Developing Gene Knockdown-Replacement Therapies for Spinocerebellar Ataxia type 7

Helen Curtis

Wolfson College



A THESIS PRESENTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY IN THE DIVISION OF MEDICAL SCIENCES

DEPARTMENT OF PHYSIOLOGY, ANATOMY AND GENETICS

UNIVERSITY OF OXFORD

HILARY 2013

Acknowledgements

This work was carried out thanks to funding awarded by the John Fell OUP Research Fund and my involvement would not have been possible without the Engineering and Physical Sciences Research Council scholarship through the Oxford University Systems Biology Doctoral Training Centre.

I extend my gratitude to Professor Matthew Wood for the opportunity to pursue this project and for the invaluable advice and support along the way. I found wonderful supervisors in Yiqi Seow, Chris Sibley, Janine Scholefield and Miguel Varela, whose efforts and guidance I couldn't have done without. I also thank others who facilitated this work, especially Lars Aagaard for the provision of plasmid, Anna-Lena Ström for the very kind and incredibly useful gift of antibodies, and my excellent proof-readers!

Everyone in the lab made it a friendly, fun and collaborative atmosphere to work in. Naomi and Lauren, it was a pleasure working on SCA7 with you. To Fiona, thank you for helping me many, many times!

In Wolfson I found an amazing home and many great people to share the experience with. Especially Emily, Alex and Andres, I was so lucky to end up living with people who became such good friends. From DTC08, Mireille, Ivor, Jen and everyone else, thank you for all the adventures! To Gwilym - I don't know how to thank you enough for all the ways you helped me. To Mum, Dad and Andrew - here's what I've been doing for the past few years!

Abstract

For many dominant diseases, conventional treatment options are limited. This makes them attractive candidates for gene therapy, which may be directed to specifically silence a disease-causing allele. However, many mutations are not easily amenable to this technique, including nucleotide repeat expansions, which cause numerous neurodegenerative diseases such as Huntington's disease and several Spinocerebellar Ataxias. Combined gene knockdown and replacement (K&R) may present a more practical approach for such conditions, whereby the gene of interest is subject to mutation-independent non-allele-specific silencing and concurrently replaced with a functional copy. Artificial mimics of naturally occurring intronic microRNAs are theoretically ideal for this purpose, since they can be nested within the replacement gene. This thesis investigates the development of mimics of two different intronic miRNA systems (mirtron miR-1224 and the miR-106b cluster) to silence Ataxin-7 and to be incorporated into novel single K&R constructs for testing *in vitro*. Artificial mirtrons and intronic miRNAs were successfully developed and shown to silence Ataxin-7 mRNA in numerous cell lines. An RNAi-resistant gene was developed and mirtrons could be successfully incorporated as introns. Patient-derived fibroblasts and iPSC-derived neuronal cells were investigated as models for testing of gene silencing therapies. This work suggests that mirtron-based K&R is achievable, and warrants further investigation.

Table of Contents

Abstract 3

Abbreviations 11

1 Introduction 14

1.1 Overview 14

1.2 Gene therapy 14

1.2.1 Gene therapy for dominant diseases 15

1.3 RNAi and mirtrons 20

1.3.1 Introduction to RNAi 20

1.3.2 Micro-RNA pathways 20

1.3.3 Synthetic RNAi in gene therapy 31

1.4 Gene therapy for polyQ disorders 35

1.4.1 Non-allele-specific silencing 35

1.4.2 Allele-specific silencing for polyQ disorders 36

1.4.3 Knockdown-replacement 38

1.5 Gene therapy for SCA7 45

1.5.1 SCA7 as a model 45

1.5.2 Genetics of SCA7 46

1.5.3 Pathology of SCA7 47

1.5.4 Mouse models for SCA7 48

1.5.5 Nuclear inclusions in SCA7 52

1.5.6 Loss of Ataxin-7 function in SCA7 58

1.5.7 The importance of non-cell-autonomous effects in SCA7 66

1.5.8 Summary 67

1.6 Potential strategies for treatment of SCA7 68

1.6.1 Regulators of Ataxin-7 protein turnover 68

1.6.2 Mediators of downstream effects 69

1.6.3 Stem cell therapy 70

1.6.4 Silencing of mutant Ataxin-7 70

1.7 Overview of project 72

2 Materials and Methods 73

2.1 Mirtron design 73

2.2 Cloning 74
2.2.1 Mirtron expression plasmids 75
2.2.2 Multiple-mirtron expression plasmids 76
2.2.3 Short hairpin RNAs, short interfering RNAs and LNAs 78
2.2.4 Dual Luciferase targets 79
2.2.5 miRNA cluster mimics 80
2.2.6 Full-length Ataxin-7 targets 80
2.2.7 Modification of Ataxin-7 and addition of introns 81
2.2.8 Lentivirus cloning 90
2.3 Cell culture and transduction 90
2.3.1 Cells and culture conditions 90
2.3.2 Patient-derived neuronal cultures 91
2.3.3 Transfection of established cell lines 92
2.3.4 Transfection of patient fibroblasts 94
2.3.5 Transfection for virus production 94
2.3.6 Viral infection 95
2.4 Protein assays 96
2.4.1 Fluorescence microscopy 96
2.4.2 EGFP and mCherry quantification 96
2.4.3 Dual Luciferase assay 96
2.4.4 Western blotting 97
2.4.5 Cell immunostaining 98
2.5 Nucleic acid assays 99
2.5.1 RNA extraction 99
2.5.2 RT-PCR 99
2.5.3 Real-time quantitative RT-PCR 100
2.5.4 Deep sequencing 102
2.6 Statistics 103
3 Silencing Ataxin-7 with intronic miRNAs 104
3.1 Introduction 104
3.2 Screening revealed a number of promising synthetic mirtrons 106
3.3 Mirtron processing 116
3.3.1 Mirtron activity is dependent upon splicing 116

3.3.2 Synthetic mirtrons are processed to mature miRNAs with biased strand selection 121
3.3.3 Investigating the potential off-target effects of mirtrons 125
3.3.4 Reducing hairpin base-pairing in mirt-18 improves splicing but compromises silencing ability 127
3.4 Multiple mirtrons may be delivered from a single transcript 131
3.5 Artificial mirtrons can silence endogenous Ataxin-7 137
3.5.1 Silencing in HEK-293 cells 137
3.5.2 Silencing in Neuro2A cells 139
3.5.3 Silencing in SH-SY5Y cells 143
3.6 Other intronic miRNAs are an alternative option for silencing Ataxin-7 147
3.6.1 Design 147
3.6.2 Individual Ataxin-7 miRNAs within a miR-106b cluster mimic effectively silenced Luciferase targets 148
3.6.3 RT-PCR confirms that the MCM7 intron was spliced correctly 148
3.7 Discussion 152
3.7.1 Artificial mirtrons silence Ataxin-7 efficiently 152
3.7.2 Artificial mirtrons are processed into mature miRNAs 153
3.7.3 Optimisation of mirtron design 154
3.7.4 Other miRNA mimics may also be used to target Ataxin-7 156
3.7.5 Conclusions 158
4 Gene knockdown-replacement systems for SCA7 159
4.1 Introduction 159
4.2 Development of RNAi-resistant Ataxin-7 162
4.3 Knockdown-replacement proof-of-principle 167
4.4 Delivering mirtrons as introns within Ataxin-7 169
4.4.1 Initial testing of splice sites in Ataxin-7 170
4.4.2 Testing of further natural and potential intron sites 172
4.4.3 Computational prediction of splice site recognition 175
4.4.4 Investigation of high-scoring splice sites 178
4.4.5 Concurrent silencing and replacement with a mirtron delivered within Ataxin-7 cDNA 183
4.4.6 Alternative strategies for combinatorial delivery 187
4.5 Discussion 192

4.5.1 Silent modifications in Ataxin-7 convey resistance to silencing by mirtrons ... 192
4.5.2 Mirtrons and a replacement gene can be delivered under a single promoter 193
4.5.3 Mirtrons can be effectively spliced from some sites in Ataxin-7 ... 194
4.5.4 Context dependence of mirtron splicing ... 194
4.5.5 Delivering mirtrons within UTRs ... 196
4.5.6 Tailed-mirtrons ... 197
4.5.7 Conclusions ... 198
5 Patient-derived cell models for assessment of novel SCA7 therapies ... 199
5.1 Introduction ... 199
5.2 Patient fibroblasts ... 201
5.2.1 Characterisation of patient fibroblast cells ... 201
5.2.2 Silencing Ataxin-7 in patient fibroblasts ... 207
5.2.3 Transcriptional rescue of genes dysregulated in SCA7 ... 217
5.3 Lentiviruses as a vector for mirtrons ... 219
5.3.1 Production of lentiviruses to silence Ataxin-7 ... 219
5.3.2 Estimation of lentiviral titre ... 222
5.3.3 Expression of shRNA and mirtrons from lentiviruses ... 224
5.3.4 Silencing in a standard neuronal cell line ... 226
5.4 Patient-derived neurons ... 231
5.4.1 Neuron production and characterisation. ... 231
5.4.2 Silencing of Ataxin-7 in neurons ... 233
5.5 Discussion ... 235
5.5.1 Patient fibroblasts as a model ... 235
5.5.2 Lentiviruses ... 236
5.5.3 Patient-derived neurons as a model ... 237
5.5.4 Conclusions ... 239
6 Discussion ... 240
6.1 Overview ... 240
6.2 Mirtron-based RNAi ... 241
6.2.1 Artificial mirtrons can silence Ataxin-7 ... 242
6.2.2 Mirtron biogenesis ... 243
6.2.3 Artificial mirtron design ... 244
6.3 Other intronic miRNAs ... 245

6.4 Mirtron-based K&R 246
6.4.1 Advantages of mirtron-based K&R therapy 246
6.4.2 Silent modifications in Ataxin-7 convey resistance to silencing by mirtrons ... 247
6.4.3 Mirtrons and a replacement gene can be co-delivered under a single promoter 247
6.4.4 Mirtrons can be effectively spliced from some sites in Ataxin-7 247
6.4.5 Further development of mirtron-based K&R 248
6.5 Patient-derived cells as models for SCA7 249
6.5.1 Patient fibroblasts as a model 250
6.5.2 Patient-derived neurons as a model 251
6.5.3 SCA7 mouse models 253
6.6 Viral delivery of mirtrons 254
6.7 Further research 255
6.8 Concluding remarks 256
7 References 258

Table of Figures

Figure 1-1 - Gene therapy strategies for dominant toxic gain-of-function conditions. 19
Figure 1-2 – The canonical miRNA, mirtron and tailed-mirtron pathways. 22
Figure 1-3 - Atypical biogenesis pathways for miRNA generation. 25
Figure 1-4 - Examples of canonical miRNAs, mirtrons and tailed mirtrons from *Drosophila melanogaster*. 27
Figure 1-5 - Potential strategies for gene knockdown-replacement (K&R) therapy. 43
Figure 2-1 - Mirtron design. 74
Figure 2-2 - pEGFP-mirt. 76
Figure 2-3 - Cloning multiple mirtrons into pEGFP. 77
Figure 2-4 - pMCM7-S1-S2-S3. 80
Figure 2-5 - Ataxin-7 cDNA was modified by codon replacement to be resistant to silencing by mirtrons. 82
Figure 2-6 – Method used for creation of RNAi-resistant Ataxin-7. 83
Figure 2-7 - Ataxin-7 cDNA was divided into segments for commercial production. 88
Figure 3-1 - Sequence and structure of synthetic mirtrons designed to target Ataxin-7. 109
Figure 3-2 – Position of mirtron target sites within Ataxin-7 cDNA. 109
Figure 3-3 - Mirtron screening results. 112
Figure 3-4 - Silencing of mCherry-tagged Ataxin-7 constructs by mirtrons and shRNAs along with representative images. 116
Figure 3-5 - RT-PCR demonstrating splicing efficiency of EGFP-mirtron constructs. 118
Figure 3-6 - Splicing and silencing activity of un-spliceable equivalents of mirtrons in HEK-293 cells. 120
Figure 3-7 – Deep sequencing reads obtained from HEK-293 cells transfected with mirt-1cN, mirt-18 or mirt-18-US. 122
Figure 3-8 – Potential off-target effects of the guide and passenger strands of artificial mirtrons. 126
Figure 3-9 - Design and characterisation of variants of mirt-18. 131
Figure 3-10 - Mirtron production from multiple-mirtron constructs 135
Figure 3-11 – Silencing ability of multiple-mirtron constructs against fluorescent or Luciferase targets in HEK-293 cells. 137
Figure 3-12 - Silencing of endogenous Ataxin-7 mRNA in HEK-293 cells. 138
Figure 3-13 - Silencing of Ataxin-7 mRNA and protein in N2A cells after siRNA treatment. 141
Figure 3-14 - Silencing of endogenous Ataxin-7 mRNA in N2A cells with mirtrons. 143
Figure 3-15 - Silencing of Ataxin-7 mRNA and protein in SH-SY5Y cells after siRNA treatment. 145
Figure 3-16 - Silencing of Ataxin-7 mRNA in SH-SY5Y cells after transfection with shRNA and mirtrons. 146
Figure 3-17 - Design and Testing of miRNA-mimics against Ataxin-7. 151
Figure 4-1 - Modified Ataxin-7 constructs were designed to be resistant to silencing by mirtrons. 165
Figure 4-2 – Further modifications were made to the mirt-18 target site. 166
Figure 4-3 – Knockdown and replacement of Ataxin-7 achieved by delivering a mirtron within the EGFP tag. 169
Figure 4-4 – Mirtrons were cloned into two natural intron locations within Ataxin-7 and tested for splicing. 172
Figure 4-5 - Introns were cloned into natural and predicted splice sites within Ataxin-7. 174

Figure 4-6 - Mirtrons were cloned into the most high-scoring natural splice sites within Ataxin-7.....180
Figure 4-7 – Local RNA structure predictions for mirt-1cN within sites 8 and 9 of Ataxin-7. 182
Figure 4-8 – Concurrent silencing and replacement of Ataxin-7 in HEK-293 cells with a single construct.....185
Figure 4-9 - Deep sequencing reads obtained from HEK-293 cells transfected with pAtaxin-7-1cN-9.....186
Figure 4-10 - Masking a cryptic splice site in Ataxin-7. 188
Figure 4-11 –The NAD intron was cloned into the 5' UTR of Ataxin-7.....191
Figure 5-1 – Characterisation of patient fibroblast cell lines. 202
Figure 5-2 – Staining of HEK-293 cells expressing Ataxin-7-Q100-EGFP.....204
Figure 5-3 – Staining of SCA7 patient versus normal control with Ataxin-7 antibodies. 206
Figure 5-4 – Silencing of Ataxin-7 mRNA in patient fibroblasts. 209
Figure 5-5 – Immunostaining of SCA7 patient fibroblasts after siRNA or mirtron transfection. 211
Figure 5-6 – Ataxin-7 protein detection in SCA7 patient fibroblasts by Western blotting. ...215
Figure 5-7 – Effects of Ataxin-7 silencing on gene expression. 218
Figure 5-8 – shRNA and mirtron lentivirus design and production.....221
Figure 5-9 – Lentiviral infections in HEK-293 cells.....223
Figure 5-10 – Mirtron expression in HEK-293 cells infected with lentivirus compared to plasmid-transfected cells.....225
Figure 5-11 – Immunostaining of Ataxin-7 in SH-SY5Y cells infected with lentiviruses. 228
Figure 5-12 – Lentiviral infection was improved with polybrene. 229
Figure 5-13 – Silencing of endogenous Ataxin-7 mRNA with lentiviruses in SH-SY5Y cells. ..230
Figure 5-14 – Ataxin-7 immunostaining in patient-derived cultures infected with lentiviruses. 234

Tables

Table 1-1 – The polyQ diseases. 17
Table 1-2 – Summary of the limitations of RNAi-based silencing therapies for polyQ diseases. 38
Table 1-3 – Studies involving concurrent gene silencing and replacement.....41
Table 1-4 – Major mouse models for SCA7 which affect the cerebellum.....49
Table 1-5 – Major SCA7 mouse models affecting the retina.....50
Table 2-1 – Sequences of siRNAs and shRNAs targeting Ataxin-7. 79
Table 2-2 – Primers used for cloning of mirtrons, shRNAs, and miRNA cluster mimics. 85
Table 2-3 – Primers used for cloning of Luciferase targets and Ataxin-7 87
Table 2-4 - Segments of Ataxin-7 produced commercially. 89
Table 2-5 – Combinations of commercially-produced segments of Ataxin-7 (A, B, C) and backbone plasmids used to make Ataxin-7 plasmids with modifications and/or introns. 90
Table 2-6 – Transfection conditions according to cell line and experimental application. 93
Table 2-7 – Primer combinations used for RT-PCR and qPCR. 102
Table 5-1 – Lentiviral titre estimation. 223

Abbreviations

A	Adenosine
AAV	Adeno-associated virus
AD	Alzheimer's disease
AGO	Argonaute
ALS	Amytrophic lateral sclerosis
ATN1	Atrophin 1
BDNF	Brain-derived neurotrophic factor
bFGF	Basic fibroblast growth factor
bp	Base pairs
C	Cytosine
C. elegans	*Caenorhabditis elegans*
CACNA1A	Calcium channel, voltage-dependent, P/Q type, alpha 1A subunit
CBP	CREB-binding protein
cDNA	Complementary DNA
CMV	Cytomegalovirus
CNS	Central Nervous System
CNTF	ciliary neurotrophic factor
CREB	cAMP response element-binding protein
CRX	Cone-rod homeobox
CTCF	CCCTC-binding factor
D.melanogaster	*Drosophila melanogaster*
DBR1	Lariat debranching enzyme
DGCR8	DiGeorge syndrome critical region 8
DMEM	Dulbecco's Modified Eagle Medium
DNA	Deoxyribonucleic acid
DNAJA1	DnaJ homolog subfamily A member 1
DRPLA	Dentatorubral pallidoluysian atrophy
dsRNA	Double-stranded RNA
EDTA	Ethylenediaminetetraacetic acid
EGFP	Enhanced green fluorescent protein
endo-siRNAs	Endogenous-siRNAs
ERG	Electroretinogram
FCS	Foetal calf serum
G	Guanine
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GDNF	Glial-derived neurotrophic factor
h	Hour(s)
h	Human
HAT	Histone acetyltransferase
HD	Huntington's disease
HDAC	Histone deacetylase
HEK	Human embryonic kidney
HIV	Human immunodeficiency virus
HSP	Heat shock protein

IGFBP5	Insulin-like growth factor-binding protein 5
iMEF	Inactivated mouse embryonic fibroblast
iPS(C)	Induced pluripotent stem (cell)
K&R	Knockdown and replacement
KI	Knock-in
KO	Knockout
LNA	Locked nucleic acids
miRNA	MicroRNA
mRNA	Messenger RNA
MT	Microtubule
N2A	Neuro-2A
NBM-A	Neurobasal Medium A
NI	Nuclear inclusion
NRL	Neural retina leucine zipper
NS	Non-specific
nt	Nucleotide
N-terminal	Amino terminal
PBS	Phosphate buffered saline
PBST	PBS supplemented with 0.1% Triton X-100
PC	Purkinje cell
PCAF	P300/CBP-associated factor
Pcp-2	Purkinje cell protein 2
PCR	Polymerase chain reaction
PD	Parkinson's disease
PDGF	Platelet-derived growth factor
PML	Promyelocytic leukaemia protein
Pol-II	Polymerase-II
Pol-III	Polymerase-III
polyQ	poly-glutamine
Pre-miRNA	Precursor microRNA
Pri-miRNA	Primary microRNA
PrP	Prion protein
PTGS	Post-transcriptional gene silencing
Q	Glutamine
qPCR	Real-time quantitative reverse-transcription PCR
RISC	RNA-induced silencing complex
RNA	Ribonucleic acid
RNAi	RNA interference
RNAse	Ribonuclease
ROS	Reactive oxygen species
rpm	Revolutions per minute
RT-PCR	Reverse transcription polymerase chain reaction
SAGA	Spt/Ada/Gcn5/acetylase
SBMA	Spinal bulbar muscular atrophy
SCA	Spinocerebellar Ataxia
SCA7	Spinocerebellar Ataxia type 7
SCAANT1	SCA7/ataxin-7 antisense RNA 1
SDS	Sodium dodecyl sulphate

shRNA	Short hairpin RNA
siRNA	Small interfering RNA
SNP	Single nucleotide polymorphism
snRNA	Small nuclear RNA
SOD1	Superoxide dismutase 1
STAGA	SPT3/TAF9/ADA/GCN5 acetyltransferase complex
Taq	Thermus aquaticus
TBP	TATA binding protein
TBS	Tris buffered saline
TBST	TBS supplemented with 0.1% Triton X-100
TFTC	TATA-binding protein (TBP)-free TAF-containing complex
Tg	Transgenic
tRNA	Transfer RNA
UBC	Ubiquitin C
UCHL1	Ubiquitin carboxy-terminal hydrolase L1
UPS	Ubiquitin-proteasome system
US	Un-spliceable
UTR	Untranslated region
VA1	Adenoviral associated RNA 1
VEGF	Vascular endothelial growth factor
VPA	Valproic acid
WT	Wild-type

1 Introduction

1.1 Overview

For many dominant diseases, in which a mutation causes a toxic gain-of-function or dominant-negative effect, conventional treatment options are limited, making them attractive candidates for gene therapy. Gene therapy aims to correct the expression of pathogenic genes and may be directed to specifically silence a disease-causing allele. However, this approach faces substantial difficulties for the treatment of many dominant diseases. Two such groups are those which can be caused by a large number of possible mutations, where the development of therapies to treat each mutation poses a significant challenge, and disorders caused by an expanded nucleotide repeat, where the mutant allele is not easily amenable to allele-specific silencing. Such conditions are candidates for combined gene knockdown and replacement therapies, whereby the gene of interest is subject to non-allele-specific silencing and concurrently replaced. This project will investigate the use of gene knockdown-replacement therapies applied to spinocerebellar ataxia type 7, a nucleotide repeat disorder.

1.2 Gene therapy

The field of gene therapy is being explored for the treatment of a wide range of conditions, many of which currently have no effective treatments. The development of viral vectors was a breakthrough allowing the delivery of recombinant nucleic acid molecules designed to alter gene expression. Early gene therapy clinical trials for X-linked severe combined immunodeficiency (SCID-X1) were effective in most patients, but provoked safety concerns concerning insertional mutagenesis and fatal inflammatory responses (Hacein-Bey-Abina et

al. 2003; Raper et al. 2003). However, both viral and non-viral vectors have since been developed and improved and many gene therapy clinical trials have been undertaken – numbering over 1800 as of June 2012 (Ginn et al. 2013). These include therapies for cancers, cardiovascular disease, infectious diseases, cystic fibrosis, haemophilia and muscular dystrophy. Efficient delivery is a limiting factor, therefore the most well-developed therapies tend to address conditions affecting the most accessible cells or tissues, including tumours, the haematopoietic compartment and retina (Ginn et al. 2013; Hauswirth et al. 2008; Jacobson et al. 2012). However, gene therapies are also in development for several conditions of the central nervous system (CNS), including Alzheimer's disease (AD) and Parkinson's disease (PD) (Naidoo and Young 2012). For example, gene delivery via intracranial injection has been attempted for several different genes to treat Parkinson's disease, demonstrating safety and some efficacy in some cases (Christine et al. 2009). As in other applications, there were also some adverse effects, particularly regarding the mode of delivery via intracranial injection (Marks et al. 2010).

Clearly, the complex field of gene therapy has challenges to overcome. However, it continues to advance in terms of safety and measurable phenotypic improvements, providing motivation for its further development.

1.2.1 Gene therapy for dominant diseases

Most gene therapies are designed to supply a therapeutically beneficial gene, for example restoring the deficit caused by a loss-of-function mutation. Such mutations tend to be recessive. However, dominant diseases generally involve a mutation which causes a toxic gain-of-function or dominant-negative effect, where supplementing cells with the wild-type protein is usually not adequate to overcome the pathogenic effects of the mutant.

Prevention of the expression of the mutant protein is required for gene therapy of most dominant diseases. The development of RNA interference (RNAi) technologies provides a mechanism for post-transcriptional gene silencing for such conditions, directed specifically to target the mutant allele. Gene silencing technologies are becoming well-developed and reaching the clinic, yet substantial barriers still remain for certain groups of dominant diseases. One group is those caused by a large number of possible mutations in a single gene, such as amyotrophic lateral sclerosis (ALS) and several forms of retinal degeneration (Xia et al. 2005; O'Neill et al. 2000), which make the development of therapies for every patient a significant technical and financial challenge.

Another problematic group are disorders caused by mutations which are not easily amenable to allele-specific silencing by RNAi, such as expanded nucleotide repeats. These occur when a repetitive region of a gene, a motif of 2-6 nucleotides (nt), becomes expanded beyond the normal range, resulting in toxicity at the messenger RNA (mRNA) or protein level, which largely affects neuronal cells. These mutations cause more than 40 neurological, neurodegenerative and neuromuscular diseases (Pearson et al. 2005). A major constituent of this group are the polyglutamine (polyQ) diseases, where a CAG-repeat region is expanded, producing an abnormally long stretch of glutamines in the protein.

1.2.1.1 PolyQ diseases

The polyQ diseases are progressive and ultimately fatal neurodegenerative or neuromuscular diseases (although other tissues can also be affected), which are largely untreatable (Scholefield and Wood 2010). There are nine known polyQ diseases: Huntington's disease (HD), spinal and bulbar muscular atrophy (SBMA), dentatorubral-pallidoluysian atrophy (DRPLA) and six of the group of dominantly inherited ataxias, known as the spinocerebellar ataxias (SCAs) (Table 1-1). Of these, HD has a high incidence,

worldwide 3-7 per 100 000 people (Hersheson et al. 2012). In the UK, this may be as high as 12.3 per 100 000 (and rising), currently amounting to around 5 700 adults over 21, thus placing a substantial burden on healthcare systems (Evans et al. 2013). SBMA occurs in 1-2 per 100 000 people (Banno et al. 2012). Most of the others are rare, but incidence varies widely by country and population (Scholefield and Wood 2010; Hersheson et al. 2012). For example, in Japan, DRPLA has an incidence of 0.8 in 100 000 but this disease rarely occurs elsewhere (Becher et al. 1997; Hersheson et al. 2012). SCA7 is amongst the most common form of SCA in Mexico, Finland and in the indigenous black population of South Africa (Bryer et al. 2003; Hersheson et al. 2012). However, the most common form of SCA worldwide is SCA3 (also known as Machado-Joseph disease) (Hersheson et al. 2012).

Table 1-1 – The polyQ diseases. The known polyglutamine (polyQ) diseases are listed, with: the genomic locus and gene/protein in which the CAG/polyQ expansion is found; the pathogenic size range of the expansion and the tissue which is primarily affected (La Spada et al. 1991; Paulson 2009; Becher et al. 1997; Trottier et al. 1995).

Disease	Locus	*Gene*/Protein	Mutation size range	Primary affected tissue
HD	4p	*HTT*/Huntingtin	40+	Striatum
SBMA	Xq	*AR*/Androgen receptor	40-62	Motor neurons
DRPLA	12p	*ATN1*/Atrophin-1	49-75	Brain
SCA1	6p	*ATXN1*/Ataxin-1	39-82	Cerebellum
SCA2	12q	*ATXN2*/Ataxin-2	33-64	Cerebellum
SCA3	14q	*ATXN3*/Ataxin-3	52-86	Cerebellum
SCA6	19p	*CACNA1*/Calcium channel	19-30	Cerebellum
SCA7	3p	*ATXN7*/Ataxin-7	37-200	Cerebellum, retina
SCA17	6q	*TBP*/TATA binding protein	50-63	Cerebellum

1.2.1.2 Mutation-independent gene therapy

Specific targeting of the mutation (Figure 1-1a) is difficult in polyQ disorders due to the limitations of RNAi, which will be discussed in detail later. Instead, a mutation-independent therapy may be required. This would also be highly desirable in conditions with many possible disease-causing mutations, as mentioned previously.

To design a mutation-independent strategy, RNAi may be used to target both alleles indiscriminately (Figure 1-1b). However, the overall deficiency in the targeted protein which results from this approach may have detrimental (or unknown) consequences in patients (Scholefield and Wood 2010). Therefore, a functional (RNAi-resistant) copy of the gene could be provided concurrently to maintain the function of the protein, forming a gene knockdown and replacement (K&R) therapy (Figure 1-1c).

In the following section, the RNAi mechanism will be discussed in detail. This will include a recently discovered participant of this pathway, mirtrons, which may be particularly suited to K&R applications. Subsequently, the application of RNAi to gene silencing therapy will be covered. In particular, the limitations for problematic dominant diseases and the design of K&R therapies will be investigated.

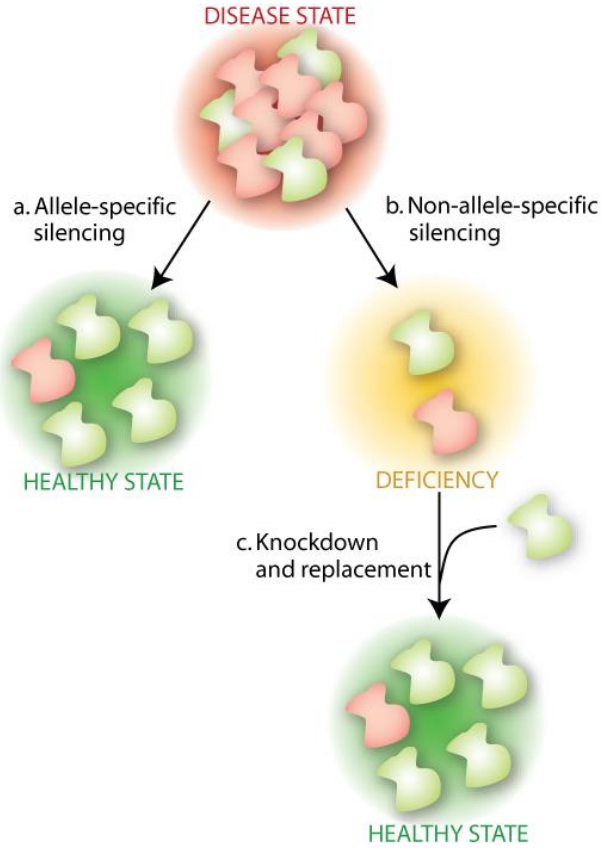


Figure 1-1 - Gene therapy strategies for dominant toxic gain-of-function conditions. In this representation of a dominant condition, in the disease state (top) the mutant (red) protein forms aggregates which may also incorporate the wild-type (green) protein. **(a)** Allele-specific silencing, where possible, reduces levels of mutant protein only, to prevent aggregation while leaving expression of the wild-type allele unaffected. **(b)** Non-allele-specific silencing may effectively reduce aggregation but lead to deficiency, which may cause known or unknown deleterious effects. **(c)** Knockdown and replacement adds functional protein alongside non-allele-specific knockdown to avoid deficiency. Published in (Curtis et al. 2012).

1.3 RNAi and mirtrons

1.3.1 Introduction to RNAi

Post-transcriptional gene silencing (PTGS) via RNA interference (RNAi) has attracted great interest as a gene regulatory mechanism in recent years due to its potential to fine-tune gene expression. The RNAi pathway appears to be well-conserved across mammals and invertebrates (Han et al. 2006; Lund et al. 2004; Liu et al. 2004; Khvorova et al. 2003; Lee et al. 2003). Through endogenous genome-encoded effectors termed microRNAs (miRNAs), RNAi is involved in development and normal physiology, whilst dysregulation has been linked to numerous pathologies ranging from cancer to neurodegenerative disease (Cullen 2009; Croce 2009; Eisenberg et al. 2009; Kim et al. 2007; Wang et al. 2008; Sibley et al. 2010). Exploitation of the sequence-specific silencing activity of RNAi has potential as a promising therapeutic strategy (Hannon and Rossi 2004).

1.3.2 Micro-RNA pathways

1.3.2.1 The canonical miRNA pathway

The canonical miRNA pathway involves step-wise processing of long genome-encoded primary-miRNA (pri-miRNA) transcripts into short (19-21 bp) double-stranded miRNA duplexes which act as substrates for the gene-silencing machine (Figure 1-2a) (Elbashir et al. 2001; Zamore et al. 2000). In the first step, a pri-miRNA is transcribed either from an intergenic region off its own separate promoter, or within an intron of a protein-coding gene. The pri-miRNA is processed by a microprocessor complex containing the nuclease Drosha and its cofactor, DiGeorge syndrome critical region 8 (DGCR8). Substrate recognition by the microprocessor requires an extended stem-loop, importantly flanked by single-stranded RNA (Han et al. 2006). This yields a shortened precursor-miRNA (pre-miRNA)

bearing just the stem-loop hairpin with a 2-nt overhang at the 3′ end, which acts as a recognition signal for Exportin-5 (Lund et al. 2004). Exportin-5 transports the pre-miRNA to the cytoplasm, where processing by the nuclease Dicer produces the double-stranded miRNA which bears 2-nt 3′ overhangs. These miRNA associate with an Argonaute (AGO)-containing protein complex to assemble the RNA-induced silencing complex (RISC) responsible for gene silencing (Liu et al. 2004). There are four AGO proteins (AGO1-4) in mammals, and two in *Drosophila* and *C.elegans*. One strand of the miRNA duplex, known as the guide strand or mature miRNA, is preferentially retained based on thermodynamic stability and the other is discarded (Khvorova et al. 2003). The activated RISC scans mRNA transcripts for sequence complementarity with the selected guide strand. If the guide strand and an mRNA have significant (but incomplete) base-pairing, translation is repressed (Baek et al. 2008). In contrast, where there is complete base-pairing (as commonly utilised in synthetic RNAi approaches using short-interfering RNAs [siRNAs]) the mRNA is cleaved in a process requiring AGO2 (Zeng et al. 2003; Rand et al. 2005). Recently, the explanation for the miRNA-directed translational repression effect in animals was primarily found to be mRNA destabilisation, rather than direct inhibition of translation. The loss of poly-adenylated tails (and hence destabilisation), was more significant than the reduction of ribosome activity. This appears to apply to both exogenously-applied miRNA (Hendrickson et al. 2009) and endogenous miRNA (Guo et al. 2010).

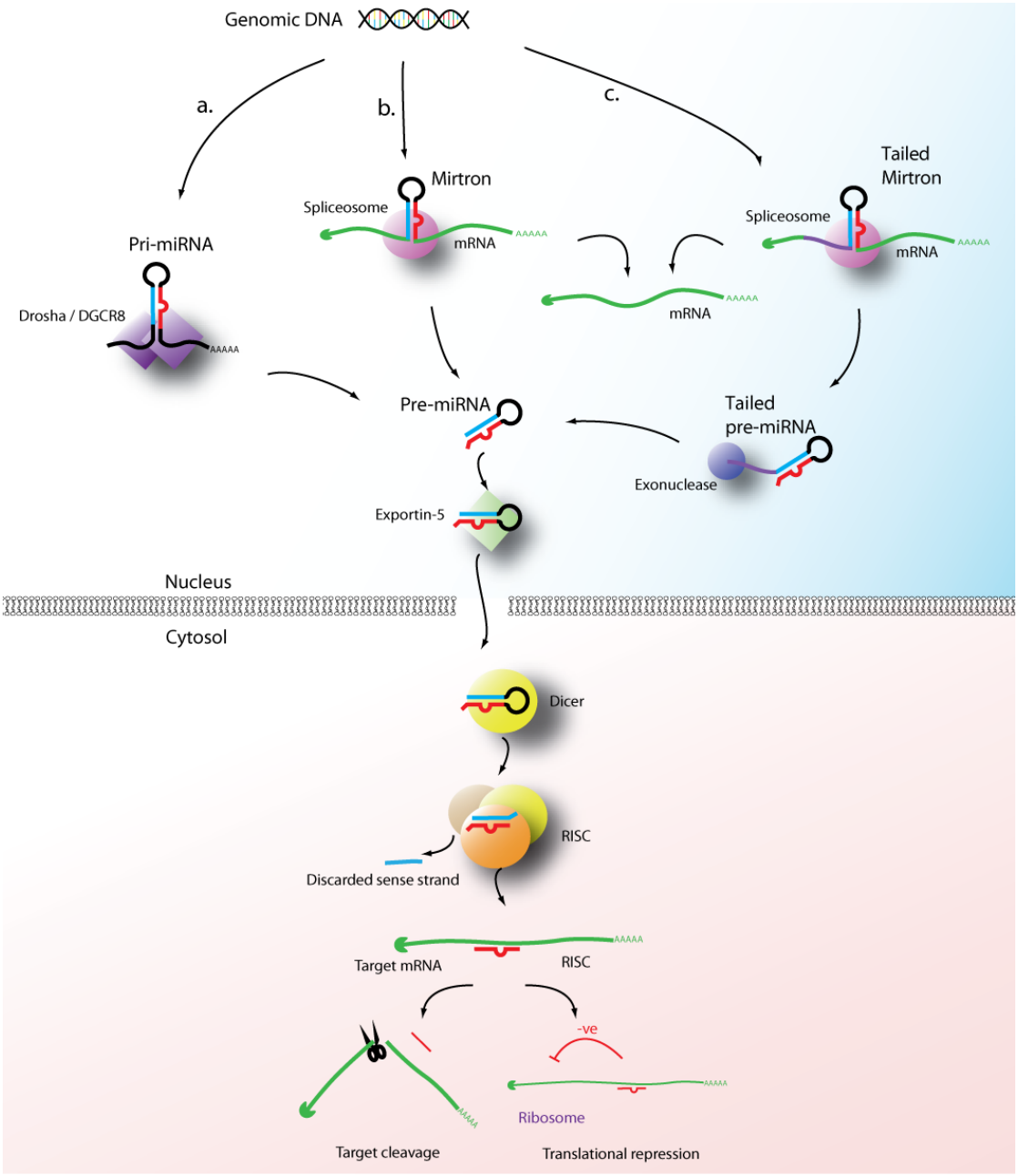


Figure 1-2 – The canonical miRNA, mirtron and tailed-mirtron pathways. (a) In the canonical mammalian miRNA pathway, pri-miRNA transcripts are processed by the microprocessor complex (Drosha/DGCR8) into pre-miRNA hairpins that are recognised by Exportin-5 for nuclear export. Following nuclear export and cleavage by Dicer, one strand is selected for recruitment into the RISC, resulting in a mature miRNA capable of directing translational repression of imperfectly matched targets, or target cleavage of perfectly matched targets. **(b)** In the mirtron pathway, pre-miRNA hairpins are produced directly from introns via splicing of pre-mRNAs, whereby functional mature mRNAs are also produced. The hairpins are then thought to join the canonical pathway at the stage of nuclear export, not requiring any processing by Drosha. **(c)** Tailed-mirtrons derive from slightly larger introns than mirtrons. Following splicing from the transcript, a single tail remains on one strand (here the 5' strand), which is removed by an exonuclease. Again these hairpins then join the canonical pathway for nuclear export. Published in (Curtis et al. 2012).

The “seed” region of the mature miRNA (positions 2-8 of the guide strand) has a key role in mRNA target silencing (Doench and Sharp 2004; Haley and Zamore 2004). It appears that near complete complementarity between the seed region and a 3′ UTR target site is necessary for the initiation of RNAi, and the pairing beyond this region regulates the rate of RISC catalysis (Haley and Zamore 2004).

Given that individual miRNAs often have the potential to repress many mRNA targets, the human mature miRNAs now identified (more than 1500) may post-transcriptionally regulate ~30% of the human mRNA transcriptome (John et al. 2004; Lim et al. 2005; Chi et al. 2009; Hafner et al. 2010). Over 20 000 mature miRNAs have been reported across 168 different organisms, hinting that the miRNA pathway is a vast post-transcriptional regulation network with important roles in both health and disease across many levels of evolution. A thorough understanding of this pathway is crucial for furthering knowledge of basic biology and in development of new therapeutics.

1.3.2.2 Micro-RNAs in disease and gene therapy

miRNAs are involved in many biological processes, including cell-cycle control (Hatfield et al. 2005), myogenesis (Eisenberg etal. 2007) and neuronal differentiation (Kim et al. 2007). Their misfunction is implicated in a number of diseases. In particular, levels of various miRNAs which regulate disease-relevant genes were found to be altered in conditions affecting the brain, including cancer (Ruan et al. 2009), PD (Kim et al. 2007) and HD (Packer et al. 2008). This potentially provides novel biomarkers for disease and potential targets for therapy, via delivery or blocking of natural miRNAs (Bader et al. 2010). Crucially for wider gene therapy applications, the pathway can also be hijacked, designing RNAi molecules to silence disease-causing genes in other conditions, which will be discussed in section 1.3.3.

1.3.2.3 The mirtron pathway

A number of natural variants of the miRNA biogenesis pathway have been revealed, including numerous Drosha-independent mechanisms (Figure 1-3). One significant example with potential applications in gene therapy (in particular K&R) is the novel splicing-dependent mirtron pathway (Figure 1-3a-c) (Seow et al. 2012; Sibley et al. 2012).

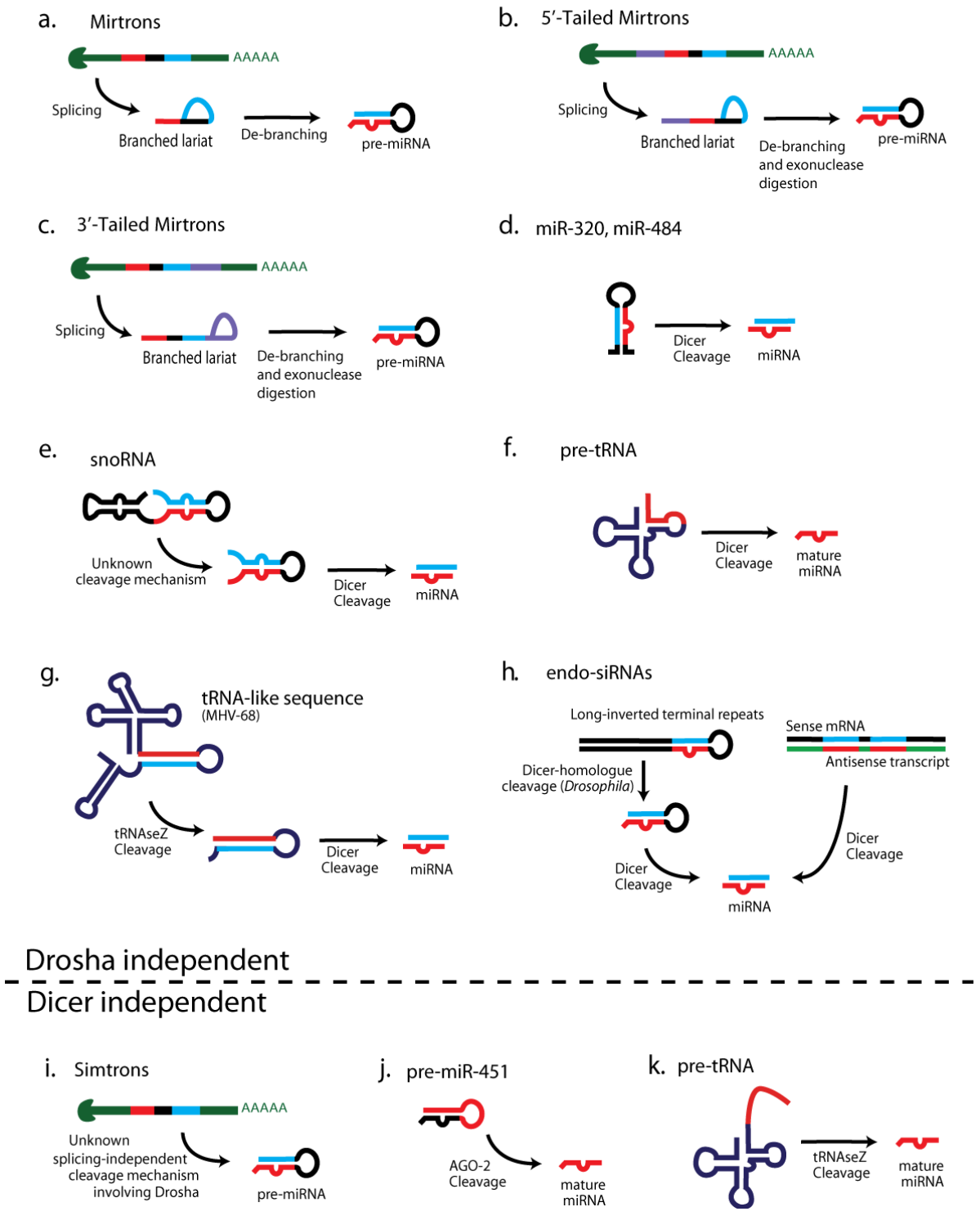


Figure 1-3 - Atypical biogenesis pathways for miRNA generation. (a) The mirtron pathway involves splicing of short introns with hairpin-forming potential. Following de-branching of the branched lariat intermediate of the splicing pathway, sequence homology between 5′ and 3′ ends of the intron allows pre-miRNA-like hairpins to form. **(b, c)** Tailed-mirtrons are very similar to mirtrons, arising from short introns with hairpin-forming potential. However, following de-branching the pre-miRNA-like hairpin has a single stranded tail on either the 5′ or 3′ end which requires exonucleolytic cleavage by a Drosha-independent mechanism. **(d)**

miR-320 and miR-484 are pri-miRNA sequences devoid of Drosha recognition characteristics. Dicer processes the hairpins directly to release mature miRNAs. **(e)** Some miRNAs map to small nucleolar RNAs (snoRNAs) containing secondary structures mimicking Dicer substrates. Following initial processing by a Drosha-independent mechanism, hairpin structures are processed by Dicer into mature miRNAs. **(f)** tRNA-like secondary structures can be processed by Dicer to release short RNA species from their 3' ends in mammals. **(g)** miRNAs encoded by the MHV-68 virus derive from complex secondary structures incorporating pol-III promoters and which are analogous to tRNAs. In initial processing, tRNAaseZ cleaves to release a hairpin sequence that is recognised by Dicer. **(h)** Endogenous siRNAs (endo-siRNAs) can derive from long inverted repeat sequences that are cleaved by a homologue of Dicer in Drosophila to shorter hairpins recognised by Dicer. Alternatively mRNA transcripts can anneal to complementary antisense transcripts and form dsRNA duplexes that are recognised and processed by Dicer. A number of pathways are also known which are independent of Dicer. **(i)** Some mirtron-like miRNAs may be processed via the simtron route, requiring Drosha but not splicing, DGCR8 or Dicer. **(j)** Pri-miR-451 is processed by Drosha into a hairpin shorter than canonical Dicer substrates, and the mature miRNA forms part of the hairpin loop. AGO2 has been found to mediate processing of the pre-miRNA into the mature sequence. **(k)** In addition to Dicer-dependent cleavage of tRNAs, tRNAseZ can also cleave 3' sequences of precursor tRNAs to release small RNAs that can enter gene silencing pathways. (Ruby et al. 2007a; Flynt et al. 2010; Havens et al. 2012; Ender et al. 2008; Lee et al. 2009; Haussecker et al. 2010; Bogerd et al. 2010; Czech et al. 2008; Tam et al. 2008; Watanabe et al. 2008; Cheloufi et al. 2010; Cifuentes et al. 2010; Yang et al. 2010). Published in (Curtis et al. 2012).

Mirtrons are transcribed as short introns within protein-coding genes. Although canonical miRNAs (Figure 1-2a) can also occur within introns, mirtrons are distinct from these in that the ubiquitous splicing and debranching process directly produces the hairpin-forming pre-miRNA (Figure 1-2b). This joins the miRNA pathway, bypassing the microprocessor, and requiring only processing in the cytosol by Dicer before integration into the RISC. The closely related tailed-mirtrons have a tail on one strand after splicing which requires removal via exonuclease digestion (Figure 1-2c) (Berezikov et al. 2007; Flynt et al. 2010). Examples of each are shown in Figure 1-4. Mirtrons may even possess tails on both ends (two-tailed mirtrons) (Li et al. 2013).

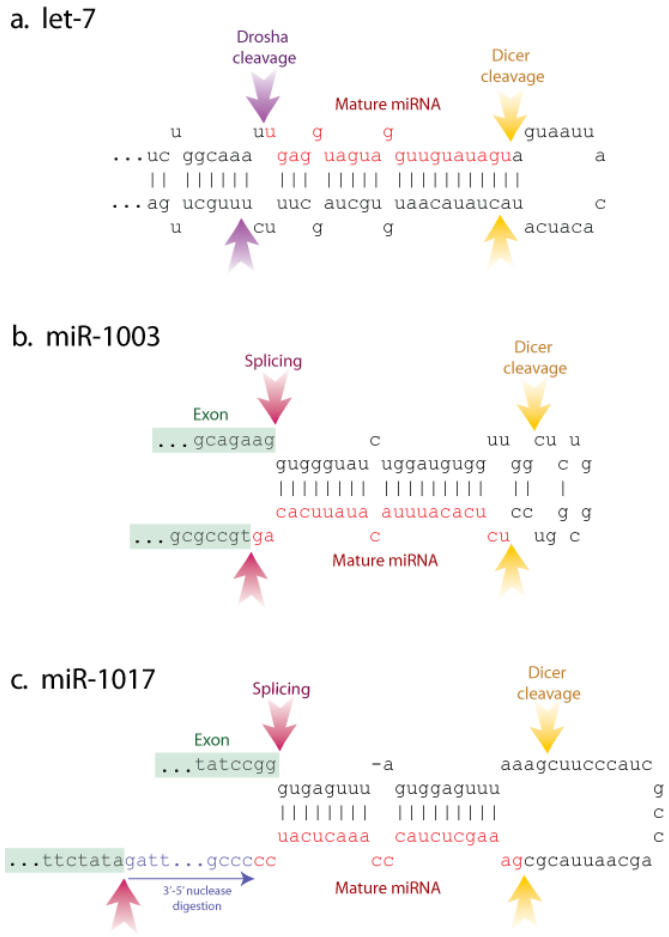


Figure 1-4 - Examples of canonical miRNAs, mirtrons and tailed mirtrons from *Drosophila melanogaster*. (a) let-7, a canonical miRNA. The primary transcript contains an extended hairpin which is first cleaved by Drosha to produce the pre-miRNA. The loop region is then removed by Dicer and the mature miRNA strand is incorporated into RISC. **(b)** miR-1003, a mirtron: a miRNA produced from an intron without Drosha cleavage. The primary transcript is spliced, forming a functional mRNA and a pre-miRNA hairpin which is cleaved by Dicer as in (a). **(c)** miR-1017, a tailed mirtron. The primary transcript is spliced as in (b), except that the pre-miRNA hairpin produced contains a tail of around 100 nucleotides on the 3' arm. This is digested by a 3'-5' exonuclease (Flynt et al. 2010). Published in (Curtis et al. 2012).

1.3.2.4 Discovery of mirtrons

The first mirtrons were identified in *D.melanogaster* and *C.elegans* (Ruby et al. 2007a), where dependence on splicing and debranching was shown. Unlike canonical miRNAs, Drosha was not an essential requirement for mirtron processing (Ruby et al. 2007a; Okamura et al. 2007; Smibert et al. 2011). *Drosophila* mirtrons were shown to produce functional miRNA, through their ability to silence a Luciferase reporter target (Ruby et al. 2007a; Okamura et al. 2007). Further, the absence of mirtron-specific exonic sequence elements was proven, since mirtrons could be processed when inserted as introns of a gene other than their endogenous host (Okamura et al. 2007).

Tailed-mirtrons were also confirmed in *Drosophila*. Only the 5′ end of the miR-1017 hairpin was derived from the end of an intron (Figure 1-2c, Figure 1-4c). The 3′ end lay within the intron, meaning that, after splicing, the hairpin would have a single-stranded tail of around 100 nucleotides; this was removed by the 3′-5′ exonuclease activity of the RNA exosome (Flynt et al. 2010).

In summary, the identification of mirtrons and tailed-mirtrons in invertebrates defined a new sub-class of miRNAs in which at least one of the hairpin ends is defined by splicing.

1.3.2.5 Identification and biogenesis of mammalian mirtrons

A systematic bioinformatics search revealed 13 well-conserved candidate mirtrons in mammals, with hairpin termini corresponding to the predicted splice sites of short introns (Berezikov et al. 2007). The most highly expressed mammalian mirtrons (miR-877 and miR-1224) were also the most conserved, suggestive of functional relevance. The mammalian brain was thought to be a likely organ to reveal mirtrons given that neurons employ much translational regulation (Ruby et al. 2007b). From this further search, 16 primate-specific mirtrons were confidently classified, including a number of 5′- or 3′-tailed mirtrons

(Berezikov et al. 2007). Removal of the tail was shown to be not carried out by Drosha, but the precise exonuclease responsible in mammals remains unidentified (Babiarz et al. 2008).

Mammalian mirtrons can produce mature species from either the 5′ or 3′ end (Berezikov et al. 2007), or even both strands simultaneously (Schamberger et al. 2012). The functionality of mammalian mirtrons was initially investigated using expression constructs where miR-877, 1224 or 1226 was cloned as an intron within the open reading frame of an enhanced green fluorescent protein (EGFP) gene (Sibley et al. 2011). The presence of EGFP fluorescence confirmed that mirtrons were accurately spliced from these transcripts. As expected, splicing was abolished by single base mutations in the 5′ splice site. All three mirtrons were found to silence Luciferase reporter constructs, in a splicing- and sequence-dependent manner (Sibley et al. 2011). Unlike canonical miRNAs (Han et al. 2009), mirtron expression from the EGFP constructs was unaffected by Drosha knockdown.

Reduced miR-877 target silencing in the presence of viral competitor VA1 suggested dependence on both Exportin-5 and Dicer (Lu and Cullen 2004). Unlike the 2-nt 3′-overhang found in canonical miRNAs and invertebrate mirtrons, human miR-877 has a 3-nt 5′ overhang and a 4-nt 3′ overhang. Indeed, numerous unusual hairpin end structures were apparent in mammalian mirtrons, the most common being a single-nucleotide overhang on both strands (Berezikov et al. 2007). Exportin-5 can bind hairpins which bear blunt ends, short (2-nt) overhangs on both strands or a long 3′ overhang (up to 5-nt) (Zeng and Cullen 2004). It cannot bind hairpins with long (5-nt) overhangs on both strands. miR-877 does not strictly fit into the above group of substrates recognised by Exportin-5. However, given the observations with VA1, it appears that miR-877 is indeed transported by Exportin-5. Conversely, miR-1224, which has single-nucleotide overhangs on both strands, could not be inhibited by competition with VA1. This suggests that this mirtron may be subject to an

alternative nuclear export route and/or Dicer independence. Although this remains to be fully investigated, it is potentially feasible for a different nuclear export factor to recognise some mirtron hairpins (Murphy et al. 2008), and Dicer-independent miRNA processing is also possible (Figure 1-3).

Finally, multiple seed-matched target sites of miR-877, identified in neuronal genes, were cloned into Luciferase constructs and shown to be repressed by miR-877 (Sibley et al. 2011). This was extended to show that levels of the endogenous FXR2 (fragile X mental retardation, autosomal homolog 2) transcript were reduced by over-expression of miR-877 *in vitro*. However, the first bona-fide target of a mammalian mirtron to be confirmed *in vivo* was vascular endothelial growth factor receptor 3 (VEGFR-3), which was suppressed by mirtron miR-1236 (Jones et al. 2012). It was further shown that injection of miR-1236 inhibited lymphangiogenesis and angiogenesis (downstream effects of VEGFR-3) *in vivo*, and also that this mirtron was up-regulated in response to inflammatory cytokines. This provides the first evidence for a major role of a mirtron in mammalian biology.

It was recently shown that two previously predicted mirtrons (miR-1225 and miR-1228) may be alternatively processed via a pathway which does not require splicing (Havens et al. 2012). These splicing-independent mirtron-like miRNAs are referred to as simtrons. These mature miRNAs were still detected despite mutations in the splice site recognition sequences. The biogenesis of simtrons was insensitive to manipulations to Dicer, AGO2 and DGCR8, but affected by a dominant-negative Drosha variant. It appears that a novel Drosha recognition motif allows a subset of mirtrons to be processed via the simtron pathway, but perhaps only when splicing is impaired. It is also possible that naturally splicing-incompetent miRNAs may exist which are predominantly processed via this pathway. Further

investigation is required to accurately define the simtron pathway and determine its significance.

