RNAi based allele-specific silencing of the disease-causing gene in Black South African patients with SCA7

Janine Scholefield (MSc)

Thesis Presented for the Degree of
DOCTOR OF PHILOSOPHY
in the Department of Human Genetics
UNIVERSITY OF CAPE TOWN

August 2008

DECLARATION

I, Janine Scholefield, hereby declare that the work on which this research project is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university. This study has been approved by the Research Ethics Committee of the Faculty of Health Sciences; number 422/2005.

I empower the university to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever.

Signature: Signed by candidate

Janine Scholefield (MSc)

Date: 21st August 2008

Dedication

For Mum and Dad

Acknowledgements

Firstly and foremostly I would like to acknowledge my supervisor, Professor Jacquie Greenberg. Under her instruction over the last few years she has imparted to me the principles and ethics of how to be a great scientist. She has been both a great mentor and a great friend…no student could ask for more.

My family, in particular my parents, whose encouragement, love and support (and enduring hope that one day I might earn a salary) has fuelled my determination.

My co-supervisors, Dr Matthew Wood, Dr Marco Weinberg and Professor Patrick Arbuthnot for their constant encouragement and criticism; most specifically for helping me realise the potential of this sexy science.

My colleagues in the Division of Human Genetics, in particular Miss Lisa Roberts and Dr George Rebello. Lisa, who has been my rock and a true friend in the lab; George, a mentor and the scientist whose opinion I value the most. Most of all, to both of you for your wicked sense of humour… the drosophila impression will be etched in my brain forever. Mary Viljoen, Latiefa Jattiem and Adielah van der Schyff for your happy faces and for bending over backwards to organise my disorganised life.

Mrs Toni Wiggins, Dr Lester Davids, Dr Sharon Prince, Miss Ingrid Baumgarten and Professor Sue Kidson for introducing me to molecular biology and for your support.

Dr Wood's lab in Oxford especially Graham, Chris, Aurelie and Yvonne who made me feel I had a home away from home.

The HBV lab for their support during my work at Wits.

Professor Anne-lise Williamson for the tissue culture facilities.

Mrs Adri Winckler and Emeritus Professor Peter Beighton for always having an open door.

Professor Raj Ramesar for providing the environment in which I have worked and his support.

Associate Professors Jeanine Heckmann and Alan Bryer and Sister Di Sklar for their assistance with the SCA7 families.

The following organisations for funding support

- The South African Spinocerebellar Ataxia Association
- Ataxia UK
- The Oppenheimer Fund (University of Oxford)
- The John Fell Fund (University of Oxford)
- The South African National Research Foundation
- South African Medical research Council
- Guy Elliot Fellowship
- UCT postgraduate funding office
  - Siri Johnson Bursary
  - KW Johnstone Scholarship
  - Harry Crossley Postgraduate Scholarship
  - Marion Beatrice Waddel Bursary
  - Twamley Scholarship
  - Benfara Scholarship
- Ernst and Ethel Eriksen Trust

Finally, my friends for staying friends with a “freelance molecular geneticist”, and trying so hard to understand what I do, for Tuesday nights, for Sundays with The Times’ 20 questions, for sailing expeditions, for African bush-wacking adventures, but most of all for the wine and laughter.

Abstract

The polyglutamine disorders are a subgroup of inherited neurodegenerative disorders with a common mutation which confers toxicity via a polyglutamine tract in the protein leading ultimately to various forms of neurodegeneration. One of these disorders, spinocerebellar ataxia 7 (SCA7) exists at a higher frequency in South Africa, than elsewhere in the world, and a founder effect has been demonstrated in South Africa, such that every patient tested thus far is linked to a common ancestor.

The manipulation of RNA interference (RNAi) has been used with increasing success to selectively knockdown the expression of disease-causing genes at the RNA level. Thus, the possibility of applying this method to SCA7 in South Africa was considered. However, the wild-type allele of ataxin-7 is likely to be necessary for cellular function therefore a form of allele-specific silencing is required, such as a SNP linked to the mutation. Given the less than desirable nature of the SNP itself, extensive screens of twenty different RNAi hairpins were performed to identify one that might distinguish between the mutant and wild-type alleles. Using different reporter genes and target lengths a hairpin was identified that showed significant selective knockdown of the mutant transcript in a heterozygous full-length assay. Furthermore, this hairpin led to the release of wild-type protein from mutant-forming aggregates, suggesting an additional benefit of allele-specific silencing over general knockdown. Finally, re-structuring of this hairpin in a primary miRNA based format yielded the same selectivity at the same position, strongly suggesting that the designs and therefore mismatch placement positions were accurate. These findings strongly support previous studies that suggest that single nucleotide mismatch placement in the 3' end of the guide strand may lead to strong discrimination.

Most importantly, as SCA7 uniquely presents with a retinal phenotype in addition to the classic cerebellar pathology, the relative accessibility of the eye lends itself to efficacious assessment of therapy. This study will therefore form the basis for *in vivo* work investigating the efficacy of RNAi based gene therapy.

Table of contents

List of Tables

List of Figures

List of Appendices

Abbreviations:

AAV	Adeno-associated virus
AcH3	acetylated histone 3
AchR	acetylcholine receptor
AD	Alzheimer's disease
ADCA	autosomal dominant cerebellar ataxia
Ago2	Argonaute2
ALS	amyotrophic lateral sclerosis
AOO	age of onset
A2RP1	ATXN2 related protein 1
atxn3	Ataxin-3 gene/ mRNA
ATXN7	Ataxin-7 protein
atxn7	Ataxin-7 gene/ mRNA
BDNF	brain-derived neurotrophic factor
bp	base pair(s)
CMV	cytomegalovirus
CNS	central nervous system
CRX	cone-rod homeobox (homo sapiens)
Crx	cone-rod homeobox (mus musculus)
DGCR8	DiGeorge critical region-8
dH_2O	distilled water
DRPLA	dentatorubral pallidoluysian atrophy
dsRNA	double stranded RNA
eGFP	enhanced green fluorescent protein
GCL	granule cell layer
GFP	green fluorescent protein
HAT	histone acetyltransferase
HEK293	human embryonic kidney cells
HD	Huntington disease
HDAC	histone deacetylase
Hsp	heat shock protein
htt	huntingtin

IGF	insulin growth factor
JNK	c-Jun NH_2-terminal kinase
Kb	kilobases
LN	larger normal
µg	micrograms
min	minutes
miRNA	microRNA
µl	microlitres
ml	millilitres
mM	millimolar
mRNA	messenger RNA
ML	molecular layer
NES	nuclear export signal
ng	nanograms
NII(s)	neuronal intranuclear inclusion(s)
NLS	nuclear localisation signal
nm	nanometre
N/P	nitrogen to phosphate ratio
nt	nucleotide(s)
ONL	outer nuclear layer
PAZ	Piwi Argonaute Zwille
PCL	Purkinje cell layer
PCR	Polymerase chain reaction
pmol	picomoles
Pol II	polymerase II
Pol III	polymerase III
pre-miRNA	preliminary miRNA
pri-miRNA	primary miRNA
RFP	red fluorescent protein
Rho	Rhodopsin
RISC	RNA induced silencing complex
RLC	RISC loading complex
RSA	Republic of South Africa
RNAi	RNA interference

SAGA	histone acetyltransferase Spt/Ada/Gcn5/acetylase
SBMA	spinobulbar muscular atrophy
SCA	spinocerebellar ataxia
SCCMS	slow-channel congenital myasthenic syndrome
sec	seconds
shRNA	short hairpin RNA
siRNA	small interfering RNA
SNP	single nucleotide polymorphism
STAGA	SPT3/TAF9/GCN5 acetyltransferase complex
T_A	annealing temperature
TAF	TBP associated factor
TBP	TATA box binding protein
TFTC	TBP-free TAF containing protein
TRBP	TAR RNA binding protein
UK	United Kingdom
UPS	Ubiquitin proteasome system
USA	United States of America
UV	Ultra violet

1. Introduction

In recent years, disorders of the nervous system have been seen to have an increasing impact on the global burden of disease (1) for the most part, due to the trend towards aging populations (2). While this is currently more of a feature in developed countries, improved medical care in developing nations will soon lead to a similar phenomenon, with better health care leading to increased life expectancy (2). Neurodegenerative diseases encompass a group of nervous system disorders presenting with progressive symptoms that last decades, placing a significant financial and emotional burden on families. The devastating phenotypes and general lack of meaningful therapies thus requires a significant effort towards the investigation of novel cures. Given the incomplete understanding of neurodegeneration, an ideal subgroup to study is that of the dominantly inherited disorders, since the disease-causing mutation sheds light on the mechanism of pathogenesis.

The dominantly inherited neurodegenerative disorders include a plethora of distinct as well as phenotypically overlapping diseases. Within this group, the subset of polyglutamine disorders constitutes a significant area of research. Conditions falling within this subset have a common type of mutation; that of an expanded CAG repeat, encoding a stretch of polyglutamines within the disease-causing protein. To date, nine polyglutamine disorders have been identified; Huntington disease (HD), spinobulbar muscular atrophy (SBMA), dentatorubral pallidoluysian atrophy (DRPLA) and the following spinocerebellar ataxias, 1, 2, 3, 6, 7 and 17 (3). As a result of the common type of mutation these disorders share a variety of features including late onset of symptoms, anticipation and instability of repeats in transmission from parent to child. The common pathogenic number of repeats in each polyglutamine disorder vary, but mostly lie in the 40-60 range, with numbers extending into the hundreds (4, 5). The polyglutamine disorders share similar threshold levels, such that numbers of 37- 40 repeats in most of these

conditions results in presentation of neurodegeneration (3). A common feature of all these disorders is that of anticipation, whereby the number of repeats beyond the threshold level is inversely proportional to the age of onset (AOO) of symptoms, such that larger expansions result in earlier and often more severe forms of the phenotype. While the mechanism of pathogenesis for each disorder has not been fully elucidated, many common features have been identified. The toxic gain-of-function caused by the mutant protein is a generally accepted view, possibly due to nuclear fragmentation and/or cytoplasmic or nuclear aggregation of the mutant proteins, of which the implications of the latter is a highly contentious issue (6). The context within which the expanded repeat lies is likely to lead to different cell-type susceptibilities based on the function of each of the protein coding genes. Thus specific disorders show degeneration of various regions of the cerebellum and cortex, consequently leading to different phenotypic symptoms. It is therefore possible that research conducted on one of these disorders may provide information on others.

With this in mind, we wanted to investigate a possible therapeutic approach targeting the mutant gene in SCA7. One of the methods showing success with other polyglutamine disorders is the manipulation of the RNA interference (RNAi) pathway in order to yield decreased expression of the disease-causing gene. Thus, in this study an RNAi based approach was investigated specifically for SCA7 patients in South Africa.

1.1. Spinocerebellar ataxia 7

1.1.1. Phenotype

The significant heterogeneity between and within families with inherited late onset ataxia led to the classification of the different forms of the disorder into three different subsets of autosomal dominant cerebellar ataxias (ADCA) based on the Harding classification (7); ADCAI, II and III. ADCAII was characterised by ataxia and the distinction of progressive blindness in the

form of pigmentary retinopathy (8). However in the mid-1990s the identification of heterogeneous disease-causing mutations in this and other types of ADCA led to a reclassification based on the genotype. Although ADCAI and III constitute a variety of genetically different SCAs, ADCAII is caused by SCA7 alone (8). A single exception to this classification has been identified in which a British family presenting with SCA7 symptoms is negative for the SCA7 mutation (9).

A more recent description of SCA7 symptoms includes ataxia, dysarthria, dysphagia, and exaggerated deep tendon reflexes, as well as a macular degeneration of cone-rod dystrophy (10, 11). In adult onset SCA7, cerebellar symptoms usually precede the visual phenotype; however other phenotypic signs are evident in the earlier onset (childhood or juvenile) forms of the disorder, usually with visual degeneration preceding the onset of cerebellar symptoms (12). Juvenile onset SCA7 presents with an early onset of symptoms and a more rapid progression of the disorder, but is otherwise similar to adult onset SCA7. Infantile onset, however, which occurs before 2 years of age, can be quite distinct phenotypically (13, 14). Features include failure to thrive, weight loss, motor regression, cardiac abnormalities, wasting and hypotonia with ataxia difficult to recognise; indeed the infantile onset cases appear to involve multiple tissues and organs (14, 15).

1.1.2. Genetics

Identification of linkage to the short arm of chromosome 3 indicated that SCA7 was genetically distinct from other ADCAs (16). Further investigation of some SCA7 families using the repeat expansion detection method (used to identify expanded repeats without knowledge of the locus of the mutated gene) indicated that an expanded triplet repeat co-segregated with the disease (4). Refinement of the previously mentioned area of chromosome 3 (17) and positional cloning techniques (18) resulted in characterisation of the SCA7 gene, *ataxin-7* (*atxn7*) (19) confirming the expanded repeat sequence

as disease-causing in five different French families with SCA7. A schematic representation of the *atxn7* is indicated in Figure 1.1.

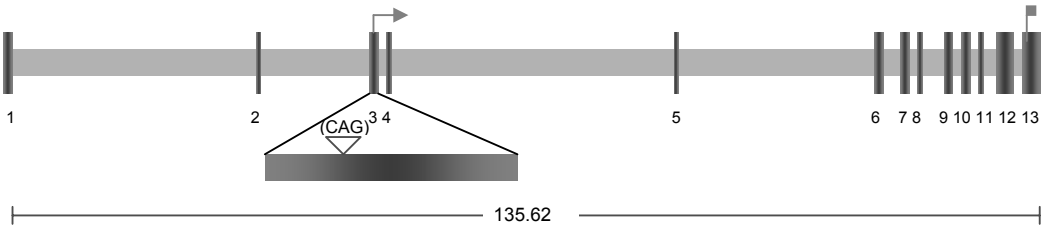


Figure 1.1: Representation of the *atxn7* transcript in Vega (BCM:ATXN7-001). This transcript represents the most common isoform containing 13 exons, 3.940 Kb with 892 residues over 135.62 Kb of genomic DNA (19). Alternative transcripts have subsequently been identified in humans with homologues in mice (20, 21). Numbers represent exons shown with approximate distances between each other.

The normal allele of *atxn7* contains CAG repeat numbers ranging from 4 to 35 with the most common normal sized alleles being $(CAG)_{10}$, (12, 22) although this differs in some population groups (23). Allele sizes within the range of 28 to 35 are now considered mutable normal alleles (previously known as intermediate alleles) because while there are no reports of SCA7 being associated with these sizes, they have been shown to expand into the pathological range in transmission from parent to child (24). The sizes of disease-causing expansions are considered to range from 37 to over 400, with the largest expansion identified being $(CAG)_{460}$ (15). The lower threshold is not perfectly defined, however, as results from a family study show phenotypic presentations in individuals with 36 repeats (25). Repeats in SCA7 show the greatest instability in transmission from parent to child than any of the polyglutamine disorders, in particular through the paternal line, indicating the predisposition of spermatogenic instability (24, 26) such as that indicated in Figure 1.2. The bias of instability in paternal transmission is not unique to SCA7 (27, 28) and may be due to the high mitotic turnover of spermatocytes in comparison to oocytes (29). However the extreme cases of instability far surpass those of the other polyglutamine disorders with a report of a $(CAG)_{49}$ expanding to $(CAG)_{460}$ in a single transmission (15). A possible

explanation may be that the genomic context within which the *atxn7* expansion lies contributes to the instability of the repeat (30). Evidence for this has been demonstrated in a mouse model of SCA7 where the removal of intronic sequences within and flanking *atxn7*, leads to reduced instability in transmission to offspring (30).

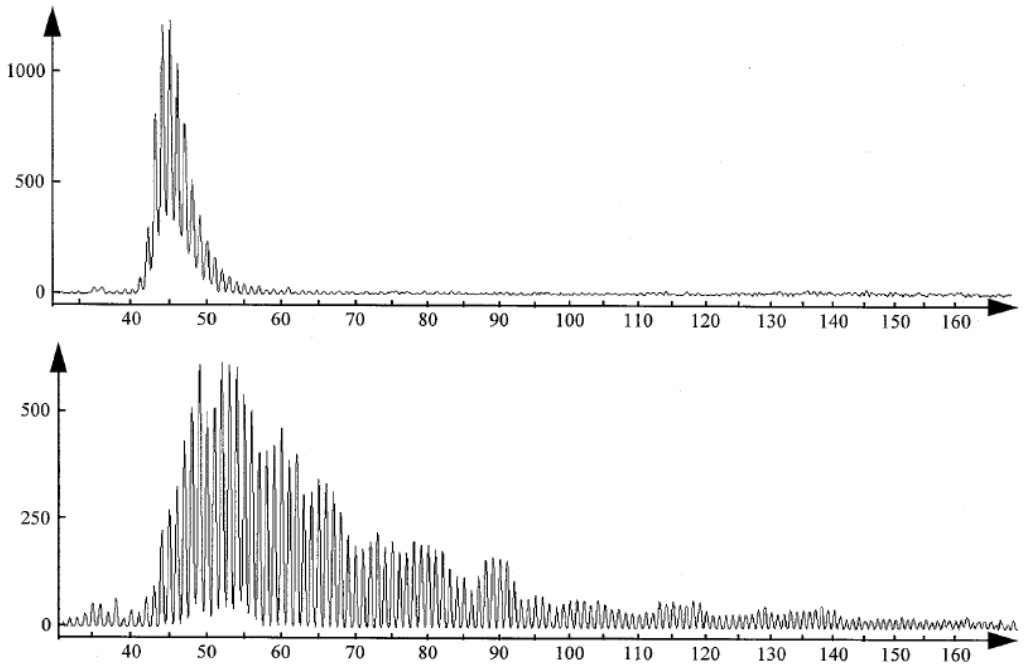


Figure 1.2: A comparison of the number of repeats identified in blood (top panel) and sperm (bottom panel) of a SCA7 patient. PCR products were analysed using GENESCAN software, where the horizontal axis represents the number of repeats and the vertical axis, fluorescent units. Taken from David et al. 1998 (22).

1.1.3. Correlations between genotype and phenotype

Anticipation is the phenomenon of a decreasing AOO in successive generations first observed in patients with HD by Heilbronner in 1903 (31). It has since been shown to be caused by an increase in the number of repeats in transmission from parent to child (32), and is striking in SCA7, due to the previously mentioned dramatic instability of the SCA7 (CAG) repeat. The strong correlation between the AOO and the number of repeats has been examined extensively (12, 22, 24) and indicates that the repeat number contributes to most of the variability in AOO, with an inversely proportional

relationship. An increase in the number of repeats has also been shown to be associated with a more severe and shorter course of disease (22). Generally, patients with smaller repeat sizes (less than 60) present with classic ataxic symptoms, later followed by progressive macular degeneration (12). Indeed, a patient with only 39 repeats and an AOO of 77 years showed no gross macular degeneration, reinforcing this observed trend (33). In contrast, patients carrying larger expansions (which constitute a smaller percentage of SCA7 patients (12, 22)), and therefore presenting with an earlier AOO, present with initial symptoms of macular degeneration, followed by progressive ataxia or a dual onset of symptoms.

1.1.4. Pathology

The pathology of SCA7 correlates with the phenotype in that both regions of the cerebellum and the retina show marked degeneration. Atrophy of the cerebellum and pons in human studies are typical of SCA7, with the latter significantly worse in SCA7 patients than for other SCAs, and found to precede atrophy of the cerebellum (11). Specifically, atrophy of the cerebellar cortex and dentate nucleus, is noted (34). Extensive loss of the Purkinje cells with minimal changes to the granular layer of the cerebellum occurs. In the brainstem, olivocerebellar, spinocerebellar, and pyramidal tracts show atrophy with the inferior olive showing noticeable neuronal loss with gliosis. Pyrimidal pathways and motor neurons in the spinal cord are also affected. Degeneration of the subthalamic nucleus, globus pallidus and substantia nigra is variable, and seen in some patients (34). Other pathologies include neuronal loss in brainstem cranial nerve motor nuclei (III, IV, XII), and demyelination of pyramidal tracts and posterior columns of the spinal cord (11). Inclusion formation has been reported in affected areas, notably in the inferior olive, however they are also found in some areas of the cerebral cortex not affected in SCA7 brains, although this study reported an early onset case (35). In contrast, cases of severe infantile onset SCA7 show significant multiorgan failure in addition to neuronal degeneration (14, 15).

The retinal phenotype occurs as a result of the degeneration of the cone and rod photoreceptor neurons in the retina – one of the few disorders characterised by cone followed by rod photoreceptor degeneration (36). The retinal phenotype appears as macular degeneration, but is caused by photoreceptor dysfunction of initially the cones, followed by the rods, leading to the disorder being incorporated into the group of cone-rod dystrophies. Central vision is initially impaired as indicated by a decrease in colour discrimination (37, 38). The loss of vision then proceeds to the entire retina, as dysfunction of the rods follows that of the cones, leading to total blindness.

1.1.5. Function of the normal protein

The normal protein has been studied by several groups with some success as to the characterisation and the identification of the function of the protein. Initial investigations of the *atxn7* gene product (ATXN7) indicated that it localised to the nucleus due to the presence of an arginine-lysine nuclear localisation signal (NLS), whose activity was demonstrated in transfected Cos-1 cells (39). In addition the protein was shown to contain a phosphate binding site (40). Both of these domains are conserved in the murine homologue of ATXN7, which has 88% homology to the human sequence (41). In addition, the protein has been found to possess a highly conserved nuclear export signal (42) which in addition to the NLS suggests a nuclear-shuttling role for ATXN7 (43). Immunoreactivity for the 95 kD protein has been found the nucleus of Purkinje cells, although both cytoplasmic and nuclear staining has been identified in non-cerebellar neurons (44). The distribution of the normal protein indicates ubiquitous expression throughout the neurons of the nervous system, as well the retina, in addition to expression of *atxn7* in several non-neuronal tissues (45). The only tissues tested that did not show immunoreactivity were the liver and kidney. Jonasson and colleagues refined these data by comparing their data with that of other studies (46). They indicated that there may be an age-dependent sub-cellar localisation of the protein in brains of control

individuals, with nuclear staining present earlier followed by cytoplasmic staining later in life. However the data for these controls is not conclusive and since the studies on the five different controls were performed by four different groups, this would need to be verified (46). Subcellular localisation shows co-localisation with cytoplasmic organelles such as the endoplasmic reticulum and nuclear envelope, but with neither the mitochondria nor the golgi (45).

Yeast 2-hybrid studies looking for ATXN7-interacting proteins identified R85, a splice variant of Cbl-associated protein as well as an ATPase subunit S4 of the 19S proteasomal complex (47, 48). The former suggested a possible role for the ubiquitination and degradation of ATXN7, whilst the latter indicated proteasome inhibition, which is known to cause aggregation and neuronal intranuclear inclusions (NIIs) in *in vitro* polyglutamine models such as SCA3 (49).

Subsequently, several studies have proposed that ATXN7 is involved in transcriptional regulation. In 2001, a study revealed that ATXN7 interacted with the cone-rod homeobox (CRX) protein – a photoreceptor cell-specific transcription factor (50). Although the exact function of CRX is unknown, mutations in the gene coding for the latter are responsible for retinal degenerative disorders (51). Biologically, many of these mutations have been shown to result in dysregulated expression of the photoreceptor genes. Furthermore a knockout mouse model of Crx (the *mus musculus* homologue of CRX) develops retinal degeneration due to impairment of photoreceptor cell development (52). In 2004 interacting domains of CRX and ATXN7 were mapped using various deletions and point mutations and indicated that the CRX/ATXN7 interaction occurs via their respective polyglutamine tracts present in the wild-type proteins (36). It was further indicated that CRX and ATXN7 co-occupy a regulatory region of the CRX target genes regulating expression of retinal genes (36).

A second hypothesis for the role of ATXN7 in transcriptional regulation came from the identification of a yeast homologue which showed significant similarity to a gene identified as a novel subunit of the histone acetyltransferase (HAT) Spt/Ada/Gcn5/acetylase (SAGA) complex - responsible for approximately 10% of transcription regulation of the yeast genome (53). The following year the ATXN7 protein was identified as a subunit of GCN5 HAT of the transcriptional co-activator complexes, TATA box binding protein (TBP)-free TBP associated factor (TAF) containing protein TFTC, and SPT3/TAF9/GCN5 acetyltransferase complex STAGA; the mammalian equivalent of the yeast SAGA complex (54). These complexes contain several proteins except TBP, making it distinct from the general transcription factor TFIID and preferentially acetylate histone H3 and activate transcription on chromatin templates (54). Indeed it has since been demonstrated that an *in vivo* interaction between STAGA and ATXN7 exists and is functional, based on the HAT activity of GCN5 (55).

1.1.6. Pathogenesis

A common point of discussion within the area of research of polyglutamine disorders is that of the mechanism of pathogenesis; whether it be solely a gain-of-function due to the accumulation of toxic protein fragments, or alternatively whether loss-of-function of the wild-type allele contributes to disease pathogenesis. Certainly, it is generally accepted that the gain-of-function accrued by the toxic mutant protein contributes towards and results in cell death in all the polyglutamine disorders. However, the precise manner in which the mutant protein which causes toxicity is unknown. Many studies have explored cell-specific vulnerability, expression, aggregate formation and other possible mechanisms of pathogenesis without definitive outcomes. In addition the extent to which loss of function may contribute towards disease pathogenesis is also unknown.

1.1.6.1. Expression and aggregation

For the most part there has been difficulty identifying a correlation between the level of expression, cell-type expression and aggregate formation with respect to the neuropathology observed in SCA7 patients. This is most strikingly demonstrated by the ubiquitous expression of *atxn7* contrasting sharply with specific sub-types of neurons affected by the disease process (45). Furthermore there is no relationship between levels of expression and the degree of degeneration suggested by the strong immunoreactivity of ATXN7 in the external pallidum compared with the weak presence of the protein in the Purkinje cells, both severely degenerated in SCA7 patient brains (see Figure 1.3, page 21, (56)). The only correlating observation made was the increase in loss of neurons with nuclear labelling. However this hypothesis is weakened due to the lack of a direct relationship between cells containing inclusions and cell death. While almost all cells with inclusions showed high levels of expression, the converse did not follow in many situations such as the lack of inclusions in the putamen and pallidum, which had shown high levels of expression.

One of the most controversial areas within the field of polyglutamine disorders research is the consequence of aggregate formation within specific neuronal cell types, i.e. are the aggregates protective, causative or simply associated with neuronal death? Aggregation in the form of NIIs has been identified in multiple neuronal cell types (35, 39, 44, 46, 56) and these inclusions appear large and at the ultra-structural level, filamentous, a feature regarded as specific to the glutamine expansion (57). These NIIs have been shown to be for the most part, ubiquitinated, indicating that the inclusions are recognised by the ubiquitin proteasome system (UPS) but are in some way resistant to degradation (35, 56). Interestingly, a patient with 41 repeats (who died at 92 years of age) was shown to have no inclusion formation, whilst a patient from the same kindred with 42 repeats did show evidence for inclusion formation, suggesting other factors were involved in the formation of inclusion bodies (44). Furthermore, inclusion formation in infantile cases of

SCA7 spreads to non-neuronal tissues such as the cardiovascular system which is known to be a component of the broad systemic phenotype of infantile cases (58). These inclusions were shown to be non-ubiquitinated. However, strongly ubiquitinated inclusions were identified in the hippocampus of the same individual; an area demonstrating a striking lack of cellular loss. ATXN7 NIIs have been shown to form in nuclear bodies and recruit proteins including those involved in the UPS as well as chaperones (47, 48, 59) although the impairment of the UPS was examined *in vivo* and shown not to be involved in disease pathology. Indeed this study strongly suggests that NIIs are protective (60).

1.1.6.2. Studies of the mutant protein

In a study by Zander and colleagues, extensive observations of ATXN7 were made using comparisons of *in vitro* cellular models (42). Using various constructs of full-length and truncated, mutant and wild-type atxn7 cDNA fused to multiple tags, the authors were able to investigate the localisation and interactions of both wild-type and mutant ATXN7. Firstly, they showed that the protein in its normal form is predominantly nuclear, confirming a functional nuclear localisation signal (NLS) initially identified by Kaytor and colleagues (39), and that no aggregates were found in human embryonic kidney cells (HEK293) transfected with the normal $(CAG)_{10}$-containing vector. However, nuclear and peri-nuclear inclusions were formed in the same cells transfected with the mutant $(CAG)_{100}$ full-length vector. Interestingly, cells sub-cloned from a human neuroblastoma cell line did not show as high expression levels of transfected ATXN7, nor as extensive aggregate formation. Ultrastructural examination provided evidence for the fact that the aggregates formed by ATXN7 were fibrillar in nature. Furthermore, the study showed that the aggregate formation in the *in vitro* cellular model did not require cleavage of the protein, in that both terminals of the protein could be identified *in vitro*, however in SCA7 brains, most of the NIIs were only reactive to the N-terminal antibody, suggesting some kind of masking or

processing of the protein had occurred *in vivo*. This is not a phenomenon unique to SCA7 (61).

Importantly, the study by Zander et al. also examined the proteins recruited by the ATXN7 inclusions (42). Experiments using co-transfection of mutant and wild-type vectors showed that the mutant ATXN7 recruits its wild-type counterpart into the aggregates. Further investigation indicated that several other proteins are recruited into the inclusions of SCA7 patient brains, mimicked by the *in vitro* model (42). The first of these, the SCA3 protein (ATXN3), reinforces previous studies that proteins containing polyglutamine stretches are recruited by inclusions caused by other polyglutamine proteins. Other examples include an abundance of heat shock protein (Hsp) -40 and ubiquitin, as well as proteasome subunits, although a discrepancy lies in the fact that the 19S subunit was detected in the SCA7 brain, whilst the 20S subunit is detected only in the transfected cells. The latter aside, the aforementioned folding proteins suggest that an effort towards refolding/ disaggregation/ degradation of the mutant proteins is made by the cell but is none-the-less unsuccessful. Low levels of activated caspase-3 were also detected in both models. Indeed comparisons of SCA7 and non-SCA7 brains indicated a substantial and statistically significant increase in the level of activated caspase-3 in the former, which would suggest the initiation of a cell-death pathway. This is pertinent given that mutant ATXN7 has been shown to cause apoptotic cell death in primary cultures of cerebellar neurons by induction of caspase-3 and -9 (62). However the Zander et al. study (42) also noted several morphological changes within cells forming ATXN7 positive inclusions such as irregular shaped nuclei as well as ultrastructural observations such as nuclei with indentations often associated with membrane disruption and multiple organelle abnormalities such as swollen mitochondria and enlarged cisternae of the Golgi complex associated with proximal autophagic vacuoles. These observations indicate early signs of cells in distress reminiscent of an autophagic (non-apoptotic) cell death illustrating the importance of differences noted between neuronal and non-

neuronal cell types. Further analysis of these cells showed 10-20 times retarded growth, compared to minimal effects by the addition of only the wild-type vector, suggesting a response of the cell to both expression of the mutant protein, as well as the effects of aggregation.

1.1.6.3. Alteration of transcription

Since the wild-type protein has been shown to be affected by the presence of the mutant, there is a suggestion that loss of function of this transcription subunit would lead to undesired phenotypic effects. Alternatively transcription could be affected by the dominant negative effects of the mutant protein.

Given that ATXN7 has been shown to recruit and interact with retinal-specific transcriptional complexes, it follows that disruption of the recruitment would lead to dysfunction of retinal-specific genes. Although ATXN7 had been shown to be a subunit of the transcriptional activator complex, STAGA (54), the introduction of the expansion did not prevent complex formation. Experiments in yeast showed that while human ATXN7 could rescue SAGA formation in the absence of yeast Sca7, expanded human ATXN7 forms a dysfunctional SAGA complex (63). This complex retains its HAT activity, but has reduced interactions of crucial proteins, allowing for acetylation of free histones, but not nucleosomes, suggesting the alteration of transcription of certain genes. In an accompanying paper this mechanism of mutant ATXN7 pathogenesis was further elucidated as indicated by the ascertainment of GCN5 HAT activity which is removed with the addition of atxn7_92Q but not a truncated form of atxn7_92Q (55). Furthermore, the authors showed an *in vitro* as well as *in vivo* association of STAGA and CRX, dependent on the presence of ATXN7. Interestingly, knockdown of *atxn7* resulted in decreased levels of GCN5 and STAF36 (another component of STAGA), and that the resultant lack of HAT activity could be rescued by the addition of histone deacetylase (HDAC). *In vivo* experiments showed a discrepancy between consistently reduced CRX occupancy on CRX-dependent promoters as opposed to a progressive reduction in acetylated histone 3 (AcH3; a

substrate of GCN5) on CRX dependent genes. This led the authors to suggest that an effect of polyglutamine ATXN7 on CRX could not be solely responsible for the transcriptional dysregulation observed. Taken together the study contributed not only to the understanding of the normal function of ATXN7, but in addition to the complex effect of the polyglutamine protein. Indeed the authors suggest that while it is not clear whether expanded ATXN7 replaces normal ATXN7, or interferes with the latter's associations, it is clear that it causes a dominant negative effect upon retinal transcription, contributing to the understanding of the unique retinal-specific degeneration in SCA7.

1.1.6.4. Nuclear localisation and cleavage

Trafficking into and out of the nucleus appears to be a common theme of polyglutamine disorders (64, 65). In their study Taylor and colleagues showed that a point mutation in the conserved nuclear export signal (NES) of ATXN7 resulted in cellular toxicity, in addition to which the mutant protein was shown to have a reduction in its ability to export from the nucleus in a polyglutamine-length dependent manner. This lends itself to the hypothesis that the accumulation of the mutant protein in the nucleus may lead to mild toxicity which is evident over a longer time period – resulting in a late onset of symptoms (43).

Cleavage of the mutant protein into truncated N-terminal fragments has been suggested as contributing to pathogenesis in other polyglutamine disorders (66-69). Hinting at a similar mechanism in SCA7, ATXN7 NIIs in animal models show the presence of truncated fragments (70) (Chapter 1.1.7.1). Recently it was shown that caspase-7 cleaves ATXN7 (71). Cleavage results in increased toxicity *in vitro* as well as the accumulation of NIIs and transcriptional repression. Interestingly though, this cleavage position results in the disruption of the NES, creating accumulation of the fragment in the nucleus. They further showed that expanded ATXN7 activates caspase-7 in the nucleus and the resulting product is the same size as that seen *in vivo*

(70) thus indicating that capase-7 would likely act *in vivo* to create the observed fragments. Prevention of this cleavage can result in reduced toxicity, as is the case with other polyglutamines suggesting a common contributing mechanism of pathogenesis for the polyglutamine diseases (71), but as has been previously shown, cleavage does not prevent aggregation *in vitro* (chapter 1.1.6.2), further reinforcing the lack of a correlation between aggregation and toxicity.

1.1.6.5. Contribution of larger normal alleles

While most studies have centred on expanded polyglutamine proteins beyond a defined threshold level, one study has investigated the implications of larger normal (LN) alleles. Investigation of a cohort from Southern India showed that a population of patients with ataxia symptoms (negative for expansions in SCA1, 2, 3, 7 and DRPLA) had a significantly higher frequency of LN alleles (23). For instance the most common normal allele in this group for the SCA7 locus was $(CAG)_{15}$. Many of these individuals had LN alleles at more than three of the aforementioned loci, indeed, over 75% had LN alleles at more than two loci, while none of the control individuals had LN alleles at more than two loci. Given that sequestration of transcription factors (theorised to give rise to pathogenesis) occurs due to accumulation of polyglutamine proteins, the authors suggest that a possible mechanism of disease could be caused by not one polyglutamine protein in the traditional defined expanded range, but by many LN polyglutamine proteins. While these results are intriguing, this study is the only investigation to hypothesise such a mechanism and has yet to be validated experimentally or by other studies of polyglutamine disorders.

1.1.7. Lessons from in vivo models

In an elegant paper in 2004, Helmlinger and colleagues posed the following, "*It remains to be determined at which stage polyglutamine expansions induce irreversible damage to neurons. This issue is crucial, considering therapeutic*

strategies interfering with polyglutamine protein expression. Finally, it is not known whether neurons can recover from the toxicity of full-length mutant proteins and whether these results can be extended to other polyglutamine disorders" (72). Hence it is of utmost importance to be aware of the detailed phenotype of animal models of the disease, in order that we might be able to assess the efficacy of interventions as well as gain a deeper understanding of the mechanism of disease.

1.1.7.1. Mouse models

Three groups have created transgenic mouse models in an attempt to mimic the SCA7 phenotype. The first designed several transgenic lines created using full-length human *atxn7* cDNA containing 10 or 90 (CAG) repeats expressed off either the photoreceptor-specific rhodopsin promoter (R7N/E) or the Purkinje cell-specific (pcp-2) promoter (P7N/E) (73). Mice expressing the expanded sequence from these two promoters developed retinal and locomotor deficient phenotypes respectively. R7E mice showed several manifestations of retinal pathology including severe disorganisation of the outer nuclear layer (ONL), and a reduction in the percentage of photoreceptor cells. Interestingly, in addition to demonstrating progressive retinal degeneration as well as profuse NII formation, the mice showed morphological changes in interneurons (postsynaptic to the targeted neurons) suggesting a possible explanation for a *trans* degeneration of neurons with minimal *atxn7* expression. In the P7E mice, Purkinje cells showed NIIs after eight months, and by 16 months, almost all these cells contained NIIs. The phenotype was less obvious, in that there were no visible abnormalities in the mice after 1 year possibly due to the low expression level of the transgene under the pcp-2 promoter. However, a rotarod performance test showed that P7E mice performed significantly worse than their wild-type littermates. The pathological changes in these mice were also less noticeable than those in the R7E mice, however, reduced dendritic arbour – indicative of unhealthy neurons - and a decreased Purkinje cell density were evident in the cerebellum. Interestingly, the NIIs were shown to

be positive only for the N-terminal ATXN7 antibody, and not the C-terminal antibody, suggesting that some form of proteolytic cleavage of the mutant protein may be occurring as is the case in SCA7 brains (42). Redesigning the transgene construct to incorporate 128 repeats expressed off a stronger promoter resulted in. the mice developing ataxia a few months after birth, with death within 3 months of the onset of symptoms (74). It was shown that localisation and processing - including N-terminal fragment accumulation and stabilisation of ATXN7 by the expansion resulting in decreased turnover of the mutant protein - was the same through spared and unspared areas of neurodegeneration. Finally it was noted that NIIs sequestered many proteins of the transcriptional pathway confirming findings from a previous *in vitro* study (42).

