embryonic kidney 293 cells
hiPSCs	patient-derived induced pluripotent stem cells
HNH	HNH endonuclease domain (referring to Cas9)
HSV-TK	herpes simplex virus thymidine kinase
HTT	huntingtin
IgG	immunoglobulin G
IL2RG	interleukin 2 receptor subunit gamma
indels	insertions and/or deletion (referring to mutations)
iPSC	induced pluripotent stem cell
Klf4	krüppel-like Factor 4
KRAB	Krüppel-associated box
LSD1	lysine-specific histone demethylase 1A
LV	lentivirus
MJD	Machado-Joseph disease
MRI	magnetic resonance imaging
mRNA	messenger RNA
MTOC	microtubule-organizing centre

NES	nuclear exporting signal
NGG	*any nucleotide*-guanine-guanine
NHEJ	nonhomologous end-joining
NLS	nuclear localization signal
nMag	photoinducible dimerization negative magnet
Oct3	octamer-binding transcription factor 3
Oct4	octamer-binding transcription factor 4
PAM	protospacer-adjacent-motif
PCR	polymerase chain reaction
PGK1	phosphoglycerate kinase 1
pMag	photoinducible dimerization positive magnet
PNKP	polynucleotide kinase 3' phosphatase
PoliQ	poliglutamina
PolyQ	polyglutamine
qPCR	quantitative PCR
RISC	RNA-induced silencing complex
Rnase III	ribonuclease III
rpm	rotation per minute
RuvC	endonuclease domain (referring to Cas9)
SCA3	spinocerebellar ataxia type 3
sgRNA	single-guide RNA
shRNA	short-hairpin RNA
SIRT1	sirtuin 1
SMA	spinal muscular atrophy
SMN1	survival of motor neuron 1
SNP	single-nucleotide polymorphism
SOX2	SRY-box transcription factor
TadA	adenosine deaminase
TALEN	transcription activator-like effector nuclease
TGA	thymine-guanine-adenine (stop codon)
TPM	trimethoprim
tracrRNA	trans-activating crRNA
TTR	transthyretin

UIM	ubiquitin-interacting motifs
UK	United Kingdom
USA	United States of America
VCP	valsolin-containing protein
VPR	VP64-p65-Rta activator domains
YAC	Yeast Artificial Chromosome
ZFNs	zinc-finger nucleases

List of images

Figure 1 - Schematic representation of brain regions affected in MJD 16

Figure 2 - Schematic representation of wild type (WT) *ATXN3* and mutant *ATXN3* gene 17

Figure 3 - Schematic representation of a normal and an expanded polyQ tract in the atxn-3 protein 18

Figure 4 - Schematic representation of the domains of the atxn-3 protein 19

Figure 5 - Schematic representation of the Type II CRISPR *locus* of *Streptococcus pyogenes* 43

Figure 6 - Schematic representation of Type II CRISPR/Cas9 mechanism of action in prokaryotic cells.. 44

Figure 7 - Schematic representation of CRISPR/Cas9 DNA double-stranded break mechanism and the variations of the CRISPR/Cas9 system adapted from *S. pyogenes* 50

Figure 8 - Schematic representation of the Donor-141 template containing 141 CAG repeats being integrated into the human *ATXN3* gene and PCR screening analysis using the primers 5'F/6R to detect an amplicon of 1956 bp distinctive of integration 79

Figure 9 - PCR screening of cells co-transfected with the Donor_141, Cas9 and sgRNAs Pre-10, Ps10_2 or both 80

Figure 10 - Detection of atxn-3 protein from cells co-transfected with the Donor_141, Cas9 and both Pre-10 and Pos10_2 sgRNAs.. 81

Figure 11 - Schematic representation of the Donor-14Q-Puro repair template, containing a selection cassette and 14 CAGs, being integrated into the human *ATXN3* gene 84

Figure 12 - Schematic representation of the PCR genotyping strategy for detection of selection cassette insertion 85

Figure 13 - PCR screening of the cells transfected with the Cas9 and the Donor-14Q-Puro, aiming at amplifying a fragment of the selection cassette of the repair template 86

Figure 14 - PCR screening of the cells transfected with the Cas9 and the Donor-14Q-Puro, aiming at amplifying a fragment that indicated a correct integration 87

Figure 15 - Schematic representation of the PCR products generated by the primers 5'F and 6R, upon a second amplification step using primers CopGFPF/CopGFPR and the respective PCR screening analysis of cells transfected with the Cas9 and the Donor-14Q-Puro 88

Figure 16 - Schematic representation of the amplicon generated by the primers PB/PB.BB and PCR screening analysis of cells transfected with the Cas9 and the Donor-14Q-Puro.. 90

Figure 17 - Schematic representation of the amplicon generated by the primers PB/5'F_2 and PCR screening analysis of the cells transfected with the Cas9 and the Donor-14Q-Puro 91

Figure 18 - Detection of human atxn-3 protein by Western blot analysis of the cells transfected with the Cas9 and the Donor-14Q-Puro 92

Figure 19 - Schematic representation of the timeline used to behaviourally test and treat MJD transgenic mice with lentiviral particles encoding dCas9-KRAB 96

Figure 20 - *ATXN3* expression profile from transgenic mice cerebella, stereotaxically-injected with lentiviral particles encoding dCas9-KRAB/sgRNA_SNP or dCas9-KRAB/sgRNA_Empty 97

Figure 21 - Evaluation of the cerebellar degeneration profile in mice injected with lentiviral particles encoding for the dCas9-KRAB/sgRNA_SNP 99

Figure 22 - Motor performance of MJD transgenic mouse model at three, six and nine weeks after stereotaxic injection with lentiviral particles encoding the dCas9-KRAB targeted at the MJD-associated SNP of human *ATXN3*... 101

List of tables

Table 1 - Comparison between the different programmable nucleases used in gene editing. 47

Table 2 - Online platforms used to design sgRNAs that target the human *ATXN3* gene. 62

Table 3 - Primers used for amplification of the 5’ fragment and 3’ fragment 64

Table 4 - Antibodies used in Western blot analysis 66

Table 5 - Primers used to genotype transfected HEK 293T cells 67

Chapter 1 – Introduction

1.1 Machado-Joseph disease

1.1.1 Disease description and clinical features

Machado-Joseph disease (MJD), also known as spinocerebellar ataxia type 3 (SCA3), was first described as a syndrome (later designated as MJD) in 1972 by Nakano and colleges. They observed a new form of autosomal dominant hereditary ataxia in immigrant communities of Portuguese people living in Massachusetts, USA, descending from Guilherme Machado, who lived in the Portuguese archipelago of Azores (Nakano et al., 1972). Later, in 1978, Coutinho & Andrade performed an extended clinical description of several families living in the Azores, affected by a neurological disorder with genetic aetiology. They observed that patients developed a large spectrum of variable, degenerative and continuously progressing symptoms, normally appearing between 20-50 years of age. The clinical signs of what would eventually be defined as MJD included lack of motor coordination and impaired gait (cerebellar ataxia), difficulties articulating words (dysarthria), double vision (diplopia), involuntary eye movements (nystagmus), swallowing difficulties (dysphagia), progressive eye muscle weakness (ophthalmoplegia), hypotonia and muscular atrophy. Interestingly, and despite its hereditary characteristics, MJD shows a broad heterogeneity of symptoms even within the same family (Coutinho and Andrade, 1978; Vale et al., 2010).

Neuropathological studies of *post-mortem* MJD patients' brains helped to understand that the neurodegeneration that accompanies the disease is, in part, due to neuronal loss of several different brain regions (**Figure 1**). The spinocerebellar tracts and the dentate nucleus (cerebellum) along with the substantia nigra are the most commonly affected areas (Koeppen, 2018). Dorsal and ventral spinocerebellar tracts appear to be demyelinated, and this is accompanied by axonal loss and gliosis, the last one being also observed in the substantia nigra and dentate nucleus. There are also other areas of the nervous system that are affected and contribute to the neurologic symptoms. They include the basal ganglia, thalamus, brainstem (pons, midbrain and medulla oblongata); cranial nerves, spinal Clarke's nucleus and the anterior horns of the spinal cord; the Purkinje cell layer of the cerebellum seems to be only mildly affected (Bettencourt and Lima, 2011; Coutinho and Andrade, 1978; Matos et al., 2011;

Mendonça et al., 2015; Nakano et al., 1972; Rosenberg, 1992; Rüb et al., 2008; Seidel et al., 2012; Yamada et al., 2004).

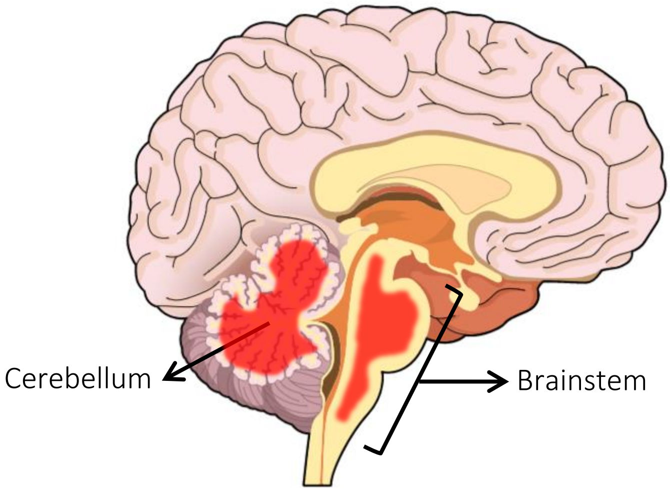

Figure 1 - Schematic representation of brain regions affected in MJD. The brain regions that are most affected in MJD are highlighted in red: cerebellum and brainstem.

1.1.2 Epidemiology

Although it has the classification of rare disease according to the European Commission definition (any disease affecting less than 1 people per 2000 inhabitants), MJD is the most common form of hereditary autosomal dominant ataxia in the world, globally affecting 0.3 to 2 people per 100 000 inhabitants. Nevertheless, in Portugal, MJD shows a prevalence of 3.1 people per 100 000 inhabitants, with foci of high prevalence in the archipelago of Azores (1 person per 2544 inhabitants), especially in the island of Flores, where the disease is estimated to affect 1 in 158 inhabitants. The disease also shows significant frequency in Brazil, Japan, Germany, Singapore and China (Araújo et al., 2016; Bettencourt et al., 2008; Bettencourt and Lima, 2011; Coutinho et al., 2013; Matos et al., 2011).

1.1.3 Disease aetiology

MJD is a monogenic disease, meaning that it is caused by a mutation in a single gene, the *ATXN3* gene, located in chromosome 14 (14q32.12; Genome Reference Consortium, GRCh38 December 2013). The mutation consists on an expansion of the cytosine-adenine-guanine (CAG) trinucleotide repeat tract present in the exon 10 of the gene and leads to the translation of an abnormally elongated polyglutamine (polyQ) tract-containing protein product - the expanded ataxin-3 (atxn-3) protein. Therefore, MJD is considered to be a polyQ disease. Eight other diseases are included in this group: dentatorubral-pallidoluysian atrophy, spinal and bulbar muscular atrophy, Huntington's disease and SCA1, 2, 6, 7 and 17. In all of these neurodegenerative disorders the underlying cause is an abnormal expansion of CAG trinucleotide repeats present in otherwise unrelated genes, originating overly expanded polyQ-containing protein products (Costa and Paulson, 2012; Ichikawa et al., 2001; Kawaguchi et al., 1994; Takiyama et al., 1993).

The numbers of CAG repeats associated with the pathogenic condition is different for each of the polyQ diseases (Orr and Zoghbi, 2007). In the case of MJD, in healthy people, exon 10 of the *ATXN3* gene normally exhibits a number of CAG repeats lower than 47, while patients normally carry above 55 CAG repeats (**Figure 2**). However, the exact threshold of CAG repeats needed for MJD to manifest is not precise, with reports of both healthy and affected people showing repeats between 45 and 51 CAGs (Matos et al., 2011; Montcel et al., 2014; Van Alfen et al., 2001). Importantly, it is well accepted that the number of CAG repeats inversely correlates with the age of MJD onset, and directly correlates with the rate of neurodegeneration, meaning that the more CAG repeats that are present, the sooner the disease manifests and the greater the symptoms severity (Abe et al., 1998; Jardim et al., 2001; Klockgether et al., 1996; Onodera et al., 1998).

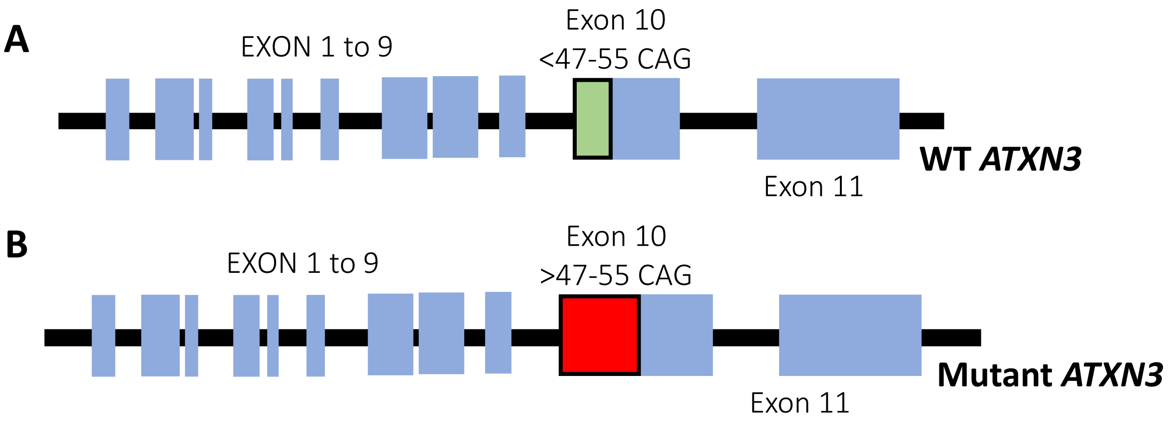

Figure 2 - Schematic representation of wild type (WT) *ATXN3* and mutant *ATXN3* gene. A) The WT *ATXN3* gene, which is not associated with the development of MJD, contains a CAG repeat number below 47-55 in the exon 10, while B) mutant *ATXN3*, which is present in people affected by MJD, contains an abnormal CAG repeat number above 47-55.

It is still not fully understood how an expanded polyQ tract in the atxn-3 protein leads to the MJD phenotypic traits. However, during the past years of research, some enlightenment has been achieved regarding the mechanisms causing the disease. In the next section we will try to cover some of those aspects.

1.1.4 Atxn-3 protein

The *ATXN3* gene product is a ubiquitously expressed 42 kDA protein named ataxin-3 (atxn-3). When the CAG repeat is overexpanded, the translated protein contains an abnormally expanded polyQ tract (**Figure 3**), that is thought to be responsible for MJD. However, the exact mechanisms underlying the expanded polyQ tract toxicity remain unclear (Bevivino and Loll, 2001; Nóbrega et al., 2018; Yang et al., 2015).

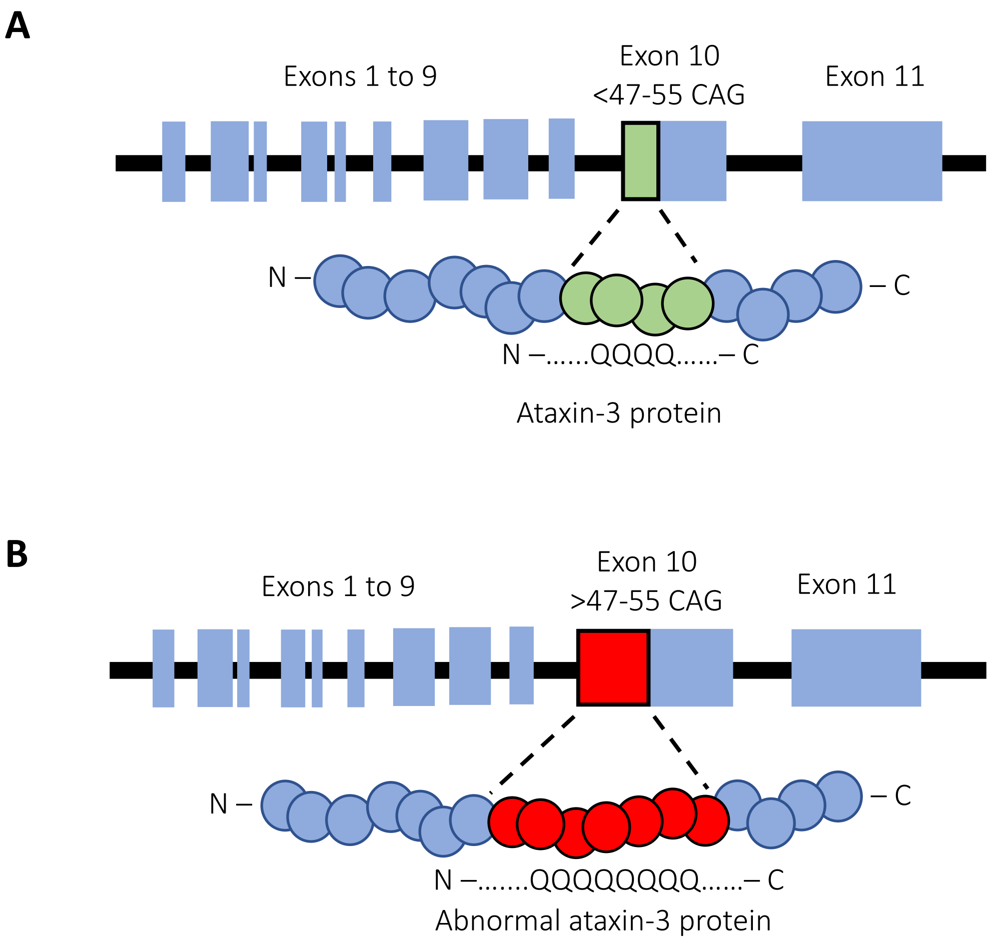
A
Exons 1 to 9
Exon 10
<47-55 CAG
Exon 11
N –
– C
N –......QQQQ......– C
Ataxin-3 protein
B
Exons 1 to 9
Exon 10
>47-55 CAG
Exon 11
N –
– C
N –.......QQQQQQQQ......– C
Abnormal ataxin-3 protein

Figure 3 - Schematic representation of a normal and an expanded polyQ tract in the atxn-3 protein. In normal conditions **A)** the *ATXN3* gene is translated as a protein containing less than 47 glutamine residues (green), while **B)** when expanded the polyQ tract contains > 47-55 glutamines (red).

Little is known about atxn-3 biological function. Studies showed that atxn-3 is composed by a Josephin domain with deubiquitinating activity and three ubiquitin-interacting motifs (UIM; **Figure 4**), suggesting an involvement in the ubiquitin-proteasome pathway. The protein also contains two nuclear exporting signals (NES) in the Josephin domain and one nuclear localization signal (NLS) between UIM2 and the polyQ tract (Burnett et al., 2003; Mao et al., 2005; Matos et al., 2016).

Other functions have been proposed for atxn-3, such as transcription regulation, through interaction with histone deacetylase 3 (HDAC3), the cAMP response element-binding-binding protein (CREBBP) and the transcription factor forkhead Box O4 (FOXO4) (Araujo et al., 2011; Evert et al., 2006; Li et al., 2002). A role in DNA damage repair responses has also been recently suggested, with atxn-3 interacting with proteins involved in genome stability and DNA damage responses such as the Chk1 protein and polynucleotide kinase 3'-phosphatase (PNKP) protein, respectively (Tu et al., 2017; Ward and La Spada, 2015).

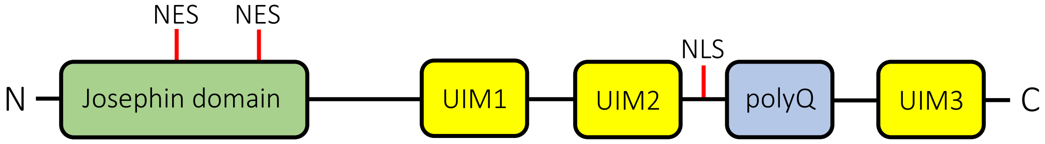

Figure 4 - Schematic representation of the domains of the atxn-3 protein. Atxn-3 is known to have a N-terminal Josephin domain with deubiquitinating activity, and two or three ubiquitin-interacting motifs (UIM) (depending on the isoforms), contained in the C-terminal tail. The polyQ tract is also represented between UIM2 and UIM3, as well as the nuclear export signals (NES) and a nuclear localization signal (NLS). Adapted from Nóbrega et al., 2018.

It is also thought that atxn-3 plays a role in cytoskeletal organization, regulating the formation of focal adhesions needed for myogenesis, via integrins signalling. Atxn-3 protein interacts with integrin subunits α5 and α7 that are expressed during myogenesis, preventing the degradation of the subunits or the integrins complex. *In vitro* studies showed that upon atxn-3 depletion, cells forced to differentiate into myocytes acquired morphologic characteristics that do not correspond to muscle cells and the cytoskeleton was disarranged. Integrin subunits α5 and α7 were also downregulated after atxn-3 depletion (do Carmo Costa et al., 2010).

Atxn-3 appears to be involved in the endoplasmic reticulum-associated degradation (ERAD). Atxn-3 binds to valosin-containing protein (VCP)/p97 decreasing the extraction of ERAD subtracts from the endoplasmic reticulum (ER). It is known that this activity is upregulated in the presence of mutant forms of atxn-3 leading to ER stress, possibly due to accumulation of ERAD subtracts. However, the reason why atxn-3 regulates ER extraction of ERAD substrates is unknown; it might be regulating the flow of misfolded proteins to degradation (Poksay et al., 2011).

Finally, aggresomes formation might be regulated by the atxn-3 protein. Aggresomes are aggregates of misfolded proteins and are formed near the microtubule-organizing centre (MTOC) as a homeostasis maintenance mechanism, when the proteasome is impaired and cannot process the degradation of proteins. Atxn-3 interacts with dyneins and the microtubule-associated deacetylase histone deacetylase 6 (HDAC6), known to form complexes responsible for transporting misfolded proteins to the MTOC, during aggresomes formation (Burnett and Pittman, 2005).

1.1.5 Mutant atxn-3 protein and MJD pathogenesis

The mechanisms by which the mutant atxn-3 leads to MJD pathology are still not fully understood. In the last few years, studies have proposed that changes in protein conformation, protein-protein interactions and even impairments in protein degradation pathways might be the main mechanisms by which the mutant atxn-3 protein leads to neurodegeneration and consequently to the MJD motor deficit traits.

The expanded polyQ tract causes atxn-3 to acquire an abnormal conformation, leading to atxn-3 intracellular aggregation in neurons. Protein aggregation is admitted to play an important role in expanded atxn-3 toxicity; for example, it was observed that atxn-3 aggregates are capable of sequestering non-mutant atxn-3 or even other proteins, possibly causing them to become dysfunctional. The presence of large, ubiquitinated, atxn-3-containing intraneuronal inclusions is a hallmark of MJD, though the link between these macromolecular structures detected in patients and neurodegeneration is still debatable. In fact, soluble amyloid atxn-3 oligomers are instead pointed out as the effectors of toxicity (Takahashi et al., 2010). Why only a specific population of neurons is affected by atxn-3 toxicity also remains a mystery. A possible answer may reside in interactions between atxn-3 and cell specific proteins that would culminate in the selective neurodegeneration. (Nóbrega et al., 2018; Paulson et al., 1997; Yang et al., 2015).

Proteolytic cleavage of mutant atxn-3 is another pathway which seems to contribute to MJD neurodegeneration. Upon cleavage, expanded atxn-3 forms smaller toxic fragments that aggregate in the nucleus and dysregulate cellular homeostasis. The players contributing for these proteolytic cleavage events are caspases and calpains. The cleavage of mutant atxn-3 some amino acids N-terminally of the nuclear localization signal (NLS) site (**Figure 4**) is thought to be responsible for the translocation of atxn-3 fragments containing polyQ tracts to the nucleus (Antony et al., 2009; Berke et al., 2004; Breuer et al., 2010; F. Gould et al., 2007; Hubener et al., 2013; Paulson et al., 1997).

Other studies established a link between mutant atxn-3 and the impairment of two mechanisms involved in protein degradation: the ubiquitin-proteasome pathway and autophagy, which is especially important in degrading proteins that are prone to aggregate. In fact, studies report that autophagic vesicles accumulate in MJD, possibly as a mechanism to clear atxn-3. Dysregulations in autophagy pathways can lead to neurodegeneration as seen in MJD. Additionally, induction of autophagy seems to be beneficial in the context of the disease, decreasing the toxic effects of atxn-3 aggregates, highlighting the importance of atxn-3 clearance in MJD (Burnett et al., 2003; Ravikumar et al., 2008; Sittler et al., 2017).

Although an increasing number of reports has shed light onto some of the functions of atxn-3 and some of the mechanisms by which its mutant form can lead to cellular dysfunction and neuronal death, much is still not known. MJD pathogenic mechanisms are not fully understood and the disease remains, to this date, incurable. A better understanding will arise from further studies, which will certainly benefit from experimental models that recapitulate the disease condition. Optimized models will allow testing new therapies and possibly the development of a definitive cure.

1.2 Models of Machado-Joseph disease

1.2.1 Importance of human disease models

From the simplest unicellular organism to complex multicellular animals, scientists have been using different biological systems to model human diseases, hoping to discover the mechanisms underlying a certain pathology. The models are the tools that allow researchers to manipulate cellular pathways, test new treatments and develop therapies for diseases.

Whatever the case, it is important to choose an appropriate model that recapitulates the relevant aspects of the human condition. Depending on the mechanism that is supposed to be investigated a suitable model must be picked in order for it to yield accurate insights into the molecular and cellular malfunctions causing a disease. Additionally, choosing a proper model that recapitulates the disease phenotype opens the possibility of developing effective therapeutics against human pathologies (Dow and Lowe, 2012; Reed et al., 2017). Ideally, disease models should have a genetic basis that is comparable to that of the humans conditions; similar physiological and anatomical features; similar disease mechanisms that recapitulate the human conditions and allow the prediction of the clinical outcomes of a therapy (Mullane and Williams, 2014).

1.2.2 Cell cultures as disease models

Since 1665, when Robert Hooke first observed cells from plants in his newly evented compound microscope, the world of science was revolutionized. He observed the cell structure of cork, which was similar to honeycombs and coined the term *cell* because it reminded him of monks' rooms (cells) in a monastery (Davidson, 2010). Robert Hook opened the first door needed to study cells, their structures and later their mechanism.

In the end of the nineteenth century, Roux performed the first cell culture attempt by maintaining the neural plate of chicken embryos in culture with saline buffer. Later in the 1900s, Harrison first cultured tissue in a test tube, using embryonic tissue from frogs placed in agar, blood clots or salt solutions. Other strategies were developed in the following years. Carrel cultured mammal tissue placed in serum from different animals and was able to serially transfer cells to new plates, and he also implemented sterile conditions and Pyrex glass (as growing surface) in his cultures. Carrel was able to subculture cells from embryo chicken heart for several years (1912-1946), creating what is considered to be the predecessor of cell lines (Jedrzejczak-Silicka, 2017; Rodríguez-Hernández et al., 2015; Taylor, 2014).

The first mammalian cell line derived from subcutaneous connective mouse tissue was established by Earle and colleagues in the 1940s. However, it was only in 1951 that the widely known HeLa cells, the first immortalized human cell line, were created, almost 300 years after Robert Hook first observations of cork cells. As a model of study, HeLa cells revolutionized the word of science, contributing to numerous breakthroughs including polio eradication and the development of vaccines against cancer-related human papilloma virus (Earle et al., 1943; Masters, 2002; Mittelman and Wilson, 2013). Nowadays, it is possible to find cell lines from over 150 different species. Just from humans nearly 4000 different cell lines ranging from fibroblasts to neuronal cells are available (ATCC, 2019; Kaur and Dufour, 2012).

Yamanaka and colleagues also revolutionized the scientific world of cell studies and cell culture. In their 2006 work, they identified four factors (c-Myc, Klf4, Oct3/4 and Sox2) capable of inducing cells to acquire a pluripotent state, reverting differentiation. Those cells were designated as induced pluripotent stem cells (iPSCs). The main advantages of iPSCs is the fact that they are a person-specific, unlimited source of cells, they can be differentiated into any other cell type by providing the appropriate factors and they can be used in cell therapy, developing low to non-immune responses when auto-transplanted (Shi et al., 2017; Takahashi and Yamanaka, 2006).

iPSCs gain even more relevance when studying diseases in which the main affected organ is inaccessible, such as in the case of polyglutamine disorders. Easily accessible fibroblasts can be collected from affected patients and reprogramed into iPSC as an unlimited source of cells, that can later be derived into neurons, for example. Patient-derived iPSCs can be used as disease models and drug discovery platforms, to study for example pathophysiological pathways and drugs responsiveness. However, one does not fit all, and results obtained from iPSC derived from one patient may differ when compared with another patient, due to differences in the genetic background (Xiaohong Xu et al., 2017).

Ideally, studies using iPSCs for disease modelling may employ isogenic iPSCs cell lines for. Isogenic lines are cell lines that have equal genotype, differing only in the expression of a limited number of relevant, disease-related genes. Since isogenic cell lines share the same genetic background, studies are safeguarded against intrinsic genetic variations in gene expression that are not caused by a disease-related gene, but that instead result from the normal genetic differences between individuals. Therefore, using isogenic cell lines helps in assuring that differences observed between patient-derived cell cultures and their respective healthy controls are disease-related and relevant as a disease-associated output (Szlachcic et al., 2017; Xiaohong Xu et al., 2017). This methodology also facilitates the study of disease-modifying variants that can be responsible for pathological differences observed between patients or therapeutic responsiveness to a drug, that can differ between patients due to their genetic background differences (Flower et al., 2019; Visschedijk et al., 2018).

In 2017, Xiaohong Xu and collaborators were able to generate an isogenic cell line for HD. They successfully corrected HD patient-derived iPSCs using CRISPR/Cas9. They observed that cells retained pluripotency, differentiation potential and, when differentiated into neurons, cells exhibited electrical activity (Xiaohong Xu et al., 2017). It is fair to assume that a similar approach can be implemented for MJD

1.2.3 Animal models of Machado-Joseph disease

Despite the obvious advantages of using cell cultures in research (less expensive cost, easy maintenance and faster results), several inter-tissue communications are lost in these types of models (although co-culture can mimic some aspects of it) and cells behave differently when organized into a living multicellular organism. Thus, the use of more complex systems for modelling diseases in the form of animal models such as *Caenorhabditis elegans*, *Drosophila melanogaster*, mice or non-human primates, offers several advantages over simple cell cultures.

Since the brain is the main affected organ in MJD, its accessibility is a limiting factor in research. For many years researchers were restricted to analyse *post-mortem* brain tissue that could only give insight into the neurodegenerative profile at late stages of the disease. Thus, MJD animal models are crucial tools that help to understand the mechanisms by which a polyQ protein with an expanded glutamine tract can lead to the observed pathology and the pathways involved and their repercussions.

After the discovery and characterization of the genetic mutation responsible for MJD, an obvious correlation between the genetic mutation and the phenotypic outcome emerged (Kawaguchi et al., 1994). Since then, disease models expressing the mutant

ATXN3 gene have been engineered and they have provided clues that have helped understand the cellular and molecular mechanisms underlying the pathology of MJD.

Much of what is known about MJD pathogenesis came from experiments conducted in animal models. They helped understand the symptoms progression, the neurodegenerative profile, the affected brain regions, atxn-3 functions and mutant atxn-3 toxicity. Assessing neuronal dysfunction such as glutamatergic and GABAergic release impairments was only possible using models of MJD (Cemal et al., 2002; Chou et al., 2008; Nóbrega et al., 2018).

Nowadays, there is a wide range of animal models of MJD available, from very simple organisms, such as *C. elegans*, to complex mammals such as the common marmoset. Other animals used for modelling MJD include the fruit fly, the zebrafish, the rat and the mouse. Mice have been used in biomedical research for more than 100 years. Not only because they share 99% of their genes with humans but also because they show similarities in their physiological and pathological conditions. Moreover, mice breed and mature quickly, they are inexpensive, easy to house and some strains have all their genomes sequenced (Justice and Dhillon, 2016; Rosenthal and Brown, 2007). For all these reasons mice are interesting organisms for human disease modelling. Nevertheless, researchers must be careful when they are analysing results obtained from mouse models in order to get an accurate translational approach to humans.

Using animal models of MJD as pre-clinical tools for therapy studies offers the possibility of testing and developing putative therapeutic strategies for the disease.

1.2.4 Caenorhabditis elegans *models*

The nematode *Caenorhabditis elegans* is an organism that has been used for more than fifty years as an animal model in research. They are easy to grow; they display a short lifespan and they are self-fertilizing. However, they are evolutionarily very distant from humans, anatomically and physiologically they are very different, and the small size makes some experimental procedures difficult (Johnson, 2003; Tissenbaum, 2015).

One group of transgenic *C. elegans* strains used to study MJD expresses full-length or truncated forms of human mutant atxn-3 under the control of a neural promoter (unc-119). In embryonic stages the full-length form did not display aggregates; only the truncated forms of atxn-3 with 64 and 127 glutamines showed perinuclear aggregations (not nuclear), in a few or in most of the neurons, respectively. Just in later stages a full-length atxn-3 isoform-1a (contains only two UIMs) containing 130 glutamines formed aggregates. Regarding neuronal impairments, similar to what happened with the aggregates, animals with full-length atxn-3 with 130 glutamines only displayed motor incoordination when they aged. The animals expressing the truncated form with 127 glutamines showed traits that went beyond motor incoordination, such as defects in egg laying and abnormal defecation, along with impairments in acetylcholine release, possibly due to interaction of the aggregates with the synaptic vesicles, since synaptotagmin co-localized with aggregates. Their neurons also showed abnormal branching, swelling and beading, but only in old animals. The ubiquitin-proteasome system was impaired in models expressing truncated atxn-3 forms with 19 glutamines (old stage) and 127 glutamines (Khan et al., 2006).

Interestingly, another group using *C. elegans* as a MJD model was able to induce symptoms using full-length atxn-3 containing 130 glutamines. This contrasts with the results from Khan and collaborators, but the promoter and the atxn-3 isoform used were different; in this study was used the isoform-3c (containing all three UIMs) and the F25B3.3 pan-neural promoter (Khan et al., 2006; Schmidt and Schmidt, 2018; Teixeira-Castro et al., 2011).

Transgenic *C. elegans* expressing full-length atxn-3 isoform-3c with 130 glutamines showed neuronal dysfunction and aggregation. The aggregates were sublocalized in the nucleus and the cytoplasm from early to late stages of the *C. elegans* life. Lines expressing fragments of atxn-3 were also generated, in order to evaluate the possible effect that proteolytic cleavage might have. Aggregations were observed in fragments containing more than 75 glutamines, reinforcing the idea that fragments are more prone to aggregate, since transgenic *C. elegans* expressing full-length forms of the atxn-3 protein with 75 glutamines fail to aggregate. Regarding motor functions, only the transgenic *C. elegans* expressing full-length atxn-3 with 130 glutamines or the truncated forms with more than 40 glutamines showed reduced movement, and they also failed to respond to chemiotaxis agents, suggesting impairment in sensory neurons, and thus neuronal dysfunction (Teixeira-Castro et al., 2011).

Later, Christie and collaborators used a different approach to develop a model that, instead of expressing atxn-3 isoform-3c in neurons, expresses the protein in *C. elegans* muscle cells under the control of the unc-54 myosin promoter. The main finding was that the pathologic signs were not age-dependent under these circumstances. High levels of aggregation could be seen in body wall muscle cells and transgenic *C elegans* showed mobility deficits (Christie et al., 2014).

1.2.5 Drosophila melanogaster *models*

The fruit fly (*Drosophila melanogaster*) is another organism that is widely used as an animal model in biomedical research. Regarding neurodegenerative diseases, they already have been used to model Alzheimer's disease, Parkinson's and polyQ diseases (Lu and Vogel, 2009). It is advantageous to use *Drosophila* because they are small, inexpensive to maintain, easy to house, and easy to manipulate genetically. However, they have some limitations that need to be considered when choosing this animal as model for disease, such as the anatomical and physiological differences they have with humans. Even so, some researchers have been using *Drosophila* as a model for MJD and in fact they were able to replicate some features of the human disease (neurodegeneration and intranuclear inclusions). Because fruit flies do not express an orthologue of the *ATXN3* gene*,* knock in and knock outs are impossible (Jung et al., 2009; Warrick et al., 2005).