In summary, there is clear evidence that the mirtron and tailed-mirtron pathways are functional in mammals. Together, they create over two hundred miRNAs in humans – one sixth of the total human miRNA repertoire (Ladewig et al. 2012). Mirtron-derived mature miRNA species can be detected, which possess silencing ability and may be derived from either arm. Their processing is independent of Drosha, but requires splicing and cleavage by Dicer. Although mirtrons are largely dependent on Exportin-5 for nuclear export, it is possible that other factors may be involved in some cases. Finally, evidence is now emerging to show that mirtrons play physiological roles in humans, and it will be important to identify further targets of mirtrons in future. Most relevant to the field of gene therapy, exogenous application of mirtrons was shown to be capable of knocking down endogenous targets (Sibley et al. 2011).

1.3.3 Synthetic RNAi in gene therapy

Following the discovery that the RNAi pathway could be hijacked to elicit silencing of a target gene in mammalian cells through the delivery of siRNA directed towards a specific sequence (Elbashir et al. 2001; Dykxhoorn et al. 2003), there has been a wide exploration of RNAi-based gene silencing for therapeutic applications. This shows promise in a number of diseases and has advanced to clinical trials.

1.3.3.1 siRNA

The application of synthetic siRNAs has been tested in a number of clinical trials, including a rare skin disorder, respiratory syncitial virus infection and solid tumours, which were well-tolerated and some silencing efficacy was shown (Davis et al. 2010; Leachman et al. 2010;

DeVincenzo et al. 2008, 2010; Strumberg et al. 2012). siRNAs have also been delivered to the eye; for example, repeated siRNA application to patients with diabetic macular edema via intravitreal injection provided small but measurable improvements in visual acuity (Nguyen et al. 2012). However, the study was terminated after one year due to its limited efficacy, attributed to the low doses used.

Most of the clinical trials involved repeated administration or short term follow-up and most were administered locally. For neurological conditions, the costs and invasiveness of this approach may be prohibitive (Soutschek et al. 2004). Modifying siRNAs increases their biostability (Whitehead et al. 2009) but delivery to the brain remains a limiting factor.

1.3.3.2 shRNA and miRNA

Short hairpin RNA (shRNAs) are designed as DNA to be transcribed in target cells under a strong, constitutive pol-III promoter, producing RNA hairpins which can be processed by Drosha and Dicer into siRNAs. shRNA expression cassettes can be successfully delivered to cells using viral vectors and achieve effective in target silencing *in vivo* (Scherr et al. 2003). However, shRNAs can also trigger an interferon response, affecting global gene expression and causing cell death (Bridge et al. 2003).

High doses can lead to saturation of the RNAi processing pathway and competition with endogenous miRNAs, in some cases resulting in fatal toxicity *in vivo* (Grimm et al. 2006; Borel et al. 2011; Grimm 2011). Toxicity was attributed to reduced activity of endogenous miRNAs, particularly in the liver, where each cell received perhaps 1000 viral particles (Borel et al. 2011; Grimm 2011). However, such effects are not limited to the liver; dose-dependent shRNA toxicity was also observed in the mouse CNS (McBride et al. 2008; Boudreau et al. 2009a), which is particularly relevant for the development of novel polyQ disease therapies. This is also common across other tissues and other species, including large mammals.

Reducing vector doses can reduce toxicity, but compromises overall transduction efficiency. However, incorporating effective shRNA guide strands into the scaffolds of endogenous miRNA sequences (i.e. creating miRNA mimics) has been shown to reduce adverse effects (McBride et al. 2008; Borel et al. 2011). The use of a weaker promoter can also be pivotal, producing lower expression levels and minimising the burden on the RNAi machinery. miRNA mimics may even be more potent than shRNAs at the same concentrations (Giering et al. 2008; Borel et al. 2011; Maczuga et al. 2012). Significantly for the prospect of treatments for SCAs (Table 1-1), cerebellar toxicity caused by shRNAs was not detected with miRNA mimics, even with comparable silencing efficacy (Boudreau et al. 2009a). Furthermore, pol-II promoters can be tissue-specific and endogenously controlled to limit off-target effects (Nielsen et al. 2009) and pol-II miRNA mimics can be effective at with a single integrated copy per cell (Stegmeier et al. 2005).

miRNA mimics are showing substantial promise as mediators of therapeutic gene silencing. Safety concerns must also be taken into consideration for miRNAs. For example, assessment of off-target effects, caused by binding of the mature species or the passenger strand to other complementary targets in other genes and 3' UTRs is important for the safety of such therapies (Boudreau et al. 2009b).

1.3.3.3 Artificial mirtrons

Like other endogenous miRNAs, mirtrons may also be mimicked to silence a gene of interest (Seow et al. 2012; Sibley et al. 2011). Mirtrons may confer additional benefits. By bypassing the microprocessor altogether, and perhaps also Exportin-5 in some cases (Okamura et al. 2007; Sibley et al. 2011), they place minimal dependence on some aspects of the saturable RNAi pathway. Indeed, it was revealed that both Drosha and its cofactor DGCR8 can be sequestered into nuclear RNA aggregates in patient brain cells in a nucleotide expansion

disorder. This led to potentially harmful decreases in processing of endogenous miRNAs (Sellier et al. 2013). Therefore, RNAi-based therapies which are not dependent on the microprocessor may be important. Additionally, Drosha processing can be subject to variability (Wu et al. 2009a). By utilising the splicing pathway, which is accurately directed by specific sequence motifs, mirtrons may be processed into more accurate RNAi species, limiting the potential off-target effects. Having well-defined hairpin ends can also be advantageous for the accuracy of Dicer processing (Park et al. 2011). Overall, artificial mirtrons are an exciting novel technology which may be applied to gene silencing therapy. As introns, they may be particularly well-suited to being delivered within a therapeutically advantageous transgene.

1.3.3.4 Limitations of RNAi in allele-specific silencing

RNAi can be sufficiently specific to distinguish two targets with just a single base difference (Elbashir et al. 2001), which is an exciting prospect for allele-specific gene therapy. However, as mentioned previously, many mutations remain problematic, for a number of reasons. Firstly, achieving substantial allele-specificity with just a single base pair difference can be challenging (Kubodera et al. 2011). Secondly, where there are many possible mutations in a single gene, the development of RNAi treatments for every patient may be prohibitively costly. Examples include retinitis pigmentosa (RP) (and similar cases of retinal degeneration), which can be caused by over 150 dominant mutations in Rhodopsin and also by mutations in rds-peripherin (Millington-Ward et al. 2011) and amyotrophic lateral sclerosis (ALS), linked to over 150 mutations in Superoxide Dismutase 1 (SOD1) (Andersen and Al-Chalabi 2011). Thirdly, diseases caused by a nucleotide repeat expansion, including the polyQ disorders, present difficulties because siRNA molecules are typically too short to individually distinguish the size of an expanded tract from a normal one, or any others in the genome. In each of

these cases, inability to achieve allele-specific silencing can be harmful where the protein has an important physiological role. The routes through which such mutations could be treated with gene therapy will be discussed in the next section, with particular reference to the polyQ diseases.

1.4 Gene therapy for polyQ disorders

The polyglutamine (polyQ) diseases, caused by expanded CAG repeats, are problematic cases for RNAi-based gene silencing therapy. Therapeutic improvements in SCA1, SCA3 and HD have been shown *in vivo* by silencing the mutant human transgene (Xia et al. 2004; Alves et al. 2008; DiFiglia et al. 2007), relying on sequence differences between the human and the mouse genes. However, in patients the alleles are not so easily distinguishable.

Two potential strategies are available to attempt to achieve allele-specific silencing in polyQ diseases, and in some cases non-allele-specific silencing may be a realistic option. Combined non-allele-specific gene knockdown-and-replacement (K&R), however, may be the most widely applicable approach. This section will cover each of these approaches as well as their advantages and pitfalls.

1.4.1 Non-allele-specific silencing

Non-allele-specific silencing of a gene is the most straightforward approach to develop in principle because the entire gene sequence can be examined for optimal RNAi target sites. Although this indiscriminate silencing could be effective in reducing the toxic gain-of-function effects of the mutant, the protein's physiological function would be largely inhibited. Therefore, this approach is only appropriate where the mutation affects a protein which does not perform an essential function. This may be the case in HD. Partial suppression of Huntingtin (30% mRNA, 45% protein) was achieved with virally delivered

shRNA via direct injection into the striatum of adult non-human primates (Grondin et al. 2012). No pathogenic effects were detected, indicating that non-allele specific therapy for HD may be well-tolerated. However, for most other polyQ diseases, a non-allele-specific approach is not deemed ideal given the known function or lack of available data (Scholefield and Wood 2010).

1.4.2 Allele-specific silencing for polyQ disorders

Allele-specificity towards CAG repeat motifs may be achieved through increased number of target sites in the expanded allele. Alternatively, disease-associated single nucleotide polymorphisms (SNPs) can allow the mutant allele to be specifically targeted (Scholefield et al. 2009).

1.4.2.1 SNP-directed therapies

Some SNPs tend to occur on the same allele as an expanded repeat mutation, and targeting these sequences can achieve allele-specific silencing *in vitro* (Alves et al. 2008). For each polyQ disorder, this may allow treatment of certain patient groups; for example, targeting three different SNPs in HD may provide treatment for 75% of HD patients (Pfister et al. 2009). However, across the polyQ disorders, some patients may not possess a targetable SNP (either having no single-base differences between their normal and disease alleles, or having only SNPs which are not easily discernible by RNAi) (Varela et al., in preparation). Furthermore, most studies on allele-specific silencing (~95%) are based exclusively on patients of European descent, therefore only covering a subset of human genetic variation. The high cost involved in the clinical development for each molecular target will tend to leave behind patients with the rarest genotypes.

1.4.2.2 Expansion-directed therapies

Targeting the mutant allele with antisense sequences complementary to the expansion itself may achieve allele-specificity. This is based on the CAG tract in the mutant allele being longer and having altered secondary structure compared to the wild-type allele, which appears to offer better access for increased binding of antisense molecules compared to the wild-type version. Single-stranded antisense oligonucleotides (ASOs) can be used as an alternative to siRNA, to cause steric blocking of transcripts or direct them to be cleaved by the nuclear RNaseH endonuclease (Wheeler et al. 2012). In some cases, silencing of long repeat-containing mRNA can be detected in mice one year after administration is ceased (Wheeler et al. 2012). This approach has been pursued in several nucleotide expansion disorders including Myotonic Dystrophy, HD and several ataxias (Mulders et al. 2009; Wheeler et al. 2009; Hu et al. 2009; de Mezer et al. 2011). Encouragingly, it was shown that these treatments do not affect other genes containing repeat sequences (Hu et al. 2009, 2010).

One major advantage of this approach is that a single antisense molecule may be applicable to all of the polyQ disorders. This could be essential for treating rare diseases. Very long repeats may be particularly amenable to silencing by this mechanism (Wheeler et al. 2012). However, it can be less effective on shorter expansions such as those occurring in the SCAs. For example, SCA6 is particularly difficult to target this way, since just 28 CAG repeats can be disease-causing (Kubodera et al. 2005). Further, the off-target silencing effects on long repeats in the **non**-protein coding transcriptome have not been fully investigated. Crucially, it has not yet been shown that targeting an expansion can prolong life in an *in vivo* model.

Whether targeting the repeat itself or a SNP, allele-specificity is lost at high effector concentrations (Hu et al. 2009; Scholefield et al. 2009), yet clearly, inadequate silencing

would be achieved at low effector levels. This concentration-dependence presents issues of dose determination for clinical administration. The major limitations of allele-specific and non-allele-specific therapies are summarised in Table 1-2.

Table 1-2 – Summary of the limitations of RNAi-based silencing therapies for polyQ diseases.

	Causes deficiency	Highly dose-sensitive	High risk of off-target effects	Limited patient coverage
Targeting the repeat		*x*	*x*	
Targeting SNPs		*x*		*x*
Non-allele-specific	*x*			

1.4.3 Knockdown-replacement

Considering the limitations of allele-specific therapies and the risks of non-allele-specific silencing, a gene knockdown-and-replacement (K&R) approach may hold the most potential for widely-applicable treatments for nucleotide expansion disorders such as the polyQ diseases (Rossmiller et al. 2012).

This approach can be achieved by targeting RNAi silencing toward a region of the gene free from any mutations, hence silencing both alleles, while concurrent delivery of a functional, RNAi-resistant copy of the gene would maintain the protein's physiological role. This could therefore provide a mutation-independent treatment for all patients affected by any dominant mutation in the same gene.

In order to be expressed in the presence of RNAi, the replacement gene must be RNAi-resistant. If the target site is within the 3′ UTR, the replacement gene can avoid silencing through complete or partial absence of the 3′ UTR. If the coding sequence is targeted,

redundancy in the genetic code allows for codon substitutions to be made within the target region, destroying complementarity with the RNAi effector without changing the amino acid sequence (Xu et al. 2009). Using a pol-II rather than pol-III promoter, perhaps the gene's endogenous promoter, avoids excessive over-expression of the transgene and allows tissue-specificity (Sibley et al. 2012).

Various approaches to gene knockdown-replacement have been tested and novel approaches may be developed. These strategies are summarised in Table 1-3 and Figure 1-5, and discussed in the following sections, with respect to proof-of-principle of the K&R approach, and the design of such therapies.

1.4.3.1 Proof of principle

Evidence that RNAi-based gene K&R is a feasible approach to gene therapy has been presented in a number of studies, which are summarised in Table 1-3 A. In one of the first attempts at RNAi-based gene K&R, the mutated calcium channel gene (CACNA1A) was successfully silenced and concurrently replaced with an RNAi-resistant codon-exchanged copy in a cell culture model of SCA6 (Kubodera et al. 2005).

Proof of this principle was also achieved *in vivo* in a K&R therapy for ALS. Mutant SOD1 suppression and concurrent replacement was demonstrated in liver after systemic administration (Kubodera et al. 2011).

Notably, K&R has become the most well-developed for forms of retinal degeneration such as RP. K&R successfully targeted WT rds-peripherin in retinal explants and also in the retina, where the replacement gene was shown to rescue the function of the repressed WT gene (Palfi et al. 2006; Petrs-Silva et al. 2012). K&R has been demonstrated for Rhodopsin in several studies (Kiang et al. 2005; O'Reilly et al. 2007; Millington-Ward et al. 2011;

Greenwald et al. 2012). Crucially, in transgenic models of retinal degeneration, phenotypic improvement was shown compared to eyes receiving only a GFP transgene (O'Reilly et al. 2007). However, greater phenotypic improvement through K&R has not been demonstrated when compared to silencing only. The transgenic mice used are in some cases more amenable to gene replacement therapy, particularly those with a *rho*-knockout background (Millington-Ward et al. 2011; Greenwald et al. 2012).

Table 1-3 – Studies involving concurrent gene silencing and replacement. **A**: Those involving silencing and replacement of the same gene; **B**: Those involving silencing and replacement of two different genes. The RNAi silencing mechanism is given, where dagger (†) indicates shRNA not delivered under a pol-III promoter. No of const = Number of DNA/RNA constructs delivered. Promoters: both number and identity are given, with the RNAi promoter stated first. Abbreviations: endo = endogenous; tg = transgenic; h = human; m = mouse; mut = mutant; cod-ex = codon-exchanged (RNAi-resistant); HIV-1 = human immunodeficiency type 1; huRevM10 = transdominant negative mutant (humanized) form of HIV-1 Rev; KO = knockout; CMV = cytomegalovirus; CBA = chicken beta-actin; H1 and U6 are both pol-III promoters.

RNAi	Application	Gene silenced	Gene delivered	Model		No of const	Promoters	Reference
A								
siRNA	SCA6	CACNA1A	CACNA1A (*cod-ex*)	*In vitro*	Transiently transfected cells	2	1 (siRNA)/CMV	Kubodera 2005
siRNA	Retinal degeneration	mRHO (endo)	mRHO (*cod-ex*)	*In vivo*	WT mouse	2	1 (siRNA)/CMV	Kiang 2005
shRNA	Retinal degeneration	hRHO (tg)	hRHO (*cod-ex*)	*In vivo*	Tg mouse (RHO-KO background)	2	2 CMV/mRho	Millington-Ward 2011
shRNA	Retinal degeneration	peripherin (endo)	peripherin (*cod-ex*)	*In vivo*	WT mouse	2	2 H1/CBA	Petrs-Silva 2012
miRNA	Retinal degeneration	hRHO (tg)	RHO (*cod-ex*)	*In vivo*	Tg mouse (RHO-KO background)	2	2 CBA/CMV	Greenwald 2012
shRNA	Retinal degeneration	hRHO (tg)	hRHO (*cod-ex*)	*In vivo*	Tg mouse (RHO+/- background)	1	2 H1/CMV	O'Reilly 2007
shRNA	ALS	SOD1	SOD1 (*cod-ex*)	*In vivo*	Tg mouse	1	2 U6/CMV	Kubodera 2011
miRNA	Retinal degeneration	peripherin (endo)	peripherin (*cod-ex*)	*Ex vivo*	WT mouse retinal explants	1	2 H1/CMV	Palfi 2006
B								
shRNA	Type I diabetes	iNOS	VEGF	*In vitro*	AD-293 cells	1	2 U6/CMV	Li and Mahato 2009
shRNA †	Sickle cell anaemia	β-globin (mut)	γ-globin	*Ex vivo*	Patient CD34(+) cells	1	1 β-globin	Samakoglu 2006
shRNA †	HIV-1	HIV-Rev	HuRevM10	*In vitro*	HIV-infected CEM T cells	1	1 HIV-1-inducible	Unwalla 2006
miRNA	Type I diabetes	iNOS	VEGF	*In vitro*	AD-293 cells	1	1 CMV	Li and Mahato 2009
miRNA	Solid tumour vaccine	Furin	GM-CSF	*Ex vivo*	Advanced cancer (Phase I Clinical)	1	3 CMV/CMV/CMV	Senzer 2012

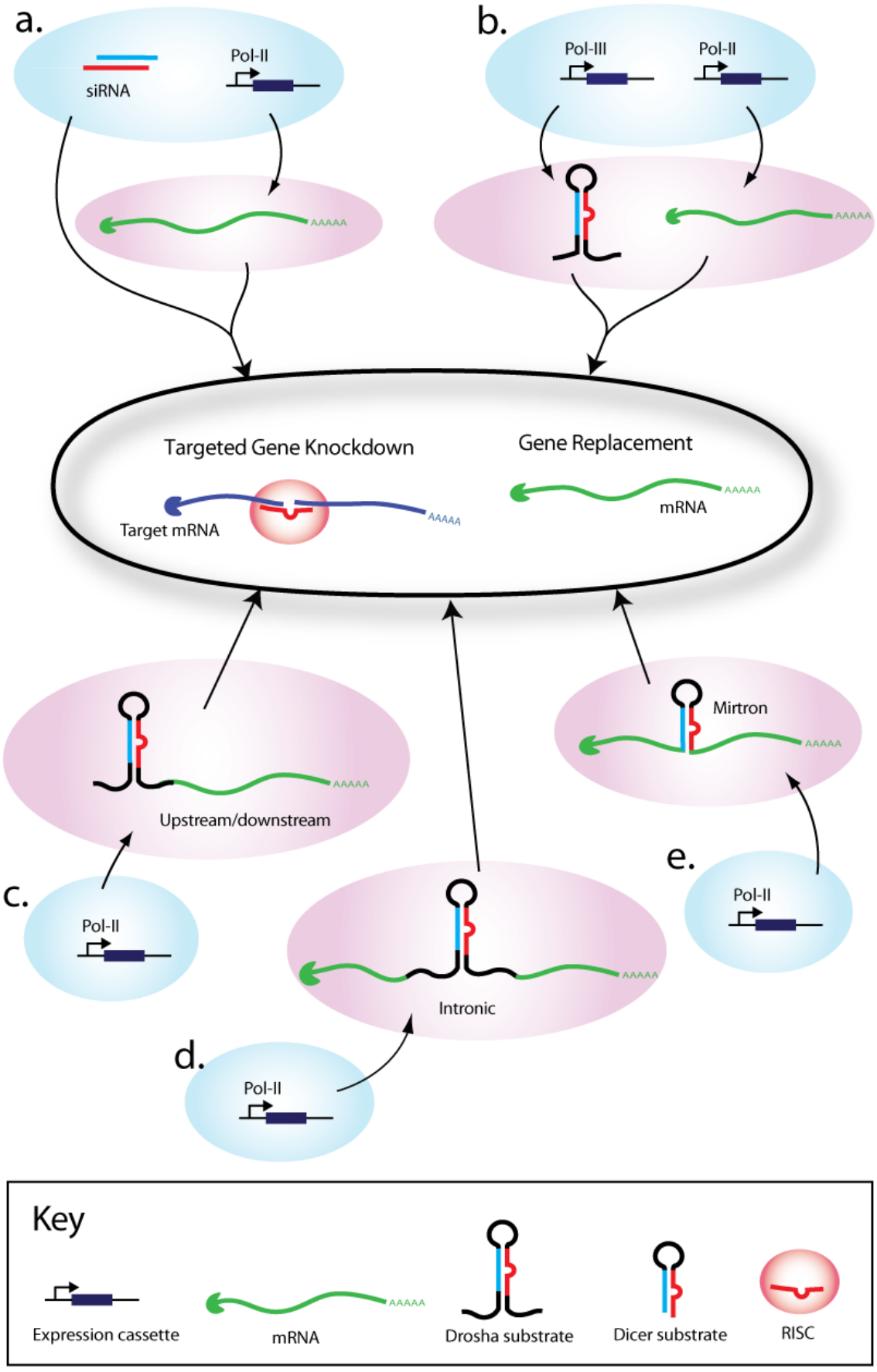

a.
siRNA
Pol-II
AAAAA
b.
Pol-III
Pol-II
AAAAA
Targeted Gene Knockdown
Target mRNA
AAAAA
Gene Replacement
mRNA
AAAAA
Upstream/downstream
AAAAA
Mirtron
AAAAA
c.
Pol-II
Intronic
AAAAA
e.
Pol-II
d.
Pol-II
Key
Expression cassette
mRNA
AAAAA
Drosha substrate
Dicer substrate
RISC

Figure 1-5 - Potential strategies for gene knockdown-replacement (K&R) therapy. Notable references are included. **(a)** Delivery of a targeting siRNA molecule and a transcriptional unit encoding a pol-II driven replacement gene (Kubodera et al. 2005; Kiang et al. 2005). **(b)** Delivery of two cassettes: a pol-II/III-driven shRNA or miRNA and pol-II driven replacement gene, either within separate viruses (Millington-Ward et al. 2011; Petrs-Silva et al. 2012; Greenwald et al. 2012) or combined into one (O'Reilly et al. 2007; Palfi et al. 2006; Li and Mahato 2009; Kubodera et al. 2011; Senzer et al. 2012). **(c)** Delivery of a single pol-II transcription unit, with a shRNA/miRNA either upstream or downstream of the replacement gene (Unwalla et al. 2006; Li and Mahato 2009). **(d)** Delivery of a single pol-II transcriptional unit where a shRNA/miRNA mimic is expressed within an intron in the replacement gene (Samakoglu et al. 2006). The shRNA/miRNA is processed by Drosha and the transcript is also processed by splicing. **(e)** As (d), with a mirtron mimic as a special case of intronic miRNA, which does not require Drosha processing.

1.4.3.2 Approaches to K&R

In order to investigate the design of K&R strategies, a number of other studies should also be considered, where the targets of silencing and replacement are two different genes (Table 1-3 B); this also highlights some of the wider applications of this approach. Any of these strategies may potentially be applied to other situations, and are discussed below.

Initial K&R (and similar) studies explored the use of siRNA (Figure 1-5a). The limitations of siRNA and advantages of stable expression systems (sh/miRNA), as discussed in sections 1.3.3.1 and 1.3.3.2, also apply here. K&R studies involving stable expression systems (Figure 1-5 b-d) are of the greatest interest as therapeutic strategies.

A clinical trial of a combination therapy assessed the safety of a novel vaccine designed to prevent recurrence of cancer following surgery (Senzer et al. 2012). The plasmid construct successfully strongly expressed a GM-CSF transgene and two shRNA structures (based on the miR-30a backbone) both targeting furin convertase, which had the indirect effect of strongly down-regulating both TGF-β2 and TGF-β3. Encouragingly, patient survival was

significantly greater in the patient group. However, the plasmid was delivered *ex vivo* by electroporation of tumour cells, which is not a viable option for the brain.

Delivering transgenes to the brain and CNS is generally limited to the use of viral vectors, such as lentiviruses, which are capable of infecting non-dividing cells (Linterman et al. 2011). For example, a lentiviral vector delivered a therapeutic transgene to PCs in mouse models of SCA3 (Torashima et al. 2008). However, the current limitations include local injection for delivery to the brain and risks of insertional mutagenesis (Lundberg et al. 2008). Certain recombinant Adeno-associated viruses (AAV) can reach the CNS through systemic delivery (Xie et al. 2011) and AAV-mediated delivery was used in some of the aforementioned K&R studies for the retina. However, AAV can package expression cassettes no larger than 5 kb (Wu et al. 2010).

Therefore, delivering each portion of a K&R therapy under separate promoters can restrict the choice of viral vectors. One solution, delivering the therapy in two separate viruses, can lead to cells receiving only one aspect of the therapy, limiting its efficacy (Greenwald et al. 2012)

Hence, in some of the above cases, the shRNA/miRNA was expressed either upstream (Unwalla et al. 2006), downstream (Li and Mahato 2009) or within an endogenous intron (Samakoglu et al. 2006) of the transgene (Figure 1-5 c and d). This led to successful expression and/or function of both aspects of the therapy. However, placing an shRNA sequence within an endogenous intron must be carefully assessed for potentially harmful interferon responses (Samakoglu et al. 2006). Further, none of these approaches were validated in animal models.

1.4.3.3 Intronic miRNA in K&R

Mimics of endogenous intronic miRNA (including mirtrons) are, in theory, ideal for K&R therapies as they can be nested within the replacement gene Figure 1-5 d and e). Towards this application, mimics of intronic miRNAs (miR-155, the miR-106b cluster, and mirtron miR-1224) have been developed, but not delivered with a therapeutic transgene (Chung et al. 2006; Aagaard et al. 2008; Seow et al. 2012; Sibley et al. 2011).

Mirtrons may have a number of processing-related advantages over other miRNAs due to their dependence on splicing rather than the microprocessor, as discussed previously (section 1.3.3.3). Additionally, they present further benefits in K&R applications, their small size making them both relatively easy to manipulate and well-suited for limited-capacity delivery systems. Endogenous mirtrons can be expressed as introns outside of their natural context and can successfully silence a gene of interest (Sibley et al. 2011; Seow et al. 2012; Sibley et al. 2012). Multi-target or multi-gene silencing may also possible with mirtrons, through delivery of multiple mirtrons within a single gene (Sibley et al. 2012).

1.5 Gene therapy for SCA7

1.5.1 SCA7 as a model

Spinocerebellar Ataxia Type 7 (SCA7) is one of the rarer polyQ diseases, as discussed in section 1.2.1.1. However, the retinal phenotype of SCA7 (Martin et al. 1994) makes it a useful model for testing of novel therapeutic approaches for polyQ disorders. The eye provides an easily-accessible site to which gene therapies have successfully reached clinical trials for recessive diseases of the retina (Hauswirth et al. 2008). Furthermore, disease

progression can be monitored through the measurement of retinal function by electroretinogram (ERG).

No treatment for SCA7 currently exists and allele-specific therapy has a number of limitations, as discussed previously. SCA7 pathogenesis is complex, apparently involving both gain- and loss-of-function. The functions of the protein in which the polyQ expansion occurs, Ataxin-7, are still emerging, hinting at the potential for undesirable effects of non-allele-specific silencing. SCA7 may therefore be an ideal candidate for gene K&R therapy, and a useful model in which to test this approach.

A number of factors are important for appropriate therapeutic design and assessment of its efficacy, such as the extent to which mutant protein removal is required for therapeutic improvement, whether loss of normal Ataxin-7 function is detrimental, and how early intervention is necessary to halt or reverse disease progression. In the following sections, SCA7 pathogenesis and evidence for the normal function of Ataxin-7 are discussed.

1.5.2 Genetics of SCA7

Normal *ataxin-7* alleles contain a sequence of 7-19 consecutive CAG repeats and SCA7 results from expansions of more than 38 repeats (David et al. 1997; Giunti et al. 1999). Although normal alleles are relatively stable, intermediate repeat numbers are at risk of expanding into the disease range during replication (Giunti et al. 1999). Repeat number therefore often increases in patient offspring and also shows somatic instability within individuals (Gouw et al. 1998). Age of onset of SCA7 inversely correlates with the repeat length (Gouw et al. 1998; Giunti et al. 1999).

1.5.3 Pathology of SCA7

SCA7 is a progressive neurodegenerative disease which presents with a classic autosomal dominant ataxia, with degeneration of the cerebellum, brain stem and spinal cord, similar to the other forms of SCA (Martin et al. 1994). Uniquely, SCA7 also presents with macular degeneration, whereby loss of cone and rod photoreceptors and atrophy of the optic pathways lead to blindness (Martin et al. 1994). Infantile-onset cases (>180 repeats) involve more wide-ranging symptoms, often leading to multi-system failure and death within one to three years (Ansorge et al. 2004; Whitney et al. 2007).

Within the cerebellum, the Purkinje cells, a single layer of neuronal cells with extensive dendritic arborisations, are severely affected, undergoing atrophy and eventual cell death (Koeppen 2005). Indeed, in the cerebellum of patients with long (85 and 106 CAG) repeats, very few Purkinje cells (PCs) remained at death (Holmberg et al. 1998; Mauger et al. 1999).

Nuclear inclusions (NIs) are protein aggregates which are common features among the polyQ diseases. Aggregates containing Ataxin-7 are found in affected tissues (David et al. 1997; Giunti et al. 1999), including the inferior olive (Holmberg et al. 1998) and pons, as well as all nuclear layers of the retina (Mauger et al. 1999). Typically, one or two NIs are present per nucleus, each on average 3 µm in diameter (Holmberg et al. 1998). However, NIs can be difficult to detect in PCs (Holmberg et al. 1998; Mauger et al. 1999).

The cerebellum is crucial for motor skills, involved in coordinating sensory and motor information, and the PCs provide the sole neuronal output from the cerebellar cortex. The inferior olive provides a powerful input: a single climbing fibre from an inferior olivary neuron activates a single PC and is thought to play a role in the precise timing of

coordinated motor events (Wu et al. 2011). The granule cells of the cerebellum also carry messages to the PCs, and receive inputs from mossy fibres from many regions of the brain, including motor-related information from the pons (Huang et al. 2013).

Ataxin-7 is widely expressed throughout the body, in both peripheral and neuronal tissues, but expression is low in PCs (Cancel et al. 2000). For pons and inferior olive, the relatively high levels and predominantly nuclear Ataxin-7 expression found in healthy brain may contribute to their susceptibility in SCA7 (Cancel et al. 2000; Lindenberg et al. 2000). Degeneration of the inferior olive and pons could cause reduced stimulation of the PCs and contribute to their dysfunction.

Before discussing the various pathogenic features of SCA7 in detail and considering prospects for therapy, it is necessary to first introduce some of the most important mouse models of this disease.

1.5.4 Mouse models for SCA7

Although post-mortem studies are essential in determining the pathological details of SCA7, there are only a small number of patients of this rare condition and information is generally limited to the terminal stages of the disease. To determine the earliest changes in pathogenesis, the functions of normal versus mutant Ataxin-7 and to test potential therapies, animal models are vital.

Table 1-4 – Major mouse models for SCA7 which affect the cerebellum. Expan size = size of polyQ tract in model mouse. Con = size of polyQ tract in control mouse. WT indicates that non-transgenic/knock-in mice were used as controls. Tg = transgenic, KI = knock-in, dKI = double knock-in. NR = Not reported. Asterisk (*) indicates that the data refers to the most well-studied mouse line of those created.

Model Name	Model type	Promoter	Expan size	Con	Expression site	Onset age (phenotype)	Abnormal gait	First gene expression changes	First aggregates	First tissue pathology	Lifespan	References
P7E	Tg	Pcp-2	90	WT	Purkinje cells	11 mo (reduced rotarod performance)	NR		8 mo	16 mo	normal	Yvert 2000, 2001
B7E*	Tg	PDGF-B	128	10	All brain regions (variable levels)	5 mo (reduced motor coordination, and activity)		NR	3 mo	NR	~8 mo	Yvert 2001
PrP*	Tg	Mo-PrP	92	24	All CNS except PCs	2 mo (reduced rotarod performance)	3-4 mo (13-15 w)		2 mo	2 mo	~8 mo (30-34 w)	La Spada 2001, Garden 2002, Custer 2006
266Q knock-in	KI	(Mo-Atxn7)	266	5	Whole body	5 w (weight loss, ataxia, tremors, etc.)	5 w	5 w	3 mo	4 mo	4-5 mo	Yoo 2003, Gatchel 2008, Noma 2012
Gfa2*	Tg	Gfa2	92	10	Bergmann glia	8-12 mo (forelimb clasping)	16-20 mo	9 mo	12 mo	10 mo (40 w)	>24 mo	Custer 2006
PrP-floxed-92Q	Tg	BAC-PrP –floxed	92	10	Many neurons - conditional inactivation	5 mo (reduced composite score from multiple tests)	NR	NR	10 mo (40 w)	10 mo (40 w)		Furrer 2011
CMV-PDGF	Tg	CMVe-PDGF-B	52	WT	Many brain regions	7 mo (forelimb clasping)	10/11 mo	6 mo	10/11 mo	11 mo	NR	Chou 2010, Wang 2010
100Q knock-in	dKI	(Mo-Atxn7)	100	5	Whole body	8-9 mo (ataxia)	8-9 mo	NR	NR	8-9 mo	12 mo	Chen 2012

Table 1-5 – Major SCA7 mouse models affecting the retina. Expan size = size of polyQ tract in model mouse. Con = size of polyQ tract in control mouse. NR = Not reported. Asterisk (*) indicates that the data refers to the most well-studied mouse line of those created.

Model Name	Promoter	Expan size	Con	Expression site	First ERG abnormalities	First gene expression changes (genes)	First aggregates	First tissue pathology	Status at death	References
R7E	Rho	90	10	Retina (rod cells)	4 w	3 w (many)	3-4 w	2 w	30% photoreceptor loss (no ERG response after 1 year)	Yvert 2000, Kizilyaprak 2011, Helmlinger, Abou-Sleymane 2004, Abou-Sleymane 2006, Helmlinger 2006, Yefimova 2010
PrP*	moPrP	92	24	All CNS except Purkinje cells	7 w	4 w (BCP, CARR, RHO, GNAT1)	11 w	12 w	(Blind at 13 w)	La Spada 2001, Palhan 2005, Young 2007
266Q knock-in	(Mo-Atx7)	266	5	Whole body	5 w	4 w (BCP, GCP)	8 w	10 w	Too weak for ERG testing beyond 15 w	Yoo 2003
100Q knock-in	(Mo-Atx7)	100	5	Whole body	NR	4 w (S-opsin)	NR	8 w	NR	Chen 2012

The most well-studied mouse models of SCA7 are summarised in Table 1-4 (relating to the cerebellum) and Table 1-5 (relating to the retina), including the key features of each model and at which age various phenotypes become observable. These models will be discussed throughout the remainder of this chapter.

The *PrP* transgenic model represents the SCA7 phenotype well at multiple levels (La Spada et al. 2001). Notably, this model lacks transgene expression specifically in PCs, yet they exhibit abnormal size and morphology (Garden et al. 2002). Conversely, the *P7E* model has expanded mutant Ataxin-7 restricted only to Purkinje cells. This resulted in a relatively mild phenotype (Yvert et al. 2000), further indicating that the mutant Ataxin-7 protein has a minimal direct effect on PCs.

A knock-in model with a very large expansion (266Q) successfully reflected the juvenile-onset disease, with an extensive range of symptoms (Yoo et al. 2003). Whereas high transgene expression levels may contribute to toxicity, this knock-in model expressed Ataxin-7 at endogenous levels. Additionally, the long expansion allows full development of the disease, not limited by the short natural lifespan of a mouse. This mouse can therefore be considered a highly relevant model for SCA7. A comparable knock-in model with 100Q demonstrates a slower disease progression and longer lifespan, as expected (Chen et al. 2012).

In the *R7E* model, expression of expanded *ataxin-7* is driven by the Rhodopsin (*rho*) promoter and therefore limited to the rod photoreceptor cells. This allows disease progression to be studied in the retina without complication from any other phenotypes or shortened lifespan (Yvert et al. 2000). Although cones are affected before rods in SCA7, the majority (97.2%) of mouse photoreceptors are rods (Jeon et al. 1998), making this an easier cell type to study. In *R7E* mice, at onset of visual deficiency (detected by ERG), tissue

pathology and prominent NIs were already visible (Yvert et al. 2000). In *PrP* and 266Q knock-in mice, ERG abnormalities were detected in cone followed by rod function before tissue abnormalities and aggregation, eventually leading to complete loss of vision (La Spada et al. 2001; Yoo et al. 2003).

1.5.5 Nuclear inclusions in SCA7

Nuclear inclusions (NIs) are protein aggregates which are common features among the polyQ diseases. The expanded proteins involved do not share homology beyond their CAG-repeat region (Wellington et al. 1998). However, there is a shared threshold at which a CAG repeat expansion becomes disease-causing, which corresponds to a conformational shift in the polyQ tract, as evidenced by a change in antibody binding affinities (Trottier et al. 1995) and a tendency towards aggregation (Huang et al. 1998). However, it is still not clear what role nuclear inclusions play in pathogenesis. They may be toxic themselves, or form as protective accumulations to limit the toxicity of polyQ-expanded proteins. It is important to determine the role of NIs in SCA7 in order to select effective therapies.

1.5.5.1 Toxicity of nuclear inclusions

A number of factors have been considered to determine whether NIs are toxic to cells. In human patients, NIs were frequently positive for activated Caspase-3, indicating a strong link with cell death (Einum et al. 2001). A possible cause of toxicity of aggregates involves oxidative stress, as implicated in other polyQ diseases (Kim et al. 2003; Browne and Beal 2006). In an inducible rat PC12 cell model for SCA7, the appearance of nuclear aggregates coincided with an increase in reactive oxygen species (ROS), followed by decreased cell viability (Ajayi et al. 2012). It is suggested that aggregates could be sites of oxidative reactions and activation of NADPH oxidase (NOX) was thought likely to be responsible.

Indeed, NOX inhibition or anti-oxidant treatment ameliorated the toxicity. It would be interesting to determine whether anti-oxidant treatment can be beneficial *in vivo*.

Conversely, many lines of evidence show that NIs are not the primary cause of toxicity. First, NI formation often does not precede disease onset in mouse models (Table 1-4, Table 1-5) (Yvert et al. 2000; Garden et al. 2002; Yoo et al. 2003). Second, patients of other polyQ diseases such as HD and DRPLA can also develop NIs in the retina, yet patients rarely exhibit visual dysfunction (Mauger et al. 1999). Third, NIs appear in unaffected cells in SCA7 patients (Ansorge et al. 2004), although they generally form earlier in affected than unaffected regions in the 266Q knock-in mouse (Yoo et al. 2003). Further, an inverse correlation was found between aggregate formation and cellular toxicity in patients (Ansorge et al. 2004), the 266Q mouse (Bowman et al. 2005) and in primary cerebellar granule culture (Taylor et al. 2006), indicating that nuclear inclusions may in fact be protective. Accordingly, signs of aggregation were seen in the *R7E* mouse before onset (Yvert et al. 2000).

These observations together support the possibility that aggregates begin as protective structures and become toxic later in disease progression.

1.5.5.2 Cleavage of mutant Ataxin-7

Protein fragments containing the expanded region are found in numerous polyQ diseases (Wellington et al. 1998). Similarly, in SCA7, N-terminal fragments of Ataxin-7 are commonly detected in patient material and animal models. In patient brain, few aggregates contained the full-length protein (Einum et al. 2001), and in some mouse models, no full-length mutant Ataxin-7 can be detected at all: only truncated fragments, not full-length protein, were incorporated into NIs in *R7E* and *PrP* mice (Yvert et al. 2000; Helmlinger et al. 2004a; Garden et al. 2002). However, protein fragments are not always found, particularly in cell models

(Wang et al. 2006; Einum et al. 2001), though these models may be too simplistic to accurately represent all disease processes.

Caspase-7 was found to cleave expanded Ataxin-7, producing two N-terminal fragments of 45 and 55 kDa at two caspase-7 recognition sites, Asp-266 and Asp-344 (Young et al. 2007). The smaller cleavage product likely corresponded to the N-terminal fragments of 45 kDa found in *PrP* mice, and patient brain and fibroblasts (Garden et al. 2002). Accordingly, activated caspase-7 was recruited into aggregates in *PrP* mouse photoreceptors (Young et al. 2007). Remarkably, in transfected HEK-293 cells, mutation of both caspase-7 cleavage sites completely ameliorated the toxicity (apoptosis and caspase activity) caused by expanded Ataxin-7 and reduced NI formation. It also ablated NI formation in cerebellar granule neuron culture (Young et al. 2007). It seems clear that Ataxin-7 cleavage is a key event in SCA7 pathogenesis and may enhance toxicity of the mutant protein. However, attempting to therapeutically inhibit this cleavage may be difficult without affecting other proteins.

1.5.5.3 Ubiquitination and degradation of Ataxin-7

In the Ubiquitin-proteasome system (UPS), ubiquitination is used to mark damaged or unnecessary proteins for degradation by the proteasome, and this may be involved in the turnover of expanded Ataxin-7.

NIs in SCA patients are often ubiquitinated (Koeppen 2005). In the brain tissue of two SCA7 patients, almost all inclusions were ubiquitinated (Einum et al. 2001; Zander et al. 2001) and this observation was particularly linked to tissues which suffer significant cell death (Holmberg et al. 1998; Ansorge et al. 2004).

Mouse models indicate that ubiquitination is not an early event in pathogenesis. In the 266Q mouse, ubiquitinated NIs were first detected long after symptoms appeared (Yoo et al.

2003). Similarly, in *R7E* mice, NIs formed well before Ubiquitin-proteasome system (UPS) components were recruited (Yvert et al. 2000). This further supports the earlier conclusion that aggregates begin as protective structures and become toxic later in disease progression.

One study proposed that Ataxin-7 is recruited into promyelocytic leukaemia protein (PML) clastosomes, protein complexes which contain UPS components and appear in the nucleus when proteolysis is particularly active (Janer et al. 2006). In cell models, Ataxin-7 (regardless of expansion size) was recruited by PML components into NIs. This indicates that PML factors may recruit expanded Ataxin-7 for degradation, until perhaps an overload is reached, leading to the formation of larger aggregates. This does not entirely support the previous observations of late recruitment of UPS factors and the results here may represent the processing mechanism for overexpressed Ataxin-7. However, PML bodies were often found co-localising with small rather than large NIs in patient cells and *B7E* mice (Takahashi et al. 2002; Janer et al. 2006), strongly indicating that they do play a role at a relatively early stage of disease.

The reason that aggregates build up may be that the UPS becomes unable to degrade them. Through the use of inducible cell models, it was concluded that the UPS is primarily responsible for the clearance of both wild-type and expanded Ataxin-7, but autophagy plays a major role in clearance of fragments of expanded Ataxin-7 and aggregated material (Yu et al. 2012). However, no increase in autophagy was detected in response to expanded Ataxin-7 and autophagy cannot act in the nucleus, where aggregates occur (Yu et al. 2012). The 45 kDa cleavage fragment may be resistant to degradation by autophagy, providing a potential explanation for its toxicity (Mookerjee et al. 2009). Cell death can occur as a consequence of autophagy, involving a pathway distinct from apoptosis (Klionsky and Emr 2000).

1.5.5.4 Dysfunction of heat shock proteins and the UPS

Heat shock proteins (HSPs) aid protein folding, prevent aggregation and protect against various cellular stresses. Dysfunction of this and the UPS pathway have been suggested as contributors to polyQ diseases via sequestration of HSPs and proteasomal factors into NIs, leading to reduced availability. Sequestration of such factors was seen in SCA7 patients (Einum et al. 2001; Zander et al. 2001) and the *PrP* mouse model (Garden et al. 2002). However, sequestration does not necessarily indicate dysfunction. For example, in *R6* mice (a model of HD), ubiquitinated NIs recruited particular HSPs but left their normal distribution unaffected (Helmlinger et al. 2002).

Significantly, at single-cell resolution in the 266Q knock-in mouse, no impairment of either UPS or proteasome function could be detected at any stage of disease in vulnerable cells (Bowman et al. 2005). Consistent with this, over-expressing the chaperone Hsp70 and/or its cofactor Hsp40 did not reduce number or size of NIs, nor improve phenotype in *R7E* mice (Helmlinger et al. 2004b). Although reduced proteasome function was found in a rat primary cerebellar cell model (Wang et al. 2006), this may represent a limitation of cell models (Bowman et al. 2005).

1.5.5.5 Sequestration of normal Ataxin-7 and other proteins into NIs

In addition to the haploinsufficiency of wild-type *ataxin-7* in SCA7 patients and the unknown functional capacity of mutant Ataxin-7, sequestration of the normal protein into NIs could contribute to a loss-of-function effect. Although over-expressed normal Ataxin-7 co-localises with aggregates *in vitro* (Zander et al. 2001), this is difficult to determine in post-mortem tissue due to limitations of antibodies (Mauger et al. 1999). However, in both patients and mouse models, as disease progresses, normal diffuse Ataxin-7 protein staining in the cytoplasm and nucleus tends to be replaced with strong nuclear staining in the form of

aggregates (Yvert et al. 2000; Garden et al. 2002; Yoo et al. 2003; Helmlinger et al. 2004a). Similarly, the distinct perinuclear staining of Ataxin-7-10Q in the *R7N* model (a transgenic mouse comparable to *R7E* but with a polyQ tract within the normal range) was lost when either Ataxin-7-90Q or expanded Huntingtin was co-expressed. These results imply that normal Ataxin-7 may be sequestered by polyQ-containing aggregates. However, in the latter example, Ataxin-7-10Q could not be detected in Huntingtin NIs, indicating that another mechanism may be responsible (Helmlinger et al. 2004a).

The phenotype of the 266Q model was not worsened by knocking out the wild-type Ataxin-7 allele (Yoo et al. 2003), implying that a partial lack of normal function of Ataxin-7 in aggregates does not contribute to disease. On the other hand, sequestration of other polyglutamine-containing proteins, including Ataxin-3 and CREB-binding protein (CBP), is commonly observed in patient brains and may cause toxicity (Einum et al. 2001; Zander et al. 2001). Indeed, incorporation of CREB-binding protein (CBP), a Histone Acetyl Transferase (HAT), into NIs was suggested as a common polyQ disease mechanism (McMahon et al. 2005). However, even if incorporated into aggregates, the cellular distribution of CBP can remain stable (Helmlinger et al. 2002). Additionally, a broad screen in the *PDGF-B* mouse model revealed no clear differences in proteins sequestered into NIs in neuronal populations affected by SCA7 versus those unaffected (Yvert et al. 2001). It is plausible, though, that different neuronal populations vary in their sensitivity to sequestration of the same proteins.

1.5.5.6 Summary

Nuclear protein aggregation is a prominent feature in SCA7. The early formation of aggregates appears to be protective for cells, but eventually leads to toxicity, loss-of-function effects and cell death. Cleavage of mutant Ataxin-7 appears to enhance toxicity of the mutant protein and perhaps drives aggregation. It is not clear whether normal Ataxin-7

is also incorporated into aggregates in SCA7, but its distribution is affected. Sequestration of common factors into NIs may explain some of the common features of the polyQ diseases. However, clinical features can appear prior to the appearance of NIs and loss-of-function effects may contribute to this.

1.5.6 Loss of Ataxin-7 function in SCA7

The polyQ expansion, combined with the altered distribution and haploinsufficiency of normal Ataxin-7, may disrupt its cellular function. This may explain why the susceptibility of specific neuronal and retinal populations contrasts to the widespread expression pattern of Ataxin-7. This can inform the development of appropriate therapeutic approaches.

The functions of Ataxin-7 are still being uncovered. Ataxin-7 possesses a zinc-binding motif but was not predicted to form a known zinc-finger structure (Helmlinger et al. 2004c). Although commonly found in the nucleus, Ataxin-7 contains both nuclear export and import signals (Taylor et al. 2006). No a*taxin-7* knockout mouse has yet been reported, but a recent study in zebrafish (*Danio rerio*) revealed that severe depletion of Ataxin-7 levels in the embryo caused early lethality, whereas partial depletion specifically impaired the differentiation of photoreceptors, PCs and cerebellar granule neurons (Yanicostas et al. 2012).

However, this doesn't necessarily imply Ataxin-7 function in adult neurons, and the effect of the polyQ expansion was not investigated. These factors will be discussed in the following sections, with particular reference to novel therapeutic strategies.

1.5.6.1 Functional binding partners of Ataxin-7

Ataxin-7 has a number of known binding partners (Kahle et al. 2011). The altered conformation of mutant Ataxin-7 may impact these interactions and their functions to cause pathogenesis before the formation of NIs.

Ataxin-7 associates with microtubules (MTs), co-immunoprecipitating with α-tubulin throughout the cell cycle (Nakamura et al. 2012). Over-expression of Ataxin-7 caused over-stabilisation of MTs and knockdown led to slower recovery of MT structure after disassembly. However, no differences were observed between the activities of mutant and normal Ataxin-7, arguing against a role in pathogenesis. However, this supports the argument for maintaining its levels during therapy.

Another binding partner of Ataxin-7 is Amyloid precursor-like protein 2 (APLP2), which, like Ataxin-7 itself, may also contribute to toxicity in SCA7 through cleavage. Caspase-3 cleavage of APLP2 was induced in the presence of expanded Ataxin-7 in the inducible PC12 model, producing cytotoxic fragments (Takahashi-Fujigasaki et al. 2011). The functions of APLP2, which include transcription, may also be perturbed by its cleavage.

Many binding partners of Ataxin-7 are known to have links to macular degeneration (Kahle et al. 2011). The levels of some of these proteins were altered in the 266Q mouse retina, for example transcription factor TRIM23 (Kahle et al. 2011). Interaction with mutant Ataxin-7 also has the potential to affect the function of these proteins, though this remains to be explored.

The functional interactions of Ataxin-7 require further investigation to determine their importance in SCA7. However, a significant role of Ataxin-7 was discovered in certain transcription regulatory complexes, which may explain the initiation of SCA7 pathogenesis.

1.5.6.2 Ataxin-7 and STAGA/TFTC

Chromatin modifications central to gene regulation are made by changing the properties of DNA-bound histones. Specifically, histone acetylation, carried out by histone acetyltransferases (HATs), is required to weaken the chromatin structure for transcriptional initiation. Ataxin-7 interacts with members of HAT-containing transcriptional coactivator/adaptor complexes, namely the TATA-binding protein (TBP)-free TAF-containing complex (TFTC) and the SPT3/TAF9/ADA/GCN5 acetyltransferase complex (STAGA) (Helmlinger et al. 2004c). GCN5 is the mediator of each complex's histone acetyltransferase (HAT) activity. When activated, these complexes acetylate histone H3 to facilitate transcription initiation.

Studies in yeast highlighted the important role of Ataxin-7 in these transcriptional complexes. Absence of the yeast homologue of Ataxin-7 (Sgf73) compromised HAT activity (McMahon et al. 2005). Sgf73 has a complex function in yeast STAGA, regulating numerous factors required for efficient transcription and also mRNA export (Köhler et al. 2008). Additionally, Ataxin-7 is involved in the histone deubiquitination activity of TFTC/STAGA, another process related to chromatin structure (Köhler et al. 2008; Zhao et al. 2008).

1.5.6.3 Dysfunction of Ataxin-7 in STAGA

Through its role as an integral component of STAGA and TFTC transcriptional complexes, Ataxin-7 misfunction in SCA7 may directly affect the transcription of many genes (Helmlinger et al. 2004c).

When yeast Sgf73 knockouts were complemented with wild-type human Ataxin-7, HAT activity of SAGA was fully rescued. With expanded Ataxin-7, however, HAT activity was compromised and SAGA complexes lacked essential components necessary to correctly regulate GCN5 (McMahon et al. 2005). Similarly, in *R7E* retina, STAGA complexes had

reduced levels of several components and severely diminished HAT activity (Palhan et al. 2005). Conversely, in patient-derived cells, expanded Ataxin-7 did not alter the composition of STAGA/TFTC complexes, although their HAT activity was not analysed (Helmlinger et al. 2004c).

Knocking out GCN5 in PCs led to a mild ataxic phenotype (Chen et al. 2012) similar to that seen in *P7E* mice (Yvert et al. 2000), implying that disease mechanisms may indeed be related. Similarly, reducing GCN5 levels in 100Q knock-in mice accelerated disease progression (Chen et al. 2012). However, surprisingly, (retinal) gene dysregulation was not significantly accelerated, which implies GCN5 dysfunction can contribute to the phenotype of SCA7 by another mechanism. Deficiencies in many Ataxin-7-dependent STAGA functions may contribute to disease (Chen et al. 2012).

In a neuronal cell line, over-expression of mutant Ataxin-7 was shown to cause a reduction in transcription mediated by either CBP or retinoid-related orphan receptor α isoform 1 (RORα1), a PC-enriched transcription factor involved in transcription of PC genes. RORα1 haploinsufficient mice display degeneration of both PC and inferior olivary neurons. However, only CBP, not RORα1 was found in aggregates (Ström et al. 2005). Further, over-expression of wild-type Ataxin-7 had the same, or even a greater effect, but this could not be attributed to sequestering of CBP into aggregates. This may in fact indicate that Ataxin-7 can negatively regulate the STAGA complex.

Interestingly, the down-regulation of a STAGA-dependent promoter-reporter construct was largely prevented when cleavage of Ataxin-7 was inhibited by mutation of its caspase-7 recognition sites (Young et al. 2007). It could be, therefore, that a vital segment of Ataxin-7 is absent in the cleaved product and affects its role in STAGA, rather than an effect of the expanded region itself.

1.5.6.4 Transcriptional dysregulation in SCA7

Several studies have investigated the changes to the transcriptional profile in SCA7, which highlight the contribution of STAGA/TFTC dysfunction to pathogenesis. This may also reveal novel therapeutic targets.

Microarray studies in the cerebellum of transgenic *PDGFB* mice after onset revealed that more genes were down-regulated than up-regulated (32 vs 10) (Chou et al. 2010). Many down-regulated genes were involved in neuron/axon maintenance (including myelination), protein translation, folding and deubiquitination, cellular and synaptic signalling.

Gene dysregulation is a common feature across the polyQ diseases and some of the similarities have also been explored. Comparing the large-expansion knock-in mouse models of SCA7 and SCA1 (which also involves PC degeneration) revealed an overlap of 31 dysregulated genes in the cerebellum at early disease stages (Gatchel et al. 2008). The majority of these were down-regulated and their functions generally concerned signalling activities. The most strongly down-regulated gene common to both models was insulin-like growth factor binding protein 5 (IGFBP5), expressed largely in granular neurons of the cerebellum. IGFBP5 can sequester IGF1, substrate for a receptor which is present on PCs (Gatchel et al. 2008). However, the change was also seen in granular neurons when expanded SCA1 was expressed only in PCs, suggesting its dysregulation in granular neurons comes as a downstream effect of polyQ-induced toxicity.

1.5.6.5 Retinal phenotype of SCA7

The retina, having a relatively small array of cell types, is particularly facultative to analysis of specific gene expression changes and their relation to tissue pathology and functional losses.

Despite having no homology to Ataxin-7 apart from the polyQ tract, over-expression of mutant Huntingtin in mice causes retinal degeneration similar to *R7E* mice, though this rarely occurs in HD patients (Helmlinger et al. 2002). This hints that in cases of very large CAG expansions, a general polyQ-based toxicity leads to common cellular stresses, progressive cellular dysfunction and related coping mechanisms. This toxicity can particularly affect the sensitive, highly complex cells of the retina, whereas disruption of retinal-specific functions of Ataxin-7 is likely to trigger retinal disease in SCA7 patients. Indeed, only 113 of the 486 genes found to be dysregulated in *R7E* after onset were also dysregulated in the retina of *R6* (HD) mice (Abou-Sleymane et al. 2006). Further, remarkably little photoreceptor loss (20-30%) was seen in some mouse SCA7 models even at advanced stages of disease, emphasising that visual losses may be largely attributable to cell dysfunction (Yvert et al. 2000; Yoo et al. 2003).