The R7E mouse model was also used to investigate the involvement of chaperones in the pathogenesis of SCA7. Previous studies of drosophila models of polyglutamine disorders and *in vitro* cellular models (75, 76) had shown that increased expression of specific heat shock proteins has beneficial effects. However, although increased expression of Hsp40 and 70 reduced aggregate formation *in vitro*, no reduction occurred *in vivo* (77). The authors suggest that while the expression of specific chaperones may not ameliorate pathogenesis, indication of a global heat shock response may be more successful. This data is intriguing in light of the fact that a later study showed decreased protein levels of Hsp27 and 70 in patient lymphoblasts (78). Given the lack of correlation between *in vivo* mouse models and *in vitro* human cell lines, this aspect of SCA7 pathogenesis warrants further study.

A second study (72) adapting the aforementioned mouse model (74) investigated at which point irreversible neuronal damage occurs after exposure to polyglutamine proteins, using double transgenic mice, R7N-R7E. These mice expressed both full-length normal (R7N) and expanded (R7N) human *atxn7* specifically in the retina under the transcriptional control of the rhodopson (79) promoter in order to investigate whether the retinal

phenotype was caused by a general polyglutamine effect due to the observation of retinopathy in an HD mouse model. Experiments using this mouse model showed that the disappearance of normal recombinant *atxn7*, when expressed in combination with R7E was not due to recruitment of normal ATXN7 by the expanded polyglutamine protein. Rather, given an associated decrease in normal *atxn7* mRNA levels, the presence of polyglutamine proteins appears to cause transcriptional dysregulation of the *Rho* promoter. Moreover it was shown that expression of the mutant *atxn7* was also reduced; indeed they noted a reduction in expression to minimal levels of 1% after just 9 weeks without alleviation of retinopathy and eventual but not complete clearance of aggregate formation.

Given that CRX is involved in the transcription of photorecepter specific genes such as Rho, and that HD polyglutamine can also suppress Rho promoter activity, this study suggests that retinal dysfunction in SCA7 cannot be completely explained by the ATXN7/CRX interaction reported by La Spada and colleagues (50). It follows that mutant ATXN7 may cause irreversible transcriptional dysregulation, and this in conjunction with the comparatively slow removal of aggregates results in continued disease progression.

The R7E mouse has been used in other studies in comparison to the transcriptional changes that occur in the SCA1 knock-in mouse model. In 2006, a study demonstrated that the neuronal differential programme of photoreceptors in SCA7 and HD mice was altered due to transcriptional alterations of certain rod-specific genes (80). This is in addition to the investigation of the effect of SCA7 pathogenesis leading to altered neuronal fate and function by way of activation of the JNK/c-Jun stress pathway (81). A third study proposed that since there are significant overlaps between SCA1 and 7 in phenotype and pathology, there may be a common pathogenic mechanism contributing to disease (82). Using microarray methods, the authors found an alteration in a signalling pathway, the insulin-

like growth factor (IGF) pathway in a response likely due to neuronal injury. Because this down-regulation occurred in the cerebellar granular layer, it is suggested that non-cell-autonomous degeneration of the Purkinje cells might occur due to this change in the granule neurons. While this is acknowledged not to be the primary cause of disease, the commonality between the two models may represent a common pathogenetic result, a finding important in the context of other polyglutamine disorders.

The second group to create a SCA7 mouse was La Spada and colleagues (50). Using a neuron-specific promoter, they were able to create transgenic mice expressing the full-length human *atxn7* containing either 24 (Prp-SCA7-c24Q) or 92 (Prp-SCA7-c92Q) repeats directed specifically to allow high levels of expression in neurons. The authors noted that the mice showed ATXN7 NIIs in the retina and demonstrated a specific con-rod dystrophy phenotype as well as a complex neurological phenotype including gait ataxia and involuntary movements. Progressive thinning of the ONL with significant loss of cone photoreceptors was observed, as well as activation of Muller cells (the retinal counterparts of the astrocytes in the central nervous system (CNS)) indicating that some injury was taking place resulting in dysfunctional neurotransmission in the retina. Positive TUNEL staining indicated that apoptosis may also contribute towards the reported ONL thinning. In summary, it is suggested that the loss of photoreceptor cells cannot completely account for this severe retinal phenotype, and hence it is suggested that *cellular dysfunction* is the primary cause of the complete blindness in these SCA7 mice.

The neurological examinations of these mice (70) indicated little to no expression of ATXN7 protein in the Purkinje cell layer although these cells were seen to degenerate rapidly. This suggests a possible non-cell-autonomous degeneration whereby deficits in neighbouring cells result in degeneration of the Purkinje cell layer (Figure 1.3, page 21). Furthermore, the study showed evidence for truncated N-terminal fragment of ATXN7

present in the NIIs. Interestingly, the appearance of these NIIs coincided with the presentation of ataxic gait symptoms in the mice indicating a possible pathogenic role for the nuclear localisation of a truncated fragment.

Given that the cell bodies of Bergmann glia surround Purkinje cells, the aforementioned transgenic mice were manipulated in order to express the full-length mutant and wild-type human *atxn7* genes only in the former cells (83). Remarkably, these mice developed ataxia and neurodegeneration. Since Bergmann glia are known to remove glutamate from Purkinje cells, the authors investigated the effects of GLAST (the Bergmann glia-specific glutamate transporter). A significant reduction in GLAST was noted in mutant mice, which impaired glutamate transport in Bergmann glia cells, cerebellar slices and cerebellar synaptosomes. In addition, the Purkinje cell degeneration that was evident suggested excitotoxic injury. This study has greatly contributed to the understanding of the degeneration of Purkinje cells – the most severely affected cell type in SCA7.

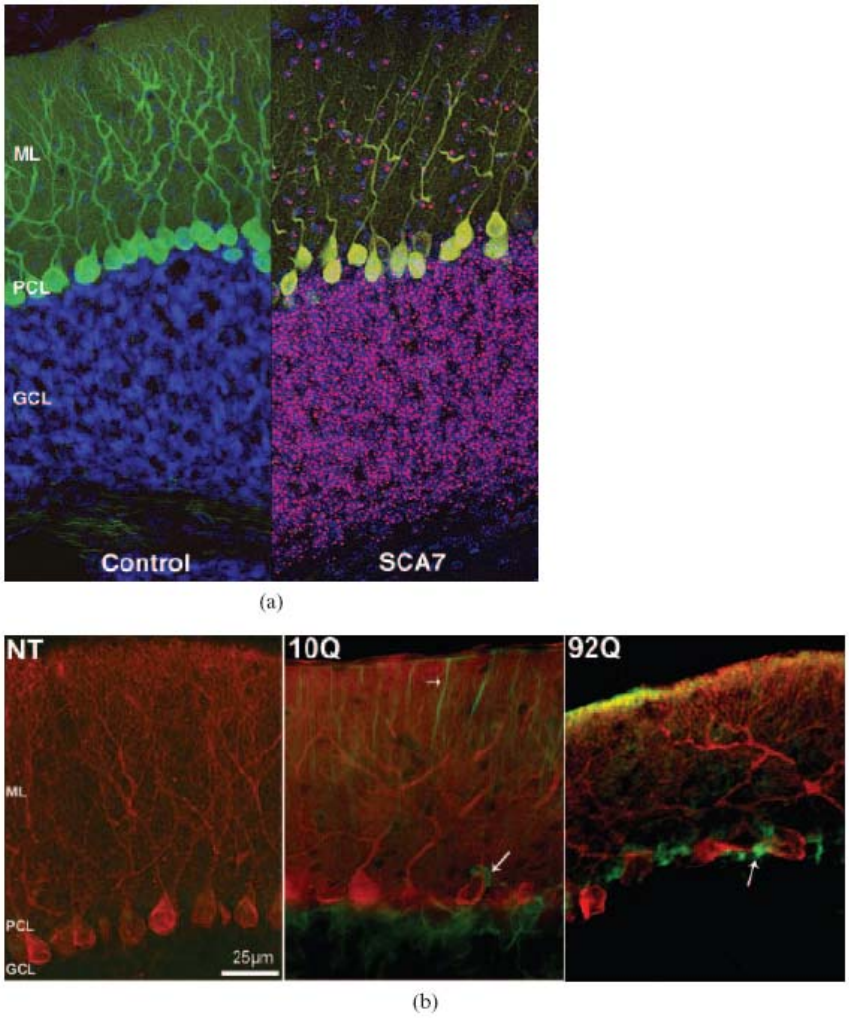


Figure 1.3: Purkinje cell degeneration in SCA7. (a) Immunohistochemistry with anti-ATXN7 (magenta), calbindin (green) and DAPI (blue) antibodies. Note the reduced dendritic arborisation and displaced Purkinje cells in the SCA7 mouse compared to the control. Neurons in the granule cell layer (GCL) and the molecular layer (ML) contain aggregates of ATXN7, while mutant ATXN7 is not present in degenerating Purkingje cells; thus indicating a non cell-autonomous process. (b) Expression of mutant ATXN7 is sufficient to induce Purkinje cell degeneration. Immunohistochemistry analysis performed using anti-ATXN7 (green) and anti-calbindin (red) antibodies. Note that both non-transgenic (NT) mice and mice transgenic for *atxn7* containing only 10 repeats (10Q) have normal cerebellar cytoarchitecture; whereas mice transfected with atxn7 containing 92 repeats (92Q) have reduced ML thickness, misaligned Purkinje cell layer (PCL) and tortuous Purkinje cell dendrites. ATXN7 expression is diffuse in 10Q PCL, whereas it becomes punctate in the 92Q PCL (arrows). From Garden and La Spada 2007 (11).

The third group to create a SCA7 mouse is that led by Huda Zoghbi. The only knock-in mouse created thus far, incorporated 266 CAG repeats in the murine Sca7 gene, which normally contains 5 repeats (84). Repeat sizes over 200 generally result in infantile onset of SCA7 symptoms and the $Sca7^{266Q/5Q}$ mouse recapitulates many of the human infantile symptoms, a broad, rapidly progressive and severe form of the late onset disorder. The progressive cone-rod dystrophy phenotype includes a decrease in several retinal-specific genes; indeed, Rho knockout mice have a similar retinal pathology (85). However cell loss is not as extensive as has been reported in other transgenic models (50, 73). Expression analysis indicated that not all downregulated genes corresponded with those affected by CRX removal, further evidence that CRX may not be the sole factor involved in the specific retinal degeneration. Further experiments reiterated the stabilisation of - and thus inability of the cellular machinery to degrade - the mutant protein also observed by Yvert and colleagues (74). Finally the study indicates that the NIIs are unlikely to be required for pathogenesis, but rather observations of accumulated mutant protein tagged by ubiquitin and Hsp70, and precursors of NIIs, correspond with increased neuronal sensitivity and death.

Using the same mouse model Bowman and colleagues examined the impairment of the UPS in vulnerable neuronal cell populations because aggregates had previously been shown to be resistant to degradation (60). The $Sca7^{266Q/5Q}$ mouse was crossed with a well documented UPS reporter mouse. High levels of the reporter gene (fused to green fluorescent protein; GFP) in the latter model indicated a dysfunctional UPS, whilst low levels presented functional UPS by way of the high degradation of the reporter protein. Although high levels of GFP were noted in the adapted knock-in model, these levels were shown to correlate with a dramatic increase in mRNA, thus exempting a dysfunctional UPS. Interestingly, the increase in transcription of the gene correlated with vulnerable neurons and the expression was shown to be directly proportional to cell specific neuropathology. Furthermore, it was demonstrated that expression of this

reporter of neuropathology was inversely related to NII formation strengthening the hypothesis that NIIs are neuroprotective, until such time as the toxicity overwhelms the cells and causes cell death.

1.1.7.2. Drosophila model

Latouche and colleagues assessed a drosophila SCA7 model in which they induced expression of expanded *atxn7* only within the neuronal cells (86), in order to distinguish between developmental defects caused by ubiquitous over-expression and specific neuronal effects. The dual aims of the study were to identify potential modifying genes affecting the SCA7 phenotype, as well as to assess the effect of inducing repression of the toxic protein in order to determine whether a normal phenotype could be restored. To allow for quicker and increased levels of expression, the study used a truncated form of *atxn7* in the fly model. The model demonstrated a reduced lifespan as well as locomotor dysfunction and limited cell death. Due to the latter, the authors suggested that the formation of NIIs (which co-immnoprecipitated with the same proteins previously linked to NIIs) may be the cause of the cell dysfunction. Removal of induction of the expanded ATXN7 protein, showed a significant time dependent increase in the life span of the flies as well as a decrease in the number of NIIs. Although it was removed after day two, this apparent clearance of NIIs only occurred after 14 to 20 days after induction, and although these flies died later than their continuously expressing counterparts and showed significant improvement in motor function, the data reiterates Helmlinger's study (72) that irreversible pathogenic events had taken place.

While this study addressed many aspects of the mechanism of degeneration, it should be noted that a model expressing a truncated form of a mutant polyglutamine protein yields only the ability to assess the effects of a toxic gain-of-function effect. The most accurate model of any disorder would be that of a transgenic model mimicking the natural situation and thus allowing for the investigation of not only a likely toxic gain-of-function, but also the

effects of a possible loss-of-function due to alteration of the protein's secondary structure as is the case with the Zoghbi mouse model (84).

1.1.8. SCA7 in South Africa

Globally, SCA7 represents one of the rarer polyglutamine disorders with a prevalence of less than 1:100 000, in comparison to the other dominantly inherited ataxias (87, 88). However the situation in South Africa is different, such that SCA7 constitutes one of the more prevalent forms of the disorders (Figure 1.4), as is the case in Scandinavia (89). While it remains a rare disorder in comparison to infectious diseases, the SCA7 situation in South Africa has further unique attributes in that it is only found in the Black African population group (Figure 1.4) (90). This is in contrast to the rest of the world where most SCA7 patients reported are of Caucasian, Arabic or North African (non-Black) descent (24). Only one family of black African descent outside of South Africa has been reported (13). The allele sizes of South African SCA7 patients range from 6 – 15 in the normal range and 42 – 92 in the expanded range in line with those documented globally (90). The mean age of onset in South Africa is 20.3 years (S.D. = 11.84, *r* = 5–39 years) (90).

Following the study by Bryer and colleagues, a second study in South Africa indicated that all the families with SCA7 shared a common ancestor (91). The presence of founder effects in South Africa is not uncommon, indeed, the variegate porphyria gene is considered one of the classic founder effects (92). Recently, another polyglutamine gene, which had previously been hypothesized to have a founder mutation in South Africa, was proved using molecular marker analysis (93). However these founder effects were identified in the Cape, a region known to have a significant sudden migrant influx in the 1600's thus explaining the presence of the common ancestor theory. In contrast, the founder effect identified for SCA7 showed that the families were from different geographical origins within the country (91). Although no further explanation of the founder effect is available at present, it

remains that all the patient samples available in the UCT database to date share the same segment of DNA around the SCA7 mutation.

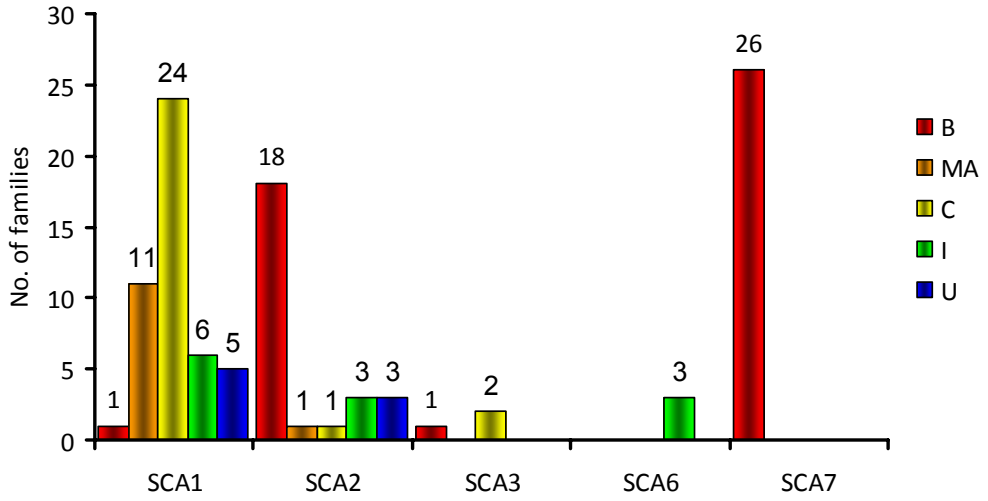


Figure 1.4: Ethnic distribution of the families in SA that tested positive for one of the five SCA expansions tested for at the UCT/NHLS laboratory in January, 2007. B = black African, MA = Mixed ancestry, C = Caucasian. I = Indian, U = Unknown. Data compiled using the University of Cape Town Human Genetics Laboratory database from de-identified and coded samples.

1.2. RNAi

RNA interference (RNAi) is a powerful endogenous gene silencing mechanism triggered by double-stranded RNAs (dsRNA) which are present in various forms including small interfering RNAs (siRNAs) and endogenously encoded microRNAs (miRNAs). These small, 19 – 25 nucleotide (nt) RNAi effectors trigger silencing in a sequence specific manner resulting in knockdown of the target mRNA. While siRNAs have traditionally been exploited for use by investigators to ascertain the functions of specific genes and as a therapeutic method to silence disease genes, miRNAs are a crucial component of the endogenous pathway responsible for the fine-tuning of gene regulation.

1.2.1. History of discovery

While the focus of this thesis is not the endogenous RNAi pathway, the context in which RNAi-based therapy falls requires a thorough understanding of this mechanism. In the years leading up to the discovery of RNAi, a mechanism of 'co-suppression' was identified in plants. Puzzling to researchers at the time, the introduction of transgenes coding for the purple pigment in flowers yielded reduced rather than increased levels of pigmentation (94, 95). This was reiterated in drosophila with the *adh* gene (96). Baulcombe and colleagues then made the important discovery that RNA was the common target inducing a similar mechanism between viral-induced gene silencing and transgene silencing, even suggesting that dsRNA may be required (97). Studies prior to 1998 had shown that antisense RNA could be used to specifically suppress levels of expression of endogenous genes but required high levels of expression themselves in order to be efficient (98). Conversely, others had shown that introducing the sense RNA strand also resulted in suppression (99). It was the Nobel laureates, Andrew Fire and Craig Mello who were able to show that gene-specific knockdown was triggered by dsRNA in animals (*C. elegans*) and that the mechanism was catalytic in nature such that the presence of only a few copies of dsRNA resulted in significant and specific knockdown that was transferred to the first generation progeny (100). As a result of the noted catalytic or amplificatory mode of action, they further suggested that "*Whatever their target, the mechanisms underlying RNA interference probably exist for a biological purpose*." However, the same long dsRNA triggered an antiviral immune response in mammals resulting in non-specific mRNA destruction and inhibition of protein synthesis (101). The suggestion that RNAi might be guided by short 21-25 nt dsRNA species (102) led to the discovery that short 21-23 nt dsRNA fragments could be used to initiate RNAi in mammals (103) and that chemically synthesized siRNAs could be used to trigger RNAi (104), launching the tool of RNAi to the world.

1.2.2. Mechanism of action

The RNAi pathway is triggered by the following mechanisms; exogenous double-stranded RNA, including viral RNA, endogenous dsRNA and selfish genetic elements such as transposons. As such there are many different small non-coding dsRNAs responsible for these different functions. For the purpose of this review, the following discussion will centre on the small dsRNAs relevant to this investigation; small interfering RNAs (siRNAs), microRNAs (miRNAs) and short hairpin RNAs (shRNAs). The differences between these two are vitally important (extensively reviewed in Rana 2007 (105). Figure 1.5 shows an example of the endogenous miRNA pathway coupled with the various points of introduction of exogenous siRNAs.

Primary miRNA (pri-miRNA) transcripts are expressed off RNA polymerase II (Pol II) promoters (although Pol III promoter driven expression has been identified (106)) from various regions of the genomic DNA, occasionally occurring in clusters (105). These transcripts resolve into a folded secondary structure, resembling a hairpin formation with a stem of approximately 33 nt with flanking single-stranded sequences. This relatively conserved general structure is then associated with the microprocessor complex, Drosha-DGCR-8 (DiGeorge critical region-8). DGCR-8 is a co-factor that recognises and binds the pri-miRNA while Drosha (an RNase III enzyme) cleaves the structure approximately 11 bases from the junction of the stem and the flanking ssRNA (107, 108). It is this precise anchoring of DGCR8 at the junction between the flanking ssRNA and dsRNA stem rather than the relative position of the terminal loop that has been shown to determine the site of cleavage by Drosha (107).

The resultant precursor miRNA (pre-miRNA) with a two nt overhang is transported across the nuclear membrane by exportin-5 (109, 110). At this point, a second RNase III enzyme, Dicer cleaves the loop approximately 22 nts from the base of the stem, resulting in a small RNA duplex with characteristic two nucleotide overhangs on the 3' ends (111). The exact cleavage is thought to occur due to the Piwi Argonaute Zwille (PAZ) domain

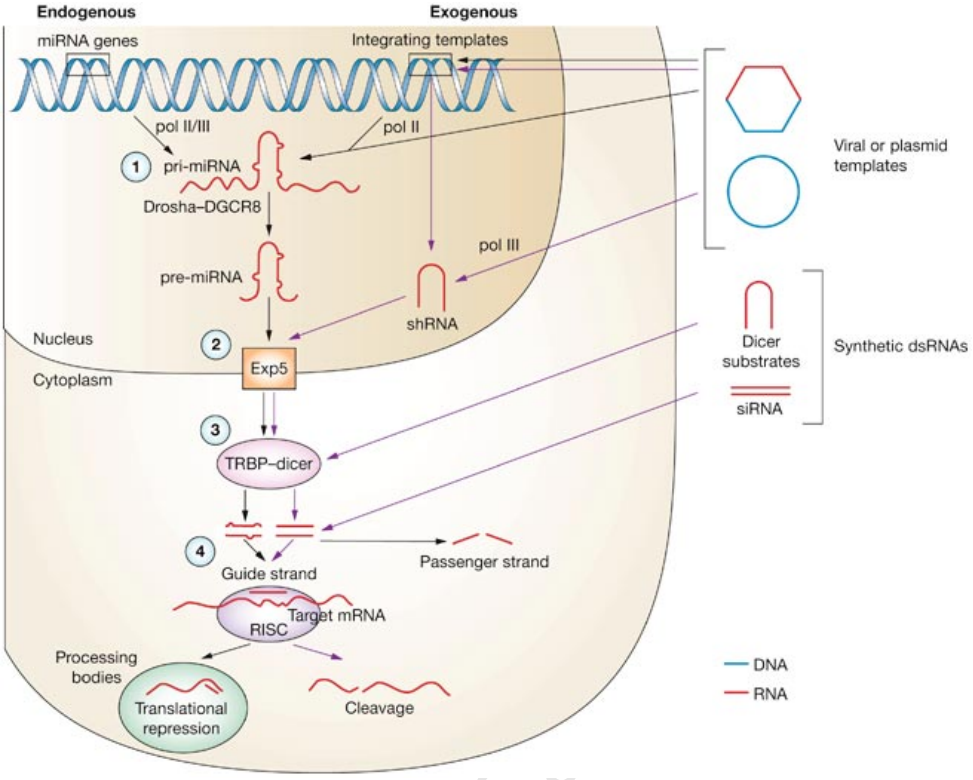


Figure 1.5: Various entry points of manipulation of RNAi in mammalian cells. The diagram shows how to manipulate the RNAi pathway in order to achieve knock down. 1. Viral or DNA plasmid vectors can be introduced into the cell nucleus to result in stable expression of hairpin structures. These can be designed to mimic pri-miRNA based hairpins which are processed by the Drosha-DGCR8 complex into pre-miRNA based hairpins expressed off Pol II promoter. Alternatively, shRNAs can be designed to resemble the pre-miRNA structure and are usually expressed off Pol III promoters. 2. Processed structures are exported into the cytoplasm and 3. processed by TRBP-dicer to yield siRNAs, which can alternatively be directly introduced into the cell. 4. The guide strand is incorporated by RISC and leads to translational repression or target mRNA cleavage. Adapted from Gonzalez-Allegre and Paulson 2007 (112).

of Dicer acting as a molecular ruler, binding the 3' overhang, resulting in the placement of the pre-miRNA stem at the precise position of cleavage (113). The mature miRNA/siRNA is then incorporated by the RNA-induced silencing complex (RISC) (114, 115). In humans, RISC is composed of Dicer, Argonaute proteins, TAR RNA binding protein (TRBP) and a dsRNA binding protein (105). The complex incorporates the miRNA/siRNA duplex and

releases one of the strands (referred to as the 'passenger' or antiguide strand) and retains the guide (or sense) strand (116). The functional asymmetry of the duplex determines which strand is incorporated into the RISC loading complex (RLC) whereby the strand with the more unstable 5' end is selectively retained, referred to as the "*characteristic thermodynamic signature*" of the duplex (117, 118).

The suggestion by Schwarz and colleagues that degradation of the passenger strand rather than separation of the duplex occurs was confirmed using drosophila and human cells (119, 120). Indeed, Matranga and colleagues suggest a model whereby Argonaute 2 (Ago2) receives the siRNA duplex, and cleaves the passenger strand, releasing the guide for formation of mature RISC (siRISC) (119). This cleavage has been confirmed in other studies where it has been shown that modification of the passenger strand can prevent the formation of a functional RISC, and that in addition, the cleavage occurs at position 10 of the passenger strand (121). This process is altered in mature miRNA RISC (miRISC) formation due to the characteristic mismatches between an miRNA guide strand and its passenger counterpart; such that the mismatches prevent Ago2 from cleaving the passenger strand, which instead is removed by a slower bypass mechanism not requiring Ago2 (119).

The mature miRISC/siRISC is then directed in an as yet, unknown manner, to a complimentary mRNA transcript, where RISC facilitates binding of the guide to the mRNA leading to translational inhibition or cleavage of the target mRNA (extensively reviewed in (122)). It is known however, that both sequence specificity as well as target site accessibility play an important role in the ability of activated RISC to exert knockdown. In the first instance, it has been shown that the 'seed' region (considered the first eight nucleotides of the 5' end of the guide strand) is enough to allow for target site recognition by the miRNA (123); and secondly, that sites with significant secondary structure formations are inaccessible and therefore prevent RISC associated knockdown (124).

1.2.3. Comparing siRNAs and miRNAs

One of the distinguishing aspects between siRNAs and miRNAs is that the former are generally perfectly complementary to the mRNA target. This exact match results in cleavage of the mRNA by the Ago2 component of RISC, yielding a transcript recognised by cellular machinery as being aberrant, and hence degraded, preventing translation (122). Ago2 has been referred to as the "*catalytic engine of mammalian RNAi*" (125) and is essential for mammalian development. Indeed, it was shown that Ago2 is distinct from other Ago proteins in being the only species with the ability to cleave mRNA. In one of their most interesting experiments, the authors show that Ago2 knockout mammalian cells are unable to knock down exactly matched targets, but that they can reduce expression of mismatched targets reinforcing the essential role of Ago2 as a component of a cleavage complex, but not in mismatched-target knockdown. Furthermore they showed that murine Ago2 knockouts resulted in embryonic lethality. These observations raise the issue of the importance of target cleavage in mammals, considering the presence of such an evolutionally conserved protein. Indeed it has been shown that miRNAs fully complementary to their targets, and siRNAs mismatched to their targets can result in cleavage and translational repression respectively (126); indeed miRNA directed cleavage has been reported in mammalian cells (127).

However, this dramatic mechanism of gene knockdown is seldom utilised by the endogenous RNAi pathway in mammals. In this scenario, mature miRNAs accompanied by RISC bind to mismatched mRNA targets. The degree of mismatched nucleotides varies, but it yields a more subtle form of regulation, in that Ago2 is prevented from cleaving the target, but the binding itself prevents the translational machinery from processing the transcript. Some studies have shown that the complex associated with the target, can be escorted to the cells' processing bodies or p-bodies, where the mature miRNA/RISC associated complex can be re-used, as well as maintain translational suppression. Extensive studies have examined the mechanism of translational inhibition and are reviewed in (122).

1.2.4. Manipulation of the endogenous pathway

The purpose of this chapter is to understand how RNAi can be manipulated for therapeutic benefit. As can be seen in Figure 1.5 (page 28) there are several points of entry of the miRNA pathway that can be used to induce a specific gene silencing response, which is required for studying effects of loss-of-function, as well as for use in therapeutic investigations. As has been previously mentioned, Tuschl and colleagues demonstrated that siRNAs could be synthetically produced to exact such a response (104). However, this response is transient; thus a mechanism for the introduction of stable inducers of RNAi was required. Hairpin RNA structures include a stem of 19-21 bases separated by a loop sequence of a few nucleotides, expressed off strong Pol III promoters which can lead to long-term expression of effectors, ideal for many applications in therapeutic approaches (128, 129). Indeed Brummelkamp and colleagues designed these short hairpins to resemble the pre-miRNA *let-7*, with 3' two nucleotide U overhangs, which result from the cleavage by Drosha and the polyT transcription stop signal (128); in addition, due the short nature of the dsRNA species, they are able to evade the mammalian immune response, which had prevented long dsRNAs from allowing a specific RNAi response. Whilst these RNAi effectors have been used with great efficiency to result in strong knockdown an improvement was made in the development of shRNAs While traditional shRNAs mimic pre-miRNAs in structure, modified hairpins were designed to incorporate features of pri-miRNAs (130-132). As a result, they have been shown to be up 12 times more effective (133). Furthermore, the use of Pol II promoters results in pri-miRNA based hairpins being more amenable to regulation e.g. tissue specificity can be achieved (134). A further advantage has been recently revealed in gene therapy studies, whereby traditional shRNAs can over-saturate the endogenous miRNA pathway and therefore cause toxicity in vivo (135, 136). The limiting factor in this over-saturation was shown to be exportin-5 (135) although ongoing studies indicate that due to the sequence specificity of the interaction, Ago2 also appears to be a limiting factor (137). The various points of entry of these different effectors are shown in Figure

1.5 (page 28) and illustrates the complexity of the RNAi pathway, in addition to the importance of knowledge of the endogenous pathway.

1.3. RNAi-based gene therapy for dominantly inherited disorders

Gene therapy has been studied as a potential treatment for inherited and other disorders. However, delivery of large sequences of recombinant DNA, with issues such as efficiency, tissue specificity and long-term expression has slowed the advancement of this sphere of study (138). More recently, the investigation of small nucleic acids in suppression of genes causing a toxic gain-of-function has spurred this field. Small molecules such as antisense oligonucleotides (139), and ribozymes (140) have been investigated therapeutically to reduce gene expression in a sequence specific manner. The most recent addition to the field of therapeutic gene silencing are RNAi effectors. The major advantage of these over the previously mentioned gene suppression molecules lies in that they can be virally delivered for long-term expression off tissue specific promoters.

The potential of RNAi in the treatment of many different disorders was recognised soon after the publication of the Fire and Mello paper (100). However, the manner in which to transfer this sequence specific knockdown to mammalian cells remained an obstacle until Tuschl and colleagues demonstrated that siRNAs could trigger RNAi without inducing an immune response in mammalian cells (141). It then became clear that it would be feasible to treat a broad range of disorders including cancer, infectious disease and genetic conditions, where partial, temporal or complete removal of the disease-causing mRNA would likely result in amelioration of the phenotype (142). Since then, a plethora of studies targeting various diseases using various mechanisms *in vitro* and *in vivo* have been investigated with encouraging results (143). In particular, the group of neurodegenerative diseases lend themselves to the investigation of an RNAi-based therapy.

Given the extensive research on these disorders, many pathogenic genetic (not necessarily inherited) defects have been identified which can be targeted using RNAi. The focus of this chapter, however, will centre on dominantly inherited disorders of neurodegeneration, with special emphasis on the polyglutamine disorders.

Studies have shown success using RNAi-based gene therapies in mouse models of neurodegenerative disorders (reviewed in (112, 144)). However, studies of some dominantly inherited conditions demonstrate that evidence of a loss of wild-type function leads to pathogenic effects due to the requirement of the normal protein in cellular functioning. As such general knockdown of both alleles would likely result in phenotypic abnormalities, thus *allele-specific* silencing of the mutant and not the wild-type transcript may be desirable (145, 146).

Given that the impact of losing SCA7 function is not well understood, the importance of preserving the normal SCA7 allele in a knockdown strategy is unknown. Thus, in order to review the likelihood of this requirement, studies of other diseases in this group have been incorporated in this chapter to infer mechanisms of action of the expanded protein on the wild-type, in addition to the importance of the wild-type protein in cellular function i.e. whether general knockdown of both alleles would be sufficient as an RNAi-based therapy, or whether allele-specific knockdown resulting in the partial expression of wild-type protein is required. The toxic gain-of-function caused by the polyglutamine tract is for the most part common to each and has been reviewed extensively (6). However, additional contributions to disease pathology have been shown to include loss-of-function of the specific wild-type protein; and the importance of this effect requires consideration when designing an allele-specific approach.

1.3.1. Effect of loss of wild-type function:

Ultimately the decision to apply knockdown of both mutant and wild-type versus allele-specific knockdown will depend to some extent on how well the total loss of wild-type function is tolerated. Thus the identification and assessment of wild-type function using *in vivo* studies such as knockout models is invaluable. The function of the normal gene product in polyglutamine disorders is mostly unknown. However, studies that indicate essential function, especially in late development, would be strong indicators of a requirement for allele-specificity. Some of these issues are summarised in Table 1.1.

Table 1.1: Assessment of the importance of the wild-type function in polyglutamine disorders

Polyglutamine disorder*	**Gene**	**Function of the wild-type**	**Knockout model severity**
SCA6	*cacna1α*	Calcium channels involved in neurotransmitter release (147)	Severe phenotype (148)
HD	*htt*	Possibly involved in transcription; trafficking and endocytosis; signalling; metabolism (149), transport of Brain-derived neurotrophic factor (BDNF) and post-Golgi-trafficking (150)	Embryonic lethality (151) Conditional knockout significant phenotypic abnormalities (152)
SCA3	*MJD1/ atxn3*	Unknown	Behavioural changes (153)
SCA1	*atxn1*	Unknown	Mild neurobehavioural abnormalities (154)
SCA7	*atxn7*	Increasing evidence as a co-factor in retinal gene expression (50, 54)	-
SCA17	*TBP*	Transcription initiation (155)	-
SCA2	*atxn2*	Unknown, possible RNA processing (156)	Mild phenotype including increased body weight and fatty acid changes in the liver (157); possible developmental vulnerability

*The disorders have been ordered starting with the strongest candidates for allele-specificity, leading to those which may not require such intervention. Since SBMA is an X-linked disorder, it has not been included in this summary, nor in the text, however loss-of-function has been implicated in SBMA (158). Furthermore, analysis of DRPLA has been excluded due to limited studies to date.

In assessing SCA7 for the requirement of allele-specific silencing, the following must be considered; will reduction of the mutant whilst retaining the remaining single wild-type allele lead to a significantly better phenotype than general knockdown of both alleles? An important distinction made here is that while loss-of-function may not contribute significantly to disease pathogenesis, complete removal of the wild-type allele may lead to detrimental defects, hence the importance of knockout models of disease.

Studies on the function of ATXN7 have been reviewed in this chapter (chapter 1.1.5, page 7). Briefly, while the toxicity caused by the mutant gene in SCA7 is well documented in vivo (50, 72, 84), less evidence for the importance of the wild-type is available in that no knockout models have been generated unlike for the causative genes of SCA1, 2, 3, 6 and HD (Table 1.1). Indeed Dragatis and colleagues have shown that a conditional knockout mouse model of HD results in a neurological phenotype suggesting the importance of the normal allele late in development (152). Conditional knockout models are invaluable, as therapy is likely to occur post-natally, therefore defects due to a lack of the protein during early development can be eliminated. Importantly in HD, evidence of a balanced translocation in a non-symptomatic individual provides the perfect evidence in a 'human model' for the fact that the remaining single allele is sufficient for normal development and function (159). This is the only reported case of a balanced translocation in a polyglutamine gene. In contrast the mild phenotype exhibited by the knockout model of SCA2 suggests that allele-specificity may not be required as the phenotype can be rescued on a low fat diet (157). This supports the idea that while the type of mutation is common to the polyglutamine diseases, the therapeutic approach may not be the same. The interaction of wild-type ATXN7 with the transcriptional activator complex TFTC/STAGA and CRX, both of which are known to regulate the expression of retinal genes suggests a dominant negative effect of the mutant protein on the wild-type leading to the distinct macular degeneration seen in SCA7 (50, 55). Furthermore, as is the case with SCA3 (160, 161) the wild-type protein is recruited by the mutant into inclusions (50) suggesting that the restriction of

normal function may contribute to pathogenesis. This indicates that retention of the wild-type is likely to be desirable.

Even with a knockout model, the impact of total loss-of-function in conjunction with a significant amount of mutant expression prior to treatment cannot be fully ascertained. Such is the case in SCA1 where a neuroprotective role for the wild-type was shown such that transgenic mutant mice lacking the wild-type allele showed an exacerbated phenotype (162). Since null Sca1 mice show no signs of ataxia and have a normal life span (154), it is likely that the wild-type protects against effects of the mutant, and may be required to limit the damage caused by the mutant. Alternatively, the removal of the mutant protein may negate the neuroprotective effects of the wild-type. The distinction between loss-of-function caused by a dominant-negative effect and neuroprotective effects of the wild-type could result in very different approaches to allele-specificity. However, the issue of minimum threshold levels of mutant verse wild-type protein tolerated by the cell can only be answered *in vivo*. With regard to SCA7, although extensive data has not been published, Yoo et al. have commented that mice homozygous for the expanded allele (i.e. lacking the wild-type allele) showed no exacerbated phenotype, suggesting neuroprotection is unlikely to be an issue in SCA7 (84).

Another facet to consider is the importance of the function of the wild-type ATXN7 protein. SCA7 is one of the few polyglutamine disorders for which the function of the wild-type has been partly elucidated and its importance as a transcription factor may indicate its essential cellular function. For example, the wild-type protein in SCA17, TBP, and its function as a part of eukaryotic transcription is well described (163); thus the importance of the protein as a ubiquitous transcription factor may pre-empt the requirement of extensive evidence obtained from knockout models, i.e. it is likely that knockout models would show embryonic lethality. Similarly, the role of ATXN7 in retinal gene transcription and the loss-of-function created by the disruption of the transcriptional complex suggest a knockout model would likely have retinal

defects at least. What this additional impact would have in cerebellar neurons is unknown without any knockout models. Furthermore there is little evidence to suggest a redundant protein may rescue the lost ATXN7, as is likely to be the case in SCA2 with the ATXN2 related protein 1 (A2RP1) (164). Taken together, it is likely that the function of the wild-type protein is required in order to prevent undesired effects caused by general knockdown and that allele-specific knockdown would be beneficial.