Warrick and collaborators generated a *Drosophila* MJD model by expressing a truncated form of the atxn-3 protein containing 78 glutamines, under the control of different promoters that would allow tissue-specific expression. However, the phenotypic severity within the transgenic *Drosophila* containing 78 glutamines was different between individuals, even though the genotype was the same, possibly due to chromosomal position of the transgene. Using a promoter targeting all the nervous system (HuC) or a promoter for muscle cells (GAL4-24B) resulted in lethality. Expressing atxn-3 with 78 glutamines in the *Drosophila* eyes (*glass* promoter) impaired eye development and consequently led to morphology and pigmentation defects; eyes underwent progressive degeneration accompanied by cell loss. This model also showed intranuclear inclusions in the eye cells that were more predominant in old flies (Warrick et al., 1998).

In 2005, Warrick and co-workers developed another MJD *Drosophila* model co-expressing normal atxn-3 with 27 glutamines and truncated or full-length pathogenic forms. Interestingly, co-expression of normal and mutant atxn-3 delayed the pathogenic effects described above, suggesting that wild-type atxn-3 may have a protective effect against polyQ-expanded mutant forms (truncated with 75 glutamines or full-length with 84 glutamines). The co-expression also delayed the formation of intranuclear inclusions. Interestingly, this protection mechanism seems to involve the ubiquitin-proteasome pathway, since disrupting the proteasome activity or mutating the UIMs within normal atxn-3 decreased its protective effects (Warrick et al., 2005).

Kretzschmar and colleagues created a transgenic MJD *Drosophila* model expressing a truncated form of atxn-3 either with 27 or 78 glutamines, under the control of glial or neuronal-specific promoters. However, the expression strength was different within the transgenics. Neuronal and glial expression was either lethal or decreased lifespan, both in the 27 and 78 glutamines truncated forms. Curiously, progressive neurodegeneration was only observed in animals with glial-specific expression under the control of the M1B promoter, but the severity was directly correlated with the extension of the polyQ tract. This model also showed atxn-3 nuclear aggregates (in glia cells and neurons) in an age-dependent manner, prominent in transgenic animals containing 78 glutamines but less visible in lines expressing the 27 glutamines. The motor functions were also affected. However, it was in the lines with glial-specific expression that these deficits were more visible (Kretzschmar et al., 2005).

In 2009, a new line expressing full-length atxn-3 protein with 84 glutamines was generated in an attempt to unveil the effect of atxn-3 cleavage in MJD pathology. In fact, animals reproduced the cleavage patterns observed in humans. Aggregates and nuclear inclusion formation were observed as well as neurodegeneration. The formation of fragments by proteolytic cleavage could be prevented using caspase inhibitors or by mutating putative sites of caspase cleavage within the atxn-3 protein. Although the presence of aggregates and nuclear inclusions was not altered in that case, the neurodegeneration progression was decreased, demonstrating a clear link between degeneration and proteolytic cleavage of the mutant atxn-3 protein (Jung et al., 2009).

1.2.6 Zebrafish model

Zebrafish (*Danio rerio*) has also been used for modelling human diseases. The advantages of using this kind of animals resides in its unexpansive cost (compared with mice), high fecundity, transparency in early developmental stages, rapid development, small size and some conserved genes with humans. However, they diverge significantly from humans and are anatomically and physiological very different. Additionally, the fact that zebrafishes are polyploid for some genes hinders gene manipulation (Guyon et al., 2007).

Taking into consideration the advantages and drawbacks of using zebrafish as a model for disease, Watchon and colleagues developed the first transgenic zebrafish model of MJD. Two lines were generated, expressing human mutant atxn-3 protein containing 23 or 84 glutamines, under the control of a neural promoter (HuC). In this model it is possible to see atxn-3 protein expression in neurons of the spinal cord, brain and cranial nerves. The lifespan in both transgenic is decreased when compared with wild-type fish but was significantly different between them (animals with 23 or 84 glutamines). Morphological changes such as neurite beading in the medulla sections were also observed. Regarding motor function, the swimming speed was not altered, but the transgenic fish expressing atxn-3 with 84 glutamines tend to move less and through shorter distances, and the swimming acceleration was also decreased (these symptoms worsen with age). Neuropathological studies showed neuronal degeneration in motor neurons characterized by a decrease in axon length. Immunoblots of transgenic zebrafish cell lysates show atxn-3 cleavage products whose formation was prevented using calpain inhibitors (calpeptin), consequently improving the motor function of the transgenic model (Watchon et al., 2017).

1.2.7 Marmoset – A non-human primate model

The first transgenic non-human primate model of MJD was developed in 2017. It expresses full-length human mutant atxn-3 isoform-3c containing 120 glutamines, under the control of the cytomegalovirus promoter. The model was generated using lentivirus to transduce common marmoset (*Callithrix jacchus*) embryos. From seven animals, six integrated the human gene, but only three showed expression accompanied with motor deficits, hypoactivity and weight loss at 3 to 4 months that progressively worsened. Magnetic resonance imaging (MRI) analysis of one symptomatic marmoset showed neurodegeneration (loss of Purkinje cells) and gliosis in the cerebellum. Intranuclear inclusions were also found in the spinal cord (anterior horn). Another marmoset showed no cerebellar pathology, but the spinal cord and brainstem were positive for intranuclear inclusions. Thus, some phenotypic variations are observed between transgenic marmosets.

Although non-human primates can display great genetic similarity with humans, making them good models of genetic diseases, their maintenance is costly, and protocols are long and time-consuming. Additionally, the conditions and licenses required to house non-human primates and experiment on them are not widely available (Tomioka et al., 2017).

1.2.8 Lentiviral rodent models

Besides the previously described animal models of MJD, the vast majority of them are rodents, especially mouse models, either expressing truncated or full-length-forms of atxn-3.

Lentiviral vectors can be used to elicit a localized expression of particular genes in adult animal tissue. Human atxn-3 expression has been lentivirally-mediated in adult rodent brain tissue in order to understand the consequences of the presence of the expanded protein in a particular brain area without directly affecting any other regions.

Alves and collaborators developed a rat model with regional pathology, by expressing the human atxn-3 cDNA using lentivirus in particular brain regions (substantia nigra, striatum and cortex). Atxn-3 contained 72 glutamines and was expressed under the control of the phosphoglycerate kinase 1 (PGK1) promoter. Expression of human mutant atxn-3 in the substantia nigra led to the formation of intranuclear inclusions that co-localized with ubiquitin, accompanied by neuronal loss of dopaminergic neurons, and motor deficit. Animals expressing mutant atxn-3 in the cortex showed intranuclear inclusions, but not signs of neurodegeneration. Striatal expression led to mutant atxn-3 aggregation and ubiquitinated intranuclear inclusion formation, as well as neuronal loss (S. Alves et al., 2008).

Later, in 2012, Nóbrega and colleagues generated an MJD lentiviral model using mice instead of rats. They performed an injection of lentivirus encoding the mutant human atxn-3 protein with 72 glutamines, directly into the cerebellum. Ten weeks after transduction, mice showed ubiquitinated intranuclear inclusions containing mutant atxn-3. Neuronal loss was also observed within the cerebellum, associated with motor impairments that progressively worsened. These included motor incoordination and lack of balance at 6 weeks. Thus, by expressing the mutant atxn-3 only in the cerebellum, Nóbrega and collaborators were able to replicate some key features of MJD (Nóbrega et al., 2013a).

1.2.9 Transgenic mouse models expressing truncated human ataxin-3

As presented above, several studies demonstrate that truncated forms of atxn-3 containing the polyQ tract are enough to induce MJD-mimicking outcomes in diverse animal models. In fact, this type of truncated forms have been repeatedly shown to be even more prone to aggregation and toxicity than the full-length mutant protein (Matos et al., 2017).

Ikeda and collaborators created the first transgenic MJD mouse model in 1996 by inducing the expression of a human mutant atxn-3 fragment. Mice expressed a C-terminal truncated form of atxn-3 with 76 glutamines, under the control of the L7 promoter (Purkinje cells). The experiments initially performed were not extensive but gave one of the first insights into some mechanisms underlying polyQ tract pathogenicity. The transgenic mice developed motor incoordination, walking difficulties and posture deficits. The severity of the phenotype directly correlated with the copy number variability of the inserted transgene. Affected mice also showed a shrunken cerebellum, and the morphology of Purkinje cells was severely affected at 8 weeks of age (Ikeda et al., 1996).

In 2007, Gould developed a mouse model expressing a mutant truncated form of atxn-3 lacking amino acids 190-200, under the control of the Purkinje cell-specific L7 promote. These mice showed motor incoordination, tremor and posture deficits, as well as weight loss. Premature death was also described (V. Gould et al., 2007). This model intended to identify putative locations where proteolytic proteins could cleave mutant atxn-3, based on the model developed by Goti and colleagues in 2004 (described below).

Torashima and collaborators created a MJD mouse model expressing a truncated form of the human atxn-3 protein containing 69 glutamines under the control of the Purkinje cell-specific L7 promoter. These mice showed a very early (3 weeks) severe ataxic phenotype, weight loss and neurodegeneration in the cerebellum. Purkinje cells nuclei, cytoplasm and axons contained aggregates that co-localized with ubiquitin at post-natal day 80 and that progressively increased with age. Overall characterization showed an early onset accompanied by a highly severe motor phenotype caused by cerebellar loss (Torashima et al., 2008).

Later, Hübener and co-workers generated an MJD mouse model expressing an N-terminal region of human atxn-3 containing the Josephin domain and one UIM, which is similar to a calpain-activity generated fragment associated with MJD. Interestingly, some MJD traits were replicated even without the polyQ tract. Motor functions were impaired at 9 months at age. Mice showed decreased motor coordination, imbalance, hindlimb clasping and weight loss, in an age-dependent manner. They also died prematurely. Aggregates were found in the cerebellum (dentate nucleus) and were accompanied by Purkinje cells degeneration. Pons and brainstem also showed aggregates that were more prominent as the mice aged. Interestingly, the expressed truncated N-terminal human protein induced endoplasmic reticulum stress and dysfunctional response to unfolded proteins. This model harbours a putative toxic function for the N-terminal region of the human atxn-3 protein, leading to traits that are characteristic of MJD (Hübener et al., 2011).

1.2.10 Knock-in mouse models

Attempts have been made to create MJD mouse models in which the CAG expansion would be introduced in the endogenous genetic context of the mouse *Atxn3* gene. This has led to the generation of two different MJD knock-in mouse models.

Switonski and co-workers developed a knock-in MJD mouse model, attempting to reproduce MJD in the endogenous genetic context of the mouse *Atxn3* gene. The Ki91 model expresses a humanized form of mouse atxn-3 under the expression of the endogenous mouse promoter. They used an approach through which a region from exon 7 to exon 11 of the *Atxn3* mouse gene would be replaced, by homologous recombination, with cDNA containing a sequence (from the exon 7 to the exon 11) of the human *ATXN3* gene, with the particularity that exon 10 would contain 91 CAG repeats, giving rise to a polyQ tract with 91 glutamines. The mutated Atxn-3 showed expression throughout the organism (cerebral cortex, cerebellum, pons, striatum, lungs, testis, liver and spleen). Intranuclear and perinuclear inclusions were observed in the cerebellum, cerebellar cortex and hippocampus. Purkinje cells showed mild neurodegeneration, that was not uniform throughout the cerebellum. Regarding motor performance, symptoms had a late onset and motor incoordination was only evident after 20 months of age. Although the model shows one of the hallmarks of the disease (aggregates), it failed to properly reproduce a pronounced neurodegeneration or replicate the motor impairments (Switonski et al., 2015).

In the same year, another knock-in mouse model was generated using a similar approach, the Atxn3Q82/Q6 model. 82 CAG repeats were inserted by homologous recombination into the *Atxn3* mouse gene, so that it would code a total of 85 glutamines. As expected, the knock-in mouse expressed the mutant atxn-3 protein in all tissues. Heterozygous mice showed small intranuclear inclusions at 10 weeks, that were more evident at 12 months of age in the striatum, hippocampus and cortex. Intracellular aggregates were also observed in the amygdala and the hippocampus. However, none of the brain regions observed demonstrated any sign of neurodegeneration, neither behavioural deficits in different experiments testing motor performance (rotarod, beam walk and open field), even after 12 months of age. Nevertheless, this model may be suitable for studying the mechanisms of aggregation and intranuclear inclusion formation, as well as some protein dynamics, since the Atxn3Q82/Q6 model showed an alternative splicing form resulting in two different isoforms of the mutated atxn-3 protein (Ramani et al., 2015).

Both the Ki91 model and the Atxn3Q82/Q6 model fail to properly reproduce motor impairments and neurodegenerative signs found in MJD patients. These knock-in mouse models raise the question of why in humans an endogenous expanded atxn-3 protein gives rise to MJD, while in mice it fails to do so. This highlights the important of the mouse models that rely on the expression of the human atxn-3 protein, under the genetic human context of the human *ATXN3* gene.

1.2.11 Transgenic mouse models expressing full-length human ataxin-3

Although mouse models expressing truncated forms of mutant human atxn-3 can replicate some of MJD traits, in humans there is expression of the full-length atxn-3 protein. In a protein context point of view, it is thus more advantageous to have models that express the full-length atxn-3 protein.

The first mouse model expressing the full-length human atxn-3 protein actually contained the entire human *ATXN3* gene *locus*. The full-length human mutant atxn-3 protein was thus expressed under the control of the native human regulators. Engineering and characterization of this mouse line was performed by Cemal and colleagues, who employed a yeast artificial chromosome (YAC) vector to modify a C57BL/6J mouse. YACs are systems with the capacity of cloning DNA fragments of up to 1000 Kbp, being able to carry entire genes or even gene clusters (Schedl et al., 1992). The particular YAC used in the generation of the model contained the whole human *ATXN3* gene, with all exons, introns and regulatory regions, thereby mimicking the human genetic context of MJD, i.e., regulation of gene expression by *cis*-acting elements; tissue-specific transcriptional regulation; alternative gene splicing; and methylation status (Cemal et al., 1999; MacKenzie, 2006; Matos et al., 2011).

The mouse model developed by Cemal and collaborators, YAC MJD-Q84.2, contains 84 CAG repeats. This number of repeats is known to be pathological in humans, causing MJD. Moreover, YAC MJD-Q84.2 mice show expression of human *ATXN3* mRNA in several different tissues (heart, skeletal muscle and brain) at levels that are approximately the same as the endogenous mouse *Atxn3* gene-derived mRNA. Regarding the human atxn-3 protein, YAC MJD-Q84.2 mice show a similar tissue-specific expression pattern when compared to the endogenous atxn-3. Neuropathology studies revealed that subcellular localization of atxn-3 was predominantly intranuclear, in the dentate and pontine nuclei, with inclusions. Regions that are not affected by the disease did not show any intranuclear inclusions (rarely observed in Purkinje cells). Neuronal loss and gliosis of the dentate nucleus, substantia nigra and pontine nuclei was observed, as well as demyelination in the peripheral nervous system. Regarding the behavioural phenotype and the motor symptoms of the model, YAC MJD-Q84.2 mice showed impaired posture, weight loss, hypoactivity, hind- and forelimbs clasping, motor incoordination and reduced pain sensitivity. Though symptoms were mild and the neurodegeneration progressed slowly, when compared with other available mouse models the YAC MJD-Q84.2 mouse model seems to be the one that better recapitulates the human MJD condition, gathering several features observed in the disease (Cemal et al., 2002; Colomer, 2012; Figiel et al., 2012).

There are other mouse models that express the full-length mutant atxn-3, however, all of them are based on the expression of cDNA. This means that they lack the endogenous human *ATXN3* regulatory elements and then cannot express all atxn-3 isoforms, limiting the translation to humans affected by MJD. Goti and collaborators created a transgenic homozygous mouse model expressing isoform-1a of human mutant atxn-3 containing 72 glutamines, under the control of the mouse prion promoter. The first symptoms appeared at 2-4 months and they included abnormal posture, motor incoordination and tremor that progressively worsened throughout aging, culminating in premature death. Strength, reflexes and body weight were also affected. Regarding neuropathology, intranuclear inclusions were selectively observed in pontine nuclei, deep cerebellar nuclei and in motor neuros of the spinal cord. A reduced number of dopaminergic neurons was also observed in the substantia nigra, but not in the dentate nucleus (one of the areas that is most affected in MJD) (Goti et al., 2004).

Later, in 2007, Bichelmeier and co-workers developed a heterozygous mouse model of MJD expressing isoform-3c of mutant human atxn-3, containing 70 or 148 CAG repeats, under the control of a mouse prion promotor fragment. These mice showed an extremely severe phenotype that worsened with the size of the polyQ tract. They displayed tremor, uncharacteristic posture and hypoactivity at 6-8 weeks. Atxn-3 and ubiquitin-atxn-3 intranuclear inclusions were present in almost all brain regions. Premature death was also observed (3 to 6 months). When human atxn-3 was intentionally targeted to the nucleus the severity of the phenotype increased, reinforcing the relation between the nuclear localization of the protein and its toxicity. However, neurodegeneration was not very marked (Bichelmeier et al., 2007; Matos et al., 2011).

One year after publication of Bichelmeier's mouse model, Chou and collaborators created a new mouse model for MJD. It expresses human mutant atxn-3 isoform-1a containing 79 glutamines, also under the control of the mouse prion promoter. Mice developed symptoms around 6-5 months of age, but showed nonsignificant neuronal loss, which led researchers to the hypothesis that ataxia was due to neuronal dysfunction of the cerebellum and not to neuronal death. The symptoms first appeared at 5 months of age and the severity increased with aging and included motor incoordination and hypoactivity. Neuropathological studies showed intranuclear inclusion in the neurons of the dentate nucleus, pontine nucleus and substantia nigra. Although no neuronal loss was observed, the morphology of Purkinje cells was altered, suggesting some degree of neurodegeneration. Gene expression studies of the cerebellum in this model showed downregulation of mRNAs involved in glutamatergic and GABAergic neurotransmission and intracellular calcium signalling, as well as upregulation of pro-apoptotic genes. In general, behavioural tests indicated phenotypic similarities with MJD patients (Chou et al., 2008).

Boy and co-workers developed another mouse model that conditionally expresses human mutant atxn-3 isoform-3c containing 77 glutamines, using the TET-off system under the control of the hamster prion promoter. This system was used to determine if symptoms could be rescued after stopping the expression of human mutant atxn-3. These mice had a strong expression of human mutant atxn-3 in the cerebellum, where intranuclear inclusions could be observed. Also, glial cells of the cerebellum showed a strong expression of atxn-3, which might be leading to the neurodegeneration observed in Purkinje cells, a feature that is not common in MJD patients. Symptoms appeared at 2 months of age, and included clasping, impaired gait and motor deficits (rotarod performance).

Interestingly, phenotypic behavioural traits could be rescued after abrogation of human mutant atxn-3 expression, with mice improving their motor performance. This model demonstrated the evident contribution of atxn-3 to the phenotypic traits of MJD, opening the possibility of developing a therapy that could directly act on mutant atxn-3 expression (Boy et al., 2009).

Later, Boy developed another mouse model of MJD, but this time using a rat huntingtin promoter to express human mutant atxn-3 isoform-3c with 148 glutamines. Symptoms started to appear at 12 months (motor deficit, impaired motor learning), but intranuclear inclusions were only observed at 18 months, supporting the idea that intracellular macroaggregates are not necessary to cause the disease. Purkinje cells showed some dysfunctionality and degeneration, a feature that is not common in MJD patients. This is a model with late onset and slow phenotype progression, mimicking the late onset of MJD patients. However, the prolonged periods that animals need to be maintained before symptoms are observed limit the usability of this model in a research context (Boy et al., 2010).

The newest mouse model of MJD expressing full-length atxn-3 protein isoform-3c was created in 2010. Atxn-3 in this model contains 94 glutamines and is ubiquitously expressed under the control of the cytomegalovirus promoter. This model only shows decreased motor activity and motor incoordination, at 4 to 22 months of age. Regarding neurodegeneration, the thalamus, and dentate and pontine nuclei were shrunk, and astrogliosis was also observed. However, intranuclear inclusions were not observed even at 19 months, again supporting the hypothesis that inclusions are not necessary for the disease to arise (Silva-Fernandes et al., 2010).

Looking at all available MJD mouse models that express mutant forms of full-length human atxn-3, it is fair to say that the YAC MJD-Q84.2 created by Cemal and collaborators is the one that better recapitulates the human condition, mimicking key features of the disease (characteristic symptoms, neurodegeneration, presence of intranuclear inclusions) and its genetic context. However, the late onset and slow progression of the disease limits its usability, forcing scientists to maintain mice at least for 1 year, in order to perform the phenotypic evaluations. Thus, refining this mouse model would be an interesting strategy to overcome this problem and generate a novel model with improved characterises. Since the number of CAG repeats inversely correlates with the age of MJD onset, and directly correlates with the rate of neurodegeneration, expanding the number of CAG repeats within the *ATXN3* gene may yield a mouse model that displays an earlier onset and a more severe phenotype.

1.3 Gene-based therapies for Machado-Joseph disease

1.3.1 Current perspectives

Presently, MJD is still an uncurable disease, for which the available therapies serve only to improve, as much as possible, the patient’s daily life (Liu et al., 2016). MJD patients rely on drugs that alleviate symptoms, but that do not treat the cause, or stop the progression of the disorder. Baclofen is used to alleviate spasticity, levodopa to reduce parkinsonism symptoms and antidepressants to treat depression or sleeping problems (Cecchin et al., 2007; Nandagopal, 2004).

Preclinical studies using other drugs have also shown promising results. Teixeira-Castro and collaborators demonstrated that serotonin reuptake inhibitors (citalopram) were able to improve some aspects MJD pathology, increasing motor coordination and decreasing atxn-3 aggregation in a *C. elegans* MJD model (Teixeira-Castro et al., 2015). Cunha-Santos and colleagues demonstrated that caloric restriction mimetics such as resveratrol can ameliorate MJD symptoms (including motor improvement) in the transgenic mouse model containing 69 glutamines, via SIRT1 activation (Cunha-Santos et al., 2016).

Nonetheless, promising new perspectives for hereditary diseases have arisen with recent advances in the field of gene therapy. One of the most noteworthy cases until now was the approval, in May 2019, of the first gene therapy treatment for spinal muscular atrophy (SMA) that requires a single intravenous dose – ZOLGENSMA® (AveXis/Novartis, 2019). SMA is a genetic disorder, in the majority of cases caused by a loss-of-function mutation in the *SMN1* gene. Affected infants have a life expectancy of 18 months, but the newly approved gene therapy can reportedly cure the disease. ZOLGENSMA® makes use of adeno-associated virus (AAV) 9 to deliver a functional copy of the gene that encodes for the SMN1 protein (AveXis/Novartis, 2019; Groen et al., 2018).

1.3.2 Machado-Joseph gene therapies based on RNA-targeting

It is evident that there is a need for new therapeutic approaches for MJD, therapies that target the causes of the disease rather than simply alleviate the associated clinical signs. Since MJD is a monogenic disorder, techniques that prevent the expression of the *ATXN3* gene would represent a promising approach.

Since 1998, when Craig Mello and Andrew Fires first published their work in RNA interference, diverse strategies have been used in order to develop molecules capable of targeting a gene's messenger RNA. Years later, in 2018, the Food and Drug Administration (FDA) agency approved the first RNA interference-based (RNAi) therapy to treat hereditary transthyretin amyloidosis, a neurodegenerative disorder caused by mutations in the *TTR* gene that leads to abnormal accumulation of amyloid plaques in the nervous system and at later stages in other organs, such as the heart, kidneys and eyes (Hoy, 2018; Sekijima et al., 2018).

Concerning MJD, though no similar strategies have been yet approved, several studies have developed strategies to target *ATXN3* mRNA and prevent its translation, thus knocking-down atxn-3 protein expression.

In 2008, Alves and colleagues used a lentiviral rat model of MJD to show that allele-specific silencing of mutant *ATXN3* using lentiviral vectors encoding short-hairpin RNA (shRNA) downregulated the mutant human *ATXN3* gene expression, leading to a decrease in atxn-3 inclusions and neurodegeneration. The shRNA targeted a SNP ($G^{987}GG \rightarrow C^{987}GG$) that is present in 70% of patients (Sandro Alves et al., 2008). Broadly, shRNAs are small synthetized double-stranded RNA molecules, 19-22 bp long, that target mRNAs by complementarity, leading to their degradation, via RNA-induced silencing complex (RISC) (Taxman et al., 2010).

Later, in 2013 and 2014, Nóbrega and collaborators showed that the allele-specific shRNA-based strategy was able to rescue neurodegeneration and motor performance both in transgenic and lentiviral cerebellar mouse models of MJD. This approach was able to preserve dendric integrity and Purkinje cell projections. The number of nuclear inclusions of mutant atxn-3 protein, which is one of the MJD hallmarks, was also reduced (Nóbrega et al., 2014, 2013b).

Other methods used to target mRNAs and prevent their translation rely on antisense oligonucleotides (ASOs). Broadly, ASOs are short and synthetic sequences (8-50 nucleotides length) of single-stranded oligodeoxynucleotides that can bind mRNAs and downregulate or modify protein expression (Rinaldi and Wood, 2018). Importantly, FDA approved an ASOs-based therapy for cytomegalovirus treatment in 1998, and since then five more ASOs-based therapies have been approved, including *SPINRAZA®*, an ASO that treats spinal muscular atrophy (Stein and Castanotto, 2017).

Several preclinical studies using ASOs as a treatment for MJD have been conducted. Evers and colleagues demonstrated that, in MDJ patients-derived fibroblast cultures, it is possible to completely skip exon 9 and 10 (which includes the CAG repeats) by targeting them with ASOs. Furthermore, the resulting atxn-3 protein was demonstrated to retain its ubiquitin-binding properties. They went further in their studies and stereotaxically-injected mice with the ASOs, confirming exon 9 skipping in the animals. Though exon 10 showed only a non-significant tendency to be skipped, this strategy constitutes a promising perspective (Evers et al., 2013).

Later, in 2017, Toonen and colleagues further demonstrate the possible therapeutic effects of ASOs. They used MJD derived-patient fibroblasts to show the ability of exon 10-targeting ASOs to induce exon 10 skipping and thereby remove the CAG repeats responsible for MJD pathology. This strategy led to the appearance of a premature stop codon (TGA) right at the end of exon 9, thus giving rise to a truncated atxn-3 variant without the polyQ. Although this truncated variant lacks UIM 3, its capability of binding ubiquitin was retained. Intracerebroventricular injections of ASOs targeting exon 10 in YAC MJD-Q84.2 mice (Cemal et al., 2002) showed similar results regarding exon 10 skipping. The amount of insoluble atxn-3 protein was reduced in cell lysates from the cerebellum, and the number of nuclear inclusions in the substantia nigra was decreased. Despite the promising results, Toonem and collaborators did not perform any behavioural tests to assess motor behaviour of treated mice (Toonen et al., 2017).

In 2018, researchers from Henry Paulson laboratory showed that ASOs can improve motor performance and increase locomotor activity of YAC MJD-Q84.2 mice, but the ASOs that they used targeted instead the 3' UTR of atxn-3 mRNA. The ASOs led to the knock-down of mutant human atxn-3 protein and the endogenous mice atxn-3, contrary to what Toonem and colleagues achieved, since they were able to maintain expression of a ASOs-generated truncated variant of the atxn-3 protein. Electrophysiological tests on Purkinje cells suggested an ameliorated firing frequency upon ASOs treatment (McLoughlin et al., 2018).

Alves and co-workers showed that allele-nonspecific atxn-3 silencing (silencing of both mutant a endogenous atxn-3) reduces striatal neuropathology, in a MJD lentiviral rat model (Alves et al., 2010). However, this approach is less desirable than a pathogenic allele-specific one, since the non-expanded atxn-3 protein may play important cellular functions (see section *Atxn-3 protein and Machado-Joseph disease*). Some of the above studies demonstrated that, by making use of the $G^{987}GG \rightarrow C^{987}GG$ SNP present in 70% of MJD patients, it is possible to selectively target the MJD-associated mutant atxn-3 mRNA (Sandro Alves et al., 2008). The exon 10 skipping strategy, in which the CAG repeats are not translated, and thus there are no polyQ tract present at atxn-3 protein sequence, also seems a promising therapeutic approach that does not compromise the cellular function of the protein.

RNA-targeting therapies have shown promising results both *in vitro* and *in vivo* and may in the future be transitioned for clinical trials.

1.3.4 Toxicity of expanded CAG sequence-bearing mRNA molecules

In the last few years several reports have shown that mRNA molecules containing abnormal CAG repeats may be toxic. In 2008, Li and colleagues showed in *Drosophila* that untranslated abnormal CAG repeats lead to neurodegeneration, which can be reverted if the CAG-CAG tract is altered to CAA-CAG repeats (also coding for glutamines) (Li et al., 2008). Later, Wang and collaborators showed in *C. elegans* that increasing the number of RNA CAG repeats increases toxicity in a length-dependent manner (Wang et al., 2011). In a mouse model of Huntington's disease, it was show that untranslated CAG repeats in huntingtin (the causative protein) mRNAs can lead to cell disfunctions and death. Studies have also been suggesting that abnormal CAG repeats in RNA can lead to defects in alternative splicing, abnormal RNA subcellular location and sequestering of siRNA proteins dysregulating gene expression (Martí, 2016; Nalavade et al., 2013; Wojciechowska and Krzyzosiak, 2011).

Despite the promising results demonstrated by gene therapy strategies that rely on RNA-targeting to treat MJD, the toxicity of expanded CAG sequence-bearing mRNA molecules highlights the need for therapeutic strategies that act at a pre-transcriptional level, blocking or altering the synthesis of expanded mRNAs. In the next section we will discuss the potential of the CRISPR/Cas9 system to directly target DNA and how it can be used to perform gene therapies pre-transcriptionally.

1.4 CRISPR/Cas9-based gene therapy tools

1.4.1 The biology of CRISPR/Cas9

Clustered regularly-interspaced short palindromic repeats (CRISPR) were first reported by Ishino in 1987, and later by Francisco Mojica in 1993. The term describes a series of tandem repeats with 30-34 bp, spaced by 35-39 bp unique sequences (spacers), that were found in the genome of *Escherichia coli and Haloferax mediterranei* (Ishino et al., 1987; Mojica et al., 1993). Initially, researchers did not know the biological relevance of those repeats or if they played any particular function.

The term CRISPR was first used just in 2002 by Ruud Jensen, who described the CRISPR *locus* and its associated genes – the CRISPR-associated (*Cas*) genes. No conclusive function for the proteins coded by the Cas genes was given, though predictions aimed for helicases and exonucleases (Jansen et al., 2002). Only in 2005, after a study that described CRISPR as sequences that derive from exogenous genetic elements, Mojica and collaborators came with the hypothesis that CRISPR *loci* might be a part of a bacterial adaptive immune system against bacteriophages, because bacteria that contained CRISPR sequences equal to genetic sequences of bacteriophages were not infected by them (Mojica et al., 2005).

Two years later, Rodolphe Barrangou and his team proved that CRISPR is an adaptive immune system of prokaryotes. In their studies they showed that bacteria were capable of integrating bacteriophage genomic material into the CRISPR *locus* after infection. This material constituted the observed spacers. They also observed that the integrated bacteriophage genomic sequences were random, without specific preference. Finally, they demonstrated that CRISPR spacers were responsible for the bacteriophage resistance, by conducting a series of experiments where spacers were added or deleted, in order to manipulate the bacterial resistance against a certain group of bacteriophages (Barrangou et al., 2007).

The CRISPR/Cas system may differ in its biogenesis pathways depending on the particular species of prokaryotic organism it is found in, and each species may have more than one system. Shortly, there are two classes of the CRISPR/Cas system. The Class 1 which contains types I, III and IV, and Class 2 which contains types II, V and VI. All of them rely on *cas1* and *cas2* proteins, on a CRISPR RNA (crRNA) molecule and a protein that assures the degradation of the exogenous DNA. Beyond that, CRISPR/Cas systems types differ in the mechanisms by which the protein complexes involved are formed, the proteins that constitute them and on how they assure immunity. CRISPR/Cas systems are grouped into those classes depending on specific genes they express along with the *cas1* and *cas2* (Mazhar Adli, 2018; Rath et al., 2015).

However, the Type II CRISPR/Cas system from *Streptococcus pyogenes* is the most well know and widely studied (**Figure 5**). The corresponding CRISPR *locus* is composed by a trans-activating crRNA (*tracrRNA*) gene, the *cas* genes (*cas9*, *cas1*, *cas2* and *csn2*) and the CRISPR array (leader, repeats and spacers). The specific function of each gene present in the CRISPR *locus* remains undetermined. However, it is known that *cas1*, *cas2* and *csn2* genes code for proteins that form a machinery complex needed for spacer acquisition and integration into the CRISPR array. The *cas9* gene codes for a protein – Cas9 – with nuclease activity capable of inducing double-stranded breaks in the exogenous DNA (Ka et al., 2016; Mali et al., 2013a; Nuñez et al., 2014). After acquisition and integration of exogenous DNA, the spacers will define targets in the foreign DNA sequence (protospacers) marked for destruction by the activity of Cas9 (**Figure 6A**). For that, all the repeats and all the spacers are transcribed into a long pre-CRISPR-RNA (pre-crRNA). The tracrRNA, which is also transcribed, binds by complementarity to repeats present in the pre-crRNA and recruits RNase III that cleaves the pre-crRNA into small mature crRNAs, every one of them formed only by one repeat-spacer sequence. The Cas9 endonuclease protein forms an endoribonuclease complex with the tracrRNA and the crRNA (tracrRNA:crRNA-Cas9), creating an unit capable of targeting the foreign protospacer by Watson and Crick base-pairing, cleaving it by inducing a double-stranded break (**Figure 6B**).

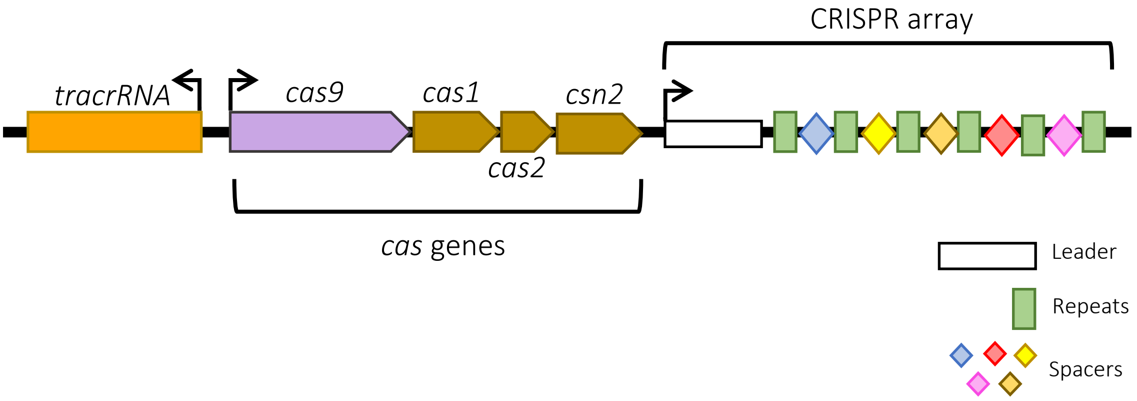

Figure 5 - Schematic representation of the Type II CRISPR *locus* of *Streptococcus pyogenes*. The CRISPR *locus* of *Streptococcus pyogenes* is composed by the *tracrRNA* gene, the *cas* genes and the CRISPR array.