The outer segment of the retina, containing the light-sensitive ends of the photoreceptors, is progressively thinned in SCA7 (Yvert et al. 2000; La Spada et al. 2001; Yoo et al. 2003). Down-regulated genes encoding structural components of the outer segment may explain this degeneration (Yoo et al. 2003). The nuclei of rod cells in these regions of thinning outer segment in *R7E* mice often had decondensed chromatin, even before disease onset (Helmlinger et al. 2006). Structurally, this was attributed to a decreased transcription of key linker histone component H1c (Kizilyaprak et al. 2011). This appears to represent a form of de-differentiation related to a loss of rod-specific gene expression (Helmlinger et al. 2006). Accordingly, most rod cells in this model reverted to a rounded morphology (Yefimova et al. 2010).

Indeed, the expression levels of many photoreceptor-specific genes decreased in *PrP, R7E* and 266Q mice (La Spada et al. 2001; Yoo et al. 2003; Helmlinger et al. 2006). For example,

12 out of 13 rod-specific genes were repressed in *R7E* mice, including Rhodopsin (Helmlinger et al. 2006), a pigment essential for light detection. Accordingly, SCA7 tissue pathology reflects that seen in Rhodopsin knockout mice (Yoo et al. 2003).

However, gene dysregulation is a complex and changing feature. A microarray study in *R7E* retina found 106 genes dysregulated at onset (3 w), 50% of which were no longer altered at a moderate stage of disease (9 w) (Abou-Sleymane et al. 2006). This difference highlights the need for careful consideration of conclusions being made from gene expression changes assessed at a single time-point.

Overall, complex gene dysregulation may account for many of the pathological developments occurring in SCA7. Many genes down-regulated at onset of visual deficiency are involved in signal transduction, cell communication and visual perception, perhaps accounting for the functional abnormalities (Abou-Sleymane et al. 2006). Further genes of similar functions were later repressed, along with genes involved in various aspects of growth, development and cell maintenance, which perhaps account for the onset of tissue pathology.

1.5.6.6 The role of Ataxin-7 in the retina

The specific retinal phenotype of SCA7 may be explained by the role of Ataxin-7 in STAGA. Ataxin-7, as a part of STAGA, appears to interact with the photoreceptor-specific transcription factor cone-rod homeobox protein (CRX) (Palhan et al. 2005). Rhodopsin is one notable CRX-regulated gene. In support of this, Ataxin-7 co-localised and co-immunoprecipitated with CRX *in vitro* (La Spada et al. 2001), and co-occupied promoter/enhancer regions of CRX-regulated genes *in vivo* (Chen et al. 2004).

The expansion mutation in Ataxin-7 may affect this function. Reduced binding of CRX to its targets was observed in pre-symptomatic *PrP* mice (La Spada et al. 2001; Palhan et al. 2005), though not in 266Q mice (Yoo et al. 2003). In *PrP* mice, five CRX-dependent genes exhibited marked reductions in acetylated histone H3 on their promoter/enhancer regions. Expanded Ataxin-7 inhibited CRX-dependent transactivation in a Rhodopsin promoter-reporter assay, although gene dosage contributed to this (La Spada et al. 2001).

However, there were significant differences between *crx* haploinsufficient mice and SCA7 models, and genes which are not targets of CRX were also down-regulated in SCA7 models (Yoo et al. 2003; Helmlinger et al. 2006). This indicates that the other roles of STAGA/TFTC are also affected in the SCA7 retina.

1.5.6.7 Summary

The expansion of Ataxin-7 may alter its interactions with various binding partners. The aforementioned cleavage products and the potential loss-of-function of wild-type Ataxin-7 may also contribute to this. One key role of Ataxin-7 is as a part of STAGA/TFTC transcriptional complexes, which facilitate transcription of a large number of genes through histone acetylation. It appears that expanded Ataxin-7 disrupts the normal function of these complexes and their recruitment to chromatin. This perhaps initiates the SCA7 disease process and may explain many of the transcriptional alterations, particularly given the relatively high frequency of down-regulated genes. Other functions of GCN5, including deacetylation and deubiquitination may also be affected in SCA7. Crucially, via CRX, Ataxin-7 is required for STAGA-mediated transcription of photoreceptor-specific gene expression. Misfunction of expanded Ataxin-7 in this role perhaps initiates the dysfunction of photoreceptors.

Identifying the pathogenic processes involved in SCA7 is important for developing an appropriate therapy. Determining the functions of Ataxin-7 helps to reveal whether it is essential to maintain wild-type Ataxin-7 expression during therapy.

1.5.7 The importance of non-cell-autonomous effects in SCA7

Beyond the cellular effects of the Ataxin-7 expansion, neuron function and survival is dependent upon receiving the correct inputs from surrounding and connected cells (Koeppen 2005). Notably, PCs do not express mutant Ataxin-7 in *PrP* mice, yet were abnormal and subject to cell death (Garden et al. 2002). In *P7E* mice, where mutant Ataxin-7 was instead restricted only to PCs, some post-synaptic neurons showed increased Ubiquitin staining (Yvert et al. 2000). Further, in *R7E* mice, only rod photoreceptors express mutant Ataxin-7, yet other retinal neurons also develop abnormal growth patterns (Yvert et al. 2000) and gene dysregulation (Abou-Sleymane et al. 2006).

There is evidence that Bergmann glia and the inferior olive may impact PC survival in SCA7. Bergmann glia are a type of astrocyte found surrounding PCs and play important roles in PC maintenance. Surviving Bergmann glia were swollen and disorganised in patient cerebella, 266Q, *PrP-BAC* and *Gfa2-Ataxin7-92Q* mice (Custer et al. 2006). *PrP-BAC* mice express expanded Ataxin-7 in all neurons; *Gfa2* mice express it only in Bergmann glia, yet exhibit PC degeneration and an ataxic phenotype. However, surprisingly, specifically excising the transgene from Bergmann glia of *PrP-BAC* mice did not rescue degeneration of either PCs or Bergmann glia, although disease onset was delayed slightly (Furrer et al. 2011). A similar effect was seen when the transgene was instead excised from both PCs and inferior olive, whereas excision from all three tissues delayed onset more significantly (Furrer et al. 2011). These findings strongly indicate that pathology in Bergmann glia, PCs and inferior olive all contribute to SCA7.

There are several pathways potentially affected in SCA7 which may contribute to this effect. For example, the inferior olive normally transmits neurotrophic factor Insulin-like growth factor-1 (IGF-1) to PCs (Nieto-Bona et al. 1993). Reelin, a factor involved in PC maintenance, was found to be repressed in a SCA7 astrocyte cell model (McCullough et al. 2012). Bergmann glia from both *Gfa2* and *PrP-BAC* mice were deficient in glutamate uptake, explained by a reduction in expression of the GLAST glutamate transporter (Furrer et al. 2011). Resultant sustained glutamate stimulation was thought to be responsible for non-apoptotic cell death observed in PCs in both models (Furrer et al. 2011).

Non-apoptotic cell death may also be involved in *R7E* retinal pathology. Although most rod cells appeared to de-differentiate and undergo apoptosis, a subset instead migrated through the retina and appeared to undergo non-apoptotic cell death, perhaps as a result of the loss of their normal input (Yefimova et al. 2010).

The complex interdependent relationship between neurons and non-neuronal cells in the brain and retina means that cell dysfunction caused by expanded Ataxin-7 in one cell population can have a knock-on effect on other populations. This may explain the sometimes confusing relationship of aggregation with cell death. Furthermore, revealing which cell populations contribute to disease highlights which brain regions and cell types require corrective treatment.

1.5.8 Summary

This section discussed the major mechanisms of pathogenesis in SCA7. Regarding the development of an effective gene therapy for SCA7, it appears to be important to treat a wide range of neuronal and even non-neuronal cells. Although PCs are severely affected in SCA7, mutant Ataxin-7 may have only a minimal direct effect on PCs and their demise may

instead be largely a downstream effect caused by the dysfunction of connected cells. Transcriptional dysregulation is often the first detectable change in the course of the disease in mouse models and may be the trigger for the chain of pathogenic events. It appears that dysfunction of expanded Ataxin-7 in its role as part of GCN5-containing transcriptional complexes may initiate these changes. Early intervention may be vital for maximising the therapeutic potential of any novel treatments.

1.6 Potential strategies for treatment of SCA7

There is no treatment for SCA7, only physiotherapy to help patients to manage symptoms. The accumulated evidence regarding the effects of expansions in Ataxin-7 and the function of the normal protein are vital for making the best choice regarding the development of novel therapeutics. A number of potential therapeutic agents or targets for treatment have been revealed in SCA7.

1.6.1 Regulators of Ataxin-7 protein turnover

Factors which affect the production or turnover of Ataxin-7 may provide possible targets for therapy.

Increased degradation of expanded but not WT Ataxin-7 was achieved through β-interferon (β-INF) treatment in a cell model (Janer et al. 2006). β-INF is already an approved drug, for the treatment of multiple sclerosis. However, given its limited efficacy when added after expanded Ataxin-7 expression, it may only be effective as a very early treatment, and its efficacy has yet to be shown in an animal model.

Ataxin-7 appears to be a target of an endogenous brain-enriched miRNA, miR-124 (Tan et al, in preparation). However, therapeutic enhancement of miR-124 or other endogenous miRNAs may not be an ideal therapeutic strategy since other targets would be also affected.

1.6.2 Mediators of downstream effects

A number of other factors may be able to reduce some of the downstream effects of the Ataxin-7 expansion.

Levels of a glutamate transporter, GLAST, were increased by over-expressing hepatocyte growth factor (HGF) in 266Q mice (Noma et al. 2012). The degeneration of PCs and Bergmann glia was reduced and behavioural phenotype was improved. Drug-based stimulation of such glutamate transporters is possible (Custer et al. 2006).

Similarly, the cytokine granulocyte–macrophage colony-stimulating factor (GM-CSF) protected photoreceptors against cell death in rats (Schallenberg et al. 2012) and a number of general neuroprotective factors, such as ciliary neurotrophic factor (CNTF) and glial cell derived neurotrophic factor (GDNF), can provide therapeutic benefit against neurodegeneration (Lipinski et al. 2011; Sari 2011). Delivering CNTF increased visual acuity in patients with retinal degeneration (Sieving et al. 2006) and prolonged survival of photoreceptors in various animal models, including Rhodopsin knockout mice (Liang et al. 2001). However, some studies have shown that CNTF and GDNF treatment has negative effects (Wen et al. 2006; Rossmiller et al. 2012).

In a SCA7 cell model, anti-oxidant treatment or inhibition of NOX reduced ROS levels and aggregation, resulting in restoration of cell viability (Ajayi et al. 2012). It is suggested that increased ROS levels are involved in numerous polyQ diseases. It would be interesting to see if anti-oxidant treatment can alleviate disease in SCA7 mouse models.

1.6.3 Stem cell therapy

Recent advances in stem cell technology allow the reprogramming of patient fibroblasts into induced pluripotent stem (iPS) cells through the delivery of key reprogramming genes (Takahashi et al. 2007). iPS cells may be differentiated into neuronal cultures through the provision of appropriate small molecule stimuli and culture conditions (Dimos et al. 2008; Ebert et al. 2009). This was recently attained for SCA7 (Luo et al. 2012; Watson 2012). As well as being useful in the study of SCA7 and testing of treatments, iPS cells have potential in the treatment of neurodegenerative disease, can be derived from a simple skin biopsy following diagnosis (even in elderly patients), corrected by gene therapy *ex vivo* and re-implanted into the patient, avoiding any immune responses. The retina is particularly amenable to this technique (Comyn et al. 2010), but many challenges yet remain in the development of such therapies (Feng and Gao 2012).

1.6.4 Silencing of mutant Ataxin-7

Mutation-directed silencing therapy for SCA7 is potentially achievable through targeting of the CAG expansion, which has a number of limitations, as discussed in section 1.4.2.2. A common SNP has been linked to the SCA7 mutation, where 50% of South African SCA7 patients are heterozygous for a single base change, having the 'A' and 'G' alleles associated with the mutant and wild-type transcripts respectively (Greenberg et al. 2006). Specifically targeting this SNP in an allele-specific knockdown therapy is achievable *in vitro* (Scholefield et al. 2009), but would treat only a subset of SCA7 patients. Particularly for such a rare disease, a more widely applicable therapy is desirable, such as non-allele specific silencing. However, the functions of Ataxin-7 are still emerging, hinting at the potential for undesirable effects to be caused by such indiscriminate silencing. SCA7 is therefore an ideal candidate for

K&R therapy. However, it is first important to determine that silencing of mutant Ataxin-7 is achievable *in vivo* and effective in reducing pathogenesis of the polyQ expansion.

There is little reported data regarding silencing of Ataxin-7 beyond cell models. However, in one mouse model, an early suppression of the exogenous mutant Ataxin-7 gene was revealed, which may be considered equivalent to a targeted silencing therapy. In *R7E* mice, the Rhodopsin-promoter-driven expression of mutant Ataxin-7 was reduced very early in disease progression (but endogenous mouse Ataxin-7 persisted) (Helmlinger et al. 2004a). The mRNA became barely detectable after a period of only 3-4 weeks of expression. However, degeneration continued, with loss of photoreceptors, retinal structure and vision accountable to loss of Rhodopsin expression, which was never rescued. Although aggregates were slowly but significantly reduced, at two years a significant proportion of rods still bore aggregates. In contrast, in a HD model mouse, the clearing process was rapid following induced shutdown of the mutant (truncated) Huntingtin transgene (Yamamoto et al. 2000); just three weeks after switching off the gene, no aggregates were detected, and the behavioural phenotype was reversed. This may highlight the limitations of making comparisons between polyQ diseases regarding treatment strategies. The persistence of aggregates indicates that Ataxin-7 NIs are difficult for cells to process and it appears that the transcriptional shutdown of Rhodopsin expression may be irreversible once triggered. These findings support the need for administering a novel therapy before onset (Helmlinger et al. 2004a). However, the high levels of transgene expression in this model may contribute to the apparent irreversibility. The *R7E* mouse has an expansion size of 90 glutamines, yet onset of disease was detected earlier than in the 266Q model (Table 1-5), indicating that high expression levels may play a substantial role. Crucially, vision was not lost completely until one year of age, 11 months after first ERG abnormalities, whereas in the *PrP* (92Q)

mouse, this process took little more than a month (Table 1-5). This indicates that transcriptional silencing of mutant Ataxin-7 in the *R7E* mouse in fact induced a substantial slowing of the expected disease progression.

1.7 Overview of project

Gene K&R therapies potentially offer an exciting widely-applicable corrective treatment for SCA7 and other polyQ diseases. Artificial mimics of naturally occurring intronic miRNAs and mirtrons are theoretically ideal for this purpose. This project will investigate the development of mimics of two different intronic miRNA systems (mirtron miR-1224 and the miR-106b cluster) to silence Ataxin-7 and to be incorporated into novel single K&R constructs for testing *in vitro*.

This will consist of three main aims:

- Achieving effective post-transcriptional gene silencing of Ataxin-7 by mimicking mirtrons and miRNA clusters.
- Silent modification of the miRNA target sites in an Ataxin-7 cDNA sequence to confer RNAi resistance, and incorporation of the most promising miRNA into this sequence for co-delivery in transfected cell models.
- Investigation of the efficacy of Ataxin-7 silencing and replacement in SCA7 patient-derived cell models and comparison of any phenotypic improvements.

2 Materials and Methods

2.1 Mirtron design

An algorithm was developed by Dr Yiqi Seow (A*STAR, Singapore) (Seow et al. 2012) to identify a set of potential mirtron target sites within a given gene or cDNA sequence. Sites are selected which can allow antisense molecules to fulfil splicing requirements when incorporated as either the 5' strand or 3' strand of a mirtron hairpin (Figure 2-1a). Each target is then scored by the algorithm according to eight key criteria to indicate their potential as targets of RNA interference (RNAi) (Reynolds et al. 2004). This was later updated with criteria from the siRNA Information Resource (sIR) (Shah et al. 2007). This algorithm was used to select putative target sites within the human Ataxin-7 cDNA (Zander et al. 2001). Following previous results in the lab, antisense strands were each incorporated into a stem-loop design based upon the endogenous mouse mirtron miR-1224 (Berezikov et al. 2007) (Figure 2-1b), after removal of the "bridge" regions (Figure 2-1c) which was previously found to enhance silencing (Sibley et al. 2012). Further details and the resulting mirtron designs are given in the next chapter (section 3.2).

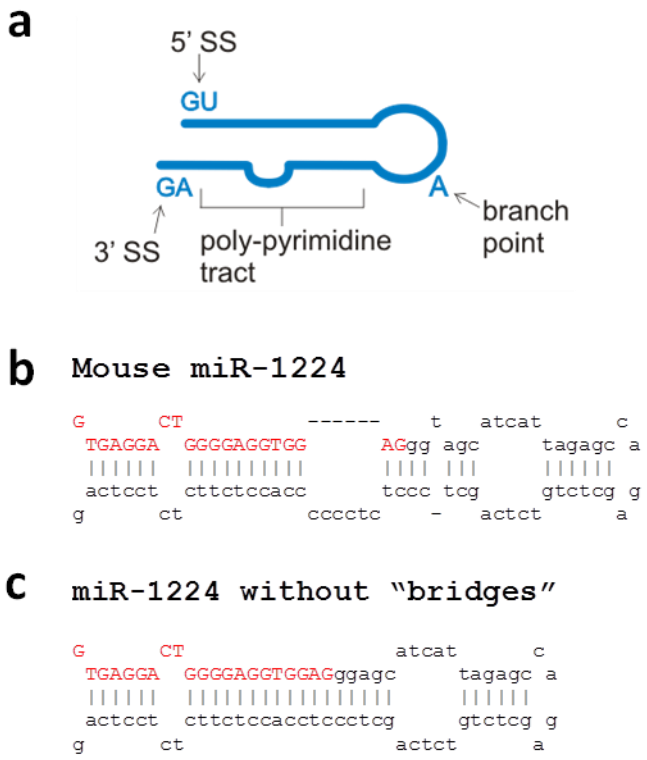


Figure 2-1 - Mirtron design. (a) Sequence requirements for mirtrons. The guide strand may be incorporated into either end of the mirtron provided that these requirements are fulfilled. The bulge represents the general feature of mismatches found between guide and passenger strands of miRNA. An A residue is the key element of the branch point. The 5' and 3' splice sites (SS) are strictly defined. The poly-pyrimidine tract must contain minimal purine interruptions. **(b)** The sequence and hairpin structure of endogenous mouse miR-1224. The guide strand is indicated in red capitals. **(c)** The sequence of miR-1224 with its "bridge" regions removed to improve silencing efficiency. Artificial mirtrons targeting Ataxin-7 were incorporated into this hairpin.

2.2 Cloning

Primers were purchased from Sigma-Aldrich (UK) or Integrated DNA Technologies (IDT; USA) and are listed in Table 2-2 and Table 2-3. PCR reactions were carried out using Taq DNA polymerase from New England Biolabs (NEB; USA) unless otherwise stated. Restriction enzymes were from NEB. Plasmids digestion products were gel-purified (unless otherwise stated) (Qiagen, USA) and PCR products were all subject to PCR cleanup (Qiagen) according to the manufacturer's instructions. Amplified plasmids were isolated and purified from

bacteria using miniprep, midiprep or maxiprep plasmid extraction kits (Qiagen) according to the manufacturer's instructions. Before use in Neon transfection (sections 2.3.3 and 2.3.4.1), plasmids were instead purified using a PureLink HiPure maxiprep kit (Life Technologies, USA). All plasmids made were sequence verified (Source Bioscience, UK).

2.2.1 Mirtron expression plasmids

To create mirtrons and their variants, long primers (Table 2-2) were annealed by slow cooling (from 95°C to 4°C) in T4 DNA ligase buffer (NEB). A Cytomegalovirus (CMV)-driven enhanced green fluorescent protein (EGFP) expression construct containing a short intronic region (pEGFP-mirt, Figure 2-2a, derived from pEGFP-C1 by Y. Seow), was digested with BbsI. Following PCR cleanup of the backbone, the annealed mirtrons (created with the appropriate 5' overhangs) were ligated in as introns within EGFP (Figure 2-2b). Hence, successful splicing of each mirtron from its transcript produces a functional EGFP transcript and therefore a measurable fluorescent signal. As a control, an identical EGFP vector containing a non-hairpin-forming intron, NADH-Coenzyme Q reductase intron 6, which is of similar length to mirtrons (91-nt) was used (henceforth referred to as "NAD"; courtesy of Y. Seow).

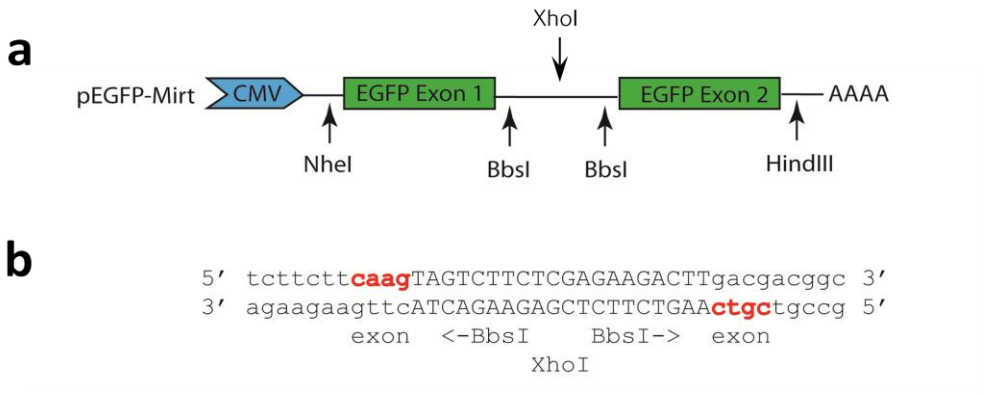


Figure 2-2 - pEGFP-mirt. **(a)** Schematic representation of the EGFP cassette of pEGFP-mirt, showing the CMV promoter, two exons of EGFP, polyA tail and restriction sites. **(b)** Sequence of the mirtron insertion site and the surrounding EGFP exon sequences. Both forward and reverse strands are shown. The exonic regions are labelled and shown in lowercase. The short intronic region is shown in uppercase and its restriction sites are indicated. BbsI digestion removes the intronic region along with the exonic bases shown in red, leaving overhanging ends to which the annealed mirtron sequences are complementary. Both figures courtesy of Y. Seow.

2.2.2 Multiple-mirtron expression plasmids

A second putative splice site was chosen in EGFP (Y. Seow; Figure 2-3a). To create EGFP constructs containing two mirtrons/introns, mirt-18, mirt-1cN and the NAD intron were each inserted into this site by PCR within the pEGFP-mirt backbone *before* insertion of a mirtron into the original site. Long primers were designed, complementary to the new splice site in EGFP at their 3′ end, with 5′ tails containing more than half of the mirtron/NAD intron to be inserted (Table 2-2). This allowed a 3-step overlapping PCR process to be used to insert each mirtron/intron into the second site (Figure 2-3 b-d). A mirtron/intron was then inserted into the original site in each construct through BbsI digestion (section 2.2.1) (Figure 2-3e), creating EFGP constructs containing two mirtrons/introns in all desired combinations (NAD-NAD, 18-18, 1cN-1cN, 1cN-18 and 18-1cN).

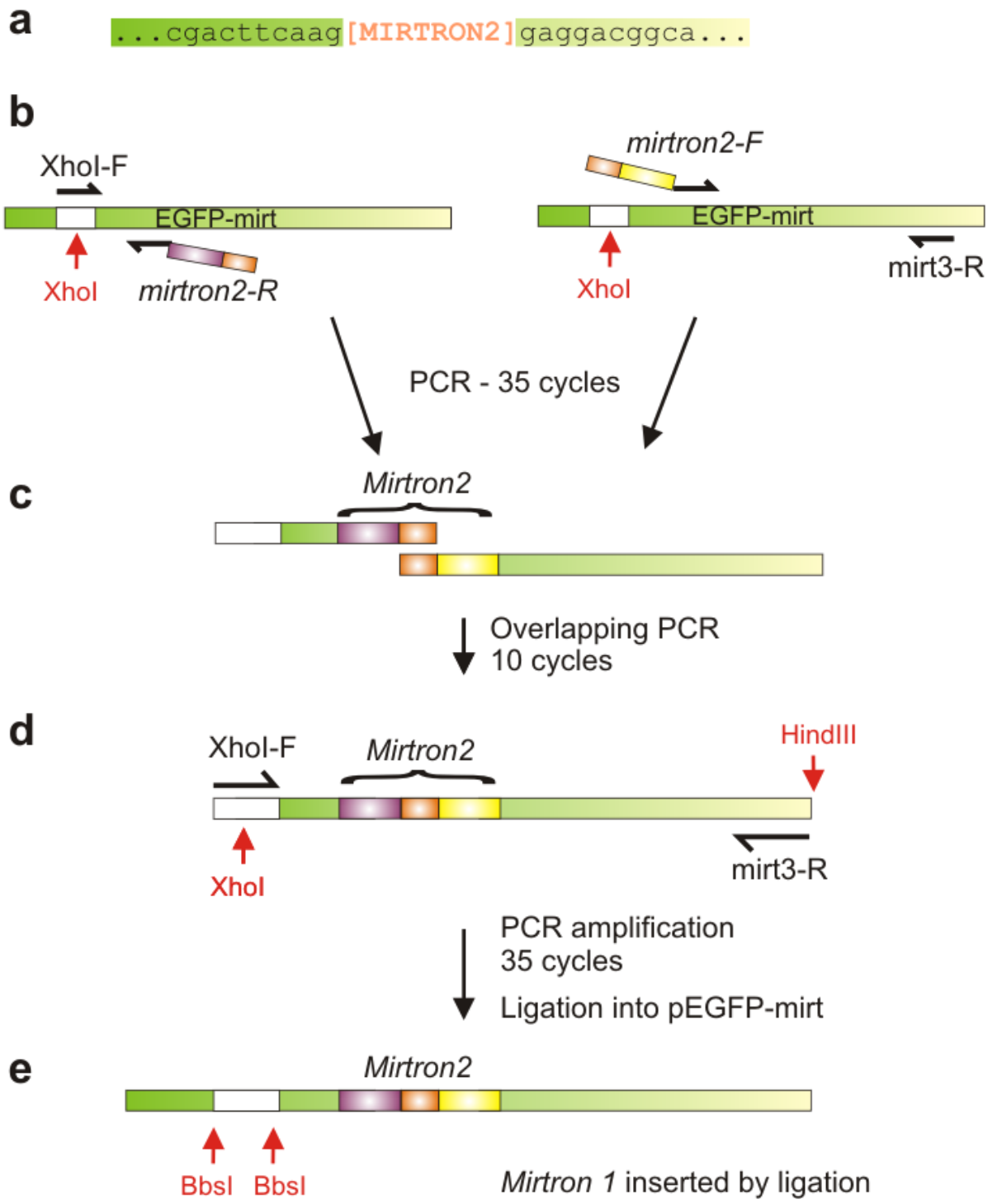


Figure 2-3 - Cloning multiple mirtrons into pEGFP. (a) Sequence of the additional site in EGFP used for insertion of mirtrons. **(b-e)** Schematic diagrams detailing the cloning procedure. The white box represents the short intronic region containing BbsI sites for mirtron insertion; red arrows indicate restriction sites; the orange boxes indicate the overlapping central section of the mirtron/intron being inserted. **(b)** For each mirtron/intron, long primers were designed, complementary to the desired insertion site in EGFP at their 3' end (black half-arrows), with 5' tails containing more than half of the mirtron or NAD intron to be inserted (coloured boxes). For each mirtron/intron, two separate PCR reactions

were carried out using EGFP-mirt as template: (1) the long forward primer with reverse primer mirt3-R and (2) the long reverse primer with XhoI-F (primer sequences are listed in Table 2-2). **(c)** The two products were combined into a 10-cycle PCR reaction, allowing the overlapping sections of the mirtron/intron (orange boxes) to act as primers against each other to create the full-length mirtron/intron-containing sequence. **(d)** This full-length product was amplified by PCR using primers XhoI-F and mirt3-R, then digested and ligated back into pEGFP-mirt between XhoI and HindIII sites. **(e)** BbsI sites were then used to insert a second mirtron into each construct.

2.2.3 Short hairpin RNAs, short interfering RNAs and LNAs

Short hairpin RNAs (shRNAs) based upon mirtron sequences were made by annealing long primers (listed in Table 2-2) as in section 2.2.1, followed by insertion into a U6 expression vector (Dr Thomas Roberts, Scripps Research Institute, CA, USA) cut with BbsI. Transcription from U6 is terminated after UU in a UUUU motif; therefore each shRNA was designed to be exactly the same size as its corresponding mirtron by exchanging the final two nucleotides of the 3' end with TT. This was not possible for mirt-1c since it contained a string of T (U) residues towards the 3' end which would cause premature termination. To circumvent this problem, an additional single base was changed in the mirtron passenger strand, for which an equivalent shRNA version could be made (mirt-1cN, shR-1cN - Table 2-2).

A short interfering RNA (siRNA), siR-P16 (Scholefield et al. 2009) was designed to silence human Ataxin-7, matching a SNP commonly associated with SCA7 (Greenberg et al. 2006) at position 16 of the guide strand (Table 2-1, highlighted). siR-P16 was supplied by Novartis and used as a positive control for Ataxin-7 silencing in human cell lines. At the concentrations used in this thesis (>20 nM), siR-P16 exhibited strong non-allele-specific silencing of endogenous *ataxin-7* (Scholefield et al., in preparation). siRNAs targeting the mouse Ataxin-7 sequence (used as positive controls for silencing in mouse cell lines, Table 2-1) and a non-

targeting fluorescent (Cy3-labelled) siRNA (used to assess transfection efficiency) were all designed and supplied by Novartis. A non-targeting control siRNA, siR-NS, was designed and supplied by Eurogentec. shR-P16 (a U6-driven shRNA based upon siR-P16) and shR-1 (a shRNA designed to target Ataxin-7 away from the SNP site) were also used as positive controls for Ataxin-7 silencing. shR-P16 and shR-1 plasmids were provided by Dr Janine Scholefield. The non-specific control shRNA, shR-NS, targeted an irrelevant gene (α-synuclein) and was supplied by Dr Christopher Sibley (University of Cambridge, UK).

Single-stranded oligonucleotides designed by (Dr Miguel Varela, University of Oxford) to target Ataxin-7 for steric blocking, containing locked nucleic acids (LNAs), were supplied by Eurogentec (Belgium). Sequences were as follows: LNA-REP - TgcTgcTgcTgcTgcT, LNA-5 - TgcTgcTgcTgcCggGcgG, LNA-NS (non-targeting control) - TgcTatAccAgcGtcGtcA; where LNA nucleotides are in uppercase.

Table 2-1 – Sequences of siRNAs and shRNAs targeting Ataxin-7. All sequences are given in the 5'-3' direction. The SNP-targeting base in siR- and shR-P16 is highlighted in yellow. Names ending in "m" target the mouse Ataxin-7 sequence. The loop section common to the shRNA sequences is indicated between hyphens.

siR-P16	GCCATGAACAATGTTCATA
shR-P16	GCCATGAACAATGTTCATA-CCTGACCCA-TGTGGACATTGTTCATGGCTT
shR-1	GCTGAAGACAATTCTAATA-CCTGACCCA-TATTGGAATTGTCTTCAGGTT
siR-4m	TTAAACTTAATATCCAGGC
siR-10m	TTTCAGTAGTGTGCCATCC
siR-11m	TAATAACACATCAAATCGT
siR-12m	ATAACCGAGAAGCTGGCTC

2.2.4 Dual Luciferase targets

Target sequences were combined into a single string for mirtrons 1-6 (Tar1) and 9-18 (Tar2). All others contained only a single target site. Sequences were either ordered in full or as overlapping primers (Table 2-3). These were annealed (or amplified by PCR and digested if

necessary) and ligated into a psiCheck 2.2 vector between XhoI and NotI restriction sites downstream of *Renilla* Luciferase (Ely et al. 2008).

2.2.5 miRNA cluster mimics

miRNA cluster mimics targeting Ataxin-7 were designed based upon the miR-106b cluster (Aagaard et al. 2008). Primers were designed based upon mirt-18, mirt-1cN or shR-P16 with sequences modified to mimic the hairpin structure of each endogenous miRNA of the cluster (Table 2-2) and with ends compatible with restriction sites for each hairpin position. Primers were annealed as in section 2.2.1 and ligated into pMCM7-S1-S2-S3 (gift from Lars Aagaard) (Figure 2-4). The pMCM7-S1-S2-S3 hairpins targeted HIV and this plasmid was used as a non-specific control. For each mimic of miR-106b (position 3), the restriction sites were initially incorporated incorrectly and rectified by PCR using primers XSFix-F and XSFix-R (Table 2-2), followed by digestion.

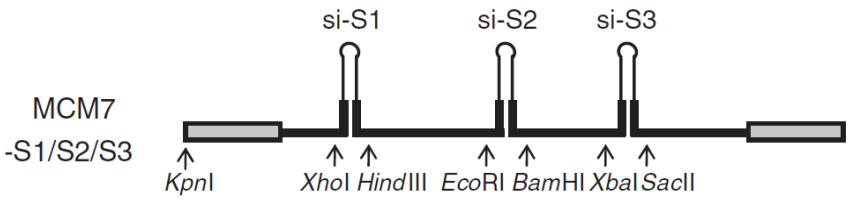


Figure 2-4 - pMCM7-S1-S2-S3. Schematic diagram indicating the relative position of hairpins and splice sites within the intron of MCM7. Around 100 bp of each exon was included in the construct (grey boxes). Image taken from (Aagaard et al. 2008).

2.2.6 Full-length Ataxin-7 targets

Ataxin-7 cDNA plasmids were supplied by J. Scholefield, containing all translated exons of Ataxin-7, driven by a CMV promoter and fused to a DsRED or EGFP tag at the N-terminus (Zander et al. 2001). The Ataxin-7 sequence contained either 10 CAG repeats (coding for 10

glutamines, 10Q) or an expansion mutation (100Q). Both wild-type and mutant constructs were also modified to create mCherry-tagged constructs: a mCherry tag was amplified by PCR (using primers mCherry-F and mCherry-R; Table 2-3) and then digested and ligated in place of the EGFP tag, between AgeI and HpaI sites.

2.2.7 Modification of Ataxin-7 and addition of introns

2.2.7.1 Sequence modification

The cDNA sequence of Ataxin-7 was modified for RNAi resistance by codon replacement (Figure 2-5). These modified constructs were produced by designing tailed primers (Table 2-3; Figure 2-6a). Overlapping PCR (as Figure 2-3) was carried out using Accuzyme DNA polymerase (Bioline, UK), using the Ataxin-7-Q10-mCherry plasmid as template (Figure 2-6b). Problematic PCRs were carried out using AccuPrime GC-rich DNA polymerase (Life Technologies). Modified fragments were inserted into the Ataxin-7-Q10-mCherry plasmid between NheI/BspEI or BspEI/HIndIII restriction sites as appropriate. The 18M2 construct was commercially synthesised (see section 2.2.7.3).

mirt-18 resistance

Target site:	CAG	GGT	AGA	AGA	AAA	CGA	TTT
Translation:	*Gln*	*Gly*	*Arg*	*Arg*	*Lys*	*Arg*	*Phe*
New sequence:	CAG	GGA	CGG	CGG	AAG	CGC	TTC

Gln	Gly	Arg (x2)	Lys	Arg	Phe
caa: 0.27	ggt: 0.16	cgt: 0.08	aaa: 0.43	cgt: 0.08	ttt: 0.46
cag: 0.73	ggc: 0.34	cgc: 0.18	aag: 0.57	cgc: 0.18	ttc: 0.54
	gga: 0.25	cga: 0.11		cga: 0.11	
	ggg: 0.25	cgg: 0.20		cgg: 0.20	
		aga: 0.21		aga: 0.21	
		agg: 0.21		agg: 0.21	

mirt-1cN resistance

Target site:	ACA	TCc	CCA	GTC	CTG	CTC	TCA	TCT
Translation:	*Thr*	*Ser*	*Pro*	*Val*	*Leu*	*Leu*	*Ser*	*Ser*
New sequence:	ACT	TCC	CCT	GTT	CTG	TTG	AGT	TCT

Thr	Ser	Pro	Val	Leu	Leu	Ser	Ser
act: 0.25	tct: 0.19	cct: 0.29	gtt: 0.18	tta: 0.08	tta: 0.08	tct: 0.19	tct: 0.19
acc: 0.36	tcc: 0.22	ccc: 0.32	gtc: 0.24	ttg: 0.13	ttg: 0.13	tcc: 0.22	tcc: 0.22
aca: 0.28	tca: 0.15	cca: 0.28	gta: 0.12	ctt: 0.13	ctt: 0.13	tca: 0.15	tca: 0.15
acg: 0.11	tcg: 0.05	ccg: 0.11	gtg: 0.46	ctc: 0.20	ctc: 0.20	tcg: 0.05	tcg: 0.05
	agt: 0.15			cta: 0.07	cta: 0.07	agt: 0.15	agt: 0.15
	agc: 0.24			ctg: 0.40	ctg: 0.40	agc: 0.24	agc: 0.24

Figure 2-5 - Ataxin-7 cDNA was modified by codon replacement to be resistant to silencing by mirtrons. The sequence of the target site for each mirtron is shown along with the translated amino acid sequence. Codon substitutions were made to prevent silencing by mirtrons while maintaining the amino acid sequence. Each base change is highlighted in green. Bases outside the target site but contained within the codons are shown in grey and the bases corresponding to the seed region of the mature miRNA are underlined. For each amino acid in the target site, all possible codons are listed along with their usage levels in humans according to the Codon Usage Database (NCBI-GenBank Flat File Release 160.0 [June 15 2007], via http://www.kazusa.or.jp/codon/). Codons present in the WT Ataxin-7 sequence are highlighted in yellow, the alternative codons chosen are highlighted in green.

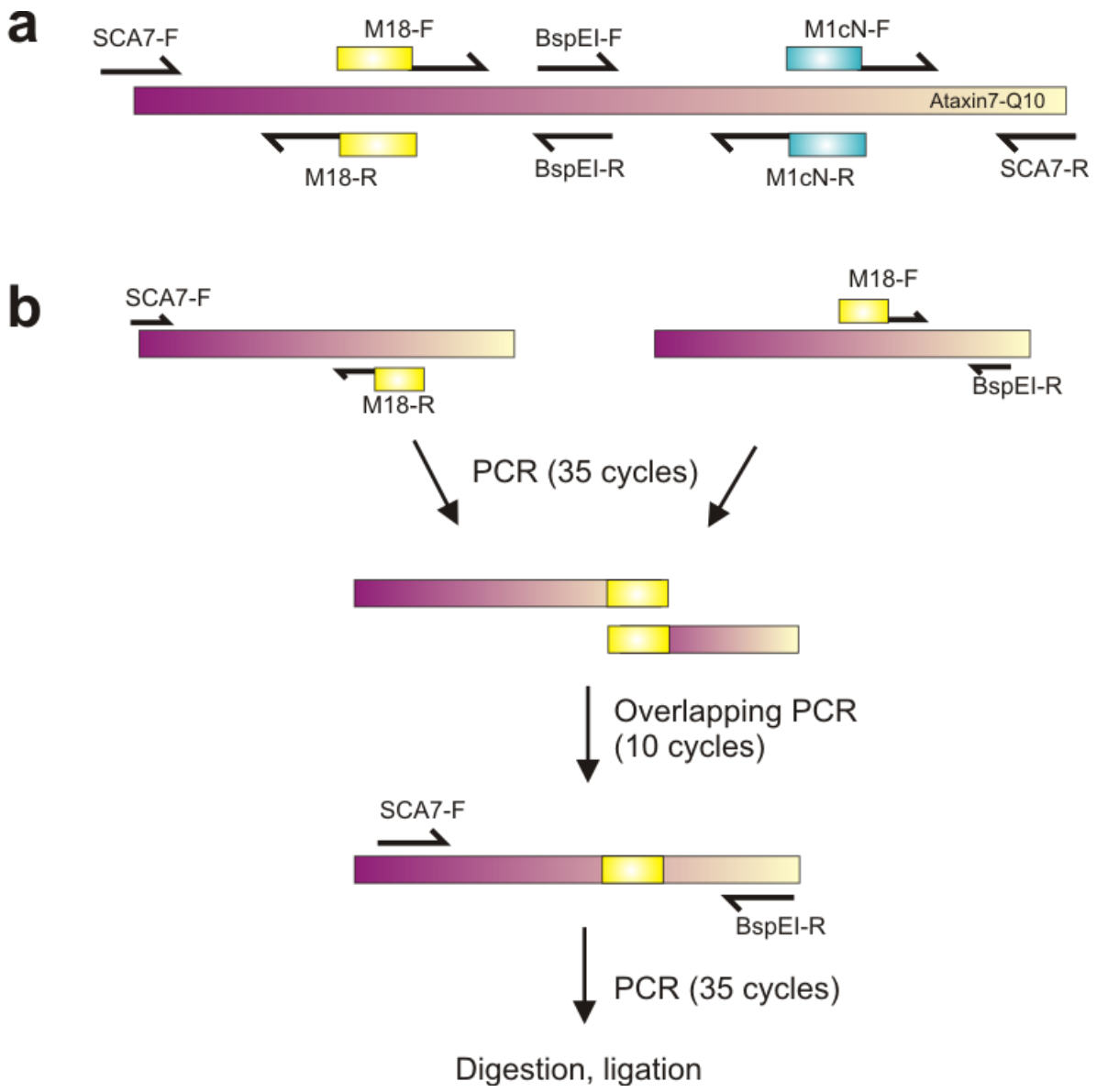


Figure 2-6 – Method used for creation of RNAi-resistant Ataxin-7. Mirtron target sites in Ataxin-7 were modified by an overlapping PCR method. **(a)** Relative positions of each primer. The coloured regions of M18 and M1cN primers indicate a sequence approximately 20 bases in length containing the base modifications for alteration of target sites. **(b)** Cloning procedure, using M18 as an example.

2.2.7.2 Ataxin-7-EGFP-mirt plasmids

To create Ataxin-7 plasmids with mirtrons contained within the EGFP tag, EGFP cassettes EGFP-NAD, EGFP-mirt-18 and EGFP-mirt-1cN were digested out and used to replace the

mCherry tag (between BsrGI and AgeI restriction sites) in pAtaxin-7-Q10-mCherry, pAtaxin-7-18M2-mCherry and pAtaxin-7-M1cN-mCherry respectively.

2.2.7.3 Cloning introns into Ataxin-7

NAD introns and were inserted into sites 1-5 in pAtaxin-7-Q10-mCherry as in section 2.2.7.1, using long overlapping primers listed in Table 2-3. When inserting mirtrons, the plasmid used as the PCR template or for ligation contained the modification for resistance to mirt-1cN or mirt-18 as appropriate, so that the resulting plasmid did not possess the unmodified target site for the mirtron contained within it.

In order to place mirtrons in sites 8-10, which lay between the NheI and BspEI sites of the Ataxin-7 cDNA (Figure 2-7), components of the desired sequences were produced commercially. To achieve this, the section of Ataxin-7 between NheI and BspEI was divided into three overlapping fragments (A, B and C; Figure 2-7). Each segment incorporated mirtrons or sequence modifications as detailed in Table 2-4. Variants of segment A (which included the GC-rich section of Ataxin-7) were produced as minigenes by Biomatik (Canada) within plasmid backbone pGH. Blunt-end restriction sites (AfeI and HpaI) were incorporated at the ends of segment A such that when digested from pGH, the ends perfectly overlapped >20 bp of the adjacent segments of Ataxin-7 with no extra bases (Table 2-4). Following digestion, these segments were gel-purified. Segments B and C were produced by IDT as double-stranded, blunt-ended DNA (G-blocks, Table 2-4). In the desired combinations (Table 2-5), using a Gibson assembly kit (NEB), 20 ng of each G-block/digested minigene was annealed along with the appropriate gel-purified backbone (Ataxin-7-Q10-mCherry or Ataxin-7-Q10-M1cN-mCherry digested with NheI and BspEI). Each mixture was then transformed into 5-α competent *Escherichia coli* (High Efficiency) (NEB) and colonies were screened for the correct plasmid sequences.

Table 2-2 – Primers used for cloning of mirtrons, shRNAs, and miRNA cluster mimics. miRNA cluster mimics are named according to the miRNA being mimicked: miR-106b, miR-25 or miR-93.

Primers for Mirtrons	
Mirt1-F	caaggtgagagcaggactggggatgaaagcatcattagagccagagctctgtctcagctttcatcctctgttcttctttttag
Mirt1-R	cgtcctaaaaagaagaacagaggatgaaagctgagacagagctctggctctaatgatgctttcatccccagtcctgctctcac
Mirt2-F	caaggtgggtgaggtggtactgggcaaagcatcattagagccagagctctgtctcagctttgcccagttcctcttcactttag
Mirt2-R	cgtcctaaagtgaagaggaactgggcaaagctgagacagagctctggctctaatgatgctttgcccagtaccacctcacccac
Mirt3-F	caaggtgaagaggacacttgtctggaaagcatcattagagccagagctctgtctcagctttccatccatgtctccttttttag
Mirt3-R	cgtcctaaaaaggagacatggatggaaagctgagacagagctctggctctaatgatgctttccagacaagtgtcctcttcac
Mirt4-F	caaggtgagaagaagaggaggaggaagcatcattagagccagagctctgtctcagcttcctcctcctctttttttttag
Mirt4-R	cgtcctaaaaaaaaagaggaggaggaagctgagacagagctctggctctaatgatgcttcctcctcctcttcttctcac
Mirt5-F	caaggtggaaaagaaacctgtaggtaaagcatcattagagccagagctctgtctcagctttgtcttcaggtttcttttcctag
Mirt5-R	cgtcctaggaaaagaaacctgaagacaaagctgagacagagctctggctctaatgatgctttacctacaggtttcttttccac
Mirt6-F	caaggtaaacaaaaccaggcaagagggagcatcattagagccagagctctgtctcagcttccttttttcctggttttgtttag
Mirt6-R	cgtcctaaacaaaaccaggaaaaaggaagctgagacagagctctggctctaatgatgctccctcttgcctggttttgtttac
Mirt1a-F	caaggtgagagcaggactggggatgaaagcatcattagagccagagctctgtctcagctttcattctcacttttttccctttag
Mirt1a-R	cgtcctaaagggaaaaagtgagaatgaaagctgagacagagctctggctctaatgatgctttcatccccagtcctgctctcac
Mirt1b-F	caaggtgagagcaggactggggatgatagcatcattagagccagagctctgtctcagcttcactcccctcctctgttcttcttttttag
Mirt1b-R	cgtcctaaaaagaagaacagaggaggggagtgaagctgagacagagctctggctctaatgatgctatcatccccagtcctgctctcac
Mirt1c-F	caaggtgagagcaggactggggatgaagcatcattaagccagagcttgtctcagcttcatcctctgttcttcttttag
Mirt1c-R	cgtcctaaaagaagaacagaggatgaagctgagacaagctctggcttaatgatgcttcatccccagtcctgctctcac
Mirt1d-F	caaggtgagagcaggactggggatgaaagcatcattgaggctggagcctcgtctcagctttcatcctctgttcttcttttttag
Mirt1d-R	cgtcctaaaaagaagaacagaggatgaaagctgagacgaggctccagcctcaatgatgctttcatccccagtcctgctctcac
Mirt1cN-F	caaggtgagagcaggactggggatgaagcatcattaagccagagcttgtctcagcttcatcctctgttcttctttcag
Mirt1cN-R	cgtcctgaaagaagaacagaggatgaagctgagacaagctctggcttaatgatgcttcatccccagtcctgctctcac
Mirt9-F	caaggtgaggatgaagaggacacttaagcatcattaagccagagcttgtctcagcttaattgtcctctttctcttcag
Mirt9-R	cgtcctgaagagaaagaggacaattaagctgagacaagctctggcttaatgatgcttaagtgtcctcttcatcctcac
Mirt12-F	caaggtggaatgagaagaagaggagaagcatcattaagccagagcttgtctcagcttctcctcttcttctcttttcag
Mirt12-R	cgtcctgaaaagagaagaagaggagaagctgagacaagctctggcttaatgatgcttctcctcttcttctcattccac
Mirt13-F	caaggtgggaaggacgggatagagtaaagcatcattaagccagagcttgtctcagcttaattctgtcccgtccttcccag
Mirt13-R	cgtcctgggaaggacgggacagaattaagctgagacaagctctggcttaatgatgctttactctatcccgtccttcccac
Mirt14-F	caaggtagaggaaagaaaaggaggaaagcatcattaagccagagcttgtctcagcttactttcttttcttttcctctag
Mirt14-R	cgtcctagaggaaagaaaagaaagtaagctgagacaagctctggcttaatgatgctttcctccttttcttttcctctac
Mirt15-F	caaggtagggaaaaggaattggtaaagcatcattaagccagagcttgtctcagcttaatcaattccttttccctag
Mirt15-R	cgtcctagggaaaaggaattgattaagctgagacaagctctggcttaatgatgctttaccaattccttttccctac
Mirt18-F	caaggtgggtagaagaaaatggttaaagcatcattaagccagagcttgtctcagcttaaatcgttttcttctacccag
Mirt18-R	cgtcctgggtagaagaaaacgatttaagctgagacaagctctggcttaatgatgctttaaccattttcttctacccac
Multiple mirtrons in EGFP	
mirt2-1cN-F	tcattaagccagagcttgtctcagcttcatcctctgttcttctttcaggaggacggcaacatcctg
mirt2-1cN-R	gacaagctctggcttaatgatgcttcatccccagtcctgctctcaccttgaagtcgatgcccttc
mirt2-18-F	tcattaagccagagcttgtctcagcttaaatcgttttcttctacccaggaggacggcaacatcctg
mirt2-18-R	agacaagctctggcttaatgatgctttaaccattttcttctacccaccttgaagtcgatgcccttc
mirt2-NAD-F	tctcagattttaattttattagcagcatgagattgactctttcataatctacttaaggaggacggcaacatcctg
mirt2-NAD-R	tgctaataaaattaaaatctgagacaaccaaaaagattgacagaattctactacataccttgaagtcgatgcccttc
XhoI-F	Tagtcttctcgagaagacttgacgacggcaactacaagacc
mirt3-R	Atcgaagcttacttgtacagctcgtccat
Mirtron Variants, Unspliceable (US) and shRNA equivalents	
mirt-18v2-F	caaggtgggtagatggaaagggtaaaagcatcattaagccagagcttgtctcagcttaaatcgttttcttctacccag
mirt-18v2-R	cgtcctgggtagaagaaaacgatttaagctgagacaagctctggcttaatgatgcttttaccctttccatctacccac
mirt-18v3-F	caaggtgggttgaaggaataggttaaagcatcattaagccagagcttgtctcagcttaaatcgttttcttctacccag
mirt-18v3-R	cgtcctgggtagaagaaaacgatttaagctgagacaagctctggcttaatgatgctttaacctattccttcaacccac
mirt-1cN-US-F	caagatgagagcaggactggggatgaagcatcattaagccagagcttgtctcagcttcatcctctgttcttctttcag
mirt-1cN-US-R	cgtcctgaaagaagaacagaggatgaagctgagacaagctctggcttaatgatgcttcatccccagtcctgctctcat
mirt-18-US-F	caagatgggtagaagaaaatggttaaagcatcattaagccagagcttgtctcagcttaaatcgttttcttctacccag
mirt-18-US-R	cgtcctgggtagaagaaaacgatttaagctgagacaagctctggcttaatgatgctttaaccattttcttctacccat
shR-18-F	caccgtgggtagaagaaaatggttaaagcatcattaagccagagcttgtctcagcttaaatcgttttcttctaccct
shR-18-R	aaaaagggtagaagaaaacgatttaagctgagacaagctctggcttaatgatgctttaaccattttcttctacccac
shR-18-US-F	caccatgggtagaagaaaatggttaaagcatcattaagccagagcttgtctcagcttaaatcgttttcttctaccct
shR-18-US-R	aaaaagggtagaagaaaacgatttaagctgagacaagctctggcttaatgatgctttaaccattttcttctacccat
shR-1cN-F	caccgtgagagcaggactggggatgaagcatcattaagccagagcttgtctcagcttcatcctctgttcttctttct
shR-1cN-R	aaaaagaaagaagaacagaggatgaagctgagacaagctctggcttaatgatgcttcatccccagtcctgctctcac

shR-1cN-US-F	caccatgagagcaggactggggatgaagcatcattaagccagagcttgtctcagcttcatcctctgttcttctttct
shR-1cN-US-R	aaaaagaaagaagaacagaggatgaagctgagacaagctctggcttaatgatgcttcatccccagtcctgctctcat
Primers for construction of miRNA cluster mimics	
18-25-F	ctagaactggccagtgtttagagggcagaagttatcgattttctggacgctgccctgggaaatcgttttcttctaccctggaagtgccg gcccaacacc
18-25-R	gcggtgttgggccggcacttccagggtagaagaaaacgatttcccagggcagcgtccagaaaatcgataacttctgccctctaaacact ggccagtt
1cN-25-F	ctagaactggccagtgtttagctttgcagtcacggtctcattctggacgctgccctgggatgagagcaggactggggatgtaagtgccg gcccaacacc
1cN-25-R	gcggtgttgggccggcacttacatccccagtcctgctctcatcccagggcagcgtccagaatgagaccgtgactgcaaagctaaacact ggccagtt
shR1-25-F	ctagaactggccagtgttgggacttaaagactttgccaattactggacgctgccctgggtattggaattgtcttcaggtttcagtgccg gcccaacacc
shR1-25-R	gcggtgttgggccggcactgaaacctgaagacaattccaatacccagggcagcgtccagtaattggcaaagtctttaagtcccaacact ggccagtt
18-106b-F	tcgagccctgccggggcaaatcgttttcttctaccctggatggtcctctccgtgtccccgtagaagtgagcgatgcagctccagcaggg cacgcaa
18-106b-R	agctttgcgtgccctgctggagctgcatcgctcacttctacggggacacggagaggaccatccagggtagaagaaaacgatttgccccg gcagggc
1cN-106b-F	tcgagccctgccggggcatgagagcaggactggggatgtatggtcctctccgtgtaccacccagtcgtgttctcgctgctccagcaggg cacgcaa
1cN-106b-R	agctttgcgtgccctgctggagcagcgagaacacgactgggtggtacacggagaggaccatacatccccagtcctgctctcatgccccg gcagggc
shR1-106b-F	tcgagccctgccggggctattggaattgtcttcaggtttctggtcctctccgtggaacatgaagactgtttcttgctgctccagcaggg cacgcaa
shR1-106b-R	agctttgcgtgccctgctggagcagcaagaaacagtcttcatgttccacggagaggaccagaaacctgaagacaattccaatagccccg gcagggc
18-93-F	aattcagtcctgggggctcaaatcgttttcttctaccctggatgtgattacccaactccgggtgaagataaacgatcccgagcccccgg
18-93-R	gatcccgggggctcgggatcgtttatcttcacccggagttgggtaatcacatccagggtagaagaaaacgatttgagcccccaggactg
1cN-93-F	aattcagtcctgggggctcatgagagcaggactggggatgtatgtgattacccaactactcccggttcttgctctccccgagcccccgg
1cN-93-R	gatcccgggggctcggggagagcaagaaccgggagtagttgggtaatcacatacatccccagtcctgctctcatgagcccccaggactg
shR1-93-F	aattcagtcctgggggctctattggaattgtcttcaggtttctgtgattacccaacgaacctgggatatattccttcccgagcccccgg
shR1-93-R	gatcccgggggctcgggaaggaatatatcccaggttcgttgggtaatcacagaaacctgaagacaattccaatagagcccccaggactg
XSfix-F	tagtctttctagaactggccagtgtt
XSfix-R	tagtcttccgcggtgttgggccgg

Table 2-3 – Primers used for cloning of Luciferase targets and Ataxin-7

Primers for construction of Luciferase targets	
Tar1-F	tcgaccagacaagtgtcctcttcatccccagtcctgctctcatcctcctcctcttcttctcataaacaaaaccagggaaaagga
Tar1-R	ggcctccttttccctggttttgtttatgagaagaagaggaggaggatgagagcaggactggggatgaagaggacacttgtctgg
Tar2-F	tacggtactcgagaagtgtcctcttcatcctcataggaaagaaaagaaagtggaatgagaagaagaggagcagggtagaagaaaac
Tar2-R	atgagtcgcggccgcaatcaattccttttccctggtaattctgtcccgtccttcccaaatcgttttcttctaccctgctcctcttctt
Tar3-F	Tcgaggaaacctgaagacaattccaata
Tar3-R	Ggccgctattggaattgtcttcaggtttc
Tar-psngr18-F	Tcgagtaaccattttcttctacccac
Tar-psngr18-R	Ggccgcgtgggtagaagaaaatggtta
Tar-psngr1cN-F	tcgactgaaagaagaacagaggatg
Tar-psngr1cN-R	Ggcccatcctctgttcttctttcag
Tar-M1cN-F	tcgacttcccctgttctgttgagtt
Tar-M1cN-R	Ggccaactcaacagaacaggggaag
Tar-M18-F	Tcgagcagggacggcggaagcgcttc
Tar-M18-R	Ggccgcgaagcgcttccgccgtccctg
Primers for manipulation of Ataxin-7 and insertion of introns/mirtrons	
mCherry-F	Ccgggatccaccggtcgccaccatggtgagcaagggcga
mCherry-R	tgcaataaacaagttaacaacaacaattgcattcatgtctacttgtacagctcgtcc
M18-F	Ggacggcggaagcgcttcgatgtgttattagccgagcacaaa
M18-R	Gaagcgcttccgccgtccctggacagccctgcgct
M1cN-F	Ttcccctgttctgttgagttctacctgcatctccccaaat
M1cN-R	Actcaacagaacaggggaagtagagacggtggctgctga
BspEI-F	tcgccatccggactctcag
BspEI-R	Ctgagagtccggatggcga
SCA7-F	Cagatccgctagcttaatacgac
SCA7-R	aattcgaagcttgggacgtg
NAD-site1-F	tctcagattttaattttattagcagcatgagattgactctttcataatctacttaaggcctccaggctgccctg
NAD-site1-R	tgctaataaaattaaaatctgagacaaccaaaaagattgacagaattctactacataccttggaagactgggggtgtg
NAD-site2-F	tctcagattttaattttattagcagcatgagattgactctttcataatctacttaagttttgcacatttgggagcc
NAD-site2-R	tgctaataaaattaaaatctgagacaaccaaaaagattgacagaattctactacatacagatgctggctgaggatgatg
1cN-site1-F	tcattaagccagagcttgtctcagcttcatcctctgttcttctttcaggcctccaggctgccctg
1cN-site1-R	gacaagctctggcttaatgatgcttcatccccagtcctgctctcaccttggaagactgggggtgtg
18-site2-F	tcattaagccagagcttgtctcagcttaaatcgttttcttctacccagttttgcacatttgggagcc
18-site2-R	agacaagctctggcttaatgatgctttaaccattttcttctacccacagatgctggctgaggatgatg
NAD-site3-F	tctcagattttaattttattagcagcatgagattgactctttcataatctacttaaggaaaatcccaccagtgccc
NAD-site3-R	tgctaataaaattaaaatctgagacaaccaaaaagattgacagaattctactacatacttccatagctgtgcattcagatg
NAD-site4-F	tctcagattttaattttattagcagcatgagattgactctttcataatctacttaaggcgaaggcgatgacaaagaa
NAD-site4-R	tgctaataaaattaaaatctgagacaaccaaaaagattgacagaattctactacataccctcctcactggataaccgaga
NAD-site5-F	tctcagattttaattttattagcagcatgagattgactctttcataatctacttaaggagtcttctggtaacagcactaactg
NAD-site5-R	tgctaataaaattaaaatctgagacaaccaaaaagattgacagaattctactacataccttgggcctcaaagatttgc
NAD-UTR4-F	tctcagattttaattttattagcagcatgagattgactctttcataatctacttaaggacgacgactcactatagggcatgtatcc
NAD-UTR4-R	tgctaataaaattaaaatctgagacaaccaaaaagattgacagaattctactacataccttgaagacgtattaagctagcggggatc

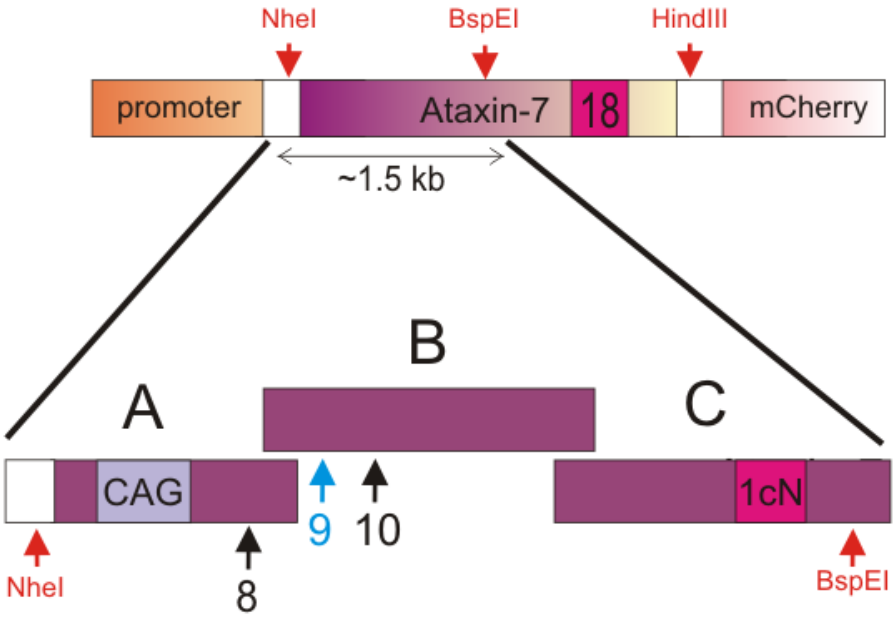


Figure 2-7 - Ataxin-7 cDNA was divided into segments for commercial production. Schematic representation of the Ataxin-7 cDNA construct and its key restriction sites (indicated in red). The section of Ataxin-7 between NheI and BspEI was divided into three 400-500 bp fragments (A, B and C). Each end overlapped the adjacent fragment (or the backbone plasmid) by more than 20 bp. The exon-exon boundaries into which mirtrons were inserted are indicated (8-10). "CAG" represents the CAG-repeat and GC-rich region. Pink boxes labelled "18" and "1cN" represent the mirtron target sites.