1.3.2. RNAi studies in polyglutamine disorders

1.3.2.1. Studies using general knockdown

There have been studies using general RNAi knockdown to treat polyglutamine disorders; specifically SBMA, SCA1 and HD (165-167). Caplen and colleagues (165) were one of the first groups to show sequence specific knockdown in invertebrates and mammalian cells. They demonstrated that by using small dsRNAs, this could be obtained without invoking the non-specific immune response of mammalian cells to dsRNA. Thus they importantly showed evidence for the use of RNAi as a therapeutic reverse genetics tool.

In 2004, this was further by investigated by assessing whether viral vectors expressing hairpins directed against a human transgene could reduce pathology in a SCA1 mouse model (167). After screening hairpins targeting 10 different regions of the full-length myc-tagged *atxn1* sequence, one showed efficient knockdown *in vitro*. Transferring this knockdown to neuronal cells by adopting a viral vector with cytomegalovirus (CMV; a Pol II) instead of an H1 (Pol III) promoter resulted in more efficient expression of the hairpin as well as successful knockdown. Using this effector plasmid *in vivo*, the study noted significantly improved motor performance, with no effects of toxicity cause by the expression of the hairpin. Furthermore, SCA1 specific pathology improved in that thickenings of the molecular layer was noted (thinning of this layer is indicative of SCA1 pathogenesis in humans and mice) to the point where it was indistinguishable from the control. The rapid

expression of the AAV1 vector resulted in a reduction of ATXN1 levels only a week after injection as well as the complete resolution of inclusions – characteristic of SCA1 and other polyglutamine disorders. Notably, accessibility to the target was clearly shown to be an issue in that over 10 constructs had to be tested before an efficacious hairpin could be identified. The following year the same group performed an equivalent study on an HD mouse model (166). This HD mouse contained an N-terminal truncated huntington gene (*htt*) fragment containing the expansion. Various hairpins were tested *in vitro* against the short N-terminal myc-tagged transgenes and full-length sequences. The most efficient hairpin was incorporated within an adeno-associated viral (AAV) vector driven off the Pol II U6 promoter for *in vivo* studies. This showed the same strong sustained hairpin expression and consequent knockdown (up to five months) as the previous study. Benefits included the absence of inclusions reactive for htt compared to abundance in the disease model, corresponding to reduced levels of soluble htt protein and mRNA. Although there was no improvement in the weight loss of the mice (also a characteristic of the human phenotype), significant improvements in stride length and rotarod tests were observed. These studies are important in the study of RNAi-based therapy for polyglutamine disorders in that removal of the mutant protein was shown to result in amelioration of the phenotype.

However, the shRNAs used targeted only the human transcript leaving the endogenous wild-type for the most part, intact; therefore the studies were unable to address the issue of expected knockdown of the wild-type gene in a therapeutic system. Currently, investigations into general but partial RNAi knockdown of murine *htt* in a knock-in HD mouse model are being assessed to create partial reduction (approximately 50%) of both mutant and wild-type protein (168).

1.3.2.2. Therapeutic benefits of allele-specific studies

There are few studies of the polyglutamine disorders which have investigated the phenotypic benefits of allele-specific silencing. One of the critical issues

associated with creating allele-specific silencing for the polyglutamine disorders is that one cannot use the difference in size between the polyglutamine tracts as a distinguishing factor, as most effectors are only 19-24 bp in length resulting in the targeting of all CAG repeat motifs (165, 169). This is in contrast with disorders such as the Swedish mutations in Alzheimers disease (AD) and the small deletion in DYT1 dystonia which allow for relatively simple designs for allele-specific selection (170-172). As such, one needs to identify another distinguishing factor between the two alleles of polyglutamine genes. One way would be to use single nucleotide polymorphisms (SNPs). This was first shown in 2003 against a SNP linked to the mutant allele of *atxn3* (169). To date, phenotypic effects of allele-specific silencing have only been reported in one polyglutamine disorder; using this SNP linked to *atxn3* (169, 173).

The study by Miller and colleagues in 2003 investigated the possibility of discriminating between the two alleles of *atxn3*, *in vitro* (169). Various siRNAs were tested including one targeting the CAG repeat itself in the hope that the expansion might lend to this transcript accessibility not achievable against the shorter repeat in the wild-type target. The results indicated that the repeat length could not be used to discriminate between the two alleles. However, a SNP linked to the mutant gene, was successful, in that many of the siRNAs tested, varying the position of the mismatch within the siRNA, showed significant selective knockdown of the full-length mutant target with minimal effects on the wild-type. This selectivity was maintained in a heterozygous assay, in which both transgenes were expressed, as well using hairpin constructs of the most selective siRNA expressed off a U6 promoter in an adenovirus backbone transfected in a neuronal cell line. That selectivity could be achieved was important. Further investigations used siRNAs targeting the red fluorescent protein (RFP) and GFP tagged reporter genes of the mutant and wild-type transcripts respectively. This experiment demonstrated that the selective removal of the mutant, released the wild-type to a normal dispersed expression from recruitment into aggregates. Although this was not demonstrated with the *atxn3* sequence targeting effectors, it

strongly suggests that this would be an additional phenotypic advantage of allele-specific silencing.

This study has further implications in that if one can identify a SNP which is not only linked to the mutant allele, but also has a significant heterozygosity; presumably a significant number of the individuals with the disorder could theoretically be treated with the same effector. This therapeutic avenue can be especially targeted towards populations in which the disease shows evidence of a founder effect.

A study published just a few months later targeting the same SNP outlined above, indicated similar success in selectivity of siRNAs, as well in a significant decrease in cell death in a neuronal cell line caused by toxicity of the mutant gene (173). However, this study did produce interesting data on secondary structure, indicating that the repeat size (using 22 and 79 repeats) could significantly affect the accessibility of the target to knockdown. Although this was a comparison of normal versus the expanded repeat size, it suggests that large differences in repeat sizes between different expansions may result in differential selectivity, and should be taken into account when assessing the effects in patients with different repeats. This may be particularly pertinent to SCA7 which has the largest difference in repeat sizes. In essence, these studies provide evidence that allele-specific silencing is a feasible and desirable approach for therapy for the polyglutamine disorders which should be further investigated.

The only *in vivo* study using an allele-specific approach in a neurological disorder showed that modest selectivity of the mutant allele delayed onset of disease symptoms in a mouse model of ALS (174). The importance of this study lies in the fact that complete discrimination between the two alleles is not essential for beneficial phenotypes, suggesting that even moderate discrimination of the two alleles may lead to phenotypic advantages. To date there has been no allele-specific study investigated for polyglutamine diseases *in vivo*.

Given the fact that even moderate discrimination is not always feasible for some single nucleotide differences, Kubodera et al. have investigated a different approach of allele-specific repression (175). Using SCA6 as a model system, they were able to show strong knockdown of both mutant and wild-type alleles of SCA6 *in vitro*. The siRNA was co-transfected with a vector expressing a modified version of wild-type *cacna1α*, which was resistant to siRNA repression. This was achieved by incorporating specific synonymous nucleotide changes in the region targeted by the siRNA such that the siRNA guide sequence would not be complementary to and thus would not knockdown the replacement wild-type transcript. This method has several advantages and disadvantages such as bypassing the numerous mutations in genes such as *SOD1* and *presenilin*; however the level of expression of the modified wild-type gene would have to be assessed and carefully controlled, although this may be said to be the case in any gene supplementation therapy. Given that these issues are being assessed in other disorders (176) this method represents a potential alternative strategy to the more canonical allele-specific silencing method.

1.3.2.3. Mechanistic aspects to consider

Initial studies investigated the sequence specificity of the guide strand to the target with opposing results. Whereas some results showed single mismatch specificity – the ability of an effector to discriminate between two sequences based on a single nucleotide difference – (177-179), others indicated that single nucleotide changes could be easily accommodated (177, 180) (Table 1.2, page 43). There have been several publications showing effective discrimination based on single nucleotide discrimination against a variety of genes involved in neurodegenerative disorders including the acetylcholine receptor (*AchR)* in slow-channel congenital myasthenic syndrome (SCCMS) (181), the London mutation in AD (170) and the SOD1 mutations that cause ALS (174, 182, 183) and HD (183). Aspects of these studies are included in the summary of all allele-specific studies performed to date in Table 1.2.

In addition several mechanistic issues remain to be considered. Most importantly strand bias, while not a concern exclusive to allele-specific studies, must be addressed. As has been discussed in the previous section, the RISC complex can retain either strand of the siRNA duplex, and this is largely determined by the thermodynamic profile of the duplex (118). It has been shown that RISC incorporates the strand with greater thermodynamic instability at its 5' end. Thus, subtle changes within the 3' end of the antiguide/passenger strand can be introduced to this effect, such as those that produce G:U wobbles with the 5' end of the guide strand, which do not interfere too much with duplex structure (118). The extent of off-target effects which result from undesired strand bias can be measured using simple reporter assays (171). This is an important consideration bearing in mind that off-target effects can result from incorporation of the passenger strand by RISC as even limited sequence similarity (e.g. less than 11 continuous complementary bases) allows for silencing (184). Indeed, off-target effects should be considered for the guide strand as even double mismatched sequences may have an effect on non-specific transcripts (185). Thus it is possible that even if the correct guide strand is incorporated; it may lead to knockdown of other genes with similar sequences.

Table 1.2: Summary of allele-specific studies

Disorder/Gene	**Effector**	**Position tested and the most effective mismatch***	**Type of mismatch to the wild-type**	**Distinctive outcomes**
Notch gene in drosophila (177)	siRNA	Central positions, using single and double mismatches	various	Decreased but not completely abrogated knockdown
Luciferase (178)	siRNA	Various, including multiple mismatches	Various	Knockdown abrogated when the mismatches were placed centrally and in the 3' region of the guide strand.
Human tissue factor (180)	siRNAs	Centrally placed	Purine:purine	Moderate selectivity
GFP (179)	siRNAs/shRNAs	Centrally placed double mismatches	Various	Double mismatches could confer specificity; single mismatches less so
Oncogenic genes (186) (187) (128)	 siRNAs shRNAs	 Position 9 Position 9 Position 11 and 18	 G:U A:G U:C and U:G respectively	First papers to show single nucleotide specificity for therapeutic purposes; Showed that it was feasible to create expression cassettes producing functional siRNAs using viral vectors
ALS (182)	siRNAs shRNAs	Position 9; 10; 11 Position 11	G:A; C:C G:G	Showed that different types of mismatches within the same sequence context and placed at the same position showed different efficacies of knockdown.
SCCMS (181)	shRNAs	Tested positions 12 and 13, 12 most effective	A:C	All positions showed some selectivity.
SCA3, *Tau* (169)	siRNAs/ shRNAs	Positions 11; 16; double mismatches P16P15; latter most effective Positions 14; double mismatches P14P15; P14P11	G:G G:U	Targeting the CAG repeat had no effect Targeted a linked SNP; Double mismatch most efficient
Tau (188)	shRNAs	'Centrally placed'	G:U	Noted difficulty with weaker mismatches

SCA3, *atxn3* (173)	siRNAs	Positions 6; 9; 12; and 15 with 9 the most effective	G:G	Targeted the same linked SNP as above; Found sequence independent knockdown likely due to 2o structure causing disruption of accessibility.
CD46 (189)	siRNAs	Central and 3' positions†	Various	A:C and G:U mismatches (in particular the latter) are well tolerated; Centrally placed positions better for discrimination
AD (170)	shRNAs	Positions 10, 11, 12, 13 with the latter working the best against the London mutation.	G:U	Additional double mismatches (placed in the 3' end) added to an initial centrally placed mismatch may aid discrimination
sickle cell anaemia (190)	siRNAs	Position 10	Multiple	Showed that bulky purine:purine residues were most effective for single nucleotide discrimination
ALS (174)	shRNA	Position 11	G:G	Low selectivity required for significant improvement using *in vivo* phenotype
ALS, *SOD1* HD, *htt* (183)	siRNAs	Positions 1 to 19; with P16 the most effective	G:G G:A	Positions 3' to the centre showed the greatest discrimination, in particular position 16. Showed limited selectivity using a G:U mismatch.
CD46 (185)	siRNAs	Positions 1- 19	Various	Double mismatches better tolerated in the 3' region of the siRNA than the 5' region
pachyonychia congenita (191)	siRNAs	Positions 1 to 19 with P4 and 12 working most efficiently	U:C	Mismatches at positions both 3' and 5' of the centre can result in selectivity

*In the analysis of publications where a different nomenclature has been used, the positions have been changed to correspond with Schwarz et al. 2006, where the position of the mismatch is given according to the first nucleotide of the 5' end of the predicted guide strand. †This is extrapolated because the target site was changed, not the siRNA itself. The allele-specific study on DYT1 Dystonia (171) has not been included because the discrimination is based on a 3 base pair (bp) difference which creates a significantly easier manner of discrimination.

In 2001 and 2002 investigations of the tolerance of mismatches to the target sequences noted abrogation of knockdown in the presence of mismatches (177-180). This led to the hypothesis that it may be possible to distinguish between two alleles for therapeutic purposes. However the rules that govern single base pair allele-specific silencing are not strongly defined. With specific reference to allele-specific studies, the following are important aspects to consider: strength of the mismatch, position of the mismatch and accessibility of the target site.

The strength of the mismatch formed against the wild-type allele can differ resulting in effective or minimal levels of discrimination (183, 190) shown by different mismatches in the same sequence context at the same position yielding different selectivity (182). Those cases where the mismatch conferred against the wild-type is predicted to be a strong destabiliser (such as bulky purine:purine mismatches) will show the greatest discrimination (190). In comparison, weaker mismatches such as purine:pyrimidine. In particular the G:U wobble pair is generally well tolerated (169, 183, 188, 189), although, one study has indicated that strong selectivity can be achieved (186). This can be modified with the addition of secondary mismatches such that the guide has two mismatches against the wild-type but only one against the mutant target (169, 170).

The position of the mismatch relative to the proposed cleavage site of Ago2 is also important in the design, as early studies suggested that given that Ago2 cleavage occurred at centrally at positions 10 and 11 (178), placing the mismatch at this position would yield the most selectivity. Numerous studies have examined centrally placed mismatches (Table 1.2) but these have been shown to not always be the most successful. Of mismatches at positions 16, 15 and 9, a double mismatch of the former was shown to be the most selective (169); with position 13 (170) and positions 12 and 4 (191) also showing the strong selectivity. A recent study has indicated that those placed 3' to the centre of the effector can be more effective (183). This study was particularly important

because it used two identical and consistent screens of siRNAs defining placement of the mismatch exactly and that mismatches placed at position 16 appear to consistently show the greatest selectivity.

Secondary structure of the target region can be an issue for siRNA accessibility (180). The target region containing the SNP should be taken into consideration rather than modified since the presence of the SNP dictates limited mobility of the placement of the effector. Furthermore, since predicting software is limited, it may not be possible to effectively predict the accessibility of a certain region. However this may alter discrimination in polyglutamine disorders since the large expansions may create significant alterations in secondary structure and therefore accessibility.

1. 4. Aims and Objectives

In the same way that allele-specific silencing was used to specifically target a linked SNP in SCA3 (169), it is proposed that SCA7 too would benefit from such an approach. The function of the wild-type protein in SCA7, although not definitive, indicates that it is likely to be essential for cellular function. Thus a method of retaining the normal protein is desirable. In South Africa, it has been shown that a founder effect exists in the cohort of SCA7 patients (91). As such, there is a well-characterised SNP (rs3774729) located in exon 12 of *atxn7*, linked to the mutant allele in 100% of the patients studied to date. In addition, over 50% of these patients are heterozygous for this SNP, which provides a single nucleotide disparity with which to discriminate between the wild-type and mutant alleles, containing the G and A alleles respectively. Finally, SCA7 is the only polyglutamine disorder that has a retinal degenerative phenotype in addition to the progressive neurological symptoms, thus providing a more accessible tissue to measure the efficacy of the treatment *in vivo*. In addition, age-related macular degeneration is currently being tested as a potential therapeutic target for RNAi (192) and most recently, successful human trials for Leber's congenital amaurosis have been performed (193) providing a therapeutic testing ground for gene therapy based approaches in the eye.

The aim of this project was to identify expressed RNAi effectors in order to specifically suppress expression of the mutant allele of the *atxn-7* gene in heterozygous mammalian cell lines, such that this effector has the potential to be used for an RNAi based gene therapy for SCA7 patients in South Africa.

In order to do this the following objectives were identified:

1. Design stably expressing RNAi effectors incorporating mismatches at multiple positions.
2. Screen these hairpins against the wild-type and mutant targets in a hemizygous cellular assay using short target sequences and identify an shRNA

which selectively reduces mutant protein levels without significantly affecting the levels of the wild-type protein.

2. Screen the aforementioned effectors in a cellular assay which more closely resembles the natural state using full-length targets in both hemizygous and heterozygous systems, and investigate any beneficial effects *in vitro.*

2. Materials and Methods

2.1. RNA isolation

RNA was isolated from mammalian cells using the RNeasy mini kit (*Qiagen*, Southern Cross Biotechnology, Republic of South Africa (RSA)) as per manufacturer's instructions. Following each method the quality and concentration of each sample was determined by spectrophotometry (NanoDrop ND-1000 spectrophotometer, *NanoDrop Technologies*, Inqaba, RSA). Samples indicating an A_{260}/A_{280} of 1.8 to 2.0 were stored at -80C for further experiments.

2.2. PCR

Amplification of the G>A SNP polymerase chain reaction (PCR) was performed according to basic conditions outlined in the Appendix (A1, page 134) with the modification of an annealing temperature (T_A) of 60^{o}C. Primers used for the PCR are indicated in Table A2 in the Appendix (page 134).

2.3. Sequencing

Total PCR product was resolved on a 1% agarose gel at 120 volts for 30 minutes (min). The bands were excised using a scalpel, placed in pre-weighed 1.5 ml tubes (*Eppendorf,* Merck, RSA) and weighed. The DNA was extracted as per the instructions in the QIAquick Gel Extraction kit (*Qiagen,* Southern Cross Biotechnology, RSA). After the clean-up, 5 µl of recovered amplicon was resolved by agarose gel electrophoresis to confirm the presence of extracted PCR product. The cycle sequencing reaction was performed in a total volume of 20 µl containing 5 µl of the PCR product, 8 pmol of primer, 1 µl of the BigDye^{TM} termination mix (*Applied Biosystems,* RSA), 1 x BigDye^{TM} Buffer (*Applied Biosystems,* RSA), made up with distilled H_2O (dH_2O). The cycling conditions were performed as follows on a Hybaid Touchdown Thermocycler:

1 cycle of 96^{o}C – 5 min

30 cycles of 96^{o}C – 30 seconds (sec)
50^{o}C – 15 sec
60^{o}C – 4 min

Unincorporated nucleotides and salts were subsequently removed from sequencing products using columns (*CENTRI-SEP*, Princeton Separations, RSA), as per the manufacturer's instructions. Sequencing analysis was performed using capillary based electrophoresis on the ABI 3100 Genetic Analyzer and analysed using the Sequence Analysis version 3.7 software (*Applied Biosystems,* RSA).

2.4. cDNA synthesis

2.4.1. DNase treatment

RNA was treated with AmbionTurbo DNAse (*Ambion*, UK) to remove remaining genomic and/or plasmid DNA contamination. The reaction was performed using a maximum of 10 µg of RNA; 1x TURBO buffer; 2 units of TURBO DNase made up to a final volume of 50 µl with RNAse-free dH_2O and incubated for 30 min at 37^{o}C. DNase was inactivated using 3µl of DNase Inactivation reagent (Ambion, UK), incubated at room temperature for two to three min, vortexing during incubation. The sample was centrifuged at 10 000 x g for two min, after which the supernatant was transferred to a fresh tube.

2.4.2. cDNA synthesis

cDNA synthesis was performed according to the PrimerDesign protocol (*PrimerDesign*, UK).

2.5. Real time PCR

PCR was performed using custom made primers (*PrimerDesign*, UK). The following reaction conditions were used to perform real time PCR; 10 µl of Precision Mastermix (*PrimerDesign*, UK); 10 pmol of each primer (*PrimerDesign*, UK); 25 ng of cDNA template in a total volume of 20 µl. A

standard curve consisting of the following dilution was made using multiple aliquots of test samples; 100 ng, 50 ng, 25 ng, 10 ng, 5 ng. The cycling conditions used were:

1 cycle of	95^{o}C – 10 min
40 cycles of	95^{o}C – 15 sec
	56^{o}C– 30 sec
	72^{o}C – 30 sec

PCR was performed using an Applied Biosystems 7000. Each sample was quantified in duplicate and a negative control was used in each experiment to control for contamination. Relative expression levels were normalised to GAPDH (Table A2, page 134).

2.6. Effector design

2.6.1. shRNA

2.6.1.1. General design

shRNA effectors were designed based on the original expression vector system by Brummelkamp and colleagues (128). The design included the use of a Pol III promoter which produces small transcripts with a well-defined start site and a five thymidine termination signal. In this study the Pol III promoter from the pTZU6+1 vector was used for small transcript production (194). The effectors were then designed with 19 nt of sequence matching the target sequence representing the sense strand followed by 19 nt of reverse complementary sequence representing the antisense strand. This semi-palindromic sequence was interrupted by a short nine nt spacer or loop sequence. Finally the antisense sequence was followed by five thymidines which results in a two nt 3' overhang because termination site cleavage occurs after the second uracil (195). The favoured thermodynamics of the transcript thus predict the formation of a short hairpin loop structure. Cleavage of this dsRNA product by Dicer has been shown to yield 21 and 22 nt sense and antisense products (128). Furthermore, functional asymmetry of the predicted duplex was induced by

creating G:U mis-pairings in the 3' end of the anti-guide strand as has been shown previously (118) in order to bias guide strand incorporation by RISC.

2.6.1.2. Mismatch placement

Based on the predicted guide strand, mismatches to the wild-type G allele were placed consecutively from position 10 to position 16 where position 1 is the first nt of the guide strand assuming a 21 nt duplex (Figure 2.1). Further shRNA mismatch designs incorporated a secondary mismatch immediately 3' to the primary mismatch (Figure 2.2). Finally, shRNAs with alternative secondary mismatch positions were designed as well as the addition of a tertiary mismatch (Figure 2.3)

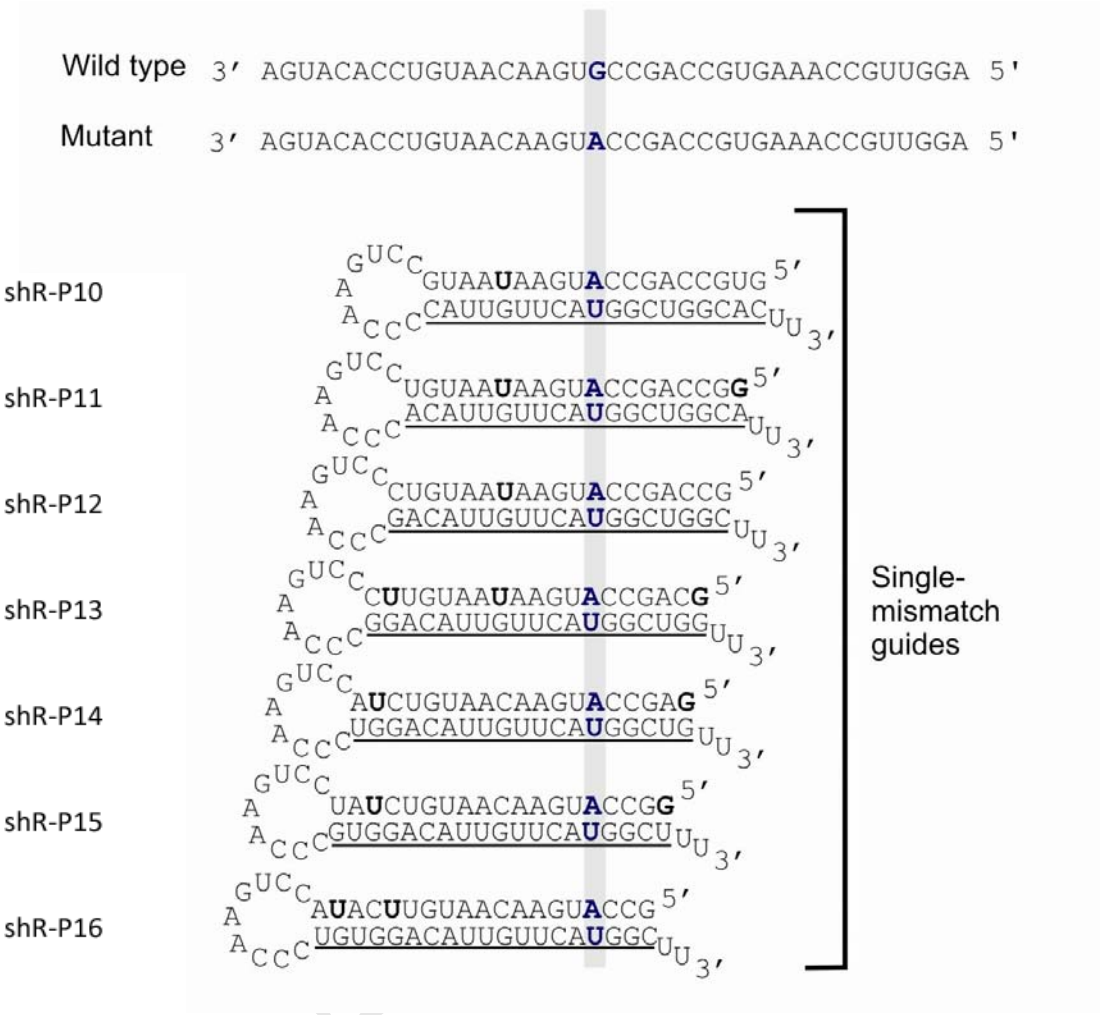


Figure 2.1. Single mismatch guide strands in shRNA format. Convention for the diagram follows Schwarz and colleagues (183). Sequences of expected single mismatch guide strands (underlined) processed from a shRNA format indicating the site of the primary mismatch to the G allele of the wild-type *atxn7* target. Sequence context of both the wild-type and mutant targets are shown. G:U mis-pairings are highlighted in bold. The shRNAs were labelled according to the position of the primary mismatch relative to the 5' end of the guide strand.

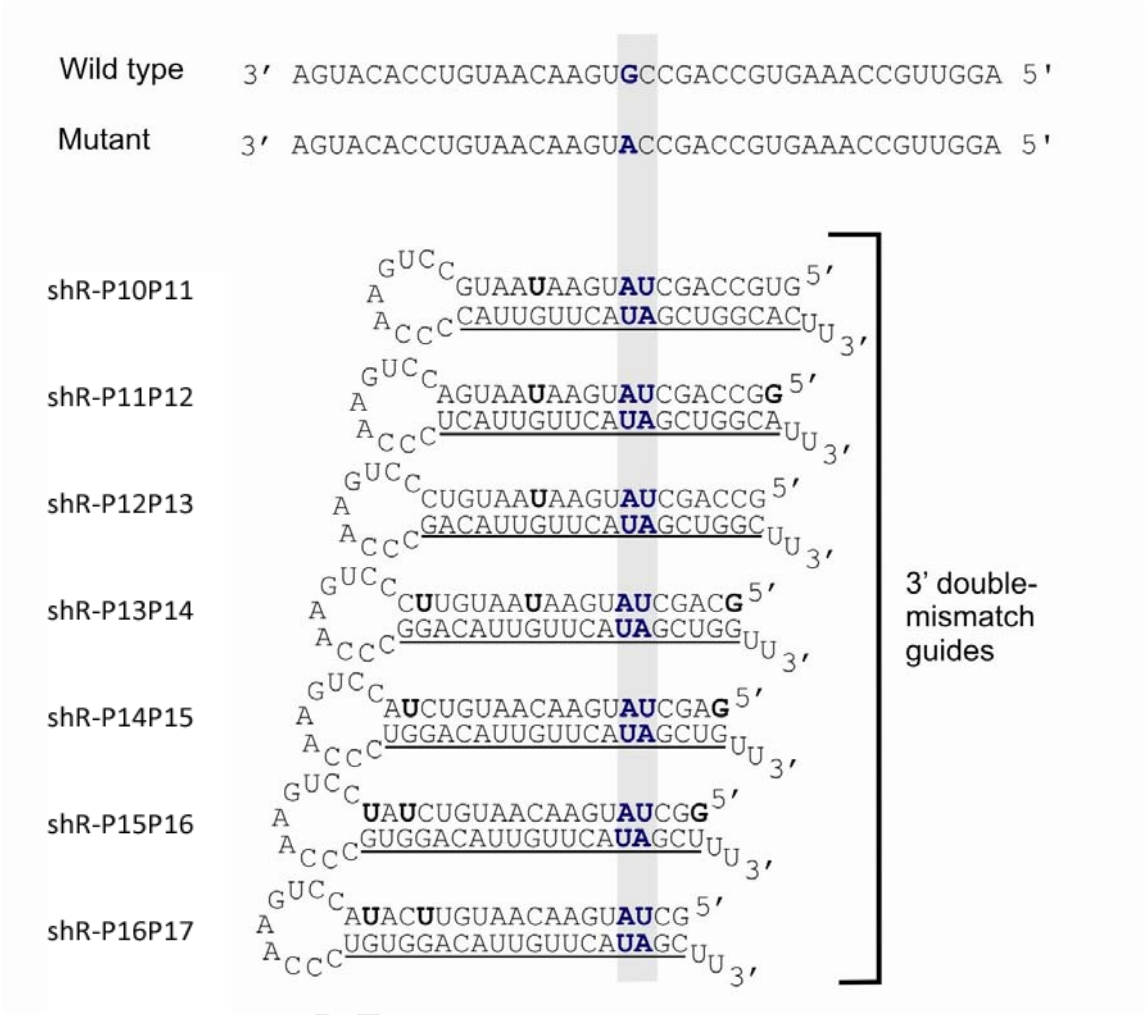


Figure 2.2. Double mismatch guide strands in shRNA format. Convention for the diagram follows Schwarz and colleagues (183). Sequences of expected double mismatch guide strands (underlined) processed from a shRNA format indicating the site of the primary mismatch to the G allele of the wild-type *atxn7* target. Sequence context of both the wild-type and mutant targets are shown. G:U mis-pairings are highlighted in bold. The shRNAs were labelled according to the position of the primary mismatch relative to the 5' end of the guide strand; followed by any additional mismatches.

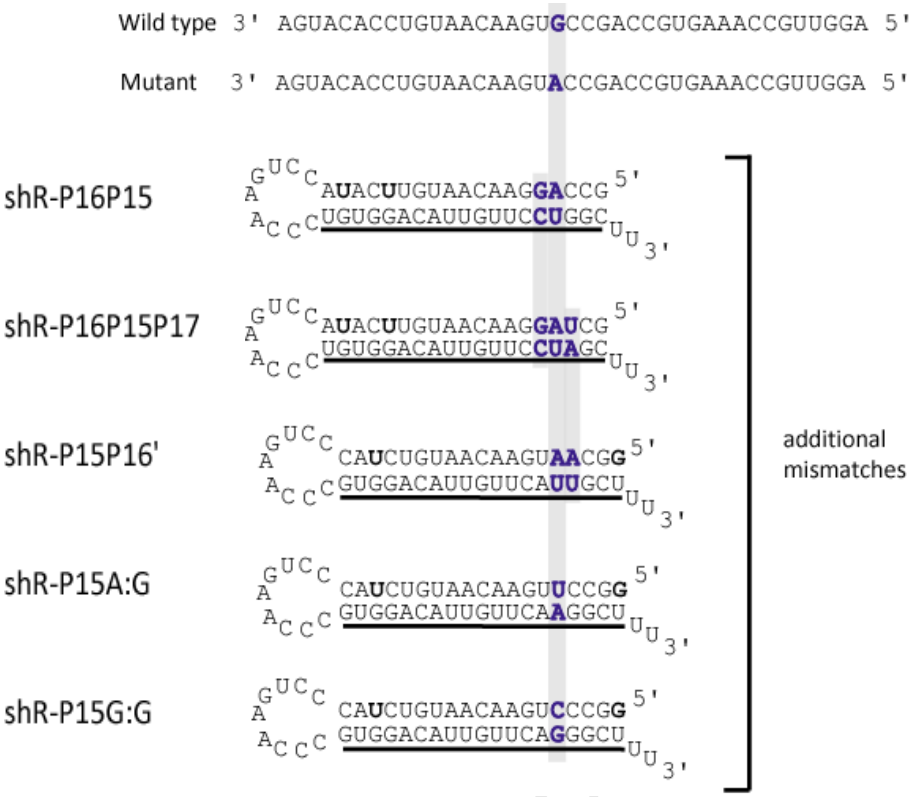


Figure 2.3. Additional mismatch guide strands in shRNA format. Convention for the diagram follows Schwarz and colleagues (183). Sequences of expected double and tertiary mismatch guide strands (underlined) processed from a shRNA format indicating the site of the primary mismatch to the G allele of the wild-type *atxn7* target. Sequence context of both the wild-type and mutant targets are shown. G:U mis-pairings are highlighted in bold. The shRNAs were labelled according to the position of the primary mismatch relative to the 5' end of the guide strand; followed by any additional mismatches.

2.6.2. miRNA

2.6.2.1. general design

The design of the pri-miRNA-based constructs was adapted from Ely and colleagues (196). The reported miR-122 structure from miRBase, http://microrna.sanger.ac.uk/cgi-bin/sequences/mirna_entry.pl?acc=MI0000442, was used as a template for the design of the constructs seen in Figure 2.4 (page 57). The miR-122 guide/antiguide sequences were replaced with 22 nt *atxn7*

sequences without modifying the overall secondary structure of the pri-miRNA and inserted within an exon of a CMV immediate early promoter enhancer expression cassette. The predicted guide sequence was designed to be perfectly complementary to the target unlike most miRNAs and thus should trigger Ago2 directed cleavage of the target. Furthermore, the bulges caused by single mismatches in the stem loop were maintained wherever possible; in terms of both position as well as mismatch strength.

2.6.2.2. Mismatch placement

Given that the pri-miR based hairpins are based on a naturally occurring and experimentally confirmed guide sequence, constructs were designed to place the mismatch at position 16. However, since Dicer processing of shRNA constructs may yield heterogeneous products, variations of the shRNA based P16 construct were accommodated by creating multiple pri-miRNA based hairpins one or two nt up and down of the predicted Drosha and Dicer cleavage sites (Figure 2.4).

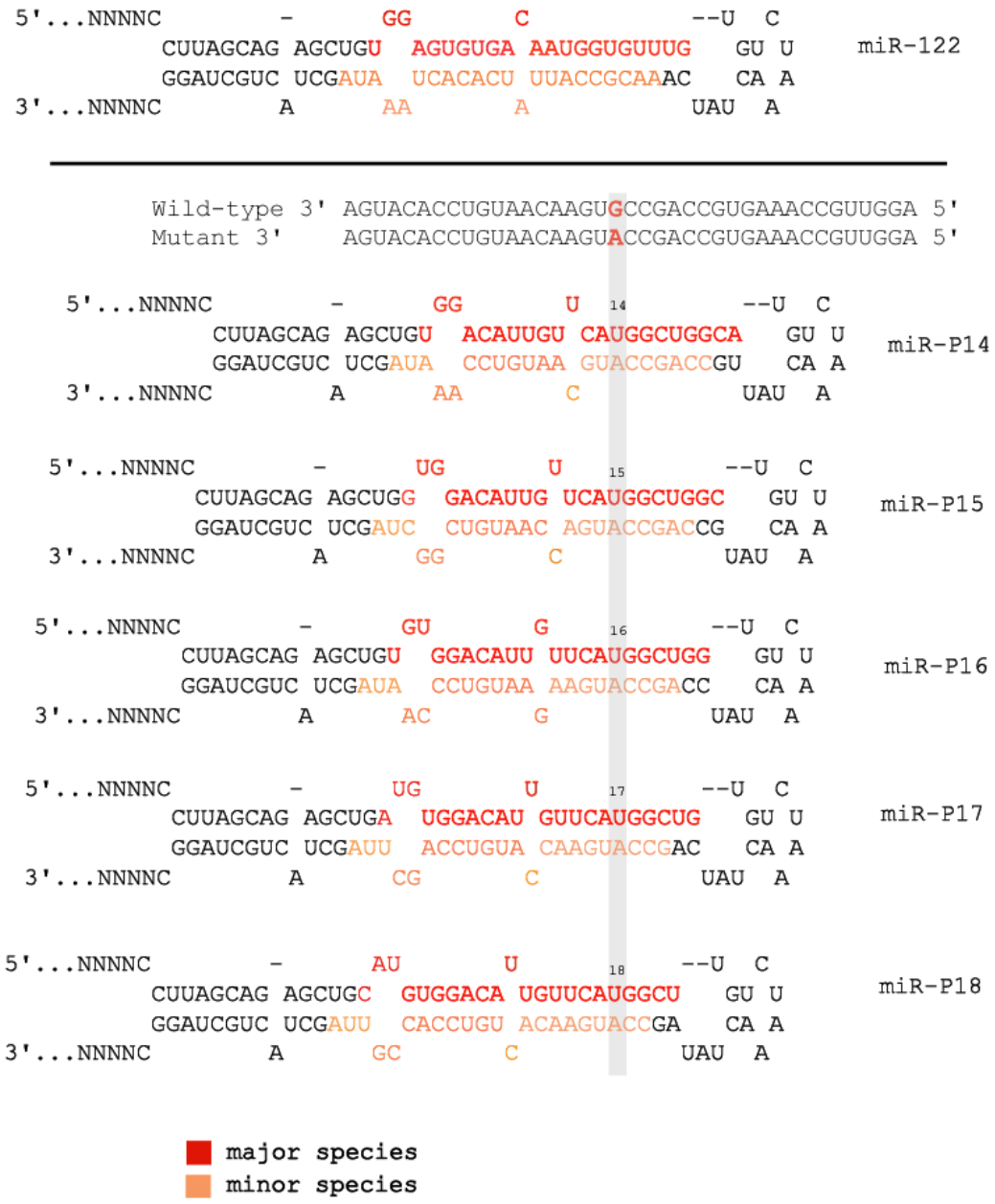


Figure 2.4. Guide strands in miRNA format. Sequences of mismatch guide strands processed from a miRNA format indicating the site of the primary mismatch to the G allele of the wild-type *atxn7* target. Sequence context of both the wild-type and mutant targets are shown. The miRNAs were labelled according to the position of the primary mismatch relative to the 5' end of the guide strand. Major species refers to the strand most likely to be incorporated by RISC, while the minor species indicates the strand less likely to be incorporated.