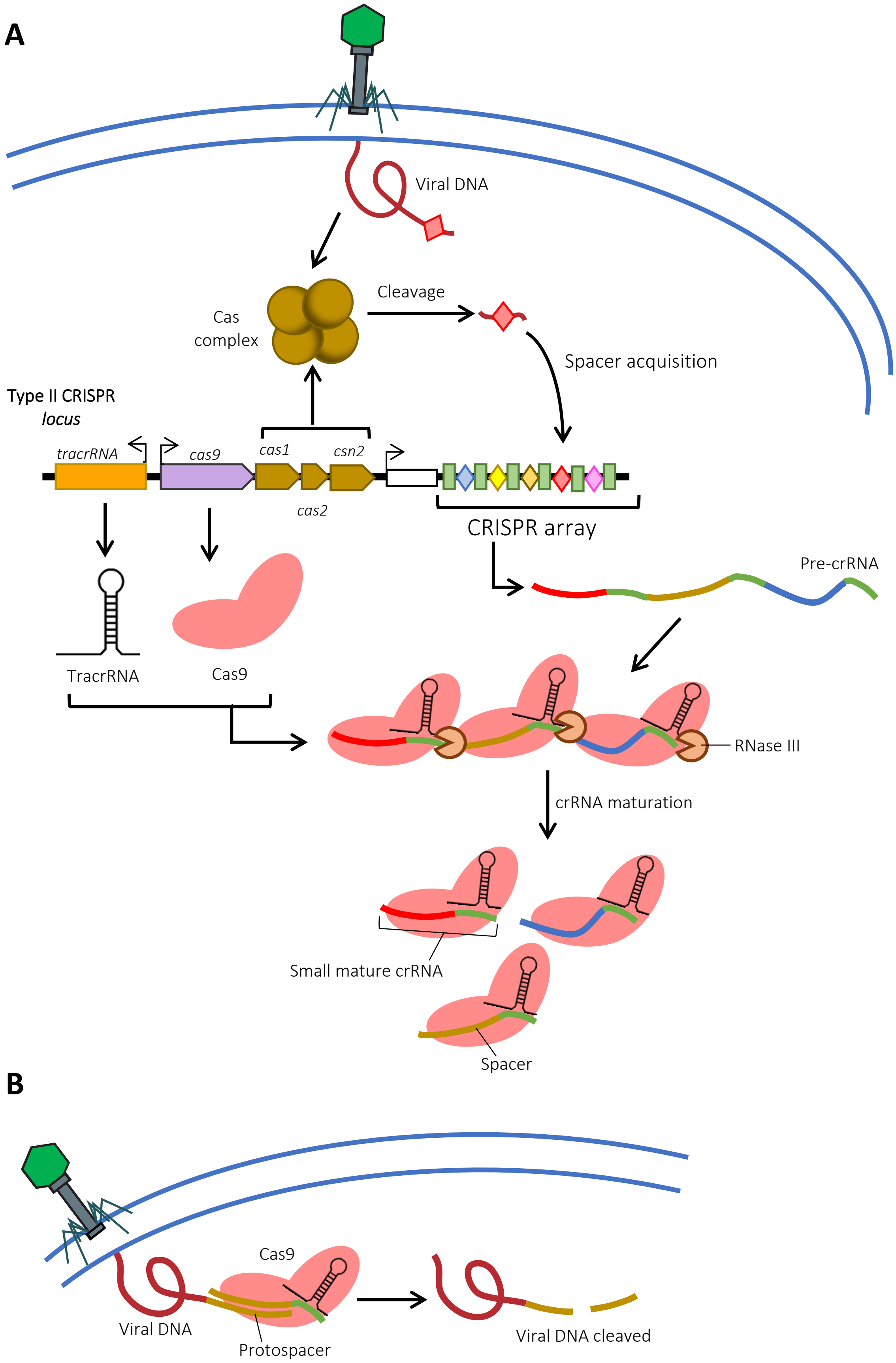
A
Viral DNA
Cas
complex
Cleavage
Spacer acquisition
Type II CRISPR
locus
tracrRNA
cas9
cas1
csn2
cas2
CRISPR array
Pre-crRNA
TracrRNA
Cas9
RNase III
crRNA maturation
Small mature crRNA
Spacer
B
Cas9
Viral DNA
Protospacer
Viral DNA cleaved

Figure 6 - Schematic representation of Type II CRISPR/Cas9 mechanism of action in prokaryotic cells. A) Immunization is the first step needed for the immune response against bacteriophages. After infection, the viral DNA is injected into the bacteria and the Cas proteins form a complex that is responsible for acquisition and integration of the spacer into the CRISPR array. The sequences in the CRISPR array (repeats and spacers) are transcribed into a pre-crRNA molecule that undergoes maturation after binding the tracrRNA and the Cas9 protein. (Figure description continues next page).

Figure 6 - With the help of RNase III, that cleaves the pre-crRNA, a single complex is formed containing a small mature crRNA, a Cas9 protein and a tracrRNA (tracrRNA:crRNA-Cas9). **B)** The second step, immunity, is carried out by the tracrRNA:crRNA-Cas9 complex. The spacer recognizes the foreign DNA (protospacer), binding it by Watson and Crick base-pair complementarity, and inducing a double-stranded break catalysed by Cas9, but only if a PAM sequence is present

1.4.2 CRISPR/Cas9 as a tool for gene editing

The ability to biotechnologically introduce precise double-stranded breaks in the DNA is invaluable for many methodologies aiming at editing genes and genomes. Double-stranded breaks can be repaired by the cells' endogenous DNA repair mechanisms, and two pathways are prominent in this process: nonhomologous end-joining (NHEJ) and homology-directed repair (HDR). These two repair pathways are fundamental to perform gene editing. NHEJ, an error-prone mechanism, frequently entails the insertion or deletion of a small number of nucleotides at the break site (indels), which can lead to frame-shift mutations and the eventual appearance of premature stop codons, allowing gene knock-outs or translation of truncated proteins. Therefore, a double-stranded break induced by simply targeting a nuclease to a desired *locus* can alter a gene sequence and generate a gene knock-out or truncate a protein. On the other hand, HDR is a more conservative repair mechanism, in which cells use a homologous template that will be integrated at the double-stranded break site to repair the break. Consequently, inducing a double-stranded break using a targeted nuclease and providing a donor homologous template with a desired sequence allows the insertions of particular sequences at a particular site of interest (Liang et al., 1998; Sander and Joung, 2014; Wang et al., 2016).

Nuclease-based gene editing approaches have been used since the 80's. At the time, meganucleases were the first proteins that were employed. These large proteins, first discovered in yeasts, are able to recognize DNA sequences of up to 40 bp, making them very specific and allowing few off-targets. However, this feature also constitutes a limitation since the sequence recognized might not be common in the target genome. Another limitation is its engineering in order to adapt them to target other DNA sequences. Researchers are forced to study the protein structure in order to redesign it and assure the protein is able to target a novel site and retain cleavage activity. Since the recognition and cleavage domain are not independent this process is hard to accomplish (Puchta et al., 1993; Puchta and Fauser, 2014).

In 1996, zinc-finger nucleases (ZFNs) were engineered. They are formed by a nuclease domain (FokI) fused with a zinc-finger array composed by three different zinc-finger units, each one of them able of recognize 3 different nucleotides, in total recognizing 12-32 nucleotides. In order to perform the double-stranded break, the system has to form a dimer. This system shows more versatility compared to meganucleases, since it can easily target different sequences by varying the zinc-finger units. However, ZFNs show some off-target activity leading to undesired cleavage sites and consequently undesired mutations. The fact that this approach also relies on protein engineering makes the design and development of new ZFNs a time-consuming process (Hsu et al., 2014; Kim et al., 1996; Smith et al., 2000).

Transcription activator-like effector nucleases (TALENs) were engineered later, in 2010. They also have a FokI nuclease domain and a DNA-binding domain (found in *Xanthomonas* bacteria) composed by an array of different DNA-binding units, each one of them capable of binding a single nucleotide for a total of 30-40 bp. Comparing with the previously described nucleases, TALENs are the most versatile, being virtually able to target any genetic *locus*. Also comparing with the ZNFs system, TALENs show fewer off-target activity, since off-target activity can be modulated by altering the size of the DNA-binding domain (Christian et al., 2010; Gaj et al., 2013; Kim and Kim, 2014; Puchta and Fauser, 2014).

It was only in 2013 that the CRISPR/Cas9 system was used for the first time to target and modify human and mice cells' DNA. This platform would soon reveal itself as the most versatile and feasible nuclease-based gene editing system. The system does not need any kind of protein engineering and targeting relies only on the design of RNA molecules complementary to the desired site. By making use of CRISPR/Cas9 ability to elicit precise targeted DNA double-strand breaks, altering genes and genomes became simpler, leading the scientific community to an unprecedented new era of gene editing. However, the system shows more off-target events than TALENs and researcher have been trying to overcome this limitations since the CRISPR/Cas9 system was first introduced (Cong et al., 2013; Zhang et al., 2015).

Table 1 - Comparison between the different programmable nucleases used in gene editing. Despite its low cost, easy design and simple use, CRISPR/Cas technology shows highly frequent off-target effects. However, the advantageous compromise between versatility, cost and feasibility in detriment of the possibility of off-target effects makes CRISPR/Cas the most appealing nuclease platform used for gene editing.

	MEGANUCLEASES	ZFNS	TALENS	CRISPR/CAS
Target	≈ 40 bp	12-32 bp	30-40 bp	20-22 bp
Mode of action (interaction)	DNA-protein	DNA-protein	DNA-protein	DNA-RNA Watson-Crick
Feasibility	Hard	Relatively hard	Moderate	Easy
Off-target	Very low	Moderate	Low	High
Cost	High	High	Moderate	Low

Le Cong and co-workers (from Feng Zhang's lab) were the pioneers of using the CRISPR/Cas9 as a tool for gene editing. They fused the Cas9 protein with two NLS to make sure that the protein would be translocated to the nucleus and remain there and created a spacer that would target a protospacer within the *EMX1* gene. By expressing the codifying complex in human embryonic kidney (HEK) 293T cells, they observed the induction of a DNA double-stranded break. Finally, they were able to integrate two restriction sequences (for HindII and NheI) into the *EMX1 locus*, and also delete a region of the gene (Cong et al., 2013).

By the same time, another group performed similar experiments. They used 293T cells that had been previously integrated with a disrupted green fluorescent protein (GFP) that did not give rise to a fluorescent signal. Using CRISPR/Cas9 targeting the flanking sides of the disrupted site and providing a non-disrupted template of the GFP sequence, they were able to restore the gene by HDR, and consequently reinstate florescence. This group also targeted two different *loci* using the CRISPR/Cas9 system and were able to induce mutation by NHEJ. Interestingly, in another set of experiments using human cells, the research team was able to integrate a DNA sequence at the AAVS1 *locus* containing a puromycin resistance gene and a GFP gene. The integration was confirmed upon incubating the cells with puromycin for 2 weeks. Similarly to what had also been described by Le Cong, they confirmed that the CRISPR system was able to target different protospacers at the same time (multiplexing) (Mali et al., 2013)

These two articles proved the possibility of successfully performing gene editing in human cells using the CRISPR/Cas9 technology, with an ease that was never seen before. Of course, the scientific community came up with optimizations and new ideas, “playing around” with the technology and adapting the CRISPR/Cas9 system for other purposes other than gene editing.

Importantly, it is worth noting that, since its discovery, the CRISPR/Cas9 system has been simplified for biotechnological use. The tracrRNA:crRNA-Cas9 complex was manipulated in order to create a single RNA molecule that could form a complex with Cas9, instead of having the tracrRNA and the crRNA separately. A chimeric RNA molecule was engineered by fusing the 3′ end of the crRNA to the 5′ end of the tracrRNA in order to form a hairpin structure that retains the complementarity between the tracrRNA and the crRNA. This resulted in the generation of a single-strand RNA called single-guide RNA (sgRNA) that is able to target the Cas9 to a particular DNA *loci* by itself (Jinek et al., 2012).

1.4.3 CRISPR/Cas9 as a tool for gene expression regulation

*S*everal CRISPR/Cas9 adaptations have been engendered since the development of this gene editing system. Possibly the first one was made by mutating the catalytic domains of the Cas9 protein: RuvC and HNH. Since each one of them cuts a different DNA strand, mutating these sites gives rise to versions of Cas9 in which one or both catalytic domains are inactivated (**Figure 7A**). The process is simple, and requires only point mutations in the Cas9 protein sequence: a substitution of an aspartate residue by an alanine at position 10 (D10A), for the RuvC site, or a substitution of histidine at position 840 by an alanine (H840A), for the HNM site. Mutation of both sites leads to the complete inactivation of Cas9 (dead Cas9; dCas9), while mutation of just one domain gives rise to nickase Cas9 proteins, which produce single-strand DNA breaks (Gasiunas et al., 2012; Jinek et al., 2012).

Since dCas9 retains its ability to target and bind DNA, fusing dCas9 with a transcriptional repressor or activator allows targeted regulation of gene expression, without induction of a double-stranded breaks (**Figure 7B**). The process is relatively simple: it requires the fusion of dCas9 with the transcriptional regulator and the design of a sgRNA with a spacer that targets the desired gene. The Krüppel-associated box (KRAB) domain is an example of repressor that is fused to the dCas9 to induce gene repression (CRISPR interference; CRISPRi). It recruits chromatin-modifying complexes that lead to gene silencing. In 2013, the first work using dCas9-KRAB to successfully induce gene repression by *in vitro* lentiviral expression of the CRISPRi system was published. The system targeted an expressing GFP plasmid and endogenous HeLa cells genes (CD71 and CXCR4) (Gilbert et al., 2013).Regarding transcription activation, the VP64-p65-Rta (VPR) complex is an example of an activator that, when fused with dCas9, is able of activate gene expression (Chavez et al., 2015).

The creativity of the scientific community did not stop with Cas9 variants that were able of regulate gene expression by fusing dCas9 to transcription repressors or activators. Scientist came up with new ideas to repurpose and create new Cas9 variants.

1.4.4 Expanding CRISPR/Cas9 potential

The imagination and curiosity led scientists to engineer alternative Cas9 applications. Even the epigenome can be subjected to editing, by using a version of dCas9 fused with a histone or DNA-modifier enzyme (**Figure 7C**). The DNA methyltransferase 3 (DNMT3a) is a protein capable of methylating the cytosine-guanine (CpG) sites in the genome. Fusing DNMT3a with dCas9 and directing the complex to a desired *locus* can increase the methylation of the targeted region, thus editing the epigenome (Stepper et al., 2017). The histone demethylase LSD1 or histone acetyltransferase p300 are examples of histone modifiers that can be fused with dCas9 to modify the epigenome at the histone level, removing methyl groups from the histone tails or adding acetyl groups, respectively (Hilton et al., 2015; Kearns et al., 2015).

Another interesting approach it to fuse dCas9 to a fluorescent protein, such as GFP. This offers the possibility of determining the position of specific genes within the nucleus and of visually following a chromosome during mitosis or meiosis, for example (Chen et al., 2013).

Nowadays, it is even possible to edit a particular base pairs without inducing DNA double-stranded breaks, by fusing dCas9 with a adenosine deaminase (TadA) or a cytosine deaminase (AID), which are able to exchange an adenine for a guanine (A•T to G•C) or a cytosine for a thymine (C•G to T•A) nucleotide, respectively (**Figure 7D**) (Gaudelli et al., 2017; Nishida et al., 2016).

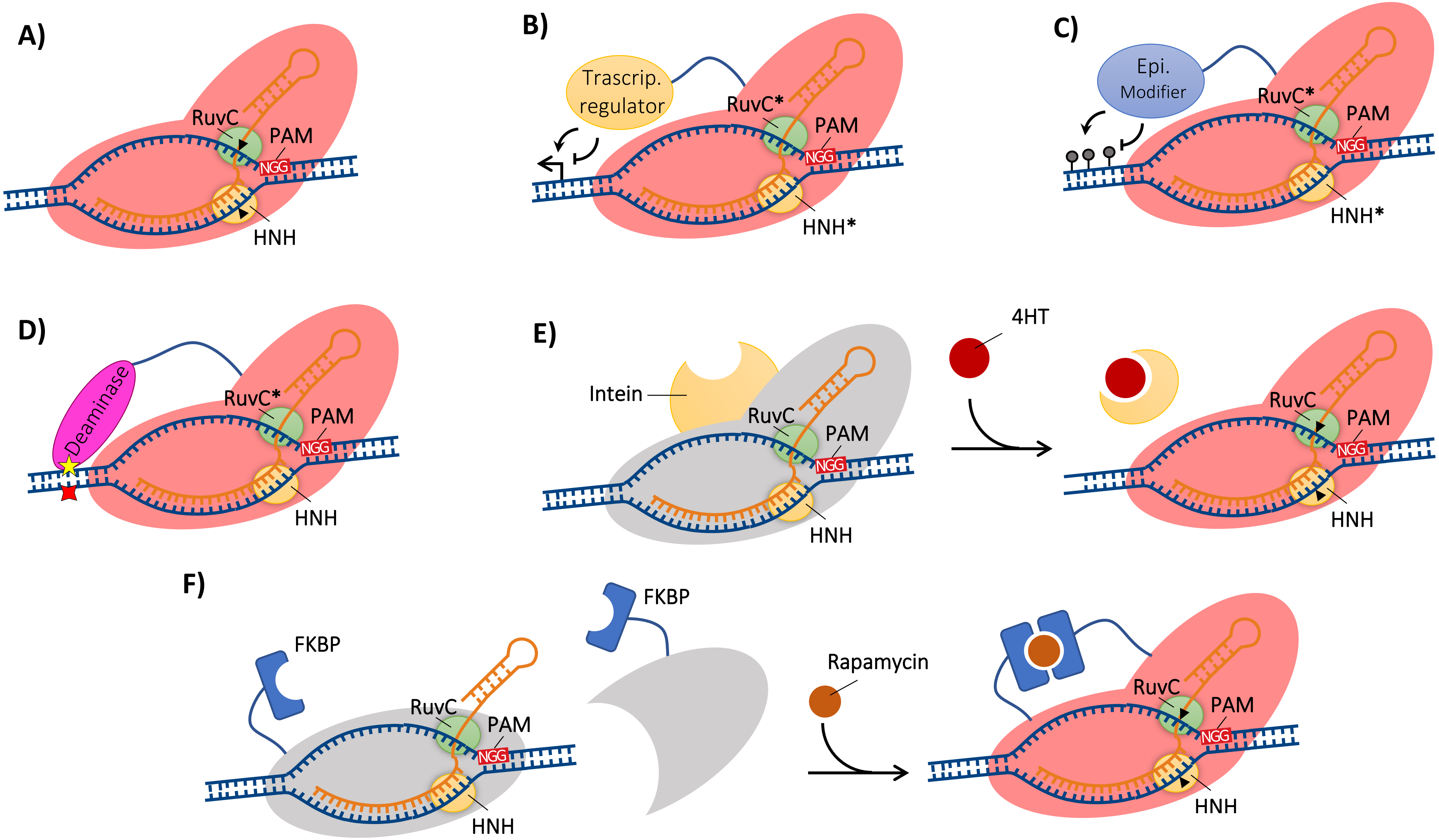
A)
RuvC
PAM
NGG
HNH
B)
Trascrip. regulator
RuvC*
PAM
NGG
HNH*
C)
Epi. Modifier
RuvC*
PAM
NGG
HNH*
D)
Deaminase
RuvC*
PAM
NGG
HNH
E)
Intein
RuvC
PAM
NGG
HNH
4HT
RuvC
PAM
NGG
HNH
F)
FKBP
RuvC
PAM
NGG
HNH
FKBP
Rapamycin
RuvC
PAM
NGG
HNH

Figure 7 - Schematic representation of CRISPR/Cas9 DNA double-stranded break mechanism and the variations of the CRISPR/Cas9 system adapted from *S. pyogenes*. A) CRISPR/Cas9 ribonucleoprotein complex, composed by the Cas9 protein (pink) and a sgRNA (dark yellow). In order to target a desired region, the protospacer (blue) needs to have a PAM sequence immediately downstream (5'-NGG-3'). (Figure description continues next page).

Figure 7 - The double-strand break (▲) is induced by two catalytic domains, RuvC (green circle) and HNH (yellow circle), 3 base pairs upstream of the PAM. **B)** dCas9, in which the catalytic domains were mutated (*), is able to activate or repress gene expression when fused with transcription regulators. Fusing dCas9 with a VP64-p65-Rta can activate gene expression, while fusing it with a KRAB domain can induce gene silencing. **C)** dCas9 fused with epigenetic modifiers alters the epigenome. dCas9 can be fused with DNMT3a to increase DNA methylation of CpG sites or be fused with histone modifiers to add methyl or acetyl groups to the histone tails (LSD1 and p300, respectively). **D)** Editing only a base pair is possible by fusing dCas9 with a deaminase. **E)** An inducible form of Cas9 is inactivated (grey) by an intein segment that can be excised after 4-HT administration, activating Cas9. **F)** A split and inactive Cas9 (grey) in which each terminal fragment is fused with FKBP or FRB domains can restore it activity after rapamycin administration, since it induces the dimerization of the Cas9 fragment pairs.

It has also been possible to engineer an inducible version of the Cas9 system, by providing an activator that temporally regulates its activity and that may consequently be used to decrease off-target activity. An example of this type of system is intein-Cas9, a chimera formed by Cas9 fused with an excisable protein segment (intein) that actively disables the complex. After addition of 4-hydroxytamoxifen (4-HT), the intein is spliced allowing Cas9 to perform its function (**Figure 7E**) (Davis et al., 2015).

A mechanism for activating Cas9 with ultraviolet light has also been developed. This approach requires the inactivation of the Cas9 through the substitution of lysine 866 by a photocaged lysine (pyrrolysine). After UV exposure the pyrrolysine is converted back into a lysine residue, restoring Cas9 activity (Hemphill et al., 2015). There is yet another strategy for Cas9 light activation, using blue light instead. It consists of a split Cas9 in which a N-terminal fragment of the protein and the separated C-terminal fragment are each fused with a photoinducible dimerization magnet (pMag and nMag, respectively). Upon blue light stimulation, the N- and C- terminals tend to fuse and Cas9 activity is restored (Nihongaki et al., 2015). Using the same split strategy, it is also possible to induce Cas9 activity by administration of rapamycin (**Figure 7F**). The Cas9 N-terminal fragment is fused with a FKBP-rapamycin binding (FRB) domain and the C-terminal fragment is fused with FK506-binding protein 12 (FKBP). After rapamycin administration the complex dimerises and Cas9 is activated (Zetsche et al., 2015). Finally, another method for regulating Cas9 activity is to manipulate its degradation rate, which can be increased by fusing it with a destabilizing dihydropholate reductase (DHFR) domain that signals the complex for degradation. Trimethoprim (TMP) treatment, which inhibits the function of the DHFR domain, allows Cas9 to persist and remain active (Maji et al., 2017).

It is evident that researchers put great efforts into optimizing the CRISPR/Cas9 system and have come up with interesting variants that adapt and repurpose the technology depending on the experimental approach that is required. This subject is thoroughly reviewed by (M. Adli, 2018) and by (Wu et al., 2018).

1.4.5 Generation of mouse models using CRISPR/Cas9

Generation of mouse models is a fundamental step in order to have a tool that allows researchers to get insight into molecular and cellular mechanisms of human diseases. Prior to the introduction of nuclease-based gene editing platforms, methods to create new animal models were necessarily long and involved several intermediate steps.

Classical transgenic mouse generation requires cultures of mouse embryonic stem (ES) cells that need to be transfected with DNA constructs containing the DNA sequence that introduces the desired mutation or DNA inserts and a selection cassette. ES cells are then selected, collected and injected directly into mouse blastocysts or aggregated with the morula. Finally, embryos are implanted in pseudopregnant female mice. However, because the method is done after oocyte and sperm pronuclei fusion and there is already some self-proliferation, the progeny is chimeric. Resulting animals need to be further crossed with wild-type individuals to ensure germline transmission and a non-chimeric next generation (Kumar et al., 2009; Longenecker and Kulkarni, 2009; Tanaka et al., 2009). Although this process has been used for many decades and with solid results, it is expensive and time-consuming. Employing nuclease-based gene editing platforms – and in particular the easily-engineered CRISPR/Cas9 systems – has turned out to be an easier and less expensive approach to generating transgenic mouse models.

The potential of the CRISPR/Cas9 system is boundless. Besides allowing gene editing in cell lines, it has also allowed gene editing of animal embryos, opening the possibility of generating and refining animal models of disease in an unprecedented simple and straightforward manner. The process only requires the Cas9 protein, a RNA molecule targeting the desired genetic *locus* and eventually a homologous template containing the genetic mutation or insert that is to be introduced. Fertilized zygotes are collected from female mice and the Cas9 protein, RNAs (crRNA and tracrRNA or, alternatively, the sgRNA) and the template are introduced into the pronucleus by microinjection or electroporation. The endogenous DNA repair mechanisms are then expected to generate the desired genetic alterations: NHEJ is able to elicit the generation of knockouts and deletions, while integration of particular mutations and inserts is mediated by HDR, in the presence of the homologous repair template. Finally, the embryos are re-implanted in pseudopregnant females and after birth pups are genotyped to verify the presence of the desired gene alterations. Even after the genetic edition has been successful confirmed, model characterization must be performed to ensure that mice recapitulate the disease in a way that can be translated to humans. Using CRISPR/Cas9 to generate transgenic mouse models spares scientists many intermediate steps and, because the procedure is preformed prior to pronuclei fusion, chimeres are not expected. Therefore, progeny will have all their cells edited and further crossings are not required (Birling et al., 2017; Wang et al., 2013).

Gene editing has already been used to edit human embryos and correct mutations associated with disease. As an example, a study concerning a heterozygous mutation in the *MYBPC3* gene associated with cardiopathy has shown that, using a sgRNA targeting the mutated region, Cas9 and a plasmid containing the corrected sequence, elicited gene correction by HDR (Ma et al., 2017).

Correction has also already been attained in the context of a polyQ disease. Recently, a research group was able to correct the expanded CAG repeat region within the *HTT* gene in Huntington's patient-derived induced pluripotent stem cells (hiPSCs). They used CRISPR/Cas9, a sgRNA and a template containing the smaller CAG sequence, that was integrated into the mutated *loci*, therefore correcting the expanded repeat. The hiPSCs retained their pluripotency and after differentiation into neurons their electrophysiology properties were maintained (X. Xu et al., 2017). This result supports the idea that it is possible to perform gene editing of CAG repeat-containing regions, in order to correct or overexpand the number of repeats.

1.4.6 CRISPR/Cas9 towards Machado-Joseph disease

Taking these ideas into consideration, CRISPR/Cas9 opens the possibility of developing new therapeutic strategies for human diseases. Using CRISPR/Cas9 to target a disease-causative gene of a genetic disorder such as MJD may allow the correction of the pathogenic mutation or the modulation of its expression. In addition, CRISPR/Cas9 may represent an inexpensive, precise, fast and simple way to generate new or refined disease models. In the case of MJD the YAC MJD-Q84.2 model may be further enhanced to display an earlier onset and a more severe phenotype by using CRISPR/Cas9 as a tool for gene editing in order to overexpand the CAG repeats of the human *ATXN3* gene, present in the transgenic mouse. Animal models that properly mimic the phenotypic traits and the cellular and molecular disease mechanisms of human disease conditions are fundamental for an accurate biomedical research aiming at understanding pathology or developing new therapeutic approaches for MJD.

Recent studies indeed support the possibility of excising the abnormal CAG repeats from genes causing polyglutamine disorders. In two different studies, one studying MJD, and another studying HD, investigators used a pair of guided RNAs targeting the vicinities of the abnormal CAG repeat. In the study targeting mutant *HTT* (the gene associated with HD), Cas9 led to depletion of the mutant gene *in vitro* (in HD patient-derived cells). For MJD, the mutant *ATXN3* gene was edited in MJD patient-derived iPSCs. The methodology used resulted in the deletion of exon 10, while exon 9 and 11 were joined together. This led to the generation of a premature stop codon right at the beginning of exon 11, which in turn resulted in the expression of a truncated variant of atxn-3 lacking UIM 3. However, the new variant retained its capacity to bind ubiquitin and edited cells maintained pluripotency properties, being able to differentiate into neural cells after CAG removal (Dabrowska et al., 2018; Ouyang et al., 2018).

The development of new platforms, simpler than animal models, to study MJD, has also become easier with the CRISPR/Cas9 advent. Isogenic iPSCs in communion with CRISPR/Cas-based genome editing tools have the potential of refining the *in vitro* studies of MJD. Easily available cells such as fibroblasts can be collected from patients and genetically corrected with CRISPR/Cas tools, yielding a corrected cell population and a MJD cell population genetically differing only in the MJD CAG expansion mutation (isogenic cell lines). iPSCs can then be reprogrammed and differentiated into cerebellar neurons. This technique can also be used to create isogenic cell containing different numbers of CAG repeats, as a platform for evaluating how this number affects cellular and molecular mechanisms of MJD (Malankhanova et al., 2017; Xiaohong Xu et al., 2017).

Gene editing has the potential to correct gene mutations and edit the known causative genes associated with particular disorders has the potential to completely cure a disease, especially in the case of monogenic disorders. However, the possibility of off-target effects is still a problem that haunts the scientific community. The case of the "Bubble boys", who have a condition called severe combined immunodeficiency (a genetic disorder cause by mutations in the *IL2RG* gene characterized by the lack of most types of immune cells) is well known. They were treated with gammaretroviral vectors encoding a functional copy of the *ILR2G* gene. Though 9 out of 10 were treated, 4 out of 9 developed leukaemia, due the random integration of the coding sequences near proto-oncogenes (Hacein-Bey-Abina et al., 2008). Patients were treated, but the side-effect weighted too much for the therapy to be viable. To date, using CRISPR/Cas9 as a tool for gene editing to treat disease may still be considered a leap of faith, because researchers cannot yet understand the full mechanisms by which this technology can induce undesired mutations, at the risk of leading to other diseases.

However, a safer approach may lay in the previous described dCas9-KRAB (see section *CRISPR/Cas9 as a tool for gene expression regulation*). dCas9-KRAB mechanism of action does not rely on DNA double-stranded breaks, thus, off-target effects are less concerning, as it does not induce unexpected mutations such as indels. In fact, previous results from our laboratory showed that dCas9-KRAB is able of specifically silence the mutant allele by coinfecting HEK cells with lentiviral particles codifying human expanded antxn-3 and particles codifying the dCas9-KRAB targeted to the pathogenic gene. The silencing elicited a robust and long-term decrease of the expanded atxn-3 protein levels (Pena, 2018).

MJD is still a uncurable disease with poor, limiting symptomatic, treatments. Further studies must be carried to gain insight into the cellular and molecular mechanisms that lead to MJD as well as studies that intend in develop therapeutic strategies aiming at delaying the disease progression or actually curing it.

1.5 Objectives

As a monogenic disorder, MJD is a well-suited candidate to employ gene therapy-based approaches. The abnormal CAG expansion is known to be the causative genetic mutation leading to the pathology of MJD, through the translation of an atxn-3 protein that gains a toxic function. Moreover, it is acknowledged that the greater the number of CAG repetitions, the earlier the onset of symptoms and the greater their severity. To prevent the pathological effects observed in MJD, several approaches have been developed to avoid or modulate atxn-3 protein translation, both *in vitro* and in *vivo*. Such strategies almost always rely on targeting atxn-3 mRNA after it is transcribed, but the fact that the CAG-expanded mRNA also seems to have toxic properties urges for approaches that act pre-transcriptionally. Additionally, to study new potential therapies and to better assess the molecular mechanisms of MJD pathology it is important to have an appropriate disease model that can recapitulate, as much as possible, the human condition. During the past years, several mouse models of MJD have been generated, but only the YAC MJD-Q84.2 mouse model expresses a full-length human *ATNX3* gene that contains all introns, exons and regulatory regions. Moreover, it is the one that betters recapitulates the human pathology. However, the YAC MJD-Q84.2 mouse model displays a late onset and a mild phenotype, hindering its applicability in laboratory.

As the potential of CRISPR/Cas9 technologies to operate genetic modifications and modulate gene expression is being increasingly established, we aimed at taking advantage of this platform in the context of MJD therapy development and disease model generation. The objectives of this work are divided into two parts a) developing a molecular strategy for editing the CAG repeat regions present at the exon 10 of the human *ATXN3*, in order to overexpand or decrease the number CAG repeats; b) assess the effect of dCas9-KRAB as a means to specifically silence the mutant *ATXN3* allele *in vivo*, in a mouse model of MJD.

The establishment of a strategy for *ATXN3* CAG overexpansion would serve as a base for the future refinement of the YAC MJD-Q84.2 mouse model; increase of the number of CAGs would generate a more suited model for use in laboratory. Previous works supports the premise that increasing the number of CAGs accelerates the symptoms onset and increases the severity of the phenotype (Abe et al., 1998; Jardim et al., 2001; Klockgether et al., 1996; Onodera et al., 1998). CAG retraction intended to establish a method for generating isogenic MJD patient-derived iPSCs lines, through correction of abnormal CAG tracts. This would contribute to the development of a robust *in vitro* system to model cellular and molecular disease mechanisms, that would also serve as a platform to study new therapeutic approaches, in which the control cells differ genotypically only in the mutation that causes the disease. Both refinement of the YAC MJD-Q84.2 mouse model and the generation of isogenic cell lines would rely on gene editing tools and HDR as main strategies to retract or overexpand the CAG repeats, and it was the aim of this work to develop those strategies, *in vitro*.

Finally, the second part of this work aimed at assessing the *in vivo* therapeutic effects of catalytic inactive Cas9 (dCas9) fused with a KRAB protein as an approach to *ATXN3* silencing, in the context of an animal model of the disease. This strategy followed up on previous results by our group, which proved the ability of a CRISPRi strategy targeting a disease-associated SNP (present in 70% of MJD patients) to specifically silence mutant *ATXN3* in cell cultures. The same molecular tools were administered to a MJD mouse model, with the aim of evaluating their effects on *ATXN3* expression, MJD-related neuropathological marks and mice motor performance.

Chapter 2 – Materials and methods

Plasmids

pLV hU6-sgRNA hUbC-dCas9-KRAB-T2a-Puro (dCas9-KRAB, simplified) was a gift from Charles Gersbach (Addgene plasmid # 71236). It codifies for a humanized form of inactive Cas9 (dCas9) C-terminally fused in frame with KRAB; a sgRNA scaffold-region in which the desired sgRNA sequence can be cloned under the control of the U6 promoter; and a puromycin resistance cassette, all inserted into a lentiviral backbone.

lentiCRISPR v2 was a gift from Feng Zhang (Addgene plasmid # 52961). It codifies for a humanized Cas9, a sgRNA scaffold-region in which the desired sgRNA sequence can be cloned under the control of the U6 promoter and a puromycin resistance cassette, all inserted into a lentiviral backbone.

Plasmid amplification

Competent *Escherichia coli* STBL3 (*E. coli*) were transformed with the Cas9-KRAB or lentiCRISPR v2 plasmids through heat shock. Cells were grown at 37 ^{o}C in lysogeny broth (LB; 10% (w/v) tryptone, 5% (w/v) yeast extract) for 1h, and then plated in dishes containing LB-agar with ampicillin. Cells were incubated at 37 ^{o}C overnight and resulting colonies were picked and grown in LB at 37 ^{o}C with rotation overnight. Plasmidic DNA was purified using a NZYMaxiprep kit (NZYTech, Portugal), following the manufacturer's recommendations. DNA concentration and purity were assessed using Nanodrop 2000 (Thermo Scientific, USA).

Single guided RNAs (sgRNA) design

Bioinformatic platforms were used to scan the DNA sequence of the human *ATXN3* gene (Accession code ENSG00000066427; European Bioinformatics Institute – EMBL-EBI). PAMs were detected and a score based on predicted off-targets was attributed to each possible sgRNAs (**Table 2**). Candidates for further testing were selected based on attribution of high scores across the largest number of platforms.

Table 2 - Online platforms used to design sgRNAs that target the human *ATXN3* gene.

PLATFORM	URL
GPP WEB PORTAL	http://www.broadinstitute.org/rnai/public/analysistools/ssgRNA-design
CRISPR DIRECT	https://crispr.dbcls.jp
CRISPR DESIGN	http://crispr.mit.edu/
SSGRNA SCORER 1.0	https://crispr.med.harvard.edu/
CRISPRSCAN	http://www.crisprscan.org/
ZIFIT	http://zifit.partners.org/ZiFiT/Introduction.aspx
E-CRISP	http://www.e-crisp.org/E-CRISP/

In the case of the sgRNA targeting the $G^{987}GG \rightarrow C^{987}GG$ SNP, we scanned the SNP region downstream the CAGs repeat in the human *ATXN3* gene.

For sgRNAs aiming at editing the CAG repeat region, we scanned a region of the human *ATXN3* gene upstream and downstream the CAGs repeat, that includes intron 9-10, exon 10 and intron 10-11 (containing 160 nucleotides upstream of the start of the exon and 123 nucleotides downstream its end). Four sequences were selected taking into account the corresponding scores and their proximity to the CAG tract – one on intron 9-10, upstream of the repeat region, and three downstream, on exon 10. A sgRNA sequence targeting *Escherichia coli* LacZ that has been validated as having no targets in human cells was used as a cut-inefficient control (Platt et al., 2014).

sgRNA cloning

sgRNAs were cloned into the expressing vectors, dCas9-KRAB or lentiCRISPR v2, in the form of annealed oligonucleotide pairs. Each strand (template and codifying) of the sgRNA was complementary to each other, with the exception of an overhang region of four nucleotides at the 5' end of each strand, which allow the cloning into a BsmBI restriction site of the vectors under the control of the hU6 promoter. In case sgRNA sequences lacked it, an additional guanine nucleotide was added to the 5' end of the codifying strand of the sgRNAs, between the overhang and the 20 nucleotides that target the human *ATXN3* gene. This procedure was required for the polymerase IIIdependent U6 promoter to function. The sgRNA oligonucleotides were obtained from *Invitrogen* (Thermo fisher).

Oligonucleotide pairs were phosphorylated and annealed using T4 polynucleotide kinase (New England Biolabs), according to the manufacturer's recommendations.