Table 2-4 - Segments of Ataxin-7 produced commercially. Segment A and its variants were produced as minigenes, containing either no intron (A), mirt-18 or mirt-1cN in site 8 (A18-8, A1cN-8), or a splice-weakening mutation of site 8 (A8w, base changes highlighted in grey). Segments B and C and their variants were produced as linear double-stranded DNA "G-blocks". Restriction sites are highlighted (green=AfeI; yellow=HpaI). Bases removed by digestion are shown in lowercase. Introns are underlined. Other deviations from the wild-type sequence are highlighted in grey: C18M2 contained an enhanced modification of the mirt-18 target site (18M2); both C and C18M2 also contained silent mutations which removed a short self-complementary region to allow these segments to be produced as G-blocks.

Segment Name	Sequence
A	agcGCTGGTTTAGTGAACCGTCAGATCCGCTAGCTTAATACGACTCACTATAGGGCATGTATCCATATGATGTTCCAGATTATGCTATGTCGGAGCGGGCCGCGGATG ACGTCAGGGGGGAGCCGCGCCGCGCGGCGGCGGCGGCGGCGGGCGGAGCAGCGGCCGCGGCCGCCCGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCGCCGCCGCCTCCGC AGCCCCAGCGGCAGCAGCACCCGCCACCGCCGCCACGGCGCACACGGCCGGAGGACGGCGGGCCCGGCGCCGCCTCCACCTCGGCCGCCGCAATGGCGACGGTCGGGG AGCGCAGGCCTCTGCCCAGTCCTGAAGTGATGCTGGGACAGTCGTGGAATCTGTGGGTTGAGGCTTCCAAACTTCCTGGGAAGGACGGGACAGAATTGGACGAAAGTT aac
A18-8	agcGCTGGTTTAGTGAACCGTCAGATCCGCTAGCTTAATACGACTCACTATAGGGCATGTATCCATATGATGTTCCAGATTATGCTATGTCGGAGCGGGCCGCGGATG ACGTCAGGGGGGAGCCGCGCCGCGCGGCGGCGGCGGCGGCGGGCGGAGCAGCGGCCGCGGCCGCCCGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCGCCGCCGCCTCCGC AGCCCCAGCGGCAGCAGCACCCGCCACCGCCGCCACGGCGCACACGGCCGGAGGACGGCGGGCCCGGCGCCGCCTCCACCTCGGCCGCCGCAATGGCGACGGTCGGGG AGCGCAGGCCTCTGCCCAGTCCTGAAGTGATGCTGGGACAGTCGTGGAATCTGTGGGTTGAGGCTTCCAAACTTCCTGGGAAGGACGGTGGGTAGAAGAAAATGGTTA AAGCATCATTAAGCCAGAGCTTGTCTCAGCTTAAATCGTTTTCTTCTACCCAGGGACAGAATTGGACGAAAGTTaac
A1cN-8	agcGCTGGTTTAGTGAACCGTCAGATCCGCTAGCTTAATACGACTCACTATAGGGCATGTATCCATATGATGTTCCAGATTATGCTATGTCGGAGCGGGCCGCGGATG ACGTCAGGGGGGAGCCGCGCCGCGCGGCGGCGGCGGCGGCGGGCGGAGCAGCGGCCGCGGCCGCCCGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCGCCGCCGCCTCCGC AGCCCCAGCGGCAGCAGCACCCGCCACCGCCGCCACGGCGCACACGGCCGGAGGACGGCGGGCCCGGCGCCGCCTCCACCTCGGCCGCCGCAATGGCGACGGTCGGGG AGCGCAGGCCTCTGCCCAGTCCTGAAGTGATGCTGGGACAGTCGTGGAATCTGTGGGTTGAGGCTTCCAAACTTCCTGGGAAGGACGGTGAGAGCAGGACTGGGGATG AAGCATCATTAAGCCAGAGCTTGTCTCAGCTTCATCCTCTGTTCTTCTTTCAGGGACAGAATTGGACGAAAGTTaac
A8w	agcGCTGGTTTAGTGAACCGTCAGATCCGCTAGCTTAATACGACTCACTATAGGGCATGTATCCATATGATGTTCCAGATTATGCTATGTCGGAGCGGGCCGCGGATG ACGTCAGGGGGGAGCCGCGCCGCGCGGCGGCGGCGGCGGCGGGCGGAGCAGCGGCCGCGGCCGCCCGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCGCCGCCGCCTCCGC AGCCCCAGCGGCAGCAGCACCCGCCACCGCCGCCACGGCGCACACGGCCGGAGGACGGCGGGCCCGGCGCCGCCTCCACCTCGGCCGCCGCAATGGCGACGGTCGGGG AGCGCAGGCCTCTGCCCAGTCCTGAAGTGATGCTGGGACAGTCGTGGAATCTGTGGGTTGAGGCTTCCAAACTTCCTGGGAAGGATTGGAACAGAATTGGACGAAAGTT aac
B	GGACAGAATTGGACGAAAGTTTCAAGGAGTTTGGGAAAAACCGCGAAGTCATGGGGCTCTGTCGGGAAGACATGCCAATATTTGGTTTCTGTCCAGCCCATGATGATT TCTACTTGGTGGTGTGTAACGACTGTAATCAGGTTGTCAAACCGCAGGCATTTCAATCACATTATGAAAGAAGACATAGCTCATCCAGCAAGCCGCCTTTGGCCGTTC CTCCCACTTCAGTATTTTCCTTCTTCCCTTCTCTGTCCAAAAGCAAAGGAGGCAGTGCAAGTGGAAGCAACCGTTCTTCCAGTGGAGGTGTTCTTAGCGCATCCTCAT CAAGTTCCAAGTTGTTGAAATCACCCAAAGAGAAACTGCAGCTCAGGGGGAACACCAGGCCAATGCATCCCATTCAGCAAAGTAGAGTTCCCCATGGTAGAAT
B18-9	GGACAGAATTGGACGAAAGTTTCAAGGTGGGTAGAAGAAAATGGTTAAAGCATCATTAAGCCAGAGCTTGTCTCAGCTTAAATCGTTTTCTTCTACCCAGGAGTTTGG GAAAAACCGCGAAGTCATGGGGCTCTGTCGGGAAGACATGCCAATATTTGGTTTCTGTCCAGCCCATGATGATTTCTACTTGGTGGTGTGTAACGACTGTAATCAGGT TGTCAAACCGCAGGCATTTCAATCACATTATGAAAGAAGACATAGCTCATCCAGCAAGCCGCCTTTGGCCGTTCCTCCCACTTCAGTATTTTCCTTCTTCCCTTCTCT GTCCAAAAGCAAAGGAGGCAGTGCAAGTGGAAGCAACCGTTCTTCCAGTGGAGGTGTTCTTAGCGCATCCTCATCAAGTTCCAAGTTGTTGAAATCACCCAAAGAGAA ACTGCAGCTCAGGGGGAACACCAGGCCAATGCATCCCATTCAGCAAAGTAGAGTTCCCCATGGTAGAA
B1cN-9	GGACAGAATTGGACGAAAGTTTCAAGGTGAGAGCAGGACTGGGGATGAAGCATCATTAAGCCAGAGCTTGTCTCAGCTTCATCCTCTGTTCTTCTTTCAGGAGTTTGG GAAAAACCGCGAAGTCATGGGGCTCTGTCGGGAAGACATGCCAATATTTGGTTTCTGTCCAGCCCATGATGATTTCTACTTGGTGGTGTGTAACGACTGTAATCAGGT TGTCAAACCGCAGGCATTTCAATCACATTATGAAAGAAGACATAGCTCATCCAGCAAGCCGCCTTTGGCCGTTCCTCCCACTTCAGTATTTTCCTTCTTCCCTTCTCT GTCCAAAAGCAAAGGAGGCAGTGCAAGTGGAAGCAACCGTTCTTCCAGTGGAGGTGTTCTTAGCGCATCCTCATCAAGTTCCAAGTTGTTGAAATCACCCAAAGAGAA ACTGCAGCTCAGGGGGAACACCAGGCCAATGCATCCCATTCAGCAAAGTAGAGTTCCCCATGGTAGAA
B1cN-10	GGACAGAATTGGACGAAAGTTTCAAGGAGTTTGGGAAAAACCGCGAAGTCATGGGGCTCTGTCGGGAAGGTGAGAGCAGGACTGGGGATGAAGCATCATTAAGCCAGA GCTTGTCTCAGCTTCATCCTCTGTTCTTCTTTCAGACATGCCAATATTTGGTTTCTGTCCAGCCCATGATGATTTCTACTTGGTGGTGTGTAACGACTGTAATCAGGT TGTCAAACCGCAGGCATTTCAATCACATTATGAAAGAAGACATAGCTCATCCAGCAAGCCGCCTTTGGCCGTTCCTCCCACTTCAGTATTTTCCTTCTTCCCTTCTCT GTCCAAAAGCAAAGGAGGCAGTGCAAGTGGAAGCAACCGTTCTTCCAGTGGAGGTGTTCTTAGCGCATCCTCATCAAGTTCCAAGTTGTTGAAATCACCCAAAGAGAA ACTGCAGCTCAGGGGGAACACCAGGCCAATGCATCCCATTCAGCAAAGTAGAGTTCCCCATGGTAGAA
C	AAGTAGAGTTCCCCATGGTAGAATCATGACACCCTCTGTGAAAGTGGAAAAGATTCATCCGAAAATGGATGGCACACTACTGAAATCTGCGGTGGGGCCAACCTGTCC TGCTACTGTGAGTTCCTTAGTCAAGCCTGGCCTTAACTGCCCCTCAATACCAAAGCCAACCTTGCCTTCACCTGGACAGATTCTGAATGGCAAAGGGCTTCCTGCACC GCCCACTCTGGAAAAGAAACCTGAAGACAATTCCAATAATAGGAAATTTTTAAATAAGAGATTATCAGAAAGAGAGTTTGATCCTGACATCCACTGTGGGGTTATTGA TCTCGACACCAAGAAACCTTGCACCCGGTCTTTGACATGCAAGACACATTCCTTAACCCAGCGCAGGGCTGTCCAGGGTAGAAGAAAACGATTTGATGTGTTATTAGC CGAGCACAAAAACAAAACCAGGGAAAAGGAATTGATTCGCCATCCGGACTCTCAGCAACCACCGCAGC
C18M2	AAGTAGAGTTCCCCATGGTAGAATCATGACACCCTCTGTGAAAGTGGAAAAGATTCATCCGAAAATGGATGGCACACTACTGAAATCTGCGGTGGGGCCAACCTGTCC TGCTACTGTGAGTTCCTTAGTCAAGCCTGGCCTTAACTGCCCCTCAATACCAAAGCCAACCTTGCCTTCACCTGGACAGATTCTGAATGGCAAAGGGCTTCCTGCACC GCCCACTCTGGAAAAGAAACCTGAAGACAATTCCAATAATAGGAAATTTTTAAATAAGAGATTATCAGAAAGAGAGTTTGATCCTGACATCCACTGTGGGGTTATTGA TCTCGACACCAAGAAACCTTGCACCCGGTCTTTGACATGCAAGACACATTCCTTAACCCAGCGCAGGGCTGTCCAGGGACGGCGCAAGAGTTCGATGTGTTATTAGC CGAGCACAAAAACAAAACCAGGGAAAAGGAATTGATTCGCCATCCGGACTCTCAGCAACCACCGCAGC

Table 2-5 – Combinations of commercially-produced segments of Ataxin-7 (A, B, C) and backbone plasmids used to make Ataxin-7 plasmids with modifications and/or introns. Backbones were digested with NheI and BspEI.

Plasmid	Backbone	A	B	C
18M2	WT	A	B	C18M2
18-site8	WT	A18-8	B	C18M2
1cN-site8	M1cN	A1cN-8	B	C
1cN-site9	M1cN	A	B1cN-9	C
1cN-site10	M1cN	A	B1cN-10	C
1cN-10-8w	M1cN	A8w	B1cN-10	C

2.2.8 Lentivirus cloning

Commercially available backbone plasmid pLKO.1 was used to insert cDNA into lentiviruses, which contained a Cytomegalovirus (CMV)-driven EGFP gene. For expression of shRNAs, a U6-promotor-driven shRNA cassette was added, containing either shR-NS or shR-p16 (plasmids both supplied by J. Scholefield). For mirtrons, the EGFP gene in pLKO.1 was replaced with EGFP-NAD-NAD or EGFP-1cN-18 using AgeI and BsrGI restriction sites.

2.3 Cell culture and transduction

2.3.1 Cells and culture conditions

HEK-293, HEK-293 T, Neuro2A (N2A), SH-SY5Y and primary adult dermal fibroblast cells were cultured in GIBCO Dulbecco's Modified Eagle Medium (DMEM)-Glutamax (Life Technologies) supplemented with 10% Foetal Calf Serum (FCS; Life Technologies) and 1% Gibco Antibiotic-Antimicotic (Life Technologies), except where otherwise stated. Cells were cultured at 37^{o}C with 5% CO_2.

"Coriell" SCA7 (line GM03561) adult dermal fibroblast cell lines were obtained from the Coriell Institute for Medical Research, USA. A control line was obtained from Lonza (Switzerland). The 1518 and 1519 fibroblast lines were established from patient skin biopsies with informed consent (J. Scholefield, in preparation). Ethics approval for their use was granted by the University of Cape Town Faculty of Health Sciences Human Research Ethics Committee (REC REF. 380/2009 and 460/2010), renewed annually, incorporating amendments to the project protocol where necessary.

2.3.2 Patient-derived neuronal cultures

Patient dermal fibroblast cells were reprogrammed into induced pluripotent stem (iPS) cells by J. Scholefield and Dr Lauren Watson (UCT, South Africa) based upon a published protocol (Park et al. 2008), adapted according to the authors' recommendations. Briefly, retroviral vectors encoding three reprogramming factors (hKlf4, hSox2, hOct3/4) were produced in Platinum-GP cells (Cell Biolabs, USA) using plasmids obtained from Addgene (USA). One and two days after plating, fibroblasts were infected with the cocktail of retroviruses. Six days after plating, media was replaced with human embryonic stem cell medium (huESC medium). Valproic acid (VPA), a HDAC inhibitor, at a final concentration of 2 mM, was included in the medium for 10 days post-infection to enhance the efficiency of reprogramming. Thereafter, medium was replaced daily with sterile-filtered iMEF-conditioned huESC medium supplemented with 10 ng/ml basic fibroblast growth factor (bFGF, R&D Systems, USA). iPSC colonies could be observed as clustered cells approximately 20-40 d post-infection. iPSC colonies were manually dissected and grown on an iMEF feeder layer. Colonies were split every 7 d by manual dissection and cultured in huESC medium. Their genomic integrity was verified and pluripotency was confirmed both through expression of pluripotency markers (determined by RT-PCR and immunofluorescence

staining) and the capability for *in vitro* differentiation of all three germ layers from embryoid bodies.

The differentiation of iPSCs into neurons was carried out by L. Watson as previously described (Dottori and Pera 2008). Briefly, iPSCs were plated on an iMEF feeder layer with 500 ng/ml human recombinant Noggin (R&D Systems). After 14 d of Noggin treatment without passage, colonies were manually dissected and transferred to individual wells of a non-adherent 96-well plate, to allow neurosphere formation. Neurospheres were cultured in suspension in Neurobasal Medium A (NBM-A) (Life Technologies) supplemented with 10 ng/ml human recombinant epidermal growth factor (EGF) (Merck Millipore) and 10 ng/ml bFGF for approximately seven days.

To promote differentiation into neurons, whole neurospheres were seeded onto plates containing coverslips, coated with 10 µg/ml poly-D-lysine (Sigma Aldrich) and 5 µg/ml natural mouse laminin (Sigma Aldrich) in NBM-A without bFGF and EGF supplements. The expression of neuron- and glial-specific markers, including βIII-tubulin, gamma-aminobutyric acid (GABA), and glial fibrillary acidic protein (GFAP) was confirmed by immunofluorescence staining.

2.3.3 Transfection of established cell lines

HEK-293, N2A and SH-SY5Y cells were plated 24 h before transfection according to the conditions in Table 2-6 , except where otherwise stated. Media was replaced with DMEM-Glutamax + 10% FCS + 1% Antibiotic-Antimicotic, 24 or 48 h post-transfection as necessary.

For siRNA, cells were transfected at 40-60% confluency using Lipofectamine RNAiMax (Life Technologies) at 1.5 µl/ml. Transfection mixes were made up in Opti-MEM (Life Technologies). siRNAs were transfected at 20 nM unless otherwise stated.

For plasmids, HEK-293 and N2A cells were transfected at 60-90% confluency using Lipofectamine 2000 (Life Technologies) at 1 µl per µg total plasmid DNA. Transfection mixes were made up in Opti-MEM (Life Technologies). For plasmid transfection of SH-SY5Y cells, a Neon electroporation 10 µl kit (Life Technologies) was used according to the manufacturer's instructions, using optimised settings to maximise transfection efficiency and minimise cell death: Voltage = 1200 V, pulse width = 20 ms, number of pulses = 2. Electroporated cells were plated at 1×10^6 cells/ml in antibiotic/antimicotic-free media (DMEM-Glutamax + 10% FCS).

Mirtron, shRNA and miRNA cluster mimic plasmids (and all derivatives thereof) were used at a final concentration of 1 µg/ml. Luciferase plasmids were transfected at 0.5 µg/ml. Ataxin-7 plasmids were transfected at 2 µg/ml in 24-well plates or 1.5 µg/ml in 12-well plates.

Table 2-6 – Transfection conditions according to cell line and experimental application. Asterisk (*) indicates two wells were combined together into single samples. Dagger (†) indicates plates were pre-treated with Poly-L-lysine (Sigma-Aldrich).

Cell line	Application	Plate type	Biological replicates
HEK293	Luciferase	24-well	Triplicate
	Fluorescence	12-well†	Triplicate
	qPCR	24-well	Duplicate
	RT-PCR	24-well	Single
	Staining	24-well	Duplicate
N2A	qPCR	24-well	Duplicate
	Staining	24-well	Duplicate
	Western blot	24-well*	Single
SH-SY5Y	qPCR	24-well	Duplicate
	Staining	24-well	Duplicate
	Western blot	24-well*	Single

2.3.4 Transfection of patient fibroblasts

Fibroblasts were transfected according to the following conditions, where an asterisk (*) indicates that four wells were combined together into a single sample:

Application	Plate type	Biological replicates
Fluorescence	24-well	Triplicate
qPCR	12-well	Duplicate
Staining	24-well	Duplicate
Western blot	6-well*	Single

2.3.4.1 Electroporation of fibroblasts

Neon electroporation kits (Life Technologies) were used for plasmid and siRNA transfection according to manufacturer's instructions using optimised settings: Voltage = 1500 V, pulse width = 20 ms, number of pulses = 2. For 24 and 12-well plates, a 10 µl Neon kit was used, with 1×10^5 and 2×10^5 cells per tip, respectively. For 6-well plates, a 100 µl Neon kit was used with 1×10^6 cells per tip. Plasmids were used at a final concentration of 1 µg/ml, and siRNA at 20 nM unless otherwise stated. Following electroporation, cells were plated in DMEM-Glutamax + 10% FCS.

2.3.4.2 LNA transfection

Fibroblasts were transfected at 2×10^5 cells/ml with LNAs at 100 nM and 1.5 µl/ml Lipofectamine RNAiMax (Life Technologies) in 24- or 6-well plates, according to the manufacturer's instructions for reverse transfection, carried out by M. Varela.

2.3.5 Transfection for virus production

HEK-293 T cells were plated in 10 cm dishes (10 per virus) in DMEM-Glutamax + 10% FBS + 1% Antibiotic-Antimicotic 24 h prior to transfection, to reach 80% confluency. Modified pLKO plasmids (section 2.2.8), were transfected along with commercially available lentiviral

packaging plasmids psPAX2 and pmD2.G (gift from E. H. Chang), in the following ratio (by mass): 4:3:1. Transfection was carried out using filter-sterilised Calcium Chloride ($CaCl_2$) at a final concentration of 200 mM. HEPES-buffered saline (pH 7.05) was added dropwise to the plasmid-$CaCl_2$ mixture while vortexing and left on ice for 10 min to allow Calcium Phosphate crystals to form. The transfection mixture was added to cells 1:10 in media. Media was replaced 5 h post-transfection. Media was collected 48 h post-transfection and replaced with fresh medium which was collected 24 h later. Collected supernatant was spun down at 2000 rpm for 7 min at 4°C and filtered through a 0.45 µm membrane to remove cells. Filtered media was spun down at 25 000 rpm for 100 min at 4°C. The supernatant was removed and the viral pellet left to dry for 10 min, then allowed to resuspend in 100 µl phosphate-buffered saline (PBS) overnight at 4°C. Resuspended pellets were pooled, aliquotted and stored at -80°C. 10 ml of the filtered supernatant for both NAD-NAD and 1cN-18 viruses was retained and stored at -80°C.

2.3.6 Viral infection

HEK-293 and SH-SY5Y cells were plated 24 h before infection in 24- or 48-well plates such that they reached ~50% confluency at the time of infection. iPS-derived neurons were plated on coverslips in 24-well plates as in section 2.3.2. Infection was carried out by adding each virus (quantities are indicated in each experiment) to the media (DMEM-Glutamax + 10% FCS + 1% Antibiotic-Antimicotic) in the presence of polybrene (Merck Millipore, USA) where indicated. Media was replaced 24 or 48 h post-infection.

2.4 Protein assays

2.4.1 Fluorescence microscopy

Fluorescence was used to assess transfection and infection efficiency. Splicing efficiency for pEGFP-mirtron, pAtaxin-7-EGFP-mirtron and pAtaxin-7-mirtron-mCherry constructs was assessed qualitatively by fluorescence microscopy of EGFP or mCherry. Images were taken 48 h post-transfection, using the same exposure for each channel for each experiment. Images were optimised for brightness and contrast for printing where necessary.

2.4.2 EGFP and mCherry quantification

Media was removed 48 h following transfection and cells washed with PBS. Cell lysis was carried out in 100/200 µl protein extraction buffer (10 mM Tris, 100 mM NaCl, 1 mM EDTA, 1% Triton X-100) for 24-well/12-well plates, respectively. Total protein concentration was determined by Micro BCA assay (Thermo Scientific), according to the manufacturer's instructions. 250 µg of total protein (unless otherwise stated) was loaded into microplates and EGFP and mCherry fluorescence quantified using a Wallac-Victor 3 plate reader. Readings were normalised by subtracting measurements for untreated cells and dividing by the average for control-transfected cells.

2.4.3 Dual Luciferase assay

Dual Luciferase assays were carried out using a Dual-Luciferase Reporter Assay System (Promega, USA) as follows. Media was removed 48 h post-transfection, and cells washed with PBS then lysed in 100 µl Passive Lysis Buffer. 20 µl of each sample was loaded into 96-well microplates, and EGFP fluorescence was quantified to give an indication of splicing efficiency. Luciferase Assay Reagent II (containing the substrate for firefly Luciferase) was

added and a Wallac-Victor 3 plate reader was used to measure firefly luminescence. Adding Stop & Glo buffer then quenches this signal and provides the substrate to allow *Renilla* Luciferase activity to be measured.

A ratio of *Renilla*:firefly luminescence was calculated, averaged across the replicates. Firefly Luciferase is constitutively expressed from the dual Luciferase plasmid, therefore dividing by firefly luminescence corrects for cell number and transfection variations. Values were normalised to controls in each experiment.

2.4.4 Western blotting

Cells were harvested with TrypLE (Life Technologies). Cell pellets were resuspended in Radioimmunoprecipitation (RIPA) buffer (50 mM Tris, 150 mM NaCl, 0.1% SDS, 0.5% sodium deoxycholate, 1% Triton X-100) with protease inhibitor (Roche), shaken on ice for 1 h and centrifuged at 16 000 g for 20 min to remove insoluble material. Protein concentration was quantified using a BCA assay (Thermo Fisher Scientific) according to the manufacturer's instructions. Protein extracts were stored at -80°C.

Equal amounts of protein were loaded (quantities as stated), diluted in NuPAGE LDS Sample Buffer (Life Technologies) with NuPAGE Sample Reducing Agent (Life Technologies) and made up to equal volumes with water. Samples were incubated at 75°C for 15 min prior to loading and HiMark protein ladder (Life Technologies) was used for estimation of protein size. Western blotting was carried out on 3-8% Tris-Acetate gels (Life Technologies) using the Novex system (Life Technologies) according to the manufacturer's instructions and transferred onto PVDF membrane. 5-10% milk in TBS-0.1% Tween 20 (TBST) was used for blocking (1 h) and 1-2% milk-TBST for antibody dilutions. TBST was used for washing (3x10 min after each antibody incubation). Antibody concentrations were optimised by first

using a 1:500 dilution and reducing the dilution as necessary for a strong signal:background ratio. Anti-Ataxin-7 antibody 1-135 (raised in rabbit against amino acids 1-135; gift from A-L Ström) (Ström et al. 2002) was used at 1:250 unless otherwise indicated, incubated overnight at 4°C. A mouse anti-vinculin antibody, hVin-1 (Sigma-Aldrich), was used at 1:100 000 as loading control (incubated for 1 h). Fluorescent secondary antibodies, IRDye 800CW Goat anti-Rabbit IgG and IRDye 680RD Goat anti-Mouse IgG (Li-Cor) were used at 1:10 000 (1-2 h incubation). Sodium dodecyl sulphate (SDS) was added at 0.01% to 800CW antibodies to reduce background staining. Blots were imaged using a Li-Cor Odyssey fluorescent imager.

2.4.5 Cell immunostaining

Cells were washed with PBS then fixed in 4% Paraformaldehyde (Alfa Aesar, UK) for 15 min and washed with PBS for 2x10 min. Cells were blocked in 10% Normal Goat Serum (NGS) in 0.1% Triton X-100/PBS (PBST) for 1-3 h. Antibodies were diluted in 1% NGS/PBST. Ataxin-7 primary antibodies (1-135 and 1-15 as indicated; gift from Anna-Lena Ström, Sweden) were used at 1:100 unless otherwise stated. The concentration was determined through reducing dilutions until the signal to background ratio was strong, but this was limited by the total amount of antibody available. Anti-polyQ antibody 1c2 was used at the stated concentrations. Wells were incubated in primary antibody overnight at 4°C then washed with 3x10 min with PBST. No-primary controls were included to exclude any non-specific binding of secondary antibodies. Wells were incubated with goat Alexa-Fluor 488 or 543 secondary antibodies (Life Technologies) at 1:500 in PBST for 1-2 h and washed with PBS 3x10 min. Patient-derived neuronal cells (which were plated on coverslips, section 2.3.2) were then mounted onto slides using Vectashield Hard Set Mounting Medium with DAPI (Vector Laboratories, UK).

2.5 Nucleic acid assays

2.5.1 RNA extraction

For HEK, N2A, and SH-SY5Y cells, RNA extraction was carried out using TRIzol (Life Technologies) according to the manufacturer's instructions, unless otherwise stated. For fibroblasts, RNA extraction was carried out using an RNeasy kit (Qiagen) according to the manufacturer's instructions.

For small RNA qPCR assays (section 2.5.3), RNA was extracted from HEK-293 and fibroblast cells using the RNeasy kit (Qiagen), adapted to retain small RNAs (70% ethanol replaced with 100% ethanol; Buffer RW1 not used).

RNA was treated with RQ1 DNase (Promega) or Turbo DNase (Ambion) where necessary (i.e. When the RNA was to be assayed for a gene which had been delivered by transfection). RNA was assessed for concentration and quality using a NanoDrop spectrophotometer (Thermo Fisher Scientific) and diluted to equal concentrations for each experiment.

2.5.2 RT-PCR

Reverse transcription was carried out using Thermoscript Reverse Transcriptase (Life Technologies) using oligo-$dT_{(12\text{-}18)}$, according to the manufacturer's instructions (65°C incubation for 30-45 min). A no-reverse transcriptase (-RT) control was used to indicate DNA contamination (plasmid or genomic). No-template controls were used to check for contamination in reagents. RT-PCR was also carried out on un-transfected (or mock transfected) cells to control for amplification of endogenous RNA or DNA. The PCR step was carried out using primer pairs listed in Table 2-7 and Taq DNA polymerase (NEB) in optimised conditions, except for RT-PCR of Ataxin-7 (with introns in sites 3-5 and 8-10), where

AccuPrime GC-rich DNA polymerase (Life Technologies) was used. Products were run on 1-2% agarose gels until separated, alongside a 1 kb or 100 bp ladder (NEB) or Hyperladder II (Bioline).

Quantification of splicing efficiency from RT-PCR gel images was carried out using ImageJ software before inverting the colours or carrying out any other image enhancement. The intensity of three points across the centre of each band was averaged, and the average background measurement for each lane was subtracted. Splicing efficiency was calculated by dividing the intensity of the "spliced" band by the total (sum of the spliced and un-spliced bands).

2.5.3 Real-time quantitative RT-PCR

For real-time quantitative RT-PCR (qPCR), cDNA synthesis was carried out using a High Capacity cDNA Reverse Transcription kit (Life Technologies) according to the manufacturer's instructions using random primers (supplied). An equal amount of RNA was used as template for each sample in the same experiment (0.5-1 µg for standard cell lines and 0.2-0.5 µg for fibroblast lines). qPCR was carried out using Power or Fast SYBR-green Mastermix (Life Technologies) except where otherwise stated, on an Applied Biosystems StepOne Plus Real-time PCR System (Life Technologies), complying with the Minimum Information for publication of Quantitative real-time PCR Experiments guidelines (Bustin 2010; Taylor et al. 2010) as far as possible. Results were analysed using StepOne Software v2.1 (Applied Biosystems). Primers were designed by Primer Design (UK) and purchased from either Primer Design or IDT. Those for which the sequence is known are listed in Table 2-7. A dilution series was used to produce a standard curve for each primer pair in each cell line, from which expression levels in each sample were quantified. Mouse assays were normalised to housekeeping gene Ubiquitin-C (UBC), human to β-actin, unless otherwise

stated. The melt curve for all products was confirmed to have a single peak indicating that a single product was formed, except in the case of UBC, which was known by the manufacturer to sometimes produce a double peak from a single product. Results were discarded and the experiment repeated wherever the Ct values for –RT samples were less than 7 cycles above the other samples, indicating that DNA contamination may impact the results significantly.

Small RNA qPCR assays were custom-designed by Life Technologies and carried out as above, except cDNA synthesis was carried out using gene-specific primers (supplied) with 100 ng total RNA and qPCR was carried out using TaqMan gene expression mastermix (Life Technologies). Values were normalised to GAPDH (TaqMan assay) (Life Technologies). A standard curve analysis was not performed and quantification was instead carried out using the delta-delta-Ct method.

Table 2-7 – Primer combinations used for RT-PCR and qPCR. Sequences are listed in the 5'-3' orientation.

Application	Primers	Sequence
EGFP splicing	EGFP-UF	atcctggtcgagctggacg
	EGFP-R2	cggttcaccagggtgtcgcc
EGFP splicing (intron 2)	EGFP-mirt2-F	ggtgaagttcgagggcgac
	EGFP-UR	ctgccgtcctcgatgttgtg
Cluster mimic splicing	MCM7-qF2	gatgccacctatacttctgcc
	MCM7-qR2	tcctttgacatctccattagcc
Ataxin-7 splicing (5'-UTR intron)	UTR4RT-F	ccgctagcttaatacgtcttcaag
	SCA7-spl-R3	ctccgacatagcataatctggaac
Ataxin-7 splicing (sites 1-5)	BspEI-F	tcgccatccggactctcag
	SCA7-spl-R	gccgccactggtactgctact
Ataxin-7 splicing (sites 8-10)	SCA7-spl-F	gcctctgcccagtcctgaa
	SCA7-spl-R2	tgatgaggatgcgctaagaaca
Ataxin-7 (human) qPCR	SCA7-hqF	gccagccgtgaacaatgtc
	SCA7-hqR	ttcctccccgtgctattttca
UCHL1 qPCR	UCHL1-qF	tgaagcagaccattgggaat
	UCHL1-qR	tgtttcagaactgatccatcct
DNAJA1 qPCR	DNAJA1-qF	aaaggaggagaacaggcaattaa
	DNAJA1-qR	tagggttactgagagctgatgt

2.5.4 Deep sequencing

High-throughput sequencing of small (~18-24-nt) RNA extracted from transfected HEK-293 cells was carried out by Y. Seow similar to as previously described (Seow et al. 2012). Briefly, HEK-293 cells were transfected in a 24-well plate with 500 ng of each plasmid. RNA was harvested and small RNA libraries were prepared with the Small RNA v1.5 Sample Prep Kit (Illumina, USA) following the manufacturer's instructions. The RNA was ligated with a 3'-RNA adaptor modified to target small RNAs with 3'-hydroxyl groups, and then with a 5'-RNA adaptor. Reverse transcription followed by PCR was performed to select for adaptor-ligated fragments. Products were barcoded with specific tags during PCR to allow later identification following pooling of samples. The bands of expected sizes were cut out from a

polyacrylamide gel (6% TBE PAGE) and dialysed to harvest the DNA. The samples were pooled into two seperate lanes (mirt-1cN constructs in one lane and mirt-18 in the other). Libraries were sequenced on a Genome Analyzer IIx for 36 cycles following the manufacturer's protocols. Image analysis and base calling were done using Illumina's GA Pipeline. Adaptors were trimmed with Biopieces remove_adapter script, and remaining sequences were aligned against full-length mirtron hairpins.

2.6 Statistics

Error bars represent standard deviation or standard error as indicated in each figure. To show significance, p-values shown represent student’s one-tailed t-test, assuming equal variation.

3 Silencing Ataxin-7 with intronic miRNAs

3.1 Introduction

Spinocerebellar Ataxia type 7 (SCA7) is caused by an expansion in the CAG-repeat region of the gene encoding Ataxin-7, leading to a long poly-glutamine stretch in the protein, which has both gain- and loss-of-function effects. The principal requirement for effective treatment of SCA7 and other expansion disorders is likely to be reducing levels of the mutant protein. As discussed in Chapter 1 (section 1.4), allele-specific silencing may be possible and in certain cases desirable, but has a number of limitations. Non-allele-specific therapy can be achieved more easily, but the resulting deficiency in the silenced protein may be problematic if it has an important cellular function. This non-allele-specific silencing may therefore be combined with delivery of a replacement wild-type gene to avoid protein deficiency, forming a gene knockdown-replacement (K&R) therapy.

In order to achieve silencing of Ataxin-7, the endogenous RNA interference (RNAi) pathway may be exploited. RNAi is an endogenous post-transcriptional gene silencing system, through which short interfering RNAs (siRNAs), 21-23 bp in length, reduce target gene expression by complementary base pairing with target mRNA and either inhibiting translation or inducing mRNA cleavage. RNAi may be used to target both alleles of a gene in a mutation-independent manner, resulting in reduction of both the wild-type and mutant protein. An RNAi-resistant copy can be concurrently delivered to maintain functional protein levels (Kubodera et al. 2005, 2011; Kiang et al. 2005; Palfi et al. 2006; O'Reilly et al. 2007; Millington-Ward et al. 2011; Greenwald et al. 2012).

As discussed in Chapter 1 (section 1.3.3), RNAi may be delivered in the form of short hairpin RNAs (shRNAs) or artificial microRNAs (miRNAs), which are produced from large stem-loop transcripts, cleaved by the microprocessor complex and Dicer into siRNAs. Mimicking endogenous miRNAs, and driving them with pol-II promoters, can demonstrate lower toxicity compared to shRNAs, yet achieve strong target silencing (Giering et al. 2008; McBride et al. 2008; Borel et al. 2011). Many natural miRNAs are found within introns (intronic miRNAs) (Kim and Kim 2007). Mirtrons are a special class of intronic miRNA where the ends of the short pre-miRNA stem-loops are defined by the splice sites and the typical Drosha processing step is not required (Ruby et al. 2007a). Natural intronic miRNA and mirtrons can be adapted to successfully silence a gene of interest *in vitro* (Chung et al. 2006; Aagaard et al. 2008; Seow et al. 2012; Sibley et al. 2012). A potential benefit of using mirtrons in gene therapy is the strict definition of pre-miRNA through the splicing reaction, perhaps minimising off-target effects by producing a well-defined seed region compared to miRNA and shRNA processed by the microprocessor.

It is desirable to deliver both miRNA and replacement gene in a single construct to ensure cells receive both aspects of the therapy (Greenwald et al. 2012). To achieve this, mirtrons and other intronic miRNAs could be nested within a replacement gene. Mirtrons may be particularly advantageous for this purpose due to their minimal size. However, artificial mirtron design can be more challenging than other miRNA approaches due to the specific sequence elements which must be incorporated for their recognition as introns by the spliceosome (Seow et al. 2012).

Towards the development of a gene K&R therapy for SCA7, it is first necessary to optimise the silencing strategy to find an efficiently processed miRNA structure and an effective antisense sequence. This chapter covers the design, screening and functional testing of

mirtrons and other intronic miRNAs designed to target Ataxin-7 for silencing. The sequence of the target site can then be altered to create a miRNA-resistant replacement gene for a K&R therapy, which will be investigated in the next chapter.

Given their stricter constraints for design, artificial mirtrons targeting Ataxin-7 were designed first and placed as introns within an enhanced green fluorescent protein (EGFP) gene. These mirtrons were then adapted into mimics of a natural intronic miRNA cluster. This chapter will cover their design, optimisation, splicing and target silencing activity.

3.2 Screening revealed a number of promising synthetic mirtrons

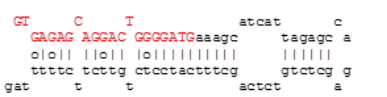
A previously published algorithm was used to identify putative target sites for artificial mirtrons in an Ataxin-7 cDNA sequence (Seow et al. 2012) (detailed in section 2.1). Taking into account the specific splicing requirements for each end of an intron, the algorithm identified two sets of target sites, those which may be targeted by the 5' arm of a mirtron, and those which may be targeted by the 3' arm. These were then ranked by the algorithm according to their suitability for RNAi silencing according to published criteria. After discarding any target sites which would require multiple (>3) interruptions to the poly-pyrimidine track of the mirtron, four of the most highly scoring non-overlapping sequences of the 5' set and two from the 3' set were selected. A Basic Local Alignment Search Tool (BLAST) was used to ensure that these sequences were not found in other human transcripts (http://blast.ncbi.nlm.nih.gov/Blast.cgi).

Based upon the selected target sites, six synthetic mirtrons were designed (mirt-1-6, Figure 3-1, List A). The position of each mirtron's target site is indicated in Figure 3-2. Each mirtron guide strand was designed with complete complementarity to its target site. Following

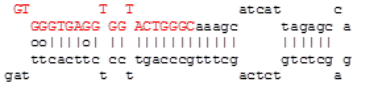
previous results in the lab, guide strands were each incorporated into a stem-loop design based upon a silencing-optimised adaptation of the endogenous mouse mirtron miR-1224 (Chapter 2; Figure 2-1) (Sibley et al. 2012; Seow et al. 2012). For mirt-1-4 where the guide strand was on the 5' arm, some complementary purines within the 3' arm were substituted with pyrimidines, reducing the number of interruptions to the poly-pyrimidine tract to improve splicing potential. Weak (G-T) mismatches were incorporated at the 5' end of each guide strand (by editing the passenger strand), in order to bias the selection of the intended mature miRNA into the RNA-induced silencing complex (RISC) (Khvorova et al. 2003).

List A

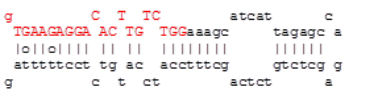
mirt-1 – Target 5′catccccagtcctgctctcat



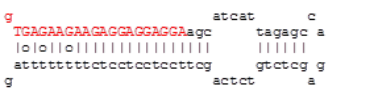
mirt-2 – Target 5′gcccagtaccacctcacccat



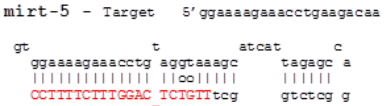
mirt-3 – Target 5′ccagacaagtgtcctcttcat



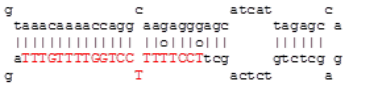
mirt-4 – Target 5′tcctcctcctcttcttctcat



mirt-5 – Target 5′ggaaaagaaacctgaagacaa

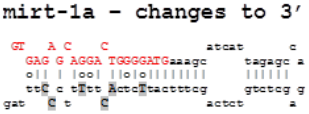


mirt-6 – Target 5′ aaacaaaaccagggaaaagga

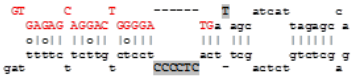


List B

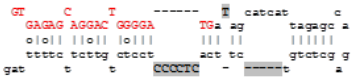
mirt-1a – changes to 3′ end



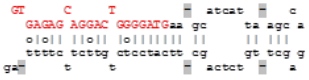
mirt-1b – addition of extra sections (as in 1224)



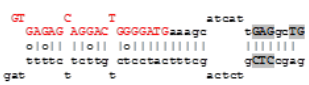
mirt-1bx – additions and deletion



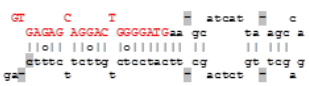
mirt-1c – deletions



mirt-1d – changes to loop section

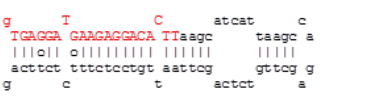


mirt-1cN

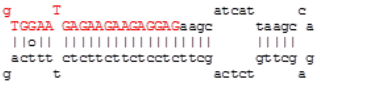


List C

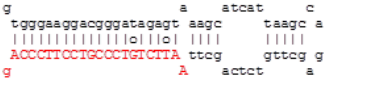
mirt-9 – Target 5′aagtgtcctcttcatcctcat



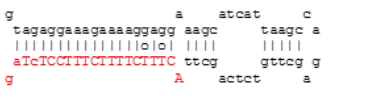
mirt-12 – Target 5′ctcctcttcttctcattccat



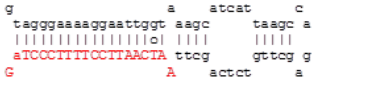
mirt-13 – Target 5′tgggaaggacgggacagaatt



mirt-14 – Target 5′tcataggaaagaaaagaaagt



mirt-15 – Target: 5′accagggaaaaggaattgatt



mirt-18 – Target: 5′cagggtagaagaaaacgattt

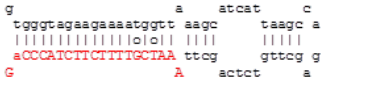


Figure 3-1 - Sequence and structure of synthetic mirtrons designed to target Ataxin-7. Six putative target sites in an Ataxin-7 cDNA sequence were selected, for which synthetic mirtrons were designed. The target sequence within Ataxin-7 is shown for each mirtron and the guide strand (the reverse complement of the target site) is indicated in red capitals. The guide strands were incorporated into an adapted mouse miR-1224 mirtron structure (Figure 2-1). Weak (G-T) mismatches are indicated with circles rather than lines. Where a G residue of the guide strand is shown in red but lowercase, this indicates that this base will only weakly base pair with a U in the target site, but strong binding was compromised in order to abide by the splicing requirements. **List A:** Initial designs of six mirtrons, mirt-1-6. **List B:** Variants of mirt-1, having the same guide strand but with modifications to the loop section and/or passenger strand as highlighted. Primers were designed to produce mirt-1b; though mirt-1bx was the resulting sequence. **List C:** Following improvement to the algorithm, a new set of target sites were identified within Ataxin-7 cDNA and mirtrons designed against six of these, based upon the hairpin structure of mirt-1cN.

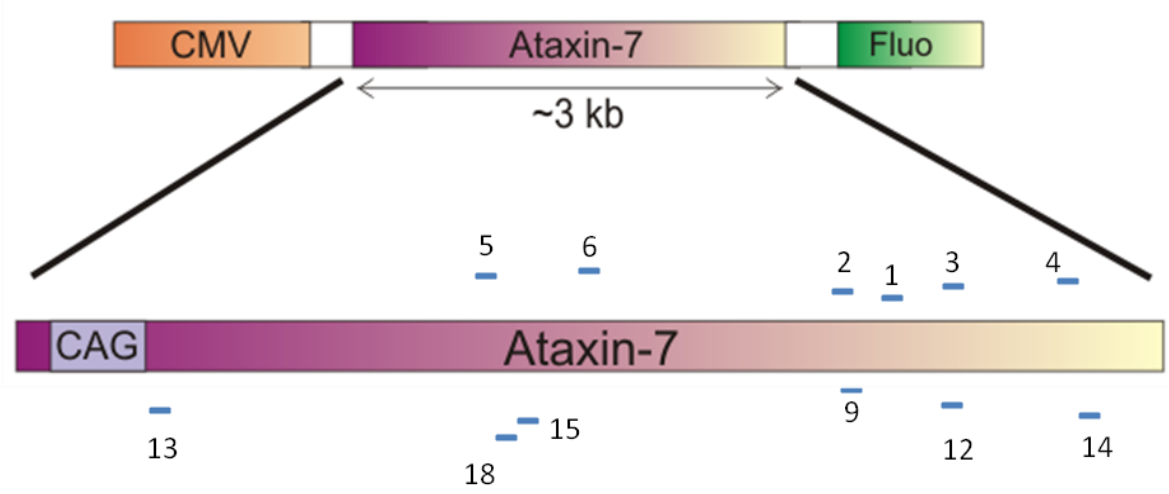


Figure 3-2 – Position of mirtron target sites within Ataxin-7 cDNA. Schematic representation of the Ataxin-7 cDNA construct, its CMV promoter and fluorescent tag (Fluo). Labels indicate the position of each mirtron target site. “CAG” represents the CAG-repeat and GC-rich region.

The initial group of six Ataxin-7-targeting mirtrons were individually cloned into a splice site within the coding sequence of EGFP, such that production of EGFP mRNA was dependent upon splicing of the mirtron (Figure 3-3a) (Seow et al. 2012). Therefore, the resulting fluorescence gives a measure of splicing efficiency. Mirtron constructs were transfected into HEK-293 cells and fluorescence was compared to an identical construct containing a non-hairpin forming intron of comparable length (91-nt), namely NADH-coenzyme Q reductase intron 6 (NAD intron). Splicing was successful to varying degrees for four of these artificial mirtrons (Figure 3-3b). Mirt-2 and -5 showed no fluorescence (data not shown).

To screen mirtrons for their silencing capability, the target sequences were all cloned into a single dual Luciferase reporter construct, within the 3' UTR of *Renilla* Luciferase. Post-transcriptional silencing activity is therefore evidenced by a reduction in the resulting enzymatic activity. This is then normalised to the activity of firefly Luciferase, which is constitutively expressed from the same plasmid. Although three of the spliced mirtrons did not cause a significant reduction in Luciferase activity, encouragingly, a 50% reduction in the relative Luciferase ratio was achieved by mirt-1 ($p<0.005$, Figure 3-3c).

Various changes were made to this construct to investigate whether some aspects of its structure may be limiting its processing, following its efficient splicing (Figure 3-1, List B). The following changes were tested for their impact on the silencing ability of this mirtron: (a) altering the 3' strand to change (weaken) the secondary structure of the stem, (b) more closely mimicking the structure of miR-1224, (c) reducing the length of the loop section and (d) changing the sequence of the loop section. For all of these changes, the resulting splicing levels were comparable to mirt-1 (Figure 3-3b). However, there was variation in silencing efficiency ranging from less than 40% to over 80% (Figure 3-3c). Weakening the secondary structure and mimicking miR-1224 led to reduced silencing activity, as seen previously

(Sibley et al. 2012), whereas shortening the loop section improved silencing. Because splicing was not significantly altered and the mature species should be identical, these effects are likely to be attributable to differences in downstream processing of the miRNA. The mirt-1cN variation was adapted from mirt-1c by making a single base change in the passenger strand, to allow creation of a comparable shRNA without premature termination (described in section 2.2.3 and discussed in the next section).

Having obtained one promising mirtron, it was desirable to design additional mirtrons against Ataxin-7 to provide a number of potential sequences for later development into K&R therapies. An updated version of the algorithm, further optimised for selection of targets for mirtrons with 3' guide strands using the *siR* algorithm (Shah et al. 2007), was used to identify a new set of target sites. Mirtrons were designed against six of these, four of which were 3' targets (Figure 3-1, List C), using the previously successful hairpin structure of mirt-1c.

Splicing levels were 40-70% of the NAD control for four of the six new mirtrons (Figure 3-3b) but very low or nil for two (mirt-12 and mirt-13). 40-50% Luciferase target silencing was observed for mirt-9 and mirt-15, and almost 90% for mirt-18 (Figure 3-3c).