2.7. Plasmid construction/ cloning

2.7.1. shRNA

Construction of shRNA-expressing vectors has been previously described (197) in which the pTZU6+1 vector (194) is used as a template to create U6/shRNA cassettes. This method was modified by using a single step PCR step. U6/shRNA cassettes (Figure 2.5) were constructed as follows; A PCR cocktail of 200 µg of pTZU6+1 vector, 10 pmol U6 universal forward primer (Table A3, Appendix, page 135), 10 pmol long reverse primer (Table A3), 0.2 mM dNTPs, 1 x Gotaq Buffer (*Promega*, Whitehead Scientific, RSA), 2 units of GoTaq (*Promega*, Whitehead Scientific, RSA), was made up to a final volume of 25 µl with dH_2O. The cycling conditions were the same as those for standard PCR as indicated in the Appendix (A1, page 134). PCR products were electrophoresed on a 2% agarose gel. The expected product size of 300 nt was excised under low Ultra violet (UV) light and cleaned using the *Qiagen* Gel Extraction kit according to manufacturers' instructions (*Qiagen*, Southern Cross Biotechnology, RSA). The PCR products were quantified using a NanoDrop ND-1000 spectrophotometer (*NanoDrop Technologies*, Inqaba, RSA. A total of 100 ng of PCR product was ligated to the TA cloning vector pGEMTeasy (*Promega*, Whitehead Scientific, RSA) according to manufacturers' instructions including ligation and background control reactions, and incubated at 4^oC overnight. Five µl of the ligation product was then used to transform competent DH5α *Esherichia coli* cells and grown on ampicillin selective agar plates. Blue/white screening was used to identify potentially positive clones. White colonies were screened using PCR with T7 and SP6 primers (*Integrated DNA Technologies*, Whitehead Scientific, RSA) according to the standard method in the Appendix (A1). Colonies producing the correct band size were selected and transformed using a miniprep kit (*Qiagen*, Southern Cross Biotechnology, RSA). Plasmid preparations were sequenced using both SP6 and T7 primers to confirm the correct sequence. It should be noted that some of the colonies representing the correct size fragment showed mismatches within the region of the U6/cassette.

Correct sequences were confirmed by manual screening as well as using Bioedit (Tom Hall, Ibis Biosciences, USA). Plasmid integrity was verified by performing either *EcoR*III or *Not*I (*Promega*, Whitehead Scientific, RSA) restriction endonuclease reactions and electrophoresed on a 2% agarose gel.

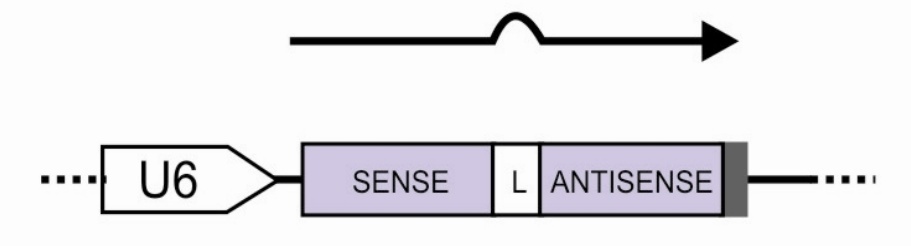


Figure 2.5. Schematic representation of the shRNA cassette. U6/shRNA expression plasmids were constructed using cassettes constructed as shown schematically where L indicates the loop.

2.7.2. miRNA

Pri-miRNA based hairpin constructs (Figure 2.6) were created by performing PCR-based primer extension of partially overlapping complementary long oligonucleotides (Table A4, 132); followed by a secondary PCR using partially flanking outer primers with *Xho*I and *Not*I restriction sites. PCR products were gel excised (*Qiagen*, Southern Cross Biotechnology, RSA), and ligated to the pGEMTeasy TA cloning vector (*Promega*, Whitehead Scientific, RSA). Positive clones were confirmed by sequencing and restriction enzyme digestion. These were then digested with *XhoI* and *NotI* restriction enzymes and ligated to complementary overhangs of the pCI_neo vector (*Promega*, Whitehead Scientific, RSA). Expression plasmids sequences were confirmed as previously described.

Figure 2.6. Schematic representation of the miRNA cassette. U6/shRNA expression plasmids were constructed using cassettes constructed as shown schematically where L indicates the loop.

2.7.3. Construction of luciferase targets

Short target sequences of *atxn7* were cloned into psiCHECK (Promega, *Whitehead Scientific*, RSA) that expresses the *Renilla* luciferase reporter gene (Figure 2.7). Seventy four bp of annealed DNA oligonucleotides harbouring each of the target sequences and *Not*I and *Xho*I sticky ends were ligated into the psiCHECK plasmids upstream of the *Renilla* reporter gene (Oligonucleotide sequences are found in Table A5, 133).

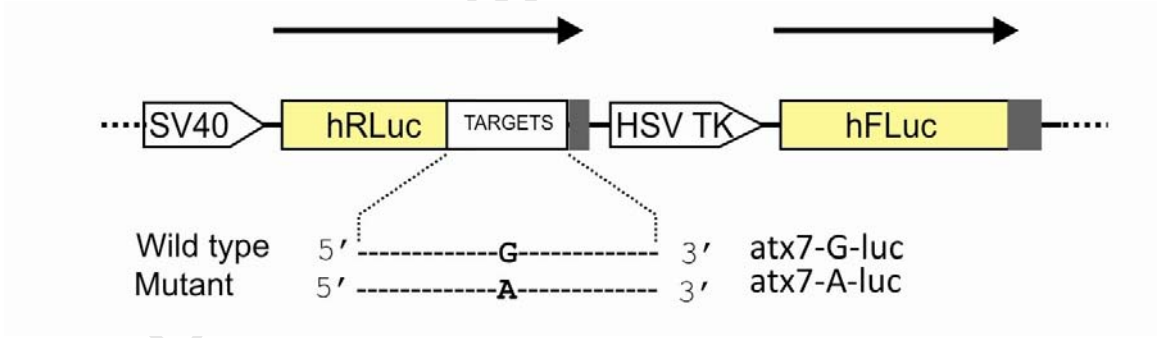


Figure 2.7. Schematic of the dual luciferase reporter targets. atx7-G-luc and atx7-A-luc are the wild-type and mutant reporter siCHECK plasmids used in the luciferase assay, and included a short region of 60bp of *atxn7* target sequence spanning the G>A SNP fused to the *Renilla* luciferase reporter gene.

2.7.4. Construction of full-length targets

Full-length *atxn7* cDNA constructs (wild-type $(CAG)_{10}$ and mutant $(CAG)_{100}$) were fused to the N-terminus of the reporter gene, enhanced green fluorescent protein (eGFP), kindly provided by Alexis Brice (42). Mutant *atxn7* $(CAG)_{100}$ was

modified to replace the G residue of the SNP with the mutant A allele. Long oligonucleotides (Table A5, Appendix, page 137) were annealed and subjected to an initial primer extension with *taq* polymerase; this was followed by a secondary primer extension to produce *Xho*I and *Hind*III sticky ends. The resultant 150 nt dsDNA cassette harbouring the A allele with *Xho*I and *Hind*III ends was cloned into the mutant vector $(CAG)_{100}$ linearised with *Xho*I and *Hind*III. The generated A allele (atx7-100-A-eGFP) was confirmed by DNA sequencing. For the heterozygous assay, the mutant atx7-100-A-eGFP was modified to substitute the eGFP with the red fluorescent protein pdsRED-monomer-N1 vector (*Clontech*, Takara Bio, USA). The dsRED-monomer-N1 plasmid was linearised with *Hpa*I and *Age* and the resulting dsRED fragment was cloned into the atx7-100-A-eGFP vector which had been linearised with the same restriction endonucleases to remove the eGFP gene. The sequence of the mutant vector (atx7-100-A-dsRED) was confirmed. pdsRED-monomer-N1 was chosen because it does not result in aggregate formation as other dsRED protein has been shown to do, therefore any aggregate formation is as a result of the fused target protein (198). Schematic diagrams of the expression cassettes are indicated in Figure 2.8.

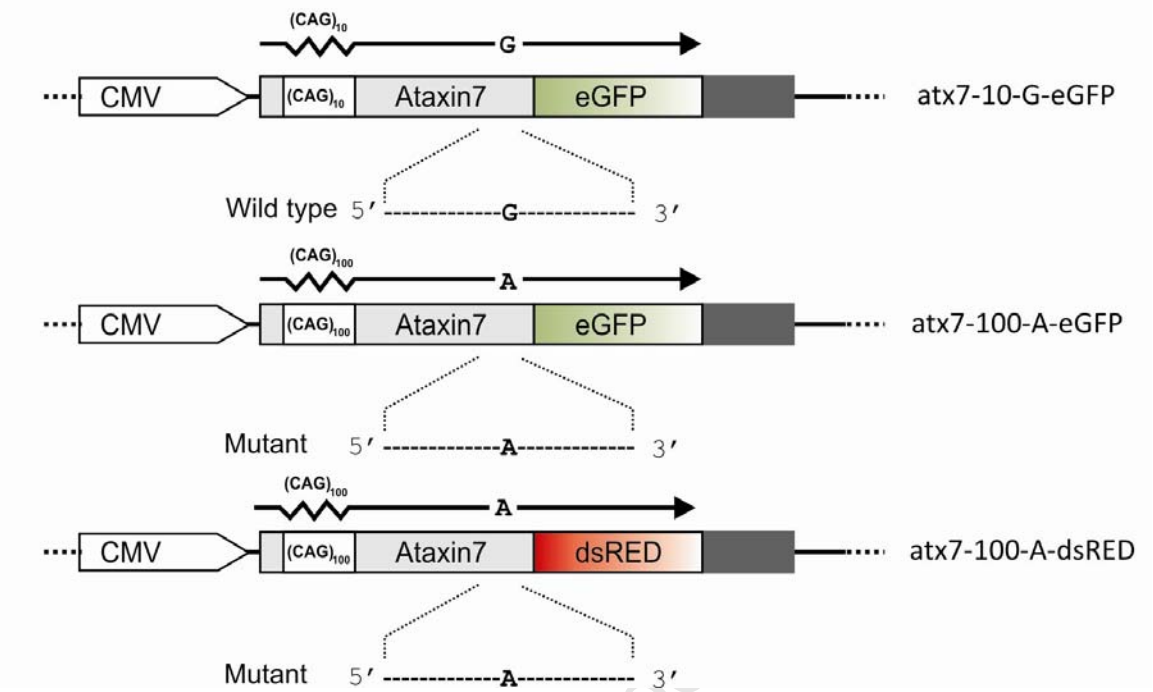


Figure 2.8. Schematic of the full-length target reporter constructs. atx7-10-G-eGFP, atx7-100-A-eGFP, atx7-100-A-dsRED are the 3 reporter plasmids used in the full-length hemi- and heterozygous assays and consist of the full-length *atxn7* cDNA. The wild-type construct (atx7-10-G-eGFP) has a CAG repeat length of 10 incorporating the G allele of the SNP, fused to eGFP. The two mutant constructs (atx7-100-A-eGFP, atx7-100-A-dsRED) have CAG repeat lengths of 100, as well as the A allele of the SNP, and are fused to eGFP and dsRED respectively. (Figure drawn by M. Weinberg)

2.8. Transfections

2.8.1. Maintenance of cell line

All the transfection experiments of this project were performed in HEK293 cell lines. This is a well known and commonly used cell line for artificial *in vitro* experiments (199). Cells were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% Fetal Calf Serum (FCS; *Sigma*, UK); L-glutamine (4 mM; *Sigma*, UK), penicillin (50 U/ml; *Sigma*, UK), and streptomycin (50 mg/ml; *Sigma*, UK). Cells were maintained in a 37^{o}C incubator and medium was replaced every three to five days. Since this cell line had been shown to be

amenable to transfection of the fusion construct plasmids obtained (42), HEK293s were used in all the transfection experiments. In addition, an antibody targeting ATXN7 had indicated no detectable levels of the protein in mock transfected HEK293 cell lines (42). However, a concern regarding the effect of endogenous levels of the ubiquitous *atxn7* transcript seen in other studies (55) was addressed by performing real-time PCR. *Atxn7* levels in cells transfected with full-length transcript plasmids were quantified relative to non-transfected cells. As can be seen in Figure 2.9, there was a 500 fold increase in the former indicating that endogenous levels of the transcript would have insignificant effects on the quantitative nature of the assay.

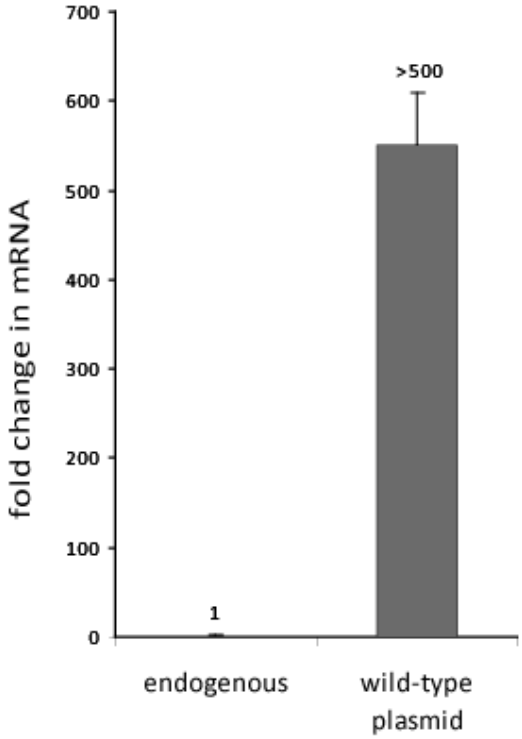


Figure 2.9. Relative abundance of *atxn7* transcript. Fold change in levels of atxn7 between endogenous levels of transcript and transfection of 1 µg of the full-length wild-type target as indicated. Each experiment was performed in triplicate, and the data is normalised to levels of GAPDH. Average ± standard deviation is shown.

2.8.2. Validation of cellular assay

All transfection experiments were normalised to a non-specific control shR/miRNA expression plasmid. In order to confirm that these sequences had no effect on the target sequence, they were compared to expression plasmids containing only the corresponding promoter with no effector sequences. The results confirm that the non-specific effectors have no effect on the target sequences, and thus can be accurately used for normalisation purposes (Figure 2.10). In each assay, miRNA and shRNA effectors were always normalised to miR-NS and shR-NS respectively, in order to take into account any effects of the different promoters and background vector sequence.

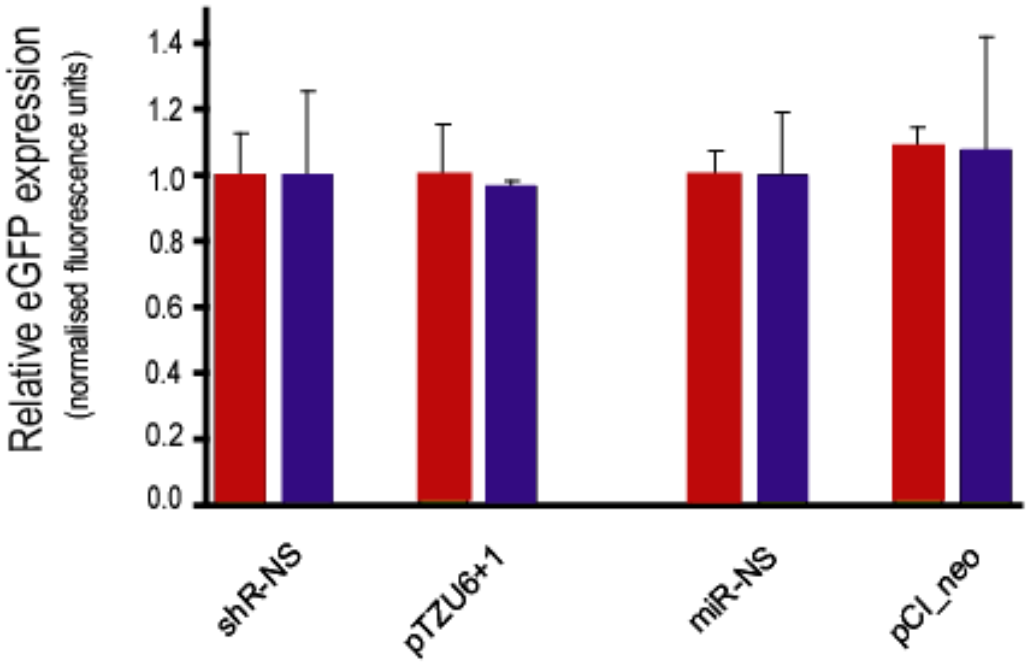


Figure 2.10: Effect of non-specific effectors on the target relative to empty vectors a full-length hemizygous assay. Quantitation of fluorescence in HEK293 cells transfected with wild-type (atx7-10-G-eGFP) or mutant (atx7-100-A-eGFP) expression plasmids and the indicated effector. Each experiment was performed in triplicate and the data is relative to that measured using a non-specific effector, NS. Average ± standard deviation is shown. Wild-type and mutant targets are represented by red and blue bars respectively. pTZU6+1 is the empty shRNA vector, and pCI_neo is the empty miRNA vector.

Further validation of this assay was demonstrated by the similar levels of knockdown achieved using E19 (Figure 2.11; kind gift from Y. Yokobayashi (200)); the eGFP-specific hairpin. In each experiment using the full-length assay, this hairpin was used, and demonstrated that a) knockdown of the transcript was feasible, and b) similar levels of knockdown were achieved against the mutant and wild-type sequences indicating that this assay was an accurate measurement of transfected target knockdown.

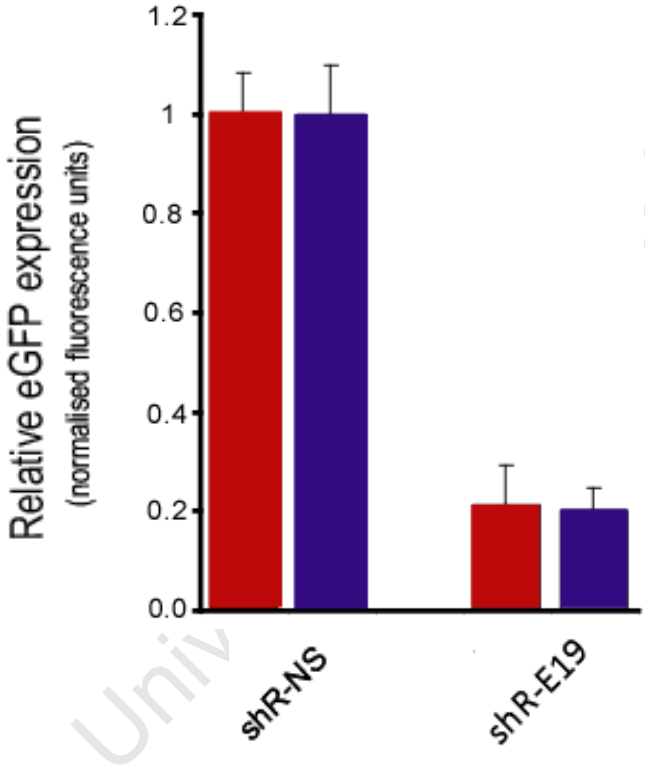


Figure 2.11: Effect of eGFP specific effector on the target in a full-length hemizygous assay. Quantitation of fluorescence in HEK293 cells transfected with wild-type (atx7-10-G-eGFP) or mutant (atx7-100-A-eGFP) expression plasmids and shR-E19. Each experiment was performed in triplicate and the data is relative to that measured using a non-specific effector, NS. Average ± standard deviation is shown. Wild-type and mutant targets are represented by red and blue bars respectively.

2.8.3. Hemizyous assays

Plasmid transfections using the luciferase targets were performed using Lipofectamine 2000 (*Invitrogen*, RSA), according to manufacturer's instructions with a total of 1µg of DNA in a 1:1 ratio of target:shRNA vector. Transfections for the hemizygous full-length target fluorescence assays were performed using jetPEI (*Polyplus*, Autogen Bioclear, UK) with a Nitrogen to Phosphate ratio (N/P) of 5, using a total of 2 µg of DNA in a 1:1 ratio of target:shRNA vector. All hemizygous assays were performed by seeding 1 x 10^5 cells in each well of a 24-well plate.

2.8.4. Heterozygous assays

2.8.4.1. shRNA transfections

The heterozygous assays were performed as above except for the following changes; 6-well plates were seeded with 2.5 x 10^5 cells per well with a total of 6µg of DNA. In cases where target vectors were transfected separately for comparison in the heterozygous experiments, pU6TZ+1 was included to equate the total amount of DNA transfected to that in the co-transfections of both targets.

2.8.4.2. miRNA transfections

The heterozygous assays were performed in a 2:1 ratio of target to effector using 6-well plates and a total of 6µg of DNA. The change in the ratio was performed to overcome problems of promoter occlusion of the target and effector CMV promoters. The combination of two CMV expressing targets as well as a CMV expressing effector (in contrast to the Pol II promoter-expressing shRNAs) likely led to a saturation of Pol III machinery recruited by the various vectors thus resulting in reduced levels of target expression. In order to accurately measure the effectors of the miRNAs, target plasmid DNA was increased to increase levels of reporter expression which could be read by the fluorimeter. In cases where target vectors were transfected separately for

comparison in the heterozygous experiments, miR-NS was included to equate the total amount of DNA transfected to that in the co-transfections of both targets. In addition, the addition of miR-NS allowed for accurate measurement since it contains the same CMV promoter as the other miRNAs.

2.9. Luciferase assay

For the purpose of this experiment, the luciferase assay was used to identify changes in mRNA levels of short targets tagged to luciferase reporter genes by measuring changes in luciferase activity. Using the target constructs shown in Figure 2.7 (page 60), expression of renilla luciferase can used to identify knockdown of the target (Figure 2.12), as well as then normalised to levels of firefly luciferase, resulting in easily replicable results. The activities of *Renilla* and firefly luciferase were measured with the Dual-Glo™ Luciferase Assay System (*Promega*, Whitehead Scientific, South Africa) and using the Veritas dual injection luminometer (*Turner BioSystems*, Sunnyvale, USA).

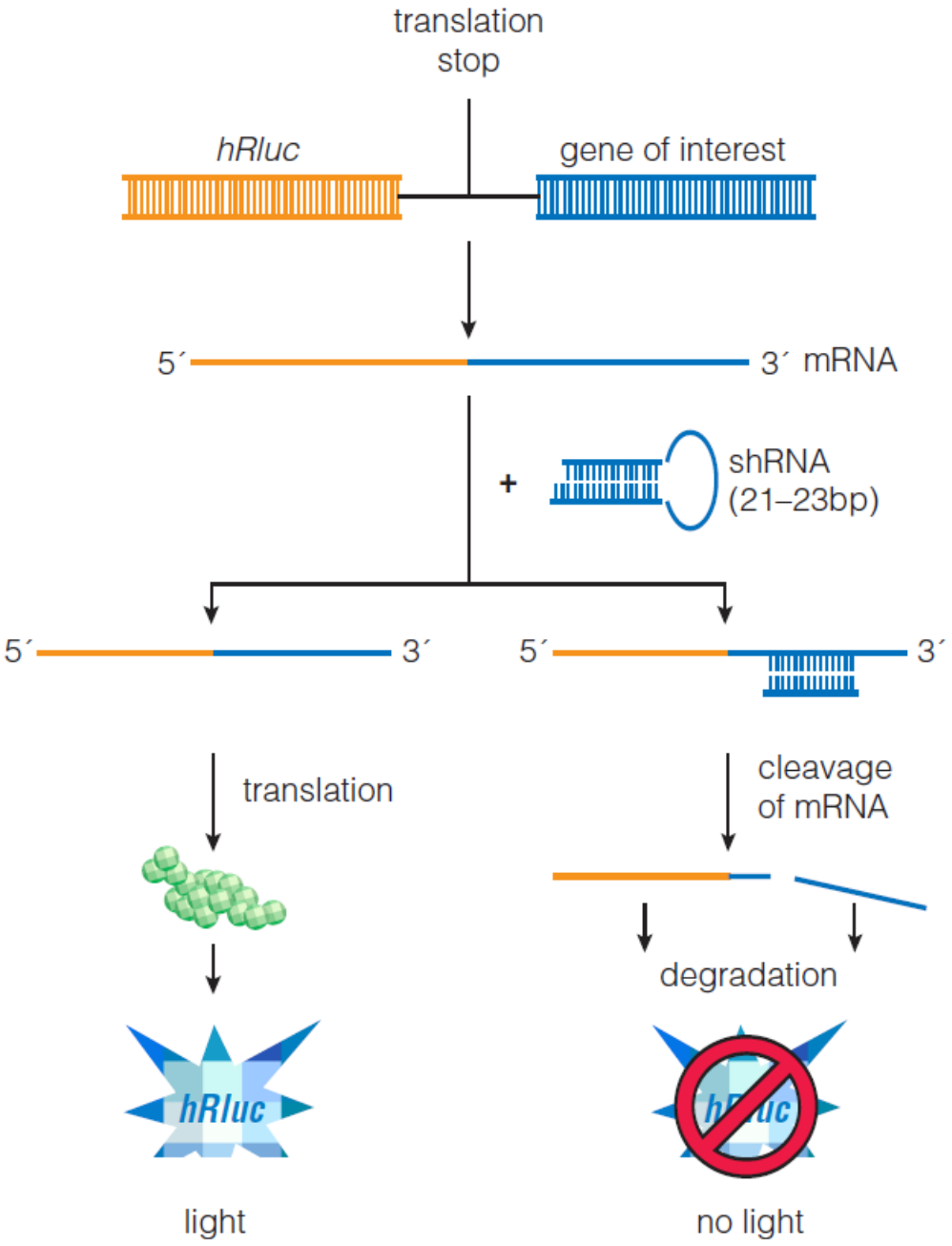


Figure 2.12. Mechanism of luciferase assay using psiCHECK vectors (adapted from Promega, www.promega.com/paguide/images/4339MA10_3A.jpg).

2.10. Fluorescence assay

2.10.1. Protein quantification

All quantitative experiments were performed by normalising to 100 µg/ml of protein using the Micro BCA Protein Kit (*Pierce*, USA).

2.10.2. Fluorimeter quantification

Fluorescence was quantified using the *Optima* software program on an Optima FLUOSTAR fluorimeter. Average values are shown from experiments in biological triplicate with values from indicated U6/shRNA expression plasmids normalized to a miss-targeted shRNA expression plasmid set at 1. Error bars indicate variation between experiments as the standard deviation of the mean. For the hemizygous experiments; equivalent quantities of mock-transfected cells seeded in 24-well plates were used to subtract the background fluorescence. Both eGFP and DsRed have an excitation wavelength 489 nm and 556 nm respectively; and emission of 508 nm and 586 nm respectively (*Clontech*, Takara Bio, USA). In measuring the fluorescence dual assay of eGFP and DsRed an overlap can occur and influence the fluorescent measurements of both. In order to avoid the overlap of green and red fluorescence in a dual hemizygous assay two modifications were performed. Firstly the excitation and emission wavelengths for both proteins were adjusted to create minimal overlap, such that eGFP and dsRed were read at excitation wavelengths of 485 nm and 544 nm respectively; and emissions of 520 nm and 590 nm respectively. Secondly because the overlap cannot be completely removed, equivalent quantities of cells transfected with 0.5 µg atxn7-10-G-eGFP + 1.5 µg of pTZU6+1 were used as the background values whilst reading through the red channel. Conversely, equivalent quantities of cells transfected with 0.5 µg atxn7-100-A-dsRED + 1.5µg of pTZU6+1 were used as the background values whilst reading through the green channel. Hemizygous assays were measured at 48 hours, while heterozygous assays were measured at 72 hours in order to maximise levels of fluorescence measured.

2.11. Aggregate counting assay

Aggregate-containing cells were counted as previously reported (201) with the following modifications. Cells were co-transfected in triplicate as described in the heterozygous assay. Non-overlapping frames of live cells from each well were captured on an AxioVision Inc. microscope using the AxioVision 4.6.3 software after 72 hours. The sum of the total number of cells in three frames was calculated using a 40x objective. Cells with aggregates and those with dispersed patterns of expression were counted as a total of the three different frames of live cells. This was performed in triplicate with the mean representing the average number of that biological triplicate. It is noted that previous studies using the tagged vectors had shown that the reporter genes had no effect on the localisation and expression pattern of the ATXN7 protein (42).

2.12. Confocal microscopy

A Zeiss LSM 510 laser scanning confocal microscope was used. Images were taken with the 40X and 63X oil immersion objective lenses. The following filter sets were used: MBS NT 80/20 or HFT UV 488/543/633 with DBS NFT 490 or NFT 545 or mirror or none. An Argon laser at 488nm was used to excite eGFP, emission collected using 505 -550 nm band pass filter. A Helium-Neon laser was used to excite dsRed at 543 nm, emission collected using band pass 560nm – 615nm. Representative images shown were analysed by the ZEISS LSM 4.1 software.

2.13. Statistical analysis

Statistical analysis utilized two-tailed t-tests, unless otherwise stated. Statistical differences were considered significant when the two-tailed p value < 0.05 using a paired t-test.

3. Results

3.1. Luciferase assay

3.1.1. Single mismatches

In order to identify RNAi effectors targeting the *atxn7* SNP, multiple shRNAs were designed to place the weak G:U mismatch created against the wild-type transcript successively in positions 10 to 16 from the 5' end of the shRNA guide strand (Figure 2.1, page 53). These predicted shRNAs were screened in a dual luciferase reporter assay (section 2.9, page 67) incorporating a short stretch (60 nt) of either the mutant or the wild-type target gene sequence (Figure 2.7, page 60). Mismatches at positions P11-P16 showed statistically significant ($p<0.05$) discrimination (Figure 3.1). Of these shRNAs, shR-P15 showed the greatest discrimination with the wild-type minimally affected at 90% and the mutant knocked down to nearly 55%, relative to a non-specific control. Mismatches placed at position 12 and 14 also showed a significant degree of selectivity with 20 – 25 % discrimination between knockdown of the mutant and wild-type, although this selectivity was achieved with strong knockdown against both the mutant and the target with more than 50% knockdown against both targets. This is in contrast to the selectivity obtained using shR-P15, where neither target was reduced to less than 50%. Finally, hairpins shR-P11, -P13 and -P16 showed significant but minimal discrimination of approximately 10 – 15 %. Selective target knockdown was the least efficient at the central shR-P10 position where no significant selectivity was obtained. Interestingly, shR-P16 which showed little mutant-specific discrimination dramatically knocked down both wild-type and mutant targets indicating that the shRNA was highly tolerant of a single nucleotide mismatch against a short target sequence at this position.

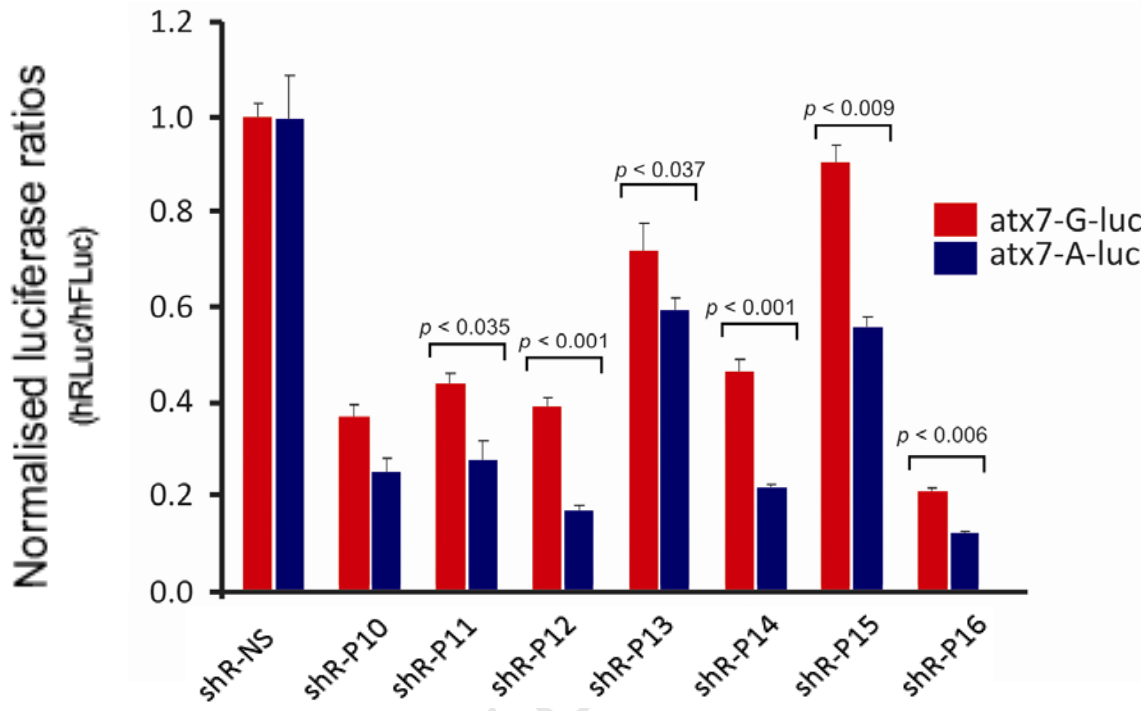


Figure 3.1. Analysis of series of mismatched shRNA guide sequences targeting the G>A SNP in *atxn7* in a dual luciferase assay. Relative levels of R*enilla* luciferase (hRluc) expression normalised to firefly luciferase (hFluc) expression. Each experiment was performed in triplicate and the data is relative to that measured using a non-specific shRNA, NS. Average ± standard deviation is shown. Statistically significant differences ($p<0.05$) between wild-type and mutant silencing are indicated by corresponding p values. Relative expression of wild-type and mutant targets is represented by red and blue bars respectively.

3.1.2. Double mismatches

Given the weak nature of the *atxn7* G:U nucleotide mismatch, efforts were made to enhance the level of wild-type:mutant discrimination by incorporating a secondary mismatch 3' to the primary mismatch (Figure 2.2, page 54). This resulted in a weak A:C mismatch against both the mutant and the wild-type targets in addition to the G:U mismatch against the mutant, thus creating a double mismatch against the mutant in comparison to the single mismatch against the wild-type. The results of the dual luciferase screen showed that discrimination could not be improved beyond that achieved using the single mismatch at shR-P15 (Figure 3.2). The greatest discrimination achieved occurred using the double mismatch shR-P13P14, where enhancement of the 10% selectivity obtained using the single mismatch at position 13 was increased to 25% with the addition of the second mismatch. This was the only position to show an enhancement of selectivity. While significant levels of discrimination were obtained for almost all the other positions, the levels were reduced or remained the same in each case. Furthermore the inclusion of these secondary mismatches resulted in decreased levels of discrimination at almost all positions tested, such that only two hairpins (shR-P10P11 and -P16P17) showed greater than 50% knockdown against both targets. However this abrogation of knockdown was minimal in the case of the double shR-P16P17 mismatch, which maintained the strong 80% knockdown against the wild-type target.

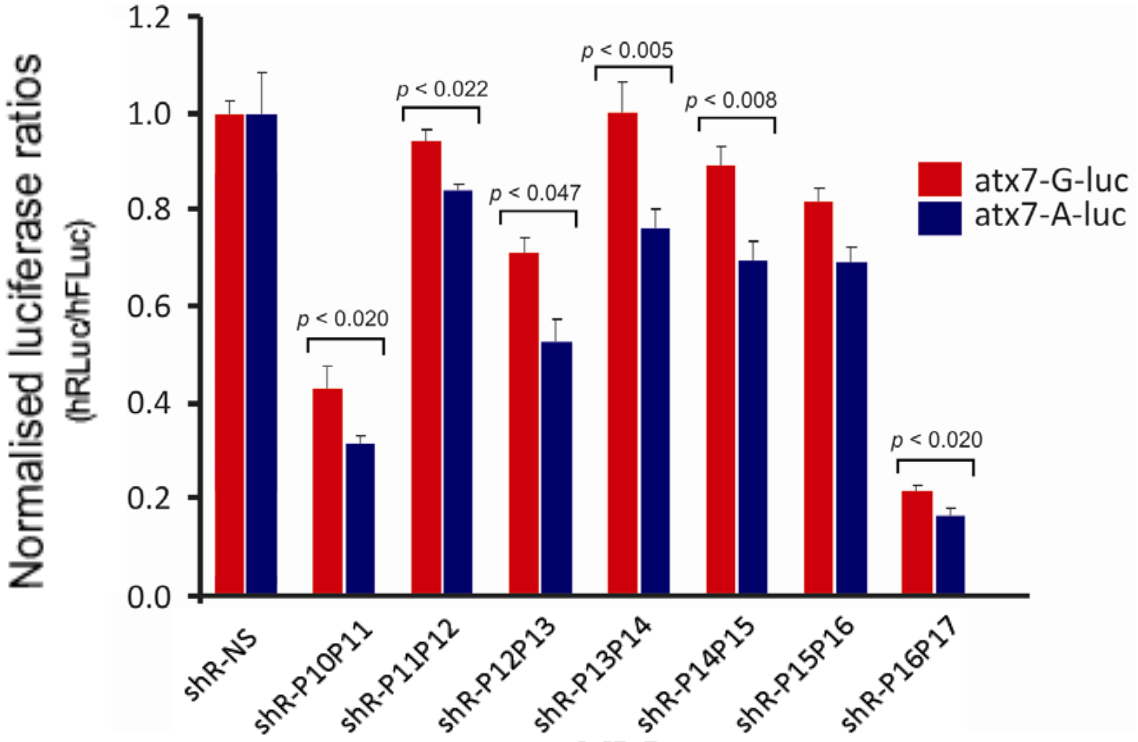


Figure 3.2. Analysis of secondary mismatched shRNA guide sequences targeting the G>A SNP in *atxn7* in a dual luciferase assay. Relative levels of R*enilla* luciferase (hRluc) expression normalized to firefly luciferase (hFluc) expression. Each experiment was performed in triplicate and the data is relative to that measured using a non-specific shRNA, NS. Average ± standard deviation is shown. Statistically significant differences ($p<0.05$) between wild-type and mutant silencing are indicated by corresponding p values. Relative expression of wild-type and mutant targets is represented by red and blue bars respectively.