Cas9-KRAB or lentiCRISPR v2 expressing plasmids were separately digested with BmsBI (Esp3I; Thermo Fisher Scientific) and the resulting product were separated by electrophoresis and then purified with the NucleoSpin® Gel and PCR Clean-up kit (Macherey-Nagel, Germany). The annealed sgRNAs were finally cloned into the open (previously digested) expressing plasmids, through ligation using T7 ligase (New England Biolabs), following the manufacturer's recommendations. The resulting products were transformed into Stbl3 cells and grown in LB containing ampicillin. Surviving cell colonies were grown in LB with ampicillin and the plasmidic DNA of each culture was purified with the NZYTech MiniPrep kit. Restriction analysis and Sanger sequencing (U6 forward sequencing primer; GATC biotech) confirmed a successful cloning.

Design and generation of donor templates

Geneious software was used to design the donor templates. A non-expanded template containing 14 CAG repeats, was synthetized by integrated DNA Technologies (IDT; USA). The plasmid contains a portion of the human *ATNX3* gene that includes the Pre-10 and Pos-10_1 cutting sites (closest to the CAG repeats) and two homology arms, one upstream and another downstream of the double-stranded break sites, of 800 bp length. 14 CAGs were chosen because that is the number of CAGs contained in the reference human *ATXN3* gene (NCBI; ENSG00000066427). To prevent double-stranded breaks in the donor plasmid, one guanine of the PAM that is targeted by the Pre-10 sgRNA was deleted from the template and seven nucleotides of the Pos-10_1 target sequence were silently mutated. This plasmid is henceforth designated 14CAG-plasmid

A selection cassette containing a copepod GFP (CopGFP) gene, a puromycin resistance gene (PuroR) and a gene codifying herpes simplex virus thymidine kinase (HSV TK) was inserted into intron 9-10 of the template, 119 nucleotides upstream the start of exon 10. This would allow selection of cells that underwent HDR. The selection cassette was flaked upstream and downstream by a TTAA motif that would allow its seamless excision upon incubation with an excision-only *piggyBac* transposase (5’ homology harm-TTAA-selection cassette-TTAA-3’ homology harm).

Insertion of the resistance cassette was performed by dividing and cloning the 14CAG repair template into a vector containing the selection cassette. Two different amplicons were produced (5’ homology arm amplicon and 3’ homology arm amplicon plus 14CAGs; **Table 3**) by PCR and purified with the NucleoSpin® Gel and PCR Clean-up kit (Macherey-Nagel, Germany), after electrophoretic separation. The primers used for amplification (**Table 3**) introduced the necessary restriction sites for cloning into the PrecisionX$^{\text{TM}}$ HR Targeting Vector PBHR100_20160411 (System Biosciences), that included the selection cassette. Restriction analysis and Sanger sequencing confirmed the successful joining of the template to the selection cassette (primer used for the 5’ fragment: 5’- TTT CCT AAG ATC AGC ACT TCC A -3’; primer used for the 3’ fragment: 5’- ACT GCT CCT TAA TCC AGG GAA -3’). The resulting plasmid was named Donor-14Q-Puro.

Table 3 - Primers used for amplification of the 5’ fragment and 3’ fragment

Fragment	Primer name	Sequence	Restriction site
5’ Fragment	5’_HA_top	5’- GAT CTA **GCG GCC GC** GCA CTT TCA TTA GCT TAC AT GC -3’	NotI
	5’_HA_bottom	5’- GAA GAA **GGT CTC** G*TT AA*A AGA ATG CAA GAG CAG TTA GTC -3’	BsaI
3’ Fragment	3’_HA_top	5’- CGC CGG **GCT CTT C***TT AA*T AAC AGT GAC TAC TTT GAT TCG -3’	LguI (BspQI)
	3’_HA_bottom	5’- GTC GTC G**AA TTC** CTG CAC TTC CCA TTG TTA TTT AG -3’	EcoRI

Underlined: represents extra nucleotides needed for proper enzymatic digestion; **Bold:** represents nucleotide inserting restriction site at the end of the amplicon; *Italic*: represent the TTAA motif.

To overexpand the number of CAG repeats a donor template (henceforth named Donor-141) containing 141 CAG repeats was designed and generated following the same rational as the 14CAG-plasmid. The template includes the region between the cutting sites targeted by the Pre-10 and Pos-10_2 sequences and homology arms of 1000 bp (this template does not have the selection cassette). To prevent double-stranded breaks in the Donor-141, the complete PAM that was targeted by the Pre-10 sgRNA was deleted and eight nucleotides targeted by the Pos-10_2 sgRNA were silently mutated. The Donor-141 was synthesized by GenScript (USA).

Cell culture and transfection

Human embryonic kidney (HEK) 293T cells were cultured in Dulbecco's Modified Eagle Medium (DMEM), pH 7.2, supplemented with 10% (v/v) fetal bovine serum (FBS), 10 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), 44 mM of $NaCO_3$) and 1% (v/v) penicillin-streptomycin. Cells were maintained in 75 cm^2 flasks in an incubator at 37 °C, with a humidified atmosphere containing 5% CO_2. Every 2/3 days, media was changed as cells were split upon reaching an average confluence of 80-90%.

Transfection with the *Lipofectamine® 3000* reagent (Thermo Fisher Scientific, USA) was carried out according to the manufacturer's indications. The day before transfection cells were diluted 1:6 and seeded onto 100 mm culture plates. Seeded cells were transfected with 9.6 µg of DNA and 36 µl of transfection reagent. 4 h later transfection medium was substituted by fresh culture medium and cells were maintained for 22-44 days. Every 2-3 days, cells were treated with puromycin 0.1 µg/ml until 19 days and 0.5 µg/ml for the remaining 25 days, before DNA or protein purification.

In the CAG retraction experiences cells were also transfected with *Lipofectamine® 3000* reagent (Thermo Fisher Scientific, USA), according to the manufacturer's indications. Cells were seeded onto 6 well culture plates diluted 1:6. Seeded cells were transfected with 1.6 µg of DNA and 12 µl of transfection reagent. After 4 h incubation transfection medium was substituted by fresh culture medium and the next day cells were treated with puromycin 0.5 µg/µl and maintained for more 48h before DNA and protein purification.

Preparation of cell lysates

Cultured cells were washed with cold phosphate buffer saline (PBS; 137 mM NaCl, 2.7 KCl, 1.8 mM K_2PO_4, 10 mM $Na_2HPO_4 \cdot H_2O$, pH 7.4) and lysates were prepared by adding radioimmunoprecipitation (RIPA) buffer supplemented with protease inhibitors (10 µg/ml chymostatin, pepstatin, antipain and leupeptin (CLAP; Roche, Germany), 1 mM dithiothreitol (DTT; Fisher Bioreagents, USA) and 0.1 mM phenylmethanesulphonyl fluoride (PMSF; Sigma-Aldrich, USA); and phosphatase inhibitors (5 mM NaF, 2 mM Na_3VO_4 (Sigma-Aldrich, USA) and 1 µM okadaic acid (OA; Sigma-Aldrich, USA). The volume of buffer added was 100 µl/welll for 6-well plates and 600 µl/plate for 100 mm dishes. Cells were scraped vigorously on ice and then sonicated with 20 pulses (20 A), each 1 second long. Samples were centrifuged at 13,000 rpm for 10 min, at 4 ºC and the supernatant recovered. Protein concentration was determined by the bicinchoninic acid (BCA) method using the Pierce® BCA Protein assay kit as recommended by the manufacturer (Thermo Scientific, UK). Lysates were kept frozen at -20 °C.

Antibodies

Antibodies used in Western blot procedures are presented on Table 4

Table 4 - Antibodies used in Western blot analysis

	Antibody name	Species	Dilution	Incubation	Manufacturer
Primary antibody	anti-atxn-3 (1H9 clone)	Mouse mono	1:1,000	ON, 4 ^{O}C	MERC/Millipore (USA)
	anti-polyQ (5TF1-1C2 clone)	Mouse mono	1:1,000	ON, 4 ^{O}C	MERC/Millipore (USA)
	anti-HA tag	Rabbit poly	1:1,000	ON, 4 ^{O}C	Thermo Scientific (USA)
Secondary antibody	anti-mouse IgG APL-C	Rabbit mono	1:10,000	2h, RT	Thermo Scientific (USA)
	anti-rabbit IgG APL-C	Goat poly	1:10,000	2h, RT	Thermo Scientific (USA)

Abbreviations: **atxn-3**, ataxina 3; **PolyQ**, polyglutamine expansions; **APL-C**, alkaline phosphatase-conjugated; **IgG**, Immunoglobulin G; **mono**, monoclonal; **Poly**, polyclonal; **ON**, overnight; **RT**, room temperature.

HEK cells genotyping

Genomic DNA from HEK 293T cell cultures was extracted and purified using the GeneJET Genomic DNA Purification Kit (Thermo Scientific), according to the manufacturer's recommendations. DNA concentration and purity were accessed using Nanodrop 2000 (Thermo Scientific, USA).

To assess integration of the Donor_141 or Donor-14Q-Puro repair templates, DNA samples were submitted to PCR using DreamTaq™ DNA Polymerase (Thermo Scientific) using the appropriate pair of primers, depending on the experiment that was being conducted (**Table 5**). PCR parameters were set according to the polymerase manufacturers. PCR-generated amplicons were submitted to an 1% (w/v) agarose gel electrophoresis during 45 minutes at 90 V. The resulting gels were scanned using the Gel Doc™ EZ system (Bio-Rad, Portugal).

When necessary, bands of interest were excised from the agarose gel using a clean scalpel and their DNA purified using a DNA purification kit (NucleoSpin Gel and PCR Clean up, Macherey-Nagel, Düren, Germany), according to the recommended protocol.

Table 5 - Primers used to genotype transfected HEK 293T cells

Experiment	Amplicon detected	Primer name
CAG overexpansion	~ 1956 bp or ~ 1567 bp	5’F
		6R
CAG retraction	~ 6174 bp or ~ 1567 bp	5’F
		6R
	~ 887 bp	CopGFP_F
		CopGFP_R
	~ 3199 bp	PB.BB
		PB
	~ 1967 bp	5’F_2
		PB

Note: primer sequence cannot be disclosed.

Lentiviral production

For each lentiviral production, a total of 20 100 mm cell plates was plated with 3,5x10^6 cells from the HEK 293T line each and incubated at 37 °C, with 5% CO2, overnight. Cells were transfected with the plasmid of interest and 3 other plasmids for lentivirus production and were left to incubate at 37°C, 3% CO_2, for 6 h. After that, cell medium was changed to allow for cell growth and to stop the transfection. Cells were incubated again at 37 °C, 5% CO2, overnight.

Cell medium containing the lentiviral particles was filtered with a 0.45 µm vacuum filterer (Stericup and Steritop Vacuum Driven Sterile Filterer, Merck Millipore, Darmstadt Germany), in order to eliminate possible cell residues. The filtered medium was ultra-centrifuged at 19,000 rpm, for 1 h 30 min, at 4 °C, in order to pellet the viruses. The pellet from each tube was resuspended in 0.5% bovine serum albumin (BSA)/PBS (Acros Organics, Thermo Fisher Scientific, Belgium) and incubated, in ice, for 1 h 30 min. Viruses were centrifuged again at 19,000 rpm, 4 °C, for 1 h 30 min and the resulting pellet resuspended in 0.5% BSA/PBS (Acros Organics, Thermo Fisher Scientific, Belgium). The pellet was left to fully resuspend, in ice, overnight. Finally, viruses were aliquoted (5 µL) into eppendorfs and frozen at -80 °C until use.

Lentiviral quantifications were performed using a RetroTek HIV-1 p24 Antigen Enzyme-Linked Immunoabsorbent Assay (ELISA) (ZeptoMetrix), according to manufacturer’s indications.

Animal experimentation

In vivo experiments were conducted using MJD mouse model with a C57Bl/6-background, generated by Torashima and collaborators and described in section *Animal models of Machado-Joseph disease* of the introduction (Torashima *et al.*, 2008). Mice express a truncated form of human atxn-3 containing 69 glutamines, fused with a Human influenza hemagglutinin (HA) epitope. Expression of atxn-3 is under the control of a Purkinje cell-specific L7 promoter.

Transgenic mice (Torashima et al., 2008) were held at the CNC animal facility. At approximately 3 weeks of age behavioural tests began and the day after the first test mice were subjected to stereotaxic surgeries. Food and water were dispensed ad libitum. All experiments were carried out according to the European Community Council directive (86/609/EEC) for the care and use of laboratory animals.

Lentiviral stereotaxic injections

Mice were injected at 2.3 mm rostral to lambda, 0.0 mm lateral to midline and 3 mm ventral from lambda, with mouth bar set at 0.0 mm, with a single 4.14 µL injection of lentivirus (LV) (corresponding to 400.000 ng of p24 antigen) encoding the dCas9-KRAB along with a sgRNA SNP-directed sequence (dCas9-KRAB/sgRNA_SNP). Alternatively, as control, mice were injected with lentivirus encoding the dCas9 and sgRNA empty (dCas9-KRAB/sgRNA_Empty), in which the variable 20 bp strand (crRNA) that is responsible for targeting Cas9 to specific DNA *loci* is lacking.

Behavioural tests

Transgenic mice were submitted to a battery of tests to access their motor performance, at 5 weeks of age. Animals were then stereotaxically-injected and tested again every three weeks until being sacrificed at eleven weeks after injection.

Motor coordination was accessed by the Rotarod test. Mice were placed in the Rotarod apparatus at a constant speed of 5 rpm. The time that mice remained in the apparatus was recorded. Four trials were performed, with 15-20 minutes interval between each trial. The mean of latency to fall of the 4 trials was used for statistical analysis.

The Beam Walking test was used to evaluate balance. Mice were placed on an elevated squared 50 cm long horizontal beam with 1.8 cm or 0.9 cm width. Mice performed 2 trials per beam type with approximately 15-20 minutes interval between trial. The means of the time taken by the animals to transverse the beams was used for statistically analysis.

Gait coordination was accessed by the Footprint test. Mice paws were coated with blue (hindfeet) or red (forefeet) non-toxic paint. Mice were placed over a white sheet of paper 50 cm long and 10 cm wide and allowed to transverse it. Four parameters were measured on the prints left by the paws: stride length (distance between each stride), hind-base width (distance between left and right hind footprints), front-base width (distance between right and left front footprints) and feet overlap (distance between front and hind footprint, measuring foot placement). Motor impaired mice tend to reduce the stride length and increase the overlap and both hind- and front-base distance. Measures were made for sequences of four consecutive steps (Nóbrega et al., 2013b; Torashima et al., 2008)

Swimming performance was assessed by placing mice at one end of a rectangular tank with 10 cm width, filled with water at room temperature. Mice freely swam a distance of 1.5 m until they reached a platform and the time taken to transverse the tank was recorded. Mice performed the trial three times, with an interval of 15-20 minutes per trial. The mean of the time taken to cross the tank in the tree trials was used for statistical analysis.

Brain tissue collection

Mice were euthanized by intraperitoneal injection of 80 mg/kg sodium pentobarbital and then transcardially perfused with 4% (w/v) paraformaldehyde in PBS, pH 7. Whole brains were removed and incubated for ~48 h in 20% (w/v) sucrose PBS. Tissue was then frozen at –80 °C.

Histological processing and cresyl violet staining

30 µm sagittal sections of the entire cerebellum were cut using a cryostat-microtome model CryoStar NX50 (Thermofisher, USA). Slices were stored at 4 °C, free-floating in 0.02% (w/v) sodium azide PBS inside 48-well plates, until further use.

Eight mid-sagittal sections from each mouse were mounted in gelatin-coated microscope slides and left drying overnight. Then, slices were stained with cresyl violet, passing them sequentially through water, ethanol 96% (v/v), ethanol 100% (v/v), xylene, ethanol 75% (v/v) and the 0.1% (w/v) cresyl violet solution. Slices were then washed in waster and sequentially passed through ethanol 75% (v/v), ethanol 96% (v/v), ethanol 100% (v/v) and xylene. Slides containing dyed slices were them mounted with Eukitt (Sigma-Aldrich, Portugal).

For histological analysis were used 12 mice (10 treated with dCas9-KRAB/sgRNA_SNP, and as control 10 treated with dCas9-KRAB/sgRNA_Empty).

Cerebellar RNA and protein extraction

Mouse cerebellar tissue was homogenised with TRIzol™ Reagent (Invitrogen, USA). Chloroform was added and the resulting aqueous phase, which was collected for RNA purification using NucleoSpin® RNA kit (Macherey-Nagel, Germany), according to the manufacturer's recommendations. RNA concentration and purity were assessed using Nanodrop 2000 (Thermo Scientific, USA).

Protein purification was performed by completely discarding any residues of the aqueous phase. Ethanol 100% was added to the remaining supernatant and then isopropanol, to precipitate the protein. Samples were then centrifuged at 13,000 rpm for 10 mim, the supernatant was discarded, and the pellet was collected, then 0.3 M guanidine-hydrochloride in ethanol 95% (v/v) was added to the pellet and this process was repeated 3 times. Ethanol 100% was added, and samples were left air-draying until no ethanol was observed. Finally, urea/DDT (10 M urea, 50 mM DTT in water) was added to the samples, and they were finally sonicated with 20 pulses (20 A), each 1 second long. Protein was quantified by the Bradford method using the Bio-Rad Protein Assay Dye Reagent Concentrate (Bio-Rad, Portugal). RNA samples were stored at -80 ^{o}C and protein samples were stored at -20 ^{o}C.

For cerebellar extracts were used 8 mice (4 treated with dCas9-KRAB/sgRNA_SNP, and as control 4 treated with dCas9-KRAB/sgRNA_Empty)

Western blot

Cell lysates were diluted in RIPA buffer, so that the total amount of protein separated by electrophoresis would be 30 µg (*in vitro* experiences) or 20 µg (*in vivo* experiences). 6x concentrated loading buffer (Tris-HCL 0.5M pH 6.8, glycerol 30% (v/v), 10% (w/v) SDS, 0.6M DTT and ~0.012% (w/v) bromophenol blue) was added and samples were heated for 5 minutes at 95 °C. Samples were loaded into 4% (w/v) bis-acrylamide stacking gels (125 mM Tris-HCl, pH 6.8 with 0.1% (w/v) SDS) and separated In 10% (w/v) running gels (375 mM Tris-HCl pH 8.8 with 0.2% (w/v) SDS) by electrophoresis until the desired separation of the NZYColour Protein Marker II was observed (NZYTech, Portugal). Voltage was initially at 50V until samples reached the running gel, and from then on at 90 V. Proteins were electrotransfered for 2 h 30 minutes at 4 °C (1000 mA) into polyvinylidene fluoride (PVDF) membranes.

Membranes were blocked with Tris-buffered saline (20 mM Tris and 137 mM NaCl, pH 7.6) containing 0.1% (v/v) Tween 20 (TBS-T) and 5% (w/v) low-fat dry milk, for 2 h with agitation. Membranes were then incubated with the primary antibodies diluted in 0.5% (w/v) low-fat dry milk TBS-T, overnight at 4 °C. Membranes were washed 3 times with 0.5% (w/v) low fat-dry milk TBS-T, during 10 minutes for each wash. Then, membranes were incubated with the adequate secondary antibody diluted in the same solution as the primary antibody. After 1 h incubation, membranes were again washed (3 times, 10 minute each), and finally resolved using Enhanced Chemifluorescence substrate (ECF; GE Healthcare, Portugal) and scanned with a ChemiDoc™ MP Imaging System (Bio-Rad, Portugal) or VersaDoc Imaging System Model 3000 (Bio-Rad, Portugal).

Antibody reprobing was performed by incubating membranes with 40% (v/v) methanol for 45 minutes and then with 0.2 mM NaOH for 15 minutes. Membranes then were blocked again and labelled as previously described. Densitometric analysis of the scanned Western blots was performed using ImageJ v1.52p software (National Institute of Health, USA).

qPCR

cDNA was obtained by reverse transcription of the purified RNA using the iScript cDNA synthesis kit (Bio-Rad, Portugal), according to the manufacturer's recommendations. qPCR was performed using the SsoAdvanced SYBR Green Supermix kit (Bio-Rad, Portugal), according to the manufacturer's recommendation. cDNA amplification was then carried out using a home-made pair of primers that target the human *ATXN3* RNA present in the transgenic mice. As control, a pair of primers that targeted the transgenic mice *Gapdh gene* were used. Primers sequence used to target *ATXN3* and *Gapdh* cannot be disclosed, patent is under approval. qPCR was performed in the StepOnePlus Real-Time PCR System (Applied Biosystems) using 96-well plate containing 10 µL reaction volume.

Microscopy

Microscopy images of cresyl violet-dyed brain slices were obtained with a Axio Imager.Z2 transmitted-light bright field microscope (Carl Zeiss, Germany), illuminated by a halogen lamp at ~ 2.5-3.5 V and using a Plan-Apochromat 20x/0.8 M27 objective

(20x; Carl Zeiss, Germany). Images were collected with a colour digital camera Axiocam HRc (Carl Zeiss, Germany). Approximately 20 tiles were taken per cerebellar section.

Images were then processed, and cerebellar layer thickness measured using ZEN 2.6 2018 (blue edition) software (Carl Zeiss, Germany). For each mouse, thickness of the molecular, granular and Purkinje cell layers of three sections of lobes IX and X were quantified (two measurements per layer). Means of each different layer thickness from each lobe were used for statistical analysis.

Statistical analysis

Statistical analysis was performed using the GraphPad 6 software (GraphPad Software, USA). Comparison between two groups was carried out using unpaired Student's t-test. Multiple comparisons between groups were performed using one-way analysis of variance (ANOVA). Results are expressed as mean ± standard error of the mean (SEM) and significance was set at *$p<0.05$, **$p<0.01$ and ***$p<0.001$.

Chapter 3 – Results and discussion

3.1 Using CRISPR/Cas9 to overexpand and retract the *ATXN3* CAG repetition

3.1.1 Overexpanding CAG trinucleotide repeats in vitro

One of the bottlenecks of studying any disease is finding a suitable model capable of mimicking its pathological profile. However, there are no perfect disease model and each one of them has its pros and cons. The refinement of previously existing models, when possible, is something that should be kept in mind.

The yeast artificial chromosome (YAC) MJD-Q84.2 (described in section *Transgenic mouse models expressing full-length human ataxin-3*) is the MJD mouse model that better recapitulates the human condition. Moreover, it is the only one that contains a full-length mutant human ATXN3 gene containing all exons, introns and regulatory regions, mimicking the human genetic context. This model is of special relevance, for example to study splicing events or gene regulation by micro RNAs (miRNA) or antisense oligonucleotides (ASOs) that bind to intronic regions of the *ATXN3* mRNA. This mouse model contains 84 CAG repeats, but despite this CAG repeat number being enough to cause the disease to be manifested in humans, mice show a mild and slowly progressive cerebellar deficit, making this model difficult to use in research due to the time and resources required to maintain the animals until they start to develop the first symptoms (Carmona et al., 2017; Cemal et al., 2002; Gould, 2012; Moore et al., 2017).

Due to the potential of the YAC MJD-Q84.2 mouse model to be used in research because of its genetic and phenotypic characteristics, we intended to establish a method for refining this mouse model in order to better suit research objectives. Therefore, taking into consideration previous studies showing that the number of CAGs inversely correlates with the time of MJD onset and directly correlates with symptom severity, we decided to develop an *in vitro* strategy to overexpand the number of CAGs of the YAC MJD-Q84.2 mouse from 84 to 141 CAGs repeats, expecting earlier onset and a more severe phenotype (Abe et al., 1998; Jardim et al., 2001). We choose 141 CAGs repeats, taking into consideration that a) the number of glutamines had to be larger than 84 and that b) the MJD mouse model expressing full-length atxn-3 with the largest expansion displays 148 CAG repeats (Boy et al., 2010).

The new mouse model would be expected to have exactly the same characteristics of the YAC MJD-Q84.2, with the exception that it would be less time-consuming to use and a less expensive to maintain.

In order to overexpand the CAG repeat region of the *ATXN3* gene, we employed a gene editing-based approach using the CRISPR/Cas9 technology. As described before, this technique consists on the insertion of a homologous DNA sequence (donor template), in this case containing 141 CAG repeats, through homology-directed repair (HDR) upon a double-stranded break induced by the Cas9 nuclease at the vicinity of the region aimed to be edited. We designed a pair of sgRNA sequences that target regions upstream (Pre-10) and downstream (Pos-10_2) of the CAG repeats at the exon 10 of the human *ATXN3* gene. Each sgRNA is codified in a plasmid containing the Cas9-codifying sequence along with a sequence that confers puromycin resistance to the transfected cells, therefore enabling selection of cells that were successfully transfected. The donor template consists on a plasmid containing a fragment of the human *ATXN3* gene with 141 CAG repeats, flanked by a 5’ homology arm and a 3’ homology arm that are homologous to exon 10 of the *ATXN3* gene or the regions at its vicinity (**Figure 8**).

HEK 293T cells were used as an *in vitro* system to test the editing strategy, since they constitute a human cell line that is easy to maintain and transfect. To induce HDR in order to overexpand the CAGs repeats of HEK cells endogenous *ATXN3* gene, cells were co-transfected with the donor template (Donor-141) that provides the 141 CAGs repeats, along with the plasmids encoding Cas9 and the two sgRNAs (Pre-10 and Pos-10_2) responsible for the double-stranded breaks. 24 h after transfection cultures were treated with puromycin (10 µg/ml) in order to select the transfected cells. Cultures were then maintained for an additional 48 h before being collected and processed for genotyping.

To perform a PCR-screening analysis we designed a pair of primers that would lead the amplification of a region of the *ATXN3* gene containing the intended editing site, indicating if the insertion of the 141 CAGs was successful: a forward primer (named F5’) that targeted intron 9-10 the *ATXN3* gene 1172 bp upstream of sgRNA Pre-10 target site, and a reverse primer (named R6) that targets exon 10 197 bp downstream of the Pos-10_2 target site. Since the expanded glutamines would increase the length of the resulting amplicon, based on the size of the DNA fragment generated by the two primers we would be able to determine if Donor-141 was inserted at the intended site. We expected DNA fragments of a 1567 bp in non-edited conditions and a fragment of 1956 bp for cells correctly edited (**Figure 8**).

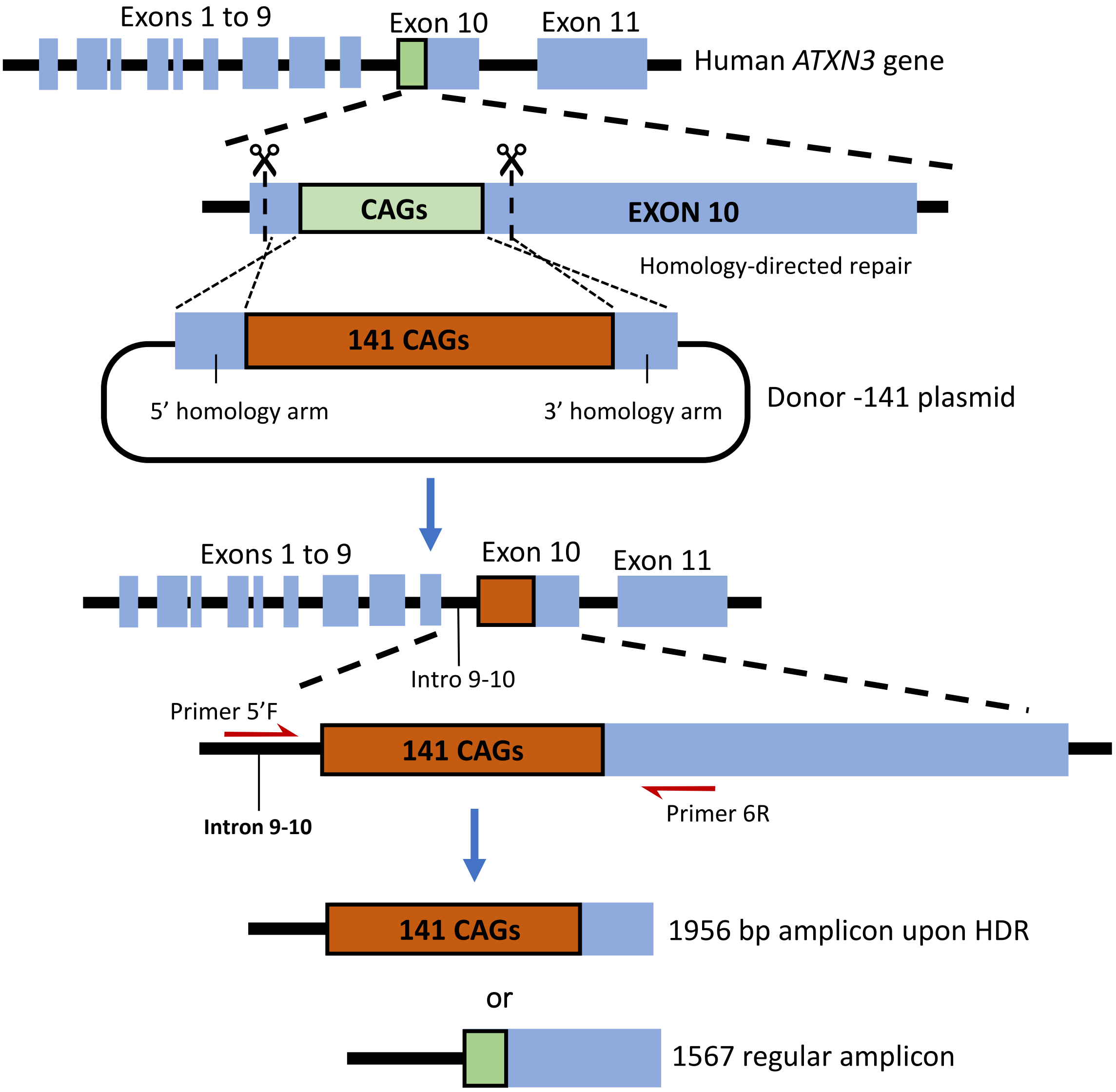
Exons 1 to 9
Exon 10
Exon 11
Human *ATXN3* gene
CAGs
EXON 10
Homology-directed repair
141 CAGs
5' homology arm
3' homology arm
Donor -141 plasmid
Exons 1 to 9
Exon 10
Exon 11
Intro 9-10
Primer 5'F
141 CAGs
Intron 9-10
Primer 6R
141 CAGs
1956 bp amplicon upon HDR
or
1567 regular amplicon

Figure 8 - Schematic representation of the Donor-141 template containing 141 CAG repeats being integrated into the human *ATXN3* gene and PCR screening analysis using the primers 5'F/6R to detect an amplicon of 1956 bp distinctive of integration. Upon the double-stranded breaks mediated by CRISPR/Cas9 in the vicinities of the CAG repeats present in the exon 10 of the human *ATXN3* gene, Donor-141 is integrated via homology-directed repair. Through a PCR analysis on agarose gel electrophoresis, using the primers 5'F/6R, we expected an amplicon of 1957 bp indicating that integration of the 141 CAGs was successful. Otherwise, the fragment generated would be 1567 bp long.

PCR-based genotyping of the transfected cells genomic DNA yielded only a ~1500 bp fragment, compatible with the size of the amplicon predicted to be produced from HEK cells *ATXN3* gene. No fragment that would indicate correct insertion of the donor template and subsequent increase in CAG repeat number was detected, suggesting that HDR did not occur or occurred only in a very limited population of cells (**Figure 9**).

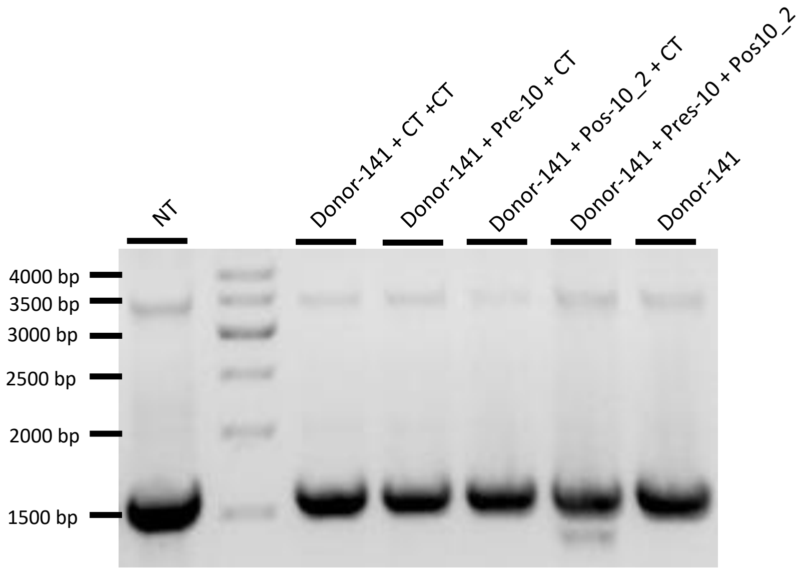

Figure 9 - PCR screening of cells co-transfected with the Donor_141, Cas9 and sgRNAs Pre-10, Ps10_2 or both. Purified DNA from HEK 293T co-transfected with a plasmid encoding for the Donor_141, a plasmid encoding for Cas9-Pre-10 and a plasmid encoding Cas9-Pos-10_2, after 72 h of expression. Exon 10-derived PCR products obtained from the purified DNA were analysed on agarose gel electrophoresis, showing a fragment of a ~ 1500 bp for all tested samples. Absence of a 1956 bp fragment suggests that insertion of the Donor-141 containing 141 CAG repeats did not occur. **CT** – control sgRNA that targets a sequence of DNA that is not present in the human genome (see section *Material and Methods*), example: Donor-141 + Pre-10 + CT means that cells were transfected with plasmid encoding for the Donor-141, a plasmid encoding the Cas9-Pre-10 and a plasmid encoding for the Cas9-sgRNA (CT) that does not target the human genome); **NT** – non-transfected cells.

The aim of this work was to establish a molecular strategy to refine the previously existing YAC MJD-Q84.2, by overexpand the CAG via gene editing. The new model would be expected to express an atxn-3 protein with a larger polyglutamine tract with 141 glutamines, aggravating the MJD phenotype. If HDR was successful in our *in vitro*, HEK cell-based platform, cells transfected with Donor-141, Cas9 and the sgRNA (Pre-10 and Pos-10_2) are expected to express a heavier protein, of approximately 66 kDa. Western blot analysis using the anti-Spinocerebellar Ataxia Type 3 Antibody (1H9) to label endogenous human atxn-3 showed no alterations in all tested samples, regarding endogenous protein apparent molecular weight (**Figure 10A**). To detect if transfected cell were expressing a protein containing an expanded polyglutamine tract that could correspond to overexpanded atxn-3, we used Anti-Polyglutamine-Expansion antibody (1C2), that detect polyglutamine-containing protein.