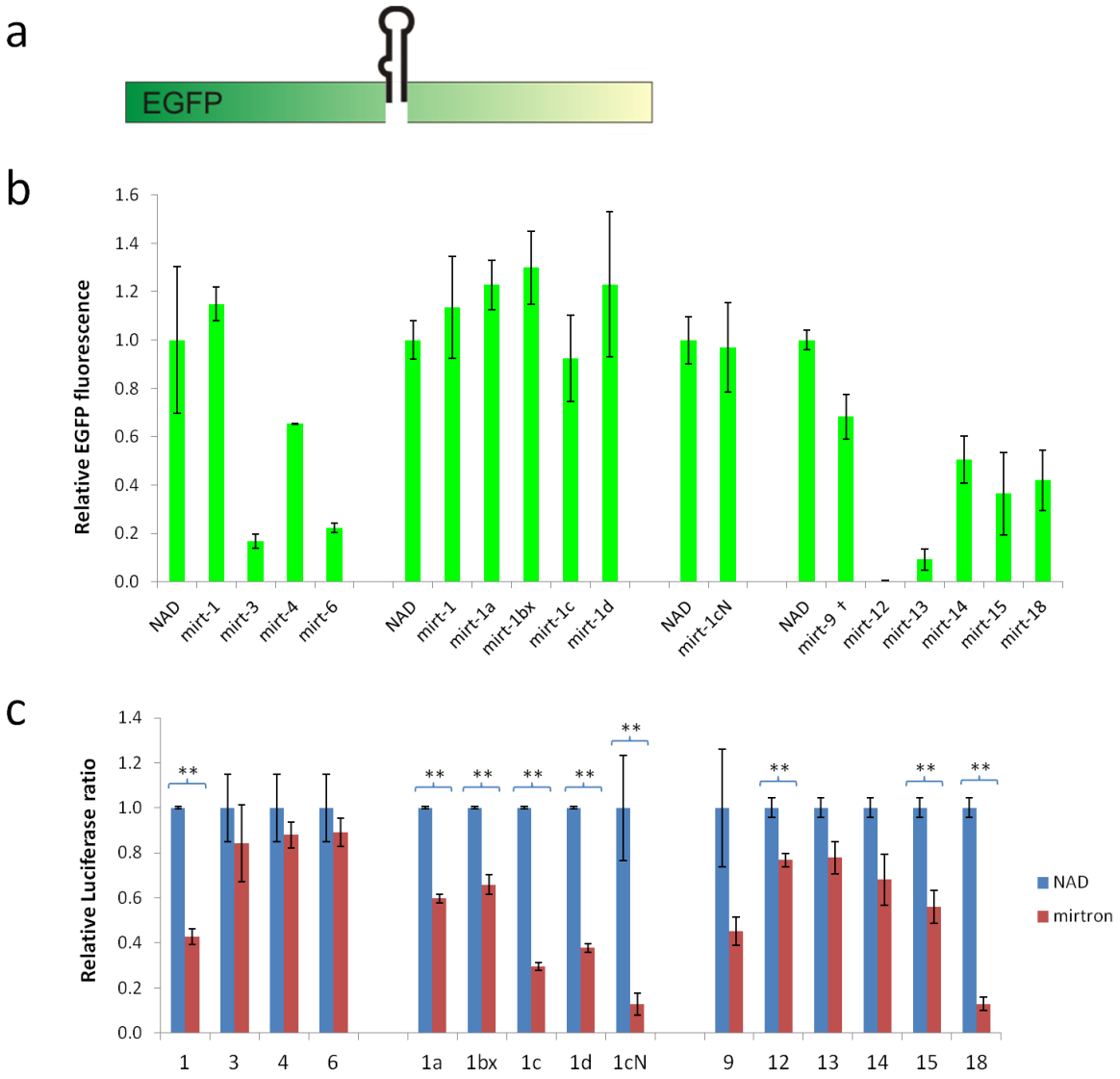
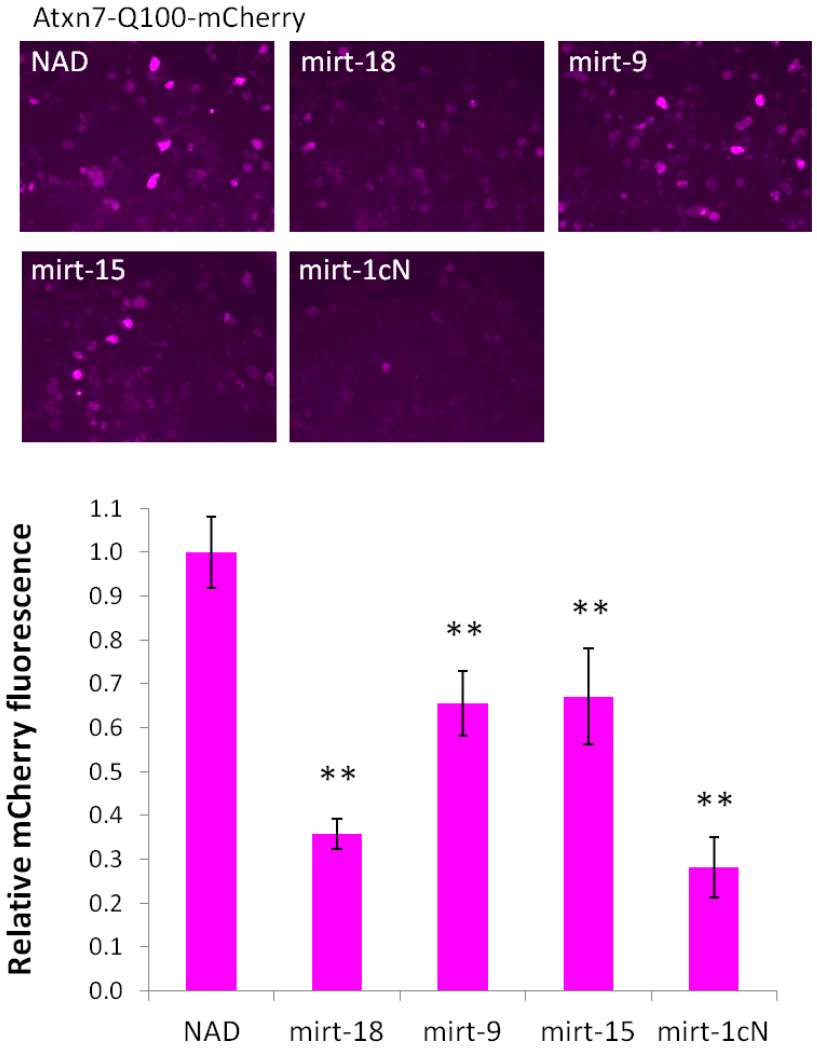


Figure 3-3 - Mirtron screening results. (a) Schematic diagram of the EGFP-mirt construct, showing the position of the artificial mirtron within the open reading frame of EGFP. **(b)** Splicing efficiency of mirtrons as represented by EGFP fluorescence in transfected HEK-293 cells after 48 h, compared to NAD control, which is normalised to 1 for each set. 50 µg total protein was loaded and fluorescence was measured using a plate reader. Background fluorescence (un-transfected cells) was subtracted from all readings. N=3 for mirtrons 1-6; N=6 for all others except where indicated (†, N=3). Error bars represent standard deviation. **(c)** Normalised *Renilla* to Firefly Luciferase ratios for dual Luciferase constructs containing mirtron target sites in the 3' UTR of *Renilla* Luciferase (Tar1 for mirt-1-6, and Tar2 for mirt-9-18). Transient co-transfections were carried out in HEK-293 cells and lysis was performed 48 h later. In each case the value obtained with the non-targeting NAD intron was normalised to 1. N=3. Error bars represent standard deviation. Asterisks (**) represent $p<0.005$ (one-tailed t-test).

Transcript secondary structure can inhibit the access of RNAi effectors to their target sites (Bohula et al. 2003). Ataxin-7 mRNA secondary structure was not necessarily accurately represented in the Luciferase screen, where only the short target region itself was included. Therefore, four of the most promising mirtrons were then tested for silencing activity against a full-length human Ataxin-7 gene, which more closely resembles the true target. Mirtrons were co-transfected with mCherry-tagged mutant Ataxin-7 containing an expansion of 100 CAG repeats (Ataxin-7-Q100-mCherry). Encouragingly, mirt-1cN and mirt-18 each achieved 65-75% silencing of the tagged mutant protein, as measured by mCherry fluorescence levels ($p<0.005$, Figure 3-4a). shRNA equivalents for these two mirtrons were also designed, to produce the same mature RNAi species (Figure 3-4b). The activity of the mirt-18 and mirt-1cN shRNAs was comparable to a positive control shRNA targeting Ataxin-7 (shR-1; Figure 3-4c). Notably, shRNA-directed silencing was also comparable to the silencing achieved by the parent mirtrons, despite the stronger promoter which drives shRNA expression. The functionality of these shRNAs provides extra confirmation of the RNAi activity of the selected mirtrons. It was also confirmed that these mirtrons silenced a non-expanded (wild-type) Ataxin-7 construct with similar efficiency, as expected (Figure 3-4d).

a



b

shR-1

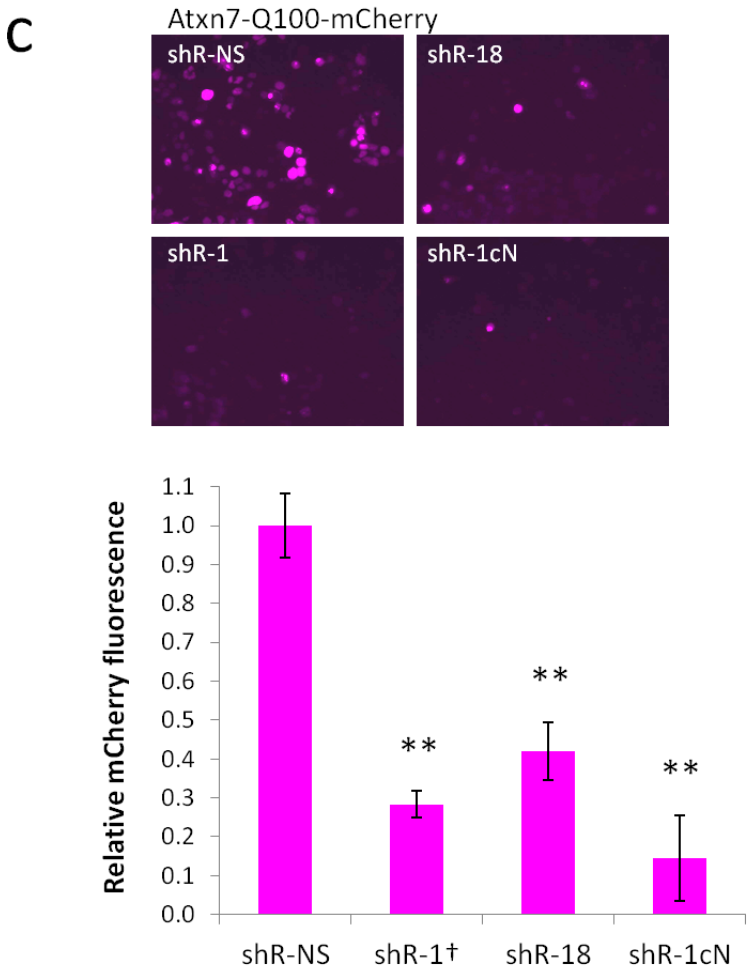
```
  g                  ccu
   ctgaagacaattctaata   g
   ||||||||||||o||||    a
   GACTTCTGTTAAGGTTAT   c
ttG                   acc
```

shR-18

```
gt                  a    atcat     c
  gggtagaagaaaatggtt aagc     taagc a
  ||||||||||||||o|o|| ||||     |||||
  CCCATCTTCTTTTGCTAA ttcg     gttcg g
tt                  A    actct     a
```

shR-1cN

```
GT     C     T           atcat     c
  GAGAG AGGAC GGGGATGaagc     taagc a
  |o|| ||o|| |o|||||||||     |||||
  ctttc tcttg ctcctacttcg     gttcg g
tt     t     t           actct     a
```


c
Atxn7-Q100-mCherry
shR-NS
shR-18
shR-1
shR-1cN
Relative mCherry fluorescence
1.1
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
**
**
**
shR-NS
shR-1†
shR-18
shR-1cN

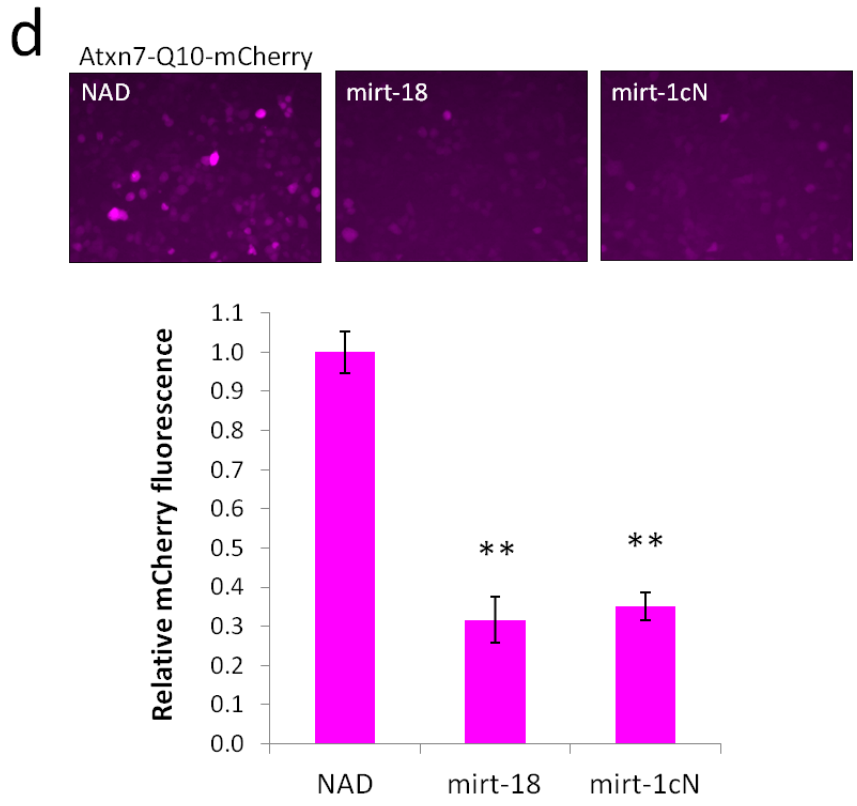

d
Atxn7-Q10-mCherry
NAD
mirt-18
mirt-1cN
Relative mCherry fluorescence
1.1
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
**
**
NAD
mirt-18
mirt-1cN

Figure 3-4 - Silencing of mCherry-tagged Ataxin-7 constructs by mirtrons and shRNAs along with representative images. Ataxin-7 levels are represented by mCherry fluorescence in co-transfected HEK-293 cells after 48 h. 250 µg total protein was loaded and fluorescence was measured using a plate reader. Background fluorescence (un-transfected cells) was subtracted from all readings and control values (NAD intron or non-specific shRNA, shR-NS) were normalised to 1. **(a)** Silencing of Ataxin-7-Q100 by mirtrons. N=6, from two independent experiments. **(b)** The sequence and structure of positive control shRNA (shR-1) and the shRNA equivalents for mirt-18 and mirt-1cN are shown (guide strands are shown in red capitals and bases altered compared to the original mirtrons are highlighted). **(c)** Images and quantification of Ataxin-7-Q100 silencing. N=6, except where indicated (†, N=5), from two independent experiments. **(d)** Silencing of wild-type Ataxin-7 (Atxn7-Q10-mCherry) by mirtrons. N=3. For (a, c and d), error bars represent standard deviation and asterisks (**) indicate $p<0.005$ compared to NAD or shR-NS (one-tailed t-test).

3.3 Mirtron processing

To consider synthetic mirtrons to be appropriate for development as intronic miRNAs for therapeutic use, it is important to investigate their processing, to confirm that mirtrons behave as predicted in terms of splicing, processing and RNAi activity.

3.3.1 Mirtron activity is dependent upon splicing

Intron-spanning RT-PCR was carried out to confirm the splicing efficiency of artificial mirtrons at the RNA level (Figure 3-5a and b). Encouragingly, the proportion of spliced mRNA closely reflected the EGFP quantification (Figure 3-3b), indicating that the splicing of transcripts translates directly to an equivalent level of production of functional protein. It was also verified that the NAD intron was spliced completely. Mirt-18 was spliced from around 60% of transcripts; mirt-1cN from around 95%. Figure 3-5a confirmed that the lower EGFP fluorescence from mirt-18 was attributable to poor splicing efficiency rather than mis-splicing through activation of cryptic splice sites.

In order to prove that the artificial Ataxin-7 mirtrons were processed via the mirtron pathway, un-spliceable (US) versions of the two most promising mirtrons were made. These were identical to the parent sequence except for a single G-A base substitution in the 5' splice site (Figure 3-6a), which prevents recognition by the spliceosome (Montell et al. 1984; Seow et al. 2012). No EGFP fluorescence was detected for these constructs (Figure 3-6b) and the silencing of mutant Ataxin-7 was substantially reduced compared to the original mirtrons (Figure 3-6c). This demonstrates a substantial dependence of mirtrons on splicing. Crucially, the single base change made in a corresponding shRNA (Figure 3-6d) did not affect silencing of the Luciferase target (Figure 3-6e). This confirms that downstream processing and mRNA silencing activity are not affected by the base change, since shRNAs are not processed by splicing. The small but significant level of silencing attributable to un-spliceable mirtrons (Figure 3-6c) may be explained by processing of the hairpin-containing transcripts by Drosha, perhaps as a result of the splicing impairment (this is investigated further in section 3.3.2).

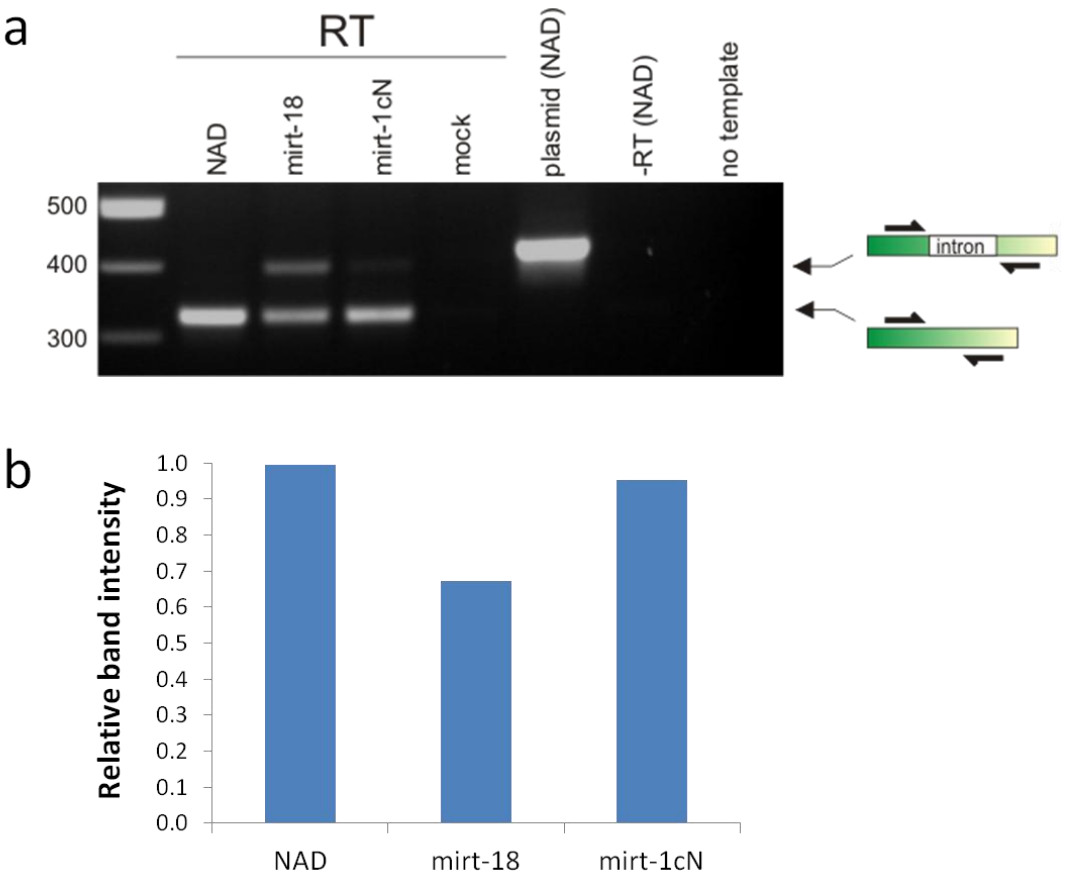


Figure 3-5 - RT-PCR demonstrating splicing efficiency of EGFP-mirtron constructs. **(a)** PCR was carried out on cDNA from HEK-293 cells transfected with EGFP-mirtron constructs, using a pair of intron-spanning primers. The products represent the presence of un-spliced and spliced transcripts for each mirtron, as indicated in the schematic diagram. **(b)** Quantification of splicing efficiency for each mirtron, based upon band intensities in (a), calculated as a proportion of spliced product to the total amount of transcript (spliced + un-spliced).

a

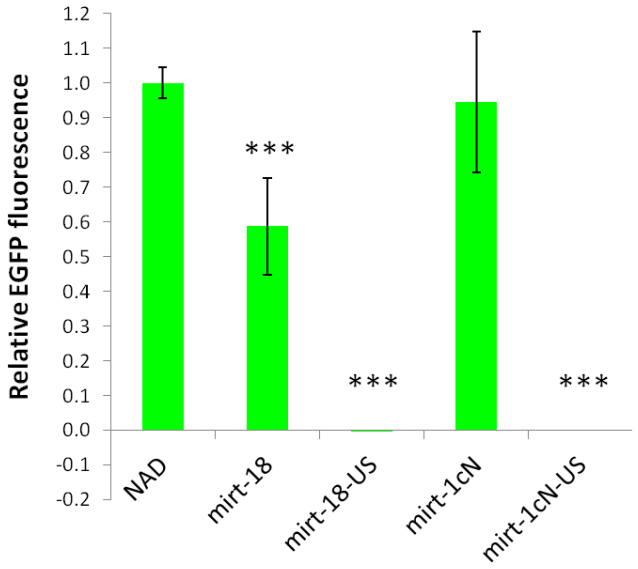
mirt-18-US

```
A                  a    atcat     c
 tgggtagaagaaaatggtt aagc     taagc a
 |||||||||||||o|o|| ||||     |||||
 aCCCATCTTCTTTTGCTAA ttcg     gttcg g
G                   A    actct     a
```

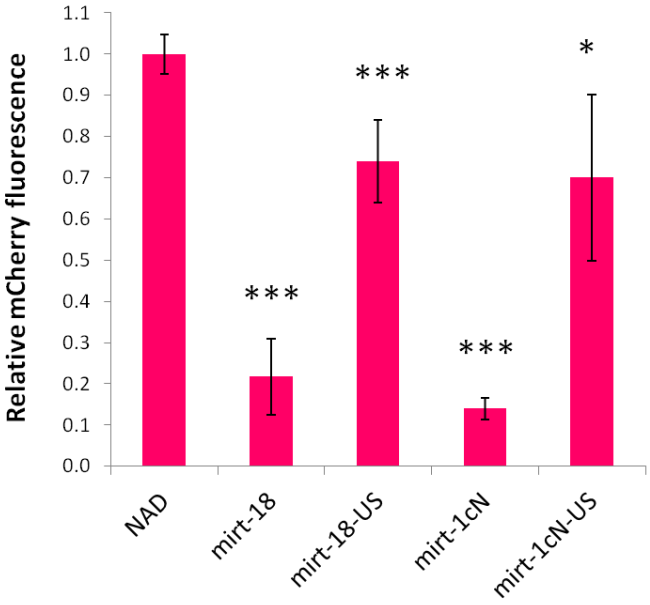
mirt-1cN-US

```
AT     C     T           atcat     c
  GAGAG AGGAC GGGGATGaagc     taagc a
  ||o|| ||o|| |o|||||||||     |||||
  ctttc tcttg ctcctacttcg     gttcg g
ga     t     t           actct     a
```

b



c



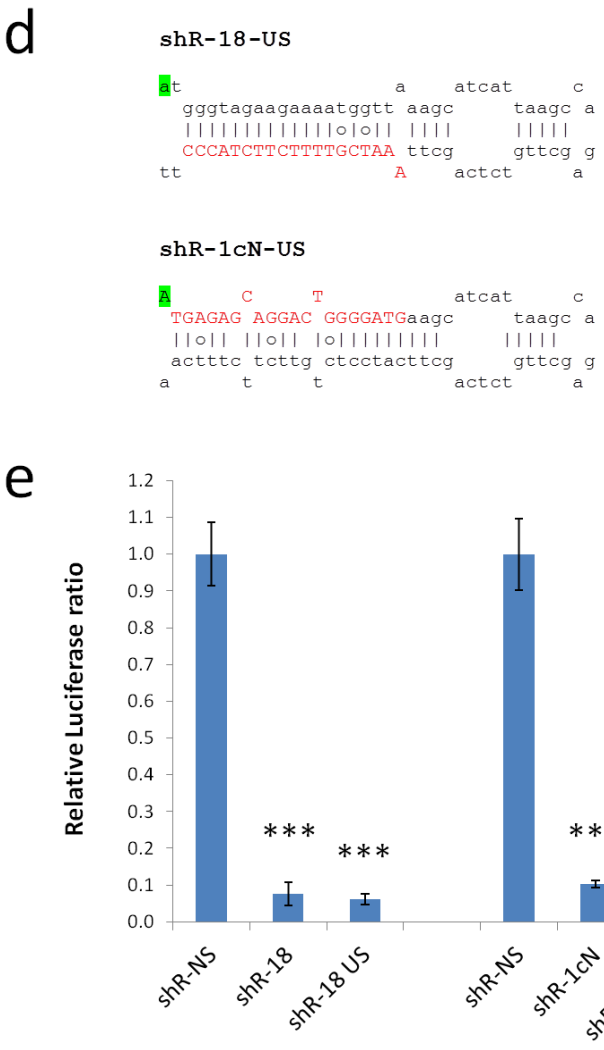


Figure 3-6 - Splicing and silencing activity of un-spliceable equivalents of mirtrons in HEK-293 cells. (a) Sequence and structure of un-spliceable (US) mirtron constructs. A single G-A mutation was made in the 5' splice site for each mirtron (highlighted). **(b)** Splicing efficiency of mirtrons and their US equivalents is represented by EGFP fluorescence and **(c)** their **s**ilencing against pAtaxin-7-Q100-mCherry is represented by mCherry fluorescence. In each case the NAD control is normalised to 1. HEK-293 cells were harvested 48 h post-co-transfection with mirtrons and Ataxin-7 plasmids. 250 µg total protein was loaded and fluorescence was measured using a plate reader. Background fluorescence (un-transfected cells) was subtracted from all readings. Asterisks (*/***) indicate $p<0.05$ and $p<0.0005$, respectively, compared to NAD, one-tailed t-test. **(d)** Sequence and structure of shRNA variants of mirtrons with the same splice-inhibiting G-A mutation in each 5' splice site as in (a) (highlighted). **(e)** Relative Luciferase ratios of matched targets by shRNA equivalents of mirtrons. Target constructs used were Tar1 for shR-1cN and Tar2 for shR-18. Asterisks (***) indicate $p<0.0005$ compared to shR-NS against the same target (one-tailed t-test). For (b, c and e), N=6 and error bars represent standard deviation.

3.3.2 Synthetic mirtrons are processed to mature miRNAs with biased strand selection

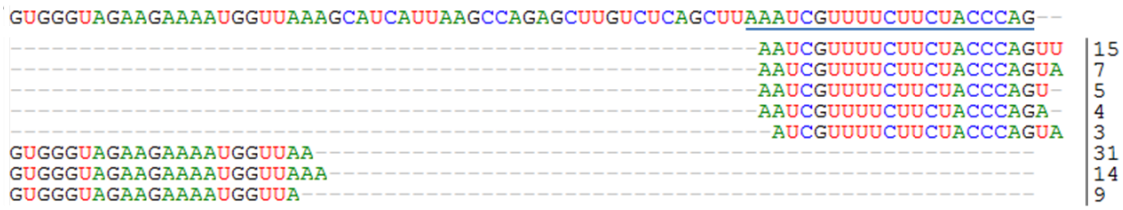
Biased selection of the intended guide strand of an artificial miRNA by the RISC is desirable since silencing of unintended targets could be caused by selection of the passenger strand. Indeed, mature miRNAs may be derived simultaneously from either arm of mirtrons (Schamberger et al. 2012). Mouse miR-1224 (the parent mirtron structure upon which the synthetic Ataxin-7 mirtrons are based) produces mature miRNA largely from the 5′ strand, but also a very small proportion from the 3′ arm (1/1350, according to miRBase.org Release 19) (Chiang et al. 2010; Kozomara and Griffiths-Jones 2011). Synthetic mirtrons were incorporated into this structure with base mismatches designed to bias selection of the guide strand, as described earlier (section 3.2). High-throughput sequencing was performed on ~18- to 25-nt RNA extracted from HEK-293 cells transfected with the Ataxin-7 mirtrons.

Of the sequencing reads which aligned to mirt-1cN, 98.6% were derived from the 5′ (guide) strand (Figure 3-7 a and c). Bases absent from the 5′ end may be explained by Dicer trimming, which can occur for pre-miRNAs which do not possess the canonical 2-nt 3′ overhang (Ando et al. 2011). For mirt-18, for which the intended guide strand is on the 3′ arm, only 38.6% of total reads derived from the guide strand (Figure 3-7 b and c). Nonetheless, this indicates that the mirtron design process substantially reversed the strand bias from the endogenous miR-1224 on which it was based, but this could perhaps be further optimised. Despite this, the comparable silencing ability of the two mirtrons (Figure 3-3c and Figure 3-4b) further confirms that the mature species derived from mirtrons can be potent even at low levels, as shown previously (Seow et al. 2012).

a mirt-1cN



b mirt-18



c

		Total	Guide	Passenger	Major product	Correct 5' base	Correct seed
mirt-1cN	Reads	1319	1300	19	592	1285	1273
	% total		98.6%	1.4%	44.9%	97.4%	96.5%
	% guide				45.5%	98.8%	97.9%
mirt-18	Reads	88	34	54	31	0	0
	% total		38.6%	61.4%	35.2%		
	% guide				91.2%		

d mirt-18-US

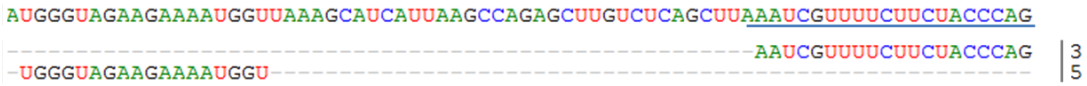


Figure 3-7 – Deep sequencing reads obtained from HEK-293 cells transfected with mirt-1cN, mirt-18 or mirt-18-US. RNA was extracted 48 h post-transfection. Sequences that were aligned to mirtrons sequences were collected and are shown in alignment with the appropriate full mirtron sequence. Bases are coloured according to identity and the designated guide strands are underlined. The number of reads detected is listed on the right,

indicating the relative abundance of each mature species. **(a)** Reads aligned to mirt-1cN. **(b)** Reads aligned to mirt-18. **(c)** Reads in (a) and (b) are summarised after post-Dicer modifications were removed. The number of reads is given for sequences derived from the guide and passenger strands, the most prevalent product (Major product), those starting with the correct 5' base and those containing the correct seed region. These are also given as percentages of the total number of reads (% reads) and of the guide strand-derived reads (% guide). **(d)** Reads aligned to mirt-18-US.

The addition of certain residues to the 3' end of the mature mirtron species appears to follow Dicer processing, as seen in some canonical miRNAs (Wu et al. 2009b) and other synthetic mirtrons (Seow et al. 2012). Adenylation of the mature species may promote its stability (Katoh et al. 2009; Burns et al. 2011) and uridylation may promote its degradation (Fernandez-Valverde et al. 2010; Wei et al. 2012). Therefore, processing of the mature strands of mirtrons may influence their turnover, affecting the number of reads obtained. In the case of mirt-18, 88.2% (30/34) of guide strand reads contained an additional 3' U (and only 32.4% contained an additional A) (Figure 3-7b), which may indicate that they are rapidly turned over and therefore present at low frequency. Interestingly, the passenger species for mirt-18 all ended in (apparently templated) A residues, which perhaps causes increased stability. These effects could contribute to the greater detection of passenger compared to guide strands. For mirt-1cN, at least 9.6% (125/1300) of 5' end (guide strand) reads were uridylated and 3.3% adenylated. Many other reads also ended in A or U, but matched the mirtron DNA sequence, such that true Dicer products were indistinguishable from modified sequences.

Rapid degradation of the guide strand may explain the low total number of reads generated for mirt-18 (88) compared to mirt-1cN (1319). Although a lower number of total reads was expected for mirt-18 given its lower splicing efficiency, the difference was surprisingly large.

Alternatively, these data may indicate that the spliced product is processed slowly. Sequences longer than ~24 bases were not detected by this method, so accumulation of the pre-miRNA could not be observed. Nonetheless, silencing activity for mirt-18 was comparable to mirt-1cN, and low miRNA levels could be advantageous in avoiding saturation of the endogenous miRNA pathway (Giering et al. 2008).

Deep sequencing results are also informative for the processing accuracy of mirtrons. The post-Dicer modifications were removed and sequences were regrouped (Figure 3-7c). The predominant mirt-1cN-derived species detected (44.9%) was 22 bases long, as expected from Dicer cleavage (Wu et al. 2009b). Crucially, 98.8% of reads derived from the 5' end of mirt-1cN (which is defined by splicing) began at the correct 5' base. The vast majority (97.9%, excluding those with errors) therefore possessed the intended seed region (bases 2-8 of the mature miRNA), the most influential region for target silencing (Haley and Zamore 2004). Although the 3' end of the mature species is not critical for target binding, some mature species were up to 24 bases in length; therefore, more target-complementary bases should perhaps be included in the design of 5'-derived mirtrons, to prevent mismatches between the target mRNA and the last few bases of the mature products, which may slow their activity.

The 5' end of the intended mature species for mirt-18 is defined by Dicer processing, not splicing. The predominant guide strand species (31 reads, 35.2% of the total, Figure 3-7c) began one base 3' of the designated base, and hence was only 20 bases in length (excluding post-Dicer additions, Figure 3-7b). However, the seed region should still be completely complementary to the target site.

The US constructs were also investigated by deep sequencing. As expected, no reads were detected for mirt-1cN-US. However, a total of eight reads were derived from mirt-18-US

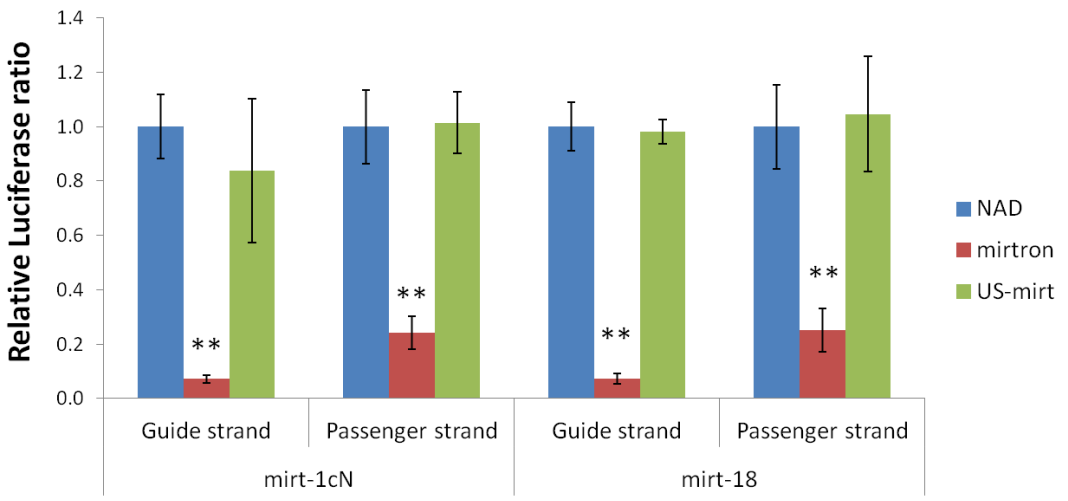
(Figure 3-7c). This may explain the significant level of silencing achieved by this construct. All of the reads derived from the 5′ end began one base 3′ of the intron, indicating that these were likely not produced by splicing, and may instead have been processed by the microprocessor.

3.3.3 Investigating the potential off-target effects of mirtrons

To investigate whether the mirtron passenger strands cause off-target effects, sequences complementary to the passenger strand for each mirtron were cloned into the 3′ UTR of *Renilla* Luciferase of dual Luciferase reporter constructs. For both mirtrons, this target was significantly silenced (Figure 3-8a), but to a lesser extent than guide strand-directed silencing. This is likely to be explained by the presence of processed passenger strands, as shown in the deep sequencing results (Figure 3-7 a-c). The Luciferase assay is very sensitive, so the ~75% silencing elicited by the passenger strand may not represent the capability for substantial silencing of endogenous transcripts, particularly given that no human transcripts were found to contain sites completely matched to the passenger strand for either mirtron (Figure 3-8b). However, many genes contain seed-matched targets, which may be sufficient for silencing and could be further investigated. However, the off-target effects of the intended mirtron guide strands should be the primary focus for investigation, given their greater efficacy in the Luciferase assay. Although mirtron target sites were selected only if they were not present in other human transcripts, nonetheless there were a number of transcripts with “significant homology” according to the BLAST search. In addition, some genes possess multiple seed-matched mirtron targets in their 3′ UTR (Figure 3-8b).

In this assay, no significant silencing was directed by either strand of the US mirtrons (Figure 3-8a), somewhat contrary to the earlier result against full-length Ataxin-7 (Figure 3-6c).

a



b

Guide strand	Transcripts with complete†/significant homology to full target site		Genes with seed-matched 3'UTR targets	Genes possessing 2 seed-matched 3'UTR targets	
18	3	ATXN7†, FILIP1L, AIPL1	399	4	CAMTA1, ENAH, ITPKB, TIMP3
18*	1	TMEM87B	63	0	
1cN	2	ATXN7†, FGF11	95	1	HIF1AN
1cN*	2	CCDC141, COL4A3BP	165	0	

Figure 3-8 – Potential off-target effects of the guide and passenger strands of artificial mirtrons. **(a)** Mirtrons and Luciferase plasmids were co-transfected into HEK-293 cells and harvested 48 h later. For each mirtron, two seperate Luciferase targets were used, one perfectly matching the guide strand and the other matching the passenger strand. The un-spliceable equivalents (US-mirts) were also included as a negative control. *Renilla*:firefly Luciferase ratios were calculated and normalised to the NAD control in each case. N=6 (two independent experiments). Error bars represent standard deviation. Asterisks (**) represent $p<0.005$ (one-tailed t-test). **(b)** Bioinformatic searches for potential off-targets of synthetic mirtrons and their passenger (*) strands. Transcripts containing sequences with "significant homology" to the matched target site were identified using BLAST (NCBI). Those with complete homology are indicated (†). Human genes with seed-matched targets in their 3' UTR were identified using TargetScan; those which possess two or more seed-matched targets are listed. Searches were carried out based upon the 22 bp product from each strand.

3.3.4 Reducing hairpin base-pairing in mirt-18 improves splicing but compromises silencing ability

The low splicing efficiency of mirt-18 compared to NAD and mirt-1cN (Figure 3-5a and b) could be attributable to the extensive complementarity between the strands of the hairpin (Figure 3-1). The resulting tight secondary structure of the mirtron may mask the splice sites or branch point. Hence, making changes to the sequence of the passenger strand to weaken the secondary structure could improve splicing efficiency and potentially increase production of mature miRNA. Particularly when incorporated into a functional replacement gene and driven by a weaker promoter, it is important that splicing is efficient and accurate, to produce enough functional mature miRNA and protein, and avoid translation of un-spliced or mis-spliced products.

Two variants of mirt-18 were created, with passenger (5') strand alterations designed to reduce the overall number of base pair matches, resulting in less negative ΔG values (Figure 3-9a). Changes could be made to this strand because sequence requirements for splicing are not stringent at the 5' end of the intron after the first five bases (Lim and Burge 2001). Base pairing at the 5' end of the guide strand was kept weaker than the 3' end, to promote correct strand selection by the RISC.

EGFP levels indicated that splicing efficiency was indeed increased for both variants compared to the original mirtron ($p<0.005$, Figure 3-9b and c) and there was no longer a significant difference with NAD (Figure 3-9c). Silencing activity against Ataxin-7-Q100-mCherry, however, was clearly compromised for both variants compared to the original mirtron ($p<0.005$, Figure 3-9b and d). Only approximately 25% silencing was achieved ($p<0.05$). RT-PCR did not show a decrease in the amplified un-spliced product, but there was

an increase in spliced product (Figure 3-9e). Together, these data show that weakening the hairpin structure of mirt-18 increased its processing as an intron but lessened its silencing activity, perhaps due to poorer recognition or processing of the pre-miRNA by Dicer or other downstream factors. This could be investigated by testing for affinity with Exportin-5, Dicer and AGO2, and using Northern blotting to compare the levels of mature miRNA.

a

ΔG

mirt-18

```
g                   a    atcat     c
 tgggtagaagaaaatggtt aagc     taagc a
 ||||||||||||||o|o|| ||||     |||||
 aCCCATCTTCTTTTGCTAA ttcg     gttcg g
G                   A    actct     a
```

-36.9

mirt-18v2

```
g        t     g   aa    atcat     c
 tgggtaga ggaaa ggt  aagc     taagc a
 |||||||| |o||| |o|  ||||     |||||
 aCCCATCT CTTTT CTA  ttcg     gttcg g
G        T     G   AA    actct     a
```

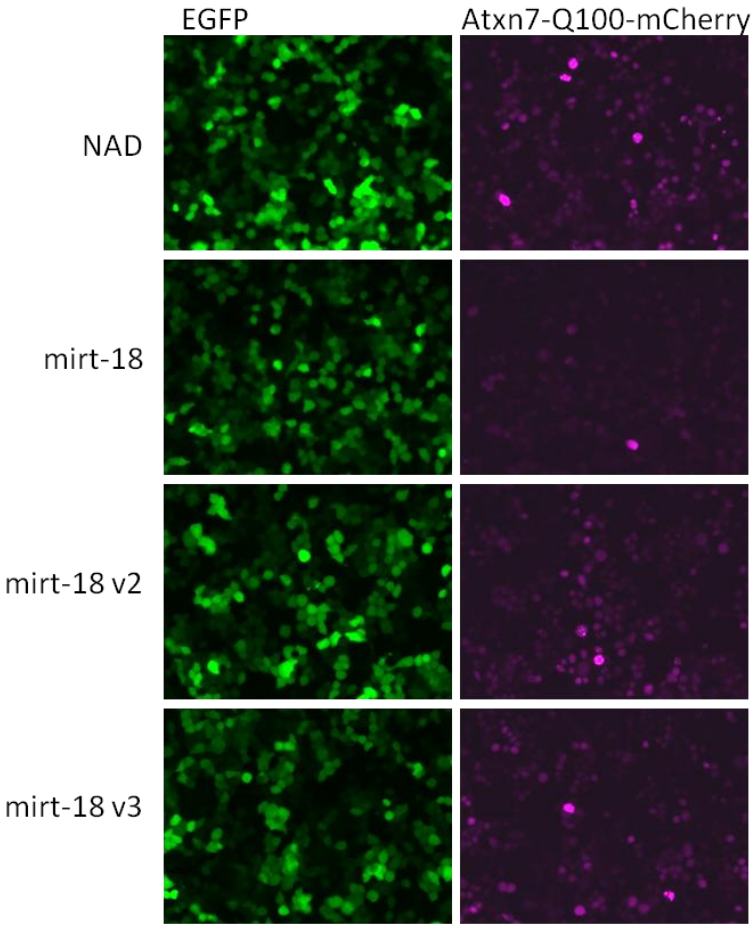
-28.5

mirt-18v3

```
g     t       ta    a    atcat     c
 tgggt gaaggaa  ggtt aagc     taagc a
 ||||| ||||o||  |o|| ||||     |||||
 aCCCA CTTCTTT  CTAA ttcg     gttcg g
G     T       TG    A    actct     a
```

-26.4

b



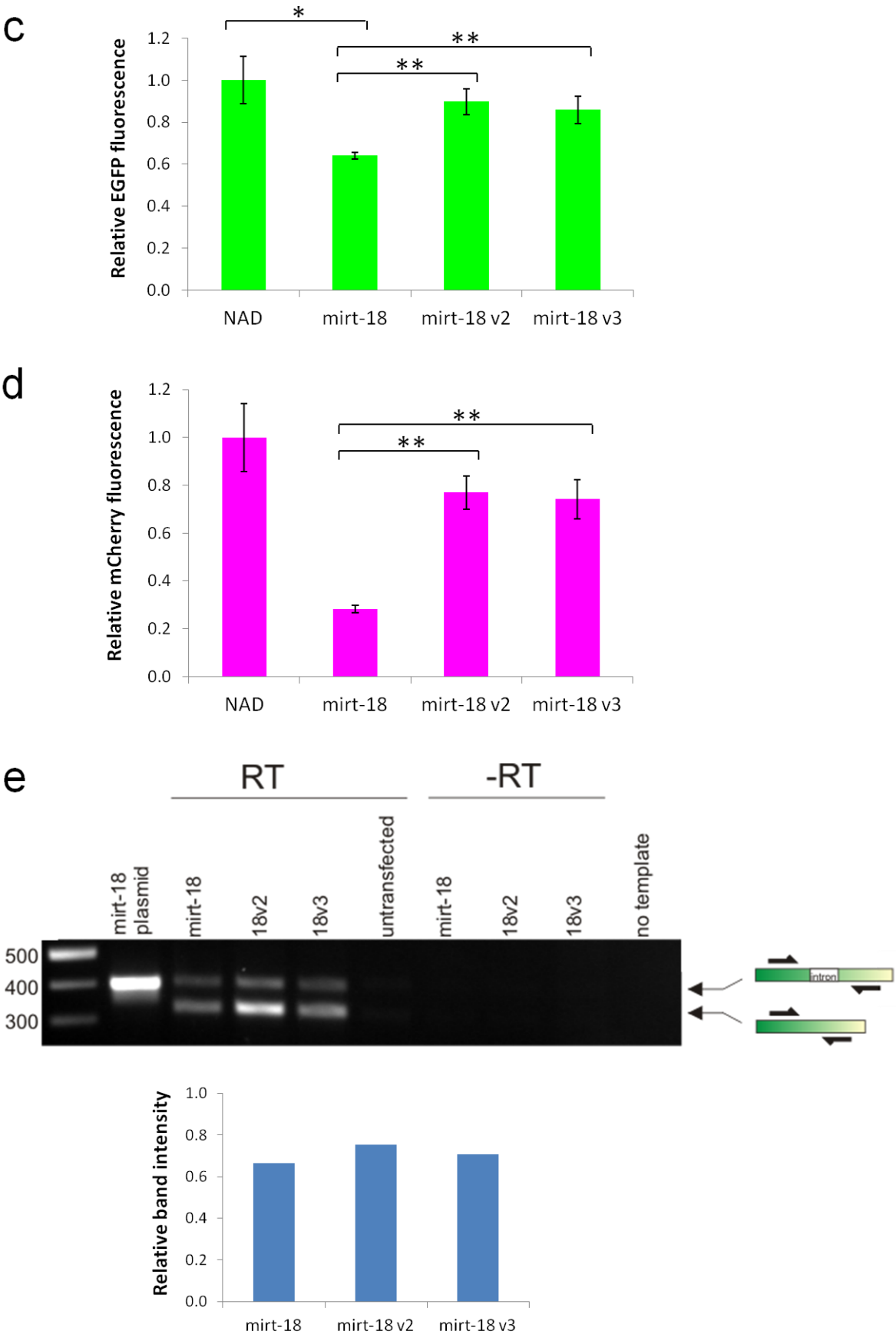

c
Relative EGFP fluorescence
1.2
1.0
0.8
0.6
0.4
0.2
0.0
*
**
**
NAD
mirt-18
mirt-18 v2
mirt-18 v3
d
Relative mCherry fluorescence
1.2
1.0
0.8
0.6
0.4
0.2
0.0
**
**
NAD
mirt-18
mirt-18 v2
mirt-18 v3
e
RT
-RT
mirt-18 plasmid
mirt-18
18v2
18v3
untransfected
mirt-18
18v2
18v3
no template
500
400
300
intron
Relative band intensity
1.0
0.8
0.6
0.4
0.2
0.0
mirt-18
mirt-18 v2
mirt-18 v3

Figure 3-9 - Design and characterisation of variants of mirt-18. (a) Sequence and structure of two variants of mirt-18, with altered bases highlighted and the guide strand shown in red uppercase. ΔG values are also indicated (kcal/mole) according to the RNAfold algorithm (Vienna RNA 2.0 package (Lorenz et al. 2011), available at http://rna.tbi.univie.ac.at/). **(b)** Fluorescent images of HEK-293 cells 48 h after co-transfection with pEGFP-mirtron constructs and pAtaxin-7-Q100-mCherry, indicating splicing and silencing efficiencies, respectively. **(c)** EGFP and **(d)** mCherry fluorescence for (b) was quantified, background fluorescence (un-transfected cells) was subtracted from all readings and NAD control was normalised to 1. 250 µg total protein was loaded and fluorescence was measured using a plate reader.. N=3. Error bars represent standard deviation. Asterisks (*/**) represent p<0.05/0.005, respectively (one-tailed t-test). Additionally, in (d), measurements for mirt-18 and both variants were significantly lower than NAD (p<0.05, one-tailed t-test). **(e)** RT-PCR was carried out for mirt-18 variants with intron-spanning primers to detect the relative abundance of un-spliced and spliced products, as indicated by the schematic diagram. Splicing efficiency for each mirtron was quantified based upon the band intensities by calculating the proportion of spliced product to the total amount of transcript (spliced + un-spliced).

3.4 Multiple mirtrons may be delivered from a single transcript

It may be necessary to co-deliver multiple miRNAs directed against different sites within a target gene to ensure efficient silencing or to avoid single nucleotide polymorphisms (SNPs) and ensure maximum patient coverage with a single treatment. To achieve this, several mirtrons could be co-delivered as separate introns within a single gene (Sibley et al. 2012). To investigate this strategy for mirtrons targeting Ataxin-7, the two most effective mirtrons were cloned within one EGFP construct as two separate introns. All four permutations were made: including two copies of each mirtron and combinations of both in each position.

EGFP fluorescence data indicated that splicing levels of the double-mirtron constructs (18-18 and 1cN-1cN) were comparable with, but slightly lower than the single mirtron equivalents (Figure 3-10a). This was expected given that splicing is an efficient process that can extract

multiple introns concurrently from a single transcript, but may take longer the greater the number of introns. When two *different* mirtrons were combined, however, splicing levels did not appear to be entirely limited by the least efficient mirtron (mirt-18). It is possible that the more well-recognised intron (mirt-1cN) can aid the recruitment of the spliceosome to mirt-18 (Crabb et al. 2010).

Intron-spanning RT-PCR was carried out to investigate the splicing efficiency of each mirtron individually (Figure 3-10b and c). Corresponding to the EGFP data, NAD was spliced completely from both positions. However, position 2 displayed lower splicing efficiency for both mirtrons, indicating it is less suited for mirtron splicing. Unfortunately, the PCR spanning intron 2 led to production of two bands for the un-spliced transcript (Figure 3-10b). Both were sequenced and the smaller (~260 bp) band corresponded to the un-spliced sequence. The sequence of the larger (~300 bp) band was not clear. Although its ends corresponded to the expected product, very little of the sequence could be determined in the position of the mirtron. This band was not produced from the plasmid templates, nor when RT-PCR was carried out on cells transfected with the plasmids containing only an intron in Position 1, ruling out a non-specific product. Hence the extra band may be the correct product, but migrating more slowly due to complex structure formed through tight folding of the mirtron hairpin. This could be confirmed on a denaturing polyacrylamide gel.

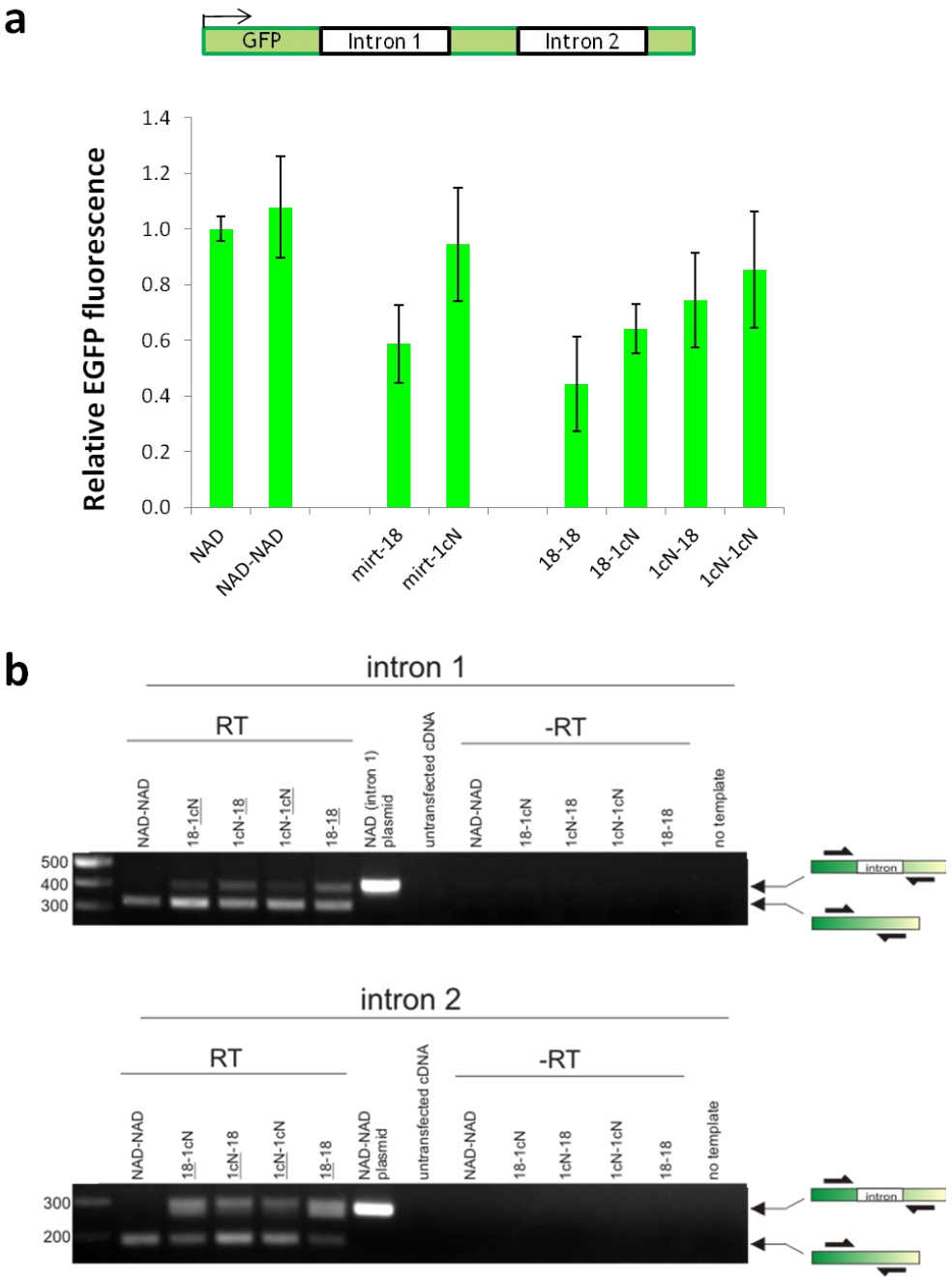
RT-PCR indicated that mirt-1cN was spliced as efficiently regardless of whether the other intron was the same or different (75-80%; Figure 3-10c). This was slightly reduced compared to when no other mirtron was present (95%, Figure 3-5b), though these were not compared in the same RT-PCR reaction. As expected from the EGFP fluorescence data, there was a small increase in splicing of mirt-18 when it was paired with mirt-1cN rather than mirt-18, particularly in site 2 (45% vs 30%). This may be explained by mirt-1cN being better

recognised as an intron and recruiting the spliceosome to the transcript, which moves in the 5'-3' direction and so could aid the splicing of mirt-18 when it is further downstream.

Quantitative RT-PCR (qPCR) was then used to investigate the production of mirtron pre-miRNA (hairpin) species (Figure 3-10d). When two copies of the same mirtron were expressed from the same transcript, the level of pre-miRNA almost doubled, as expected from the splicing data. When the mirtrons were different, however, a decrease in expression of both pre-miRNAs was detected. This also supports the findings from Figure 3-10 a-c, that expression of two mirtrons from a single transcript reduces their relative expression by decreasing the splicing efficiency. However, it does not suggest increased splicing of mirt-18 when combined with mirt-1cN. Interestingly, this assay produced higher Ct values (relative to GAPDH) for the mirt-1cN hairpin compared to mirt-18, indicating lower abundance. Taken with the deep sequencing data (Figure 3-7), this indicates that the mirt-18 pre-miRNA may accumulate, being processed slowly into mature species.

The multiple-mirtron constructs were tested for their silencing ability against Ataxin-7-Q100-mCherry (Figure 3-11a), as well as the Luciferase targets containing either the mirt-1cN or mirt-18 target site (Figure 3-11b and c, respectively). In either case, no significant improvement can be seen on the high level of silencing already achieved by a single mirtron. A difference may perhaps be seen if delivered under a weaker promoter. However, it is clear that combining multiple mirtrons into one gene does not interfere with their action in these assays, despite the reduced levels of pre-miRNA production seen in Figure 3-10d.