3.1.3. Additional mismatches

Since the introduction of a double mismatch had been less successful in enhancing the initial discrimination obtained, further modifications were attempted. A secondary mismatch was incorporated immediately 3' to the initial primary mismatch at position 15 as before. However the mismatch itself was designed to create greater destabilisation than the previous A:C combination by yielding a U:C mismatch (shR-P15P16'). Two further modifications were made to the P15 hairpins which including changing the nucleotide at this position from a U to an A or G, changing the nature of the primary mismatch itself. This resulted in an A:G mismatch to the wild-type corresponding with a G:G mismatch to the mutant (shR-P15A:G); and a G:G mismatch against the wild-type corresponding with an A:G mismatch against the mutant target (shR-P15G:G). Although the level of discrimination remained significant for all three of these positions, none led to discrimination greater than 10% and appeared to reduce the ability of the hairpin structures to exert knockdown (Figure 3.3). Further investigation into the tolerance of the mismatch at position 16 was also attempted. A secondary mismatch was placed 5' to the original shR-P16 mismatch yielding a C:U mismatch against both the mutant and the wild-type (shr-P16P15); in addition, a second hairpin was generated by placing additional mismatches on both 3' and 5' to the primary mismatch (shR-P16P15P17) (Figure 2.3, page 55). The results showed that neither hairpin was able to create enhanced discrimination (Figure 3.3). Interestingly, however, both hairpins retained the ability to exert strong knockdown approaching 80% against both targets.

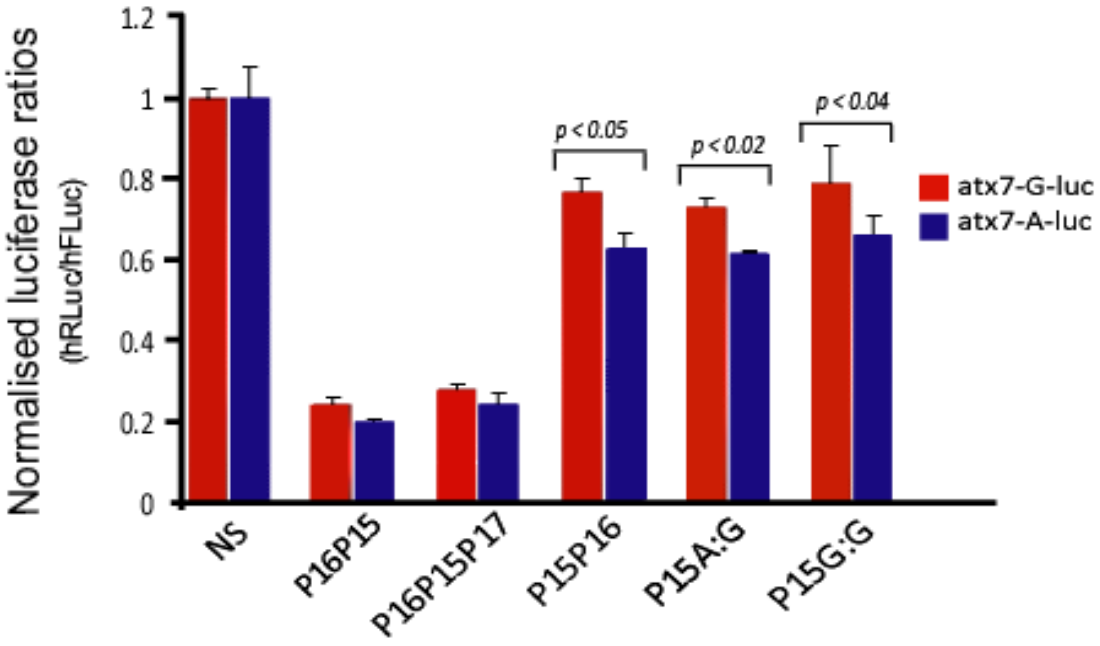


Figure 3.3. Analysis additional mismatched shRNA guide sequences targeting the G>A SNP in *atxn7* in a dual luciferase assay. Relative levels of R*enilla* luciferase (hRLuc) expression normalized to firefly luciferase (hFLuc) expression. Each experiment was performed in triplicate and the data is relative to that measured using a non-specific shRNA, NS. Average ± standard deviation is shown. Statistically significant differences ($p<0.05$) between wild-type and mutant silencing are indicated by corresponding p values. Relative expression of wild-type and mutant targets is represented by red and blue bars respectively.

3.2. Full-length assay

3.2.1. Hemizygous assay

3.2.1.1. Single mismatches

To replicate the SCA7 disease state with greater precision, shRNAs were tested against the full-length *atxn7* gene fused to eGFP in a hemizygous cell system. The assay was performed using the HEK293 cell line which had been validated for this study by confirming minimal levels of endogenous *atxn7* transcript in comparison to the transfected full-length transcripts (Materials and Methods, Figure 2.9, page 63). The results of this assay indicated that target knockdown was reduced across the panel of shRNAs in comparison to the knockdown seen in the luciferase assay (Figure 3.4). Again, shRNAs mismatched at shR-P12 and -P14 showed some level of selectivity (although not significant). However shR-P15 showed no ability to knockdown either the full-length mutant or wild-type gene, contrasting with the results which showed selectivity in the luciferase assay. Surprisingly shR-P16, which showed the most effective knockdown against both targets in the luciferase assay and thus minimal discrimination, demonstrated the most selectivity against the full-length targets with minimal effect to the wild-type (75%) and strong selectivity for the mutant target (40%; $p<0.05$).

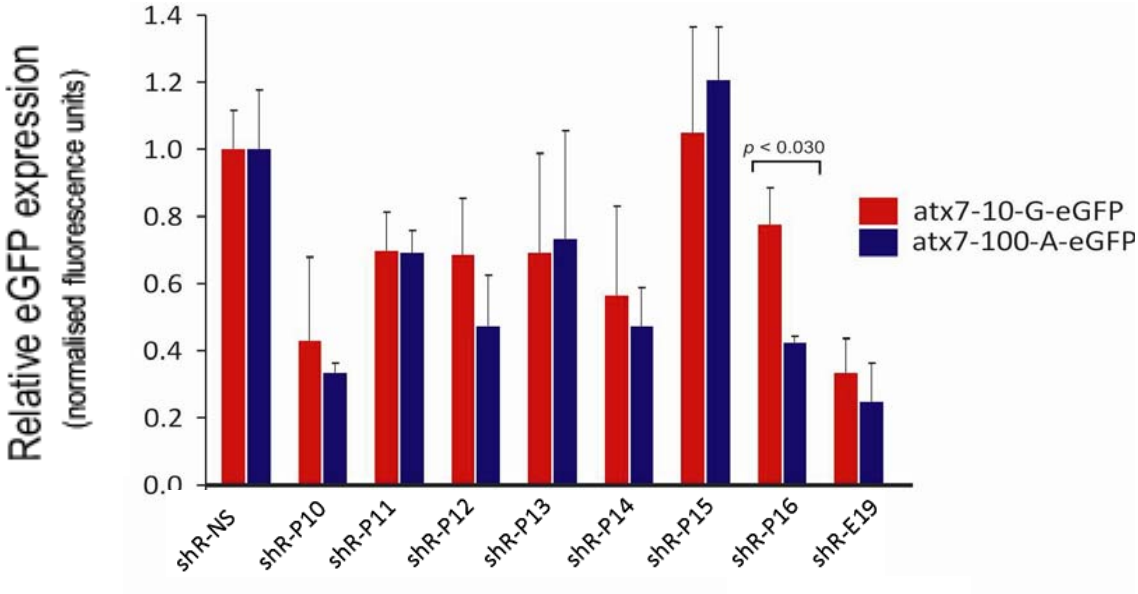


Figure 3.4: Single mismatched shRNAs targeting the *atxn7* G>A SNP in a full-length hemizygous assay. Quantitation of fluorescence in HEK293 cells transfected with wild-type (atx7-10-G-eGFP) or mutant (atx7-100-A-eGFP) expression plasmids and the indicated shRNA. Each experiment was performed in triplicate and the data is relative to that measured using a non-specific shRNA, NS. The anti-eGFP shRNA, used was pE19; a published U6/shRNA expression plasmid targeting eGFP (200). Average ± standard deviation is shown. Statistically significant differences ($p<0.05$) between wild-type and mutant silencing are indicated by corresponding p values Wild-type and mutant targets are represented by red and blue bars respectively.

3.2.1.2. Double mismatches

In a similar manner to P16 the double mismatch P16P17 showed moderate selectivity in the full-length target assay (Figure 3.5). In general, as was the case with the luciferase assay, the level of knockdown across the screen was reduced in comparison to the single mismatches to the point where none of the effectors demonstrated any ability to knock down the targets. However, E19, the hairpin designed to target eGFP specifically showed similar levels (70-80%) of

knockdown in both assays indicating that the assay accurately reflected knockdown ability of each shRNA.

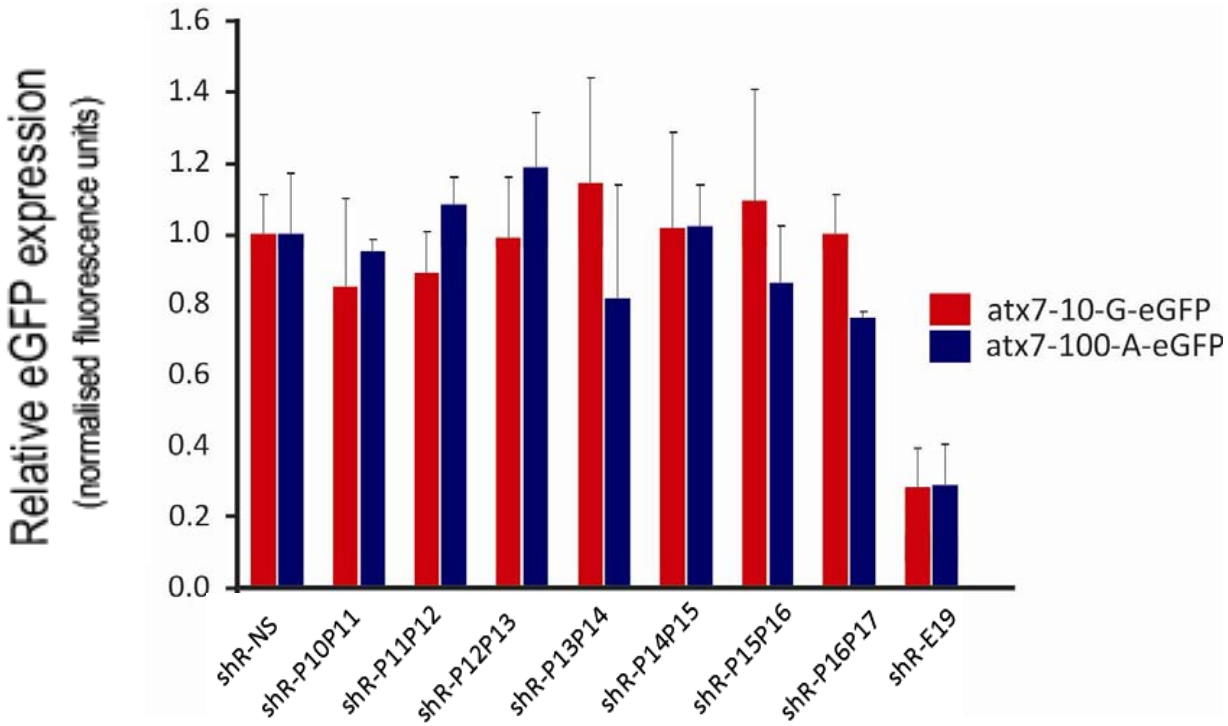


Figure 3.5: Double mismatched shRNAs targeting the *atxn7* G>A SNP in a full-length hemizygous assay. Quantitation of fluorescence in HEK293 cells transfected with wild-type (atx7-10-G-eGFP) or mutant (atx7-100-A-eGFP) expression plasmids and the indicated shRNA. Each experiment was performed in triplicate and the data is relative to that measured using a non-specific shRNA, NS. The anti-eGFP shRNA, used was pE19; a published U6/shRNA expression plasmid targeting eGFP (200). Average ± standard deviation is shown. Statistically significant differences ($p<0.05$) between wild-type and mutant silencing are indicated by corresponding p values Wild-type and mutant targets are represented by red and blue bars respectively.

3.2.1.3. Additional mismatches

Since P16 had been shown to be the most effective hairpin in this assay, the previously mentioned modified versions of the primary mismatch hairpin were assessed in this assay in the hope that discrimination might be enhanced. Neither P16P15 which included shifting of the position of the secondary mismatch, nor P16P15P17 which incorporated additional mismatches showed any ability to selectively knockdown the mutant target (Figure 3.6). Indeed neither of these hairpins showed any ability to knockdown either target reinforcing the data from the luciferase assay.

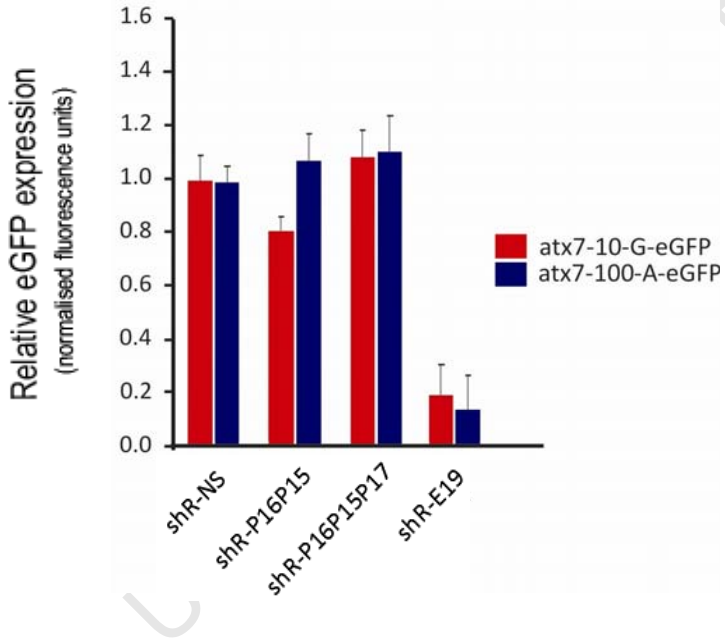


Figure 3.6: Additional mismatched shRNAs targeting the *atxn7* G>A SNP in a full-length hemizygous assay. Quantitation of fluorescence in HEK293 cells transfected with wild-type (atx7-10-G-eGFP) or mutant (atx7-100-A-eGFP) expression plasmids and the indicated shRNA. Each experiment was performed in triplicate and the data is relative to that measured using a non-specific shRNA, NS. The anti-eGFP shRNA, used was pE19; a published U6/shRNA expression plasmid targeting eGFP (200). Average ± standard deviation is shown. Statistically significant differences ($p<0.05$) between wild-type and mutant silencing are indicated by corresponding p values. Wild-type and mutant targets are represented by red and blue bars respectively.

3.2.2. Heterozygous assay

A significant test of RNAi-based single nucleotide discrimination is to use a heterozygous system (169). Therefore a full-length assay was designed to measure levels of knockdown following co-transfection of both full-length wild-type and mutant DNA encoding transcripts into HEK293 cells tagged to GFP and dsRED respectively. The P16 shRNA was chosen for study based on its high degree of target selectivity in the full-length hemizygous assay. The discriminatory effects seen with this shRNA were comparable to those in the earlier hemizygous full-length assay (Figure 3.7). Indeed, shR-P16 showed strong selectivity for the mutant target, efficiently knocking down the mutant to 7% whilst minimally affecting the wild-type with knockdown at only 74% of the control levels ($p<0.05$). shR-P15 again showed minimal selectivity and knockdown against both targets reinforcing the results of the full-length hemizygous assay. The eGFP-specific hairpin showed specific and consistent knockdown of the wild-type target tagged to eGFP, with no ability to recognise and therefore knock down the mutant transcript tagged to dsRED demonstrating the specificity of this assay. Finally, shR-P15 showed similar results to that obtained in the hemizygous assay, further demonstrating the accuracy of the assay.

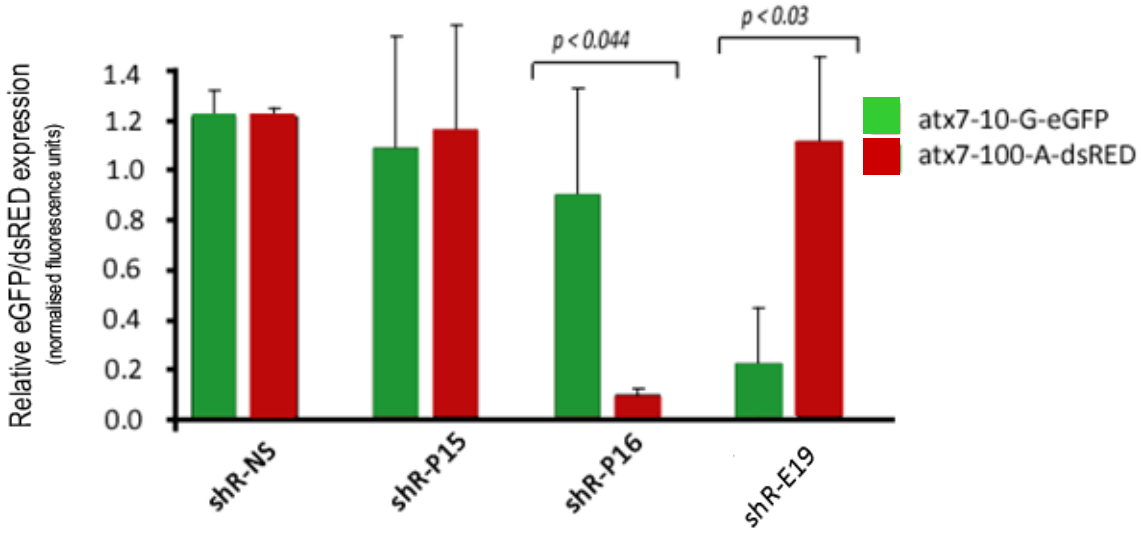


Figure 3.7. shRNA targeting the *atxn7* G>A SNP in a full-length heterozygous assay. A. Quantitation of fluorescence in HEK293 cells co-transfected with wild-type atx7-10-G-eGFP and mutant atx7-100-A-dsRED expression plasmids and the indicated shRNA. Each experiment was performed in triplicate and the data is relative to that measured using a non-specific shRNA, NS. Average ± standard deviation is shown. Statistically significant differences ($p<0.05$) between wild-type and mutant silencing are indicated by using the one-tailed t-test. Wild-type and mutant targets are represented by green and red bars respectively.

3.3. Study of aggregate formation

3.3.1. Confocal images

It is well documented that mutant polyglutamine proteins form aggregates which are associated with, if not directly causative of disease pathogenesis (202). In addition, it has been shown that mutant ATXN7 recruits wild-type ATXN7 into intracellular inclusion-like aggregates (42). In the process of performing the heterozygous assay, live cells were viewed using fluorescence microscopy prior to protein extraction to confirm fluorescent gene expression. To obtain more enhanced images, confocal microscopy was used to observe the pattern of

expression of the mutant and wild-type protein. As has previously been demonstrated (42) the wild-type protein indicates a dispersed pattern of expression with minimal aggregate formation, while in contrast the mutant protein appears in the form of inclusion bodies or large aggregates (Figure 3.8). Co-transfection of both wild-type and mutant *atxn7* genes replicates previously reported results (42) in that the wild-type protein co-localizes to these aggregates of mutant protein indicating recruitment of the wild-type protein by the mutant protein into the inclusions (Figure 3.9). However, with the introduction of the P16 hairpin, the remaining wild-type protein was observed to revert to the normal pattern of expression, and the number of intracellular aggregates of mutant protein appeared significantly reduced (Figure 3.9).

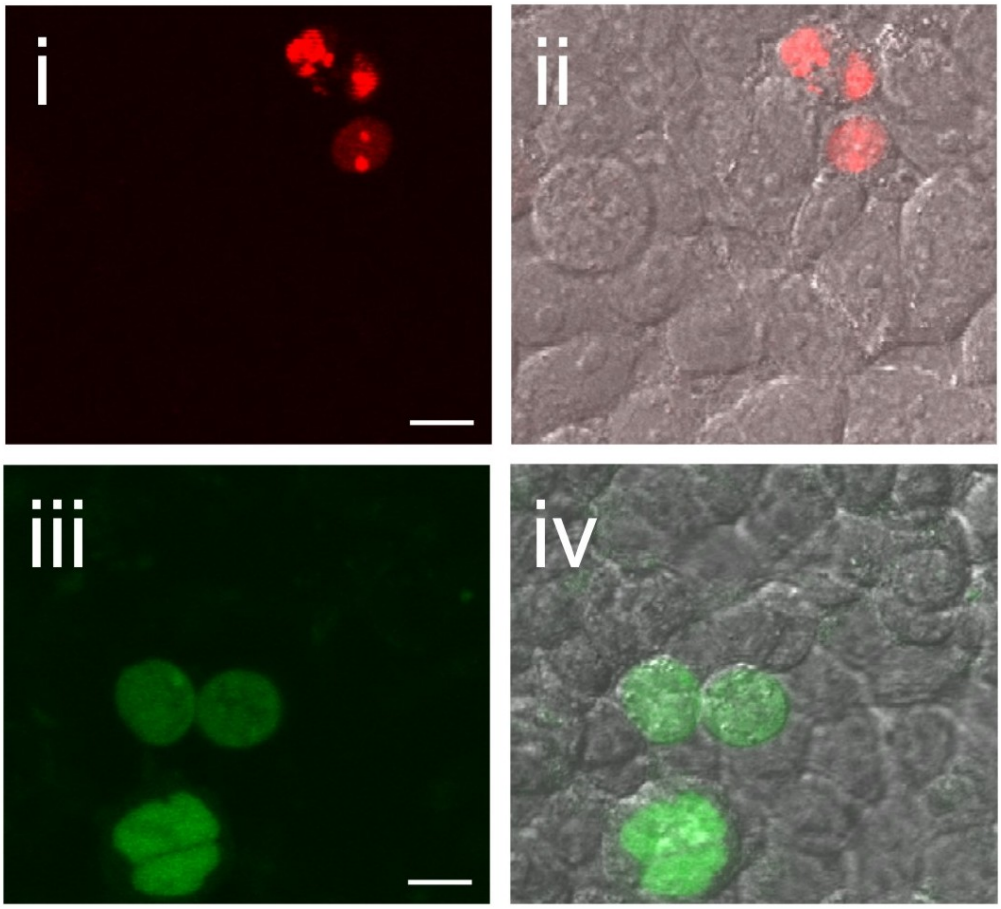


Figure 3.8. Representative confocal images of HEK293 cells from the heterozygous assay. Confocal images of cells transfected with (i, ii) mutant (atx7-100-A-dsRED) alone (iii, iv) wild-type (atx7-10-G-eGFP) alone; (i) Image visualized under red fluorescence (ii) red fluorescence merged with the bright field, (iii) green fluorescence and (iv) green fluorescence merged with the bright field. Scale bar represents 10μm.

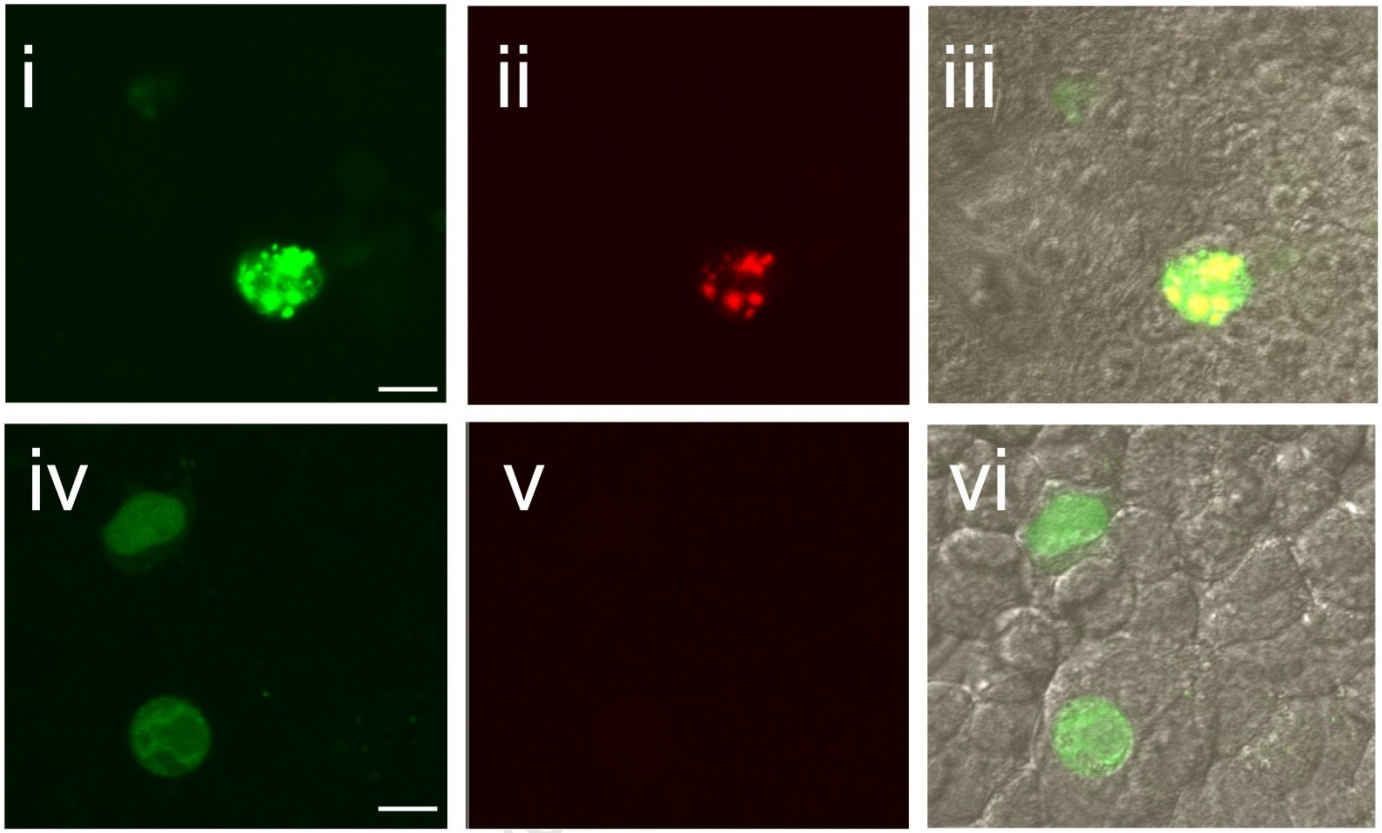


Figure 3.9. Representative confocal images of HEK293 cells from the heterozygous assay. Confocal images of cells co-transfected with wild-type and mutant expression plasmids in addition to (i - iii) shR-NS or (iv - vi) shR-P16. (i) and (iv) show images under green fluorescence to reveal wild-type protein, (180) and (v) show images under red fluorescence to reveal mutant protein, and (iii) and (vi) show images then merged under green and red fluorescence and bright field. Scale bar represents 10μm.

3.3.2. Aggregate counting assay

In order to quantify whether the level of wild-type release from aggregation to a normal dispersed pattern of expression was significant, the number of cells containing aggregates or presenting a normal pattern of expression was counted following shRNA treatment and compared to that of the non-specific control (Figure 3.10). This data shows a substantial decrease in the number of cells expressing mutant aggregates in the presence of shR-P16, corroborating data from the fluorescent heterozygous assay. More importantly, these results further support the original observation that, not only does the combination of mutant and wild-type protein result in a significantly increased percentage of aggregates containing wild-type protein ($p<0.05$), but that treatment with shR-P16 promotes release of the wild-type protein to the dispersed distribution seen in cells expressing the wild-type alone.

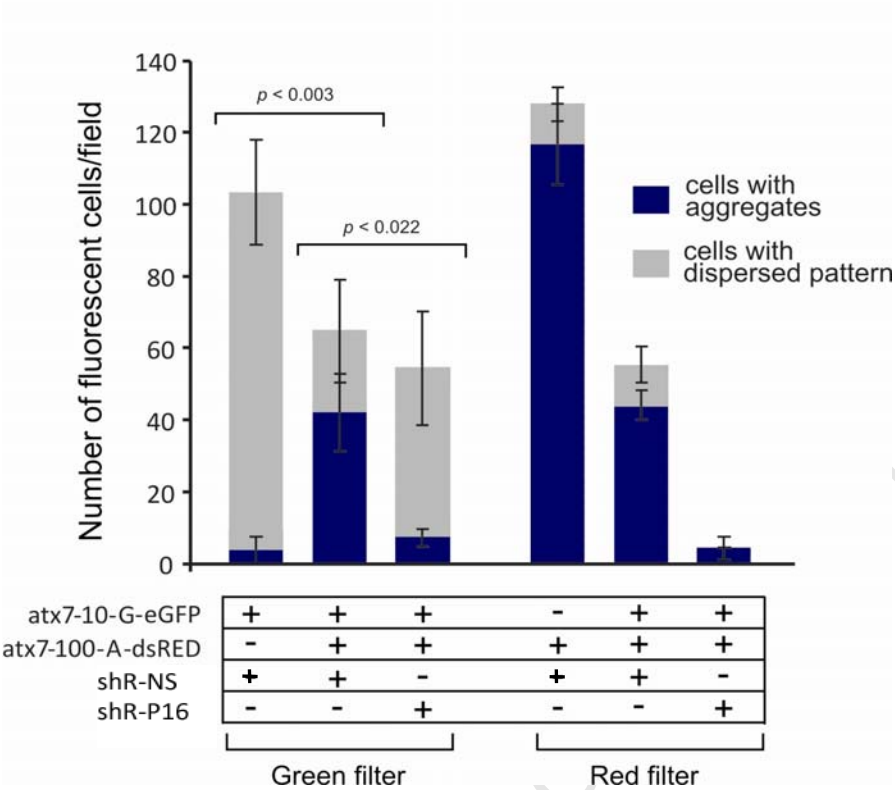


Figure 3.10. shRNA P16 reduces wild-type *atxn7* aggregates in a heterozygous assay. Cells expressing eGFP and/or dsRED were counted according to whether they contained aggregates or a dispersed pattern of expression of mutant and wild-type ATXN7. Cells were transfected with wild-type target (atx7-10-G-eGFP) alone; mutant target (atx7-100-A-dsRED) alone; mutant (atx7-100-A-dsRED), wild-type (atx-10-G-eGFP) and NS (non-specific shRNA); mutant, wild-type and shR-P16. Cells were counted separately in the red and the green filter by collecting 3 representative images from each well and combining the total number. This was performed for each indicated combination in biological triplicate, yielding standard deviations. The bars comprise the total number of cells counted in each transfection; separated according to whether they contained aggregates (blue) or a dispersed pattern of expression (grey). Note that the decrease in expression from target vectors transfected alone to co-transfected cells is likely to be a result of the promoter occlusion effect of co-expression of targets and not due to the addition of the shR-NS which has no effect upon the target vectors (Figure 2.10, page 64). Statistically significant differences in % of aggregate containing cells are indicated by corresponding p values.

3.4. miRNA assay

3.4.1. Hemizygous assay

Given that shRNAs have been shown to be toxic *in vivo*, effectors were constructed in a pri-miRNA based structure and screened in a full-length assay in order to identify an equivalent allele-specific effector in a format predicted to be less toxic (Figure 2.6, page 60). Due to the possibility of alternative processing of the two different types of effectors, multiple miRNAs were designed to incorporate this potential variation in the hope that one would mimic the mature shR-P16 guide strand. Thus miRNAs were screened in the same manner as previously described in a full-length hemizygous assay and tested simultaneously with the shR-P16 effector. Results show that miR-P16 is capable of discriminating as efficiently as shR-P16 (Figure 3.11). Indeed, miR-P16 shows almost exactly the same levels of knockdown against the mutant and wild-type effectors (with 80% and 40% remaining respectively) as shR-P16.

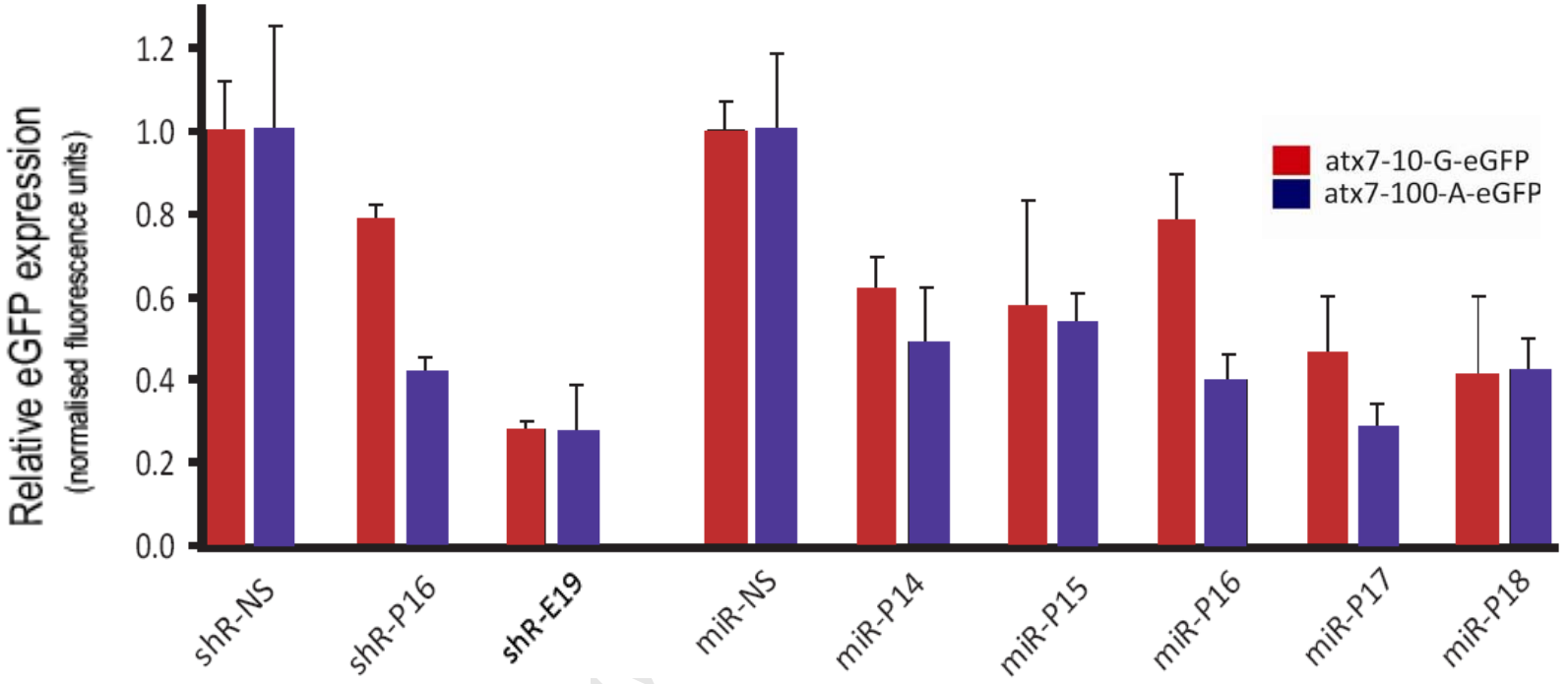


Figure 3.11: miRNAs targeting the *atxn7* G>A SNP in a full-length hemizygous assay. Quantitation of fluorescence in HEK293 cells transfected with wild-type (atx7-10-G-eGFP) or mutant (atx7-100-A-eGFP) expression plasmids and the indicated miR/ shRNA. Each experiment was performed in triplicate and the data is relative to that measured using a non-specific shRNA, shR-NS or non-specific miRNA, miR-NS. The anti-eGFP shRNA, used was pE19; a published U6/shRNA expression plasmid targeting eGFP (200). Average ± standard deviation is shown. Statistically significant differences ($p<0.05$) between wild-type and mutant silencing are indicated by corresponding p values Wild-type and mutant targets are represented by red and blue bars respectively.

3.4.2. Heterozygous assay

To compare the efficiency of the miRNAs with the selective shR-P16 in a heterozygous system, the miRNAs were screened as previously outlined. The results indicate similarities between the selectivity obtained between the shR-P16 (Figure 3.7, page 82) and the miR-P16 (Figure 3.12) with the latter demonstrating the same ability to show increased selectivity in a heterozygous system. Remarkably, miR-P16 results in almost complete ablation of the mutant target (to less than 10%) with minimal effect on the wild-type (retaining 75% expression).

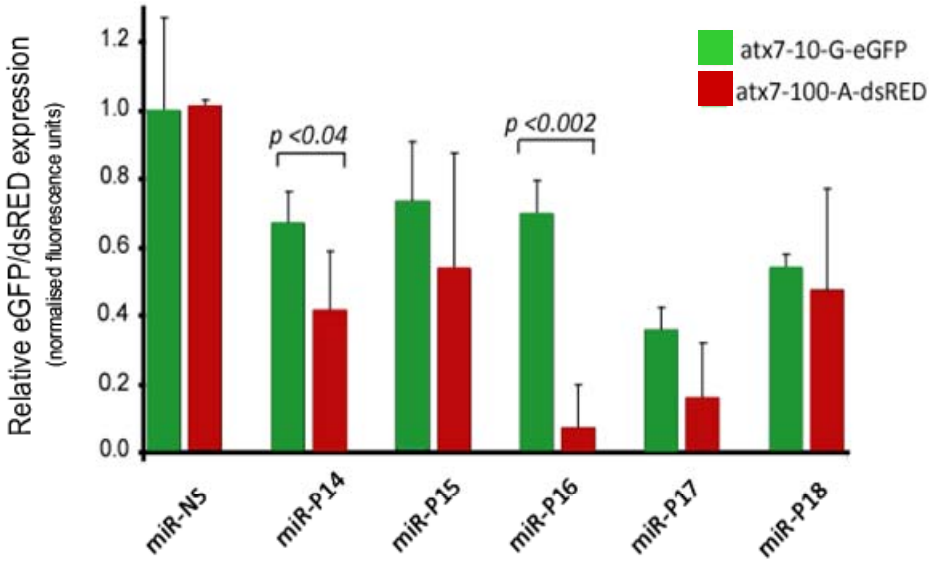


Figure 3.12. miRNAs targeting the *atxn7* G>A SNP in a full-length heterozygous assay. Quantitation of fluorescence in HEK293 cells co-transfected with wild-type atx7-10-G-eGFP and mutant atx7-100-A-dsRED expression plasmids and the indicated shRNA. Each experiment was performed in triplicate and the data is relative to that measured using a non-specific miRNA, miR-NS. Average ± standard deviation is shown. Statistically significant differences ($p<0.05$) between wild-type and mutant silencing are indicated. Wild-type and mutant targets are represented by green and red bars respectively.