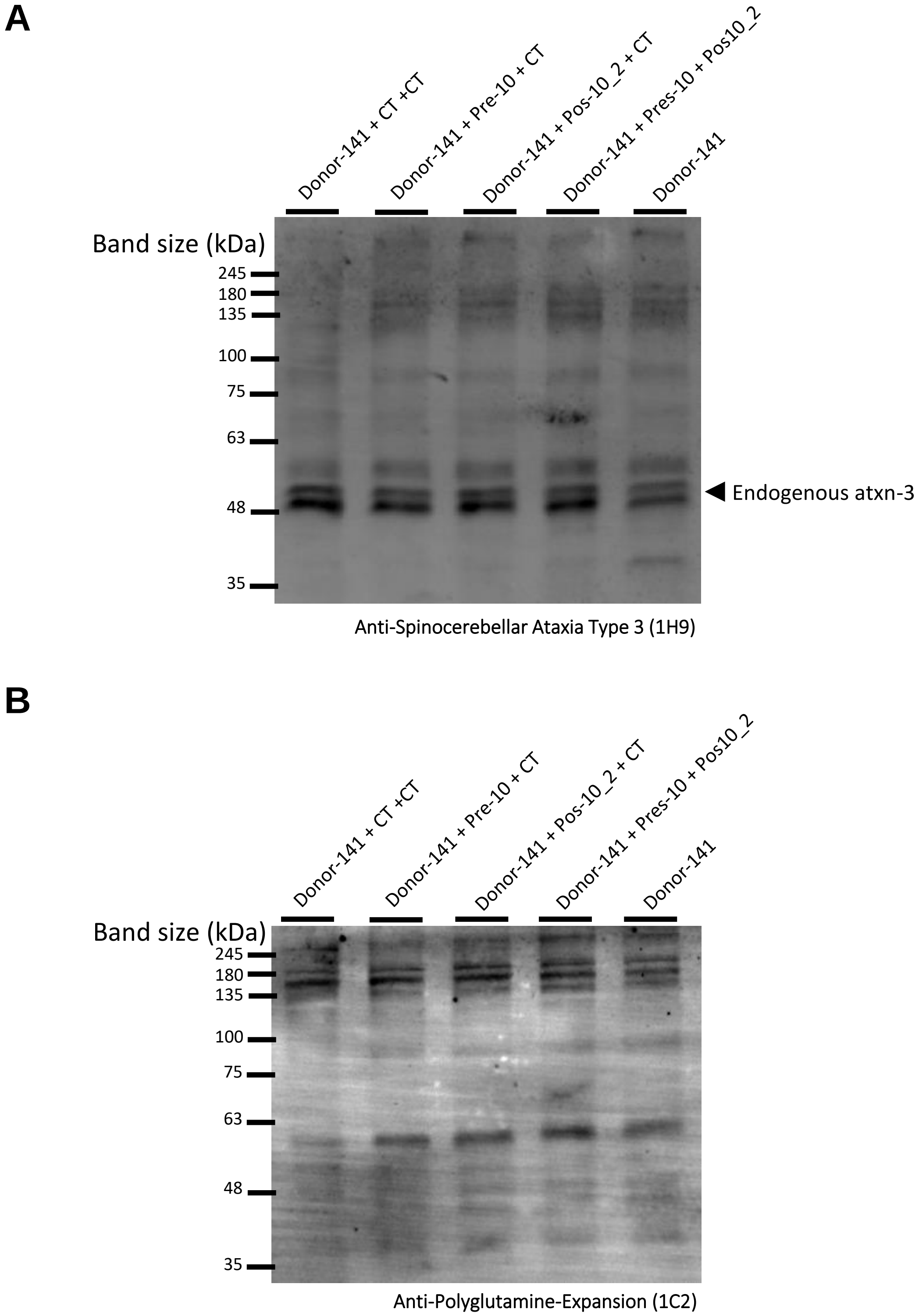
A
Donor-141 + CT +CT
Donor-141 + Pre-10 + CT
Donor-141 + Pos-10_2 + CT
Donor-141 + Pres-10 + Pos10_2
Donor-141
Band size (kDa)
245
180
135
100
75
63
48
35
Endogenous atxn-3
Anti-Spinocerebellar Ataxia Type 3 (1H9)
B
Donor-141 + CT +CT
Donor-141 + Pre-10 + CT
Donor-141 + Pos-10_2 + CT
Donor-141 + Pres-10 + Pos10_2
Donor-141
Band size (kDa)
245
180
135
100
75
63
48
35
Anti-Polyglutamine-Expansion (1C2)

Figure 10 - Detection of atxn-3 protein from cells co-transfected with the Donor_141, Cas9 and both Pre-10 and Pos10_2 sgRNAs. Protein lysates from HEK 293Tcells co-transfected with a plasmid encoding for the Donor_141, a plasmid encoding for Cas9-Pre-10 and a plasmid encoding Cas9-Pos-10_2 were prepared after 72 h of expression. **(A)** Anti-Spinocerebellar Ataxia Type 3 (1H9) antibody staining revealed a band of between of ~ 48 kDa (arrowhead), the molecular size of endogenous HEK cell atxn-3, in all tested samples, suggesting that the atxn-3 polyglutamine tract did not increase. **(B)** A band of ~ 66 kDa, corresponding to the atxn-3 containing the overexpanded polyglutamine tract, was not detected in any sample tested using the Anti-Polyglutamine-Expansion (1C2) antibody. None of the bands seem to correspond to atxn-3. **CT** – control sgRNA that targets a sequence of DNA that is not present in the human genome, thus acting as a control sgRNA.

No band of ~ 66 kDa weight, that could correspond to overexpanded atxn-3, was detected, suggesting that transfected cells were not expressing the expected modified atxn-3 protein (**Figure 10B**).

Taking all together, genotyping results did not reveal the presence of *ATXN3*-derived species that could result from an overexpanded form of the gene. Endogenous *ATXN3*-derived DNA amplicons and protein samples maintain the expected length and molecular weight, respectively, suggesting that the number of CAG repeats was not increased. Results thus seem to indicate that Donor-141 was not integrated into the exon 10 of the human *ATXN3* gene (PCR results, **Figure 9**) and that HDR was not induced after the double-stranded breaks mediated by CRISPR/Cas9.

3.1.2 Correcting the **ATXN3** *gene* **in vitro**

Cells collected directly from patients to study cellular and molecular mechanisms of certain diseases is a tremendous advantage, comparing to the use of cell lines such as HeLa or HEK. However, due to the lack of accessibility of some tissues and many health-related constraints, it is impossible to collect certain cells samples from patients, which is the case of MJD (Bettencourt and Lima, 2011; Nakano et al., 1972). Patient-derived induced pluripotent steam cells (iPSC) are powerful tools for *in vitro* disease modelling, representing an (virtually) infinite source of cells that can be later reprogrammed into any cell type (for example neurons) to study cellular and molecular mechanism of disease. However, iPSC-derived cells from different patients and from healthy individuals have different genetic background, affecting the differentiation potential and even the disease phenotype (Kajiwara et al., 2012; X. Xu et al., 2017). Therefore, a methodology through which iPSCs with the same genetic background could be produced would be advantageous – generation of control isogenic patient-derived iPSCs lines.

Isogenic patient-derived iPSCs lines can admittedly be created by correcting the genetic mutation that causes the disease. As a gene editing tool, CRISPR/Cas9 may be used to repair a pathogenic gene mutation.

In order to develop an *in vitro* strategy to generate an isogenic patient-derived iPSCs of MJD patients containing the corrected *ATXN3* gene, we employed the CRISPR/Cas9 technology to induce double-stranded breaks and, as a consequence, HDR of the disease-related site. We used the same methodology and the same paradigm described before for the integration of 141 CAGs – using HEK cells, we would induce a double-stranded break upstream and another downstream of the CAGs of the human *ATXN3* gene and correction would be mediated by a provided template containing the desired alteration.

Since we were not able of detect the integration of the Donor-141 in the previously set of experiments, we decided on a different approach, which entailed a method for selecting cells that underwent HDR. The template provided for gene correction (Donor-14Q-Puro) was composed by a region of *ATXN3* containing exon 10 with 14Q, to which selection cassette had been inserted (**Figure 11**). The selection cassette was composed by a copepod GFP (CopGFP) gene, a puromycin resistance gene (PuroR; for positive selection) and a Herpes simplex virus thymidine kinase gene (HSV-TK; for negative selection). The selection cassette was flanked upstream and downstream by a TTAA motif, which is recognized by a *piggyBac* transposase allowing the seamless excision of the selection cassette (Xiaohong Xu et al., 2017; Yusa, 2013). Upon double-strand break induction, we expect the selection cassette to integrate in the intron 9-10. At the same time, we also expect the 14 CAG-bearing tract to substitute its corresponding region in exon 10 of the ATXN3 gene. The reasons to choose a total of 14 CAG repeats are two-fold: first, 14 CAGs is the number of repetitions for the *ATXN3* gene present in the reference database (European Bioinformatics Institute – EMBL-EBI); and second, it is the most common repetition present in the human population (Gan et al., 2015)

Firstly, HEK 293T cells were co-transfected with the Donor-14Q-Puro and the previously mentioned constructs codifying for the Cas9 and a sgRNA – Pre-10 or Pos-10_1 (or with two plasmids, encoding both). After ~48 h of expression, cells were cultured under puromycin resistance-selective conditions (puromycin 0.1 µg/ml). Cells were maintained for 22-44 days and their growth and morphology were monitored. We expected that cells where integration of the selection cassette did not occur would not survive after three weeks (Mok and Lever, 2006; Xiaohong Xu et al., 2017; Zhang et al., 2017). However, even in control situations in which no double-stranded break was induced, i.e. in the conditions where cell were transfected only with the Donor-14Q-Puro; or with the Donor-14Q-Puro and the control sgRNA, cells remained alive and cultures were visually similar to those in which Cas9-induced double-stranded breaks were expected to occur.

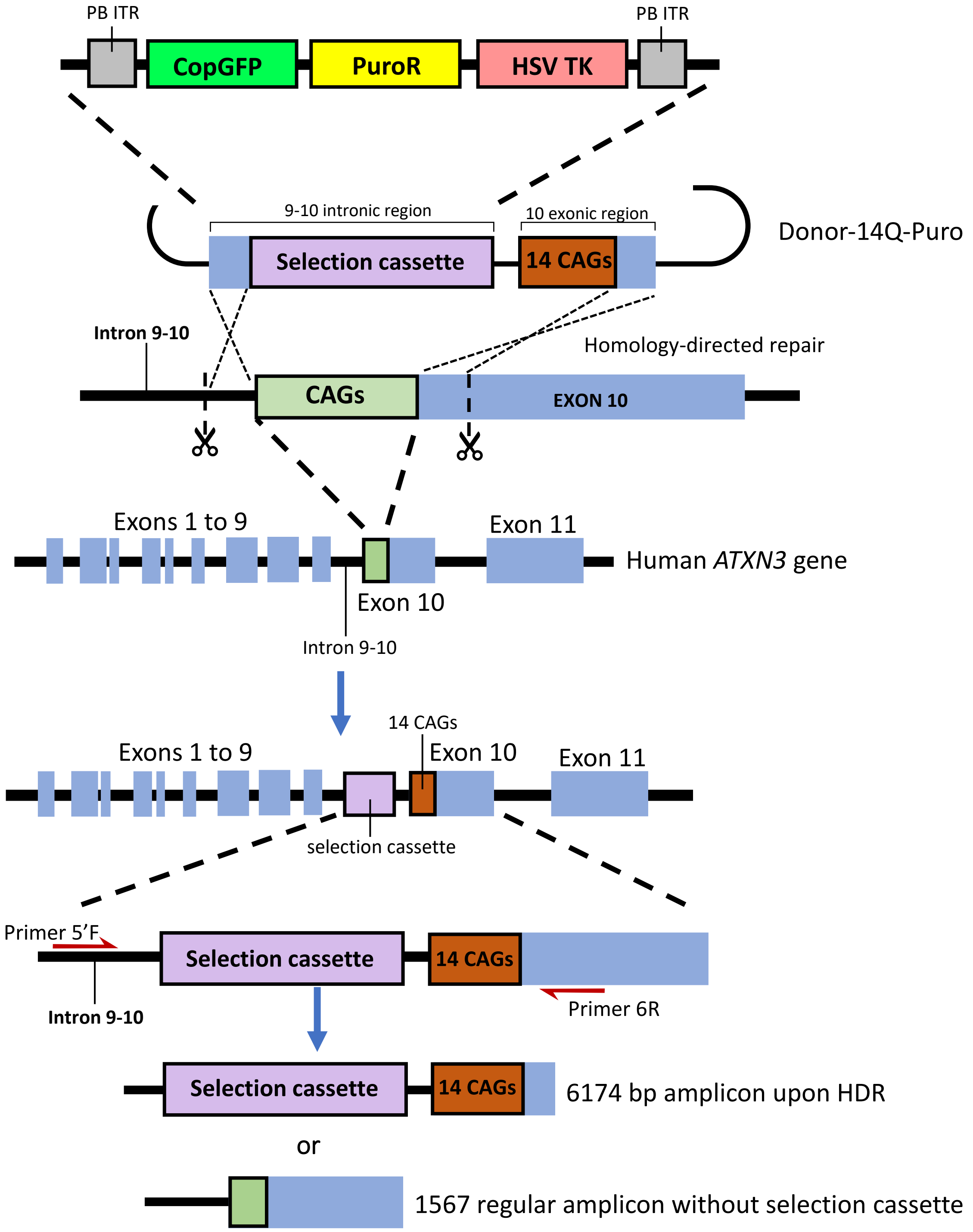
PB ITR
CopGFP
PuroR
HSV TK
PB ITR
9-10 intronic region
10 exonic region
Selection cassette
14 CAGs
Donor-14Q-Puro
Intron 9-10
Homology-directed repair
CAGs
EXON 10
Exons 1 to 9
Exon 11
Human ATXN3 gene
Exon 10
Intron 9-10
14 CAGs
Exons 1 to 9
Exon 10
Exon 11
selection cassette
Primer 5'F
Selection cassette
14 CAGs
Intron 9-10
Primer 6R
Selection cassette
14 CAGs
6174 bp amplicon upon HDR
or
1567 regular amplicon without selection cassette

Figure 11 - Schematic representation of the Donor-14Q-Puro repair template, containing a selection cassette and 14 CAGs, being integrated into the human *ATXN3* gene. Upon the double-stranded breaks mediated by the CRISPR/Cas9 system at the vicinity of the CAG region present in the exon 10 of the human *ATXN3* gene, the Donor-14Q-Puro is integrated via homology-directed repair. Through a PCR analysis on agarose gel electrophoresis we expected an amplicon of 6174 bp to be generated, indicating that integration of the *piggyBac*-flanked selection cassette was successful; otherwise, the amplicon generated would have 1567 bp. The selection cassette would enable positive and negative selection of edited cells. The positive selection would be conferred by the puromycin resistance gene if the cassette was successfully integrated; the negative selection would be conferred by the Herpes simplex virus thymidine kinase (HSV-TK) if the future excision of the cassette was successfully completed, upon piggyBac transposase treatment

These results indicate that cells were able to retain or acquire puromycin resistance presumably without the double-stranded break occurring. Consequently, resistance was possibly acquired in the absence of the HDR phenomenon that would lead to the integration of the desired DNA fragment at the target *ATXN3* site that the procedure aimed at producing. Resistance developed as a result of transfection, since non-transfected cell cultures did not survive after 24 h of puromycin treatment.

After 22 days of puromycin selection, cells were collected and genomic DNA was purified for genotyping by a PCR-screening analysis aiming at detecting the presence of the selection cassette, which could be indicative of insertion of the repair template. We used a pair of primers that would amplify a fragment of the selection cassette, containing part of the CopGFP gene and part of the PuroR gene present inside the *piggyback*-flanked selection cassette (**Figure 12**). The amplicon generated was expected to have 877 bp and to be detectable in every sample in which the selection cassette would have been inserted.

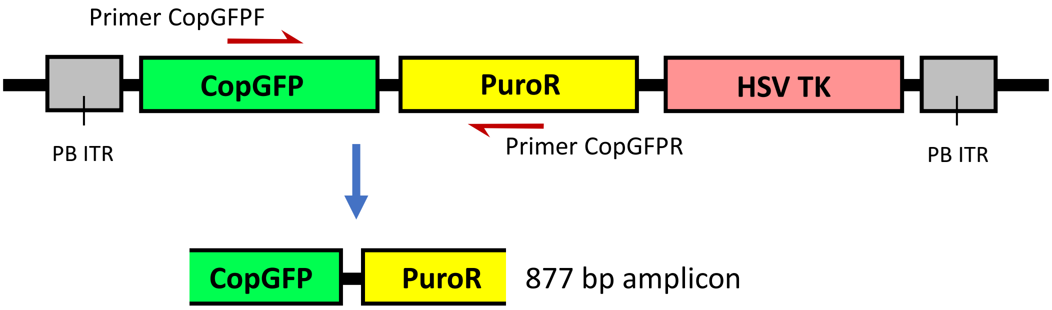

Figure 12 - Schematic representation of the PCR genotyping strategy for detection of selection cassette insertion. Amplification using primers CopGFPF and CopGFPR, which target CopGFP and the puromycin resistance gene, respectively, was expected to give rise to an amplicon of 877 bp in cells in which the cassette was inserted.

An amplicon of the desired length was detected in all conditions tested, even those where the double-stranded break was not expected – in cells transfected only with the repair template, or with the repair template, Cas9 and the control sgRNA sequence. Although the results were puzzling, they are in accordance with the fact that cells in all the different conditions tested remained resistant to puromycin (**Figure 13**).

The production of an amplicon resulting from the CopGFP and the PuroR genes present in the selection cassette is in line with the observed resistance to puromycin that all conditions tested demonstrated even after twenty-two days of puromycin treatment. Results suggest that the puromycin resistance gene remains inside the nucleus of cells, at the end of that period. However, since this seems to be independent from Cas9-induced double-stranded break, a random integration of the selection cassette may have occurred. Another less likely hypothesis is the possibility that the Donor-14Q-Puro plasmid resists to the rounds of cell division through the culture period, and the resistance gene keeps on being episomally expressed.

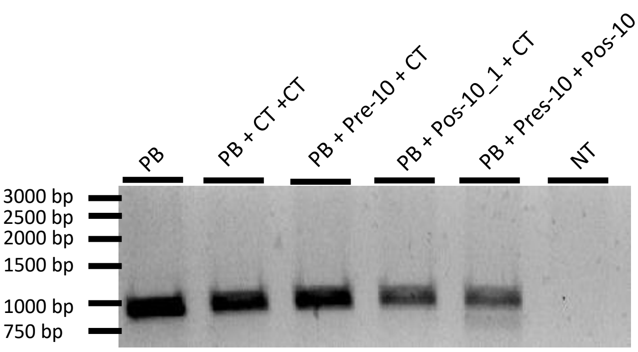

Figure 13 - PCR screening of the cells transfected with the Cas9 and the Donor-14Q-Puro, aiming at amplifying a fragment of the selection cassette of the repair template. Genomic DNA of HEK 293T cells transfected with the Donor-14Q-Puro and the Cas9 targeted at the vicinities the CAG repeats, was purified and amplified by PCR using the pair of primers CopGFPF and CopGFPR that target the Donor-14Q-Puro selection cassette. PCR products were analysed upon agarose gel electrophoresis, showing that, in all conditions tested, an amplicon of approximately 877 bp was present. The size of the amplicon was in accordance with the expected fragment of the CopGPF-PuroR present in the Donor-14Q-Puro selection cassette. Interestingly in the condition where cells were transfected with a control sgRNA (CT) that cannot target the human genome (PB+CT+CT), and thus no double-stranded breaks are expected to be induced by Cas9, cells also retained the puromycin resistance. Furthermore, in the condition where cells were only transfected with the Donor-14Q-Puro (PB), without Cas9 or sgRNA, the puromycin resistance was also retained. CT –sgRNA that targets a sequence of DNA that is not present in the human genome, thus acting as a control sgRNA; NT – non-transfected cells. PB – cells transfected only with Donor-14Q-Puro plasmid construct.

To test whether, regardless of a possible random insertion, the repair template was inserted at the intended *ATXN3* site, with which the template shares homology, another PCR screening was performed. The analysis used a forward primer that annealed at the intron 9-10 of the human *ATXN3*, 370 bp upstream of the region homologous to the 5' homology arm and a reverse primer that annealed at the region homologous to the 3' homology arm. Since the forward primer does not display any complementarity to the sequences present in the Donor-14Q-Puro we are safeguarded from amplifying the plasmid. Therefore, the length of the amplicon would inform if the selection cassette was successfully integrated, since an amplification product of 1567 bp would correspond to the regular *ATXN3*-derived fragment, whereas an amplicon of 6174 bp (1567 bp +

5607 bp of the selection cassette) would correspond to a product derived from efficiently edited cells (**Figure 11**).

We were nonetheless unable to amplify an amplicon of 6174 bp in any of the conditions tested (**Figure 14**). Alone, this result suggests that the Dono-14Q-Puro was not being integrated into exon 10 of the human HEK cell *ATXN3* gene, at the desired location.

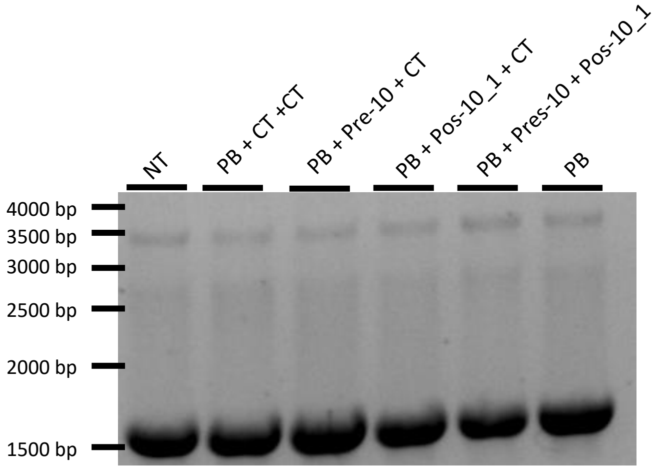

Figure 14 - PCR screening of the cells transfected with the Cas9 and the Donor-14Q-Puro, aiming at amplifying a fragment that indicated a correct integration. Genomic DNA of HEK 293T cells transfected with the Donor-14Q-Puro and the Cas9 targeted at the vicinities the CAG repeats, was purified and amplified by PCR using the pair of primers 5'F and 6R that target 370 bp upstream of the region homologous to the 5' homology arm and at the region homologous to the 3' homology arm. Agarose gel electrophoresis analysis of the PCR products showed an amplicon of approximately 1567 bp expected from non-edited cells. Despite the fact that, in all conditions, cells retained resistance to puromycin, none of the situations yielded an amplicon of 6174 bp, distinctive of a correct integration of the Donor-14Q-Puro construct at the exon 10 of the human *ATXN3* gene. **CT** – sgRNA that targets a sequence of DNA that is not present in the human genome, thus acting as a control sgRNA; **NT** – non-transfected cells. **PB** – cells transfected only with Donor-14Q-Puro plasmid construct.

However, the manufacturer of the polymerase used in the experiment warns for the difficulty of amplifying DNA products above 6000 bp, leading us to the hypothesise that, possibly, the desired fragment of 6174 bp may have been produced, but in such a small amount that could not be detected. Therefore, we performed another PCR screening, but instead of directly using the genomic DNA, we purified all PCR product generated by primers 5'F/6R as template and used them for a second amplification, now using primers CopGFPF/CopGFPR in order to detect the selection cassette (**Figure 15A**). In this case, the selection cassette-derived fragments would result from *ATXN3*-derived PCR-products amplified from the intended insertion site, indicated that insertion was successful.

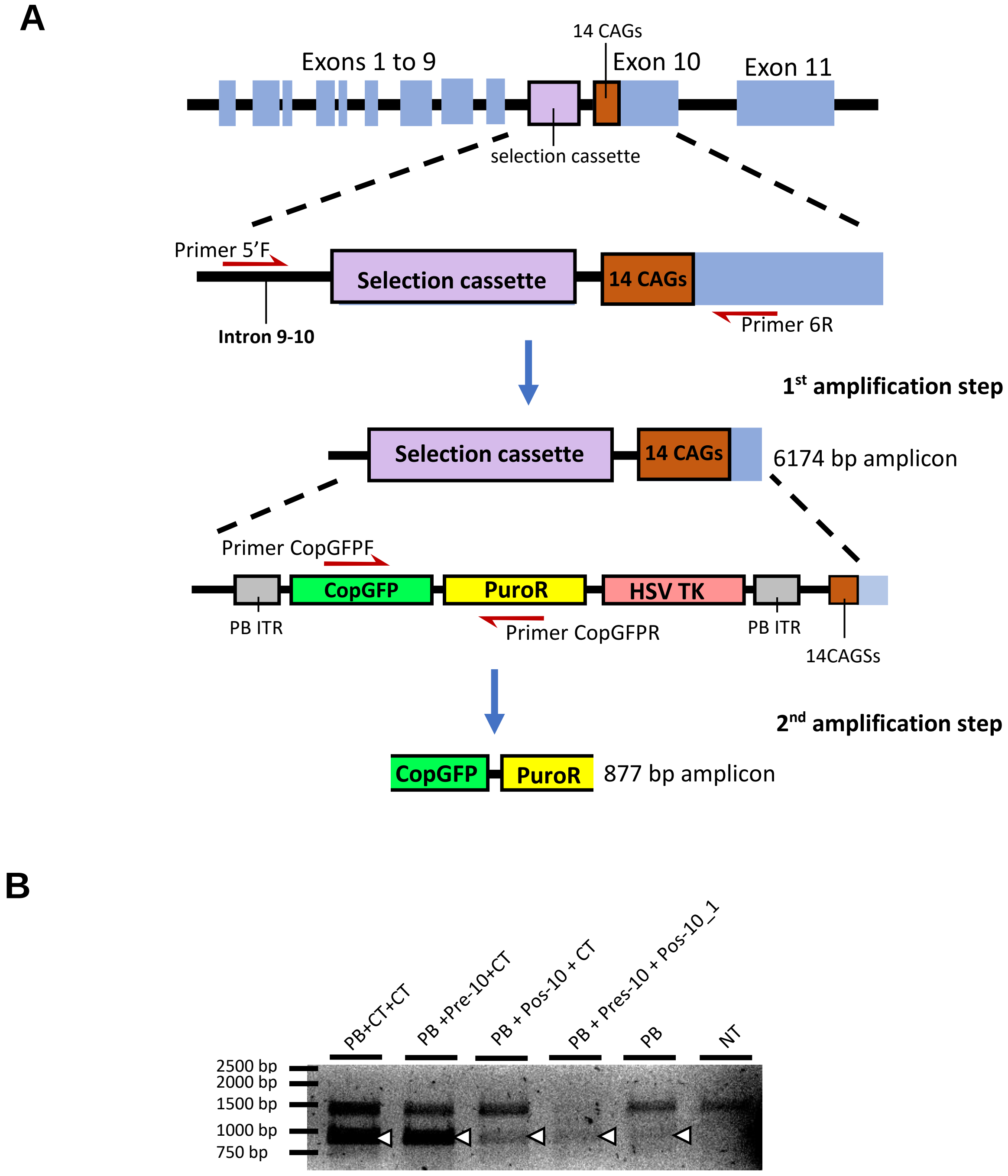
A
14 CAGs
Exons 1 to 9
Exon 10
Exon 11
selection cassette
Primer 5'F
Selection cassette
14 CAGs
Intron 9-10
Primer 6R
1st amplification step
Selection cassette
14 CAGs
6174 bp amplicon
Primer CopGFPF
CopGFP
PuroR
HSV TK
PB ITR
Primer CopGFPR
PB ITR
14CAGSs
2nd amplification step
CopGFP
PuroR
877 bp amplicon
B
PB+CT+CT
PB +Pre-10+CT
PB + Pos-10 + CT
PB + Pres-10 + Pos-10_1
PB
NT
2500 bp
2000 bp
1500 bp
1000 bp
750 bp

Figure 15 - Schematic representation of the PCR products generated by the primers 5'F and 6R, upon a second amplification step using primers CopGFPF/CopGFPR and the respective PCR screening analysis of cells transfected with the Cas9 and the Donor-14Q-Puro. A) Schematic representation of the amplicons generated by the CopGFPF/CopGFPR using as template the PCR fragments generated by the primers 5'F/6R, in a second amplification step. The PCR fragments from the second step of amplification are expected to have approximately 877 bp. **B)** Genomic DNA of HEK 293T cells transfected with the Donor-14Q-Puro and the Cas9 targeted at the vicinities the CAG repeats, was purified and amplified by PCR using the pair of primers 5'F and 6R. PCR product were purified and subjected to a second amplification step, but this time were used the primers CopGFPF and CopGFPR to amplify a fragment of the Donor-14Q-Puro selection cassette**.** Agarose gel electrophoresis analysis of the resulting PCR products showed that, in all conditions tested, an amplicon of approximately 877 bp was generated, corresponding to the expected selection cassette-derived fragment. Results are in accordance with the presence of the selection cassette in the PCR products resulting from *ATXN3*. **Arrowheads** – expected fragment of approximately 877 bp; **CT** – sgRNA that targets a sequence of DNA that is not present in the human genome, thus acting as a control sgRNA; **NT** – non-transfected cells **PB** – cells transfected only with Donor-14Q-Puro plasmid construct.

Interestingly, in all condition yielded a fragment of 877 bp, corresponding to the presence of the selection cassette in the vicinity of *ATXN3* exon 10 (**Figure 15B**). Once again, the putative integration occurred in a double-stranded break/Cas9-independent manner, since the amplicon could be detected even in the condition in which a control sgRNA was used (PB+CT+CT), and in the condition in which no sgRNA was used at all (PB).

Together, these results suggest that the selection cassette present in the repair template of the Donor-14Q-Puro construct was integrated at the vicinity of exon 10 of the human *ATXN3,* possibly at the intended site. Nevertheless, results also demonstrate that integration may occur independently from the double-stranded breaks induced by the Cas9 activity, through undisclosed, possibly homology-mediated mechanisms. Observations are not sufficient to exclude the possibility that that integration may have also occurred at other genomic sites, independently from the homology shared between the template and the human *ATXN3* gene of HEK cells.

Interestingly, even after 44 days of exposure to puromycin, a period 22 days longer than that of the previous experiments, cells were able to survive in culture. In order to determine if the donor plasmid might have been episomally maintained in the cells during this period, another PCR probing experiment was performed, using a primer that targeted the selection cassete (PB primer) and another that targets the donor backbone (PB.BB; **Figure 16A**). The assay would amplify a Donor-14Q-Puro plasmid-specific amplicon, i.e, an amplicon that would only by generated if the plasmid was still present in the cells. The amplification product would have 3199 bp and would not be generated by the selection cassette if insertion occurred as intended, through homology arm-mediation. Since the region that would presumably be integrated by HDR was in between the 5’ and 3’ homology arms, detection of any amplicon using the PB/PP.BB would mean that the plasmid remained episomal or that, somehow, regions from the plasmid beyond those between the homology arms were also inserted into the genome. The results obtained support the hypothesis that at least part of the Donor-14Q-Puro plasmid may have been integrated randomly in a double-stranded break-independent way, and/or that it still remains episomally in cells (**Figure 16B**).

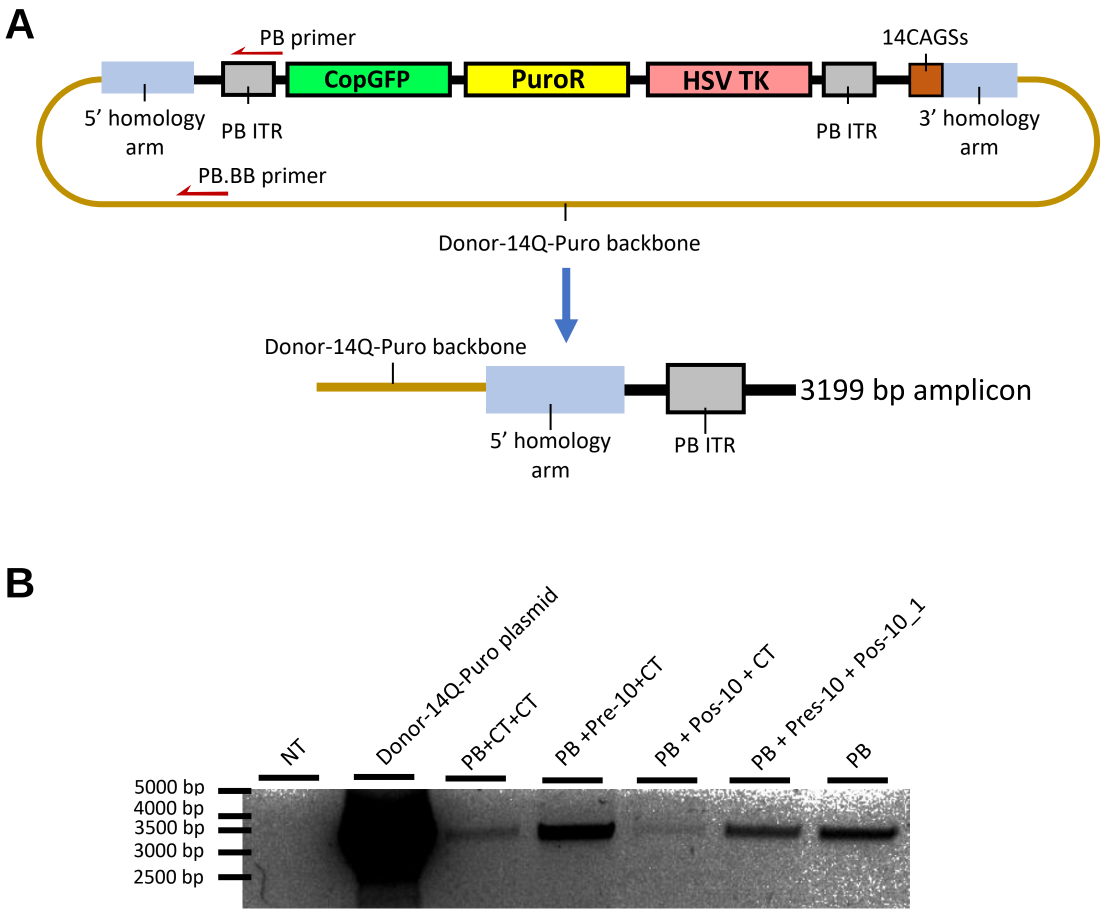
A
PB primer
14CAGSs
CopGFP
PuroR
HSV TK
5’ homology arm
PB ITR
PB ITR
3’ homology arm
PB.BB primer
Donor-14Q-Puro backbone
Donor-14Q-Puro backbone
3199 bp amplicon
5’ homology arm
PB ITR
B
NT
Donor-14Q-Puro plasmid
PB+CT+CT
PB +Pre-10+CT
PB + Pos-10 + CT
PB + Pres-10 + Pos-10_1
PB
5000 bp
4000 bp
3500 bp
3000 bp
2500 bp

Figure 16 - Schematic representation of the amplicon generated by the primers PB/PB.BB and PCR screening analysis of cells transfected with the Cas9 and the Donor-14Q-Puro. Genomic DNA of HEK 293T cells transfected with the Donor-14Q-Puro and the Cas9 targeted at the vicinities the CAG repeats, was purified and amplified by PCR using the pair of primers PB and PB.BB. **(A)** The designed primer pair targets a region of the piggyback-flanked selection cassette and the backbone of the plasmid, originating an amplicon of 3199 bp. **(B)** Agarose gel electrophoresis revealed an amplicon of approximately 3199 bp in all samples obtained from cells transfected with Donor-14Q-Puro. **CT** –sgRNA that targets a sequence of DNA that is not present in the human genome, thus acting as a control sgRNA; **NT** – non-transfected cells. **PB** – cells transfected only with Donor-14Q-Puro plasmid construct.

Regardless of this unintended consequences of transfection with the Donor-14Q-Puro plasmid, in order to get further evidence that the template was still being correctly inserted at the vicinity the exon 10 of the human *ATXN3* gene in at least a part of the cell population, we used another approach, employing a new pair of primers (**Figure 17**). To overcome the limitation of amplifying > 6000 bp described above, we designed another forward primer (5’F_2), that anneals at the intron 9-10 of the human *ATXN3*, 517 bp upstream the region homologous to the 5’ homology arm. Using as reverse primer the previously used PB primer, a fragment of 1967 bp was expected to be generated from cells that underwent succesful integration. Since the fragment includes regions contained both inside the Donor-14Q-Puro template and at a *ATXN3* genomic region, amplicons would only be detected if the the repair template was correctly inserted (**Figure 17A**).