Delivering multiple mirtrons concurrently is potentially advantageous in terms of target silencing, but also increases the risk of off-target effects resulting from silencing of other seed-matched targets of the guide or passenger strands. It would therefore be important to investigate this before further development of this approach.


a
GFP
Intron 1
Intron 2
Relative EGFP fluorescence
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
NAD
NAD-NAD
mirt-18
mirt-1cN
18-18
18-1cN
1cN-18
1cN-1cN
b
intron 1
RT
-RT
NAD-NAD
18-1cN
1cN-18
1cN-1cN
18-18
NAD (intron 1) plasmid
untransfected cDNA
no template
500
400
300
intron
intron 2
NAD-NAD plasmid
200
c
Relative band intensity
Intron 1
Intron 2
18 (18)
18 (1cN)
1cN (1cN)
1cN (18)

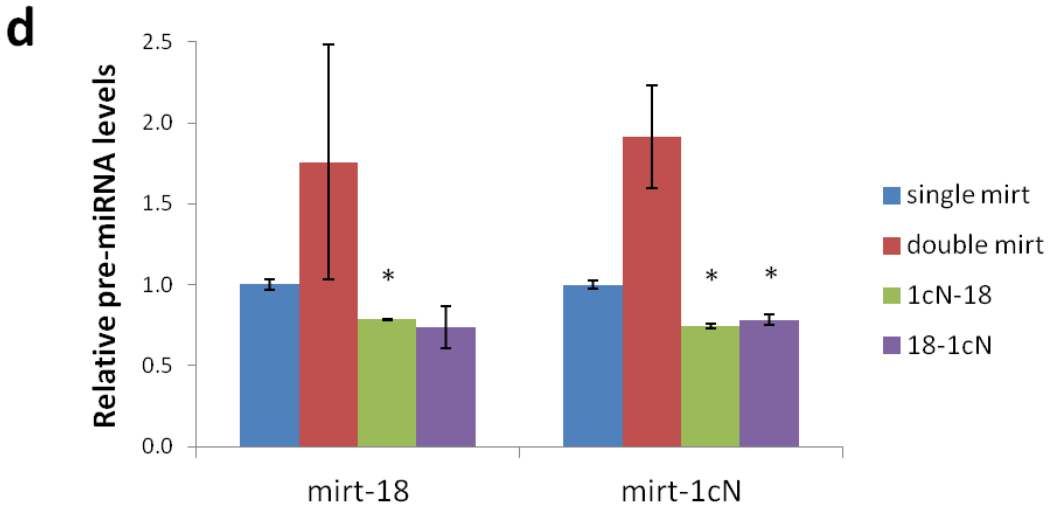


Figure 3-10 - Mirtron production from multiple-mirtron constructs. (a) Quantification of EGFP fluorescence in HEK-293 cells compared to NAD control (normalised to 1). For the multiple-mirtron constructs, the first mirtron given in the name is Intron 2 according to the schematic. "NAD-NAD" has the NAD intron in both sites. Cells were harvested 48 h post-transfection, 100 µg total protein was loaded and fluorescence was measured using a plate reader. Background fluorescence (un-transfected cells) was subtracted from all readings. N=6. Error bars represent standard deviation. **(b)** RT-PCR was carried out on RNA extracted from transfected HEK-293 cells, using primers spanning Intron 1 (top image) or Intron 2 (bottom image). The mirtron in the position being tested is underlined. **(c)** Quantification of splicing efficiency in (b) for each mirtron in each position, with the mirtron in the other position given in brackets. The average band intensity for the spliced product is given as a proportion of the total (spliced + un-spliced). For Intron 2, an average of intensities across the double band was taken. **(d)** qPCR was used to detect mirt-18 and mirt-1cN pre-miRNA in RNA extracted transfected HEK-293 cells. Mirtron expression from the multiple-mirtron constructs was compared to the original pEGFP-mirt constructs (single mirt), normalised to 1. "Double mirt" indicates the construct containing two copies of the mirtron being detected. N=2. Error bars represent standard error. Equal amounts of cDNA were used but the quantities were not normalised to an internal control gene. Asterisks (*) represent $p<0.05$ compared to the appropriate single mirt (one-tailed t-test).

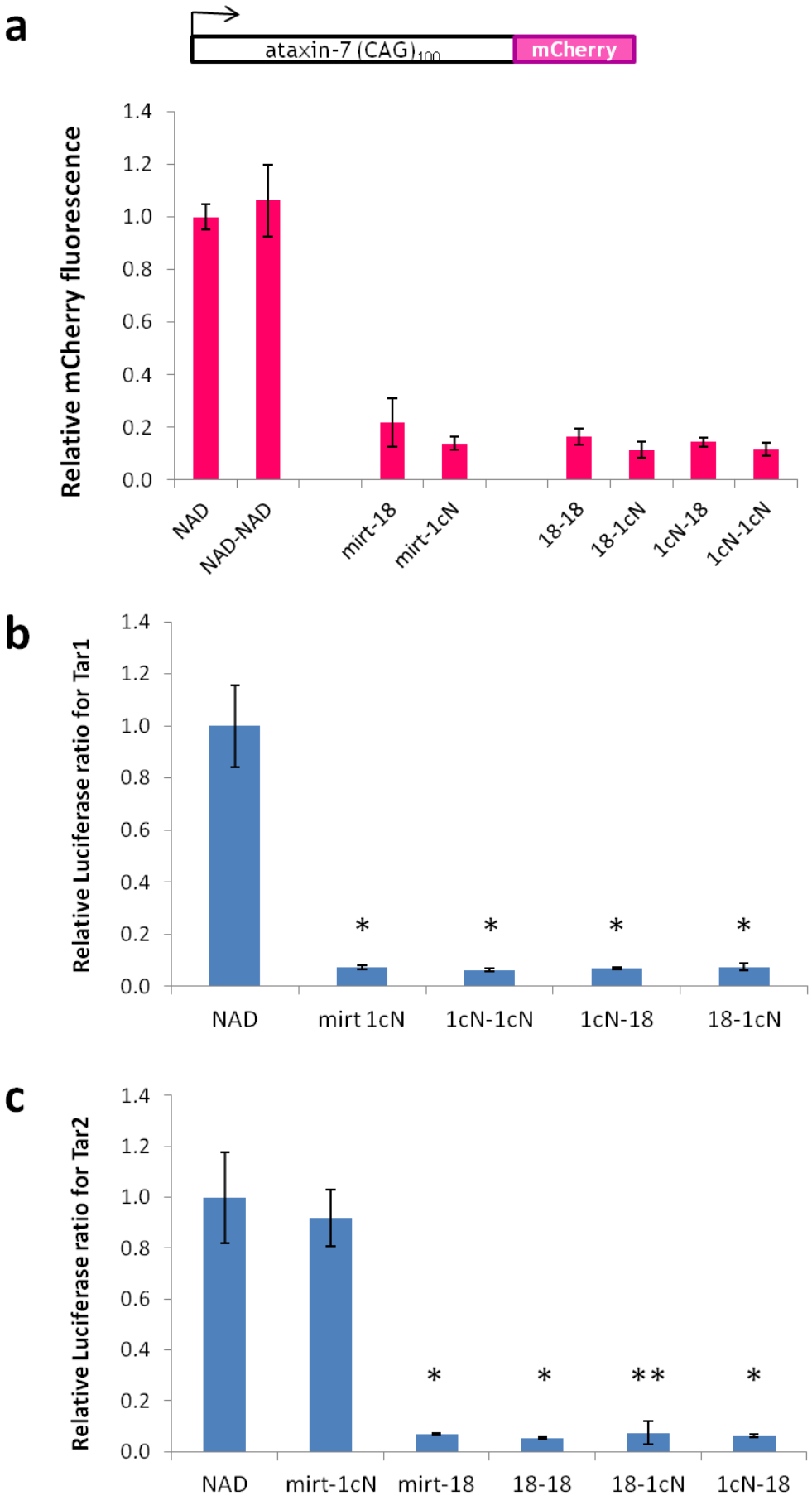

a
ataxin-7 (CAG)100
mCherry
Relative mCherry fluorescence
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
NAD
NAD-NAD
mirt-18
mirt-1cN
18-18
18-1cN
1cN-18
1cN-1cN
b
Relative Luciferase ratio for Tar1
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
*
*
*
*
NAD
mirt 1cN
1cN-1cN
1cN-18
18-1cN
c
Relative Luciferase ratio for Tar2
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
*
*
**
*
NAD
mirt-1cN
mirt-18
18-18
18-1cN
1cN-18

Figure 3-11 – Silencing ability of multiple-mirtron constructs against fluorescent or Luciferase targets in HEK-293 cells. (a) Relative mCherry fluorescence, indicating the silencing ability of each mirtron and multiple-mirtron against mCherry-tagged mutant Ataxin-7 (indicated in the schematic, top). Multiple-mirtron constructs are named as in Figure 3-10a. Protein was harvested 48 h post-transfection and fluorescence was measured using a plate reader with 250 µg total protein from each sample. Background fluorescence (un-transfected cells) was subtracted from all readings and control values (NAD intron or non-specific shRNA, shR-NS) were normalised to 1. N=6 (two independent experiments) and error bars represent standard deviation. mCherry levels were significantly reduced (p<0.0005) for all mirtron constructs (not including NAD-NAD), compared to NAD (one-tailed t-test). Multiple-mirtrons were also tested against dual Luciferase constructs containing the target site for **(b)** mirt-1cN and **(c)** mirt-18. N=3. Error bars represent standard deviation and asterisks (*/**) represent p<0.05/0.005 respectively, compared to NAD (one-tailed t-test).

3.5 Artificial mirtrons can silence endogenous Ataxin-7

It is vital to confirm that the artificial mirtrons are also capable of silencing endogenously expressed Ataxin-7, rather than over-expressed constructs which are transfected concurrently. To investigate this, mirtrons were transfected into a number of cell lines and levels of *ataxin-7* mRNA were measured by qPCR. Protocols were also developed for protein immunostaining and Western blotting.

3.5.1 Silencing in HEK-293 cells

Mirtrons were initially tested in HEK-293 cells due to the ease of transfection (Figure 3-4b). 50-60% silencing of the endogenous *ataxin-7* transcript was achieved by mirtrons and their shRNA equivalents (Figure 3-12). A range of plasmid concentrations was tested, yet no dose response was seen. Although transfection efficiency was good for the conditions tested, plasmid copy number per cell likely increased at higher concentrations. Therefore, it appeared that the plasmid-based continuous expression system saturated the miRNA pathway.

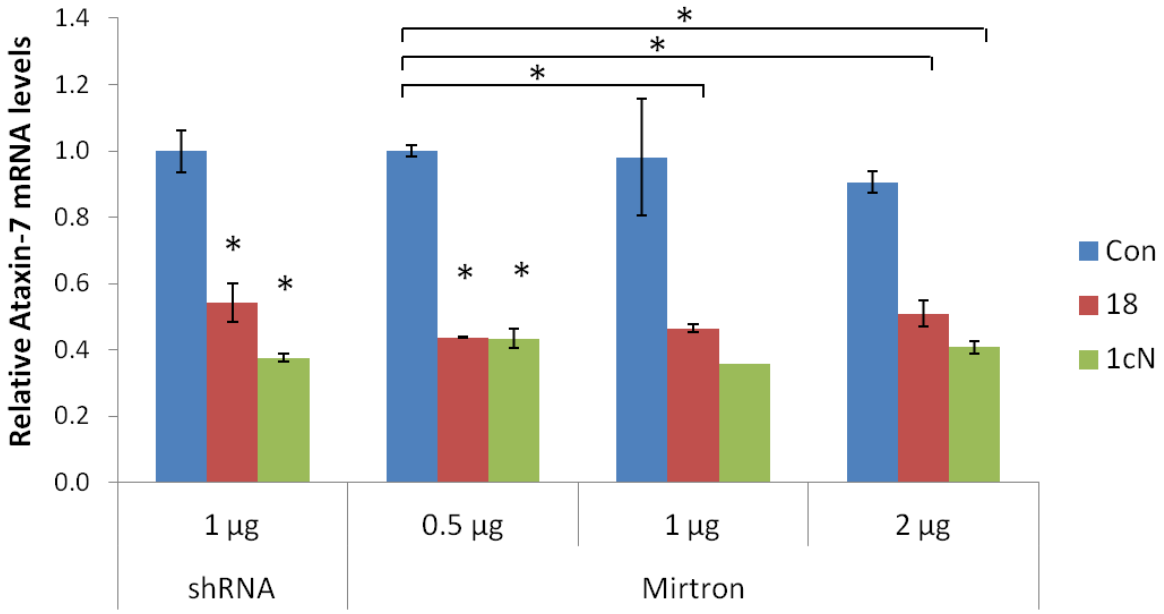


Figure 3-12 - Silencing of endogenous Ataxin-7 mRNA in HEK-293 cells. Levels of *ataxin-7* mRNA were measured by qPCR relative to the reference gene (β-actin) for each sample. RNA was harvested 48 h post-transfection with mirtrons or their shRNA equivalents Quantities represent the amount of plasmid transfected (24-well plates). The control (con) was shR-NS for shRNAs and NAD for mirtrons. Data for shRNAs was normalised to shR-NS; data for mirtrons was normalised to NAD (at 0.5 µg). Error bars represent standard error. Asterisks (*) represent $p<0.05$ compared to the shR-NS or NAD control for the same plasmid quantity (unless otherwise indicated). Experiments were carried out in biological and technical duplicate, except for mirt-1cN at 1 µg, where one technical duplicate did not produce a signal.

3.5.2 Silencing in Neuro2A cells

A mouse neuroblastoma cell line, Neuro2A (N2A), was chosen for its neuronal origin, and relative ease of transfection. Further, it is important to validate mirtrons in a mouse cell line before development towards *in vivo* application. In order to assess the expected maximum level of knockdown achievable in this cell line, detection methods were tested after siRNA treatment, because this acts more quickly (not requiring processing in the cell) and higher transfection efficiency can be observed compared to plasmids. 80-90% transfection was achieved (Figure 3-13a). Silencing of *ataxin*-7 mRNA was indeed seen in this cell line after treatment with a number of previously validated siRNAs (Figure 3-13b). However, the highest level of silencing achieved was only 60% (siR-4m), even at a high concentration (50 nM). A Western blot was then carried out on protein extracts from N2A cells 72 h after transfection (allowing an extra 24 h for protein turnover) with either siR-12 or siR-4, at two different concentrations (Figure 3-13c). However, no reduction in total protein levels could be seen. It may be that in this experiment the transfection efficiency was lower or the mRNA reduction of 60% was not sufficient for protein levels to be affected after just 3 days.

N2A cells were then transfected with mirt-18 and its shRNA equivalent. Mirt-1cN was not included because it is not complementary to the mouse *ataxin-7* sequence. Good transfection efficiency was achieved (70-80%, Figure 3-14a). 48 h post-transfection, *ataxin-7* levels appeared to be variable even with control plasmids. Compared to cells transfected with shR-NS, levels of *ataxin-7* were higher with NAD ($p<0.05$), and NAD-NAD (non-significant) (Figure 3-14b). Increased *ataxin-7* transcript levels with NAD transfection compared to shR-NS were also observed in HEK-293 cells (data not shown). The differences between the mirtron and shRNA plasmid backbones, such as the expression of EGFP, may have differing effects on the cells.

At 72 h, *ataxin*-7 expression was on average 50% lower compared to the same treatment at 48 h ($p<0.05$ for all cases except mirt-18-18 and mock treated cells). This was largely explained by an increase in levels of the housekeeping gene, perhaps indicating that the rapidly dividing N2A cells became overgrown and therefore less healthy. Despite these issues, *ataxin-7* silencing was achieved for shR-18, mirt-18 and mirt-18-18 by 40-50% after 48 h (Figure 3-14b). It again appears that treating with the multiple-mirtron construct did not improve silencing compared to a single mirtron, indicating that a factor downstream of transcription was limiting mirtron action. Normalising the data to each control revealed that the level of knockdown of endogenous *ataxin-7* (when compared to control treatments) was similar at 48 and 72 h (Figure 3-14c).

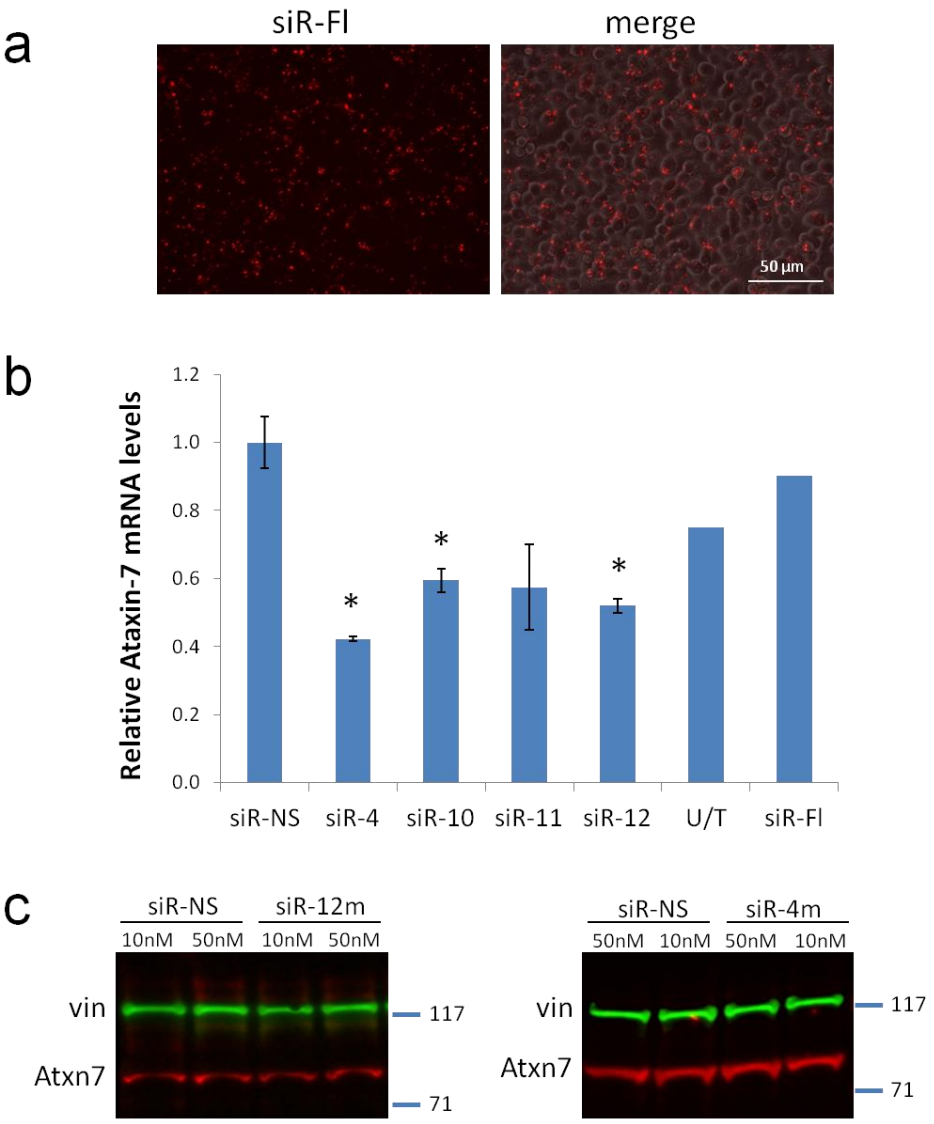


Figure 3-13 - Silencing of Ataxin-7 mRNA and protein in N2A cells after siRNA treatment. **(a)** Representative image of N2A cells 48 h post-transfection with 10 nM Cy3-tagged non-targeting siRNA (siR-Fl), also shown merged with brightfield image (merge). **(b)** *Ataxin-7* mRNA levels were measured by qPCR and normalised to internal control ubiquitin-C (UBC), 48 h after transfection with siR-NS or one of four Ataxin-7-targeting siRNAs at 50 nM. N=2, except for the untreated (U/T) and siR-Fl controls, where N=1. Error bars represent standard error. Asterisks (*) represent $p<0.05$ compared to siR-NS. **(c)** Western blots for Ataxin-7 protein probed with Ataxin-7 antibody 1-135 (labelled Atxn7) and anti-vinculin (vin), 72 h after transfection with siRNA at 10 nM or 50 nM. 25 µg total protein was loaded in each lane.

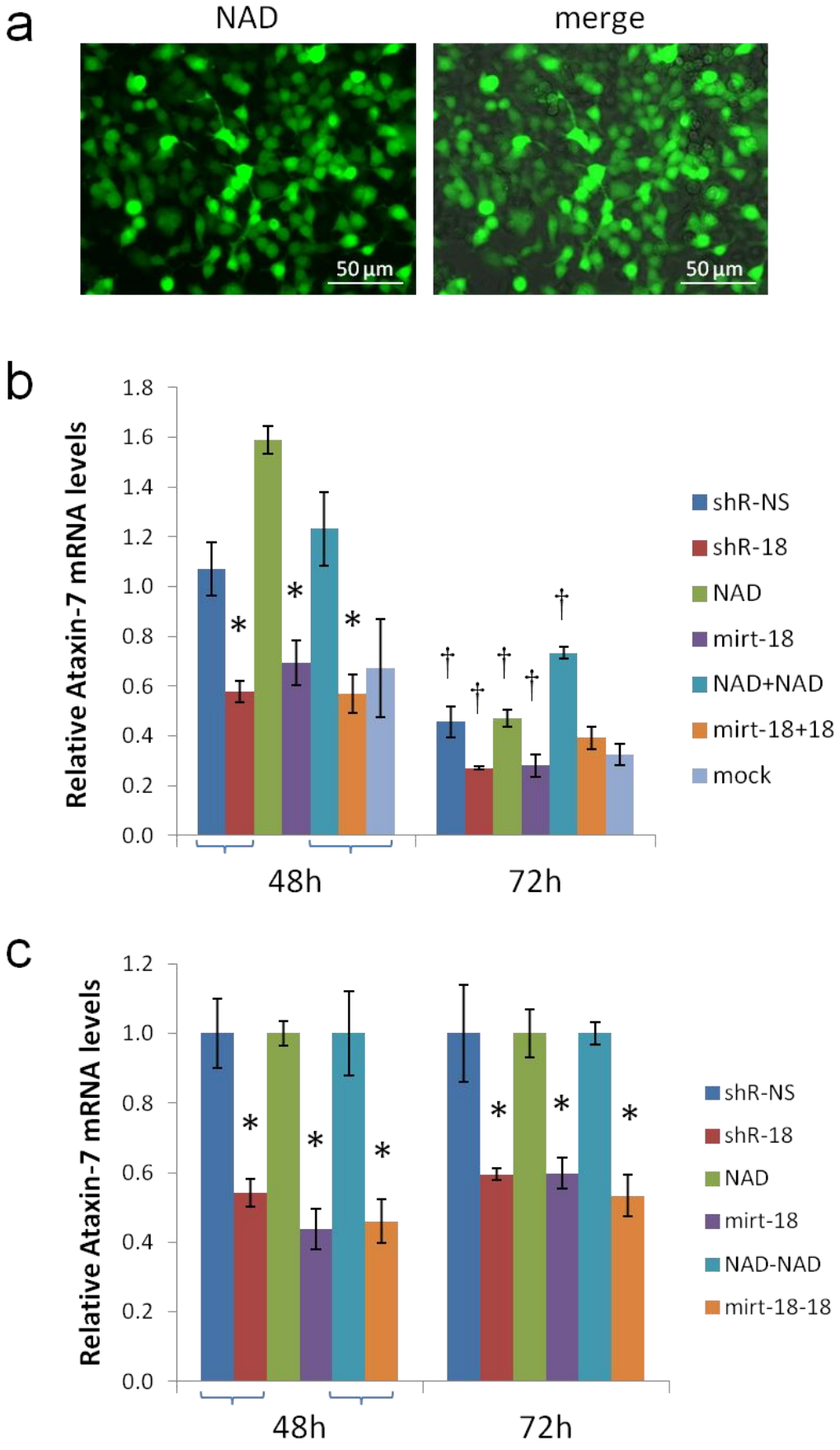

a
NAD
merge
50 µm
50 µm
b
Relative Ataxin-7 mRNA levels
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
48h
72h
shR-NS
shR-18
NAD
mirt-18
NAD+NAD
mirt-18+18
mock
c
Relative Ataxin-7 mRNA levels
1.2
1.0
0.8
0.6
0.4
0.2
0.0
48h
72h
shR-NS
shR-18
NAD
mirt-18
NAD-NAD
mirt-18-18

Figure 3-14 - Silencing of endogenous Ataxin-7 mRNA in N2A cells with mirtrons. (a) Representative image of N2A cells 48 h post-transfection with pEGFP-NAD, also shown merged with brightfield image (merge). **(b)** Levels of Ataxin-7 mRNA detected by qPCR 48 and 72 h post-transfection, relative to the internal control (UBC) for each sample. shR-NS was normalised to 1. Asterisks (*) represent $p<0.05$ compared to shR-NS (48 h); daggers (†) represent $p<0.05$ with respect to the same treatment at 48 h. **(c)** The same data is presented with all values normalised to the relevant control. Asterisks (*) represent $p<0.05$ compared to shR-NS (48 h). For both (b) and (c), experiments were carried out in biological and technical duplicate, and where indicated with brackets, results were combined from two independent experiments (i.e. N=4). Error bars represent standard error. All p-values were calculated with a one-tailed t-test.

3.5.3 Silencing in SH-SY5Y cells

Mirtrons were also tested in a human neuroblastoma cell line, SH-SY5Y, to test whether mirtrons can be effective in a relevant human cell line. Using a human cell line allowed both mirt-18 and mirt-1cN to be tested. In addition, being slower-dividing than N2A, these cells could be cultured for 96 h post-transfection. As for N2A cells, Ataxin-7 detection methods were tested first with siRNA treatment, for which 80-90% transfection was achieved (Figure 3-15a). After transfection with a positive control siR-P16 (Scholefield et al. 2009), *ataxin-7* mRNA was reduced by 50% after 48 h ($p<0.05$, Figure 3-15b). There was a general (but small) increase in *ataxin-7* mRNA levels between 48-96 h post-transfection. Silencing efficacy was reduced slightly at the later time points, perhaps due to dilution of the siRNA through cell division.

To investigate protein-level silencing, cells treated with siR-P16 were stained for Ataxin-7 protein after 96 h. siR-P16-treated cells appear to be stained more heterogeneously, with patches of more weakly stained cells compared to the control, perhaps indicating that transfected cells exhibited silencing of Ataxin-7 (Figure 3-15c). A Western blot was also carried out for SH-SY5Ys, with protein extracted 72 and 96 h post-transfection (Figure

3-15d). No reduction in protein levels could be seen, perhaps because the ~50% silencing of mRNA was not sufficient to cause a decrease in protein levels. Contrary to the general increase in mRNA levels (Figure 3-15b), there appeared to be a decrease in Ataxin-7 protein levels between 72 and 96 h post-transfection.

Using electroporation treatment, 60-70% plasmid transfection efficiency was achieved (Figure 3-16a). Positive control shRNA (shR-1) achieved 55% reduction in *ataxin-7* mRNA after 48 h (Figure 3-16b). This silencing effect was maintained at the later time point of 96 h, as expected from a sustained expression system. Mock electroporation treatment gave higher *ataxin-7* measurements compared to plasmid-transfected cells, which may be explained by the toxicity of high concentrations of DNA.

Mirt-1cN achieved a 65% silencing effect 48 h post-transfection in SH-SY5Y cells ($p<0.05$, Figure 3-16c), whereas mirt-18 achieved only 35%. Contrary to shRNA treatment, this effect was reduced at the later (96 h) time point. These effects should be confirmed by repeating the experiment, but both the low silencing elicited by mirt-18, and the decreased efficacy over time could perhaps be attributable to variations in splicing efficiency in this cell line. It is well known that splicing can be alternately regulated in different cell types (Keppetipola et al. 2012). This should be tested by RT-PCR.

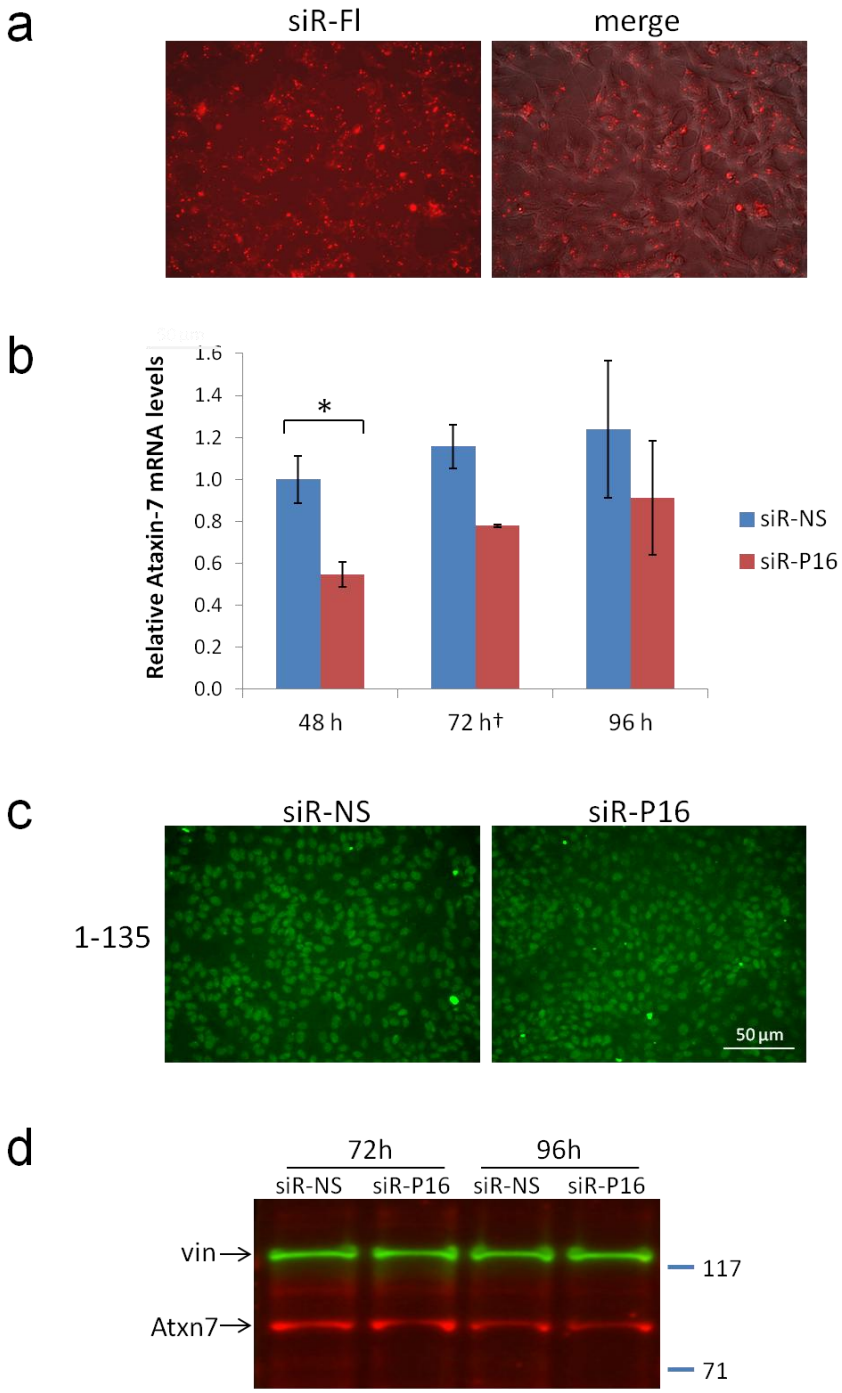


Figure 3-15 - Silencing of Ataxin-7 mRNA and protein in SH-SY5Y cells after siRNA treatment. **(a)** Representative image of SH-SY5Y cells 48 h post-transfection with 20 nM Cy3-tagged non-targeting siRNA (siR-Fl), also shown merged with brightfield image (merge). **(b)** Ataxin-7 mRNA levels were measured by qPCR and normalised to β-actin, 48, 72 and 96 h after transfection with siRNA transfected at 20 nM. N=4 (two independent experiments), except where indicated (†, N=2, one experiment). Error bars represent standard error. **(c)** Representative images of SH-SY5Y cells immunostained with Ataxin-7 antibody 1-135, 96 h after transfection with siR-NS or siR-P16 at 50 nM. **(d)** Western blots probed with anti-Ataxin-7 1-135 (labelled Atxn7) and anti-vinculin (vin). Cells were harvested 72 h and 96 h after transfection with siRNA at 10 nM or 50 nM. 25 µg total protein was loaded.

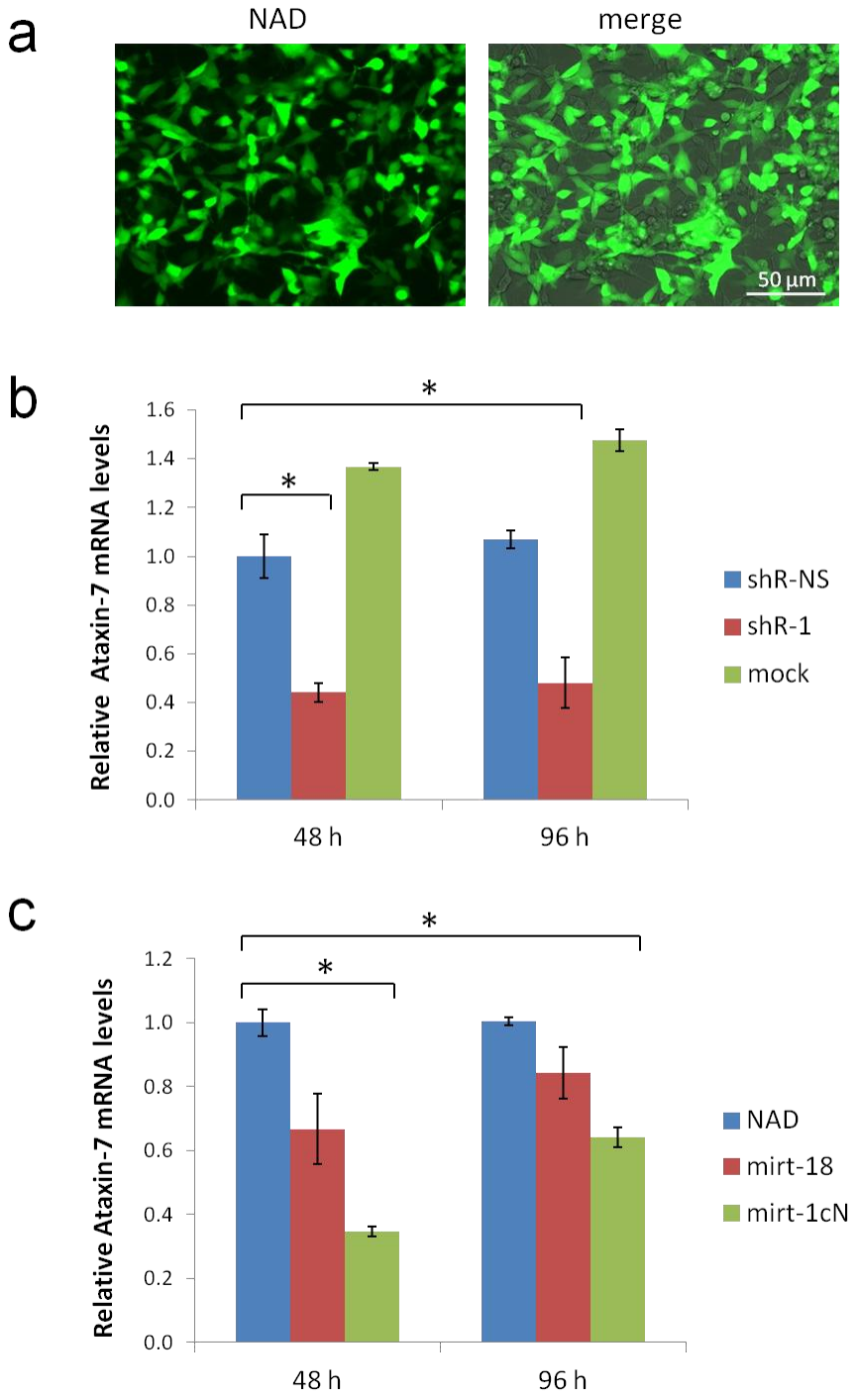


Figure 3-16 - Silencing of Ataxin-7 mRNA in SH-SY5Y cells after transfection with shRNA and mirtrons. (a) Representative image of SH-SY5Y cells 48 h post-transfection with pEGFP-NAD by electroporation, also shown merged with brightfield image (merge). For SH-SY5Y cells transfected with **(b)** shRNA and **(c)** mirtrons, Ataxin-7 mRNA levels were measured by qPCR and normalised to β-actin, 48 and 96 h after transfection. Values were all normalised to the appropriate control (NAD/shR-NS) at 48 h in each case. A mock transfection (electroporation without nucleic acid) was also included in (b). N=2. Error bars represent standard error. Asterisks (*) represent $p<0.05$ (one-tailed t-test).

3.6 Other intronic miRNAs are an alternative option for silencing Ataxin-7

Intronic miRNAs are a potential alternative to mirtrons in K&R therapy. Aagaard et al. demonstrated that an intronic miRNA cluster could be mimicked to deliver three artificial miRNAs to achieve anti-HIV activity (Aagaard et al. 2008). The naturally-occurring miR-106b cluster is found within an intron of approximately 800 bases in the MCM-7 gene. This system is interesting for K&R applications since each miRNA could target different regions within the target gene, allowing a single therapy to be applied to patients with different SNPs, and only a single effective splice site would need to be identified within the replacement gene. However, the expression of this cluster in an exogenous exonic setting has not been tested. To investigate this, miRNA cluster mimics were developed for Ataxin-7.

3.6.1 Design

The expression system used contains the whole miR-106b cluster along with the surrounding exonic regions of MCM7. This had been re-engineered such that the base of each hairpin contained unique restriction sites, allowing simple replacement with artificial miRNAs (Figure 3-17a) (Aagaard et al. 2008). To develop a miRNA cluster against Ataxin-7, mature miRNAs known to be effective against Ataxin-7 were adapted into hairpins that could be processed within the context of this cluster. It was shown by Aagaard et al. that artificial miRNAs were processed most efficiently in this context when designed to mimic the secondary structure of each of the endogenous miRNAs. This is achieved by altering the sequence of the passenger strand to incorporate mismatches in similar places. The mature miRNAs chosen were mirt-18, mirt-1cN and shR-1, and three variations of each were designed, to mimic the structure of each of the three endogenous miRNAs (example Figure 3-17b). Nine different

plasmids were therefore initially constructed, each expressing a single Ataxin-7-targeting miRNA and two irrelevant (HIV-targeting) miRNAs. The strongest hairpin for each position would then be independently identified and combined into a single intron.

3.6.2 Individual Ataxin-7 miRNAs within a miR-106b cluster mimic effectively silenced Luciferase targets

Each hairpin was tested against its corresponding Luciferase target (as previously discussed) (Figure 3-17c). This demonstrated very strong (>95%) silencing for mirt-18 and shR-1 in Position 2, and for mirt-18 in Position 3. For Position 1, however, none of the hairpins tested could silence the target more than 45%. In order to confirm these results, each plasmid was co-transfected with pAtaxin-7-Q100-mCherry and fluorescence was measured. The results reflected those achieved by Luciferase assay (Figure 3-17d). However, analysing mCherry data was problematic given the low Ataxin-7-mCherry expression, which led to some readings being close to background levels, and satisfactory data could not be obtained for the hairpins in Position 3.

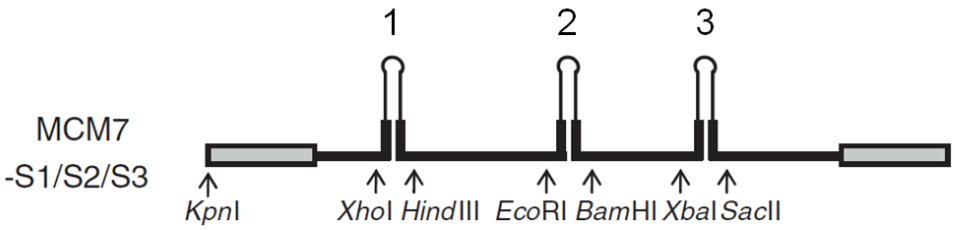
3.6.3 RT-PCR confirms that the MCM7 intron was spliced correctly

It has been shown that splicing is not essential for successful processing of canonical intronic miRNAs (Kim and Kim 2007). However, it is useful to show that splicing occurs accurately across the exons if this approach is to be applied to K&R therapy, since the spliced product would be required to perform a function. Although removal of the flanking exons was observed to reduce the silencing ability of the artificial intronic miRNAs in the miR-106b cluster, accurate splicing was not shown (Aagaard et al. 2008).

RT-PCR was therefore carried out on HEK-293 cells transfected with the original MCM7 plasmid and those containing an Ataxin-7-targeting miRNA using intron-spanning primers

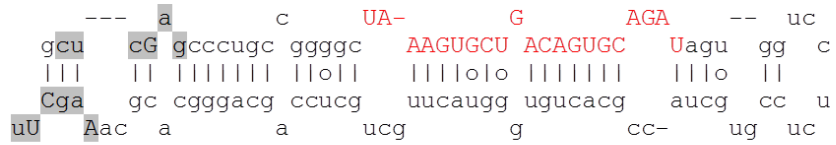
(Figure 3-17e). Encouragingly, no un-spliced band could be detected for any of the MCM7 constructs. However, such a test would also need to be carried out after placing this intron within a different exonic context for a K&R application.

a

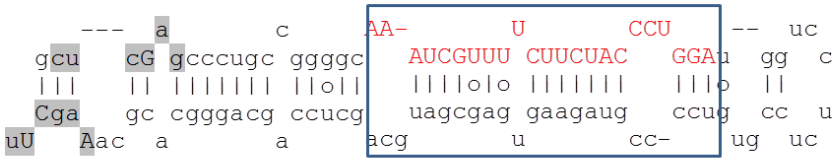


b

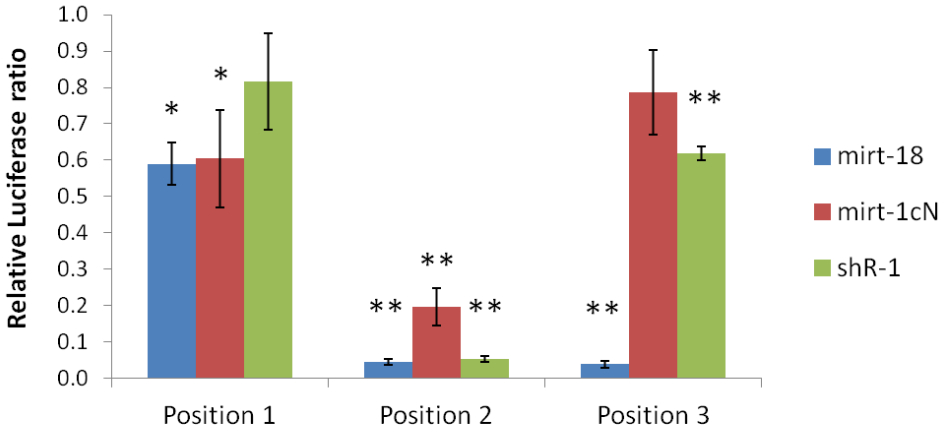
miR-106b



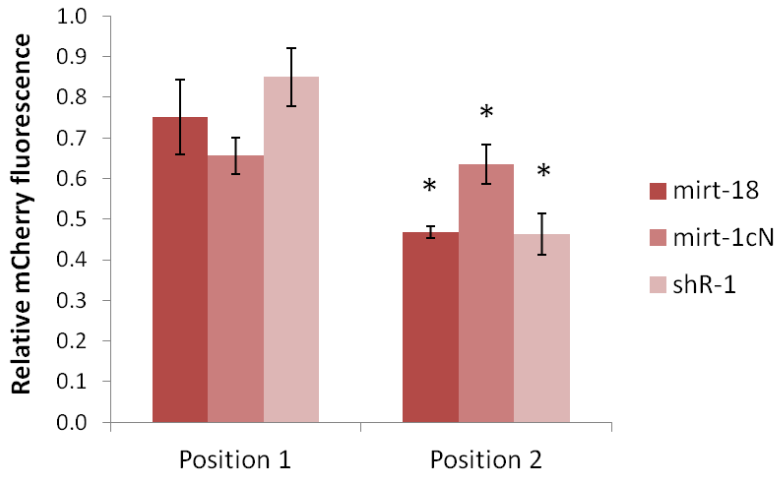
mirt-18-miR-106b mimic



c



d



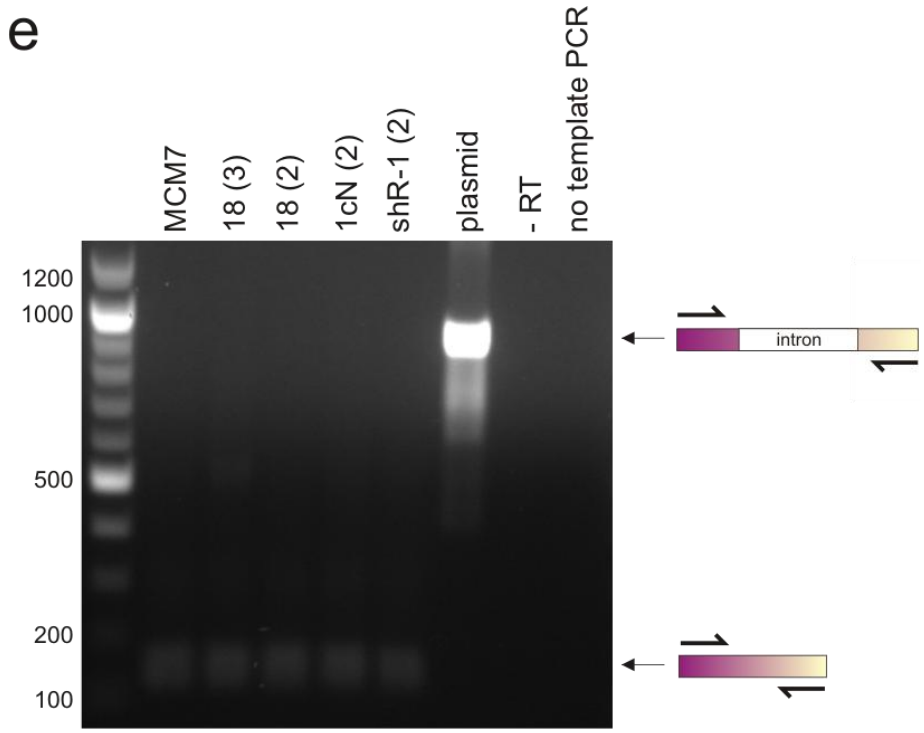


Figure 3-17 - Design and Testing of miRNA-mimics against Ataxin-7. (a) Basic structure of the miR-106b mimic (Aagaard et al. 2008) and the restriction sites used for cloning each hairpin. Initially a single targeting hairpin was cloned into each position and tested independently. Grey boxes indicate the flanking exons of MCM7. **(b)** An example of the design of miRNAs against Ataxin-7, demonstrating how the guide strand (of mirt-18 in this case) is incorporated into the miRNA hairpin (here miR-106b), mimicking its secondary structure. Guide strands are shown in red capitals. Restriction sites (XhoI and HindIII) are highlighted in grey; bases altered from the endogenous miR-106b sequence by Aagaard et al. to create these restriction sites are also in capitals. **(c)** Relative *Renilla* to firefly Luciferase ratios indicate the silencing activity of each construct against matched targets (Tar2, Tar1, or Tar3 by mirt-18, mirt-1cN and shR-1, respectively) in HEK-293 cells 48 h post-transfection. Luciferase ratio for the non-targeting control (original S1-S2-S3 plasmid, containing irrelevant miRNAs) was normalised to 1 for each target (not shown). N=3 and error bars represent standard deviation. Asterisks (*/**) represent p<0.05/0.005 respectively, compared to the non-targeting control for the same target (one-tailed t-test). **(d)** Fluorescence quantification of Ataxin-7-Q100-mCherry co-transfected with Ataxin-7-targeting cluster mimics, compared to non-targeting control (normalised to 1, not shown). 250 µg total protein was loaded and fluorescence was measured using a plate reader. Background fluorescence (un-transfected cells) was subtracted from all readings. N=3, error bars represent standard deviation and asterisks (*) represent p<0.05 compared to the non-targeting control (one-tailed t-test). **(e)** RT-PCR was carried out on cDNA from HEK-293 cells transfected with cluster mimic constructs, using a pair of intron-spanning primers. The position of the targeting hairpin is indicated in brackets. The sizes of the bands representing the un-spliced and spliced transcripts are indicated by the schematic diagrams.

3.7 Discussion

The work in this chapter focussed on developing non-allele-specific RNAi-based silencing mechanisms against Ataxin-7, in which a CAG nucleotide repeat expansion causes the polyQ disorder SCA7. Gene silencing mechanisms were chosen which could theoretically be adapted into a combination K&R therapy, by concurrently providing a "replacement" copy of the gene to prevent deficiency of Ataxin-7. K&R is an attractive therapeutic option for nucleotide expansion disorders (such as SCA7) and many other dominant conditions, where allele-specific silencing is problematic and the protein deficiency resulting from non-allele-specific silencing may be deleterious (as discussed in sections 1.2.1 and 1.4). Intronic miRNA mimics may be ideal for this purpose, since they can theoretically be incorporated as introns within the replacement gene and delivered as a single stable expression system, to provide a long-term therapy without re-administration.

The main hypothesis investigated was that naturally-occurring intronic miRNAs may be mimicked to direct silencing of Ataxin-7 by RNAi. The main focus was upon mirtrons, in which functional pre-miRNA stem-loops are formed via intronic splicing rather than the microprocessor. This may offer benefits in the accuracy of processing and minimal competition with the endogenous miRNA pathway. Additionally, their small size (<100 bp) may be beneficial in K&R, given the limited nucleic acid capacity of delivery systems. A miRNA cluster mimic was also adapted to target Ataxin-7, comprising three miRNAs contained within an 800 bp intron.

3.7.1 Artificial mirtrons silence Ataxin-7 efficiently

Using an algorithm to select appropriate target sites within Ataxin-7, a number of mirtrons were designed. They were screened for GFP expression (indicating splicing capability) and for

silencing of a matched dual Luciferase construct. Although a wide variety of results were seen across the set, several were promising mirtrons. It was demonstrated by qPCR that two artificial mirtrons elicited silencing of the endogenous *ataxin-7* transcript in HEK-293 cells and both human and mouse neuroblastoma cell lines (SH-SY5Y and N2A) at comparable levels to siRNA and shRNAs.

It was not possible to detect protein-level silencing, even using siRNAs, indicating that the extent of silencing and the turnover of Ataxin-7 protein do not give an observable change in 3-4 days. Although a single band was seen around the expected size for both cell lines tested by Western blot, it is also possible that the Ataxin-7 antibody was not specifically detecting Ataxin-7. Endogenous Ataxin-7 protein is notoriously difficult to detect, even in brain tissue, and its stability is not known. However, in a stable-inducible rat neuronal cell (PC12) model, after inhibition of Ataxin-7-Q10-EGFP synthesis the protein was not detectable at all after 12 hours (Yu et al. 2012). Its half-life was 4.5 hours. This indicates that the timescales used should be sufficiently long to observe protein-level silencing, and perhaps the limiting factor was the inability to detect small changes in the already very low levels of the endogenous protein. It is also possible that endogenous Ataxin-7 is more stable than indicated above, since over-expressed GFP-tagged proteins may be turned over more rapidly than endogenous proteins (Yewdell et al. 2011). Alternatively, Ataxin-7 may be stabilised by post-translational modification.

3.7.2 Artificial mirtrons are processed into mature miRNAs

Deep sequencing demonstrated that mature miRNAs were produced from artificial mirtrons. Mirt-18 produced a lower number of reads than mirt-1cN, yet it achieved a similar level of silencing, indicating that the mature species derived from mirtrons can be potent even at low levels. For mirt-1cN, strand selection was strongly biased towards the intended (5')

guide strand, and the bias was successfully reversed towards the 3' guide strand to some extent for mirt-18. However, for both mirtrons, mature products formed from the passenger strand silenced a matched Luciferase target, though encouragingly this was less potent than guide strand-directed silencing. The ends of the mature species defined by Dicer were more variable than those defined by splicing. For this reason it may be beneficial to design mirtrons with the mature species on the 5' arm of the hairpin such that the seed region is defined by splicing rather than Dicer.

By making mutations to the 5' splice site of mirt-18 and mirt-1cN, it was shown that the silencing activity of mirtrons was largely dependent upon splicing. However, for both of these un-spliceable mirtrons, a small but significant level of silencing of full-length Ataxin-7, but not Luciferase targets, was detected. The lack of a response in the sensitive Luciferase assay indicates that the mCherry readings may be inaccurate, perhaps due to the low expression level of this tagged construct. However, mature miRNAs were detected by deep sequencing for mirt-18-US. It may be that when splicing is inhibited, the hairpin-containing transcript can become available to processing by Drosha. In further support of this, a small but significant level of Luciferase silencing occurred for mirt-12, despite the fact that it did not produce EGFP fluorescence. However, it is likely that when splicing is possible, the mirtron pathway is dominant, since splicing is strongly linked with transcription (Montes et al. 2012).

3.7.3 Optimisation of mirtron design

The mirtrons investigated here can inform future mirtron design. The first five bases of the mirtron were important for its splicing ability. Artificial mirtrons beginning *GTGAG* were generally spliced well (at least 50% compared to NAD), whereas those beginning *GTGGA* (mirt-5 and mirt-12) were not spliced at all. This was also a general feature of other artificial

mirtrons (Y. Seow and C. Sibley). Furthermore, many naturally-occurring human mirtrons begin *GTGGG* or *GTGAG* (Berezikov et al. 2007; Ladewig et al. 2012). The consensus sequence for the 5' end of an intron is *GTRAG* (Lim and Burge 2001).

Although most introns in vertebrates are identified via exonic splice recognition sequences, small introns are generally recognised by intronic sequences (Lim and Burge 2001). However, the 5' and 3' splice sites alone are not sufficient for short intron recognition. The branch site plays a small role but pentanucleotide motifs and other unknown factors (perhaps motifs with precise spacings) are more significant (Lim and Burge 2001). The pentanucleotide sequences which occur most frequently in human short introns contain a *GGG* motif. This is a complex relationship, however, since GGG motifs are favoured beyond 8 bp into the intron, whereas GGGG motifs are favoured throughout the 5' end of the intron (Dogan et al. 2007). There did not appear to be a clear correlation between these motifs and the splicing efficiency of Ataxin-7 mirtrons. To improve mirtron design, a pentanucleotide motif could perhaps be incorporated into the loop section of mirtrons to enhance splicing efficiency.

The hairpin structure of mirtrons may mask their splice recognition motifs. Indeed, reducing the complementarity between the guide and passenger strand of mirt-18 led to an increase in splicing. Interruptions to the poly-pyrimidine tract may also reduce the splicing efficiency. Once mirt-5 and mirt-12 are removed from the set for their unrecognised 5' splice site, having 2 or more G residues in the last 20 bases of the intron was generally a good predictor of poor splicing ability.

Despite the importance of intronic sequences, splicing of mirtrons was still context-dependent, as demonstrated by the multiple-mirtron constructs, where the same hairpins were incorporated into a second splice site within the EGFP gene. Here both mirtrons were

less efficiently spliced. The flanking exonic sequences perhaps contribute to the formation of secondary structures which mask the splice sites.

Various sequence alterations were made to mirt-1, to attempt to improve its post-splicing processing and hence its silencing capability. Shortening the hairpin seemed to make the biggest impact. Since splicing was not improved by this modification, the increased silencing activity may be due to a downstream processing step. Perhaps the smaller hairpin size improved recognition by Dicer or enhanced its export to the cytoplasm. Weakening the hairpin structure for mirt-18 reduced its silencing capability despite the increase in splicing, presumably through impaired downstream processing. It appeared that mirt-18 may accumulate as pre-miRNA and be processed slowly into mature species. A careful balance must be achieved between these factors when designing artificial mirtrons. The rules for optimal mirtron design are yet to be fully established, as many avenues remain unexplored and will be pursued in future studies.

3.7.4 Other miRNA mimics may also be used to target Ataxin-7

Canonical intronic miRNAs may also be mimicked to silence a target of interest, and this was investigated for Ataxin-7. The naturally-occurring miR-106b cluster contains three miRNAs and is found within an intron of approximately 800 bp in the MCM-7 gene. This system was developed to co-express three artificial miRNAs against HIV (Aagaard et al. 2008) and later adapted to target vascular endothelial growth factor (VEGF), for which silencing was demonstrated *in vivo* (Pihlmann et al. 2012). This system may be advantageous over mirtrons, because the hairpins are not subject to splicing-related constraints, allowing a less restricted choice of miRNA target sites, and hence potentially greater silencing. In addition, three hairpins are incorporated into a single intron, which may be more straightforward to apply to a K&R therapy.