3.4.3. Ratio assay

Given that the miRNA screen performed using the heterozygous assay resulted in allele-specificity at a 2:1 ratio of target:effector, an attempt was made to ascertain whether lower ratios of target:effector might still be effective. The hairpin shR-P16 maintained the previously established high level of allele-specificity with the wild-type minimally affected to over 80% relative to the non-specific control; while the mutant target was reduced to less than 10% expression (Figure 3.13). In comparison, the selectivity of miR-P16 was significantly reduced (Figure 3.14). Interestingly, while the level of knockdown of the wild-type remained the same as that at a 2:1 ratio at 70%, the ability of the pri-miRNA-based hairpin to knockdown the mutant was drastically reduced, to the extent that 30% of the mutant target remained in comparison to the 10% previously exhibited.

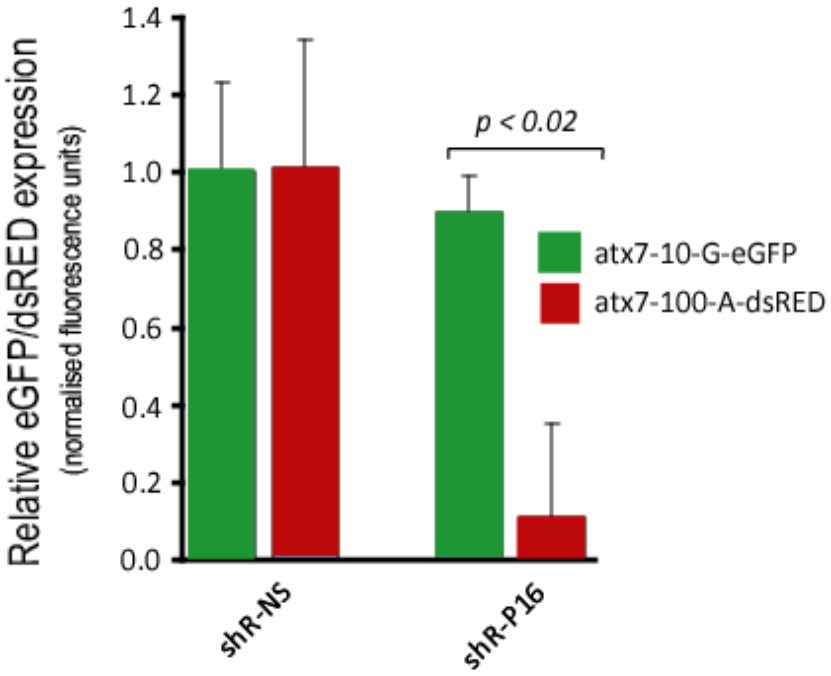


Figure 3.13. shRNAs targeting the *atxn7* G>A SNP in a full-length heterozygous assay with decreased effector. Quantitation of fluorescence in HEK293 cells co-transfected with wild-type atx7-10-G-eGFP and mutant atx7-100-A-dsRED expression plasmids and the indicated shRNA. Each experiment was performed in triplicate and the data is relative to that measured using a non-specific shRNA, shR-NS. Average ± standard deviation is shown. Statistically significant differences ($p<0.05$) between wild-type and mutant silencing are indicated. Wild-type and mutant targets are represented by green and red bars respectively.

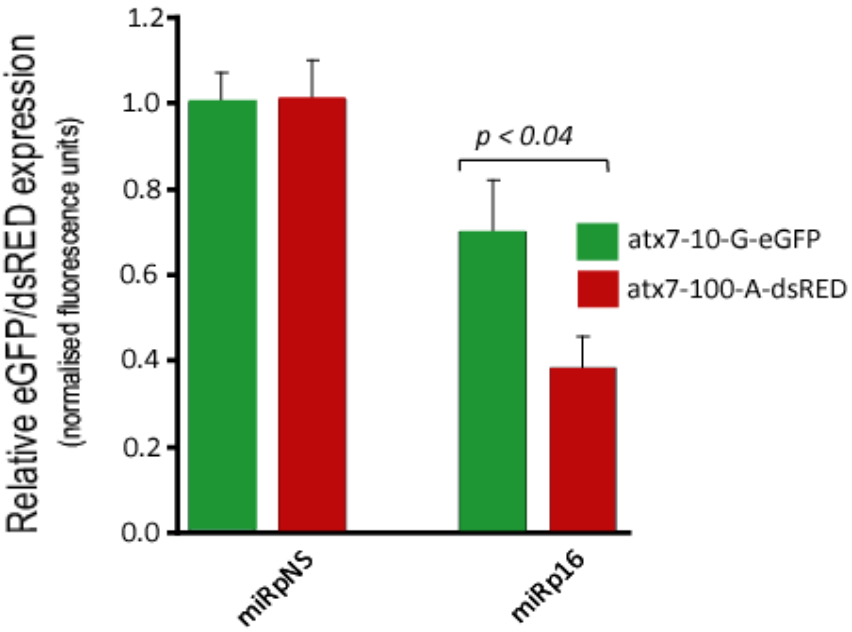


Figure 3.14. miRNA targeting the *atxn7* G>A SNP in a full-length heterozygous assay with decreased effector. Quantitation of fluorescence in HEK293 cells co-transfected with wild-type atx7-10-G-eGFP and mutant atx7-100-A-dsRED expression plasmids and the indicated shRNA. Each experiment was performed in triplicate and the data is relative to that measured using a non-specific miRNA, miR-NS. Average ± standard deviation is shown. Statistically significant differences ($p<0.05$) between wild-type and mutant silencing are indicated. Wild-type and mutant targets are represented by green and red bars respectively.

4. Discussion

In this study, the efficacy of RNAi hairpins to target the mutant *atxn7* transcript using a linked SNP, with the intention of ameliorating the pathogenicity of SCA7 in a cellular model was investigated. Using an extensive screen of various designs of expressed hairpins, multiple cellular models of target knockdown were tested. Comparisons of the short luciferase and full-length eGFP reporter assays led to conclusions regarding secondary structure. While the luciferase assay may be useful for mechanistic investigations, the full-length assay more accurately represents the endogenous gene response. Full-length heterozygous assays indicate that an allele-specific effector has been identified, which has little knockdown effect on the wild-type, and strong silencing of the mutant. Structural modifications of this effector to resemble a pri-miRNA structure yielded remarkably similar selectivity, which has importance in future work in that *in vivo* studies would require a non-toxic effector. Finally, hairpin selectivity in the heterozygous system indicated a further benefit of allele-specific silencing due to the reduction of total mutant protein causing reduced aggregate formation. Consequently, the wild-type protein is released into the more dispersed pattern of expression indicative of that observed with the normal protein alone.

4.1. Mechanistic implications

Effectors used to screen multiple positions of the SNP in order to obtain selectivity were based on the shRNA structure. Despite possible toxic effects *in vivo* (135), they are easy to design. Since they have been more extensively studied than miR-based hairpins, the rules governing their construction yield fairly predictable outcomes in terms of guide sequence production. They require only a single processing step by Dicer, pre-empting the complications of the Drosha, exportin-5, and Dicer processing that the miR-based hairpins require. Should the structure of the miR-based hairpins

be in-accurate, they may not be recognized by Drosha, or be cut in alternative positions.

In designing guide sequences several issues were taken into account, reviewed in the introduction. These include the position of the mismatch within the guide strand, the strength of the mismatch, and target accessibility (Figure 4.1, page 96). Secondary structure of the target region and the strength of the mismatch cannot be controlled, as these are defined by the single nucleotide change itself; however they should be taken into consideration as they have significant effects on the discrimination. Mismatch placement, however, can be modified; thus we were able to address this issue extensively.

4.1.1. Mismatch placement

In order to achieve selective target knockdown, a sequence change must be amenable to exploitation. Given that the nature of the mutation in polyglutamine disorders is not suitable for discrimination by a short RNA guide sequence, an alternative approach is required. Highly heterozygous SNPs present in the targeted transcripts can be used as an alternative for the mutation itself. In SCA7 patients in South Africa, it has already been demonstrated that one such SNP exists such that over 50% of the patients have been genotyped with an A and G allele on the mutant and wild-type transcripts respectively. This single nucleotide change was investigated as a template for selective silencing in this study.

The question of where to position this single base pair change within the guide sequence of a predicted shRNA effector to achieve optimal selectivity in allele-specific RNAi is debatable and has been examined in various studies (Table 1.2, page 43). Traditionally, it has been hypothesised that centrally placed mismatches i.e. those placed near position 10 or 11, with position 1 being the first nucleotide of the 5' end of the guide strand, allow for the greatest discrimination because they would interrupt the site of Ago2 cleavage (178). Since cleavage is the predominant form of post-

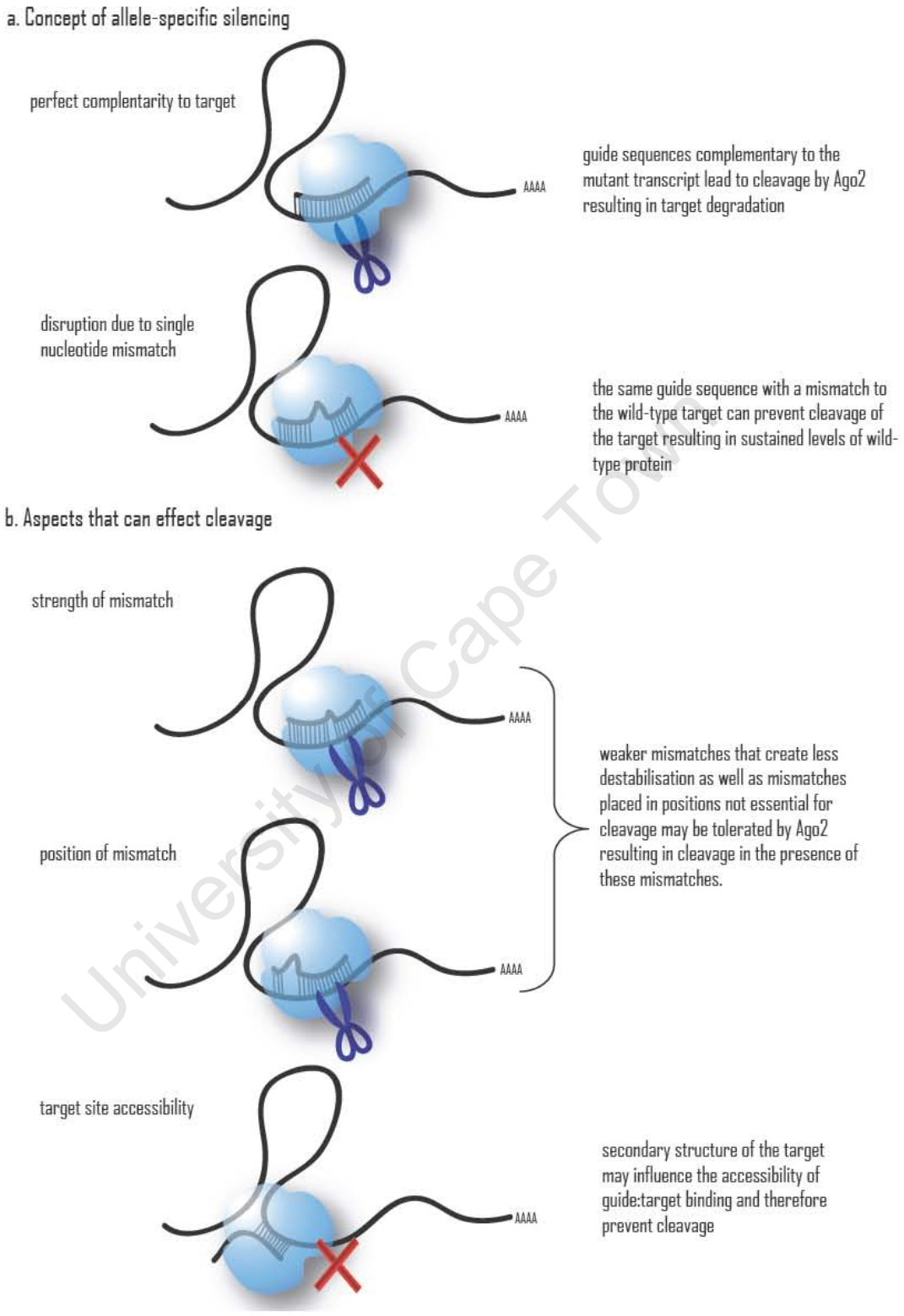


Figure 4.1: Aspects to consider in allele-specific silencing

transcriptional suppression in perfectly complementary sequences (124, 125) knockdown of the mismatched target would be avoided. However, the comprehensive study by Schwarz and colleagues (183) showed extensive evidence suggesting that mismatches in the 3' region of guide strands could allow greater selectivity than those placed centrally. In addition, the same study indicated that while mismatches placed in the seed region result in discrimination, they did not yield the most efficient selectivity in a full screen of two different genes, indicating that interfering with catalysis rather than binding, would lead to greater discrimination (183). With the view to screening as many likely discriminating sequences as efficaciously as possible, we screened mismatches placed at positions 10 to 16 of the 3' end of the guide strand, which Schwarz and colleagues had shown to have greatest selectivity in their screen. In contrast to the siRNAs investigated in the Schwarz study, expressed hairpins were used as effectors in the data presented here. Expressed hairpins are inexpensive in comparison to the RNA oligomers of siRNAs. Placement of a Pol III promoter upstream of these hairpin sequences allows for stable expression of these short sequences with specific termination without the requirement for complex termination regulatory sequence elements. For the purpose of addressing mechanistic issues, the initial part of this discussion will focus on the data obtained from the luciferase assay due to the reduced impact of secondary structure. It is acknowledged that expressed hairpins predicted to place mismatches at specific positions may yield heterozygous heterogeneous species, with the major species not representing the predicted mismatch. Thus it is emphasized that the observations made are based on predicted structure and not necessarily definitive. However, it is likely that should there be alternative species, they will be at most one or two nucleotides up or downstream of what is predicted. Further, given the extensive published studies using this design and the similarity of the results presented in this thesis, the conclusions drawn from predicted positions are likely to be correct.

Results from the luciferase assay, in which a short target sequence of 50 nucleotides is tagged to the luciferase reporter gene, indicated various single mismatches placed 3' to the centrally placed shR-P10 appear to be the discriminatory, including those at positions 12, 14, and 15. The single mismatches placed at positions 10 and 11 showed minimal selectivity conflicting with the hypothesis that interfering with the site of Ago2 cleavage at these positions leads to the greatest discrimination. Indeed, the results reinforce the Schwarz study suggesting an important role for the 3' region of the guide strand (183). To date, no such role has been proposed. Current theories on the regional structure of guide sequences indicate an important role for the 5' nucleotides of the guide, known as the seed region. This region has been shown to be essential in binding and recognition of the target sequence and have been shown in structural studies of Ago2 in *Pyrococcus furiosus* to directly interact with the slicer protein (203). Studies have shown that anchoring of the 5' nucleotide of the guide strand to a binding pocket in the PIWI domain of Ago2 in *Archaeoglobus fulgidus* (204) leads to direct interaction of the seed region within a small groove of the PIWI domain due to the conserved positive charges of the latter interacting with the negative phosphate backbone of the RNA (203). A possible reason for the importance of the 3' end may lie in the requirement of an alpha helix for slicer activity. Studies of Ago2 suggest that the PAZ domain which recognises the 3' end of the guide strand may provide additional constraints on the dsRNA helix formation (205). Nucleotide interactions have previously suggested that mismatches within nucleic acid duplexes can be more stable if surrounded by strong G:C pairings, whereas weaker adjacent pairs such as A:U allow the mismatch to maintain its maximum destabilisation (206). Similarly, since the seed region is maintained within a tight small groove within the Piwi domain of Ago2, destabilising effects may be minimised, while in comparison, it has been suggested that the 3' region bound to the target may extend into a less constrained region of Ago2, possibly even into solvent (204). Thus enabling even weak mismatches to create maximum local widening of the duplex preventing Ago2 slicing (207), possibly due to disruption of alpha-helix formation. Indeed, the alpha-helical conformation of the target:guide duplex

has been shown to be essential for binding stability of the PAZ domain (208) Furthermore thermodynamic destabilisation has been shown to be less important than distortion of the helix to the point where A:C mismatches are more tolerated by Ago2 than G:U mismatches (207). If this is the case, it may explain why traditionally weak mismatches correctly placed in a specific region have such a strong effect on slicer efficiency.

4.1.2. Nature of the mismatch

Most studies based on a single nucleotide change have exploited purine:purine or pyrimidine:pyrimidine mismatches. In this study, the nature of the SNP resulted in a weak G:U mismatch between the wild-type target and the guide strand. As such, complete discrimination was unlikely to occur, although minimal selectivity has been demonstrated in other studies (Table 1.2, page 43). The surprising selectivity of shR-P16, indicates that G:U wobbles can be used to discriminate, even though it is not considered a true mismatch due to the formation of a single hydrogen bond. This is likely in part due to the sequence context of the mismatch. It has been well established, from the complex nearest neighbour model used to identify thermodynamic differences within RNA hybrids, that different sequences surrounding the same mismatch significantly change the thermodynamic profile of the mismatch. Thus it is unlikely that every G:U mismatch will be amenable to selectivity. Only a few studies have looked at using G:U mismatches across different positions with varying degrees of success (Table 1.2, page 43). However, most acknowledge that some modifications of the guide sequence may be required to enhance selectivity. Not all of these studies have clearly defined positions, while they label a mismatch as being “central”, some have used alternative nomenclature such that a position 9 is in fact equivalent to position 12 from the 5’ end. One study has shown that a G:U mismatch at position 10 from the 5’ end of the guide strand leads to minimal selectivity (183). This corroborates the data presented in our study. However the same study did show significant selectivity at position 10 using a strong purine:purine mismatch. The notion of relaxed conformation leading to maximum destabilisation in the 3’ region may further explain the significant

selectivity obtained even with a G:U wobble. It is possible that strong mismatches lead to strong destabilisation even in tight small grooves, such as that where cleavage occurs thus resulting in the strong selectivity obtained both at positions 10 and 16. However, a weaker mismatch may be more easily constrained at the tight cleavage junction, minimising local widening of the duplex, and thus allowing slicing of the target.

4.1.3. Modifications

Given the nature of the weak mismatch, several attempts were made to enhance the level of selectivity identified in the 3' region. Initial modifications included introducing a second mismatch immediately 3' to the primary mismatch. This resulted in a single weak purine:pyrimidine (A:C) mismatch against the mutant, with two mismatches the A:C mismatch and the initial G:U mismatch against the mutant. Surprisingly, only one of the double mismatches resulted in selectivity greater than that achieved using the corresponding single mismatch; that of shR-P13P14. Other studies using weak (and strong) mismatches have shown that the addition of a secondary mismatch may enhance the selectivity by increasing the difference in destabilisation (169, 170) (Table 1.2, page 43). Indeed a double mismatch at position 14 has previously been shown to enhance mismatch specificity with a weak G:U wobble against the normal tau transcript (169). The data presented in this thesis suggest that this is not always the case, and may reinforce the role of sequence context. The general abrogation of knockdown against both targets is not surprising given that additional mismatches are expected to hamper alpha helix formation, and therefore cleavage efficiency. However, further attempts to enhance selectivity were made by changing the position of the secondary mismatch in shR-P15 which had shown the greatest selectivity in the luciferase assay. The changes that were tested included,

- shR-P15P16': changing the nature of the secondary mismatch from a weak A:C to a stronger pyrimidine:pyrimidine U:C to enhance destabilisation.

- shR-P15A:G: changing the nature of the primary mismatch. According to thermodynamics one of the most destabilizing mismatches possible is a purine:purine mismatch, in particular, A:G. Purine:purine mismatches are considered the strongest, because the strength of these bonds is decreased by incompatible hydrogen bonding positions (209, 210). Furthermore different purine: purine mismatches show different destabilisation, thus it may be possible to increase the difference in destabilisation such that, although there are single mismatches to both mutant and wild-type targets, the A:A mismatch to the mutant is less destabilizing than the A:G against the wild-type. The hypothesis is that the difference in destabilization between A:A and A:G may be greater than U:A and U:G and therefore possibly more discriminatory.
- shR-P15G:G: changing the nature of the primary mismatch. Although the above rules are true stereotypically, mismatch destabilisation is sequence specific. However, it has been suggested that G:A mismatches might form ionic bonds and that tandem G:A pairs could well in fact *stabilize* RNA duplex structure (211, 212). As a result, it may be possible that G:A mismatches are *more* stable than A:A, or G:G mismatches. With that in mind another hairpin was created to yield a G:G mismatch with the wild-type, and a G:A with the mutant.

Unfortunately, none of the modifications showed increased specificity for the wild-type. As was expected, the general level of knockdown against the mutant in all three hairpins was less than that of the perfectly complementary shR-P15 hairpin.

While investigating the modifications of the selective shR-P15, an attempt was made to modify the strong shR-p16 effector in the hope that weakening the efficacy may lead to enhanced selectivity since this effector had maintained the ability to knockdown the wild-type with a double mismatch.

Placing the second mismatch on the 5' side of the primary mismatch had no effect; while the introduction of a triple mismatch (shR-P16P15P17) released knockdown by less than 5%. The tolerance for a triple mismatch at this position is surprising given that two of these mismatches overlap with shR-P15P16 which showed minimal knockdown efficiency in contrast to the lack of tolerance to a mismatch at this position in plant genes such as *PHABULOSA* (79). However, it does reinforce the previous observation that a shift in a single base pair may have significant repercussions for allele-specific knockdown.

The use of the tagged reporter genes which may or may not confer differential secondary structure to the target region that are not present in endogenous *atxn7* transcripts, as well as the use of expressed hairpins as opposed to siRNAs, may limit the ability to draw specific conclusions about the mechanisms involved in single nucleotide mismatch design. Short targets are more likely to yield accurate information regarding base by base rules, in that they negate the effects of secondary structure. Therefore most of the conclusions drawn in this study regarding mechanistic rules are taken from the luciferase assay. Comparisons of similar types of effectors are required to make assumptions regarding specific selective positions. Here, expressed hairpins which showed selectivity were identified at positions similar to the data presented in the Schwarz et al. 2006 study, i.e. positions 12, 14, and 15 (183). These may not hold true for every sequence, given the nature of the mismatch and sequence context. Furthermore, expressed hairpins predicted to place the mismatch at each position may not be as specific as siRNAs used in the Schwarz et al 2006 screen. While every effort was made in the design of these hairpins to place the mismatches, studies have indicated that a heterogeneous species of guide strands may result from Dicer cleavage of the hairpin, such that shR-P15 may result in the presence of P13. P14, P15 P16 and P17 placed mismatches (128).

4.1.4. Comparisons of short verses full-length hemizygous assays

Given that the luciferase tagged assay had yielded results demonstrating successful allele-specificity, full-length targets were constructed to more accurately represent the effect on endogenous *atxn7*. Comparisons of the short target luciferase and full-length assays indicated that the former failed to identify or predict the specific guide sequence with the highest degree of selectivity in the full-length assay. Knockdown was generally diminished from the luciferase assay to the full-length assay. Since full-length mRNA secondary structure can affect the efficiency of RISC-associated cleavage (213) these results are not surprising. Significant energy may be required to disrupt the target region in order to accommodate RISC binding resulting in decreased knockdown efficiency (213). Given this effect of secondary structure, it was still unexpected to observe the distinct changes in efficient effectors, such as that of shR-P15. This hairpin had resulted in strong discrimination in the luciferase assay, while in the full-length assay, demonstrated no ability to knockdown either the mutant or the wild-type. A possible reason for this is that while shR-P15 had shown the highest selectivity, the levels of knockdown had been less efficient (remaining levels of expression at 90% for the wild-type and 55% for the mutant) than all the other single mismatch hairpins, most of which were able to show less than 50% of remaining levels of both targets. The effect of secondary structure would thus interfere with target binding to such an extent that knockdown of either the mutant or the wild-type was no longer possible, hence abolishing the resultant selectivity. The observation that all but one of the double mismatched hairpins, which had shown decreased knockdown but still limited selectivity, in the luciferase assay, now demonstrated no ability to knockdown either target in the full-length assay reinforces this theory. The converse of this hypothesis could be used to explain the resultant selectivity obtained using shR-P16 in the full-length assay. This hairpin had previously shown strong levels of knockdown against both targets in the luciferase assay, but became selective in the full-length assay. Secondary structure could reduce the accessibility of the target, resulting in significant release of the

mismatched wild-type target, while the perfectly matched mutant target duplex retained knockdown ability to a greater degree. Interestingly, the double mismatch, shR-P16P17, which had also shown strong knockdown of both targets in the luciferase assay, showed limited (although statistically insignificant) selectivity against the full-length targets and may represent a subtle replication of this hypothesis.

Although the full-length target assay is likely to represent more accurately the endogenous system, the question of the repeat size should be taken into account. In this study, the wild-type allele contained 10 CAG repeats – the most common normal size in most populations (12, 22). The size of the expansion in SCA7 patients varies from over 39 to 406; in this study, only one mutant allele size was examined – at 100 repeats. To place this in context, as of 2003, the largest repeat size of SCA7 in South Africa was reported to be 92 repeats (90). Thus it could be argued that given the extensive evidence of an effect of target structure, a large change in the repeat size of the expanded allele may confer significant secondary structure alterations from one mutant allele to another, thus effecting selectivity. This may be the case with the SNP used to selectively target *atxn3,* located immediately 3' to the SCA3 CAG repeat, (169). Furthermore, it may confer improved target accessibility to the wild-type allele, leading to unwanted selective knockdown of the wild-type, although this has yet to be shown. However, an advantage of the SNP used here, is the significant distance between the expansion and the SNP, with the former in exon 3 and the latter in exon 12, over 2.5kb away. Therefore, local secondary structure around the SNP is less likely to be affected by a change in the size of the expansion, and almost certainly not by a small change in the size of the normal allele.

With reference to the full-length assays, It is noted that a lower than expected transfection efficiency of the full-length targets was observed as seen in Figure 3.8 and 3.9, page 84 and 85. One of the main reasons for this was thought to be the introduction of a fusion protein in cells which express the endogenous gene at much lower levels. In order to maximise the level of

transfection, several transfection reagents were tested. Only jetPEI (*Polyplus*, UK) showed reasonable levels of transfection efficiency. In addition, half of the total amount of DNA introduced into the cells included a non-specific shRNA expressing plasmid to ensure that the same amount of DNA, and ratio of target to transfection reagent was added to each well. Despite the low level of transfection, good standard deviations were achieved by using normalising to protein levels, thus validating the reliability of the assay.

4.1.5. Strand bias

In designing any effector, the aspect of strand bias should be taken into account. RISC has been shown to have the ability to incorporate both passenger and guide strands of a small RNA duplex. As a result, those effectors which lack the ability to knockdown both the mutant and the wild-type target may not represent inefficient selectivity, but rather be due to the preferential incorporation of RISC of the passenger strand (117, 118). This could explain the lack of selectivity of some of the double mismatch hairpins, in particular shR-P11P12 and –P13P14. Both hairpins showed a significant decrease in knockdown of the wild-type gene to the point where silencing was completely removed. It is therefore possible that improved design incorporating enhanced duplex instability may result in increased incorporation of the guide, with an improvement in mutant knockdown, without significant effects on the wild-type. However, while the introduction of double mismatches at positions 12 and 14 (shR-P12P13 and –P14P15) resulted in decreased knockdown as well as decreased selectivity, silencing was not completely removed indicating that this is a true reflection of double mismatches unable to enhance selectivity. As such, it is likely that strand bias would reduce/ enhance levels of knockdown of both targets, and have little effect on the selectivity obtained. Furthermore, each hairpin in this investigation was designed with the intention of incorporating a thermodynamic signature, such that the duplex would be significantly destabilised using G:U wobble pairs in the 3’ end of the passenger strand, facilitating preferential incorporation of the guide strand.

4.1.6. Heterozygous assay

In an attempt to replicate the cellular system, the knockdown efficiency of each of the effectors was investigated. In order to distinguish between levels of knockdown of mutant verses wild-type, the eGFP reporter from the full-length mutant target was replaced with the dsRED reporter gene. The results of this assay corroborated well with that of the full-length hemizygous assay in that both shR-P16 and shR-P15 demonstrated either the ability or lack thereof to selectively knockdown the mutant target. Indeed shR-P16 showed significantly decreased levels of mutant relative to wild-type protein indicating strong silencing of the toxic protein with a difference of nearly 80% in selectivity. This was greater even than that observed in the hemizygous assays.

The reasons for this increase in selectivity of the mutant are perplexing. Before suggesting mechanistic reasons for this enhancement the alteration in the assay itself should be addressed. Changes in levels of discrimination over short periods of time hours (1-2 hours (183)) and days (24 hours through to 72 hours; personal communication C. Sibley) have been observed. Although the heterozygous assay was performed at 72 hours in comparison to the hemizygous assay which was performed after 48 hours of transfection, 72 hour incubation of the hemizygous assay showed no increase in selectivity excluding the extended incubation as a possible reason for increased selectivity (comparison between shR-P16 in Figure 3.4, page 78 and 3.11, page 89; measured at 48 and 72 hours respectively).

In addition, the possibility of a difference in mutant verses wild-type expression from the full-length targets has been examined. While the level of knockdown against the wild-type target remains the same from hemizygous to heterozygous conditions, it is the level of mutant expression which changes significantly from approximately 40% to 10%. Thus, the expression of the two different types of mutant target vectors could vary, with greater levels of expression with the eGFP tagged target, and lower levels with the dsRed-tagged target. It is unlikely, that an increase in eGFP mutant

expression would significantly affect knockdown by a hairpin, since the eGFP-specific hairpin shows equal levels of knockdown against both mutant and wild-type targets in the hemizygous assay. Furthermore, a significant reduction is unlikely to be the cause in the heterozygous assay, since shR-P15 shows the same effect on the hemi- and heterozygous assays. Although the same amount of target plasmid was added in each transfection, quantitative PCR indicates that there are some changes in the levels of target expression (A6, Appendix, page 138). The data indicates that the levels of expression of the mutant target tagged to eGFP are greater than the equivalent target tagged to dsRED in the heterozygous assay. Thus increased levels of mutant target in the hemizygous assay may contribute to the difference in hairpin selectivity of the mutant target. This indicates that the actual level of discrimination is likely to be greater than that indicated in the hemizygous assay.

A second argument is that of altered secondary structure due to the dsRED in comparison with tagged eGFP, which were tagged to the 3' end of the full-length construct, thus less than 100 nucleotides from the SNP. That there is limited sequence homology between dsRED and eGFP (approximately 50%) suggests that this may contribute to altered selectivity in this assay. However, without knowing which reporter confers greater inaccessibility, if at all, little can be drawn from this theory. A possible way in which to test this hypothesis would be to switch reporter genes in the hemizygous assays. This may clarify whether the secondary structure effect proposed to cause changes in allele-specific positions is created by the reporter tags or the length of the repeat. Additionally, it may be worthwhile to widen the target sequence from 60 nt in the luciferase assay, to assess the importance of proximity of local secondary structure. This would significantly add to the importance of this work by the creation of a predictable *in vitro* system. These surprising results do further reiterate the importance of choice of reporter genes.

It is tempting to tentatively suggest that the reason for enhanced discrimination when both targets are present may be due to competition of

the targets. To the best of our knowledge, no data has been produced to substantiate this hypothesis. Although the exact mechanism by which mature RISC finds the complementary target is not completely understood, this theory of competition is unlikely to be explained by active recruitment of the target for the shRNA. A model for target recognition has been proposed by Ameres and colleagues whereby activated RISC binds to single stranded RNA non-specifically, and cleavage depends on the thermodynamics of the complementary base pairing, where unfavourable base pairing results in dissociation of RISC (124). Indeed it is even suggested that this may facilitate the coming together of potential target sites for testing with the guide strand. Whatever the mechanism, if competition were the reason for the enhanced knockdown, it is likely that rather than the level of mutant knockdown increasing, the level of wild-type knockdown would be expected to decrease, such that guide strands in a heterozygous system would be preferentially bound to the mutant, and therefore less likely to act against the wild-type.

It is possible that promoter occlusion has some effect in this matter. Given the strong nature of the CMV promoter expressing the targets, co-expression leads to reduced levels of both targets, resulting in an optimal threshold ratio of target to effector for selectivity. No matter what the explanation, it is clear that this effector has the ability to result in significant selectivity in heterozygous assay.

4.1.7. miRNA formatted effectors

It is now well established that shRNA effectors can lead to toxicity *in vivo* by the saturation of the endogenous miRNA pathway (135). However, pri-miR structured hairpins have been shown not to result in this toxic effect, due to their more endogenous structure, resulting in more efficient processing (214, 215). In particular, the miR-122 shuttle has been shown to produce highly homogenous mature sequences of 21 bases in length as opposed to the heterogeneous products identified using the equivalent miR-31 shuttle (196). In designing a pri-miRNA, alternative processing by Drosha and Dicer was taken into account. The data indicate that miR-P16 which is predicted to

place the mismatch against the wild-type at the same position as shR-P16 showed the greatest selectivity in a full-length hemizygous assay. Indeed in an assay in which both shR-P16 and miR-P16 were tested simultaneously, the degree of knockdown of the mutant and wild-type targets is to all intent and purposes, exactly the same. Furthermore, the results of miR-P16 in a heterozygous assay reveal the same level of knockdown as shR-P16 with 70% of the wild-type remaining and less than 10% of the mutant target. None of the other pri-miR hairpins demonstrated a similar reduction in mutant targeting, although the selectivity using miR-p14 became significant, albeit minimal. This suggests that the enhanced selectivity identified using both these effectors may be specific to the guide sequence used, rather than the solely due to the difference between the hemi- and heterozygous assays.

In the absence of the use of a primer extension sequencing method of the guides themselves, the remarkable similarity of the data from heterozygous assays may indicate that the design of both hairpins is accurate in predicting the placement of the mismatch. At the very least, they are predicted to result in the same mismatch position. As has been discussed, the anchoring of the first nucleotide of the 5' end of the guide defines the position of the guide strand within the groove between the Piwi and PAZ domains of Ago2, while the 3' end is more loosely bound. This would allow for two effectors of approximately the same length (~21 nucleotides) to contain a mismatch at the same stereotypic position within the mature RISC complex. Thus, it is likely that the mismatch of these two effectors, which result in exactly the same selectivity in two different transfection conditions, must be placed in the same position relative to the first 5' nucleotide in order to create the same destabilisation. Furthermore, observations of shR-P15 and miR-P15 in both full-length hemi- and heterozygous assays indicate that the pri-miR formatted hairpin performs better, with stronger knockdown efficiencies in both cases. This is in agreement with recent studies indicating that miRNAs are processed more efficiently, thus resulting in stronger knockdown; and further, that the use of a Pol II promoter prevents saturation of the endogenous miRNA pathway (214).

4.1.8. Reduced ratios of target to effectors

Due to the issue of promoter occlusion created by the combination of three Pol II promoters in the pri-miRNA assay in a heterozygous state, the assay was performed using a ratio of 2:1 of target:effector, in contrast to all the other assays which had been performed at a 1:1 ratio. This allowed increased levels of target, in order to maximise the low level of target expression already discussed to obtain reliable data. Given that selectivity was achieved using less effector, an attempt was made to identify whether a lower ratio could result in the same specificity. It was hoped that this may indicate that reduced amounts of effector could maintain discrimination, translating into reduced amounts required *in vivo*. These experiments indicated that while shR-P16 maintained a high level of selectivity, miR-P16 selectivity was moderately reduced. While this is not the result hoped for, it is not unexpected. shRNAs result in significantly higher levels of guide sequences in part due to the expression off a Pol III promoter. Thus, there may be excess guide, at a 1:1 ratio, such that at 5:1, enough effector remains. However, pri-miRNAs are already more economically regulated by expression from a Pol II promoter. Thus, the lower ratio may be below the threshold for optimal selectivity. In any event, these various conditions are more likely to be informative *in vivo*.

4.2. Therapeutic implications of successful allele-specific knockdown: aggregate removal

The benefits of allele-specific knockdown in SCA7 are for the most part, conjectures. It is assumed that the function of the wild-type is required at the very least for retinal specific genes, such that additional loss of the one remaining wild-type allele even without the toxic gain of function of the expanded polyglutamine protein would result in a more severe retinal phenotype i.e. a full knockout model would be expected to have a retinal phenotype. Without data from *in vivo* investigations of loss-of-function of this gene, *in vitro* evidence suggests that recruitment of the wild-type protein into

mutant aggregates (42) causes a loss-of-function contributing to pathogenesis and been discussed in the Introduction of this thesis. This was reiterated in the data presented here using confocal microscopy, where co-expression of the mutant and wild-type targets indicated co-localisation of the mutant and wild-type protein. In the normal state, wild-type protein shows a dispersed pattern of expression, while the mutant protein aggregates into large inclusion-like bodies. It was observed that the introduction of an allele-specific hairpin (shR-P16) resulted in decreased mutant aggregates, due to the reduced levels of mutant expression. As a result, the numbers of cells containing wild-type aggregates were also reduced. The protein however, appeared to retain the normal dispersed expression pattern seen when the wild-type was expressed alone, thus not as a result of decreased wild-type protein. This was quantified as being significant by counting the cells showing aggregate formation in comparison to those with dispersed expression. In the wild-type state, the ratio of aggregates to dispersed cells is minimal. However with the addition of the mutant protein, the ratio shifts to a majority of cells showing aggregates of wild-type protein. The addition of the shRNA results in a significant change back to the normal state, such that the ratio of cells containing aggregates to dispersed expression in this condition is not significantly different from that in the normal state.

Some studies suggest that inclusion formation is protective and thus inhibition of these aggregates may lead to detrimental effects (216). In our study it is important to note that the removal of large aggregates correlates with the associated removal of toxic mutant protein and thus is very likely to be beneficial. Furthermore we observed as a consequence of the reduction of mutant aggregate protein, the release of wild-type ATXN7 from aggregates to a distribution of expression resembling the pattern seen in the normal state. The importance of this lies in the fact that the function of the wild-type protein in transcription (50, 54, 55) may be retarded by this sequestration into aggregate formation. One of the advantages of the full-length assay used in this experiment is that eGFP has been shown not to interfere with the localisation of ATXN7 (42), and further that the monomeric nature of the

dsRED reporter gene used in our investigation prevents formation of dsRED aggregates. Thus, localisation of the tagged wild-type protein is likely, even in this artificial system, to be representative of that occurring in the natural state, and that the aggregates that do form can only be as a result of the mutant ATXN7 protein, not the tagged reporter gene. As has been discussed, there are questions that remain regarding the contributing secondary structure of these reporter genes to *atxn7* mRNA. It may be necessary to also investigate untagged ATXN7 levels, using Western blot analysis with an anti ATXN7 antibody indicating levels of total ATXN7 (both mutant and wild-type) and a 1C1 antibody (specific for expanded polyglutamine tracts) to reveal levels of remaining mutant ATXN7. However this may not allow accurate comparisons of wild-type knockdown, which the fluorescent tags have accomplished.