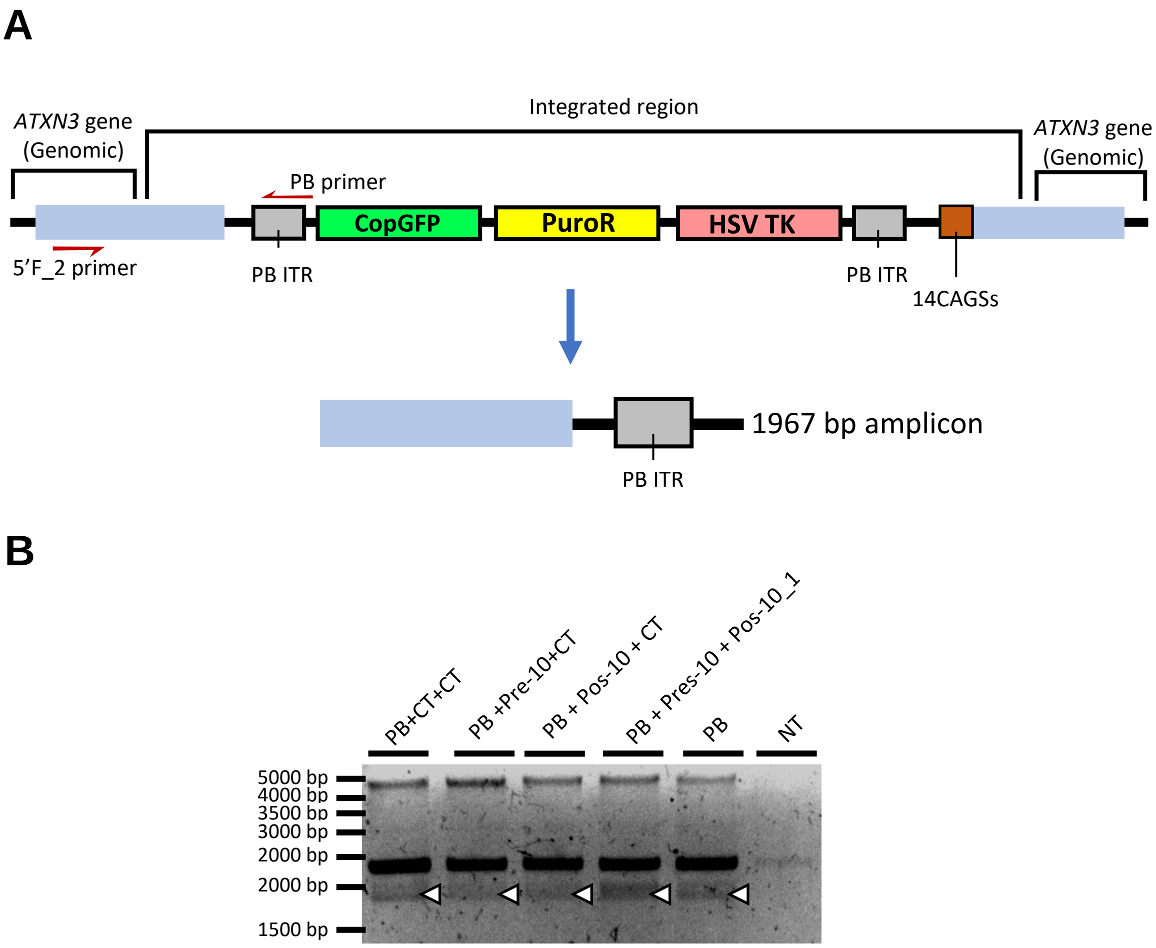
A
ATXN3 gene (Genomic)
Integrated region
ATXN3 gene (Genomic)
PB primer
CopGFP
PuroR
HSV TK
5'F_2 primer
PB ITR
PB ITR
14CAGSs
1967 bp amplicon
PB ITR
B
PB+CT+CT
PB +Pre-10+CT
PB + Pos-10 + CT
PB + Pres-10 + Pos-10_1
PB
NT
5000 bp
4000 bp
3500 bp
3000 bp
2000 bp
2000 bp
1500 bp

Figure 17 - Schematic representation of the amplicon generated by the primers PB/5'F_2 and PCR screening analysis of cells transfected with the Cas9 and the Donor-14Q-Puro. Genomic DNA of HEK 293T cells transfected with the Donor-14Q-Puro and the Cas9 targeted at the vicinities the CAG repeats, was purified and amplified by PCR using the pair of primers PB and 5'F_2. **(A)** The designed primers (PB/5'F_2) target a region of the piggyBac-flanked selection cassette, and the genomic DNA, namely a region of the *ATXN3* human gene, generating an amplicon of 1967 bp. **(B)** Agarose gel electrophoresis showed an amplicon with the expected length - approximately 1967 bp - as resulting from samples obtained from every condition of transfected cells. The only cells that did not showed the intended amplicon were those that were not transfected. **CT** – sgRNA that targets a sequence of DNA that is not present in the human genome, thus acting as a control sgRNA; **NT** – non-transfected cells. **PB** – cells transfected only with Donor-14Q-Puro plasmid construct.

A PCR amplification product compatible with integration of the repair template at the vicinity of exon 10 of the human *ATXN3* gene was detected in every sample transfected with the Donor-14Q-Puro plasmid. Furthermore, as observed before (**Figure 15**), integration was independent from double-stranded breaks induced by Cas9 at the

ATXN3 gene, since the amplicon of 1967 bp was detected in two conditions in which no cut was induced: in cells transfected only with the Donor-14Q-Puro and cells transfected with the repair template, Cas9 and the control sgRNA sequence (**Figure 17B**). Western blot analysis of cells transfected with the Donor-14Q-Puro and the constructs codifying for Cas9 and the sgRNAs - Pre-10 or Pos-10_1 - showed no differences in the ataxn-3 molecular weight between the different conditions tested. This result indicates that, although we were able to integrate the repair template at the intended *locus*, integration had no effect on the protein weight (**Figure 18**). The presence of the selection cassette at intron 9-10 admittedly leds to no splicing aberrations that may be detected at the protein level.

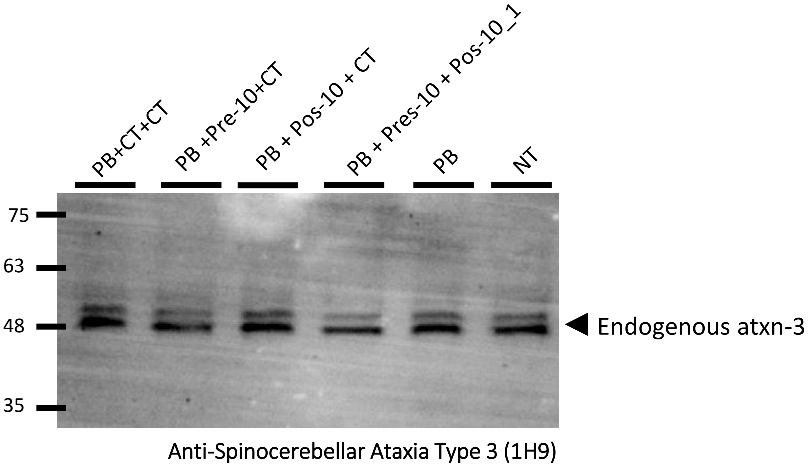

Figure 18 - Detection of human atxn-3 protein by Western blot analysis of the cells transfected with the Cas9 and the Donor-14Q-Puro. Protein extracts of HEK 293T cells transfected with the Donor-14Q-Puro and the Cas9 targeted at the vicinities the CAG repeats, were purified and analysed by Western blot using Anti-Spinocerebellar Ataxia Type 3 Antibody (1H9) antibody to stain human atxn-3 protein. Bands corresponding to HEK cells endogenous atxn-3 presented the expected molecular weight in all of the conditions (arrowhead), and no aberrant atxn-3-derived species were detected. **CT** –sgRNA that targets a sequence of DNA that is not present in the human genome, thus acting as a control sgRNA; **NT** – non-transfected cells. **PB** – cells transfected only with Donor-14Q-Puro plasmid construct.

In summarizing, cells transfected with the Donor-14Q-Puro plasmid and the constructs codifying for Cas9 and the sgRNA - Pre-10 or Pos-10_1 - acquired puromycin resistance and were able to maintain it for more than 22 days (up to 44, at least). This puromycin resistance can be explained by the permanence of the Donor-14Q-Puro plasmid, even after 44 days in culture. However, results also indicate that integration of the selection cassette compatible with the intended editing of the target *ATXN3* gene at the exon 10 *locus* also occurs, at least in a fraction of transfected cells (**Figure 15 and 17)**.

Whether the selection cassette or also other regions of the Donor-14Q-Puro were also integrated at unspecific regions (which explains why resistance was maintained for such a long culture period after transfection) remains to be determined (**Figure 16**). Importantly, integration of the selection cassette, and possibly other regions of the repair template, was independent from the Cas9 activity.

Curiously, we observed that cells transfected with the empty Donor-14Q-Puro donor plasmid (lacking the 5' and 3' homology arms and the 14 CAG repeats) remained resistant to puromycin for more than 5 weeks (experiments not shown). This observation, based on the ability of the cell cultures to survive to puromycin exposure, indicates that the selection cassette and/or the whole donor plasmid was somehow maintained by HEK cells weeks upon transfection. Since this occurred independently from the presence of homology arms, the simple development of puromycin resistance when using this type of repair template construct cannot be taken to guarantee that correct repair template insertion occurred. This constitutes a daunting limitation of the use of this type of construct as a means to edit cells and select them. The strategy does not allow selection of correctly edited cells because there is no way to a distinguish between properly edited cells from the cell in which the Donor-14Q-Puro plasmid was randomly integrated or remained episomally.

Finally, while results suggest that the designed repair template might have been correctly integrated into transfected HEK cells, this methodology is not reliable for the creation of isogenic cell lines for MJD, and will consequently need further refinement.

The cause for the lack of HDR under the above experimental conditions (both in the overexpansion and retraction of the *ATXN3* CAGs experiments) is unknown, though strategies may be developed to increase the changes of it occurring. It is important to note that HDR is a rare event that occurs only during the S or G2 phases of the cell cycle (Heyer et al., 2010; Lin et al., 2014). Therefore, arresting and synchronizing cells at a specific cycle phase, followed by the delivery of the donor template and the CRISPR/Cas9 tools may be a possible approach to overcome limitations of our experimental paradigm. Other studies have shown that using a pair of Cas9 nickases reduces the off-target DSB effects and increases HDR events frequency (Gopalappa et al., 2018; X. Xu et al., 2017). It also seems that finding an optimal ratio between Cas9 and the sgRNAs is crucial to increase HDR. Previous reports showed that a 5:1 ratio (Cas9/sgRNA) can increase the percentage of HDR (Pinder et al., 2015). Gou and co-workers showed that delivery of the Cas9 protein as RNA by lipids in conjunction with a cold-shock during 24 h - 48 h at 32 ^{o}C after transfection can increase HDR up to 2-fold (Guo et al., 2018).

Other strategies to enhance HDR rely on using proteins that modulate the cell cycle or proteins that favour the HDR pathway over the NHEJ. Studies showed that arresting cells at G2/M with Nocodazole or ABT-751 (they inhibit microtubules polymerization) can increase HDR efficiency up to 3- and 6-fold (Yang et al., 2016). Inhibition of the NHEJ pathway with NU7441 (an inhibitor of the DNA-dependent protein kinases, involved in early stages of the NHEJ pathway) and using homology arms longer than 1500 bp seems to increased HDR events up to 13.4-fold (Aksoy et al., 2019). Chu and collaborators showed that suppression of the DNA ligase IV either by shRNAs or through its degradation using E1B55K and E4orf6 adenovirus 4 proteins, increases HDR up to 5-fold (Chu et al., 2015). A recent study also reported that fusing Cas9 with an N-terminal fragment (296 aa) of CtIP (CtIP is a protein involved in early stages of HDR) enhances HDR up to 2-fold (Charpentier et al., 2018).

Further studies and a new methodology must be adopted in order to improve our results and accomplish efficient HDR. The intended generation of promising MJD models – both the refinement of the YAC MJD-Q84.2 mouse or the production of patient-derived cell lines - will entail reliable methods for altering CAG tract length of the human *ATXN3*.

3.2 dCas9-KRAB as a tool for in vivo silencing of the *ATXN3* gene

Until now, CRISPR/Cas9 has proved to be an unprecedent tool regarding human gene editing. It is a versatile system that can target virtually any desired genetic *locus*, it is inexpensive and relatively easy to use (Aryal et al., 2018). However, and despite the advances in research, CRISPR/Cas9 still remains a fallible, prone to error tool. CRISPR/Cas9 can lead to off-target effects that can induce double-stranded breaks at undesired *loci*, leading to unwanted mutations such as deletions or insertions that may cause frameshifts or knockouts, disrupting the normal function of genes (X.H. Zhang et al., 2015). When aiming at gene silencing, CRISPR interference (CRISPRi) is a process considered to be potentially safer than nuclease activity-based gene editing, as it does not involve the introduction of DNA double-stranded breaks. This process uses a catalytically dead form of Cas9 (dCas9) that can be fused to a repressor domain such as the Krüppel-associated box (KRAB) (Qi et al., 2013). dCas9 harbours the possibility of repressing and thereby prevent the production of toxic CAG-expanded RNA species that may contribute to polyQ disease pathogenesis (Martí, 2016).

Previous work by your group showed that dCas9-KRAB-mediated silencing can efficiently downregulate the endogenous human atxn-3 protein in HEK 293T cells. Furthermore, allele-specific silencing of mutant atxn-3 can also be achieved by targeting dCas9-KRAB to the disease-associated $G^{987}GG \rightarrow C^{987}GG$ SNP of *ATXN3*. Lentiviral-mediated expression of the necessary silencing molecular tools (dCas9-KRAB and the SNP-specific sgRNA) showed a robust and long-term decrease of expanded protein expression (Pena, 2018).

To assess the potential of this allele-specific strategy to silence human *ATXN3* in vivo, as well as the accompanying putative beneficial effects on MJD-related phenotypic traits, we administered the previously developed silencing molecular tools to an animal model of MJD (Pena, 2018). We used a MJD transgenic mouse model in which a truncated form of human *ATXN3* containing 68 CAG repeats and bearing the disease-associated SNP is expressed in the cerebellum (Torashima et al., 2008).

At 5 weeks of age, lentiviral particles encoding the dCas9-KRAB fusion protein and the sgRNA targeting the MJD-associated SNP (sgRNA_SNP; dCas9-KRAB/sgRNA_SNP) were administered to the mice cerebella by stereotaxic injection. Control animals were injected with lentiviral particles encoding the dCas9-KRAB fusion protein and an incomplete sgRNA sequence (sgRNA_Empty; dCas9-KRAB/sgRNA_Empty), i.e. a sgRNA without the 20 bp variable strand (crRNA) that is responsible for targeting Cas9 to specific DNA sequences.

Since one of the main features of MJD pathology – and a prominent characteristic of the animal model (Torashima et al., 2008) – is the motor impairment, transgenic mice were submitted to behavioural test in order to access motor performance of the animals and the putative effect of human *ATXN3* expression repression. Mice were submitted to behavioural tests in order to access motor performance of the animals and the putative effect of human *ATXN3* expression repression. Mice were first behaviourally tested at 5 weeks of age (prior to injection) and afterwards, three more times every three weeks until sacrifice eleven weeks after injection (**Figure 19**). Mice were then sacrificed to assess the protein and RNA expression profile of the mutant human *ATNX3*. Cerebellar thickness of the molecular, granular Purkinje cell layers of the lobes IX and X was also assessed as an indicator of neurodegeneration.

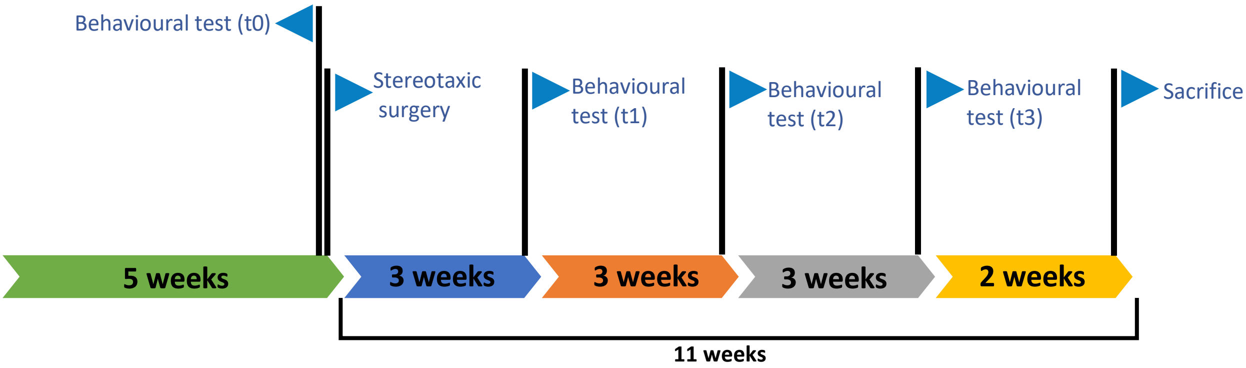

Figure 19 - Schematic representation of the timeline used to behaviourally test and treat MJD transgenic mice with lentiviral particles encoding dCas9-KRAB. At 5 weeks of age (t0), mice were submitted to a set of behavioural tests (Rotarod test, Swimming test, Footprint analysis and Beam Walking). The day after, mice cerebella was stereotaxically-injected with the lentiviral particles, and tests were repeated every three weeks. Finally, mice were sacrificed for neuropathological studies in order to assess mutant RNA protein and levels. Cerebellar thickness of the molecular, granular and Purkinje cell layers of the lobes IX and X was also assessed as an indicator of neurodegeneration.

As the dCas9-KRAB strategy aimed at pre-transcriptionally repressing the mutant human *ATXN3* expressed in the transgenic mice, we investigated whether the administration of the silencing tools downregulated atxn-3 mRNA and protein levels, in mice cerebella – the most affected brain region in MJD.

Expression levels of the mutant atxn-3 mRNA from cerebellar lysates of mice injected with lentiviral particles encoding for dCas9-KRAB/sgRNA_SNP of, as a control, injected with lentiviral particles encoding for dCas9-KRAB/sgRNA_Empty, were assessed by quantitative PCR (qPCR) analysis. Since our strategy targets acted at a pre-transcriptional level, a reduction of the mutant atxn-3 mRNA levels was expected in treated mice, comparing with controls.

However, despite the silencing strategy used, mice treated with dCas9-KRAB/sgRNA_SNP displayed no differences regarding atxn-3 mRNA levels comparing with controls (KRAB/sgRNA_Empty; **Figure 20B**). We further investigated possible differences in mutant atxn-3 protein levels by Western blot, also in the cerebellar lysates from the same transgenic mice treated with dCas9-KRAB. We probed the Western blot membrane with an Anti-HA tag antibody which binds to the human influenza hemagglutinin (HA) epitope that is fused with the truncated form of the mutant atxn-3 expressed in the transgenic mouse model (Torashima et al., 2008). However, results from protein level assessment were in accordance with atxn-3 mRNA level analysis – Western blot analysis showed no differences regarding mutant atxn-3 protein, between dCas9-KRAB/sgRNA_SNP and control-treated mice (**Figure 20C**).

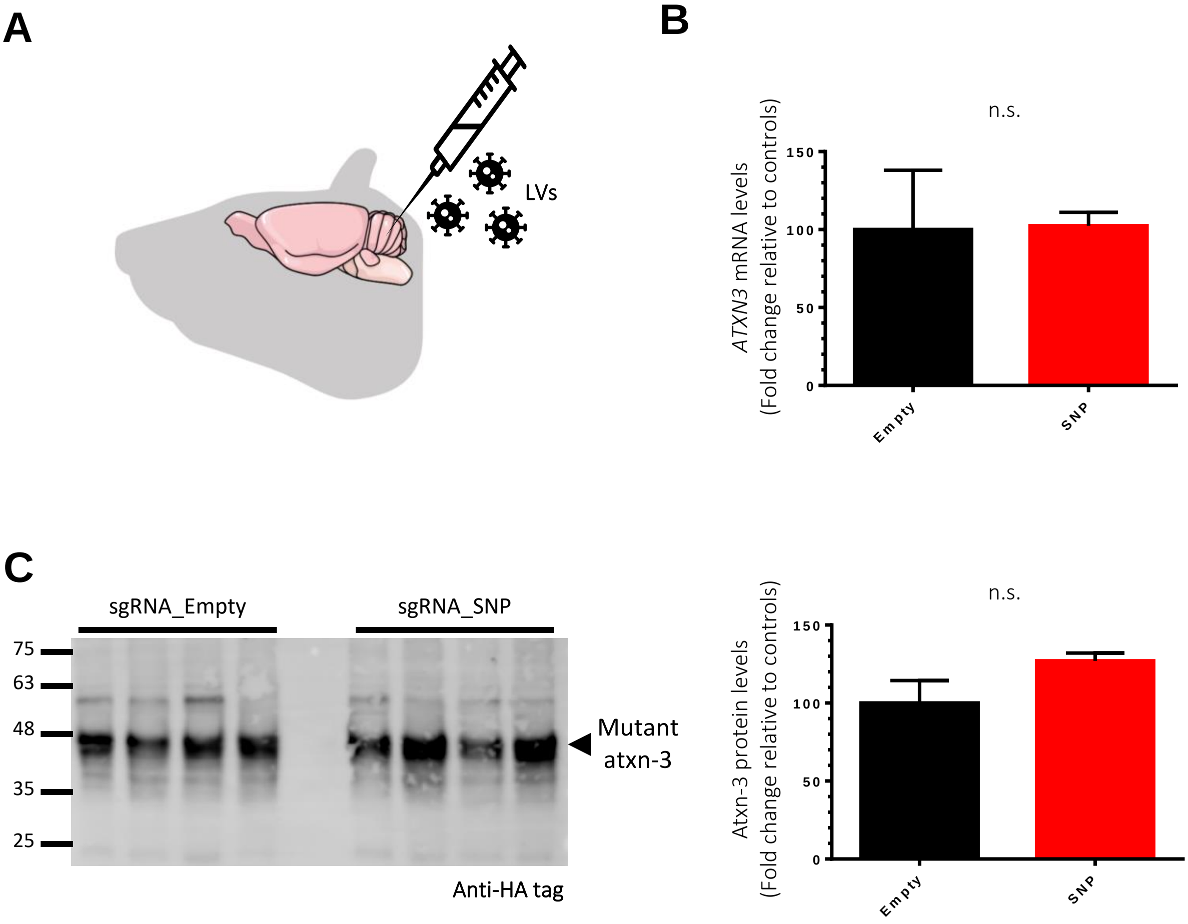

Figure 20 - *ATXN3* expression profile from transgenic mice cerebella, stereotaxically- injected with lentiviral particles encoding dCas9-KRAB/sgRNA_SNP or dCas9-KRAB/sgRNA_Empty. Mice were sacrificed eleven weeks after stereotaxic injection with lentiviral particles encoding for the dCas9-KRAB/sgRNA_SNP targeted at the disease-associated SNP or, as control, with the dCas9-KRAB/sgRNA_Empty. Cerebellar lysates were used to assess the expression levels of both atxn-3 mRNA and protein. **A)** Schematic representation of transgenic mice stereotaxically-injected at the cerebellum with the silencing molecular tools. **B)** qPCR analysis revealed no difference in atxn3 mRNA levels between mice treated with dCas9-KRAB/sgRNA_SNP and the respective controls injected with dCas9-KRAB/sgRNA_Empty **C)** Anti-HA tag antibody Western boot staining revealed a band with ~ 45 kDa corresponding to the mutant atxn-3 protein. Comparison between mice treated with dCas9-KRAB/sgRNA_SNP and the respective controls, revealed no significative differences regarding protein levels. **LVs**: lentivirus. (n=8; treated=4, controls= 4). Data is presented as mean ± SEM. n.s.: p>0.05; Unpaired Student's t test.

qPCR and Western blot and analysis revealed that both mutant atxn-3 mRNA and protein expression levels, respectively, were not altered by the silencing molecular tools. This means that, despite previous results obtained *in vitro,* in which dCas9-KRAB targeting the disease- associated SNP was able to successfully reduce the mutant *ATNX3* expression levels (Pena, 2018), the same dCas9-KRAB/sgRNA_SNP silencing strategy was not efficient at silencing the *ATXN3* pre-transcriptionally *in vivo*.

Importantly, mice were subjected to a single injection in the cerebellum, with lentiviral particles, at the coordinates corresponding to lobules IX/X. However, the entire cerebellum was processed to obtain lysates for mutant protein and mRNA analysis. Possibly, the lack of differences observed was due to a dilution effect when analysing the whole cerebellum. If a confined area corresponding to the lentiviral spreading profile was analysed instead, success of the strategy may have yielded observable differences in mutant antx-3 protein and mRNA expression levels.

The transgenic mice used is characterized by a persistent neurodegeneration that can be observed by measuring cerebellar lobules thickness. Therefore, we used cresyl violet to stain mice brain slices and assess whether the treatment with dCas9-KRAB targeted at the disease associated SNP could preserve cerebellar integrity (**Figure 21A–B**). We measured the thickness of the molecular, granular and Purkinje cell layers of cerebellar lobules IX and X, expecting treated mice to display thicker cerebellar layers compared to control, as an indicator of neurodegeneration amelioration.

Cresyl violet staining revealed that cerebellar thickness of granular, molecular and Purkinje cell layers of lobules IX and X from mice treated with the dCas9-KRAB/sgRNA_SNP displayed the same width as the respective controls (**Figure 21C**). Results indicate that targeting dCas9-KRAB to mutant *ATXN3* in the model was not able to prevent neurodegeneration. Possible, the absence of differences observed is due to the lack of *ATNX3* silencing by the dCas9-KRAB/sgRNA_SNP, observed by the qPCR and Western blot analysis. In this context, the silencing strategy and its putative neuroprotective effects would have been unsuccessful.

As described before one of the main features of MJD is the motor impairment displayed by patients (see section *Disease description and clinical features*). Foremost, a therapeutic approach must be able to ameliorate motor deficits in order to provide a better lifestyle to patients. Therefore, we evaluated whether the expected gene silencing using dCas9-KRAB targeted at the disease-associated SNP alleviated the ataxic phenotype displayed by the transgenic mice used in this study.

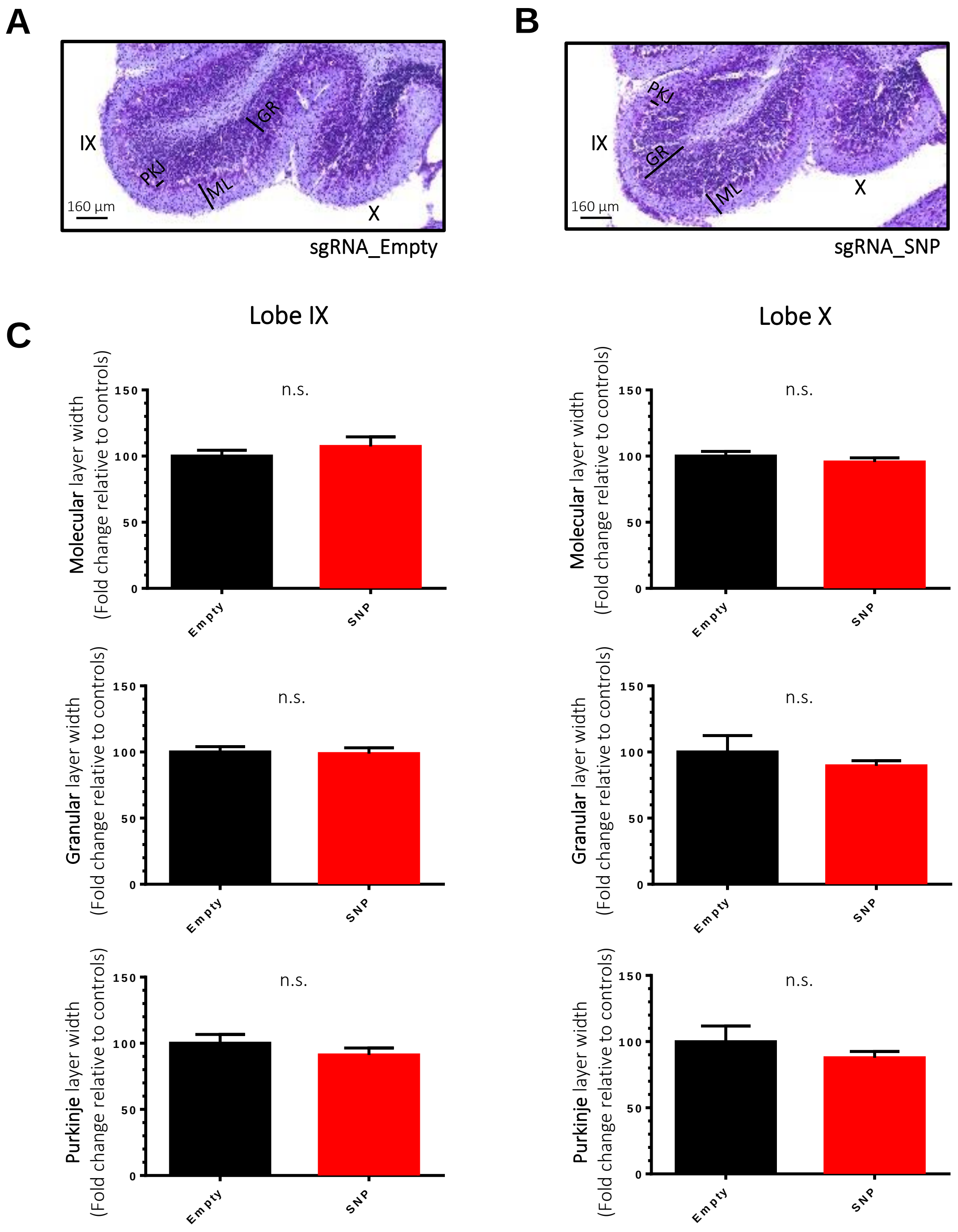

Figure 21 - **Evaluation of the cerebellar degeneration profile in mice injected with lentiviral particles encoding for the dCas9-KRAB/sgRNA_SNP.** As an indicator of neurodegeneration, brain slices were stained with cresyl violet and cerebellar molecular, granular and Purkinje cell layers of lobes IX and X from stereotaxically-injected mice with **A)** dCas9-KRAB_SNP or **B)** dCas9-KRAB_Empty were measured. **C)** Measurement of cerebellar layers thickness revealed no significative differences between mice treated with the dCas9-KRAB/sgRNA_SNP and the respective controls injected with dCas9-KRAB/sgRNA_Empty. (n=12; treated=6, controls=6), **IX** – lobule IX; **X** – lobule **X**; **ML** – molecular layer, **GR** – granular layer, **PKJ** – Purkinje layer. Data is presented as mean ± SEM. **n.s.**: $p>0.05$; Unpair Student's t test.

As mentioned above, in order to assess the motor performance of model mice submitted to stereotaxic injection with the dCas9-KRAB/sgRNA_SNP, animals were submitted to a battery of behavioural tests, including Rotarod – measuring the latency of the animals to fall from a revolving platform; Swimming test – measuring the time taken to transverse a water- filled tank; Beam Walking – measuring the time taken to transverse a wide or narrow horizontal beam; and Footprint test – measuring motor coordination based on feet positioning when walking.

Mice were injected at 5 weeks of age and at that point the transgenic animals used already present a prominent motor deficit (Torashima et al., 2008). After stereotaxic injection, mice were tested three times with three weeks apart from each test (**Figure 19**). As indicators of motor improvement, it was expected that mice treated with the dCas9-KRAB/sgRNA_SNP would be faster at crossing both beams and transverse the water-filler tank. Regarding the Footprint test, treated mice were expected to have an increased stride length, decreased distance between right and left stride, and have a paw overlap measure close to zero. Finally, in the Rotarod test, we expected that treated mice would have an increase in the latency to fall from the apparatus (Carter et al., 2001; Nóbrega et al., 2014, 2013b).

We observed that motor deficits in the Beam Walking, Swimming and Footprint tests did not improve over the 11 weeks of the study (**Figure 22A–F**). Rotarod results showed that in the first three-week period post-injection both mice groups displayed similar performances. However, in the second and third periods of tests, a motor performance improvement was evident (**Figure 22H**). Mice Stereotaxically-injected with the dCas9-KRAB/sgRNA_SNP performed better comparing to the controls injected with the dCas9-KRAB/sgRNA_Empty, at six and nine week after surgery, being able to remain on the revolving Rotarod apparatus for about double the time of the controls. These results suggest that targeting dCas9-KRAB to the disease-associated SNP of human *ATXN3* of the transgenic model mice was able to ameliorate the motor coordination deficit displayed by the MJD mouse model.

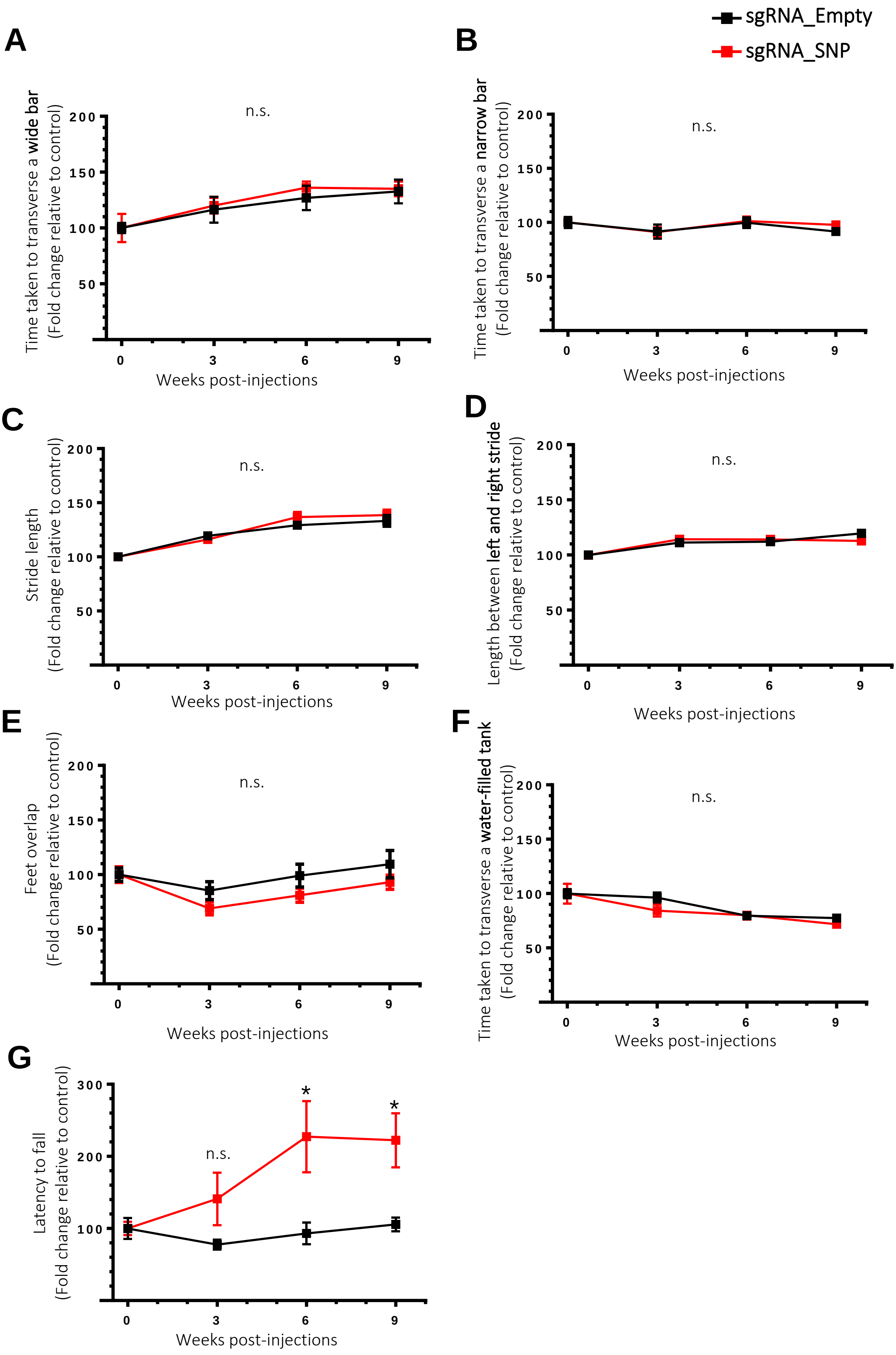

Figure 22 - Motor performance of MJD transgenic mouse model at three, six and nine weeks after stereotaxic injection with lentiviral particles encoding the dCas9-KRAB targeted at the MJD-associated SNP of human *ATXN3*. Prior to stereotaxic injection with lentiviral particles encoding for dCas9-KRAB/sgRNA_SNP or, alternatively as control, with dCas9-KRAB-sgRNA_Empty, transgenic mice were submitted to a battery of behavioural tests. (Figure description continues next page)

Figure 22 - After injection with the molecular silencing tools, mice were tested three more times every three weeks. Mice were submitted to the Beam Walking test and the time taken to transverse a **A)** wide beam (50x1.8 cm) or **B)** a narrow beam (50x0.9 cm) was measured. Mice were also submitted to the footprint analysis and **C)** stride length, **D)** length between left and right stride and **E)** feet overlap distances were measured. **F)** We also measured how much time mice took to transverse a 1.5 m long water-filler tank. **A) – F)** No differences were observed in the Beam Walking, Swimming, and Footprint tests between treated mice and controls, during the 9 weeks of testing. **G)** Regarding the Rotarod test, six weeks after treatment, mice treated with dCas9-KRAB/sgRNA_SNP already showed a statistically-significant > 2-fold change in their balance and latency to fall, compared with non-treated mice. That fold change was maintained at least until nine weeks after injection. Data is presented as mean ± SEM. **n.s.**: $p>0.05$; *****$p<0.05$; Unpaired Student's t test; n=20 (treated=10, controls=10).

Taking both behavioural tests and neuropathological analysis into consideration, we observed that results show some disparity. While motor performance was improved upon dCas9-KRAB treatment, no amelioration was observed regarding neuropathology. A possible explanation for the results divergence might be the fact that a severe neuropathological and motor phenotype was already established at the time of treatment. At 3 weeks of age, transgenic mice already display a distinctive MJD phenotype, which includes severe ataxia and severe cerebellar atrophy (Torashima et al., 2008). Since mice were stereotaxically-injected with dCas9-KRAB_sgRNA_SNP at 5 weeks of age, possibly, at this timepoint, molecular and cellular disease profile cannot be reverted using the CRISPRi strategy that was tested. Motor performance may have a wider window of opportunity for amelioration, using this type of silencing strategy.