Artificial mirtrons and a shRNA hairpin targeting Ataxin-7 were individually incorporated into the three miRNA positions within a construct mimicking the miR-106b cluster. Each hairpin was re-designed, keeping the targeting sequence but creating mismatches in the passenger strand in order to mimic the secondary structure of the miRNA which occurs in the natural cluster for each position. miRNAs within a miR-106b cluster mimic were effective against Luciferase targets and mCherry-tagged full-length mutant Ataxin-7. However, this was only true for two positions of the three available in the cluster. For one position, all three miRNAs tested were unable to produce substantial silencing. Interestingly, the second and third hairpins were previously problematic for producing good silencing efficacy (Aagaard et al. 2008; Pihlmann et al. 2012), indicating this is a sequence- or structure-specific effect rather than a positional effect. This may represent a limitation for the development of this approach, since the non-functional hairpin may place an unnecessary burden on the RNAi machinery and may cause off-target effects.

If this approach were to be applied to a combination therapy, it would be essential for splicing to occur accurately across the exons, because the resultant spliced mRNA would be required to produce a functional protein. Splicing cannot be inferred from intronic miRNA production, since splicing is not essential for successful processing for canonical intronic miRNAs (Kim and Kim 2007). Therefore splicing of this intron was confirmed here; however, it will be important to test its splicing ability in the context of a therapeutically relevant gene.

Despite their promise, when application to combination therapy is considered, these miRNA cluster mimics lead to greater difficulties in molecular manipulation compared to mirtrons. Mirtrons, at less than 100 bp, can be cloned into specific splice sites within a chosen gene cheaply by using appropriately designed long primers. However, this cloning procedure is

much more difficult and/or expensive for introns of 800 bp, especially if a number of splice sites have to be tested.

3.7.5 Conclusions

Mirtrons and intronic miRNA clusters were designed and achieved silencing of Ataxin-7. Despite their greater design restrictions, mirtrons are perhaps preferable over other intronic miRNAs, for their ease of manipulation and potential advantages related to their processing pathway. Both approaches investigated here currently have drawbacks and require further optimisation. Intronic miRNAs, including mirtrons, may be ideal for application to a combination gene therapy where they are delivered as introns within a transgene. For SCA7, it is likely that Ataxin-7 function is important, particularly in the retina. Therefore, delivering a replacement gene along with silencing the endogenous gene may be beneficial for a successful therapy. The next chapter will investigate the development of mirtron-based K&R therapies for SCA7.

4 Gene knockdown-replacement systems for SCA7

4.1 Introduction

A promising strategy for the treatment of expansion disorders such as SCA7 is to combine non-allele-specific silencing with delivery of a replacement gene, i.e. knockdown and replacement (K&R). The major advantage of this is to retain protein function while avoiding the difficulties involved in allele-specific therapy, discussed in section 1.4. The K&R approach may have wider applications to many dominant diseases where either the mutation is difficult to specifically target, or where wide heterogeneity means that any allele-specific therapy could only treat a sub-group of patients, such as in retinitis pigmentosa (RP) (Kiang et al. 2005). A K&R approach could theoretically provide a single treatment for all patients affected by any dominant mutation in the same gene. Furthermore, the potential applications are expanded by the option of silencing one gene and delivering a different therapeutically relevant gene ("combination therapy").

RNA interference (RNAi) strategies may be designed to specifically silence the gene of interest, as discussed in section 1.3.3 and Chapter 3. These may be incorporated into K&R therapies through co-delivery with a RNAi-resistant copy of the gene, the "replacement gene", which does not contain the RNAi target site. To achieve RNAi resistance, silent modifications can be made within the target region to prevent binding of the RNAi effector without changing the amino acid sequence (Xu et al. 2009). Alternatively, if the target site is within the gene's 3' untranslated region (UTR), the transgene may simply be delivered without its 3' UTR.

Short interfering RNAs (siRNAs) have been explored for gene K&R therapies for dominant diseases SCA6 and RP, where successful silencing and codon-modified transgene expression were achieved *in vitro* and *in vivo*, respectively (Kubodera et al. 2005; Kiang et al. 2005). However, the need for re-administration of siRNAs presents a major limitation for long-term gene therapy.

A small number of groups therefore adapted shRNA or miRNA expression systems for combination therapy applications tested in culture (Samakoglu et al. 2006; Unwalla et al. 2006; Palfi et al. 2006; Li and Mahato 2009). In a clinical trial, a plasmid-based vaccine therapy for cancer (delivered *ex vivo* and expressing a transgene and two shRNAs) also demonstrated some degree of success in terms of safety and increased patient survival. Furthermore, phenotypic improvements were observed in the mouse retina when K&R therapies towards Rhodopsin or rds-peripherin were applied (O'Reilly et al. 2007; Millington-Ward et al. 2011; Greenwald et al. 2012). Together this proves that combination therapies for silencing and gene delivery are feasible, capable of achieving phenotypic improvements in animal models, and may be used in humans.

Delivering transgenes to the brain and central nervous system (CNS) is limited to the use of vectors which are capable of infecting non-dividing cells (Linterman et al. 2011). Lentiviruses are one such example. Encouragingly, a lentiviral vector was used for delivery of a therapeutic transgene to Purkinje cells in mouse models of SCA3 (Torashima et al. 2008). Lentiviruses have a relatively large packaging capacity of 8-10 kb (Sinn et al. 2005). However, this and other studies used local injection for delivery to the brain, which is invasive and not ideal for clinical application (Lundberg et al. 2008). Certain recombinant Adeno-Associated viruses (AAV) can reach the CNS through systemic delivery (Xie et al. 2011). However, AAV can package expression cassettes no larger than 5 kb (Wu et al. 2010). Therefore, delivering

each portion of a K&R therapy under separate promoters can restrict the choice of viral vectors. One solution, delivering the therapy in two separate viruses, can lead to cells receiving only one aspect of the therapy, limiting its efficacy (Greenwald et al. 2012).

Hence, in some of the above cases, the shRNA/miRNA was expressed either upstream (Unwalla et al. 2006), downstream (Li and Mahato 2009) or within an endogenous intron (Samakoglu et al. 2006) of the transgene. This led to successful expression of both aspects of the therapy. However, placing a shRNA within an intron must be carefully assessed for negative interferon responses (Samakoglu et al. 2006).

Numerous safety benefits over shRNAs have been found by mimicking endogenous miRNAs (Grimm 2011). Many naturally-occurring miRNAs are intronic, and some may be mimicked to silence a gene of interest (Aagaard et al. 2008). Intronic miRNAs are, in theory, ideal RNAi effectors for K&R applications, as they can be nested within the replacement gene. Mirtrons, a type of intronic miRNA where the ends of the pre-miRNA are defined by splicing, may be particularly useful, their small size making them relatively easy to manipulate and well suited for limited-capacity delivery systems. Endogenous mirtrons can be expressed as introns outside of their natural context and can be successfully adapted to silence a gene of interest (Sibley et al. 2011; Seow et al. 2012; Sibley et al. 2012) and Chapter 3. In addition, they appear to be potent silencers even at low levels, potentially avoiding saturation of the RNAi pathway (Grimm et al. 2006; Seow et al. 2012). Multi-gene targeting is also possible, through delivery of multiple mirtrons within a single gene (Sibley et al. 2012) (section 3.4). For mirtrons derived from the 5' arm, their seed region is strictly defined by splicing, avoiding the potential for variation caused by microprocessor cleavage and hence limiting off-target effects (Wu et al. 2009b). For 3' arm mirtrons, having well-defined ends can also be

advantageous for accuracy of the seed region by informing Dicer processing (Park et al. 2011).

In this chapter, the anti-Ataxin-7 mirtrons developed in Chapter 3 will be investigated for development into a K&R therapy for SCA7. Firstly, the production of mirtron-resistant Ataxin-7 is investigated. The prospects for achieving concurrent silencing and replacement of Ataxin-7 with a single expression construct, and delivering mirtrons as introns within the replacement Ataxin-7 gene, are then explored.

4.2 Development of RNAi-resistant Ataxin-7

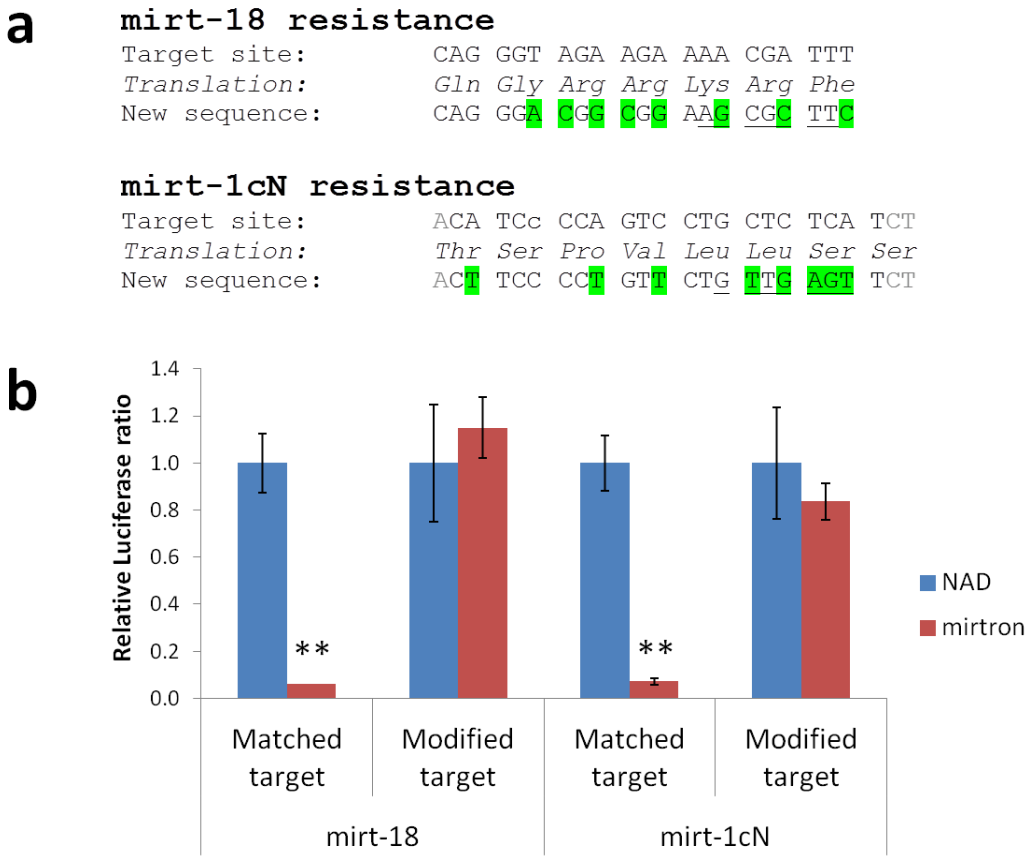
The mirtrons developed in the previous chapter target the coding region of Ataxin-7. Changing a number of bases within a mirtron target site will destroy the complementarity of the mature miRNA species, and hence inhibit its silencing capability (Xu et al. 2009). In order for this RNAi-resistant gene to produce a functional protein, the wild-type amino acid sequence must be preserved, which can be achieved by making only silent mutations in each target site, i.e. codon replacement (Kiang et al. 2005; Cashman et al. 2005; Palfi et al. 2006; O'Reilly et al. 2007; Greenwald et al. 2012) (Figure 4-1a). Although a single base change can be sufficient to inhibit silencing, including a number of mismatches can ensure more complete resistance (Kiang et al. 2005). Care must be taken to avoid rare codons, as some transfer RNAs (tRNAs) have limited availability; this can impede translation and affect the folding, and hence the function, of the resultant protein (Kimchi-Sarfaty et al. 2007). Therefore, each codon in the target site was, as far as possible, replaced with an alternative codon with comparable usage levels in humans (detailed in section 2.2.7.1).

The modified target site sequences were each cloned into a separate dual Luciferase reporter construct and confirmed to be resistant to silencing by mirtrons (Figure 4-1b). Each

modification was then incorporated into a full-length wild-type Ataxin-7 gene (Ataxin-7-10Q-mCherry) and mirtron activity against them was tested. Each modification was tested separately in this context (M18, M1cN) and they were also incorporated together to make a construct resistant to both mirtrons (M18-M1cN). For each construct, a high degree of resistance to silencing by the appropriate mirtron(s) was demonstrated (Figure 4-1c). It was confirmed that the expression level of each modified construct was not significantly different to the wild-type (Figure 4-1c, when co-transfected with the pEGFP-NAD control). Fluorescent microscopy indicated that the cellular distribution of the modified constructs was similar to the wild-type (Figure 4-1d). These data suggest that these modified versions of Ataxin-7 may be suitable for use as replacement genes in combination with one or both mirtrons in a K&R therapy.

Although reduced compared to the wild-type sequence, significant silencing was observed for M18 (Figure 4-1c). This was contrary to the result seen in the Luciferase assay (Figure 4-1b) which is generally more sensitive. Nonetheless, the reduced levels of the full-length construct may lead to difficulties interpreting later results involving development of K&R therapies. Attempts were therefore made to improve its resistance. Binding of the seed region (bases 2-8) can be sufficient for silencing of a target gene by a miRNA (Haley and Zamore 2004). In M18, only two bases were altered within the seven bases corresponding to the seed region of the mature miRNA (underlined, Figure 4-1a), one of which allowed weak G-U base pairing with the mature miRNA. The other was at base 4 of the guide strand; a mismatch at this position did not confer strong RNAi resistance in a similar study (Kiang et al. 2005). Further changes were therefore made to increase the number of mismatches ("18M2", Figure 4-2a). The limits of codon replacement and using only codons with similar abundances allowed only one additional base change to be made within the seed region and

one just beyond the seed region. The mismatch at position 4 had to be converted into a weaker G-U mismatch. These modifications were incorporated into the full-length Ataxin-7-10Q-mCherry plasmid and the ability of mirtrons to silence them was tested (Figure 4-2b). Comparable levels of silencing (almost 20%) were observed for both M18 and 18M2, but error bars were large and the results were not statistically significant. Fluorescence measurements were only 40-80% above background readings which may lead to inaccuracy. Although improved resistance to silencing could not be proven, the 18M2 construct was favoured for use in the development of K&R therapies to minimise any possible silencing.



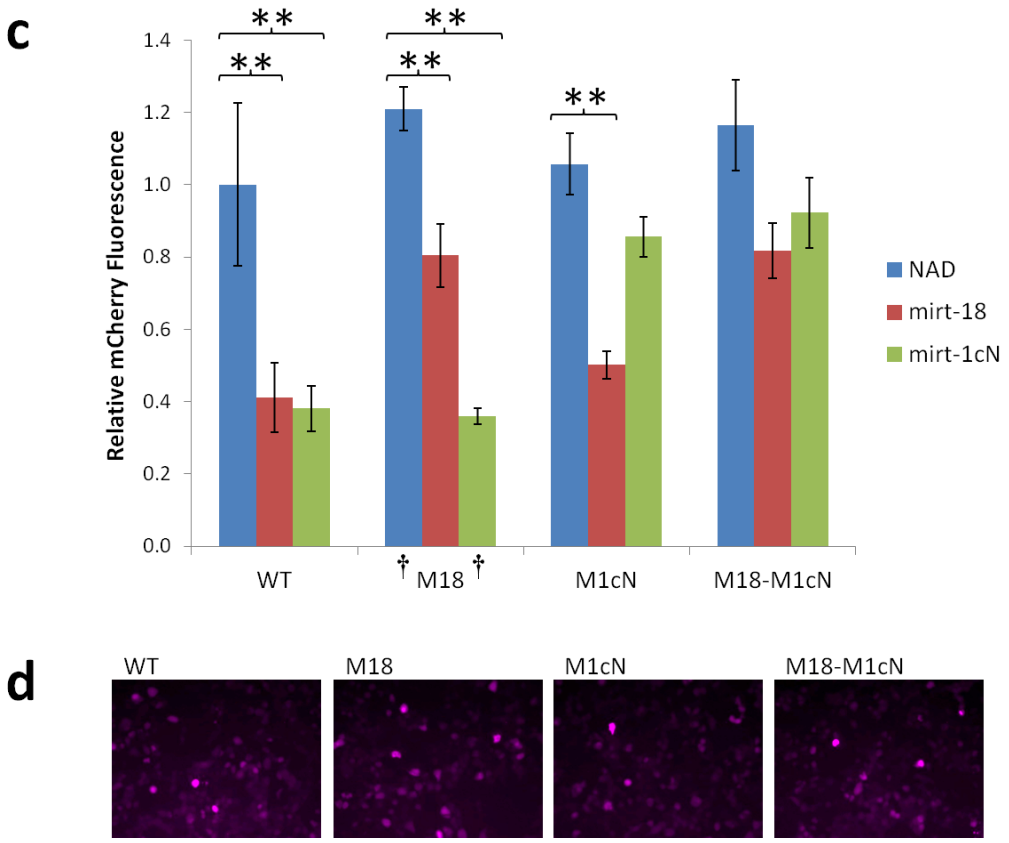


Figure 4-1 - Modified Ataxin-7 constructs were designed to be resistant to silencing by mirtrons. (a) The sequence of the target site for each mirtron and the translated amino acid sequence. Codon substitutions were made to prevent silencing by mirtrons while maintaining the amino acid sequence. Each base change is highlighted. Bases outside the target site but contained within the codons are shown in grey. The bases corresponding to the seed region of the mature miRNA are underlined. **(b)** Normalised *Renilla*:firefly Luciferase ratios for mirtrons against their matched targets and the modified targets in HEK-293 cells 48 h post-transfection. Values were normalised to the control (NAD) for each target. N=6, combined from two independent experiments. Error bars represent standard deviation. Asterisks (**) represent $p<0.005$ compared to NAD for the same target (one-tailed t-test). **(c)** mCherry fluorescence quantification for protein extracted from HEK-293 cells 48 h post-transfection with full-length Ataxin-7 constructs and mirtrons. The Ataxin-7 constructs contained either: no modification (WT), a single modified target site for resistance to silencing by either mirt-18 or mirt-1cN (M18/M1cN), or both modifications (M18-M1cN). N=6, combined from two independent experiments, except where indicated (†, N=5). 200 µg total protein was loaded for each sample. Background readings (un-transfected cells) were subtracted and fluorescence for the wild-type construct plus NAD intron was normalised to 1. Error bars represent standard deviation. Asterisks (**) represent $p<0.005$ (one-tailed t-test). **(d)** Representative fluorescent images of HEK-293 cells 48 h post-transfection with the modified Ataxin-7 constructs.

a

```
Target site:    CAG GGT AGA AGA AAA CGA TTT
Translation:    Gln Gly Arg Arg Lys Arg Phe
M18:            CAG GGA CGG CGG AAG CGC TTC
18M2:           CAG GGA CGG CGC AAG AGG TTC
```

b

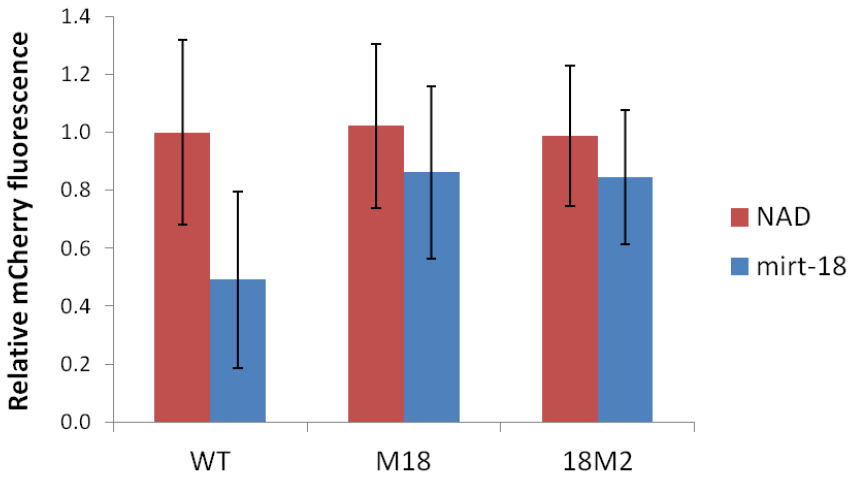


Figure 4-2 – Further modifications were made to the mirt-18 target site. (a) Sequence of the mirt-18 target site, the translated sequence and the alternative codon substitutions made, with the original base changes (M18) highlighted in green and the new changes (18M2) highlighted in yellow. The bases corresponding to the seed region of mature mirt-18 are underlined. **(b)** mCherry fluorescence quantification for protein extracted from HEK-293 cells transfected with Ataxin-7 constructs and mirt-18. The Ataxin-7 constructs contained either: no modification (WT), M18 or 18M2 as detailed in (a). N=4, from 2 independent experiments. 500 µg total protein was loaded. Background readings (un-transfected cells) were subtracted and fluorescence for the wild-type Ataxin-7 construct (WT) plus NAD intron was normalised to 1. Error bars represent standard deviation. No differences were statistically significant (one-tailed t-test).

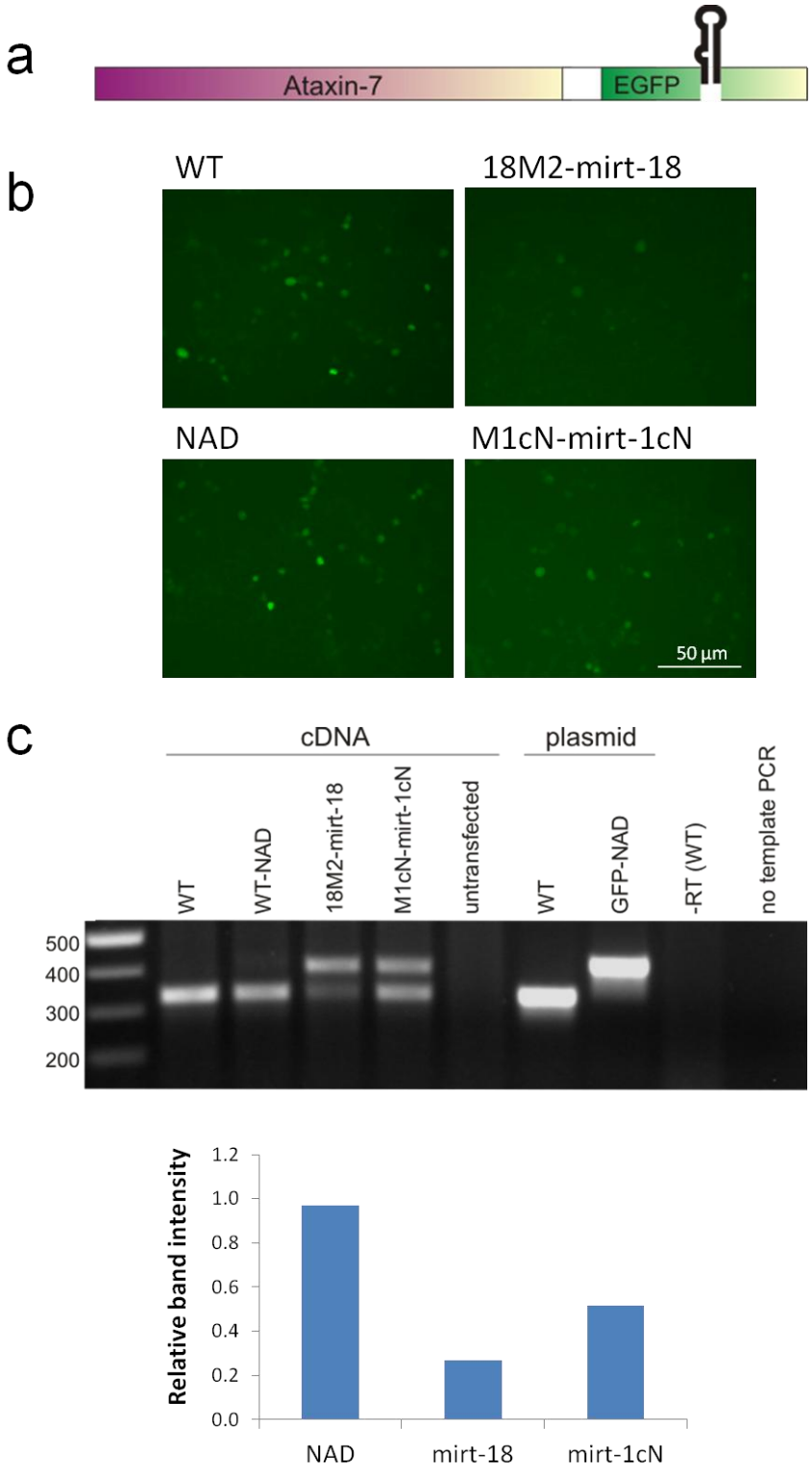
4.3 Knockdown-replacement proof-of-principle

Having created Ataxin-7 genes which are resistant to silencing by mirtrons, it was then desired to confirm that Ataxin-7 may be concurrently silenced and replaced, to verify that a mirtron-based K&R approach is valid for SCA7.

EGFP-mirtron genes (either EGFP-mirt-18 or mirt-1cN) were cloned in place of the mCherry tag after the appropriate replacement Ataxin-7 gene (18M2 or M1cN respectively; Figure 4-3a). This combined replacement Ataxin-7 and a mirtron into a single construct, driven by a common promoter, without the need for identifying splice sites within Ataxin-7. As a control, the EGFP-NAD gene was cloned into the unmodified Ataxin-7-10Q plasmid (creating pAtaxin-7-EGFP-NAD).

Each plasmid was transfected into HEK-293 cells and EGFP levels were compared to the wild-type (intron-less) Ataxin-7-10Q-EGFP construct. EGFP fluorescence indicates both the expression of the Ataxin-7-EGFP fusion protein and successful splicing of the intron from EGFP. Although fluorescence for Ataxin-7-EGFP-NAD appeared similar to the intron-less construct, it was clearly reduced for the mirtron-containing plasmids, particularly mirt-18 (Figure 4-3b). Unfortunately, EGFP fluorescence could not be reliably quantified at such low expression levels. Splicing efficiency was demonstrated by RT-PCR (Figure 4-3c). As expected, the NAD intron was very efficiently spliced. However, efficiencies were only approximately 25% and 50% for mirt-18 and mirt-1cN respectively, substantially lower than in the original EGFP-mirtron constructs (60% and 95%, respectively; section 3.3.1). Despite this, significant silencing (80%) was achieved against matched Luciferase targets for both mirtron-containing constructs (Figure 4-3d). They were not as effective as the same mirtrons in the original EGFP context (section 3.2), likely to be explained by a combination of the inefficient splicing

and the weaker promoter leading to very low levels of the mature species. It is encouraging that substantial target knockdown was seen, and at comparable levels for both mirtrons despite the lower levels of production for mirt-18.



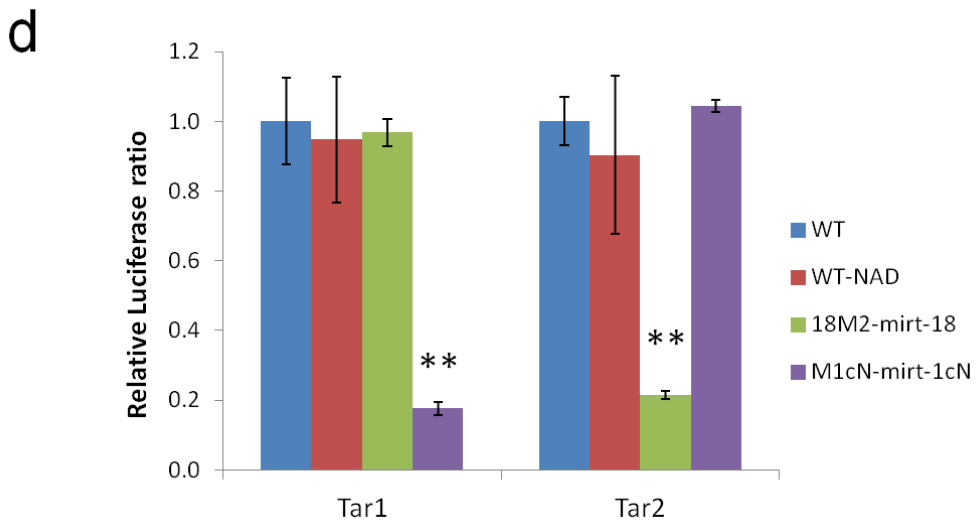


Figure 4-3 – Knockdown and replacement of Ataxin-7 achieved by delivering a mirtron within the EGFP tag. (a) Schematic of Ataxin-7 constructs, containing wild-type/mirtron-resistant Ataxin-7 cDNA and an EGFP tag containing a mirtron. **(b)** Representative fluorescent images of HEK-293 cells, 48 h post-transfection with each Ataxin-7-EGFP-mirt construct. **(c)** RT-PCR was carried out using intron-spanning primers to demonstrate splicing efficiency (larger product = un-spliced; smaller product = spliced). The quantification of the splicing efficiencies for the NAD intron and mirtrons are also shown. The average band intensity for the spliced product is given as a proportion of the total (spliced + un-spliced). **(d)** Relative *Renilla*:firefly Luciferase ratios indicating the silencing capability of each Ataxin-7-EGFP-mirtron construct, 48 h post-transfection in HEK-293 cells. The Tar1 dual Luciferase plasmid contains the mirt-1cN target site within the 3' UTR of *Renilla* Luciferase; Tar2 is the equivalent for mirt-18. The ratio was normalised to 1 for the wild-type construct for each target. N=3. Error bars represent standard deviation. Asterisks (**) represent $p<0.005$ (one-tailed t-test).

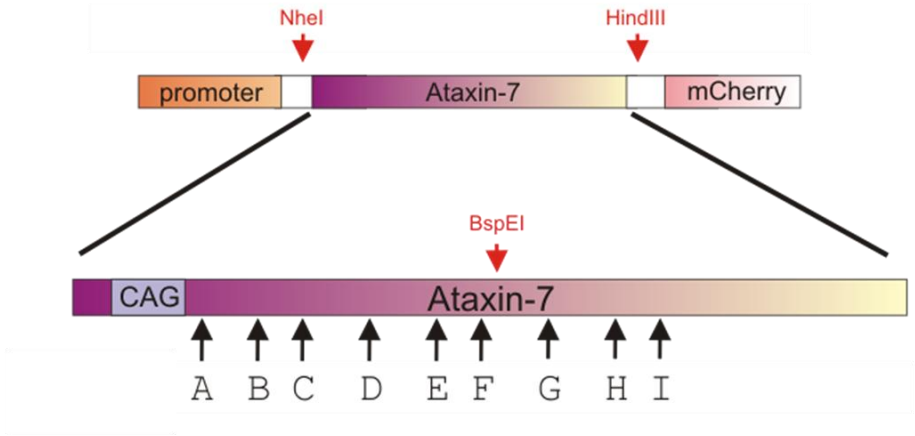
4.4 Delivering mirtrons as introns within Ataxin-7

Having established that mirtrons targeting Ataxin-7 may be delivered alongside an RNAi-resistant Ataxin-7 gene under a single promoter, it was then desired to improve this K&R system, by instead delivering mirtrons as introns within the RNAi-resistant replacement Ataxin-7 gene (M1cN or M18/18M2). This would minimise the size of the construct and eliminate the necessity for a fluorescent tag. Mirtrons and the NAD intron were inserted as introns within a number of different sites in Ataxin-7 and in the 5' UTR.

4.4.1 Initial testing of splice sites in Ataxin-7

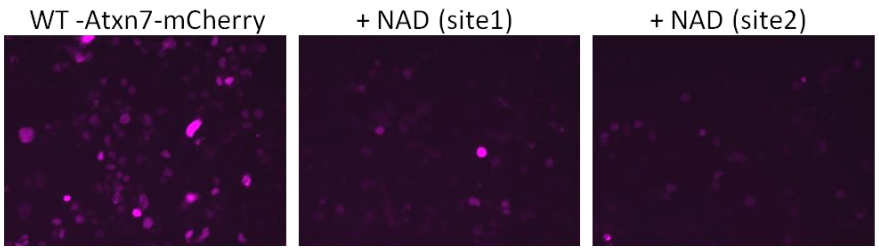
The Ataxin-7 construct used to create the RNAi-resistant genes was based upon Ataxin-7 cDNA, containing all translated exons. The natural exon-exon junctions (which would border introns in the genomic context) were identified in the cDNA construct (Figure 4-4a) and investigated for suitability for mirtron splicing. Those located between BspEI and HindIII restriction sites were first considered (G, H and I), to avoid performing PCR across the CAG-repeat and its surrounding GC-rich region, which tends to introduce errors. Of these three sites, the two which normally contain the shortest introns (G and H, 1851 and 363 bp respectively) were chosen (Figure 4-4a, sites 1 and 2). This tested the hypothesis that exonic sequences surrounding short introns may be naturally best-suited to the splicing of artificial short introns. A mirtron or NAD intron was cloned into each site, within the background of an appropriate resistance modification (M1cN or M18) such that the transcript would not be silenced. Since neither the NAD intron nor either mirtron consist of a multiple of 3 bases, the mCherry tag following Ataxin-7 is in-frame only when the intron is spliced, therefore its fluorescence gives an indication of splicing efficiency. With the NAD intron in either site 1 or 2, mCherry fluorescence was clearly compromised compared to the wild-type intron-less control (Figure 4-4b and c). Intron-spanning RT-PCR confirmed that this was attributable to intron retention, for both mirtrons and the NAD intron (Figure 4-4d). Since the NAD intron is not hairpin-forming, poor splicing could not be attributable to the formation of intron-based secondary structures. Maximal splice efficiency is desirable to ensure sufficient levels of transgene and mirtron expression while avoiding production of aberrant splice products, which could trigger interferon responses (Samakoglu et al. 2006). It was concluded that these sites may not be ideal for the recognition of these artificial short introns, perhaps due to differences in their sequences and/or size compared to the natural introns.

a

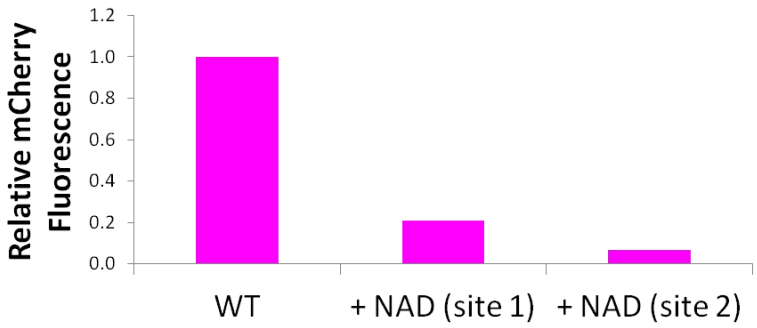


Intron site	Intron length	5′ SS	3′ SS	
A	233	GACG	GGAC	
B	39153	GAAG	ACAT	
C	27431	TATG	AAAG	
D	2018	GAAT	CATG	
E	493	TCAG	AAAG	
F	5037	CAAG	ACAC	
G	1851	CAAG	GCCT	→ Site 1
H	363	ATCT	TTTT	→ Site 2
I	4645	GGAA	GAAA	

b



c



d

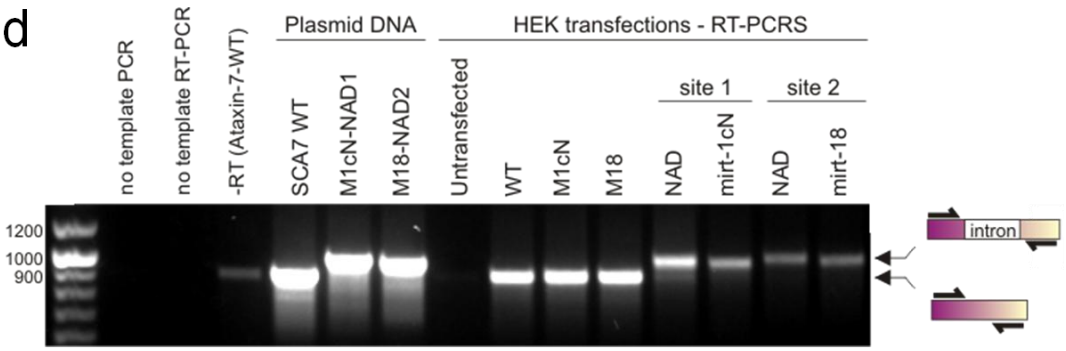


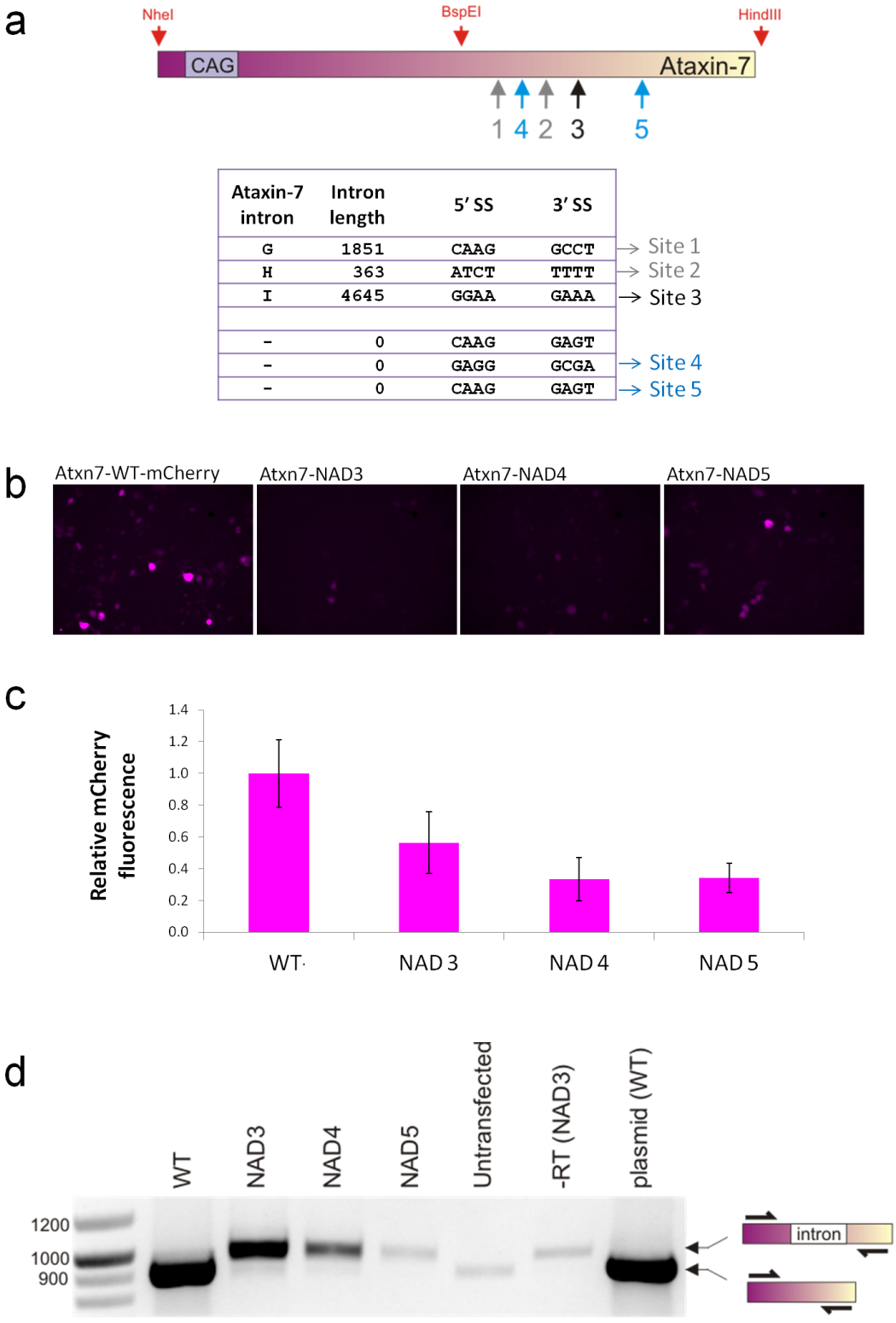
Figure 4-4 – Mirtrons were cloned into two natural intron locations within Ataxin-7 and tested for splicing. (a) Schematic representation of the Ataxin-7 cDNA construct, with its key restriction sites (red arrows) and natural exon-exon boundaries (A-I) indicated. "CAG" represents the CAG-repeat and GC-rich region. The table lists each of the natural Ataxin-7 introns, their length (bp), the last four bases of the upstream exon (5' splice site; 5' SS) and the first four bases of the downstream exon (3' SS). The chosen sites for insertion of mirtrons (1 and 2) are indicated. **(b)** Representative images of mCherry fluorescence of HEK-293 cells transfected with Ataxin-7 constructs. **(c)** mCherry fluorescence was quantified for protein extracted from HEK-293 cells transfected with Ataxin-7 constructs. N=1. 250 µg total protein was loaded. Background readings (un-transfected cells) were subtracted and fluorescence for the wild-type construct was normalised to 1. **(d)** RT-PCR was carried out using intron-spanning primers (which amplify Ataxin-7 cDNA across both sites 1 and 2), to demonstrate splicing efficiency. RNA was extracted from HEK-293 cells 48 h post-transfection with intron-containing and intron-less Ataxin-7 constructs. Negative controls (no template PCR/RT-PCR, -RT) and plasmid controls were also included. Bands represent un-spliced (retained intron) and spliced products as indicated by the schematic diagram.

4.4.2 Testing of further natural and potential intron sites

The remaining exon-exon boundary sequences in Ataxin-7 were further investigated. Between the BspEI and HindIII restriction sites which facilitate the cloning procedure, only one natural boundary remained (Figure 4-5a, site 3). However, other sites which resemble exon-exon boundaries (but do not naturally contain introns) may also be amenable to splicing. Certain boundary sequences were found in other contexts, e.g. EGFP, to allow the splicing of artificial mirtrons (Y. Seow; upstream CAAG or CAGG, downstream GACG, GAGT or GCGA). Three such sites were identified in Ataxin-7 cDNA, (listed in Figure 4-5a). Two of these lay between the BspEI and HindIII recognition sequences and were therefore selected for testing (sites 4 and 5).

The NAD intron was cloned into these three sites (3, 4 and 5) to test the splicing of short introns. mCherry fluorescence was clearly lower than the intron-less gene for all three constructs (Figure 4-5b), indicating that splicing was not efficient. Quantification of the

fluorescence confirmed this result (Figure 4-5c). However, the fluorescence measurements did not reflect the pattern seen in Figure 4-5b, likely due to inaccuracy caused by the low readings obtained; even for the wild-type construct, the raw measurement was on average only 40% above the background (un-transfected) readings. Intron-spanning RT-PCR verified that splicing of the NAD intron was poor in all three cases, with very little spliced product detected (Figure 4-5d). This was perhaps more severe than expected given the fluorescence levels observed in Figure 4-5b and c. It is possible that the fluorescence was attributable to mis-splicing, which may by chance produce mRNAs in which mCherry is in-frame. Such products would not be detected by RT-PCR if an activated splice site lay outside the primer binding sites. The PCR reaction may have been unsuccessful for NAD5, since only a faint band was visible for the un-spliced product, which is indistinguishable from the plasmid carryover demonstrated by the –RT lane. The same result was also obtained upon repetition. Some endogenous Ataxin-7 transcript was also detected, as seen in the un-transfected lane.



Ataxin-7 intron	Intron length	5′ SS	3′ SS	
G	1851	CAAG	GCCT	→ Site 1
H	363	ATCT	TTTT	→ Site 2
I	4645	GGAA	GAAA	→ Site 3
-	0	CAAG	GAGT	
-	0	GAGG	GCGA	→ Site 4
-	0	CAAG	GAGT	→ Site 5

Figure 4-5 - Introns were cloned into natural and predicted splice sites within Ataxin-7. (a) Schematic diagram showing the Ataxin-7 cDNA construct and the first two exon-exon boundary sites tested (labelled 1 and 2). One further natural boundary site (3) and two predicted sites (shown in blue; labelled 4 and 5) are also indicated. "CAG" represents the CAG-repeat and GC-rich region. Key restriction sites are also shown (red arrows). The table lists the three natural Ataxin-7 exon-exon boundaries between BspEI and HindIII sites, their length (bp), the last four bases of the upstream exon (5′ splice site; 5′ SS) and the first four

bases of the downstream exon (3' SS). The three sites identified in Ataxin-7 cDNA which resemble exon-exon boundaries are also listed. The sites selected for insertion of mirtrons (1-5) are indicated. The NAD intron was cloned into each of the newly selected sites to create NAD3, NAD4 and NAD5. **(b)** Representative images of mCherry fluorescence for HEK-293 cells 48 h post-transfection with the WT or the new NAD intron-containing Ataxin-7 constructs (NAD3, NAD4 and NAD5). **(c)** Quantification of fluorescence for protein extracted from HEK-293 cells in (b). N=3. 250 µg total protein was loaded. Error bars represent standard deviation. Background measurements (un-transfected cells) were subtracted and fluorescence for the wild-type construct was normalised to 1. Results were not statistically significant (one-tailed t-test). **(d)** RT-PCR was carried out using primers spanning the intron, as indicated in the schematic diagram, to demonstrate splicing efficiency. RNA was extracted from HEK-293 cells 48 h post-transfection. Bands represent un-spliced and spliced products as indicated by the schematic diagram.

4.4.3 Computational prediction of splice site recognition

The consensus for the exonic portion of the 5' splice site is (C/A)AG, with little definition beyond these bases (Mount 1982). This is explained by the complementary base pairing of the U1 small nuclear RNA (snRNA) across the exon-intron boundary (Krämer et al. 1984; Zhuang and Weiner 1986). Of the five sites tested, only sites 1 and 5 matched this consensus (AAG), and site 4 was similar, with just one purine difference (AGG). Regarding the 3' splice site, only a single base of the exon is significantly well-defined: a G residue immediately following the intron. All sites tested did fit this consensus with the exception of site 2. Other, less well-defined exonic sequences also make an important contribution to splice site recognition (Lim and Burge 2001). This may explain why even exonic sites that fit the consensus did not produce efficient splicing. A computational analysis was carried out to determine whether any other sites in Ataxin-7 may be more suitable for artificial mirtrons.

An algorithm, NetGene2 (Brunak et al. 1991), was used to predict the splicing efficiency of each mirtron in each site within Ataxin-7. It performs well in prediction of natural human

splice sites in comparison to other freely available algorithms (Pertea et al. 2001). NetGene2 detects all sites in a given sequence which resemble 5' and 3' splice sites and scores each based upon the likelihood of splicing. All natural and predicted Ataxin-7 cDNA splice sites were assessed, as well as the primary splice site used in EGFP. The scores assigned to the 5' and 3' splice sites in each case are listed in Table 4-1. Surprisingly, within EGFP, the 3' splice site of the NAD intron scored only 0.85, but in practice was efficiently spliced (section 3.3.1). This indicates that scores of 0.85 and above may be good indicators for splicing using this algorithm. However, it must be noted that secondary structure is not considered by NetGene2. Although the incorporation of RNA secondary structure was shown to improve accuracy of splice site prediction, integrating such analysis leads to long processing time (Marashi et al. 2006). No freely available algorithms were found to incorporate this feature. This may explain why the predicted chance of recognition of both splice sites of mirt-18 within the context of EGFP (0.95 and 1.00) does not reflect the 50-60% splicing efficiency seen in Chapter 3 (Figure 3-5).

The results appear to confirm that sites 1-5 are not ideal for splicing of mirtrons or the NAD intron. Indeed, the 3' splice site of the NAD intron had a poor recognition score in all contexts in Ataxin-7, perhaps due to the number of Adenine interruptions in the polypyrimidine tract – greater than that of either mirtron. This implies that the NAD intron is not a good test for splice sites in Ataxin-7 and their suitability for mirtrons. However, encouragingly, sites 7- 10 scored relatively highly (>0.8) for both ends of mirt-1cN, and sites 8 and 9 scored moderately well (>0.6) for mirt-18.

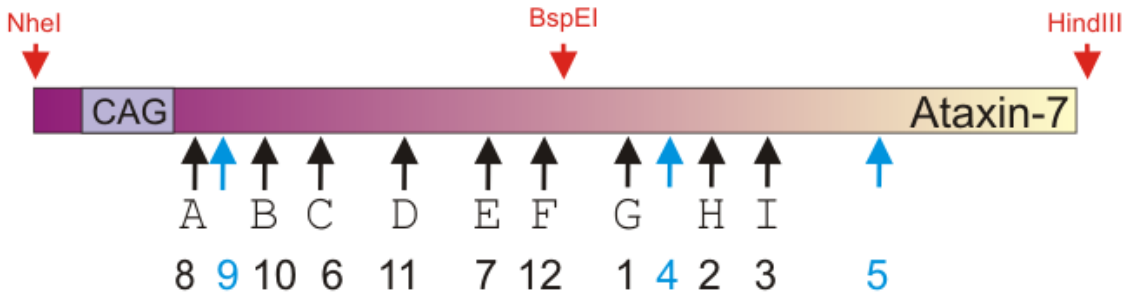
Table 4-1 – *In silico* analysis of splice sites within Ataxin-7 and EGFP. Sites in Ataxin-7 and EGFP are listed along with splice confidence scores for both 5' and 3' splice sites, according to NetGene2 (Brunak et al. 1991), when each site contains the NAD intron, mirt-1cN or mirt-18. Natural Ataxin-7 exon-exon boundaries are shown (A-I) along with those which do not naturally contain introns (†), named according to the diagram below the table. The initial five sites tested are given first (1-5), followed by the other potential sites (6-12), in order of occurrence in the cDNA (5'-3'). A score of 0 indicates the splice site was not detected by the algorithm at all. Cells are highlighted green for scores >0.85. The surrounding four bases of the 5' and 3' exons adjacent to each splice site are also given (5'/3' SS).

Site ID	G (1)		H (2)		I (3)		4†		5†	
5'/3' SS	CAAG/GCCT		ATCT/TTTT		GGAA/GAAA		GAGG/GCGA		CAAG/GAGT	
	5'	3'	5'	3'	5'	3'	5'	3'	5'	3'
NAD	0.47	0.27	0	0	0.6	0	0.62	0.14	0.83	0
mirt-1cN	0.49	0.85	0	0.85	0.76	0.85	0.79	0.97	0.99	0.56
mirt-18	0.41	0.97	0	0.36	0.55	0.43	0	0.89	0.88	0.25

Site ID	A (8)		9†		B (10)		C (6)		D (11)	
5'/3' SS	GACG/GGAC		CAAG/GAGT		GAAG/ACAT		TATG/AAAG		GAAT/CATG	
	5'	3'	5'	3'	5'	3'	5'	3'	5'	3'
NAD	0.67	0.18	0.86	0.17	0.93	0	0.54	0	0	0
mirt-1cN	0.81	0.97	0.85	1.00	0.94	0.90	0.83	0.82	0	0.96
mirt-18	0.62	0.96	0.67	0.97	0.86	0.55	0.41	0.71	0	0.52

Site ID	E (7)		F (12)	
5'/3' SS	TCAG/AAAG		CAAG/ACAC	
	5'	3'	5'	3'
NAD	0.57	0.15	0.93	0
mirt-1cN	0.81	0.97	0	0
mirt-18	0.47	1.00	0.89	0.26

EGFP	
CAAG/GACG	
5'	3'
1.00	0.85
1.00	1.00
0.95	1.00



4.4.4 Investigation of high-scoring splice sites

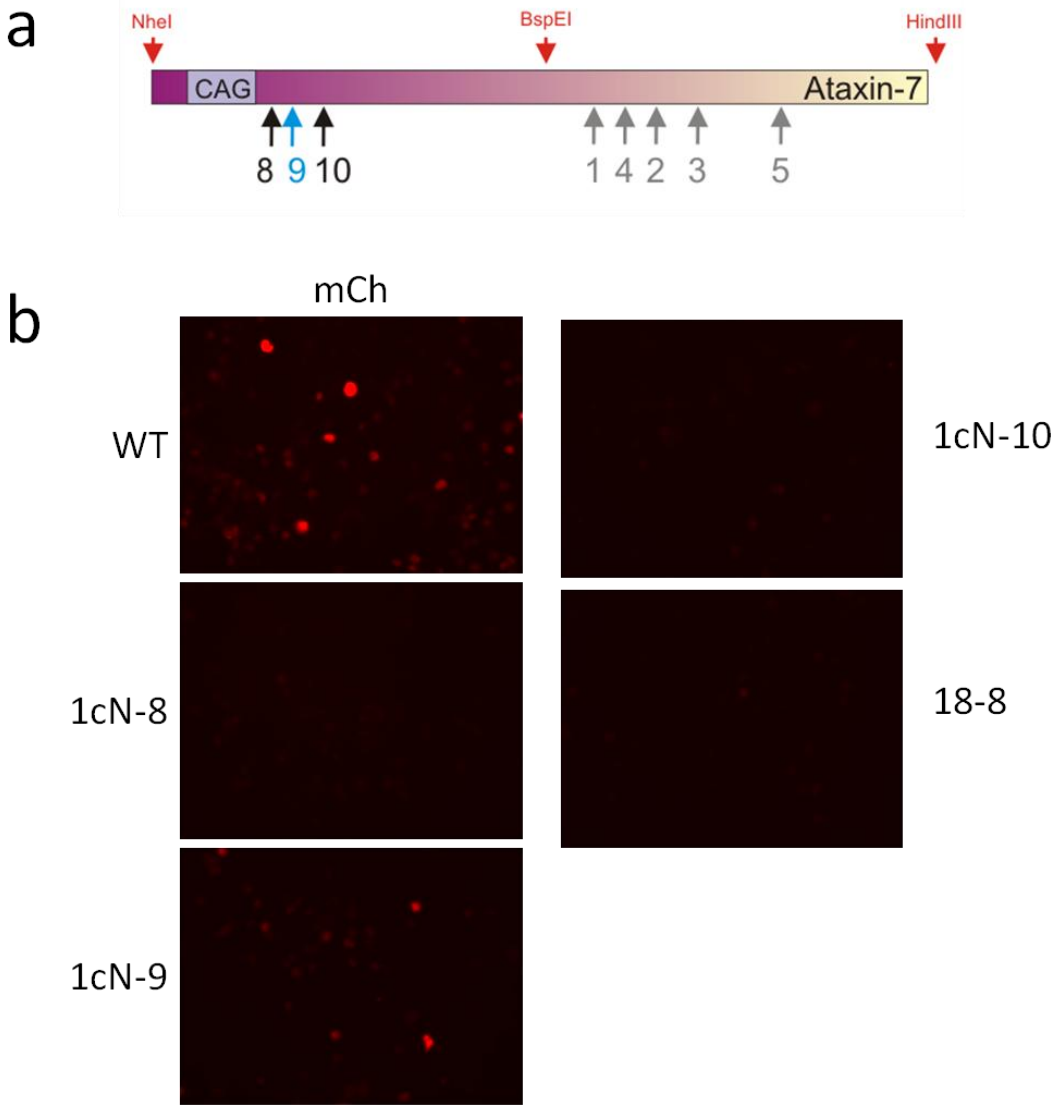
Following the computational search in the previous section, some of the most high-scoring mirtron-splice site combinations (mirt-18 in sites 8 and 9; mirt-1cN in sites 8, 9 and 10) were selected for investigation. All three of these sites lay between the NheI and BspEI sites in the Ataxin-7 construct (Figure 4-6a), therefore difficulties encountered in performing PCR across the CAG-repeat and GC-rich region of the gene were overcome by using commercially synthesised DNA. Nonetheless, mirt-18 in site 9 could not be produced.

Fluorescence was again distinctly lower for the intron-containing constructs than the intron-less Ataxin-7 construct, and was minimal for 1cN-8, 1cN-10 and 18-8 (Figure 4-6b). However, mirt-1cN in site 9 (1cN-9) produced substantial levels of fluorescence. This was confirmed by RT-PCR, which showed a splicing efficiency of approximately 50% (Figure 4-6c). Unfortunately, product detection was very weak for the intron-less (WT) and 18-8 transcripts, and no product was detected for 1cN-8. For 1cN-10, less than 10% of transcripts were processed correctly and a mis-spliced product was formed (Figure 4-6c, arrow).

Crucially, a dual Luciferase assay revealed that 1cN-9 could achieve a substantial level of target silencing (over 70%, $p<0.005$) (Figure 4-6d). Thus, an artificial mirtron may be successfully processed when spliced from Ataxin-7 cDNA. Over 30% silencing was also elicited by 1cN-8, despite the minimal splicing implied by the low fluorescence. No significant silencing was observed for 1cN-10.