Given the striking change in the cellular distribution of wild-type ATXN7, it is expected to have a significant impact on the ability of the protein to perform its function. The *in vivo* implications could be quite extensive. One of the issues to consider is when an RNAi based allele-specific therapy should be started. Some studies of SCA7 mouse and drosophila models suggest that initiation of therapeutic intervention in the form of suppression of mutant protein must take place as early as possible; in order to precede what is likely to be a threshold point of irreversible damage. The question is therefore raised as to how early this is required, and importantly, what follows is the issue of what the effects of a long-term gene therapy would be? These questions will no doubt be answered eventually, but in the interim, can assumptions be made regarding the benefits of retaining the wild-type allele? The SCA7 models studied removed only the mutant protein, leaving the wild-type for the most part intact, yet still, they indicated an early threshold level of toxicity. However, these studies used extreme expansion sizes with truncated mutant proteins and significantly high levels of expression of the mutant protein, thus creating drastic phenotypes unlikely to be representative of SCA7 patients with lower expanded repeat sizes – constituting the majority of patients. It is therefore possible that a threshold of untreatable toxicity occurs

later in the human SCA7 condition, and that retaining the wild-type allele may lead to significant beneficial effects. A further argument for an allele-specific approach stems from the mild neurological phenotypes noted in knockout models, which may not accurately reflect the equivalent in humans, such that 'mild' phenotypes may correlate to highly undesirable side-effects in human patients. Thus, as is the case with SCA1 (154), this may depend upon the severity of the symptoms, in that even mild symptoms that are apparently well-tolerated in a murine model may be undesirable in the human condition.

As with the example of the redistribution of wild-type expression shown here, any therapy that can remove the mutant allele of a disease-causing dominant gene while retaining the endogenous functional allele would be ideal. However, finding a SNP linked to the disease in the polyglutamine disorders could prove difficult as is the case with HD. Interestingly, this difficulty is partly due to the heterogeneous populations studied. This is not the case in South Africa where the presence founder effects provide a more homogeneous patient cohort, and an increased opportunity of finding a single SNP that may be useful for a significant proportion of the group. Without this advantage, the mechanistic complexity involved in assembling an allele-specific therapy may be too great. Therefore, it is noted that where not possible, general knockdown of the target gene, both wild-type and mutant, may delay onset of the disease to such a time, that significant quality of life may be allowed for those patients who have inherited the expansion. This is especially the case when taking into account the polyglutamine disorders, for which a significant delay and not necessarily total elimination of disease pathogenesis, may push the disease onset to beyond a natural life expectancy. Incomplete knockdown of both alleles may lead to sufficient wild-type protein remaining to perform its function, and little enough mutant to delay the onset of progression to beyond the normal life expectancy resulting in a significant phenotypic benefit, originally suggested by Harper and colleagues (166).

Although our data has not been tested *in vivo*, the release of the wild-type protein suggests that allele-specificity may show significant phenotypic advantages over a general knockdown approach of both alleles.

4.3. Concluding remarks

In summary, this study has identified an RNAi effector sequence that can selectively knockdown the expression of the disease-causing gene in SCA7 patients. Extensive modifications to the design of the effector, as well as the various assays tested, indicate that allele-specific silencing using a weak G:U mismatch is attainable, but careful consideration must be made in this design and in the types of assays performed. While many hypotheses regarding the exploitation of single base pair changes can be drawn from the short target luciferase assay, it is likely that for the purposes of therapeutic approaches, full-length target assays should be tested to more accurately replicate the effects of the endogenous expression and sequence context. Further investigations of the effects of allele-specific silencing suggest that the additional benefit of removal of aggregates will release the wild-type protein to its normal expression pattern. It is likely that this reversal of a characteristic cellular polyglutamine phenotype will confer a significant advantage to allele-specific silencing in comparison to a general knockdown gene therapy approach. It is important to note that the unique macular degeneration phenotype of SCA7 model would provide an ideal model for studying the principle of allele-specific knockdown in patients with neurodegenerative disease. Given the data reported here, showing that even a moderate degree of mutant selectivity may have beneficial phenotypic effects through decreasing mutant aggregates and restoring wild-type protein to its native distribution pattern; such a study may have implications for many, if not all, of the currently untreatable polyglutamine disorders.

4.4. Future studies

Many of the questions regarding the mechanistics of single nucleotide discrimination effector design have yet to be answered. Although these have less impact on the major direction of this work, they would none-the-less be of great value to the study of guide strand structure and RISC function and requirements. Firstly, what other, if any, modifications could be incorporated to produce greater selectivity? Although extensive modifications were made in this study, screening the 5' region may show additional selectivity. No screen of mismatches in the 5' region of the guide strand has yet been performed using shRNAs, in particular targeting a G:U mismatch. Data from such an experiment would test the hypothesis that stronger structural constraints around the seed region transfer increased stability to mismatches by reducing local widening of the RNA duplex.

Determining the exact species of the guide strand produced by the shRNAs and miRNAs would also contribute to a better understanding of expressed hairpin design. While primer extension may serve to identify these sequences, given the remarkably similar success with shR-P16 and miR-P16, it is likely that our predictions are correct.

A further aspect to consider would be the biological consequences of introduction of such effectors. There are multiple avenues to investigate here, including the effect of the vector, the presence of modified hairpin constructs on the endogenous RNAi machinery, and the off-target effects of the guide sequence on the expression of other genes. The ideal way in which to test these would be to look at gene expression changes within patient cell lines using expression based micro-arrays, though thorough consideration of interpretations of this type of experiment would be required given the plethora of changes that could be identified.

The main focus of this study will now turn towards the therapeutic implications of successful allele-specific silencing and the phenotypic

advantages conferred by this approach. This is likely to require two different branches of investigation. Firstly, testing a primary cell line such as transformed lymphoblasts can be relatively easily obtained from SCA7 patients with the correct heterozygous G>A genotype, thus, the effects of these effectors can be measured in patient cell lines. It is likely that this will require the construction of viral vector mediated transductions given the difficulty in transfecting this cell line. Western blot analysis would enable a clear evaluation of the level specific knockdown. In addition, levels of heat shock proteins could be investigated, given that these have been shown to be reduced in SCA7 lymphoblast cell lines (78). A return to normal levels may not necessarily indicate the direct involvement of these proteins in pathogenesis, but may be an indicator of a return to the normal state, as seen with the release of the wild-type protein from aggregates.

More recently, the development of induced pluripotent stem (iPS) cell technology should allow for significantly improved patient-specific *ex vivo* cell-based investigations (217, 218). These and other studies indicate that fibroblasts from patients could be re-programmed into embryonic stem cells, differentiated into different cell types such as neurons, and investigated both for pathogenic mechanisms of disease, but also assessed for different therapeutic strategies. This would be particularly useful for polyglutamine disorders in that several aspects previously mentioned could be addressed, such as assessing the effect of different sizes of repeats as well as investigating the very cell types that degenerate.

Finally, the acme of gene therapy investigations would be to investigate the amelioration of pathogenic phenotypes using *in vivo* mouse models of SCA7. An ideal model would include the use of the transgenic SCA7 mouse which has an expanded repeat sequence within one of the endogenous mouse alleles, such that the effects of general knockdown versus allele-specific knockdown could be investigated. Improved delivery using viral vectors such as lentivirus and AAV would need to be investigated. Extensive phenotypic characterisation of this model would be crucial for this study especially with

reference to the macular phenotype. Intra-ocular delivery of an efficient vector would be more conducive to assessment, given the well-established method of gene therapy to this organ, and the natural presence of an internal control in the untreated eye.

In summary, the data from this study should pave the way for an ideal model system of investigating allele-specific silencing for the polyglutamine disorders. Although this has been applied to other polyglutamine disease models, the work described here represents the only study on SCA7 to date. Given the unique ocular involvement of SCA7, investigation of an allele-specific approach for this disorder holds the greatest potential for the realisation of attainable gene therapy for inherited neurodegenerative diseases.

References

References have been listed in the format of Human Molecular Genetics journal.

1. Menken, M., Munsat, T.L. and Toole, J.F. (2000) The global burden of disease study: implications for neurology. *Arch Neurol,* **57,** 418-20.
2. Bergen, D.C. and Silberberg, D. (2002) Nervous system disorders: a global epidemic. *Arch Neurol,* **59,** 1194-6.
3. Orr, H.T. and Zoghbi, H.Y. (2007) Trinucleotide repeat disorders. *Annu Rev Neurosci,* **30,** 575-621.
4. Lindblad, K., Savontaus, M.L., Stevanin, G., Holmberg, M., Digre, K., Zander, C., Ehrsson, H., David, G., Benomar, A., Nikoskelainen, E. *et al.* (1996) An expanded CAG repeat sequence in spinocerebellar ataxia type 7. *Genome Res,* **6,** 965-71.
5. Rosenblatt, A., Brinkman, R.R., Liang, K.Y., Almqvist, E.W., Margolis, R.L., Huang, C.Y., Sherr, M., Franz, M.L., Abbott, M.H., Hayden, M.R. *et al.* (2001) Familial influence on age of onset among siblings with Huntington disease. *Am J Med Genet,* **105,** 399-403.
6. Shao, J. and Diamond, M.I. (2007) Polyglutamine diseases: emerging concepts in pathogenesis and therapy. *Hum Mol Genet,* **16 Spec No. 2,** R115-23.
7. Harding, A.E. (1993) Clinical features and classification of inherited ataxias. *Adv Neurol,* **61,** 1-14.
8. Duenas, A.M., Goold, R. and Giunti, P. (2006) Molecular pathogenesis of spinocerebellar ataxias. *Brain,* **129,** 1357-70.
9. Giunti, P., Stevanin, G., Worth, P.F., David, G., Brice, A. and Wood, N.W. (1999) Molecular and clinical study of 18 families with ADCA type II: evidence for genetic heterogeneity and de novo mutation. *Am J Hum Genet,* **64,** 1594-603.
10. Aleman, T.S., Cideciyan, A.V., Volpe, N.J., Stevanin, G., Brice, A. and Jacobson, S.G. (2002) Spinocerebellar ataxia type 7 (SCA7) shows a cone-rod dystrophy phenotype. *Exp Eye Res,* **74,** 737-45.
11. Garden, G.A. and La Spada, A.R. (2007) Molecular pathogenesis and cellular pathology of spinocerebellar ataxia type 7 neurodegeneration. *Cerebellum,* 1-12.
12. Johansson, J., Forsgren, L., Sandgren, O., Brice, A., Holmgren, G. and Holmberg, M. (1998) Expanded CAG repeats in Swedish spinocerebellar ataxia type 7 (SCA7) patients: effect of CAG repeat length on the clinical manifestation. *Hum Mol Genet,* **7,** 171-6.
13. Benton, C.S., de Silva, R., Rutledge, S.L., Bohlega, S., Ashizawa, T. and Zoghbi, H.Y. (1998) Molecular and clinical studies in SCA-7 define a broad clinical spectrum and the infantile phenotype. *Neurology,* **51,** 1081-6.
14. Whitney, A., Lim, M., Kanabar, D. and Lin, J.P. (2007) Massive SCA7 expansion detected in a 7-month-old male with hypotonia, cardiomegaly, and renal compromise. *Dev Med Child Neurol,* **49,** 140-3.

15. van de Warrenburg, B.P., Frenken, C.W., Ausems, M.G., Kleefstra, T., Sinke, R.J., Knoers, N.V. and Kremer, H.P. (2001) Striking anticipation in spinocerebellar ataxia type 7: the infantile phenotype. *J Neurol,* **248,** 911-4.
16. Benomar, A., Krols, L., Stevanin, G., Cancel, G., LeGuern, E., David, G., Ouhabi, H., Martin, J.J., Durr, A., Zaim, A. *et al.* (1995) The gene for autosomal dominant cerebellar ataxia with pigmentary macular dystrophy maps to chromosome 3p12-p21.1. *Nat Genet,* **10,** 84-8.
17. David, G., Giunti, P., Abbas, N., Coullin, P., Stevanin, G., Horta, W., Gemmill, R., Weissenbach, J., Wood, N., Cunha, S. *et al.* (1996) The gene for autosomal dominant cerebellar ataxia type II is located in a 5-cM region in 3p12-p13: genetic and physical mapping of the SCA7 locus. *Am J Hum Genet,* **59,** 1328-36.
18. David, G., Abbas, N., Stevanin, G., Durr, A., Yvert, G., Cancel, G., Weber, C., Imbert, G., Saudou, F., Antoniou, E. *et al.* (1997) Cloning of the SCA7 gene reveals a highly unstable CAG repeat expansion. *Nat Genet,* **17,** 65-70.
19. Michalik, A., Del-Favero, J., Mauger, C., Lofgren, A. and Van Broeckhoven, C. (1999) Genomic organisation of the spinocerebellar ataxia type 7 (SCA7) gene responsible for autosomal dominant cerebellar ataxia with retinal degeneration. *Hum Genet,* **105,** 410-7.
20. Einum, D.D., Clark, A.M., Townsend, J.J., Ptacek, L.J. and Fu, Y.H. (2003) A novel central nervous system-enriched spinocerebellar ataxia type 7 gene product. *Arch Neurol,* **60,** 97-103.
21. Strom, A.L., Forsgren, L. and Holmberg, M. (2005) A role for both wild-type and expanded ataxin-7 in transcriptional regulation. *Neurobiol Dis,* **20,** 646-55.
22. David, G., Durr, A., Stevanin, G., Cancel, G., Abbas, N., Benomar, A., Belal, S., Lebre, A.S., Abada-Bendib, M., Grid, D. *et al.* (1998) Molecular and clinical correlations in autosomal dominant cerebellar ataxia with progressive macular dystrophy (SCA7). *Hum Mol Genet,* **7,** 165-70.
23. Alluri, R.V., Komandur, S., Wagheray, A., Chaudhuri, J.R., Sitajayalakshmi, Meena, A.K., Jabeen, A., Chawda, K., Subhash, K., Krishnaveni, A. *et al.* (2007) Molecular analysis of CAG repeats at five different spinocerebellar ataxia loci: correlation and alternative explanations for disease pathogenesis. *Mol Cells,* **24,** 338-42.
24. Stevanin, G., David, G., Durr, A., Giunti, P., Benomar, A., Abada-Bendib, M., Lee, M.S., Agid, Y. and Brice, A. (1999) Multiple origins of the spinocerebellar ataxia 7 (SCA7) mutation revealed by linkage disequilibrium studies with closely flanking markers, including an intragenic polymorphism (G3145TG/A3145TG). *Eur J Hum Genet,* **7,** 889-96.
25. Nardacchione, A., Orsi, L., Brusco, A., Franco, A., Grosso, E., Dragone, E., Mortara, P., Schiffer, D. and De Marchi, M. (1999) Definition of the smallest pathological CAG expansion in SCA7. *Clin Genet,* **56,** 232-4.
26. Monckton, D.G., Cayuela, M.L., Gould, F.K., Brock, G.J., Silva, R. and Ashizawa, T. (1999) Very large (CAG)(n) DNA repeat expansions in the sperm of two spinocerebellar ataxia type 7 males. *Hum Mol Genet,* **8,** 2473-8.
27. Mandel, J.L. (1994) Trinucleotide diseases on the rise. *Nat Genet,* **7,** 453-5.

28. Trottier, Y., Biancalana, V. and Mandel, J.L. (1994) Instability of CAG repeats in Huntington's disease: relation to parental transmission and age of onset. *J Med Genet,* **31,** 377-82.
29. Yoon, S.R., Dubeau, L., de Young, M., Wexler, N.S. and Arnheim, N. (2003) Huntington disease expansion mutations in humans can occur before meiosis is completed. *Proc Natl Acad Sci U S A,* **100,** 8834-8.
30. Libby, R.T., Monckton, D.G., Fu, Y.H., Martinez, R.A., McAbney, J.P., Lau, R., Einum, D.D., Nichol, K., Ware, C.B., Ptacek, L.J. *et al.* (2003) Genomic context drives SCA7 CAG repeat instability, while expressed SCA7 cDNAs are intergenerationally and somatically stable in transgenic mice. *Hum Mol Genet,* **12,** 41-50.
31. Hayden, M.R. (1981) *Huntington's Chorea*. Springer-Verlag, Berlin Heidelberg.
32. La Spada, A.R., Roling, D.B., Harding, A.E., Warner, C.L., Spiegel, R., Hausmanowa-Petrusewicz, I., Yee, W.C. and Fischbeck, K.H. (1992) Meiotic stability and genotype-phenotype correlation of the trinucleotide repeat in X-linked spinal and bulbar muscular atrophy. *Nat Genet,* **2,** 301-4.
33. Rub, U., Brunt, E.R., Gierga, K., Seidel, K., Schultz, C., Schols, L., Auburger, G., Heinsen, H., Ippel, P.F., Glimmerveen, W.F. *et al.* (2005) Spinocerebellar ataxia type 7 (SCA7): first report of a systematic neuropathological study of the brain of a patient with a very short expanded CAG-repeat. *Brain Pathol,* **15,** 287-95.
34. Yamada, M., Sato, T., Tsuji, S. and Takahashi, H. (2008) CAG repeat disorder models and human neuropathology: similarities and differences. *Acta Neuropathol,* **115,** 71-86.
35. Holmberg, M., Duyckaerts, C., Durr, A., Cancel, G., Gourfinkel-An, I., Damier, P., Faucheux, B., Trottier, Y., Hirsch, E.C., Agid, Y. *et al.* (1998) Spinocerebellar ataxia type 7 (SCA7): a neurodegenerative disorder with neuronal intranuclear inclusions. *Hum Mol Genet,* **7,** 913-8.
36. Chen, S., Peng, G.H., Wang, X., Smith, A.C., Grote, S.K., Sopher, B.L. and La Spada, A.R. (2004) Interference of Crx-dependent transcription by ataxin-7 involves interaction between the glutamine regions and requires the ataxin-7 carboxy-terminal region for nuclear localization. *Hum Mol Genet,* **13,** 53-67.
37. To, K.W., Adamian, M., Jakobiec, F.A. and Berson, E.L. (1993) Olivopontocerebellar atrophy with retinal degeneration. An electroretinographic and histopathologic investigation. *Ophthalmology,* **100,** 15-23.
38. Gouw, L.G., Digre, K.B., Harris, C.P., Haines, J.H. and Ptacek, L.J. (1994) Autosomal dominant cerebellar ataxia with retinal degeneration: clinical, neuropathologic, and genetic analysis of a large kindred. *Neurology,* **44,** 1441-7.
39. Kaytor, M.D., Duvick, L.A., Skinner, P.J., Koob, M.D., Ranum, L.P. and Orr, H.T. (1999) Nuclear localization of the spinocerebellar ataxia type 7 protein, ataxin-7. *Hum Mol Genet,* **8,** 1657-64.
40. Mushegian, A.R., Vishnivetskiy, S.A. and Gurevich, V.V. (2000) Conserved phosphoprotein interaction motif is functionally interchangeable between ataxin-7 and arrestins. *Biochemistry,* **39,** 6809-13.

41. Strom, A.L., Jonasson, J., Hart, P., Brannstrom, T., Forsgren, L. and Holmberg, M. (2002) Cloning and expression analysis of the murine homolog of the spinocerebellar ataxia type 7 (SCA7) gene. *Gene,* **285,** 91-9.
42. Zander, C., Takahashi, J., El Hachimi, K.H., Fujigasaki, H., Albanese, V., Lebre, A.S., Stevanin, G., Duyckaerts, C. and Brice, A. (2001) Similarities between spinocerebellar ataxia type 7 (SCA7) cell models and human brain: proteins recruited in inclusions and activation of caspase-3. *Hum Mol Genet,* **10,** 2569-79.
43. Taylor, J., Grote, S.K., Xia, J., Vandelft, M., Graczyk, J., Ellerby, L.M., La Spada, A.R. and Truant, R. (2006) Ataxin-7 can export from the nucleus via a conserved exportin-dependent signal. *J Biol Chem,* **281,** 2730-9.
44. Einum, D.D., Townsend, J.J., Ptacek, L.J. and Fu, Y.H. (2001) Ataxin-7 expression analysis in controls and spinocerebellar ataxia type 7 patients. *Neurogenetics,* **3,** 83-90.
45. Cancel, G., Duyckaerts, C., Holmberg, M., Zander, C., Yvert, G., Lebre, A.S., Ruberg, M., Faucheux, B., Agid, Y., Hirsch, E. *et al.* (2000) Distribution of ataxin-7 in normal human brain and retina. *Brain,* **123 Pt 12,** 2519-30.
46. Jonasson, J., Strom, A.L., Hart, P., Brannstrom, T., Forsgren, L. and Holmberg, M. (2002) Expression of ataxin-7 in CNS and non-CNS tissue of normal and SCA7 individuals. *Acta Neuropathol,* **104,** 29-37.
47. Lebre, A.S., Jamot, L., Takahashi, J., Spassky, N., Leprince, C., Ravise, N., Zander, C., Fujigasaki, H., Kussel-Andermann, P., Duyckaerts, C. *et al.* (2001) Ataxin-7 interacts with a Cbl-associated protein that it recruits into neuronal intranuclear inclusions. *Hum Mol Genet,* **10,** 1201-13.
48. Matilla, A., Gorbea, C., Einum, D.D., Townsend, J., Michalik, A., van Broeckhoven, C., Jensen, C.C., Murphy, K.J., Ptacek, L.J. and Fu, Y.H. (2001) Association of ataxin-7 with the proteasome subunit S4 of the 19S regulatory complex. *Hum Mol Genet,* **10,** 2821-31.
49. Chai, Y., Koppenhafer, S.L., Shoesmith, S.J., Perez, M.K. and Paulson, H.L. (1999) Evidence for proteasome involvement in polyglutamine disease: localization to nuclear inclusions in SCA3/MJD and suppression of polyglutamine aggregation in vitro. *Hum Mol Genet,* **8,** 673-82.
50. La Spada, A.R., Fu, Y.H., Sopher, B.L., Libby, R.T., Wang, X., Li, L.Y., Einum, D.D., Huang, J., Possin, D.E., Smith, A.C. *et al.* (2001) Polyglutamine-expanded ataxin-7 antagonizes CRX function and induces cone-rod dystrophy in a mouse model of SCA7. *Neuron,* **31,** 913-27.
51. Sohocki, M.M., Sullivan, L.S., Mintz-Hittner, H.A., Birch, D., Heckenlively, J.R., Freund, C.L., McInnes, R.R. and Daiger, S.P. (1998) A range of clinical phenotypes associated with mutations in CRX, a photoreceptor transcription-factor gene. *Am J Hum Genet,* **63,** 1307-15.
52. Furukawa, T., Morrow, E.M., Li, T., Davis, F.C. and Cepko, C.L. (1999) Retinopathy and attenuated circadian entrainment in Crx-deficient mice. *Nat Genet,* **23,** 466-70.
53. Scheel, H., Tomiuk, S. and Hofmann, K. (2003) Elucidation of ataxin-3 and ataxin-7 function by integrative bioinformatics. *Hum Mol Genet,* **12,** 2845-52.
54. Helmlinger, D., Hardy, S., Sasorith, S., Klein, F., Robert, F., Weber, C., Miguet, L., Potier, N., Van-Dorsselaer, A., Wurtz, J.M. *et al.* (2004) Ataxin-7 is a subunit

of GCN5 histone acetyltransferase-containing complexes. *Hum Mol Genet*, **13,** 1257-65.

55. Palhan, V.B., Chen, S., Peng, G.H., Tjernberg, A., Gamper, A.M., Fan, Y., Chait, B.T., La Spada, A.R. and Roeder, R.G. (2005) Polyglutamine-expanded ataxin-7 inhibits STAGA histone acetyltransferase activity to produce retinal degeneration. *Proc Natl Acad Sci U S A*, **102,** 8472-7.
56. Mauger, C., Del-Favero, J., Ceuterick, C., Lubke, U., van Broeckhoven, C. and Martin, J. (1999) Identification and localization of ataxin-7 in brain and retina of a patient with cerebellar ataxia type II using anti-peptide antibody. *Brain Res Mol Brain Res*, **74,** 35-43.
57. Latouche, M., Fragner, P., Martin, E., El Hachimi, K.H., Zander, C., Sittler, A., Ruberg, M., Brice, A. and Stevanin, G. (2006) Polyglutamine and polyalanine expansions in ataxin7 result in different types of aggregation and levels of toxicity. *Mol Cell Neurosci*, **31,** 438-45.
58. Ansorge, O., Giunti, P., Michalik, A., Van Broeckhoven, C., Harding, B., Wood, N. and Scaravilli, F. (2004) Ataxin-7 aggregation and ubiquitination in infantile SCA7 with 180 CAG repeats. *Ann Neurol*, **56,** 448-52.
59. Takahashi, J., Fujigasaki, H., Zander, C., El Hachimi, K.H., Stevanin, G., Durr, A., Lebre, A.S., Yvert, G., Trottier, Y., de The, H. *et al.* (2002) Two populations of neuronal intranuclear inclusions in SCA7 differ in size and promyelocytic leukaemia protein content. *Brain*, **125,** 1534-43.
60. Bowman, A.B., Yoo, S.Y., Dantuma, N.P. and Zoghbi, H.Y. (2005) Neuronal dysfunction in a polyglutamine disease model occurs in the absence of ubiquitin-proteasome system impairment and inversely correlates with the degree of nuclear inclusion formation. *Hum Mol Genet*, **14,** 679-91.
61. Wheeler, V.C., White, J.K., Gutekunst, C.A., Vrbanac, V., Weaver, M., Li, X.J., Li, S.H., Yi, H., Vonsattel, J.P., Gusella, J.F. *et al.* (2000) Long glutamine tracts cause nuclear localization of a novel form of huntingtin in medium spiny striatal neurons in HdhQ92 and HdhQ111 knock-in mice. *Hum Mol Genet*, **9,** 503-13.
62. Wang, H.L., Yeh, T.H., Chou, A.H., Kuo, Y.L., Luo, L.J., He, C.Y., Huang, P.C. and Li, A.H. (2006) Polyglutamine-expanded ataxin-7 activates mitochondrial apoptotic pathway of cerebellar neurons by upregulating Bax and downregulating Bcl-x(L). *Cell Signal*, **18,** 541-52.
63. McMahon, S.J., Pray-Grant, M.G., Schieltz, D., Yates, J.R., 3rd and Grant, P.A. (2005) Polyglutamine-expanded spinocerebellar ataxia-7 protein disrupts normal SAGA and SLIK histone acetyltransferase activity. *Proc Natl Acad Sci U S A*, **102,** 8478-82.
64. Riley, B.E., Zoghbi, H.Y. and Orr, H.T. (2005) SUMOylation of the polyglutamine repeat protein, ataxin-1, is dependent on a functional nuclear localization signal. *J Biol Chem*, **280,** 21942-8.
65. Xia, J., Lee, D.H., Taylor, J., Vandelft, M. and Truant, R. (2003) Huntingtin contains a highly conserved nuclear export signal. *Hum Mol Genet*, **12,** 1393-403.
66. Ellerby, L.M., Andrusiak, R.L., Wellington, C.L., Hackam, A.S., Propp, S.S., Wood, J.D., Sharp, A.H., Margolis, R.L., Ross, C.A., Salvesen, G.S. *et al.* (1999)

Cleavage of atrophin-1 at caspase site aspartic acid 109 modulates cytotoxicity. *J Biol Chem,* **274,** 8730-6.

67. Ellerby, L.M., Hackam, A.S., Propp, S.S., Ellerby, H.M., Rabizadeh, S., Cashman, N.R., Trifiro, M.A., Pinsky, L., Wellington, C.L., Salvesen, G.S. *et al.* (1999) Kennedy's disease: caspase cleavage of the androgen receptor is a crucial event in cytotoxicity. *J Neurochem,* **72,** 185-95.
68. Gafni, J., Hermel, E., Young, J.E., Wellington, C.L., Hayden, M.R. and Ellerby, L.M. (2004) Inhibition of calpain cleavage of huntingtin reduces toxicity: accumulation of calpain/caspase fragments in the nucleus. *J Biol Chem,* **279,** 20211-20.
69. Kim, Y.J., Yi, Y., Sapp, E., Wang, Y., Cuiffo, B., Kegel, K.B., Qin, Z.H., Aronin, N. and DiFiglia, M. (2001) Caspase 3-cleaved N-terminal fragments of wild-type and mutant huntingtin are present in normal and Huntington's disease brains, associate with membranes, and undergo calpain-dependent proteolysis. *Proc Natl Acad Sci U S A,* **98,** 12784-9.
70. Garden, G.A., Libby, R.T., Fu, Y.H., Kinoshita, Y., Huang, J., Possin, D.E., Smith, A.C., Martinez, R.A., Fine, G.C., Grote, S.K. *et al.* (2002) Polyglutamine-expanded ataxin-7 promotes non-cell-autonomous purkinje cell degeneration and displays proteolytic cleavage in ataxic transgenic mice. *J Neurosci,* **22,** 4897-905.
71. Young, J.E., Gouw, L., Propp, S., Sopher, B.L., Taylor, J., Lin, A., Hermel, E., Logvinova, A., Chen, S.F., Chen, S. *et al.* (2007) Proteolytic cleavage of ataxin-7 by caspase-7 modulates cellular toxicity and transcriptional dysregulation. *J Biol Chem,* **282,** 30150-60.
72. Helmlinger, D., Abou-Sleymane, G., Yvert, G., Rousseau, S., Weber, C., Trottier, Y., Mandel, J.L. and Devys, D. (2004) Disease progression despite early loss of polyglutamine protein expression in SCA7 mouse model. *J Neurosci,* **24,** 1881-7.
73. Yvert, G., Lindenberg, K.S., Picaud, S., Landwehrmeyer, G.B., Sahel, J.A. and Mandel, J.L. (2000) Expanded polyglutamines induce neurodegeneration and trans-neuronal alterations in cerebellum and retina of SCA7 transgenic mice. *Hum Mol Genet,* **9,** 2491-506.
74. Yvert, G., Lindenberg, K.S., Devys, D., Helmlinger, D., Landwehrmeyer, G.B. and Mandel, J.L. (2001) SCA7 mouse models show selective stabilization of mutant ataxin-7 and similar cellular responses in different neuronal cell types. *Hum Mol Genet,* **10,** 1679-92.
75. Opal, P. and Zoghbi, H.Y. (2002) The role of chaperones in polyglutamine disease. *Trends Mol Med,* **8,** 232-6.
76. Sakahira, H., Breuer, P., Hayer-Hartl, M.K. and Hartl, F.U. (2002) Molecular chaperones as modulators of polyglutamine protein aggregation and toxicity. *Proc Natl Acad Sci U S A,* **99 Suppl 4,** 16412-8.
77. Helmlinger, D., Bonnet, J., Mandel, J.L., Trottier, Y. and Devys, D. (2004) Hsp70 and Hsp40 chaperones do not modulate retinal phenotype in SCA7 mice. *J Biol Chem,* **279,** 55969-77.

78. Tsai, H.F., Lin, S.J., Li, C. and Hsieh, M. (2005) Decreased expression of Hsp27 and Hsp70 in transformed lymphoblastoid cells from patients with spinocerebellar ataxia type 7. *Biochem Biophys Res Commun,* **334,** 1279-86.
79. Mallory, A.C., Reinhart, B.J., Jones-Rhoades, M.W., Tang, G., Zamore, P.D., Barton, M.K. and Bartel, D.P. (2004) MicroRNA control of PHABULOSA in leaf development: importance of pairing to the microRNA 5' region. *Embo J,* **23,** 3356-64.
80. Abou-Sleymane, G., Chalmel, F., Helmlinger, D., Lardenois, A., Thibault, C., Weber, C., Merienne, K., Mandel, J.L., Poch, O., Devys, D. *et al.* (2006) Polyglutamine expansion causes neurodegeneration by altering the neuronal differentiation program. *Hum Mol Genet,* **15,** 691-703.
81. Merienne, K., Friedman, J., Akimoto, M., Abou-Sleymane, G., Weber, C., Swaroop, A. and Trottier, Y. (2007) Preventing polyglutamine-induced activation of c-Jun delays neuronal dysfunction in a mouse model of SCA7 retinopathy. *Neurobiol Dis,* **25,** 571-81.
82. Gatchel, J.R., Watase, K., Thaller, C., Carson, J.P., Jafar-Nejad, P., Shaw, C., Zu, T., Orr, H.T. and Zoghbi, H.Y. (2008) The insulin-like growth factor pathway is altered in spinocerebellar ataxia type 1 and type 7. *Proc Natl Acad Sci U S A,* **105,** 1291-6.
83. Custer, S.K., Garden, G.A., Gill, N., Rueb, U., Libby, R.T., Schultz, C., Guyenet, S.J., Deller, T., Westrum, L.E., Sopher, B.L. *et al.* (2006) Bergmann glia expression of polyglutamine-expanded ataxin-7 produces neurodegeneration by impairing glutamate transport. *Nat Neurosci,* **9,** 1302-11.
84. Yoo, S.Y., Pennesi, M.E., Weeber, E.J., Xu, B., Atkinson, R., Chen, S., Armstrong, D.L., Wu, S.M., Sweatt, J.D. and Zoghbi, H.Y. (2003) SCA7 knockin mice model human SCA7 and reveal gradual accumulation of mutant ataxin-7 in neurons and abnormalities in short-term plasticity. *Neuron,* **37,** 383-401.
85. Humphries, M.M., Rancourt, D., Farrar, G.J., Kenna, P., Hazel, M., Bush, R.A., Sieving, P.A., Sheils, D.M., McNally, N., Creighton, P. *et al.* (1997) Retinopathy induced in mice by targeted disruption of the rhodopsin gene. *Nat Genet,* **15,** 216-9.
86. Latouche, M., Lasbleiz, C., Martin, E., Monnier, V., Debeir, T., Mouatt-Prigent, A., Muriel, M.P., Morel, L., Ruberg, M., Brice, A. *et al.* (2007) A conditional pan-neuronal Drosophila model of spinocerebellar ataxia 7 with a reversible adult phenotype suitable for identifying modifier genes. *J Neurosci,* **27,** 2483-92.
87. Storey, E., du Sart, D., Shaw, J.H., Lorentzos, P., Kelly, L., McKinley Gardner, R.J., Forrest, S.M., Biros, I. and Nicholson, G.A. (2000) Frequency of spinocerebellar ataxia types 1, 2, 3, 6, and 7 in Australian patients with spinocerebellar ataxia. *Am J Med Genet,* **95,** 351-7.
88. Bird, T.D. (1998) Hereditary Ataxia Overview
89. Jonasson, J., Juvonen, V., Sistonen, P., Ignatius, J., Johansson, D., Bjorck, E.J., Wahlstrom, J., Melberg, A., Holmgren, G., Forsgren, L. *et al.* (2000) Evidence for a common Spinocerebellar ataxia type 7 (SCA7) founder mutation in Scandinavia. *Eur J Hum Genet,* **8,** 918-22.

90. Bryer, A., Krause, A., Bill, P., Davids, V., Bryant, D., Butler, J., Heckmann, J., Ramesar, R. and Greenberg, J. (2003) The hereditary adult-onset ataxias in South Africa. *Journal of the neurological sciences,* **216,** 47-54.
91. Greenberg, J., Solomon, G.A., Vorster, A.A., Heckmann, J. and Bryer, A. (2006) Origin of the SCA7 gene mutation in South Africa: implications for molecular diagnostics. *Clin Genet,* **70,** 415-7.
92. Meissner, P.N., Dailey, T.A., Hift, R.J., Ziman, M., Corrigall, A.V., Roberts, A.G., Meissner, D.M., Kirsch, R.E. and Dailey, H.A. (1996) A R59W mutation in human protoporphyrinogen oxidase results in decreased enzyme activity and is prevalent in South Africans with variegate porphyria. *Nat Genet,* **13,** 95-7.
93. Scholefield, J. and Greenberg, J. (2007) A common SNP haplotype provides molecular proof of a founder effect of Huntington disease linking two South African populations. *Eur J Hum Genet,* **15,** 590-5.
94. Napoli, C., Lemieux, C. and Jorgensen, R. (1990) Introduction of a Chimeric Chalcone Synthase Gene into Petunia Results in Reversible Co-Suppression of Homologous Genes in trans. *Plant Cell,* **2,** 279-289.
95. van der Krol, A.R., Mur, L.A., Beld, M., Mol, J.N. and Stuitje, A.R. (1990) Flavonoid genes in petunia: addition of a limited number of gene copies may lead to a suppression of gene expression. *Plant Cell,* **2,** 291-9.
96. Pal-Bhadra, M., Bhadra, U. and Birchler, J.A. (1997) Cosuppression in Drosophila: gene silencing of Alcohol dehydrogenase by white-Adh transgenes is Polycomb dependent. *Cell,* **90,** 479-90.
97. Ratcliff, F., Harrison, B.D. and Baulcombe, D.C. (1997) A similarity between viral defense and gene silencing in plants. *Science,* **276,** 1558-60.
98. Fire, A., Albertson, D., Harrison, S.W. and Moerman, D.G. (1991) Production of antisense RNA leads to effective and specific inhibition of gene expression in C. elegans muscle. *Development,* **113,** 503-14.
99. Guo, S. and Kemphues, K.J. (1995) par-1, a gene required for establishing polarity in C. elegans embryos, encodes a putative Ser/Thr kinase that is asymmetrically distributed. *Cell,* **81,** 611-20.
100. Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E. and Mello, C.C. (1998) Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature,* **391,** 806-11.
101. Clemens, M.J. (1997) PKR--a protein kinase regulated by double-stranded RNA. *Int J Biochem Cell Biol,* **29,** 945-9.
102. Hamilton, A.J. and Baulcombe, D.C. (1999) A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science,* **286,** 950-2.
103. Caplen, N.J., Parrish, S., Imani, F., Fire, A. and Morgan, R.A. (2001) Specific inhibition of gene expression by small double-stranded RNAs in invertebrate and vertebrate systems. *Proc Natl Acad Sci U S A,* **98,** 9742-7.
104. Elbashir, S.M., Lendeckel, W. and Tuschl, T. (2001) RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev,* **15,** 188-200.
105. Rana, T.M. (2007) Illuminating the silence: understanding the structure and function of small RNAs. *Nat Rev Mol Cell Biol,* **8,** 23-36.
106. Borchert, G.M., Lanier, W. and Davidson, B.L. (2006) RNA polymerase III transcribes human microRNAs. *Nat Struct Mol Biol,* **13,** 1097-101.