Finally, another aspect to take into consideration is the injected dose of lentiviral particles encoding dCas9-KRAB and sgRNA_SNP. The dose used might be too low to produce expressive differences. An earlier administration of a higher dose of lentiviral particles may produce a more significant amelioration of the MJD-related phenotypic features of the model.

Chapter 4 – Conclusion

Amigos há vinte anos! Não há muita gente com a nossa idade que possa dizer o mesmo. Desde sempre o culto do conhecimento esteve entranhado no teu Eu-Criança, no teu Eu-Adolescente e perdurou no tempo até ao teu Eu-Adulto. Lembro com carinho todos os complexos projetos de engenharia que criávamos nas nossas cabeças e que chegamos algumas vezes a passar para o papel em forma de algo muito parecido com figuras rupestres. Lembro ainda todas as vezes em que, perante uma qualquer dúvida, criávamos as mais fantásticas teorias de solução de um problema ou justificação de um facto nas nossas cabeças. Eram horas e horas a de reflexão acompanhadas de leite e cereais. Quem sabe se juntos não estivemos perto de descobrir a cura para o cancro ou uma forma de acabar com a guerra no mundo. Quem sabe!?

Agora as coisas ficaram (ainda) mais sérias para o teu lado. Eu sei que o adulto em que te tornaste nunca vai esquecer a paixão que o teu Eu-Criança e o teu Eu-Adolescente sentia na busca pelo conhecimento. Foi essa paixão que te tornou o que estás perto de ser agora. Nunca a percas! Coloca essa paixão em todos os teus projetos e investigações e eu sei que vais ser parte da História. No futuro serei como que um amuleto e estarei cá para te aplaudir como sempre estive nestes vinte anos.

— J.D. Santos, um Amigo

"The oldest and strongest emotion of mankind is fear, and the oldest and strongest kind of fear is fear of the unknown"

— H.P Lovecraft

Conclusion and future perspectives

Currently, there is no effective curative treatment for MJD and, furthermore, the cellular and molecular mechanisms underlying the disorder are still poorly understood. It is known that an abnormal CAG expansion in the exon 10 of the *ATXN3* causes the disease, but the implications of having an aberrant protein with an expanded polyQ tract are not yet clear. Consequently, further research must be conducted before a complete picture of disease pathogenesis can be drawn.

The work hereby presented had the objective of assessing the potential of new gene editing approaches to refine MJD experimental models and contribute to the development of rational MJD therapies. Using CRISPR/Cas9 we were able of both edit the endogenous *ATXN3* gene in a human cell culture and ameliorate the characteristic motor impairment of a transgenic MJD mouse model. The results reveal that CRISPR/Cas9-based gene editing tools can be promising tools to treat MJD.

The first part of this work intended at establishing a molecular strategy to overexpand and retract the CAG repeats in the exon 10 of the human *ATXN3* gene. The overexpansion had the objective of being applied to the refinement of a previously existing YAC MJD-Q84.2 mouse model, in order to create a more suited model to work in laboratory, in which the disease phenotype would displays an earlier onset and more severe symptoms. The CAG retraction strategy aimed at being employed to generate patient-derived isogenic iPSC lines as an *in vitro* model of MJD, in which the only genotype difference would be the gene that is causing the disease. Isogenic lines derived from patients may be generated, for example, by collecting fibroblasts from affected people and correcting the mutation via gene editing (Torashima et al., 2008).

With the objective of provoking *ATXN3* CAG overexpansion in HEK cells, we used a pair of sgRNAs that targeted the vicinities of the CAG repeats present in the exon 10 of the human *ATXN3* gene. We also provided a DNA donor template containing 141 CAG repeats (Donor-141), that would be integrated by HDR upon double-stranded break, thus expanding the number of CAG repeats of the endogenous *ATXN3* gene. Despite the efforts, we were neither able to detect an expanded form of the *ATXN3* gene, nor an expanded atxn-3 protein. However, regarding the CAG retraction experiments, results suggest that the provided DNA donor template (Donor-14Q-Puro), containing 14 CAG repeats and a selection cassette, was correctly integrated, in at least a fraction of the transfected cell population. The main objective of the selection cassette was to allow for positive selection of cells that underwent HDR upon the double stranded breaks in the vicinities of the CAG repeats. Nevertheless, we observed that integration of the Donor-14Q-Puro was independent from the double-stranded breaks induced by Cas9. Cells that were transfected only with the donor template without any double-stranded breaks inducer were able to survive for long periods in culture under puromycin selection. This suggested that the provided donor has the ability to be maintained after rounds of cellular division, either epissomally or after being integrated, while it would be expected that the plasmid, presumably unable to be replicated in HEK cells, would be diluted among cells throughout the culture period.

The mechanisms by which the donor template was integrated in the absence of the targeted double-stranded breaks are not easily fathomed. Considering that the template includes a selection cassette flanked by transposon-derived sequences (a transposon is a genetic element that can transpose between *loci*), we hypothesised that those flanking regions might be recognized by HEK 293T cells transposases and be transposed from the plasmid to the genome. In fact, Henssen and collaborators show that the human PGBD5 protein is responsible for DNA transpositions and insertions at TTAA sites in a way similar to the action of the *piggyBac* transposon (Henssen et al., 2015). Furthermore, it appears that the human genome has at least 328 *piggyBac*-like elements with TTAA target sites. In this study they transfected HEK cells with a donor plasmid containing a gene for neomycin resistance flanked by *piggyBac* ITRs and show that PGBD5 activity can transpose the resistance gene from the plasmid to the genome (Henssen et al., 2015). These results might explain why the naked Donor-14Q-Puro (lacking the 5’ and 3’ homology arms and insert) was able to confer long-term puromycin resistance to transfected cells (the experiment was not continued any longer than 5 weeks).

An approach that might enhance HDR events in order to integrate donor templates by HDR, that we intend to employ in the future, is to electroporate cells with both the Cas9 ribonucleoprotein (RNP), conjugated with gRNAs, along with the donor plasmids, via nucleofection using a 4D-Nucleofector™ System (Aird et al., 2018; Lin et al., 2014). Standardized protocols and kits containing HDR enhancers that have been optimized to display the best levels of gene editing via HDR have been developed and are available to the scientific community. Similarly, to the approach used in the present study, this RNP-base strategy also relies on DNA cuts to induce HDR events. Another difference we intend to employ is the use of a Cas9 variant, in which only one catalytic nuclease site is active – Cas9 nickase (nCas9). Published results suggest that both use of nucleofection to electroporate Cas9 RNP and the use of the nCas9 variants can increase the HDR efficiency while decreasing off-target events (Liu et al., 2019; Ran et al., 2013; Sentmanat and Shondra, 2016).

In the second part of the project presented in this work we assessed the beneficial effects of targeting catalytically dead Cas9 fused with the KRAB repressor domain (dCas9-KRAB) to mutant *ATXN3* as a therapeutic approach to treat MJD. The strategy aimed silencing the mutant human *ATXN3* gene present in the MJD transgenic mice. We targeted the dCas9-KRAB to a disease associated SNP in order to perform an allele-specific silencing, in which only the mutant *ATXN3* would be repressed. Since this strategy acts at the pre-transcriptional level, it would prevent the putative toxic effect of the CAG expanded *ATXN3* transcripts. We stereotaxically-delivered the silencing molecular tools to the cerebellum of 5 weeks old MJD transgenic mice. Neuropathological results suggested that this therapy did not prevented neurodegeneration. The strategy was not able to revert the shrinkage of the thickness of the layers of cerebellar lobes X and IX. Moreover, while as the main effect of the therapy would be to repress mutant *ATXN3* gene expression (and consequently prevent its toxic effects), qPCR analysis and Western blot releveled that human atxn-3 mRNA and protein levels of mutant atxn-3 seem not to be altered by the treatment. However, we observed that treated mice displayed increased motor coordination, assessed by the Rotarod test, when compared with non-treated animals. This result is of important relevance since motor impairment is the most limiting factor in MJD patients.

One fact that may be contributing for the lack of *ATXN3* expression repression and neuropathological amelioration is the possibility of an insufficient expression of the dCas9-KRAB/sgRNA_SNP therapy. This may might be overcome by trying to determine a dose-response curve to understand if the injected dosage was too low and whether it should be increased.

Importantly, the transgenic mice employed in this study displayed a distinctive severe phenotype of MJD as early as 3 weeks of age. Possibly, in order to detect more expressive motor and neuropathological improvements, mice must have been injected sooner. However, due to animal welfare regulations it is not indicated to subject mice to surgical interventions earlier. Nevertheless, to better understand the beneficial effect of the CRISPRi strategy, further studies will be conducted. For example, immunofluorescence studies to assess the expression and dispersion throughout the cerebellar lobules of the dCas9-KRAB protein as well as the expression of the mutant atxn-3 to comprehend the extension of the therapy would be advantageous. Furthermore, atxn-3 protein aggregation - a hallmark of MJD pathology – can be investigated by observation of ubiquitin-positive intranuclear inclusions (S. Alves et al., 2008; Mendonça et al., 2015; Paulson et al., 1997).

Further studies will be necessary in order to definitely establish a strategy for editing the number of CAG repeats of human *ATXN3* and validate the developed CRISPRi strategy as an effective therapeutic approach for MJD *in vivo*. Currently, the drugs for available MJD treatment only help to alleviate the symptoms. Until now, no effective therapy that could delay the symptoms or cure the disease were developed. However, with the advent of new gene therapies tools, such as CRISPR/Cas9-based systems, the once utopic dream of definitely cure a genetic disorder has become more and more reachable.

References

Abe, Y., Tanaka, F., Matsumoto, M., Doyu, M., Hirayama, M., Kachi, T., Sobue, G., 1998. CAG repeat number correlates with the rate of brainstem and cerebellar atrophy in Machado-Joseph disease. Neurology 51, 882–884. https://doi.org/10.1212/WNL.51.3.882

Adli, Mazhar, 2018. The CRISPR tool kit for genome editing and beyond. Nat. Commun. 9, 1911. https://doi.org/10.1038/s41467-018-04252-2

Adli, M., 2018. The CRISPR tool kit for genome editing and beyond. Nat. Commun. 9, 1911. https://doi.org/10.1038/s41467-018-04252-2

Aird, E.J., Lovendahl, K.N., St. Martin, A., Harris, R.S., Gordon, W.R., 2018. Increasing Cas9-mediated homology-directed repair efficiency through covalent tethering of DNA repair template. Commun. Biol. 1, 54. https://doi.org/10.1038/s42003-018-0054-2

Aksoy, Y.A., Nguyen, D.T., Chow, S., Chung, R.S., Guillemin, G.J., Cole, N.J., Hesselson, D., 2019. Chemical reprogramming enhances homology-directed genome editing in zebrafish embryos. Commun. Biol. 2, 198. https://doi.org/10.1038/s42003-019-0444-0

Alves, Sandro, Nascimento-Ferreira, I., Auregan, G., Hassig, R., Dufour, N., Brouillet, E., Pedroso de Lima, M.C., Hantraye, P., Pereira de Almeida, L., Déglon, N., 2008. Allele-Specific RNA Silencing of Mutant Ataxin-3 Mediates Neuroprotection in a Rat Model of Machado-Joseph Disease. PLoS One 3, 1–12. https://doi.org/10.1371/journal.pone.0003341

Alves, S., Nascimento-Ferreira, I., Dufour, N., Hassig, R., Auregan, G., Nóbrega, C., Brouillet, E., Hantraye, P., Pedroso de Lima, M.C., Déglon, N., de Almeida, L.P., 2010. Silencing ataxin-3 mitigates degeneration in a rat model of Machado–Joseph disease: no role for wild-type ataxin-3? Hum. Mol. Genet. 19, 2380–2394. https://doi.org/10.1093/hmg/ddq111

Alves, S., Régulier, E., Nascimento-Ferreira, I., Hassig, R., Dufour, N., Koeppen, A., Carvalho, A., Simões, S., de Lima, M., Brouillet, E., Gould, V., Déglon, N., de Almeida, L., 2008. Striatal and nigral pathology in a lentiviral rat model of Machado-Joseph disease. Hum. Mol. Genet. 17, 2071–83. https://doi.org/10.1093/hmg/ddn106

Antony, P., Mäntele, S., Mollenkopf, P., Boy, J., Kehlenbach, R., Riess, O., Schmidt, T., 2009. Identification and functional dissection of localization signals within ataxin-3. Neurobiol. Dis. 36, 280–292. https://doi.org/10.1016/j.nbd.2009.07.020

Araujo, J., Breuer, P., Dieringer, S., Krauss, S., Dorn, S., Zimmermann, K., Pfeifer, A., Klockgether, T., Wuellner, U., Evert, B., 2011. FOXO4-dependent upregulation of superoxide dismutase-2 in response to oxidative stress is impaired in spinocerebellar ataxia type 3. Hum. Mol. Genet. 20, 2928–2941. https://doi.org/10.1093/hmg/ddr197

Araújo, M., Raposo, M., Kazachkova, N., Vasconcelos, J., Kay, T., Lima, M., 2016. Trends in the Epidemiology of Spinocerebellar Ataxia Type 3/Machado-Joseph Disease in the Azores Islands, Portugal. JSM Brain Sci 1, 1001.

Aryal, N.K., Wasylishen, A.R., Lozano, G., 2018. CRISPR/Cas9 can mediate high-efficiency off-target mutations in mice in vivo. Cell Death Dis. 9, 1099. https://doi.org/10.1038/s41419-018-1146-0

ATCC, 2019. ATCC - Cell lines [WWW Document]. Anim. Cell Lines/Human Cell Lines. URL https://www.lgcstandards-atcc.org/Products/Cells_and_Microorganisms/Cell_Lines.aspx

AveXis/Novartis, 2019. ZOLGENSMA® [WWW Document]. Suspens. Intraven. Infus. URL https://www.zolgensma.com/

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D., Horvath, P., 2007. CRISPR provides acquired resistance against viruses in prokaryotes. Science (80-.). 315, 1709–12. https://doi.org/10.1126/science.1138140

Berke, S., Schmied, F., Brunt, E., Ellerby, L., Paulson, H., 2004. Caspase-mediated proteolysis of the polyglutamine disease protein ataxin-3. J. Neurochem. 89, 908–918. https://doi.org/10.1111/j.1471-4159.2004.02369.x

Bettencourt, C., Lima, M., 2011. Machado-Joseph Disease: from first descriptions to new perspectives. Orphanet J. Rare Dis. 6, 35. https://doi.org/10.1186/1750-1172-6-35

Bettencourt, C., Santos, C., Kay, T., Vasconcelos, J., Lima, M., 2008. Analysis of segregation patterns in Machado–Joseph disease pedigrees. J. Hum. Genet. 53, 920–923. https://doi.org/10.1007/s10038-008-0330-y

Bevivino, A., Loll, P., 2001. An expanded glutamine repeat destabilizes native ataxin-3 structure and mediates formation of parallel -fibrils. Proc. Natl. Acad. Sci. 98, 11955–11960. https://doi.org/10.1073/pnas.211305198

Bichelmeier, U., Schmidt, T., Hübener, J., Boy, J., Rüttiger, L., Häbig, K., Poths, S., Bonin, M., Knipper, M., Schmidt, W., Wilbertz, J., Wolburg, H., Laccone, F., Riess, O., 2007. Nuclear localization of ataxin-3 is required for the manifestation of symptoms in SCA3: in vivo evidence. J. Neurosci. 27, 7418–28. https://doi.org/10.1523/JNEUROSCI.4540-06.2007

Birling, M., Herault, Y., Pavlovic, G., 2017. Modeling human disease in rodents by CRISPR/Cas9 genome editing. Mamm. Genome 28, 291–301. https://doi.org/10.1007/s00335-017-9703-x

Boy, J., Schmidt, T., Schumann, U., Grasshoff, U., Unser, S., Holzmann, C., Schmitt, I., Karl, T., Laccone, F., Wolburg, H., Ibrahim, S., Riess, O., 2010. A transgenic mouse model of spinocerebellar ataxia type 3 resembling late disease onset and gender-specific instability of CAG repeats. Neurobiol. Dis. 37, 284–293. https://doi.org/10.1016/j.nbd.2009.08.002

Boy, J., Schmidt, T., Wolburg, H., Mack, A., Nuber, S., Böttcher, M., Schmitt, I., Holzmann, C., Zimmermann, F., Servadio, A., Riess, O., 2009. Reversibility of symptoms in a conditional mouse model of spinocerebellar ataxia type 3. Hum. Mol. Genet. 18, 4282–95. https://doi.org/10.1093/hmg/ddp381

Breuer, P., Haacke, A., Evert, B., Wüllner, U., 2010. Nuclear Aggregation of Polyglutamine-expanded Ataxin-3. J. Biol. Chem. 285, 6532–6537. https://doi.org/10.1074/jbc.M109.036335

Burnett, B., Fusheng, L., Randall, P., 2003. The polyglutamine neurodegenerative protein ataxin-3 binds polyubiquitylated proteins and has ubiquitin protease activity. Hum. Mol. Genet. 12, 3195–3205. https://doi.org/10.1093/hmg/ddg344

Burnett, B., Pittman, R., 2005. The polyglutamine neurodegenerative protein ataxin 3 regulates aggresome formation. Proc. Natl. Acad. Sci. U. S. A. 102, 4330–5. https://doi.org/10.1073/pnas.0407252102

Carter, R.J., Morton, J., Dunnett, S.B., 2001. Motor Coordination and Balance in Rodents. Curr. Protoc. Neurosci. 15, 1–14. https://doi.org/10.1002/0471142301.ns0812s15

Cecchin, C.R., Pires, A.P., Rieder, C.R., Monte, T.L., Silveira, I., Carvalho, T., Saraiva-Pereira, M.L., Sequeiros, J., Jardim, L.B., 2007. Depressive Symptoms in Machado-Joseph Disease (SCA3) Patients and Their Relatives. Community Genet. 10, 19–26. https://doi.org/10.1159/000096276

Cemal, K., Carroll, C., Lawrence, L., Lowrie, M., Ruddle, P., Al-Mahdawi, S., King, R., Pook, M., Huxley, C., Chamberlain, S., 2002. YAC transgenic mice carrying pathological alleles of the MJD1 locus exhibit a mild and slowly progressive cerebellar deficit. Hum. Mol. Genet. 11, 1075–94. https://doi.org/10.1093/hmg/11.9.1075

Cemal, K., Huxley, C., Chamberlain, S., 1999. Insertion of expanded CAG trinucleotide repeat motifs into a yeast artificial chromosome containing the human Machado–Joseph disease gene. Gene 236, 53–61. https://doi.org/10.1016/S0378-1119(99)00273-5

Charpentier, M., Khedher, A.H.Y., Menoret, S., Brion, A., Lamribet, K., Dardillac, E., Boix, C., Perrouault, L., Tesson, L., Geny, S., De Cian, A., Itier, J.M., Anegon, I., Lopez, B., Giovannangeli, C., Concordet, J.P., 2018. CtIP fusion to Cas9 enhances transgene integration by homology-dependent repair. Nat. Commun. 9, 1133. https://doi.org/10.1038/s41467-018-03475-7

Chavez, A., Scheiman, J., Vora, S., Pruitt, B., Tuttle, M., P R Iyer, E., Lin, S., Kiani, S., Guzman, C., Wiegand, D., Ter-Ovanesyan, D., Braff, J., Davidsohn, N., Housden, B., Perrimon, N., Weiss, R., Aach, J., Collins, J., Church, G., 2015. Highly efficient Cas9-mediated transcriptional programming. Nat. Methods 12, 326–8. https://doi.org/10.1038/nmeth.3312

Chen, B., Gilbert, L., Cimini, B., Schnitzbauer, J., Zhang, W., Li, G., Park, J., Blackburn, E., Weissman, J., Qi, L., Huang, B., 2013. Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. Cell 155, 1479–91. https://doi.org/10.1016/j.cell.2013.12.001

Chou, A., Yeh, T., Ouyang, P., Chen, Y., Chen, S., Wang, H., 2008. Polyglutamine-expanded ataxin-3 causes cerebellar dysfunction of SCA3 transgenic mice by inducing transcriptional dysregulation. Neurobiol. Dis. 31, 89–101. https://doi.org/10.1016/j.nbd.2008.03.011

Christian, M., Cermak, T., Doyle, E., Schmidt, C., Zhang, F., Hummel, A., Bogdanove, A.J., Voytas, D.F., 2010. Targeting DNA double-strand breaks with TAL effector nucleases. Genetics 186, 757–61. https://doi.org/10.1534/genetics.110.120717

Christie, N.T.M., Lee, A.L., Fay, H.G., Gray, A.A., Kikis, E.A., 2014. Novel Polyglutamine Model Uncouples Proteotoxicity from Aging. PLoS One 9, e96835. https://doi.org/10.1371/journal.pone.0096835

Chu, V.T., Weber, T., Wefers, B., Wurst, W., Sander, S., Rajewsky, K., Kühn, R., 2015. Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. Nat. Biotechnol. 33, 543–548. https://doi.org/10.1038/nbt.3198

Colomer, V., 2012. Mouse Models of Spinocerebellar Ataxia Type 3 (Machado-Joseph Disease). Neurotherapeutics 9, 285–296. https://doi.org/10.1007/s13311-012-0117-x

Cong, L., Ran, F., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P., Wu, X., Jiang, W., Marraffini, L., Zhang, F., 2013. Multiplex genome engineering using CRISPR/Cas systems. Science (80-.). 339, 819–23. https://doi.org/10.1126/science.1231143

Costa, M., Paulson, H., 2012. Toward understanding Machado–Joseph disease. Prog. Neurobiol. 97, 239–257. https://doi.org/10.1016/j.pneurobio.2011.11.006

Coutinho, P., Andrade, C., 1978. Autosomal dominant system degeneration in Portuguese families of the Azores Islands: A new genetic disorder involving cerebellar, pyramidal, extrapyramidal and spinal cord motor functions. Neurology 28, 703–703. https://doi.org/10.1212/WNL.28.7.703

Coutinho, P., Ruano, L., Loureiro, J., Cruz, V., Barros, J., Tuna, A., Barbot, C., Guimarães, J., Alonso, I., Silveira, I., Sequeiros, J., Neves, J., Serrano, P., Silva, M., 2013. Hereditary Ataxia and Spastic Paraplegia in Portugal. JAMA Neurol. 70, 746. https://doi.org/10.1001/jamaneurol.2013.1707

Cunha-Santos, J., Duarte-Neves, J., Carmona, V., Guarente, L., Pereira de Almeida, L., Cavadas, C., 2016. Caloric restriction blocks neuropathology and motor deficits in Machado–Joseph disease mouse models through SIRT1 pathway. Nat. Commun. 7, 11445. https://doi.org/10.1038/ncomms11445

Dabrowska, M., Juzwa, W., Krzyzosiak, W.J., Olejniczak, M., 2018. Precise Excision of the CAG Tract from the Huntingtin Gene by Cas9 Nickases. Front. Neurosci. 12, 1–8. https://doi.org/10.3389/fnins.2018.00075

Davidson, M.W., 2010. Robert Hooke. Oxford Acad. -Laboratory Med. 41, 180–182. https://doi.org/10.1309/LMQ8H3HQHZQKECZZ

Davis, K., Pattanayak, V., Thompson, D., Zuris, J., Liu, D., 2015. Small molecule-triggered Cas9 protein with improved genome-editing specificity. Nat. Chem. Biol. 11, 316–8. https://doi.org/10.1038/nchembio.1793

do Carmo Costa, M., Bajanca, F., Rodrigues, A., Tomé, R., Corthals, G., Macedo-Ribeiro, S., Paulson, H., Logarinho, E., Maciel, P., 2010. Ataxin-3 plays a role in mouse myogenic differentiation through regulation of integrin subunit levels. PLoS One 5, e11728. https://doi.org/10.1371/journal.pone.0011728

Dow, L., Lowe, S., 2012. Life in the Fast Lane: Mammalian Disease Models in the Genomics Era. Cell 148, 1099–1109. https://doi.org/10.1016/j.cell.2012.02.023

Earle, W.R., Schilling, E.L., Stark, T.H., Straus, N.P., Brown, M.F., Shelton, E., 1943. The mouse fibroblast cultures and changes seen in the living cells. J. Natl. Cancer Inst. 4, 165–212. https://doi.org/10.1093/jnci/4.2.165

Evers, M.M., Tran, H.-D., Zalachoras, I., Pepers, B.A., Meijer, O.C., den Dunnen, J.T., van Ommen, G.-J.B., Aartsma-Rus, A., van Roon-Mom, W.M.C., 2013. Ataxin-3 protein modification as a treatment strategy for spinocerebellar ataxia type 3: Removal of the CAG containing exon. Neurobiol. Dis. 58, 49–56. https://doi.org/10.1016/j.nbd.2013.04.019

Evert, B., Araujo, J., Vieira-Saecker, A., de Vos, R., Harendza, S., Klockgether, T., Wullner, U., 2006. Ataxin-3 Represses Transcription via Chromatin Binding, Interaction with Histone Deacetylase 3, and Histone Deacetylation. J. Neurosci. 26, 11474–11486. https://doi.org/10.1523/JNEUROSCI.2053-06.2006

Figiel, M., Szlachcic, W., Switonski, P., Gabka, A., Krzyzosiak, W., 2012. Mouse Models of Polyglutamine Diseases: Review and Data Table. Part I. Mol. Neurobiol. 46, 393–429. https://doi.org/10.1007/s12035-012-8315-4

Flower, M., Lomeikaite, V., Ciosi, M., Cumming, S., Morales, F., Lo, K., Hensman Moss, D., Jones, L., Holmans, P., Monckton, D.G., Tabrizi, S.J., Kraus, P., Hoffman, R., Tobin, A., Borowsky, B., Keenan, S., Whitlock, K.B., Queller, S., Campbell, C., Wang, C., Langbehn, D., Axelson, E., Johnson, H., Acharya, T., Cash, D.M., Frost, C., Jones, R., Jurgens, C., ‘t Hart, E.P., van der Grond, J., Witjes- Ane, M.-N.N., Roos, R.A.C., Dumas, E.M., van den Bogaard, S.J.A., Stopford, C., Craufurd, D., Callaghan, J., Arran, N., Rosas, D.D., Lee, S., Monaco, W., O’Regan, A., Milchman, C., Frajman, E., Labuschagne, I., Stout, J., Campbell, M., Andrews, S.C., Bechtel, N., Reilmann, R., Bohlen, S., Kennard, C., Berna, C., Hicks, S., Durr, A., Pourchot, C., Bardinet, E., Nigaud, K., Valabre, R., Gue, `, Lehericy, S., Marelli, C., Jauffret, C., Justo, D., Leavitt, B., Decolongon, J., Sturrock, A., Coleman, A., Santos, R.D., Patel, A., Gibbard, C., Whitehead, D., Wild, E., Owen, G., Crawford, H., Malone, I., Lahiri, N., Fox, N.C., Hobbs, N.Z., Scahill, R.I., Ordidge, R., Pepple, T., Read, J., Say, M.J., Landwehrmeyer, B., Daidj, F., Bassez, G., Lignier, B., Couppey, F., Delmas, S., Deux, J.-F., Hankiewicz, K., Dogan, C., Minier, L., Chevalier, P., Hamadouche, A., Catt, M., van Hees, V., Catt, S., Schwalber, A., Dittrich, J., Kierkegaard, M., Wenninger, S., Schoser, B., Schüller, A., Stahl, K., Künzel, H., Wolff, M., Jellinek, A., Moreno, C.J., Gorman, G., Lochmüller, H., Trenell, M., van Laar, S., Wood, L., Cassidy, S., Newman, J., Charman, S., Steffaneti, R., Taylor, L., Brownrigg, A., Day, S., Atalaia, A., Raaphorst, J., Okkersen, K., van Engelen, B., Nikolaus, S., Cornelissen, Y., van Nimwegen, M., Maas, D., Klerks, E., Bouman, S., Knoop, H., Heskamp, L., Heerschap, A., Rahmadi, R., Groot, P., Heskes, T., Kapusta, K., Glennon, J., Abghari, S., Aschrafi, A., Poelmans, G., Treweek, S., Hogarth, F., Littleford, R., Donnan, P., Hapca, A., Hannah, M., McKenzie, E., Rauchhaus, P., Cumming, S.A., Monckton, D.G., Adam, B., Faber, C., Merkies, I., 2019. MSH3 modifies somatic instability and disease severity in Huntington’s and myotonic dystrophy type 1. Brain 142, 1876–1886. https://doi.org/10.1093/brain/awz115

Gaj, T., Gersbach, C., Barbas, C., 2013. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. Trends Biotechnol. 31, 397–405. https://doi.org/10.1016/j.tibtech.2013.04.004

Gan, S.-R., Ni, W., Dong, Y., Wang, N., Wu, Z.-Y., 2015. Population Genetics and New Insight into Range of CAG Repeats of Spinocerebellar Ataxia Type 3 in the Han Chinese Population. PLoS One 10, 1–12. https://doi.org/10.1371/journal.pone.0134405

Gasiunas, G., Barrangou, R., Horvath, P., Siksnys, V., 2012. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proc. Natl. Acad. Sci. U. S. A. 109, E2579-86. https://doi.org/10.1073/pnas.1208507109

Gaudelli, N., Komor, A., Rees, H., Packer, M., Badran, A., Bryson, D., Liu, D., 2017. Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. Nature 551, 464–471. https://doi.org/10.1038/nature24644

Gilbert, L., Larson, M., Morsut, L., Liu, Z., Brar, G., Torres, S., Stern-Ginossar, N., Brandman, O., Whitehead, E., Doudna, J., Lim, W., Weissman, J., Qi, L., 2013. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell 154, 442–51. https://doi.org/10.1016/j.cell.2013.06.044

Gopalappa, R., Suresh, B., Ramakrishna, S., Kim, H. (Henry), 2018. Paired D10A Cas9 nickases are sometimes more efficient than individual nucleases for gene disruption. Nucleic Acids Res. 46, 1–12. https://doi.org/10.1093/nar/gky222

Goti, D., Katzen, S., Mez, J., Kurtis, N., Kiluk, J., Ben-Haïem, L., Jenkins, N., Copeland, N., Kakizuka, A., Sharp, A., Ross, C., Mouton, P., Colomer, V., 2004. A mutant ataxin-3 putative-cleavage fragment in brains of Machado-Joseph disease patients and transgenic mice is cytotoxic above a critical concentration. J. Neurosci. 24, 10266–79. https://doi.org/10.1523/JNEUROSCI.2734-04.2004

Gould, F., Goti, D., Pearce, D., Gonzalez, G., Gao, H., Leon, M., Jenkins, N., Copeland, N., Ross, C., Brown, R., 2007. A Mutant ataxin-3 fragment results from processing at a site N-terminal to amino acid 190 in brain of Machado–Joseph disease-like transgenic mice. Neurobiol. Dis. 27, 362–369. https://doi.org/10.1016/j.nbd.2007.06.005

Gould, V., Goti, D., Pearce, D., Gonzalez, G.A., Gao, H., Bermudez de Leon, M., Jenkins, N., Copeland, N., Ross, C., Brown, D., 2007. A mutant ataxin-3 fragment results from processing at a site N-terminal to amino acid 190 in brain of Machado-Joseph disease-like transgenic mice. Neurobiol. Dis. 27, 362–9. https://doi.org/10.1016/j.nbd.2007.06.005

Groen, E.J.N., Talbot, K., Gillingwater, T.H., 2018. Advances in therapy for spinal muscular atrophy: promises and challenges. Nat. Rev. Neurol. 14, 214–224. https://doi.org/10.1038/nrneurol.2018.4

Guo, Q., Mintier, G., Ma-Edmonds, M., Storton, D., Wang, X., Xiao, X., Kienzle, B., Zhao, D., Feder, J.N., 2018. 'Cold shock' increases the frequency of homology directed repair gene editing in induced pluripotent stem cells. Sci. Rep. 8, 2080.

https://doi.org/10.1038/s41598-018-20358-5

Guyon, J., Steffen, L., Howell, M., Pusack, T., Lawrence, C., Kunkel, L., 2007. Modeling human muscle disease in zebrafish. Biochim. Biophys. Acta - Mol. Basis Dis. 1772, 205–215. https://doi.org/10.1016/j.bbadis.2006.07.003

Hacein-Bey-Abina, S., Garrigue, A., Wang, G.P., Soulier, J., Lim, A., Morillon, E., Clappier, E., Caccavelli, L., Delabesse, E., Beldjord, K., Asnafi, V., MacIntyre, E., Dal Cortivo, L., Radford, I., Brousse, N., Sigaux, F., Moshous, D., Hauer, J., Borkhardt, A., Belohradsky, B.H., Wintergerst, U., Velez, M.C., Leiva, L., Sorensen, R., Wulffraat, N., Blanche, S., Bushman, F.D., Fischer, A., Cavazzana-Calvo, M., 2008. Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. J. Clin. Invest. 118, 3132–3142. https://doi.org/10.1172/JCI35700

Hemphill, J., Borchardt, E., Brown, K., Asokan, A., Deiters, A., 2015. Optical Control of CRISPR/Cas9 Gene Editing. J. Am. Chem. Soc. 137, 5642–5. https://doi.org/10.1021/ja512664v

Henssen, A.G., Henaff, E., Jiang, E., Eisenberg, A.R., Carson, J.R., Villasante, C.M., Ray, M., Still, E., Burns, M., Gandara, J., Feschotte, C., Mason, C.E., Kentsis, A., 2015. Genomic DNA transposition induced by human PGBD5. Elife 4, 1–20. https://doi.org/10.7554/eLife.10565

Heyer, W.-D., Ehmsen, K.T., Liu, J., 2010. Regulation of Homologous Recombination in Eukaryotes. Annu. Rev. Genet. 44, 113–139. https://doi.org/10.1146/annurev-genet-051710-150955

Hilton, I., D'Ippolito, A., Vockley, C., Thakore, P., Crawford, G., Reddy, T., Gersbach, C., 2015. Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. Nat. Biotechnol. 33, 510–7. https://doi.org/10.1038/nbt.3199

Hoy, S.M., 2018. Patisiran: First Global Approval. Drugs 78, 1625–1631. https://doi.org/10.1007/s40265-018-0983-6

Hsu, P., Lander, E., Zhang, F., 2014. Development and applications of CRISPR-Cas9 for genome engineering. Cell 157, 1262–78. https://doi.org/10.1016/j.cell.2014.05.010

Hübener, J., Vauti, F., Funke, C., Wolburg, H., Ye, Y., Schmidt, T., Wolburg-Buchholz, K., Schmitt, I., Gardyan, A., Driessen, S., Arnold, H.-H., Nguyen, H.P., Riess, O., 2011. N-terminal ataxin-3 causes neurological symptoms with inclusions, endoplasmic reticulum stress and ribosomal dislocation. Brain 134, 1925–42. https://doi.org/10.1093/brain/awr118

Hubener, J., Weber, J., Richter, C., Honold, L., Weiss, A., Murad, F., Breuer, P., Wullner, U., Bellstedt, P., Paquet-Durand, F., Takano, J., Saido, T., Riess, O., Nguyen, H., 2013. Calpain-mediated ataxin-3 cleavage in the molecular pathogenesis of spinocerebellar ataxia type 3 (SCA3). Hum. Mol. Genet. 22, 508–518. https://doi.org/10.1093/hmg/dds449

Ichikawa, Y., Goto, J., Hattori, M., Toyoda, A., Ishii, K., Jeong, S.-Y., Hashida, H., Masuda, N., Ogata, K., Kasai, F., Hirai, M., Maciel, P., Rouleau, G.A., Sakaki, Y., Kanazawa, I.,

2001. The genomic structure and expression of MJD , the Machado-Joseph disease gene. J. Hum. Genet. 46, 413–422. https://doi.org/10.1007/s100380170060

Ikeda, H., Yamaguchi, M., Sugai, S., Aze, Y., Narumiya, S., Kakizuka, A., 1996. Expanded polyglutamine in the Machado-Joseph disease protein induces cell death in vitro and in vivo. Nat. Genet. 13, 196–202. https://doi.org/10.1038/ng0696-196

Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., Nakata, A., 1987. Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. J. Bacteriol. 169, 5429–33.

Jansen, R., Embden, J., Gaastra, W., Schouls, L., 2002. Identification of genes that are associated with DNA repeats in prokaryotes. Mol. Microbiol. 43, 1565–75.