107. Han, J., Lee, Y., Yeom, K.H., Nam, J.W., Heo, I., Rhee, J.K., Sohn, S.Y., Cho, Y., Zhang, B.T. and Kim, V.N. (2006) Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. *Cell,* **125,** 887-901.
108. Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Radmark, O., Kim, S. *et al.* (2003) The nuclear RNase III Drosha initiates microRNA processing. *Nature,* **425,** 415-9.
109. Lund, E., Guttinger, S., Calado, A., Dahlberg, J.E. and Kutay, U. (2004) Nuclear export of microRNA precursors. *Science,* **303,** 95-8.
110. Yi, R., Qin, Y., Macara, I.G. and Cullen, B.R. (2003) Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev,* **17,** 3011-6.
111. Bernstein, E., Caudy, A.A., Hammond, S.M. and Hannon, G.J. (2001) Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature,* **409,** 363-6.
112. Gonzalez-Alegre, P. and Paulson, H.L. (2007) Technology insight: therapeutic RNA interference--how far from the neurology clinic? *Nat Clin Pract Neurol,* **3,** 394-404.
113. Macrae, I.J., Zhou, K., Li, F., Repic, A., Brooks, A.N., Cande, W.Z., Adams, P.D. and Doudna, J.A. (2006) Structural basis for double-stranded RNA processing by Dicer. *Science,* **311,** 195-8.
114. Hammond, S.M., Bernstein, E., Beach, D. and Hannon, G.J. (2000) An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. *Nature,* **404,** 293-6.
115. Nykanen, A., Haley, B. and Zamore, P.D. (2001) ATP requirements and small interfering RNA structure in the RNA interference pathway. *Cell,* **107,** 309-21.
116. Martinez, J., Patkaniowska, A., Urlaub, H., Luhrmann, R. and Tuschl, T. (2002) Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. *Cell,* **110,** 563-74.
117. Khvorova, A., Reynolds, A. and Jayasena, S.D. (2003) Functional siRNAs and miRNAs exhibit strand bias. *Cell,* **115,** 209-16.
118. Schwarz, D.S., Hutvagner, G., Du, T., Xu, Z., Aronin, N. and Zamore, P.D. (2003) Asymmetry in the assembly of the RNAi enzyme complex. *Cell,* **115,** 199-208.
119. Matranga, C., Tomari, Y., Shin, C., Bartel, D.P. and Zamore, P.D. (2005) Passenger-strand cleavage facilitates assembly of siRNA into Ago2-containing RNAi enzyme complexes. *Cell,* **123,** 607-20.
120. Rand, T.A., Petersen, S., Du, F. and Wang, X. (2005) Argonaute2 cleaves the anti-guide strand of siRNA during RISC activation. *Cell,* **123,** 621-9.
121. Leuschner, P.J., Ameres, S.L., Kueng, S. and Martinez, J. (2006) Cleavage of the siRNA passenger strand during RISC assembly in human cells. *EMBO Rep,* **7,** 314-20.
122. Valencia-Sanchez, M.A., Liu, J., Hannon, G.J. and Parker, R. (2006) Control of translation and mRNA degradation by miRNAs and siRNAs. *Genes Dev,* **20,** 515-24.
123. Doench, J.G. and Sharp, P.A. (2004) Specificity of microRNA target selection in translational repression. *Genes Dev,* **18,** 504-11.

124. Ameres, S.L., Martinez, J. and Schroeder, R. (2007) Molecular basis for target RNA recognition and cleavage by human RISC. *Cell,* **130,** 101-12.
125. Liu, J., Carmell, M.A., Rivas, F.V., Marsden, C.G., Thomson, J.M., Song, J.J., Hammond, S.M., Joshua-Tor, L. and Hannon, G.J. (2004) Argonaute2 is the catalytic engine of mammalian RNAi. *Science,* **305,** 1437-41.
126. Zeng, Y., Yi, R. and Cullen, B.R. (2003) MicroRNAs and small interfering RNAs can inhibit mRNA expression by similar mechanisms. *Proc Natl Acad Sci U S A,* **100,** 9779-84.
127. Yekta, S., Shih, I.H. and Bartel, D.P. (2004) MicroRNA-directed cleavage of HOXB8 mRNA. *Science,* **304,** 594-6.
128. Brummelkamp, T.R., Bernards, R. and Agami, R. (2002) A system for stable expression of short interfering RNAs in mammalian cells. *Science,* **296,** 550-3.
129. Paddison, P.J., Caudy, A.A. and Hannon, G.J. (2002) Stable suppression of gene expression by RNAi in mammalian cells. *Proc Natl Acad Sci U S A,* **99,** 1443-8.
130. Zeng, Y. and Cullen, B.R. (2003) Sequence requirements for micro RNA processing and function in human cells. *Rna,* **9,** 112-23.
131. Zeng, Y., Wagner, E.J. and Cullen, B.R. (2002) Both natural and designed micro RNAs can inhibit the expression of cognate mRNAs when expressed in human cells. *Mol Cell,* **9,** 1327-33.
132. Zhou, H., Xia, X.G. and Xu, Z. (2005) An RNA polymerase II construct synthesizes short-hairpin RNA with a quantitative indicator and mediates highly efficient RNAi. *Nucleic Acids Res,* **33,** e62.
133. Silva, J.M., Li, M.Z., Chang, K., Ge, W., Golding, M.C., Rickles, R.J., Siolas, D., Hu, G., Paddison, P.J., Schlabach, M.R. *et al.* (2005) Second-generation shRNA libraries covering the mouse and human genomes. *Nat Genet,* **37,** 1281-8.
134. Dickins, R.A., Hemann, M.T., Zilfou, J.T., Simpson, D.R., Ibarra, I., Hannon, G.J. and Lowe, S.W. (2005) Probing tumor phenotypes using stable and regulated synthetic microRNA precursors. *Nat Genet,* **37,** 1289-95.
135. Grimm, D., Streetz, K.L., Jopling, C.L., Storm, T.A., Pandey, K., Davis, C.R., Marion, P., Salazar, F. and Kay, M.A. (2006) Fatality in mice due to oversaturation of cellular microRNA/short hairpin RNA pathways. *Nature,* **441,** 537-41.
136. Snove, O., Jr. and Rossi, J.J. (2006) Toxicity in mice expressing short hairpin RNAs gives new insight into RNAi. *Genome Biol,* **7,** 231.
137. Kay, M. (2008) Functional Evidence for Restricting MicroRNA Targets to the 3' Untranslated Region of mRNAs *RNAi, MicroRNA, and Non-Coding RNA* Whistler Resort, British Columbia, Canada.
138. Caplen, N.J. (1998) Gene therapy: different strategies for different applications. American Society of Gene Therapy: First Annual Meeting, Seattle, Washington, USA, 28-31 May 1998. *Mol Med Today,* **4,** 374-5.
139. Smith, R.A., Miller, T.M., Yamanaka, K., Monia, B.P., Condon, T.P., Hung, G., Lobsiger, C.S., Ward, C.M., McAlonis-Downes, M., Wei, H. *et al.* (2006) Antisense oligonucleotide therapy for neurodegenerative disease. *The Journal of clinical investigation,* **116,** 2290-6.

140. Hayashita-Kinoh, H., Yamada, M., Yokota, T., Mizuno, Y. and Mochizuki, H. (2006) Down-regulation of alpha-synuclein expression can rescue dopaminergic cells from cell death in the substantia nigra of Parkinson's disease rat model. *Biochemical and biophysical research communications,* **341,** 1088-95.
141. Elbashir, S.M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K. and Tuschl, T. (2001) Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature,* **411,** 494-8.
142. Hannon, G.J. and Rossi, J.J. (2004) Unlocking the potential of the human genome with RNA interference. *Nature,* **431,** 371-8.
143. Kim, D.H. and Rossi, J.J. (2007) Strategies for silencing human disease using RNA interference. *Nat Rev Genet,* **8,** 173-84.
144. Lingor, P. and Bahr, M. (2007) Targeting neurological disease with RNAi. *Mol Biosyst,* **3,** 773-80.
145. Rodriguez-Lebron, E. and Paulson, H.L. (2006) Allele-specific RNA interference for neurological disease. *Gene Ther,* **13,** 576-81.
146. Wood, M.J., Trulzsch, B., Abdelgany, A. and Beeson, D. (2003) Therapeutic gene silencing in the nervous system. *Hum Mol Genet,* **12 Spec No 2,** R279-84.
147. Saegusa, H., Wakamori, M., Matsuda, Y., Wang, J., Mori, Y., Zong, S. and Tanabe, T. (2007) Properties of human Cav2.1 channel with a spinocerebellar ataxia type 6 mutation expressed in Purkinje cells. *Mol Cell Neurosci,* **34,** 261-70.
148. Jun, K., Piedras-Renteria, E.S., Smith, S.M., Wheeler, D.B., Lee, S.B., Lee, T.G., Chin, H., Adams, M.E., Scheller, R.H., Tsien, R.W. *et al.* (1999) Ablation of P/Q-type Ca(2+) channel currents, altered synaptic transmission, and progressive ataxia in mice lacking the alpha(1A)-subunit. *Proc Natl Acad Sci U S A,* **96,** 15245-50.
149. Li, S.H. and Li, X.J. (2004) Huntingtin-protein interactions and the pathogenesis of Huntington's disease. *Trends Genet,* **20,** 146-54.
150. del Toro, D., Canals, J.M., Gines, S., Kojima, M., Egea, G. and Alberch, J. (2006) Mutant huntingtin impairs the post-Golgi trafficking of brain-derived neurotrophic factor but not its Val66Met polymorphism. *J Neurosci,* **26,** 12748-57.
151. Nasir, J., Floresco, S.B., O'Kusky, J.R., Diewert, V.M., Richman, J.M., Zeisler, J., Borowski, A., Marth, J.D., Phillips, A.G. and Hayden, M.R. (1995) Targeted disruption of the Huntington's disease gene results in embryonic lethality and behavioral and morphological changes in heterozygotes. *Cell,* **81,** 811-23.
152. Dragatsis, I., Levine, M.S. and Zeitlin, S. (2000) Inactivation of Hdh in the brain and testis results in progressive neurodegeneration and sterility in mice. *Nat Genet,* **26,** 300-6.
153. Schmitt, I., Linden, M., Khazneh, H., Evert, B.O., Breuer, P., Klockgether, T. and Wuellner, U. (2007) Inactivation of the mouse Atxn3 (ataxin-3) gene increases protein ubiquitination. *Biochem Biophys Res Commun,* **362,** 734-9.
154. Matilla, A., Roberson, E.D., Banfi, S., Morales, J., Armstrong, D.L., Burright, E.N., Orr, H.T., Sweatt, J.D., Zoghbi, H.Y. and Matzuk, M.M. (1998) Mice lacking ataxin-1 display learning deficits and decreased hippocampal paired-pulse facilitation. *J Neurosci,* **18,** 5508-16.

155. van Roon-Mom, W.M., Reid, S.J., Faull, R.L. and Snell, R.G. (2005) TATA-binding protein in neurodegenerative disease. *Neuroscience,* **133,** 863-72.
156. Ralser, M., Albrecht, M., Nonhoff, U., Lengauer, T., Lehrach, H. and Krobitsch, S. (2005) An integrative approach to gain insights into the cellular function of human ataxin-2. *J Mol Biol,* **346,** 203-14.
157. Kiehl, T.R., Nechiporuk, A., Figueroa, K.P., Keating, M.T., Huynh, D.P. and Pulst, S.M. (2006) Generation and characterization of Sca2 (ataxin-2) knockout mice. *Biochem Biophys Res Commun,* **339,** 17-24.
158. Thomas, P.S., Jr., Fraley, G.S., Damian, V., Woodke, L.B., Zapata, F., Sopher, B.L., Plymate, S.R. and La Spada, A.R. (2006) Loss of endogenous androgen receptor protein accelerates motor neuron degeneration and accentuates androgen insensitivity in a mouse model of X-linked spinal and bulbar muscular atrophy. *Human molecular genetics,* **15,** 2225-38.
159. Ambrose, C.M., Duyao, M.P., Barnes, G., Bates, G.P., Lin, C.S., Srinidhi, J., Baxendale, S., Hummerich, H., Lehrach, H., Altherr, M. *et al.* (1994) Structure and expression of the Huntington's disease gene: evidence against simple inactivation due to an expanded CAG repeat. *Somat Cell Mol Genet,* **20,** 27-38.
160. Fujigasaki, H., Uchihara, T., Koyano, S., Iwabuchi, K., Yagishita, S., Makifuchi, T., Nakamura, A., Ishida, K., Toru, S., Hirai, S. *et al.* (2000) Ataxin-3 is translocated into the nucleus for the formation of intranuclear inclusions in normal and Machado-Joseph disease brains. *Exp Neurol,* **165,** 248-56.
161. Uchihara, T., Fujigasaki, H., Koyano, S., Nakamura, A., Yagishita, S. and Iwabuchi, K. (2001) Non-expanded polyglutamine proteins in intranuclear inclusions of hereditary ataxias--triple-labeling immunofluorescence study. *Acta Neuropathol,* **102,** 149-52.
162. Lim, J., Crespo-Barreto, J., Jafar-Nejad, P., Bowman, A.B., Richman, R., Hill, D.E., Orr, H.T. and Zoghbi, H.Y. (2008) Opposing effects of polyglutamine expansion on native protein complexes contribute to SCA1. *Nature,* **452,** 713-8.
163. Kornberg, R.D. (2007) The molecular basis of eucaryotic transcription. *Cell Death Differ,* **14,** 1989-97.
164. Figueroa, K.P. and Pulst, S.M. (2003) Identification and expression of the gene for human ataxin-2-related protein on chromosome 16. *Exp Neurol,* **184,** 669-78.
165. Caplen, N.J., Taylor, J.P., Statham, V.S., Tanaka, F., Fire, A. and Morgan, R.A. (2002) Rescue of polyglutamine-mediated cytotoxicity by double-stranded RNA-mediated RNA interference. *Hum Mol Genet,* **11,** 175-84.
166. Harper, S.Q., Staber, P.D., He, X., Eliason, S.L., Martins, I.H., Mao, Q., Yang, L., Kotin, R.M., Paulson, H.L. and Davidson, B.L. (2005) RNA interference improves motor and neuropathological abnormalities in a Huntington's disease mouse model. *Proc Natl Acad Sci U S A,* **102,** 5820-5.
167. Xia, H., Mao, Q., Eliason, S.L., Harper, S.Q., Martins, I.H., Orr, H.T., Paulson, H.L., Yang, L., Kotin, R.M. and Davidson, B.L. (2004) RNAi suppresses polyglutamine-induced neurodegeneration in a model of spinocerebellar ataxia. *Nat Med,* **10,** 816-20.
168. Davidson, B.L. (2008) RNAi and Neurodegenerative Diseases *Keystone Symposia: RNAi, MicroRNA, and Non-Coding RNA* Whistler Resort, British Columbia, Canada.

169. Miller, V.M., Xia, H., Marrs, G.L., Gouvion, C.M., Lee, G., Davidson, B.L. and Paulson, H.L. (2003) Allele-specific silencing of dominant disease genes. *Proc Natl Acad Sci U S A,* **100,** 7195-200.
170. Feng, X., Zhao, P., He, Y. and Zuo, Z. (2006) Allele-specific silencing of Alzheimer's disease genes: the amyloid precursor protein genes with Swedish or London mutations. *Gene,* **371,** 68-74.
171. Gonzalez-Alegre, P., Bode, N., Davidson, B.L. and Paulson, H.L. (2005) Silencing primary dystonia: lentiviral-mediated RNA interference therapy for DYT1 dystonia. *J Neurosci,* **25,** 10502-9.
172. Gonzalez-Alegre, P., Miller, V.M., Davidson, B.L. and Paulson, H.L. (2003) Toward therapy for DYT1 dystonia: allele-specific silencing of mutant TorsinA. *Ann Neurol,* **53,** 781-7.
173. Li, Y., Yokota, T., Matsumura, R., Taira, K. and Mizusawa, H. (2004) Sequence-dependent and independent inhibition specific for mutant ataxin-3 by small interfering RNA. *Ann Neurol,* **56,** 124-9.
174. Xia, X., Zhou, H., Huang, Y. and Xu, Z. (2006) Allele-specific RNAi selectively silences mutant SOD1 and achieves significant therapeutic benefit in vivo. *Neurobiol Dis,* **23,** 578-86.
175. Kubodera, T., Yokota, T., Ishikawa, K. and Mizusawa, H. (2005) New RNAi strategy for selective suppression of a mutant allele in polyglutamine disease. *Oligonucleotides,* **15,** 298-302.
176. O'Reilly, M., Millington-Ward, S., Palfi, A., Chadderton, N., Cronin, T., McNally, N., Humphries, M.M., Humphries, P., Kenna, P.F. and Farrar, G.J. (2008) A transgenic mouse model for gene therapy of rhodopsin-linked Retinitis Pigmentosa. *Vision Res,* **48,** 386-91.
177. Boutla, A., Delidakis, C., Livadaras, I., Tsagris, M. and Tabler, M. (2001) Short 5'-phosphorylated double-stranded RNAs induce RNA interference in Drosophila. *Curr Biol,* **11,** 1776-80.
178. Elbashir, S.M., Martinez, J., Patkaniowska, A., Lendeckel, W. and Tuschl, T. (2001) Functional anatomy of siRNAs for mediating efficient RNAi in Drosophila melanogaster embryo lysate. *Embo J,* **20,** 6877-88.
179. Yu, J.Y., DeRuiter, S.L. and Turner, D.L. (2002) RNA interference by expression of short-interfering RNAs and hairpin RNAs in mammalian cells. *Proc Natl Acad Sci U S A,* **99,** 6047-52.
180. Holen, T., Amarzguioui, M., Wiiger, M.T., Babaie, E. and Prydz, H. (2002) Positional effects of short interfering RNAs targeting the human coagulation trigger Tissue Factor. *Nucleic Acids Res,* **30,** 1757-66.
181. Abdelgany, A., Wood, M. and Beeson, D. (2003) Allele-specific silencing of a pathogenic mutant acetylcholine receptor subunit by RNA interference. *Hum Mol Genet,* **12,** 2637-44.
182. Ding, H., Schwarz, D.S., Keene, A., Affar el, B., Fenton, L., Xia, X., Shi, Y., Zamore, P.D. and Xu, Z. (2003) Selective silencing by RNAi of a dominant allele that causes amyotrophic lateral sclerosis. *Aging Cell,* **2,** 209-17.
183. Schwarz, D.S., Ding, H., Kennington, L., Moore, J.T., Schelter, J., Burchard, J., Linsley, P.S., Aronin, N., Xu, Z. and Zamore, P.D. (2006) Designing siRNA that

distinguish between genes that differ by a single nucleotide. *PLoS Genet,* **2,** e140.

184. Jackson, A.L., Bartz, S.R., Schelter, J., Kobayashi, S.V., Burchard, J., Mao, M., Li, B., Cavet, G. and Linsley, P.S. (2003) Expression profiling reveals off-target gene regulation by RNAi. *Nat Biotechnol,* **21,** 635-7.
185. Dahlgren, C., Zhang, H.Y., Du, Q., Grahn, M., Norstedt, G., Wahlestedt, C. and Liang, Z. (2008) Analysis of siRNA specificity on targets with double-nucleotide mismatches. *Nucleic Acids Res,* **36,** e53.
186. Martinez, L.A., Naguibneva, I., Lehrmann, H., Vervisch, A., Tchenio, T., Lozano, G. and Harel-Bellan, A. (2002) Synthetic small inhibiting RNAs: efficient tools to inactivate oncogenic mutations and restore p53 pathways. *Proceedings of the National Academy of Sciences of the United States of America,* **99,** 14849-54.
187. Brummelkamp, T.R., Bernards, R. and Agami, R. (2002) Stable suppression of tumorigenicity by virus-mediated RNA interference. *Cancer Cell,* **2,** 243-7.
188. Miller, V.M., Gouvion, C.M., Davidson, B.L. and Paulson, H.L. (2004) Targeting Alzheimer's disease genes with RNA interference: an efficient strategy for silencing mutant alleles. *Nucleic Acids Res,* **32,** 661-8.
189. Du, Q., Thonberg, H., Wang, J., Wahlestedt, C. and Liang, Z. (2005) A systematic analysis of the silencing effects of an active siRNA at all single-nucleotide mismatched target sites. *Nucleic Acids Res,* **33,** 1671-7.
190. Dykxhoorn, D.M., Schlehuber, L.D., London, I.M. and Lieberman, J. (2006) Determinants of specific RNA interference-mediated silencing of human beta-globin alleles differing by a single nucleotide polymorphism. *Proc Natl Acad Sci U S A,* **103,** 5953-8.
191. Hickerson, R.P., Smith, F.J., Reeves, R.E., Contag, C.H., Leake, D., Leachman, S.A., Milstone, L.M., McLean, W.H. and Kaspar, R.L. (2008) Single-nucleotide-specific siRNA targeting in a dominant-negative skin model. *J Invest Dermatol,* **128,** 594-605.
192. Leung, R.K. and Whittaker, P.A. (2005) RNA interference: from gene silencing to gene-specific therapeutics. *Pharmacol Ther,* **107,** 222-39.
193. Bainbridge, J.W., Smith, A.J., Barker, S.S., Robbie, S., Henderson, R., Balaggan, K., Viswanathan, A., Holder, G.E., Stockman, A., Tyler, N. *et al.* (2008) Effect of gene therapy on visual function in Leber's congenital amaurosis. *N Engl J Med,* **358,** 2231-9.
194. Lee, N.S., Dohjima, T., Bauer, G., Li, H., Li, M.J., Ehsani, A., Salvaterra, P. and Rossi, J. (2002) Expression of small interfering RNAs targeted against HIV-1 rev transcripts in human cells. *Nat Biotechnol,* **20,** 500-5.
195. Baer, M., Nilsen, T.W., Costigan, C. and Altman, S. (1990) Structure and transcription of a human gene for H1 RNA, the RNA component of human RNase P. *Nucleic Acids Res,* **18,** 97-103.
196. Ely, A., Naidoo, T., Mufamadi, S., Crowther, C. and Arbuthnot, P. (2008) Expressed anti-HBV primary microRNA shuttles inhibit viral replication efficiently in vitro and in vivo. *Mol Ther,* **16,** 1105-12.
197. Castanotto, D., Li, H. and Rossi, J.J. (2002) Functional siRNA expression from transfected PCR products. *Rna,* **8,** 1454-60.

198. Campbell, R.E., Tour, O., Palmer, A.E., Steinbach, P.A., Baird, G.S., Zacharias, D.A. and Tsien, R.Y. (2002) A monomeric red fluorescent protein. *Proc Natl Acad Sci U S A,* **99,** 7877-82.
199. Graham, F.L., Smiley, J., Russell, W.C. and Nairn, R. (1977) Characteristics of a human cell line transformed by DNA from human adenovirus type 5. *J Gen Virol,* **36,** 59-74.
200. An, C.I., Trinh, V.B. and Yokobayashi, Y. (2006) Artificial control of gene expression in mammalian cells by modulating RNA interference through aptamer-small molecule interaction. *Rna,* **12,** 710-6.
201. Stenoien, D.L., Cummings, C.J., Adams, H.P., Mancini, M.G., Patel, K., DeMartino, G.N., Marcelli, M., Weigel, N.L. and Mancini, M.A. (1999) Polyglutamine-expanded androgen receptors form aggregates that sequester heat shock proteins, proteasome components and SRC-1, and are suppressed by the HDJ-2 chaperone. *Hum Mol Genet,* **8,** 731-41.
202. Janer, A., Martin, E., Muriel, M.P., Latouche, M., Fujigasaki, H., Ruberg, M., Brice, A., Trottier, Y. and Sittler, A. (2006) PML clastosomes prevent nuclear accumulation of mutant ataxin-7 and other polyglutamine proteins. *J Cell Biol,* **174,** 65-76.
203. Song, J.J., Smith, S.K., Hannon, G.J. and Joshua-Tor, L. (2004) Crystal structure of Argonaute and its implications for RISC slicer activity. *Science,* **305,** 1434-7.
204. Parker, J.S., Roe, S.M. and Barford, D. (2005) Structural insights into mRNA recognition from a PIWI domain-siRNA guide complex. *Nature,* **434,** 663-6.
205. Heale, B.S., Soifer, H.S., Bowers, C. and Rossi, J.J. (2005) siRNA target site secondary structure predictions using local stable substructures. *Nucleic Acids Res,* **33,** e30.
206. Kierzek, R., Burkard, M.E. and Turner, D.H. (1999) Thermodynamics of single mismatches in RNA duplexes. *Biochemistry,* **38,** 14214-23.
207. Li, F., Pallan, P.S., Maier, M.A., Rajeev, K.G., Mathieu, S.L., Kreutz, C., Fan, Y., Sanghvi, J., Micura, R., Rozners, E. *et al.* (2007) Crystal structure, stability and in vitro RNAi activity of oligoribonucleotides containing the ribo-difluorotoluyl nucleotide: insights into substrate requirements by the human RISC Ago2 enzyme. *Nucleic Acids Res,* **35,** 6424-38.
208. Ma, J.B., Ye, K. and Patel, D.J. (2004) Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. *Nature,* **429,** 318-22.
209. Miles, K., McAuliffe, L., Ayling, R.D. and Nicholas, R.A. (2004) Rapid detection of Mycoplasma dispar and M. bovirhinis using allele specific polymerase chain reaction protocols. *FEMS Microbiol Lett,* **241,** 103-7.
210. Ye, S., Dhillon, S., Ke, X., Collins, A.R. and Day, I.N. (2001) An efficient procedure for genotyping single nucleotide polymorphisms. *Nucleic Acids Res,* **29,** E88-8.
211. Morse, S.E. and Draper, D.E. (1995) Purine-purine mismatches in RNA helices: evidence for protonated G.A pairs and next-nearest neighbor effects. *Nucleic Acids Res,* **23,** 302-6.
212. Tolbert, B.S., Kennedy, S.D., Schroeder, S.J., Krugh, T.R. and Turner, D.H. (2007) NMR structures of (rGCUGAGGCU)2 and (rGCGGAUGCU)2: probing

the structural features that shape the thermodynamic stability of GA pairs. *Biochemistry,* **46,** 1511-22.

213. Shao, Y., Chan, C.Y., Maliyekkel, A., Lawrence, C.E., Roninson, I.B. and Ding, Y. (2007) Effect of target secondary structure on RNAi efficiency. *Rna,* **13,** 1631-40.
214. McBride, J.L., Boudreau, R.L., Harper, S.Q., Staber, P.D., Monteys, A.M., Martins, I., Gilmore, B.L., Burstein, H., Peluso, R.W., Polisky, B. *et al.* (2008) Artificial miRNAs mitigate shRNA-mediated toxicity in the brain: implications for the therapeutic development of RNAi. *Proceedings of the National Academy of Sciences of the United States of America,* **105,** 5868-73.
215. Boudreau, R.L., Monteys, A.M. and Davidson, B.L. (2008) Minimizing variables among hairpin-based RNAi vectors reveals the potency of shRNAs. *RNA (New York, N.Y,* **14,** 1834-44.
216. Arrasate, M., Mitra, S., Schweitzer, E.S., Segal, M.R. and Finkbeiner, S. (2004) Inclusion body formation reduces levels of mutant huntingtin and the risk of neuronal death. *Nature,* **431,** 805-10.
217. Park, I.H., Zhao, R., West, J.A., Yabuuchi, A., Huo, H., Ince, T.A., Lerou, P.H., Lensch, M.W. and Daley, G.Q. (2008) Reprogramming of human somatic cells to pluripotency with defined factors. *Nature,* **451,** 141-6.
218. Dimos, J.T., Rodolfa, K.T., Niakan, K.K., Weisenthal, L.M., Mitsumoto, H., Chung, W., Croft, G.F., Saphier, G., Leibel, R., Goland, R. *et al.* (2008) Induced pluripotent stem cells generated from patients with ALS can be differentiated into motor neurons. *Science (New York, N.Y,* **321,** 1218-21.

Appendix

A1. Basic PCR Cocktail and cycling conditions

PCR was performed in 0.2ml tubes (*Eppendorf,* Merck) in a total volume of 25 μl containing 200 ng of DNA; 10 pmol of each primer; 0.2mM of each of the 4 dNTPs (*Bioline,* Celtic Molecular Diagnostics, SA), 1 x PCR Buffer (*Promega*, Whitehead Scientific, SA); 0.5 Unit of *Taq* polymerase (*Promega* GoTaq, Whitehead Scientific, SA), made up with H_2O. PCR was carried out on a Perkin Elmer DNA thermal cycler as follows:

- 1 cycle of 95^oC – 5 min
- 30 cycles of 95^oC – 30 sec
  50^oC – 30 sec
  72^oC – 40 sec
- 1 cycle of 72^oC – 7 min

A2. Table of primers used for PCR

Primer	Sequence (5' – 3')
G>A SNP_F	TCCAATGAACTGCCTGTCAA
G>A SNP_R	GCTGATGAAGGAGGGAACTG
GAPDH_F	AAGGTGAAGGTCGGAGTCAA
GAPDH_R	GAAGATGGTGATGGGATTTC

A3. Long reverse oligonucleotides required for cloning shRNAs

Primer	Sequence 5' – 3'
U6+1	GATCGTCGACAAGGTCGGGCAGGAAGAGGGCCT
shR-NS	Kind gift of C Sibley.
shR-P10	AAAAAAGTGCCAGCCATGAACAATGTGGGTCAGGCACTGCCCATGGCTGGCACGGTGTTTCGTCCTTTCCACAA
shR-P11	AAAAAATGCCAGCCATGAACAATGTTGGGTCAGGACATTATTCATGGCTGGCCGGTGTTTCGTCCTTTCCACAA
shR-P12	AAAAAAGCCAGCCATGAACAATGTCTGGGTCAGGGACATTATTCATGGCTGGCGGTGTTTCGTCCTTTCCACAA
shR-P13	AAAAAACCAGCCATGAACAATGTCCTGGGTCAGGGAACATTATTCATGGCTGCGGTGTTTCGTCCTTTCCACAA
shR-P14	AAAAAACAGCCATGAACAATGTCCATGGGTCAGGTAGACATTGTTCATGGCTCGGTGTTTCGTCCTTTCCACAA
shR-P15	AAAAAAAGCCATGAACAATGTCCACTGGGTCAGGATAGACATTGTTCATGGCCGGTGTTTCGTCCTTTCCACAA
shR-P16	AAAAAAGCCATGAACAATGTCCACATGGGTCAGGTATGAACATTGTTCATGGCGGTGTTTCGTCCTTTCCACAA
shR-P10P11	AAAAAAGTGCCAGCTATGAACAATGTGGGTCAGGCACTGCCCATAGCTGGCACGGTGTTTCGTCCTTTCCACAA
shR-P11P12	AAAAAATGCCAGCTATGAACAATGATGGGTCAGGTCATTATTCATAGCTGGCCGGTGTTTCGTCCTTTCCACAA
shR-P12P13	AAAAAAGCCAGCTATGAACAATGTCTGGGTCAGGGACATTATTCATAGCTGGCGGTGTTTCGTCCTTTCCACAA
shR-P13P14	AAAAAACCAGCTATGAACAATGTCCTGGGTCAGGGAACATTATTCATAGCTGCGGTGTTTCGTCCTTTCCACAA
shR-P14P15	AAAAAACAGCTATGAACAATGTCCATGGGTCAGGTAGACATTGTTCATAGCTCGGTGTTTCGTCCTTTCCACAA
shR-P15P16	AAAAAAAGCTATGAACAATGTCCACTGGGTCAGGATAGACATTGTTCATAGCCGGTGTTTCGTCCTTTCCACAA
shR-P16P17	AAAAAAGCTATGAACAATGTCCACATGGGTCAGGTATGAACATTGTTCATAGCGGTGTTTCGTCCTTTCCACAA
shR-P16P15	AAAAAAGCCAGGAACAATGTCCACATGGGTCAGGTATGAACATTGTTCCTGGCGGTGTTTCGTCCTTTCCACAA
shR-P16P15P17	AAAAAAGCTAGGAACAATGTCCACATGGGTCAGGTATGAACATTGTTCCTAGCGGTGTTTCGTCCTTTCCACAA
shR-P15P16'	AAAAAAAGCAATGAACAATGTCCACTGGGTCAGGATAGACATTGTTCATTGCCGGTGTTTCGTCCTTTCCACAA
shR-P15A:G	AAAAAAAGCCTTGAACAATGTCCACTGGGTCAGGATAGACATTGTTCAAGGCCGGTGTTTCGTCCTTTCCACAA
shR-P15G:G	AAAAAAAGCCCTGAACAATGTCCACTGGGTCAGGATAGACATTGTTCAGGGCCGGTGTTTCGTCCTTTCCACAA

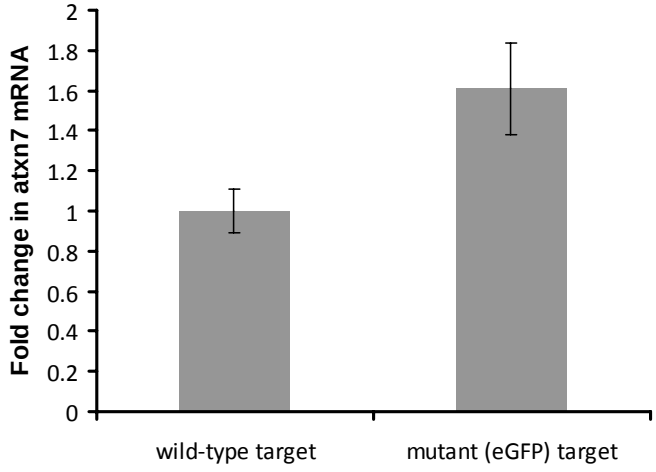
A4. Long reverse oligonucleotides required for cloning miRNAs

Primer	Sequence 5' – 3'
miR-122 universal F	GATACACTCGAGTGGAGGTGAAGTTAACACCTTCGTGGCTACAGAGTTTCCTTAGCAGAGCTG
miR-122 universal R	TGTATCGCGGCCGCAAGCAAACGATGCCAAGACATTTATCGAGGGAAGGATTGCCTAGCAGTAGCTA
miR-P14 F	GTTTCCTTAGCAGAGCTGTGGACATTGTTCATGGCTGGCATGTCTAAACTATTG
miR-P14 R	CTAACGGATCGTCATCGATAGATGTAACACGTACCGACCGTTATCAAATCTGTAC
miR-P15 F	GTTTCCTTAGCAGAGCTGGTGGACATTGTTCATGGCTGGCTGTCTAAACTATGC
miR-P15 R	CTAACGGATCGTCATCGATCCGCTGTAACCAGTACCGACCGTATCAAATCTGTCG
miR-P16 F	GTTTCCTTAGCAGAGCTGTGTGGACATTGTTCATGGCTGGTGTCTAAACTATCC
miR-P16 R	CTAACGGATCGTCATCGATAACCCTGTAAGAAGTACCGACCTATCAAATCTGTGG
miR-P17 F	GTTTCCTTAGCAGAGCTGATGTGGACATTGTTCATGGCTGTGTCTAAACTATCA
miR-P17 R	CTAACGGATCGTCATCGATTCGACCTGTACCAAGTACCGACTATCAAATCTGTGT
miR-P18 F	GTTTCCTTAGCAGAGCTGCATGTGGACATTGTTCATGGCTTGTCTAAACTATAG
miR-P18 R	CTAACGGATCGTCATCGATGGCCACCTGTCACAAGTACCGATATCAAATCTGTTC

A5. Oligonucleotides required for cloning target plasmids

Primer name	Primer sequence 5' -3'	Purpose
Atx7-G-luc_F	TCGAGATATCACAAAGGTTGCCAAAGTGCCAGCCGTGAACAATGTCCACA TGAAACACACGC	Oligonucleotide sequences used for cloning the wild-type luciferase target vector
Atx7-G-luc_R	GGCCGCGTGTGTTTCATGTGGACATTGTTCACGGCTGGCACTTTGGCAAC CTTTGTGATATC	
Atx7-A-luc_F	TCGAGATATCACAAAGGTTGCCAAAGTGCCAGCCATGAACAATGTCCACA TGAAACACACGC	Oligonucleotide sequences used for cloning the mutant luciferase target vector
Atx7-A-luc_R	GGCCGCGTGTGTTTCATGTGGACATTGTTCATGGCTGGCACTTTGGCAAC CTTTGTGATATC	
Atx-A-s	CTCGAGCAGCATCAACAACAGCAGCAGCAAACCCACAAAGGTTGCCAAAG TGCCAGCCATGAACAATGTCCACATGAAACACACAGGCAC	Oligonucleotide sequences required to replace the G allele of the full-length target
Atx-A-as	AAGCTTGGGACGTGCCTTTGGCTGATGAAGGAGGGAACTGTTCATCAGTC CTTGTGCCCCTGGGATGGTGCCTGTGTGTTTCATGTG	
Atx-A-s-F-p	GATACAACTCGAGCAGCATC	
Atx-A-as-R-p	TTGTATAAAGCTTGGGACGT	

A6. Levels of atxn7 in transfected cells



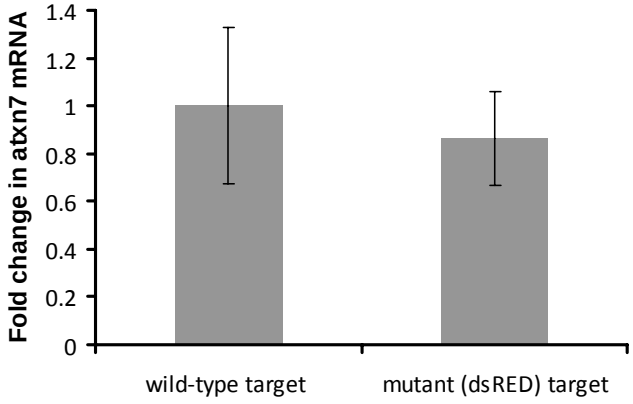


Figure A6: Levels of mutant atxn7 mRNA relative to wild-type. The top graph indicates an increased fold change of mutant target expression relative to the wild-type when the target is tagged to the eGFP. The bottom graph indicates that the level of mutant target when tagged to dsRED is slightly less relative to the levels of wild-type target. These experiments were performed in duplicate. Standard deviations are indicated.

“There is no medicine like hope. No incentive so great and no tonic so powerful as expectation of something new tomorrow. Janine, in your research, expect something new every day”

- Mr Joe Andrews, Chairman of the South African Spinocerebellar Ataxia Association.

“It is not enough that we do our best;
sometimes we have to do what is required.”

- Winston Churchill.