Jardim, L., Pereira, M., Silveira, I., Ferro, A., Sequeiros, J., Giugliani, R., 2001. Neurologic Findings in Machado-Joseph Disease. Arch. Neurol. 58, 899. https://doi.org/10.1001/archneur.58.6.899

Jedrzejczak-Silicka, M., 2017. History of Cell Culture, in: New Insights into Cell Culture Technology. InTech, pp. 1–43. https://doi.org/10.5772/66905

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J., Charpentier, E., 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science (80-.). 337, 816–21. https://doi.org/10.1126/science.1225829

Johnson, T., 2003. Advantages and disadvantages of Caenorhabditis elegans for aging research. Exp. Gerontol. 38, 1329–32.

Jung, J., Xu, K., Lessing, D., Bonini, N., 2009. Preventing Ataxin-3 protein cleavage mitigates degeneration in a Drosophila model of SCA3. Hum. Mol. Genet. 18, 4843–4852. https://doi.org/10.1093/hmg/ddp456

Justice, M., Dhillon, P., 2016. Using the mouse to model human disease: increasing validity and reproducibility. Dis. Model. Mech. 9, 101–103. https://doi.org/10.1242/dmm.024547

Kajiwara, M., Aoi, T., Okita, K., Takahashi, R., Inoue, H., Takayama, N., Endo, H., Eto, K., Toguchida, J., Uemoto, S., Yamanaka, S., 2012. Donor-dependent variations in hepatic differentiation from human-induced pluripotent stem cells. Proc. Natl. Acad. Sci. 109, 12538–12543. https://doi.org/10.1073/pnas.1209979109

Kaur, G., Dufour, J.M., 2012. Cell lines. Spermatogenesis 2, 1–5. https://doi.org/10.4161/spmg.19885

Kawaguchi, Y., Okamoto, T., Taniwaki, M., Aizawa, M., Inoue, M., Katayama, S., Kawakami, H., Nakamura, S., Nishimura, M., Akiguchi, I., Kimura, J., Narumiya, S., Kakizuka, A., 1994. CAG expansions in a novel gene for Machado-Joseph disease at chromosome 14q32.1. Nat. Genet. 8, 221–228. https://doi.org/10.1038/ng1194-221

Kearns, N., Pham, H., Tabak, B., Genga, R., Silverstein, N., Garber, M., Maehr, R., 2015. Functional annotation of native enhancers with a Cas9-histone demethylase fusion. Nat. Methods 12, 401–403. https://doi.org/10.1038/nmeth.3325

Khan, L., Bauer, P., Miyazaki, H., Lindenberg, K., Landwehrmeyer, B., Nukina, N., 2006. Expanded polyglutamines impair synaptic transmission and ubiquitin-proteasome system in Caenorhabditis elegans. J. Neurochem. 98, 576–87. https://doi.org/10.1111/j.1471-4159.2006.03895.x

Kim, H., Kim, J.-S., 2014. A guide to genome engineering with programmable nucleases. Nat. Rev. Genet. 15, 321–334. https://doi.org/10.1038/nrg3686

Kim, Y., Cha, J., Chandrasegaran, S., 1996. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. Proc. Natl. Acad. Sci. U. S. A. 93, 1156–60.

Klockgether, T., Kramer, B., Lüdtke, R., Schöls, L., Laccone, F., 1996. Repeat length and disease progression in spinocerebellar ataxia type 3. Lancet 348, 830. https://doi.org/10.1016/S0140-6736(05)65255-5

Koeppen, A.H., 2018. The Neuropathology of Spinocerebellar Ataxia Type 3/Machado-Joseph Disease, in: Polyglutamine Disorders. pp. 233–241. https://doi.org/10.1007/978-3-319-71779-1_11

Kretzschmar, D., Tschäpe, J., Bettencourt Da Cruz, A., Asan, E., Poeck, B., Strauss, R., Pflugfelder, G., 2005. Glial and neuronal expression of polyglutamine proteins induce behavioral changes and aggregate formation in Drosophila. Glia 49, 59–72. https://doi.org/10.1002/glia.20098

Kumar, T., Larson, M., Wang, H., McDermott, J., Bronshteyn, I., 2009. Transgenic mouse technology: principles and methods. Methods Mol. Biol. 590, 335–62. https://doi.org/10.1007/978-1-60327-378-7_22

Li, F., Macfarlan, T., Pittman, R., Chakravarti, D., 2002. Ataxin-3 Is a Histone-binding Protein with Two Independent Transcriptional Corepressor Activities. J. Biol. Chem. 277, 45004–45012. https://doi.org/10.1074/jbc.M205259200

Li, L.-B., Yu, Z., Teng, X., Bonini, N.M., 2008. RNA toxicity is a component of ataxin-3 degeneration in Drosophila. Nature 453, 1107–1111. https://doi.org/10.1038/nature06909

Liang, F., Han, M., Romanienko, P., Jasin, M., 1998. Homology-directed repair is a major double-strand break repair pathway in mammalian cells. Proc. Natl. Acad. Sci. U. S. A. 95, 5172–7.

Lin, S., Staahl, B.T., Alla, R.K., Doudna, J.A., 2014. Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. Elife 3, 1–13. https://doi.org/10.7554/eLife.04766

Liu, H., Li, X., Ning, G., Zhu, S., Ma, X., Liu, X., Liu, C., Huang, M., Schmitt, I., Wüllner, U., Niu, Y., Guo, C., Wang, Q., Tang, T.-S., 2016. The Machado–Joseph Disease Deubiquitinase Ataxin-3 Regulates the Stability and Apoptotic Function of p53. PLOS Biol. 14, 1–31. https://doi.org/10.1371/journal.pbio.2000733

Liu, M., Rehman, S., Tang, X., Gu, K., Fan, Q., Chen, D., Ma, W., 2019. Methodologies for Improving HDR Efficiency. Front. Genet. 9, 1–9. https://doi.org/10.3389/fgene.2018.00691

Longenecker, G., Kulkarni, A., 2009. Generation of gene knockout mice by ES cell microinjection. Curr. Protoc. cell Biol. 19, 1–36. https://doi.org/10.1002/0471143030.cb1914s44

Lu, B., Vogel, H., 2009. Drosophila Models of Neurodegenerative Diseases. Annu. Rev. Pathol. Mech. Dis. 4, 315–342. https://doi.org/10.1146/annurev.pathol.3.121806.151529

Ma, H., Marti-Gutierrez, N., Park, S., Wu, J., Lee, Y., Suzuki, K., Koski, A., Ji, D., Hayama, T., Ahmed, R., Darby, H., Van Dyken, C., Li, Y., Kang, E., Park, A., Kim, D., Kim, S., Gong, J., Gu, Y., Xu, X., Battaglia, D., Krieg, S., Lee, D., Wu, D., Wolf, D., Heitner, S., Belmonte, J., Amato, P., Kim, J., Kaul, S., Mitalipov, S., 2017. Correction of a pathogenic gene mutation in human embryos. Nature 548, 413–419. https://doi.org/10.1038/nature23305

MacKenzie, A., 2006. Production of Yeast Artificial Chromosome Transgenic Mice by Pronuclear Injection of One-Cell Embryos, in: YAC Protocols. Humana Press, New Jersey, pp. 139–150. https://doi.org/10.1385/1-59745-158-4:139

Maji, B., Moore, C., Zetsche, B., Volz, S., Zhang, F., Shoulders, M., Choudhary, A., 2017. Multidimensional chemical control of CRISPR-Cas9. Nat. Chem. Biol. 13, 9–11. https://doi.org/10.1038/nchembio.2224

Malankhanova, T.B., Malakhova, A.A., Medvedev, S.P., Zakian, S.M., 2017. Modern Genome Editing Technologies in Huntington’s Disease Research. J. Huntingtons. Dis. 6, 19–31. https://doi.org/10.3233/JHD-160222

Mali, P., Yang, L., Esvelt, K., Aach, J., Guell, M., DiCarlo, J., Norville, J., Church, G., 2013. RNA-guided human genome engineering via Cas9. Science (80-.). 339, 823–6. https://doi.org/10.1126/science.1232033

Mao, Y., Senic-Matuglia, F., Di Fiore, P., Polo, S., Hodsdon, M., De Camilli, P., 2005. Deubiquitinating function of ataxin-3: Insights from the solution structure of the Josephin domain. Proc. Natl. Acad. Sci. 102, 12700–12705. https://doi.org/10.1073/pnas.0506344102

Martí, E., 2016. RNA toxicity induced by expanded CAG repeats in Huntington’s disease. Brain Pathol. 26, 779–786. https://doi.org/10.1111/bpa.12427

Masters, J.R., 2002. HeLa cells 50 years on: the good, the bad and the ugly. Nat. Rev. Cancer 2, 315–319. https://doi.org/10.1038/nrc775

Matos, C., Macedo-Ribeiro, S., Carvalho, A., 2011. Polyglutamine diseases: The special case of ataxin-3 and Machado–Joseph disease. Prog. Neurobiol. 95, 26–48. https://doi.org/10.1016/j.pneurobio.2011.06.007

Matos, C., Nóbrega, C., Louros, S., Almeida, B., Ferreiro, E., Valero, J., Pereira de Almeida, L., Macedo-Ribeiro, S., Carvalho, A., 2016. Ataxin-3 phosphorylation decreases neuronal defects in spinocerebellar ataxia type 3 models. J. Cell Biol. 212, 465–480. https://doi.org/10.1083/jcb.201506025

Matos, C., Pereira de Almeida, L., Nobrega, C., 2017. Proteolytic Cleavage of Polyglutamine Disease-Causing Proteins: Revisiting the Toxic Fragment Hypothesis. Curr. Pharm. Des. 23, 753–775.

https://doi.org/10.2174/1381612822666161227121912

McLoughlin, H.S., Moore, L.R., Chopra, R., Komlo, R., McKenzie, M., Blumenstein, K.G., Zhao, H., Kordasiewicz, H.B., Shakkottai, V.G., Paulson, H.L., 2018. Oligonucleotide therapy mitigates disease in spinocerebellar ataxia type 3 mice. Ann. Neurol. 84, 64–77. https://doi.org/10.1002/ana.25264

Mendonça, L., Nóbrega, C., Hirai, H., Kaspar, B., Pereira de Almeida, L., 2015. Transplantation of cerebellar neural stem cells improves motor coordination and neuropathology in Machado-Joseph disease mice. Brain 138, 320–335. https://doi.org/10.1093/brain/awu352

Mittelman, D., Wilson, J.H., 2013. The fractured genome of HeLa cells. Genome Biol. 14, 111. https://doi.org/10.1186/gb-2013-14-4-111

Mojica, F., Díez-Villaseñor, C., García-Martínez, J., Soria, E., 2005. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. J. Mol. Evol. 60, 174–82. https://doi.org/10.1007/s00239-004-0046-3

Mojica, F., Juez, G., Rodríguez-Valera, F., 1993. Transcription at different salinities of Haloferax mediterranei sequences adjacent to partially modified PstI sites. Mol. Microbiol. 9, 613–21.

Mok, H.P., Lever, A., 2006. A method to estimate the efficiency of gene expression from an integrated retroviral vector. Retrovirology 3, 51. https://doi.org/10.1186/1742-4690-3-51

Montcel, S., Durr, A., Bauer, P., Figueroa, K., Ichikawa, Y., Brussino, A., Forlani, S., Rakowicz, M., Schöls, L., Mariotti, C., van de Warrenburg, C., Orsi, L., Giunti, P., Filla, A., Szymanski, S., Klockgether, T., Berciano, J., Pandolfo, M., Boesch, S., Melegh, B., Timmann, D., Mandich, P., Camuzat, A., Goto, J., Ashizawa, T., Cazeneuve, C., Tsuji, S., Pulst, S., Brusco, A., Riess, O., Brice, A., Stevanin, G., 2014. Modulation of the age at onset in spinocerebellar ataxia by CAG tracts in various genes. Brain 137, 2444–2455. https://doi.org/10.1093/brain/awu174

Mullane, K., Williams, M., 2014. Animal models of asthma: Reprise or reboot? Biochem. Pharmacol. 87, 131–139. https://doi.org/10.1016/j.bcp.2013.06.026

Nakano, K., Dawson, M., Spence, A., 1972. Machado disease: A hereditary ataxia in Portuguese emigrants to Massachusetts. Neurology 22, 49–49. https://doi.org/10.1212/WNL.22.1.49

Nalavade, R., Griesche, N., Ryan, D., Hildebrand, S., Krauß, S., 2013. Mechanisms of RNA-induced toxicity in CAG repeat disorders. Cell Death Dis. 4, e752–e752. https://doi.org/10.1038/cddis.2013.276

Nandagopal, R., 2004. Dramatic levodopa responsiveness of dystonia in a sporadic case of spinocerebellar ataxia type 3. Postgrad. Med. J. 80, 363–365. https://doi.org/10.1136/pgmj.2003.015297

Nihongaki, Y., Kawano, F., Nakajima, T., Sato, M., 2015. Photoactivatable CRISPR-Cas9 for optogenetic genome editing. Nat. Biotechnol. 33, 755–60. https://doi.org/10.1038/nbt.3245

Nishida, K., Arazoe, T., Yachie, N., Banno, S., Kakimoto, M., Tabata, M., Mochizuki, M., Miyabe, A., Araki, M., Hara, K., Shimatani, Z., Kondo, A., 2016. Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. Science (80-.). 353. https://doi.org/10.1126/science.aaf8729

Nóbrega, C., Nascimento-Ferreira, I., Onofre, I., Albuquerque, D., Conceição, M., Déglon, N., De Almeida, L.P., 2013a. Overexpression of mutant ataxin-3 in mouse cerebellum induces ataxia and cerebellar neuropathology. Cerebellum 12, 441–455. https://doi.org/10.1007/s12311-012-0432-0

Nóbrega, C., Nascimento-Ferreira, I., Onofre, I., Albuquerque, D., Déglon, N., Pereira de Almeida, L., 2014. RNA Interference Mitigates Motor and Neuropathological Deficits in a Cerebellar Mouse Model of Machado-Joseph Disease. PLoS One 9, 1–14. https://doi.org/10.1371/journal.pone.0100086

Nóbrega, C., Nascimento-Ferreira, I., Onofre, I., Albuquerque, D., Hirai, H., Déglon, N., de Almeida, L.P., 2013b. Silencing Mutant Ataxin-3 Rescues Motor Deficits and Neuropathology in Machado-Joseph Disease Transgenic Mice. PLoS One 8, 1–11. https://doi.org/10.1371/journal.pone.0052396

Nóbrega, C., Simões, A., Duarte-Neves, J., Duarte, S., Vasconcelos-Ferreira, A., Cunha-Santos, J., Pereira, D., Santana, M., Cavadas, C., de Almeida, L., 2018. Molecular Mechanisms and Cellular Pathways Implicated in Machado-Joseph Disease Pathogenesis, in: Polyglutamine Disorders. pp. 349–367. https://doi.org/10.1007/978-3-319-71779-1_18

Onodera, O., Idezuka, J., Igarashi, S., Takiyama, Y., Endo, K., Takano, H., Oyake, M., Tanaka, H., Inuzuka, T., Hayashi, T., Yuasa, T., Ito, J., Miyatake, T., Tsuji, S., 1998. Progressive atrophy of cerebellum and brainstem as a function of age and the size of the expanded CAG repeats in theMJD1 gene in Machado-Joseph disease. Ann. Neurol. 43, 288–296. https://doi.org/10.1002/ana.410430305

Orr, H., Zoghbi, H., 2007. Trinucleotide repeat disorders. Annu. Rev. Neurosci. 30, 575–621. https://doi.org/10.1146/annurev.neuro.29.051605.113042

Ouyang, S., Xie, Yingjun, Xiong, Z., Yang, Y., Xian, Y., Ou, Z., Song, B., Chen, Y., Xie, Yuhuan, Li, H., Sun, X., 2018. CRISPR/Cas9-Targeted Deletion of Polyglutamine in Spinocerebellar Ataxia Type 3-Derived Induced Pluripotent Stem Cells. Stem Cells Dev. 27, 756–770. https://doi.org/10.1089/scd.2017.0209

Paulson, H., Perez, M., Trottier, Y., Trojanowski, J., Subramony, S., Das, S., Vig, P., Mandel, J., Fischbeck, K., Pittman, R., 1997. Intranuclear Inclusions of Expanded Polyglutamine Protein in Spinocerebellar Ataxia Type 3. Neuron 19, 333–344. https://doi.org/10.1016/S0896-6273(00)80943-5

Pena, F., 2018. CRISPR-mediated pre-transcriptional silencing of ATXN3, the gene involved in Machado-Joseph disease. In: *Departamento de Ciência da Vida, Faculdade de Ciências e Tecnologia*. Universidade de Coimbra

Pinder, J., Salsman, J., Dellaire, G., 2015. Nuclear domain ‘knock-in’ screen for the evaluation and identification of small molecule enhancers of CRISPR-based genome editing. Nucleic Acids Res. 43, 9379–9392. https://doi.org/10.1093/nar/gkv993

Platt, R.J., Chen, S., Zhou, Y., Yim, M.J., Swiech, L., Kempton, H.R., Dahlman, J.E., Parnas, O., Eisenhaure, T.M., Jovanovic, M., Graham, D.B., Jhunjhunwala, S., Heidenreich, M., Xavier, R.J., Langer, R., Anderson, D.G., Hacohen, N., Regev, A., Feng, G., Sharp, P.A., Zhang, F., 2014. CRISPR-Cas9 Knockin Mice for Genome Editing and Cancer Modeling. Cell 159, 440–455. https://doi.org/10.1016/j.cell.2014.09.014

Poksay, K., Madden, D., Peter, A., Niazi, K., Banwait, S., Crippen, D., Bredesen, D., Rao, R., 2011. Valosin-containing protein gene mutations: cellular phenotypes relevant to neurodegeneration. J. Mol. Neurosci. 44, 91–102. https://doi.org/10.1007/s12031-010-9489-8

Puchta, H., Dujon, B., Hohn, B., 1993. Homologous recombination in plant cells is enhanced by in vivo induction of double strand breaks into DNA by a site-specific endonuclease. Nucleic Acids Res. 21, 5034–40.

Puchta, H., Fauser, F., 2014. Synthetic nucleases for genome engineering in plants: prospects for a bright future. Plant J. 78, 727–41. https://doi.org/10.1111/tpj.12338

Ramani, B., Harris, G., Huang, R., Seki, T., Murphy, G., Costa, M., Fischer, S., Saunders, T., Xia, G., McEachin, R., Paulson, H., 2015. A knockin mouse model of spinocerebellar ataxia type 3 exhibits prominent aggregate pathology and aberrant splicing of the disease gene transcript. Hum. Mol. Genet. 24, 1211–24. https://doi.org/10.1093/hmg/ddu532

Ran, F.A., Hsu, P.D., Lin, C.-Y., Gootenberg, J.S., Konermann, S., Trevino, A.E., Scott, D.A., Inoue, A., Matoba, S., Zhang, Y., Zhang, F., 2013. Double Nicking by RNA-Guided CRISPR Cas9 for Enhanced Genome Editing Specificity. Cell 154, 1380–1389. https://doi.org/10.1016/j.cell.2013.08.021

Rath, D., Amlinger, L., Rath, A., Lundgren, M., 2015. The CRISPR-Cas immune system: biology, mechanisms and applications. Biochimie 117, 119–28. https://doi.org/10.1016/j.biochi.2015.03.025

Ravikumar, B., Sarkar, S., Rubinsztein, D., 2008. Clearance of Mutant Aggregate-Prone Proteins by Autophagy. pp. 195–211. https://doi.org/10.1007/978-1-59745-157-4_13

Reed, L., Baer, C., Edison, A., 2017. Considerations when choosing a genetic model organism for metabolomics studies. Curr. Opin. Chem. Biol. 36, 7–14. https://doi.org/10.1016/j.cbpa.2016.12.005

Rinaldi, C., Wood, M.J.A., 2018. Antisense oligonucleotides: the next frontier for treatment of neurological disorders. Nat. Rev. Neurol. 14, 9–21. https://doi.org/10.1038/nrneurol.2017.148

Rodríguez-Hernández, C., Torres-Garcia, S., Olvera-Sandoval, C., Ramirez-Castillo, F., Muro, A., Avelar-Gonzalez, F., Guerrero-Barrera, A., 2015. Cell Culture: History, Development and Prospects. Int. J. Curr. Res. Acad. Rev. 3, 252–263.

Rosenberg, R., 1992. Machado-Joseph disease: An autosomal dominant motor system degeneration. Mov. Disord. 7, 193–203. https://doi.org/10.1002/mds.870070302

Rosenthal, N., Brown, S., 2007. The mouse ascending: perspectives for human-disease models. Nat. Cell Biol. 9, 993–999. https://doi.org/10.1038/ncb437

Rüb, U., Brunt, E., Deller, T., 2008. New insights into the pathoanatomy of spinocerebellar ataxia type 3 (Machado–Joseph disease). Curr. Opin. Neurol. 21, 111–116. https://doi.org/10.1097/WCO.0b013e3282f7673d

Sander, J., Joung, J., 2014. CRISPR-Cas systems for editing, regulating and targeting genomes. Nat. Biotechnol. 32, 347–55. https://doi.org/10.1038/nbt.2842

Schedl, A., Beermann, F., Thies, E., Montoliu, L., Kelsey, G., Schütz, G., 1992. Transgenic mice generated by pronuclear injection of a yeast artificial chromosome. Nucleic Acids Res. 20, 3073–7.

Schmidt, J., Schmidt, T., 2018. Animal Models of Machado-Joseph Disease, in: Polyglutamine Disorders. pp. 289–308. https://doi.org/10.1007/978-3-319-71779-1_15

Seidel, K., Siswanto, S., Brunt, E., den Dunnen, W., Korf, H., Rüb, U., 2012. Brain pathology of spinocerebellar ataxias. Acta Neuropathol. 124, 1–21. https://doi.org/10.1007/s00401-012-1000-x

Sekijima, Y., Ueda, M., Koike, H., Misawa, S., Ishii, T., Ando, Y., 2018. Diagnosis and management of transthyretin familial amyloid polyneuropathy in Japan: red-flag symptom clusters and treatment algorithm. Orphanet J. Rare Dis. 13, 6. https://doi.org/10.1186/s13023-017-0726-x

Sentmanat, M., Shondra, M., 2016. Efficient CRISPR/Cas9 Delivery Using Nucleofector^{TM} Technology: Comparison of Plasmid- and RNP-based Editing. BioResearch Lonza 1, 1–14.

Shi, Y., Inoue, H., Wu, J.C., Yamanaka, S., 2017. Induced pluripotent stem cell technology: a decade of progress. Nat. Rev. Drug Discov. 16, 115–130. https://doi.org/10.1038/nrd.2016.245

Silva-Fernandes, A., Costa, M., Duarte-Silva, S., Oliveira, P., Botelho, C., Martins, L., Mariz, J., Ferreira, T., Ribeiro, F., Correia-Neves, M., Costa, C., Maciel, P., 2010. Motor uncoordination and neuropathology in a transgenic mouse model of Machado–Joseph disease lacking intranuclear inclusions and ataxin-3 cleavage products. Neurobiol. Dis. 40, 163–176. https://doi.org/10.1016/j.nbd.2010.05.021

Sittler, A., Muriel, M., Marinello, M., Brice, A., den Dunnen, W., Alves, S., 2017. Deregulation of autophagy in postmortem brains of Machado-Joseph disease patients. Neuropathology 38, 113–124. https://doi.org/10.1111/neup.12433

Smith, J., Bibikova, M., Whitby, F., Reddy, A., Chandrasegaran, S., Carroll, D., 2000. Requirements for double-strand cleavage by chimeric restriction enzymes with zinc finger DNA-recognition domains. Nucleic Acids Res. 28, 3361–9.

Stein, C.A., Castanotto, D., 2017. FDA-Approved Oligonucleotide Therapies in 2017. Mol. Ther. 25, 1069–1075. https://doi.org/10.1016/j.ymthe.2017.03.023

Stepper, P., Kungulovski, G., Jurkowska, R., Chandra, T., Krueger, F., Reinhardt, R., Reik, W., Jeltsch, A., Jurkowski, T., 2017. Efficient targeted DNA methylation with chimeric dCas9-Dnmt3a-Dnmt3L methyltransferase. Nucleic Acids Res. 45, 1703–1713. https://doi.org/10.1093/nar/gkw1112

Switonski, P., Szlachcic, W., Krzyzosiak, W., Figiel, M., 2015. A new humanized ataxin-3 knock-in mouse model combines the genetic features, pathogenesis of neurons and glia and late disease onset of SCA3/MJD. Neurobiol. Dis. 73, 174–88. https://doi.org/10.1016/j.nbd.2014.09.020

Szlachcic, W.J., Wiatr, K., Trzeciak, M., Figlerowicz, M., Figiel, M., 2017. The Generation of Mouse and Human Huntington Disease iPS Cells Suitable for In vitro Studies on Huntingtin Function. Front. Mol. Neurosci. 10, 13. https://doi.org/10.3389/fnmol.2017.00253

Takahashi, K., Yamanaka, S., 2006. Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. Cell 126, 663–676. https://doi.org/10.1016/j.cell.2006.07.024

Takahashi, T., Katada, S., Onodera, O., 2010. Polyglutamine Diseases: Where does Toxicity Come from? What is Toxicity? Where are We Going? J. Mol. Cell Biol. 2, 180–191. https://doi.org/10.1093/jmcb/mjq005

Takiyama, Y., Nishizawa, M., Tanaka, H., Kawashima, S., Sakamoto, H., Karube, Y., Shimazaki, H., Soutome, M., Endo, K., Ohta, S., Kagawa, Y., Kanazawa, I., Mizuno, Y., Yoshida, M., Yuasa, T., Horikawa, Y., Oyanagi, K., Nagai, H., Kondo, T., Inuzuka, T., Onodera, O., Tsuji, S., 1993. The gene for Machado–Joseph disease maps to human chromosome 14q. Nat. Genet. 4, 300–304. https://doi.org/10.1038/ng0793-300

Tanaka, M., Hadjantonakis, A., Vintersten, K., Nagy, A., 2009. Aggregation chimeras: combining ES cells, diploid, and tetraploid embryos. Methods Mol. Biol. 530, 287–309. https://doi.org/10.1007/978-1-59745-471-1_15

Taxman, D.J., Moore, C.B., Guthrie, E.H., Huang, M.T.-H., 2010. Short Hairpin RNA (shRNA): Design, Delivery, and Assessment of Gene Knockdown. Methods Mol. Biol. 1, 139–156. https://doi.org/10.1007/978-1-60761-657-3_10

Taylor, M.W., 2014. A History of Cell Culture. Viruses Man A Hist. Interact. 1, 41–52. https://doi.org/10.1007/978-3-319-07758-1_3

Teixeira-Castro, A., Ailion, M., Jalles, A., Brignull, H., Vilaça, J., Dias, N., Rodrigues, P., Oliveira, J., Neves-Carvalho, A., Morimoto, R., Maciel, P., 2011. Neuron-specific proteotoxicity of mutant ataxin-3 in C. elegans: rescue by the DAF-16 and HSF-1 pathways. Hum. Mol. Genet. 20, 2996–3009. https://doi.org/10.1093/hmg/ddr203

Teixeira-Castro, A., Jalles, A., Esteves, S., Kang, S., da Silva Santos, L., Silva-Fernandes, A., Neto, M.F., Brielmann, R.M., Bessa, C., Duarte-Silva, S., Miranda, A., Oliveira, S., Neves-Carvalho, A., Bessa, J., Summavielle, T., Silverman, R.B., Oliveira, P., Morimoto, R.I., Maciel, P., 2015. Serotonergic signalling suppresses ataxin 3 aggregation and neurotoxicity in animal models of Machado-Joseph disease. Brain 138, 3221–3237. https://doi.org/10.1093/brain/awv262

Tissenbaum, H., 2015. Using C. elegans for aging research. Invertebr. Reprod. Dev. 59, 59–63. https://doi.org/10.1080/07924259.2014.940470

Tomioka, I., Ishibashi, H., Minakawa, E., Motohashi, H., Takayama, O., Saito, Y., Popiel, H., Puentes, S., Owari, K., Nakatani, T., Nogami, N., Yamamoto, K., Noguchi, S., Yonekawa, T., Tanaka, Y., Fujita, N., Suzuki, H., Kikuchi, H., Aizawa, S., Nagano, S., Yamada, D., Nishino, I., Ichinohe, N., Wada, K., Kohsaka, S., Nagai, Y., Seki, K., 2017. Transgenic Monkey Model of the Polyglutamine Diseases Recapitulating Progressive Neurological Symptoms. eneuro 4, ENEURO.0250-16.2017. https://doi.org/10.1523/ENEURO.0250-16.2017

Toonen, L.J.A., Rigo, F., van Attikum, H., van Roon-Mom, W.M.C., 2017. Antisense Oligonucleotide-Mediated Removal of the Polyglutamine Repeat in Spinocerebellar Ataxia Type 3 Mice. Mol. Ther. - Nucleic Acids 8, 232–242. https://doi.org/10.1016/j.omtn.2017.06.019

Torashima, T., Koyama, C., Iizuka, A., Mitsumura, K., Takayama, K., Yanagi, S., Oue, M., Yamaguchi, H., Hirai, H., 2008. Lentivector-mediated rescue from cerebellar ataxia in a mouse model of spinocerebellar ataxia. EMBO Rep. 9, 393–399. https://doi.org/10.1038/embor.2008.31

Tu, Y., Liu, H., Zhu, X., Shen, H., Ma, X., Wang, F., Huang, M., Gong, J., Li, X., Wang, Y., Guo, C., Tang, T., 2017. Ataxin-3 promotes genome integrity by stabilizing Chk1. Nucleic Acids Res. 45, 4532–4549. https://doi.org/10.1093/nar/gkx095

Vale, J., Bugalho, P., Silveira, I., Sequeiros, J., Guimarães, J., Coutinho, P., 2010. Autosomal dominant cerebellar ataxia: frequency analysis and clinical characterization of 45 families from Portugal. Eur. J. Neurol. 17, 124–128. https://doi.org/10.1111/j.1468-1331.2009.02757.x

Van Alfen, N., Sinke, R., Zwarts, M., Gabreëls-Festen, A., Praamstra, P., Kremer, P., Horstink, M., 2001. Intermediate CAG repeat lengths (53,54) for MJD/SCA3 are associated with an abnormal phenotype. Ann. Neurol. 49, 805–808. https://doi.org/10.1002/ana.1089

Visschedijk, M.C., Spekhorst, L.M., Cheng, S.-C., van Loo, E.S., Jansen, B.H.D., Blokzijl, T., Kil, H., de Jong, D.J., Pierik, M., Maljaars, J.P.W.J., van der Woude, C.J., van Bodegraven, A.A., Oldenburg, B., Löwenberg, M., Nieuwenhuijs, V.B., Imhann, F., van Sommeren, S., Alberts, R., Xavier, R.J., Dijkstra, G., Nico Faber, K., Aldaz, C.M., Weersma, R.K., Festen, E.A.M., 2018. Genomic and Expression Analyses Identify a Disease-Modifying Variant for Fibrostenotic Crohn’s Disease. J. Crohn’s Colitis 12, 582–588. https://doi.org/10.1093/ecco-jcc/jjy001

Wang, H., La Russa, M., Qi, L., 2016. CRISPR/Cas9 in Genome Editing and Beyond. Annu. Rev. Biochem. 85, 227–64. https://doi.org/10.1146/annurev-biochem-060815-014607

Wang, H., Yang, H., Shivalila, C., Dawlaty, M., Cheng, A., Zhang, F., Jaenisch, R., 2013. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. Cell 153, 910–8. https://doi.org/10.1016/j.cell.2013.04.025

Wang, L.-C., Chen, K.-Y., Pan, H., Wu, C.-C., Chen, P.-H., Liao, Y.-T., Li, C., Huang, M.-L., Hsiao, K.-M., 2011. Muscleblind participates in RNA toxicity of expanded CAG and CUG repeats in Caenorhabditis elegans. Cell. Mol. Life Sci. 68, 1255–1267.

https://doi.org/10.1007/s00018-010-0522-4

Ward, J., La Spada, A., 2015. Ataxin-3, DNA Damage Repair, and SCA3 Cerebellar Degeneration: On the Path to Parsimony? PLOS Genet. 11, e1004937. https://doi.org/10.1371/journal.pgen.1004937

Warrick, J., Morabito, L., Bilen, J., Gordesky-Gold, B., Faust, L., Paulson, H., Bonini, N., 2005. Ataxin-3 Suppresses Polyglutamine Neurodegeneration in Drosophila by a Ubiquitin-Associated Mechanism. Mol. Cell 18, 37–48. https://doi.org/10.1016/j.molcel.2005.02.030

Warrick, J., Paulson, H., Gray-Board, G., Bui, Q., Fischbeck, K., Pittman, R., Bonini, N., 1998. Expanded polyglutamine protein forms nuclear inclusions and causes neural degeneration in Drosophila. Cell 93, 939–49.

Watchon, M., Yuan, K., Mackovski, N., Svahn, A., Cole, N., Goldsbury, C., Rinkwitz, S., Becker, T., Nicholson, G., Laird, A., 2017. Calpain Inhibition Is Protective in Machado–Joseph Disease Zebrafish Due to Induction of Autophagy. J. Neurosci. 37, 7782–7794. https://doi.org/10.1523/JNEUROSCI.1142-17.2017

Wojciechowska, M., Krzyzosiak, W., 2011. Cellular toxicity of expanded RNA repeats: focus on RNA foci. Hum. Mol. Genet. 20, 3811–3821. https://doi.org/10.1093/hmg/ddr299

Wu, W., Lebbink, J., Kanaar, R., Geijsen, N., van der Oost, J., 2018. Genome editing by natural and engineered CRISPR-associated nucleases. Nat. Chem. Biol. 14, 642–651. https://doi.org/10.1038/s41589-018-0080-x

Xu, Xiaohong, Tay, Y., Sim, B., Yoon, S.-I., Huang, Y., Ooi, J., Utami, K.H., Ziaei, A., Ng, B., Radulescu, C., Low, D., Ng, A.Y.J., Loh, M., Venkatesh, B., Ginhoux, F., Augustine, G.J., Pouladi, M.A., 2017. Reversal of Phenotypic Abnormalities by CRISPR/Cas9-Mediated Gene Correction in Huntington Disease Patient-Derived Induced Pluripotent Stem Cells. Stem Cell Reports 8, 619–633. https://doi.org/10.1016/j.stemcr.2017.01.022

Xu, X., Tay, Y., Sim, B., Yoon, S., Huang, Y., Ooi, J., Utami, K., Ziaei, A., Ng, B., Radulescu, C., Low, D., Ng, A., Loh, M., Venkatesh, B., Ginhoux, F., Augustine, G., Pouladi, M., 2017. Reversal of Phenotypic Abnormalities by CRISPR/Cas9-Mediated Gene Correction in Huntington Disease Patient-Derived Induced Pluripotent Stem Cells. Stem cell reports 8, 619–633. https://doi.org/10.1016/j.stemcr.2017.01.022

Yamada, M., Tan, C., Inenaga, C., Tsuji, S., Takahashi, H., 2004. Sharing of polyglutamine localization by the neuronal nucleus and cytoplasm in CAG-repeat diseases. Neuropathol. Appl. Neurobiol. 30, 665–675. https://doi.org/10.1111/j.1365-2990.2004.00583.x

Yang, D., Scavuzzo, M.A., Chmielowiec, J., Sharp, R., Bajic, A., Borowiak, M., 2016. Enrichment of G2/M cell cycle phase in human pluripotent stem cells enhances HDR-mediated gene repair with customizable endonucleases. Sci. Rep. 6, 21264. https://doi.org/10.1038/srep21264

Yang, H., Li, J., Liu, S., Zhao, J., Jiang, Y., Song, A., Hu, H., 2015. Aggregation of polyglutamine-expanded ataxin-3 sequesters its specific interacting partners into inclusions: Implication in a loss-of-function pathology. Sci. Rep. 4, 6410. https://doi.org/10.1038/srep06410

Yusa, K., 2013. Seamless genome editing in human pluripotent stem cells using custom endonuclease–based gene targeting and the piggyBac transposon. Nat. Protoc. 8, 2061–2078. https://doi.org/10.1038/nprot.2013.126

Zetsche, B., Volz, S., Zhang, F., 2015. A split-Cas9 architecture for inducible genome editing and transcription modulation. Nat. Biotechnol. 33, 139–42. https://doi.org/10.1038/nbt.3149

Zhang, X., Tee, L., Wang, X., Huang, Q., Yang, S., 2015. Off-target Effects in CRISPR/Cas9-mediated Genome Engineering. Mol. Ther. Nucleic Acids 4, e264. https://doi.org/10.1038/mtna.2015.37

Zhang, Y., Wu, S., Song, S., Lv, J., Feng, C., Lin, X., 2017. Generation and characterization of a potentially applicable Vero cell line constitutively expressing the Schmallenberg virus nucleocapsid protein. Cytotechnology 69, 145–156. https://doi.org/10.1007/s10616-016-0046-3