The poor splicing efficiency of site 8 was surprising given its similarity to the consensus exonic splice recognition sequences and to site 9 (section 4.4.3). RNA secondary structure can be an important factor in splicing and even small (6-nt stem) hairpin structures can mask the 5' splice site in eukaryotes (Goguel et al. 1993). Therefore, secondary structure

prediction was used to investigate the likely local folding of the Ataxin-7 transcript, when mirt-1cN was included in either site 8 or 9 (Figure 4-7). This indicated that for site 8, all three key bases of the exon/intron boundary (G/GU) were incorporated into hairpin structures, whereas for 1cN-9, only one of the equivalent three bases was base-paired Thus, the secondary structure of the transcript around site 8 may contribute to the observed low splicing efficiency.



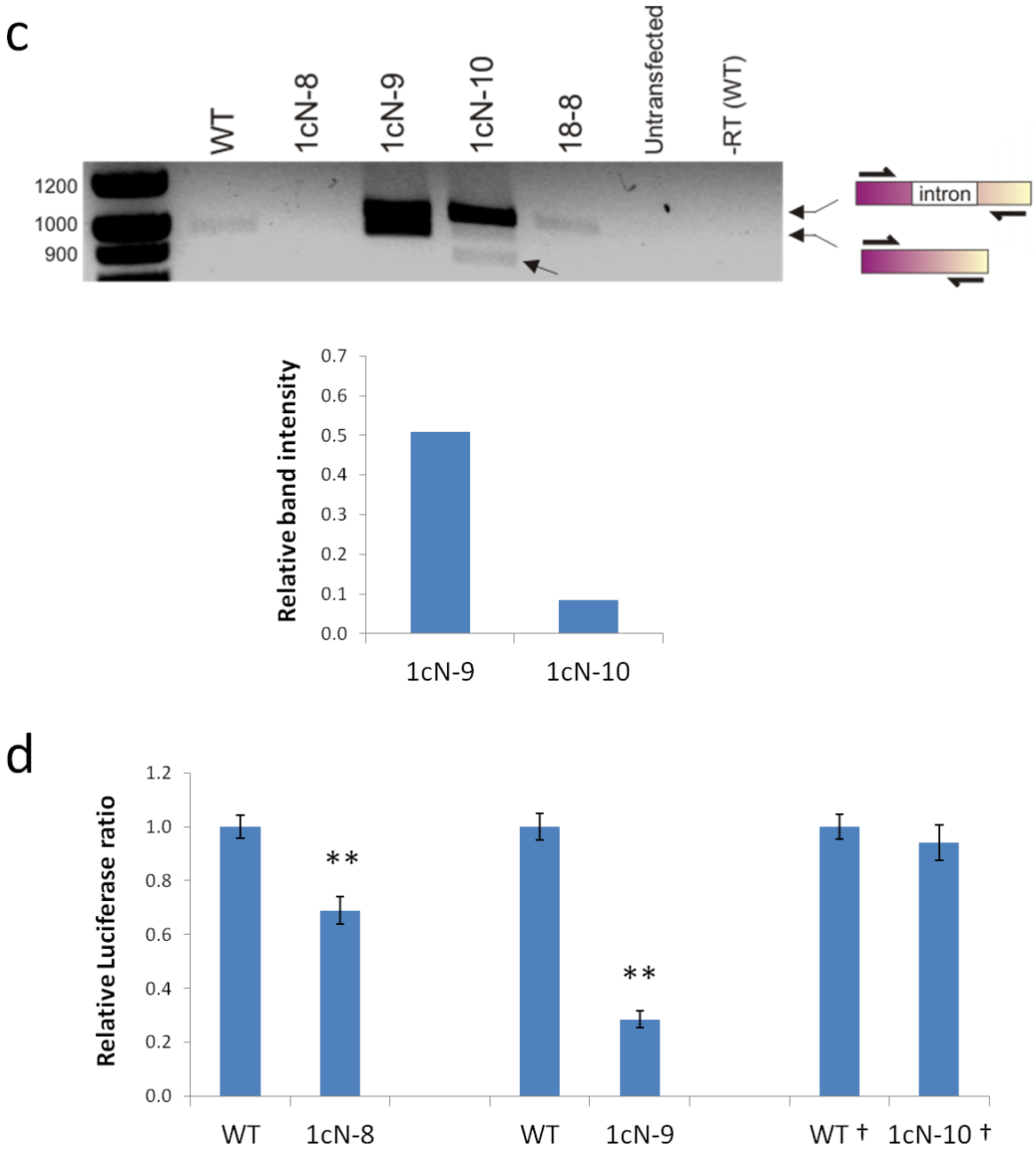


Figure 4-6 - Mirtrons were cloned into the most high-scoring natural splice sites within Ataxin-7. (a) Schematic diagram showing the Ataxin-7 cDNA construct and the exon-exon boundary sites tested (labelled 1-5 and 8-10). The predicted site (9) is shown in blue. "CAG" represents the CAG-repeat and GC-rich region. Key restriction sites are also shown (red arrows). **(b)** Representative fluorescent images for HEK-293 cells 48 h post-transfection with the mCherry-tagged WT and modified Ataxin-7 constructs (mCh), also shown merged with the corresponding brightfield images (merge). **(c)** RT-PCR was carried out using primers spanning all relevant introns, to demonstrate splicing efficiency. RNA was extracted from HEK-293 cells 48 h post-transfection. Bands represent un-spliced and spliced products as indicated by the schematic diagram. The arrow indicates a mis-spliced product. The image was enhanced for visibility of the weakest bands. Quantification of splicing efficiency for two

of the RT-PCRs is also shown, calculated as a proportion of spliced product to the total amount of transcript (spliced + un-spliced). Note: quantification was carried out before image enhancement, when pixels were not saturated. **(d)** *Renilla*:firefly ratios for matched Luciferase targets when co-transfected with Ataxin-7 plasmids containing mirtrons compared to the intron-less wild-type construct (WT, normalised to 1). HEK-293 cells were harvested 48 h post-transfection. N=6 from two independent experiments except where indicated (†, N=3). Error bars represent standard deviation. Asterisks (**) indicate $p<0.005$ (one-tailed t-test).

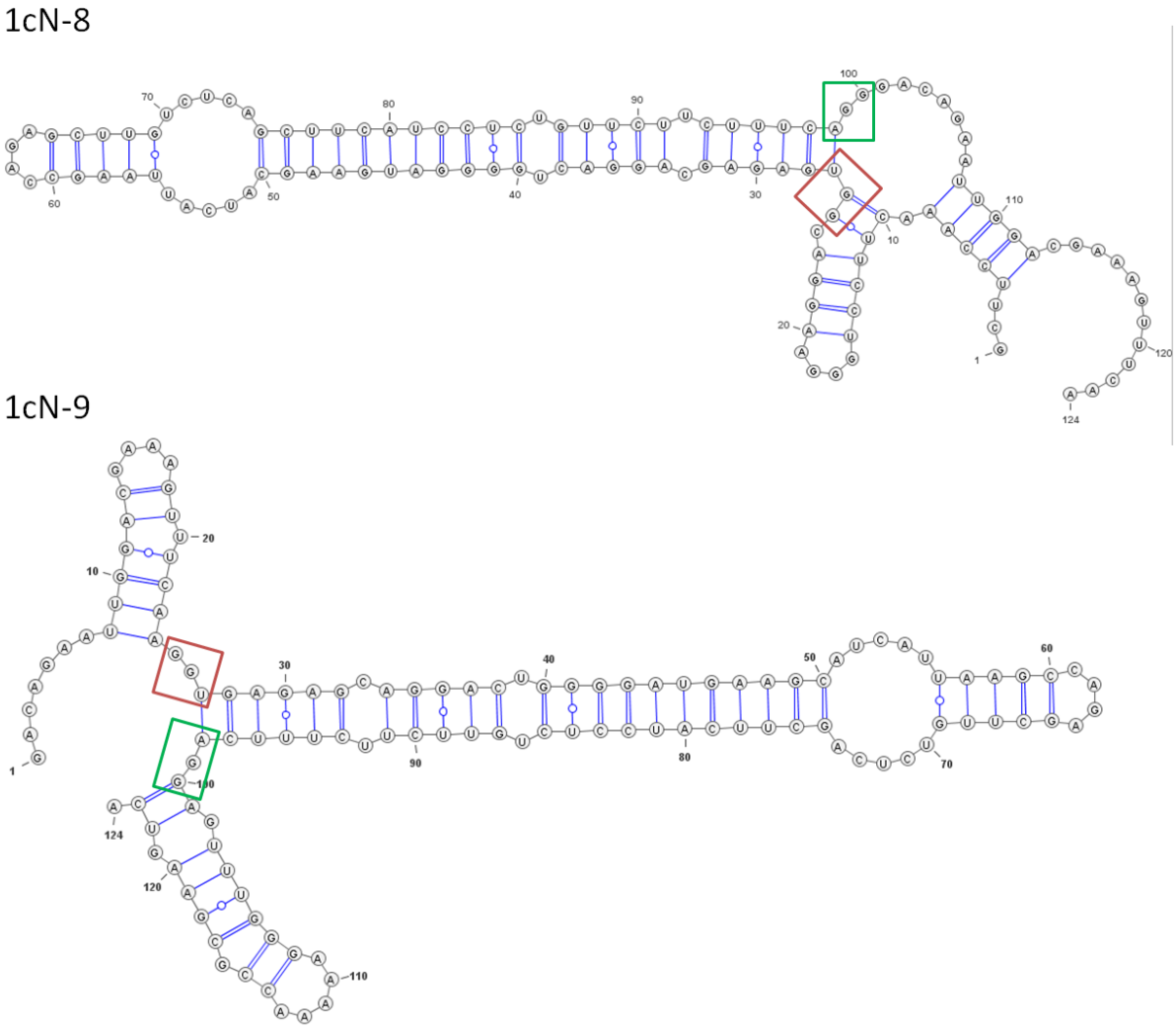


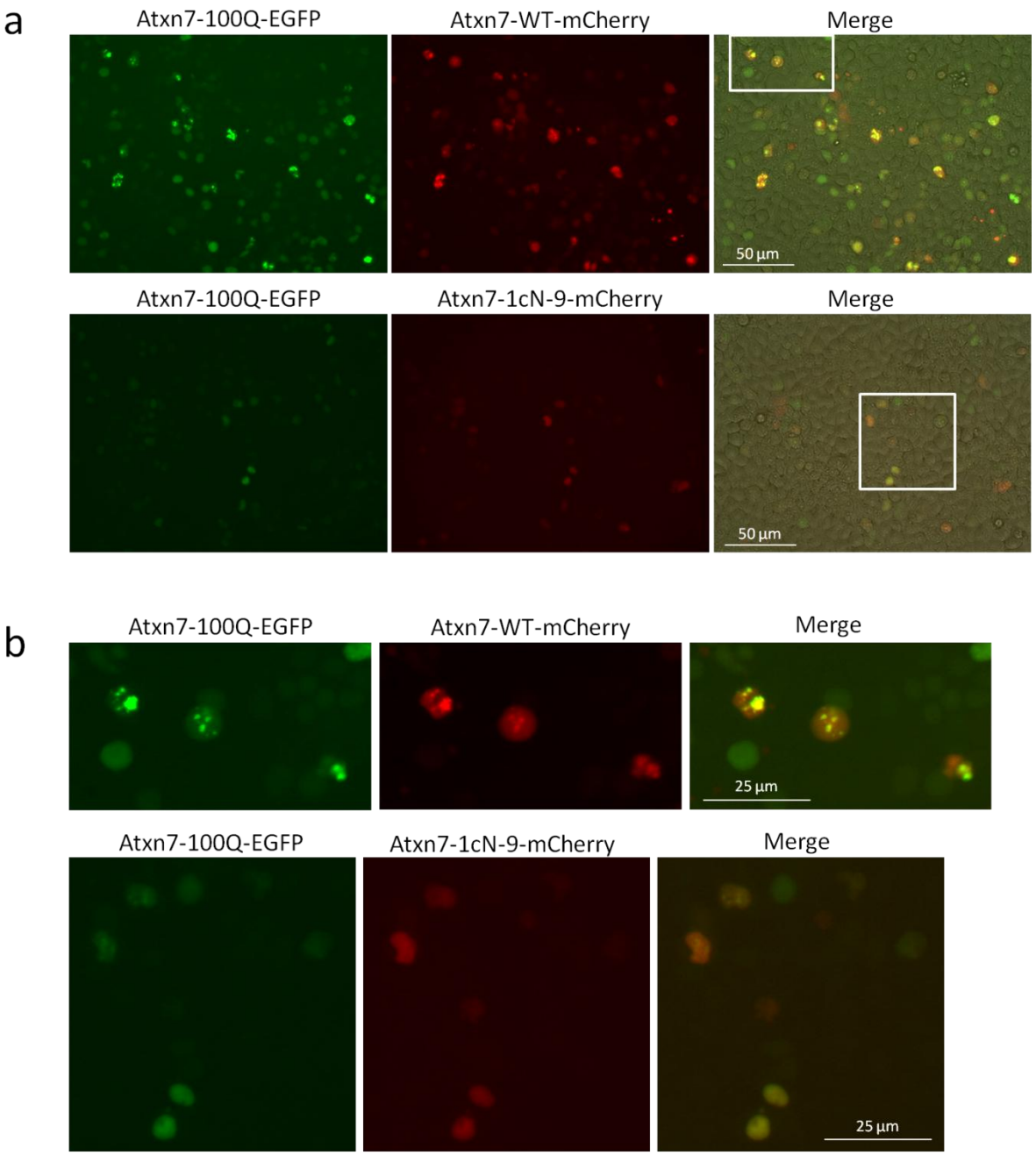
Figure 4-7 – Local RNA structure predictions for mirt-1cN within sites 8 and 9 of Ataxin-7. The RNAfold algorithm (Vienna RNA 2.0 package (Lorenz et al. 2011), available at http://rna.tbi.univie.ac.at/) was used with default settings to predict the secondary structure that contributes a minimum of free energy (Zuker and Stiegler 1981) for the transcript containing mirt-1cN in the context of either site 8 or 9 in Ataxin-7. Including 25 bases of each flanking exon was found to represent the local structure well, compared to including a longer sequence. The structures were visualised using the freely available VARNA applet (Darty et al. 2009). Bases are numbered (labelled every 10 bases) in the 5'-3' direction. Double lines indicate strong G-C base pairs, single lines indicate A-U base pairs and circles indicate weak G-U mismatches. The red boxes indicate the G/GU exon/intron boundary of the 5' splice site, and the green box indicates the AG/G intron/exon boundary of the 3' splice site.

4.4.5 Concurrent silencing and replacement with a mirtron delivered within Ataxin-7 cDNA

To further probe the knockdown-and-replacement ability of 1cN-9, this construct was co-transfected into HEK-293 cells with Ataxin-7-Q100-EGFP. By fluorescence microscopy, levels of mutant Ataxin-7 were visibly reduced in the presence of 1cN-9 than with an equivalent intron-less Ataxin-7-mCherry construct (Figure 4-8a). The presence of aggregates was also visibly reduced (Figure 4-8b). Silencing was quantified as a 40% reduction in total EGFP fluorescence ($p<0.005$) (Figure 4-8c). Although mirtrons could not be compared in the same experiment due to their expression of EGFP, an shRNA expressing the same mirt-1cN mature species (Figure 3-4) produced an 80% reduction in pAtaxin-7-Q100-EGFP (Figure 4-8c) and in similar experiments, mirtrons achieved 50-60% silencing of Ataxin-7-Q100-mCherry (Figure 3-4). The silencing activity of 1cN-9 is likely to be limited by its approximate 50% splicing efficiency, and weaker cytomegalovirus (CMV) promoter compared to the U6-driven shRNA promoter (Lebbink et al. 2011). Nonetheless, these results show that this construct is capable of concurrently silencing mutant Ataxin-7 and replacing it with a wild-type, RNAi-resistant Ataxin-7 cDNA.

In order to verify that the mature mirt-1cN species was produced from the 1cN-9 construct, high-throughput sequencing was performed on ~18- to 25-nt RNA extracted from HEK-293 cells transfected with the 1cN-9 construct. A total of 326 sequencing reads aligned to the mirt-1cN sequence (Figure 4-9a). This was lower than that obtained for pEGFP-mirt-1cN in the same experiment, by approximately 4-fold (Figure 3-7), which was expected given its lower expression levels and poorer splicing efficiency. 95.1% of reads were derived from the 5’ (guide) strand, which was comparable to pEGFP-mirt1cN (98.6%) (Figure 4-9b). 88.7% of guide strand sequences began with the intended residue, which is defined by splicing. This

was lower than pEGFP-mirt1cN (98.8%; Figure 4-9b). Given the accuracy of the splicing reaction, this reduced definition of the 5' end may indicate that some transcripts were processed by the microprocessor (Wu et al. 2009b); the lower splicing efficiency could perhaps lead to more un-spliced, mirtron-containing transcripts being available for recognition by the microprocessor.



C

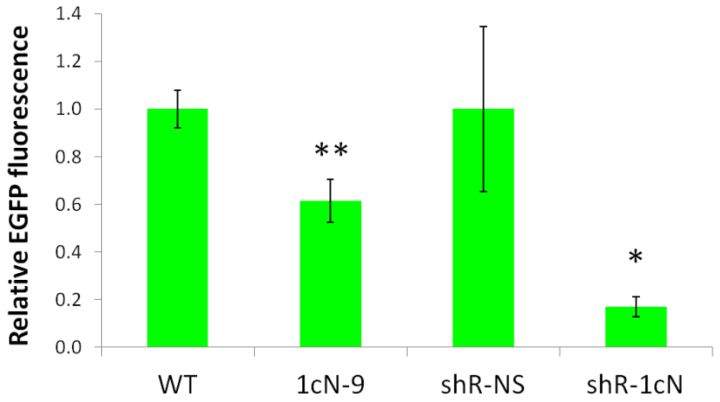


Figure 4-8 – Concurrent silencing and replacement of Ataxin-7 in HEK-293 cells with a single construct. (a) Representative images of HEK-293 cells, 48 h post co-transfection with Ataxin-7-Q100-EGFP (green channel) and an Ataxin-7-WT-mCherry (Q10) construct either containing no intron, or containing the M1cN modification and mirt-1cN in site 9 (red channel). Images are also shown merged with brightfield view. **(b)** Enlargements of the areas indicated in (a). Red and green channels are shown, along with red-green merged images. **(c)** Ataxin-7-Q100 levels are represented by EGFP fluorescence in protein extracted from HEK-293 cells 48 h post co-transfection as in (a), to quantify silencing achieved by the 1cN-9 construct and an equivalent shRNA. 200 µg total protein was loaded and fluorescence was measured using a plate reader. Background fluorescence (un-transfected cells) was subtracted from all readings and control values (WT and shR-NS) were normalised to 1. N=6 for Ataxin-Q10 transfections (two independent experiments) and N=3 for shRNAs. Error bars represent standard deviation. Asterisks (*/**) indicate $p<0.05$/$p<0.005$ respectively, compared to WT (one-tailed t-test).

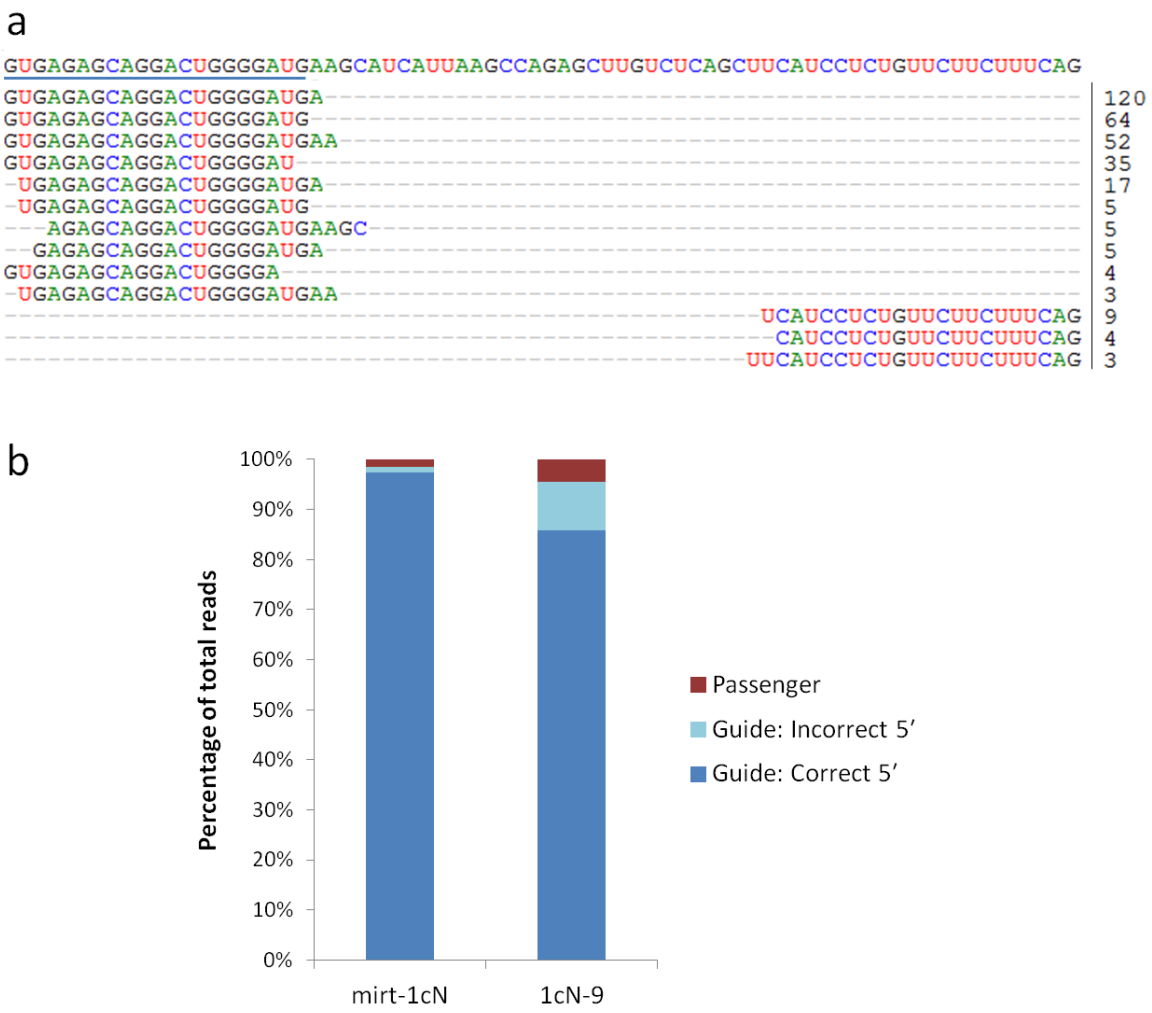


Figure 4-9 - Deep sequencing reads obtained from HEK-293 cells transfected with pAtaxin-7-1cN-9. RNA was extracted 48 h post-transfection and high-throughput sequencing was performed on ~18-24 nt sequences. **(a)** Sequencing reads that matched the mirt-1cN sequence were collected and are shown in alignment with the mirtron sequence. Non-templated U and A bases (added post-Dicer processing; discussed in section 3.3.2) were removed. Bases are coloured according to identity and the designated guide strands are underlined. The number of reads detected is listed on the right to indicate the abundance of each species. (b) For pEGFP-mirt-1cN (data presented in full in Figure 3-7) and for 1cN-9, the number of reads derived from the guide and passenger strands is presented as a proportion of the total number of reads aligned to mirt-1cN. Guide strand reads are divided according to whether or not they begin at the correct 5′ base.

4.4.6 Alternative strategies for combinatorial delivery

Given the complexities involved in identifying efficient splice sites for mirtrons within the coding region of the Ataxin-7 cDNA, alternative strategies were also investigated for combining mirtrons and gene replacement.

4.4.6.1 Masking cryptic splice sites

When mirt-1cN was inserted into site 10, RT-PCR demonstrated that one splice product formed was smaller than expected (Figure 4-6c, arrow). Sequencing revealed that this product was the result of a mis-splicing event, as it lacked a stretch of 69 bases immediately upstream of the mirtron (Figure 4-10a). The cryptic splice site being activated corresponded to site 8. It was postulated that this site could be masked by making silent mutations to prevent it from being recognised over the intended site (1cN-10-8w, Figure 4-10a). However, this did not result in any improvement in mCherry fluorescence (Figure 4-10b) but there was a small increase in silencing activity (Figure 4-10c). RT-PCR indicated that the cryptic splice product was no longer being formed, but no increase in the desired product was seen (Figure 4-10d). This may represent reduced PCR efficiency, given that the un-spliced band was also weaker for 1cN-10-8w than 1cN-10. It appears likely that other factors remain limiting for correct splicing of mirt-1cN from this site.

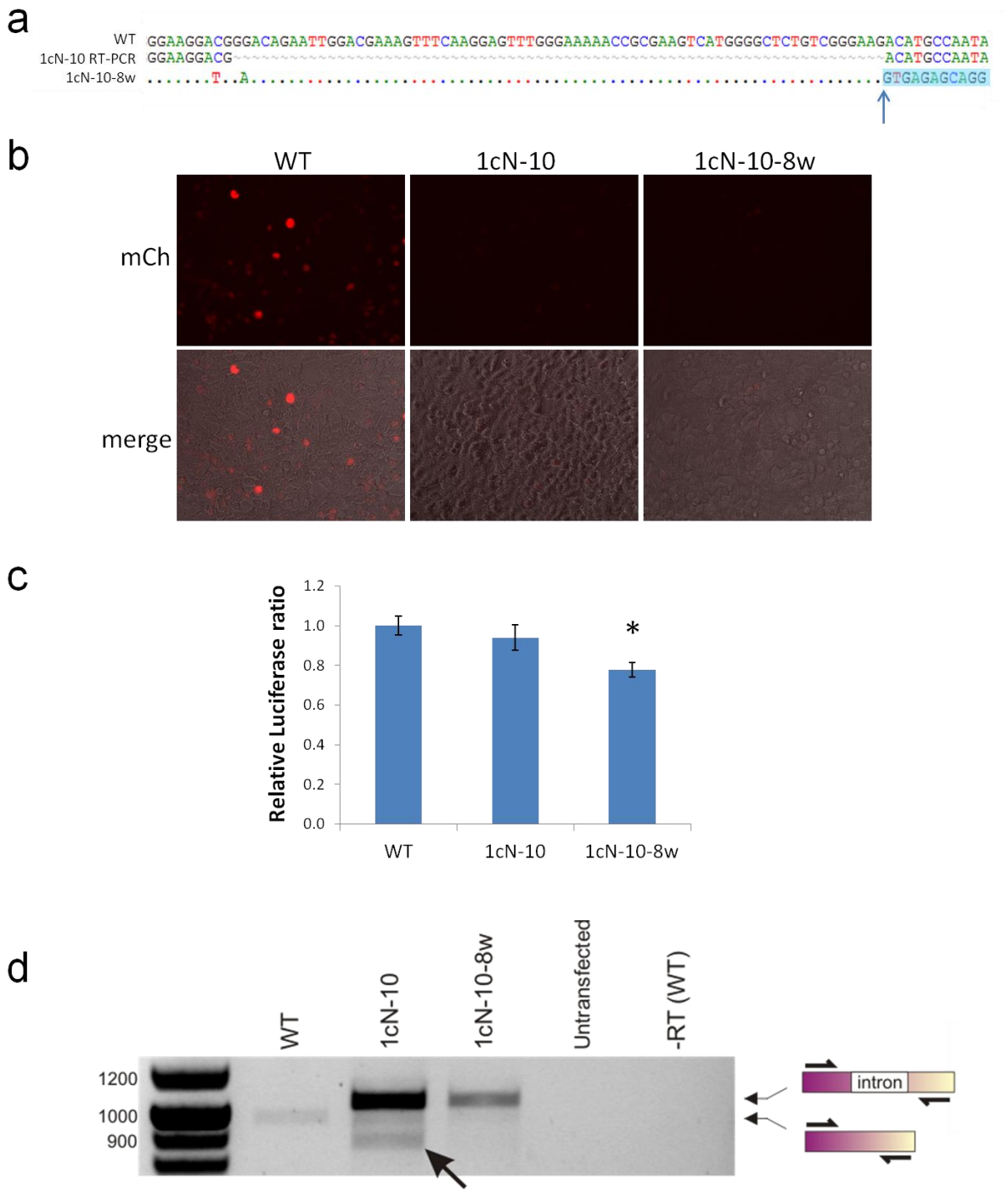


Figure 4-10 - Masking a cryptic splice site in Ataxin-7. (a) Sequence alignment of wild-type Ataxin-7 with the major 1cN-10 splice product indicating the missing bases (~). The position of site 10 is indicated (arrow). The base changes made to 1cN-10 to weaken the recognition of the cryptic splice site are also shown (1cN-10-8w). Dots indicate perfect alignment with the wild-type sequence. The start of the mirtron sequence is highlighted. **(b)** Representative images of HEK-293 cells 48 h post-transfection with 1cN-10-8w, its parent (1cN-10) or the wild-type (WT) construct, represented by mCherry fluorescence, also shown merged with the brightfield images. **(c)** Normalised *Renilla*:firefly Luciferase ratios for K&R constructs against their matched target (Tar1), measured in HEK-293 cells 48 h post-co-transfection.

Values were normalised to the control (WT). N=3. Error bars represent standard deviation. Asterisk (*) represents $p<0.05$ compared to WT (one-tailed t-test). **(d)** RT-PCR was carried out using intron-spanning primers on cDNA from transfected HEK-293 cells, to demonstrate splicing efficiency. RNA was extracted 48 h post-transfection. Bands represent un-spliced (intron-retained) and spliced products, as indicated by the schematic diagram. The mis-spliced product is also indicated (arrow).

4.4.6.2 Placing introns in the 5' UTR

Many genes carry introns within their UTRs, particularly the 5' UTR (Hong et al. 2006). However, adding introns to the 3' UTR of a transgene may cause transcripts to be targeted for nonsense-mediated decay (Nott et al. 2003). The 5' UTR may therefore theoretically be used for the delivery of artificial mirtrons. The benefit of this is that splice sites do not need to be identified within the coding sequence of the gene, yet the length of the construct need not be substantially increased. The consensus sequences for splice recognition sites within the 5' UTR are similar to those of the coding region (Eden and Brunak 2004). Therefore, once it is known that an artificial mirtron splices well from EGFP, a short section of the flanking exons may be taken into the 5' UTR of the desired transgene, to provide well-recognised splice sites.

To test this hypothesis, the NAD intron was cloned into the 5' UTR of Ataxin-7, along with 8 bp from each flanking exon of EGFP (Figure 4-11a). 8 bp may be considered a generous allowance since exonic base identities are not strictly defined by splicing constraints beyond 2-3 bp upstream and 1 bp downstream of an intron, as discussed in section 4.4.3. However, for the case of mirtrons, given the context-dependence seen in sections 4.4.1-0, taking a slightly larger sequence may provide a more favourable context for splicing. Limiting the increase in size to only 16 bp maintains ease of cloning.

mCherry fluorescence for the intron-containing construct was comparable to the WT construct (data not shown); however, this did not imply successful splicing of the NAD intron

from the 5’ UTR, because the intron did not disrupt the open reading frame (ORF). Splicing was therefore analysed by RT-PCR (Figure 4-11b). Although it was difficult to distinguish the spliced product (84 bp) from primer-dimer, the un-spliced product appeared to be the dominant product for the NAD-containing construct (Atxn7-NAD-UTR4), and no smaller band was seen. There was some plasmid carryover, as seen through the band visible in the -RT reaction, but this band was much weaker than that from the positive reaction, therefore likely has little impact on the result. It was concluded that this strategy was not successful for mirtron delivery.

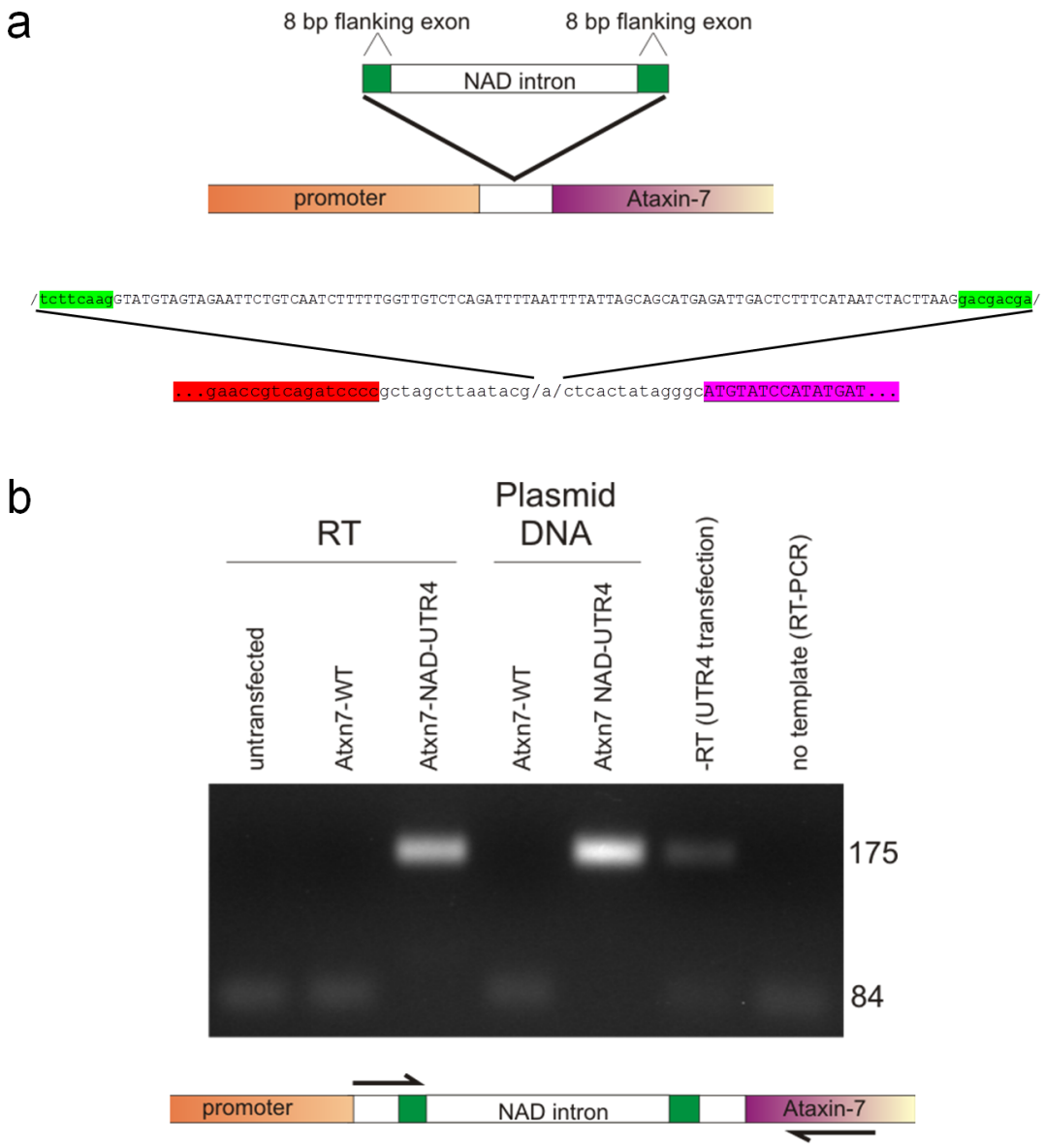


Figure 4-11 –The NAD intron was cloned into the 5' UTR of Ataxin-7. (a) The NAD intron was cloned into the 5' UTR of the Ataxin-7-10Q-mCherry plasmid along with a total of 16 bp of the flanking exons from the EGFP construct, shown in both schematic and sequence formats. **(b)** RT-PCR demonstrating splicing efficiency. Primers spanning the intron (as indicated in the schematic diagram, bottom) were used on cDNA from transfected HEK-293 cells. RNA was extracted 48 h post-transfection. The forward primer was not expected to bind to the wild-type endogenous or plasmid sequence. The predicted RT-PCR product sizes are indicated: 175 bp for the NAD-containing sequence before splicing, and 84 bp after splicing).

4.5 Discussion

The work in this chapter investigated the prospects for development of a K&R therapy for SCA7, whereby mirtrons silencing all endogenous Ataxin-7 are co-delivered as introns within an RNAi-resistant wild-type gene which serves as a replacement.

4.5.1 Silent modifications in Ataxin-7 convey resistance to silencing by mirtrons

The redundancy within the genetic code was exploited to design RNAi-resistant versions of the Ataxin-7 gene which may be delivered as replacement genes alongside the Ataxin-7-targeting mirtrons developed in the previous chapter. Through codon-replacement in each mirtron target site, the modified sequences retained the wild-type amino acid sequence, while silencing was substantially reduced. Surprisingly, mirt-18 still appeared to produce significant (though lesser) silencing of the full-length transcript despite eight base changes in the mirt-18 target site. However, only two of these changes fell within the critical seed region for mirt-18. Further changes were therefore made to the target site, but no improvement in resistance to silencing could be confirmed. It is desirable to avoid any silencing of the replacement gene, because as well as reducing the resulting protein levels, some of the mature mirtron species would be sequestered, potentially reducing its intended activity (Ebert et al. 2007). It is therefore important to make the maximum possible number of base changes within the target side to ensure strong resistance (Kiang et al. 2005).

Although the functionality of the replacement proteins was not investigated, some evidence towards this aim was acquired: sequencing of the plasmids indicated that the modifications were present, *in silico* translation produced the correct amino acid sequence and the detection of the mCherry tag indicated that modified Ataxin-7 was successfully transcribed

in-frame (and at comparable levels to the unmodified gene). Although the endogenous functions of Ataxin-7 appear to be complex, some options are available for the development of functional assays, such as interaction with cone-rod homeobox protein (CRX) (Chen et al. 2004). For further development of this therapeutic approach it would be valuable to verify that the modified gene can perform the role of the wild-type gene and has no unexpected effects.

4.5.2 Mirtrons and a replacement gene can be delivered under a single promoter

Mirtrons were co-delivered with Ataxin-7 under a single promoter by cloning an EGFP-mirtron as a tag after the ORF of mirtron-resistant Ataxin-7. Both the expression of replacement Ataxin-7 and successful splicing of the mirtrons were indicated by EGFP fluorescence and RT-PCR. The mirtrons produced from these constructs showed significant silencing activity against Luciferase constructs. This demonstrated that mirtrons can be delivered as introns under the same promoter as a replacement gene and effectively repress a target. This further supports their potential for development into single-construct knockdown-replacement systems.

Mirtrons were less efficiently spliced from EGFP in this context compared to the original EGFP plasmid. Being located several thousand bases from the transcriptional start site (TSS) may be a contributing factor. Splicing occurs co-transcriptionally and the 5’ cap helps recruit spliceosome factors to the first 5’ splice site (Pandya-Jones 2011); decreasing proximity of the mirtron to the TSS may hinder the chances of intron recognition. This is perhaps confounded by secondary structure around the required splice sites, since the NAD intron was still spliced efficiently.

For clinical development it would also be necessary to remove the fluorescent tag and also to produce a construct of minimal size for ease of packaging to a suitable delivery system.

4.5.3 Mirtrons can be effectively spliced from some sites in Ataxin-7

As noted in Chapter 3 (section 3.4), the splicing of mirtrons from a host transcript is context-dependent. This led to difficulties in finding sites within Ataxin-7 from which mirtrons could be spliced efficiently.

For one site (9), mirt-1cN was spliced from the RNA-resistant Ataxin-7 transcript at approximately 50% efficiency. This was successful in achieving silencing of a Luciferase target (70%) and a full-length EGFP-tagged Ataxin-7 cDNA (40%). High-throughput sequencing confirmed that the mature mirt-1cN species was produced, with strand selection comparable to pEGFP-mirt-1cN, but slightly less accurate processing. Its lower splicing efficiency may lead to some mirtron-containing transcripts being available for Drosha processing. Before further development, it will be important to confirm mirtron splicing from this construct in a neuronal cell line, since splice sites may be selectively activated or repressed according to cell type (Keppetipola et al. 2012).

4.5.4 Context dependence of mirtron splicing

Factors influencing splicing are complex (Buratti and Baralle 2004) and computational analysis of splice sites proved a useful tool, yet some sites, despite scoring well, did not produce good splicing efficiency.

It appeared that placing a mirtron several kilobases from the promoter reduced its splicing efficiency, as seen in the pAtaxin-7-EGFP-mirtron constructs (section 4.3). This occurred only for mirtrons and not the NAD intron, hinting that secondary structure could also be involved. Interestingly, the most successful site within the coding region of Ataxin-7 (site 9) was the

second most promoter-proximal site tested, approximately 400 bases downstream of the TSS. However, neither mirtron was spliced efficiently from the site closest to the TSS (site 8) despite having exonic splice sequences which matched the consensus and scoring well *in silico*. Nonetheless, placing mirtrons within promoter-proximal sites may be a beneficial strategy.

The formation of secondary structures can affect splice site recognition (Goguel et al. 1993; Buratti and Baralle 2004; Marashi et al. 2006). The hairpin structure of a mirtron, by bringing the neighbouring exons together, may alter local secondary structure of the transcript and therefore potentially mask splice sites in a context-dependent manner. Computational prediction indicated that this may account for the difference in splicing efficiencies of mirt-1cN from sites 8 and 9. It may be crucial, therefore, to incorporate such analysis in the design of mirtron-based K&R therapies. However, interpretation of secondary structures and the potential effects on splicing is not straightforward, and results can vary based upon the length of transcript included in the analysis.

Mutations weakening the secondary structure of mirt-18 in the context of EGFP led to increased splicing efficiency (section 3.3.4). This approach could be used to improve the potential for mirtron splicing in the context of Ataxin-7. However, this would require extensive optimisation in order to maintain efficient miRNA processing and silencing ability. Further, given that mirt-1cN splices almost completely from the original site within EGFP (Figure 3-5), it would be difficult to identify in this way any artificial mirtrons which were more likely to splice well from different contexts.

An additional option may be to strengthen certain splice sites (perhaps those which give favourable secondary structures) by improving their similarity to the consensus sequence. It is also possible to mask cryptic splice sites which are unintentionally activated. This was

attempted in the case of 1cN-10, but did not improve splicing at the desired site. For both of these strategies, the limitations of codon replacement apply.

This complexity highlights the need for better understanding and prediction of mirtron splicing efficiencies. This currently makes mirtron-based K&R a difficult strategy to apply in practise. However, more thorough computational prediction may improve this situation and the constantly decreasing cost of gene synthesis allows much easier screening of different intron locations.

4.5.5 Delivering mirtrons within UTRs

Inserting introns into the UTRs could be an alternative strategy to co-deliver mirtrons along with a transgene without needing to identify exon-exon-boundaries which are amenable to mirtron splicing. Once a mirtron is known to be efficiently spliced, it may even be transferred to a UTR along with a stretch of each flanking exon to ensure a good context for splicing. The 5' UTR was chosen to minimise the chances of triggering nonsense-mediated decay and this also avoids the detrimental effects of placing an intron several thousand bases from the TSS. An additional benefit of choosing the 5' UTR may be the capacity to increase transgene expression without using a stronger promoter (Palmiter et al. 1991).

To test this, the NAD intron was taken along with 8 bp of each flanking exon from EGFP and tested for splicing by RT-PCR. However, no splicing appeared to occur. Perhaps the 16 bp total flanking sequence was not sufficient to recreate the well-recognised context in the EGFP plasmid. However, secondary structure should not be an issue for the NAD intron. The intron may not be a sufficient distance from the end of the promoter, not allowing a large enough length of transcript for the assembly of spliceosome components.

Canonical miRNAs also occur in UTRs (Morlando et al. 2008), therefore delivering artificial canonical miRNAs within UTRs may be an alternative strategy to construct K&R therapies of minimal size.

4.5.6 Tailed-mirtrons

Tailed-mirtrons are another potential option for K&R applications. They are related to mirtrons in that they are dependent upon splicing for their biogenesis, but the splice product contains a tail on either the 5' or 3' end of the hairpin, which requires digestion by an exonuclease other than Drosha. Tailed-mirtrons are more widespread in the human genome than mirtrons (Ladewig et al. 2012), but the use of artificial tailed-mirtrons has not been investigated. They would require addition of only a small number of bases (<100) to the intron, hence still allowing construction of a K&R system of close to minimal size, with a relatively straightforward cloning procedure.

Artificial tailed-mirtrons may be designed either by mimicking endogenous tailed-mirtrons or potentially by adapting artificial mirtrons by adding a tail. However, both options are theoretical strategies and the exact processing result may be hard to predict. Given that functional artificial mirtrons have been identified, adding a tail onto one end of a mirtron may increase its splicing efficiency by preventing the hairpin structure from masking at least one splice site. Mirtrons could be adapted either by adding the tail from an endogenous tailed-mirtron, or retaining one end of the endogenous intron which is being replaced. A number of tailed variants of mirt-18 and mirt-1cN were investigated *in silico* by adding different 5' and 3' tails and measuring the splicing score in the context of various sites. Adding a 5' tail to a 5'-derived mirtron (mirt-1cN) was avoided, since it was thought that this may compromise the accuracy of the seed region. However, none of the options tested gave

splicing scores substantially higher than the parent mirtron. Nonetheless, this remains a promising option which could be further investigated.

4.5.7 Conclusions

Gene knockdown and replacement as a potential therapy for SCA7 was found to be achievable using mirtrons as RNAi effectors, delivered as introns within the replacement gene and effectively silencing reporter genes. Finding sites from which artificial mirtrons were efficiently and accurately spliced was problematic and may lead to difficulties in applying this approach effectively to other genes. The highly context-dependent mirtron splicing means that computational analysis is necessary to identify the best possible intron locations. As well as sequence features, secondary structures and distance from the TSS are also important factors. The development of tailed-mirtrons could help to overcome some of these difficulties by providing one strong splice site. The therapeutic potential of mirtrons for SCA7 will be further investigated in the next chapter.

5 Patient-derived cell models for assessment of novel SCA7 therapies

5.1 Introduction

SCA7, a neurodegenerative disease caused by an expanded CAG repeat in *ataxin-7*, is a potential candidate for gene silencing or gene knockdown-replacement (K&R) therapy. The use of standard cell lines as in Chapters 3 and 4 can indicate the direct effects of such technologies on the expression of transgenic or endogenous genes. However, these cells do not naturally carry the disease-causing mutation, and transgene over-expression does not faithfully represent disease-related phenotypes. Therefore, the therapeutic benefits of novel therapies may be better explored in cells which naturally carry the disease-causing mutation and show measurable phenotypes compared to controls. This could allow greater optimisation and more well-informed selection of novel therapies to test in animal models.

In transfected cells, it has been established that silencing Ataxin-7 can reduce the presence of a key disease feature, nuclear aggregates (Scholefield et al. 2009), as seen in Chapters 3 and 4 (Figure 3-4 and Figure 4-8). In other polyQ diseases such as Huntington's disease (HD), silencing of the mutant protein can provide measurable phenotypic improvements *in vivo* (Yamamoto et al. 2000; DiFiglia et al. 2007; Kordasiewicz et al. 2012). However, there is little data on the efficacy of gene silencing treatment against the phenotype of SCA7 *in vivo*, except in one case, where a cellular response comparable to silencing therapy was found unexpectedly. In the *R7E* retinal SCA7 mouse model, the transcription of transgenic mutant Ataxin-7 was inhibited, attributed to repression of its *rhodopsin* promoter (Helmlinger et al.

2004a). After only 3-4 weeks of expression, transgenic Ataxin-7 mRNA and protein could no longer be detected. Importantly, transcriptional dysregulation and progressive photoreceptor dysfunction persisted, leading to neuronal loss. The nuclear inclusions (NIs) were not cleared in all photoreceptors even after one year. High transgene expression compared to endogenous levels may have contributed to this irreversibility by causing rapid disease progression. Therefore, this does not necessarily represent the potential for disease irreversibility in patients. Nonetheless, once triggered, the transcriptional silencing of photoreceptor-specific genes was not reversed by silencing Ataxin-7. This indicates that it is crucial to test Ataxin-7 silencing therapies very early in the disease progression, which should also be considered in cell models.

Patient skin biopsies are a relatively non-invasive way of obtaining cells from patients, from which it is possible to culture fibroblast populations. Fibroblasts may be transfected with nucleic acids, potentially making them a useful model for testing novel gene therapies for SCA7 (Scholefield et al. 2009). Some phenotypic features may also be present, as in HD fibroblasts (Seo et al. 2004; de Mezer et al. 2011). However, SCA7 primarily affects neuronal cells and phenotypic effects are rarely observed in peripheral tissues. Therefore, disease processes are likely to be more accurately represented in neuronal cell types, but these are difficult to obtain and culture from patients.

Advances in stem cell technology may allow an alternative route for SCA7 patient-derived neuronal cell models. The reprogramming of dermal fibroblasts into induced pluripotent stem cells (iPSCs) through the viral delivery of key reprogramming factors has been demonstrated (Takahashi et al. 2007; Yu et al. 2007). iPSCs may even be derived from autopsy material (Hjelm et al. 2011). iPSCs can be differentiated into specific cell types, including neuronal cultures, through the provision of appropriate small molecule stimuli and

culture conditions (Dimos et al. 2008; Ebert et al. 2009; Dottori and Pera 2008). Encouragingly, disease-associated phenotypes were identified in neurons derived from HD and SCA3 patient cells (Koch et al. 2011; Feyeux et al. 2012). Neuronal cultures have recently been derived from SCA7 patients (Luo et al. 2012; Watson 2012). This provides a human cell model for SCA7 naturally containing an Ataxin-7 expansion but of a more relevant cell type than fibroblast cell lines. This potentially allows the effects of therapies on phenotype to be investigated more extensively than in fibroblasts.

In this chapter, mirtrons targeting Ataxin-7 will be tested in patient-derived fibroblasts and iPSC-derived neurons, in comparison with other Ataxin-7 silencing technologies. In order to transduce patient-derived neuronal cells, lentiviral vectors delivering shRNA and mirtrons were produced. Phenotypic features and their restoration are also investigated in these cell models.

5.2 Patient fibroblasts

5.2.1 Characterisation of patient fibroblast cells

Four adult dermal fibroblast cell lines were available, derived from three different SCA7 patients and one normal control individual (Figure 5-1a). The patient labelled 1518 was the parent of patient 1519; as expected, the expansion size of the child was larger than that of the parent. Expanded alleles and their relative sizes in these cell lines were confirmed by PCR (Figure 5-1b). The normal allele(s) in each case had 10-12 CAG repeats.

a

Fibroblast line	# Q
Control	10
1518	42
1519	51
Coriell	55

b

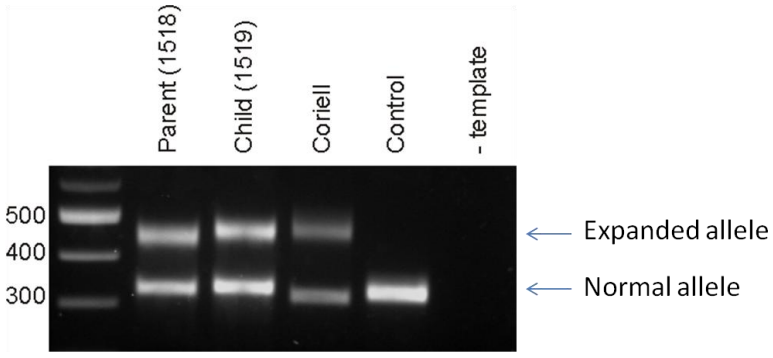


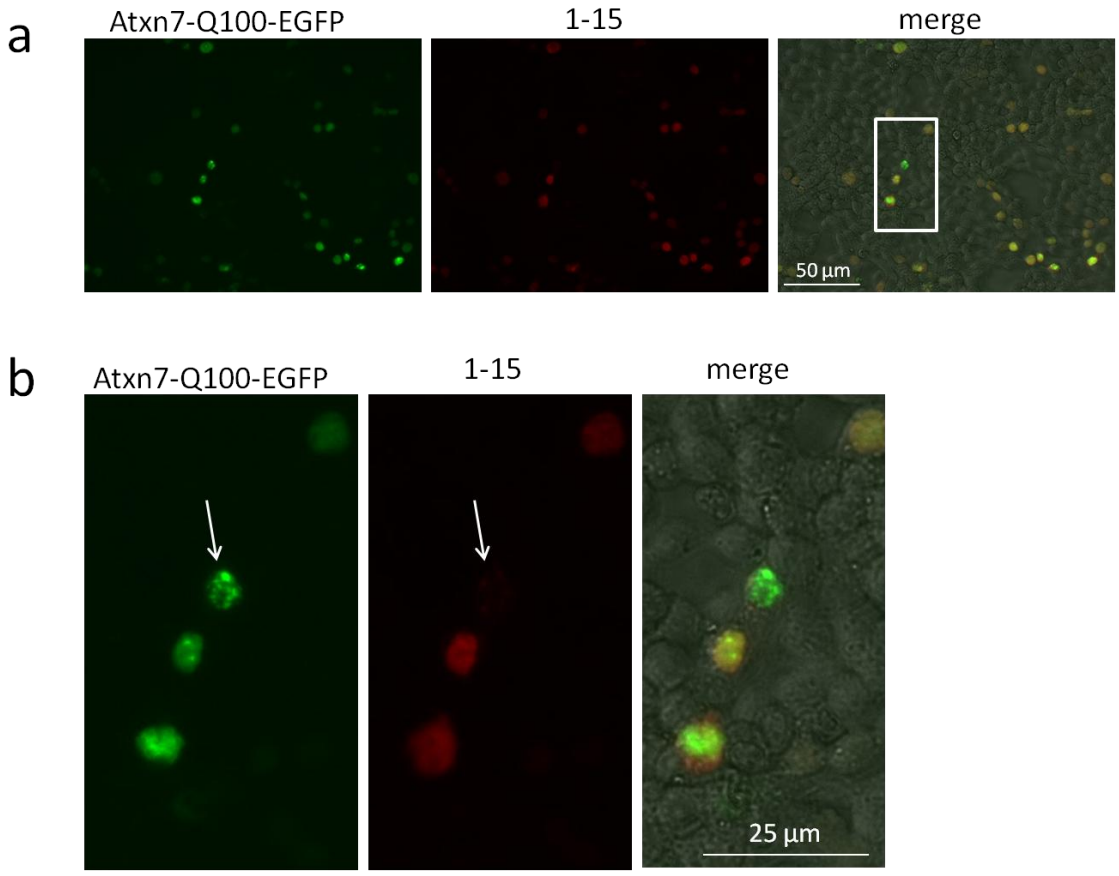
Figure 5-1 – Characterisation of patient fibroblast cell lines. (a) Size of the poly-CAG sequence of the expanded allele in each cell line available, i.e. number of glutamine (#Q) residues in the resultant protein. **(b)** PCR was carried out on genomic DNA extracted from each fibroblast cell line using primers spanning the poly-CAG stretch. The position of the products amplified from expanded and non-expanded alleles are indicated.

5.2.1.1 Antibodies for detection of mutant Ataxin-7

It was hypothesised that patient fibroblasts may exhibit the typical SCA7 feature of NIs containing mutant Ataxin-7, which would provide a useful phenotypic property to monitor. In order to detect expression of Ataxin-7 protein in patient fibroblasts, two anti-Ataxin-7 antibodies (1-15 and 1-135) and the anti-polyglutamine 1c2 antibody (Trottier et al. 1995) were tested. Although the 1-135 antibody was previously confirmed to bind to wild-type Ataxin-7, it was also necessary to confirm that these antibodies can bind the mutant protein, which may have an altered conformation.

All three antibodies were tested for their affinity to expanded Ataxin-7 by immunostaining HEK-293 cells transfected with Ataxin-7-Q100-EGFP plasmids. For 1-15, the immunostaining pattern closely matched the Ataxin-7-EGFP expression pattern, confirming that mutant Ataxin-7 can be detected (Figure 5-2a). However, where EGFP fluorescence appeared in small foci, indicating dense aggregates, immunostaining was weak (Figure 5-2b, arrow). This indicated that the 1-15 antibody was only able to bind soluble Ataxin-7 but not the

aggregated material. Similar results were observed for 1-135 (Figure 5-2c and d), but staining with this antibody generally gave a higher signal. The 1c2 antibody also clearly demonstrated affinity to transfected Ataxin-7. This antibody produced higher background staining compared to the Ataxin-7 antibodies, which appeared as weak nuclear staining (Figure 5-2e), perhaps due to its affinity to TATA-box binding protein (TBP) (Trottier et al. 1995). However, like the 1-15 and 1-135 antibodies, 1c2 did not stain dense aggregates of Ataxin-7 (Figure 5-2f, arrows). It was confirmed separately that the antibodies do not have affinity to EGFP and that 1c2 did not show affinity to Ataxin-7-Q10-EGFP.



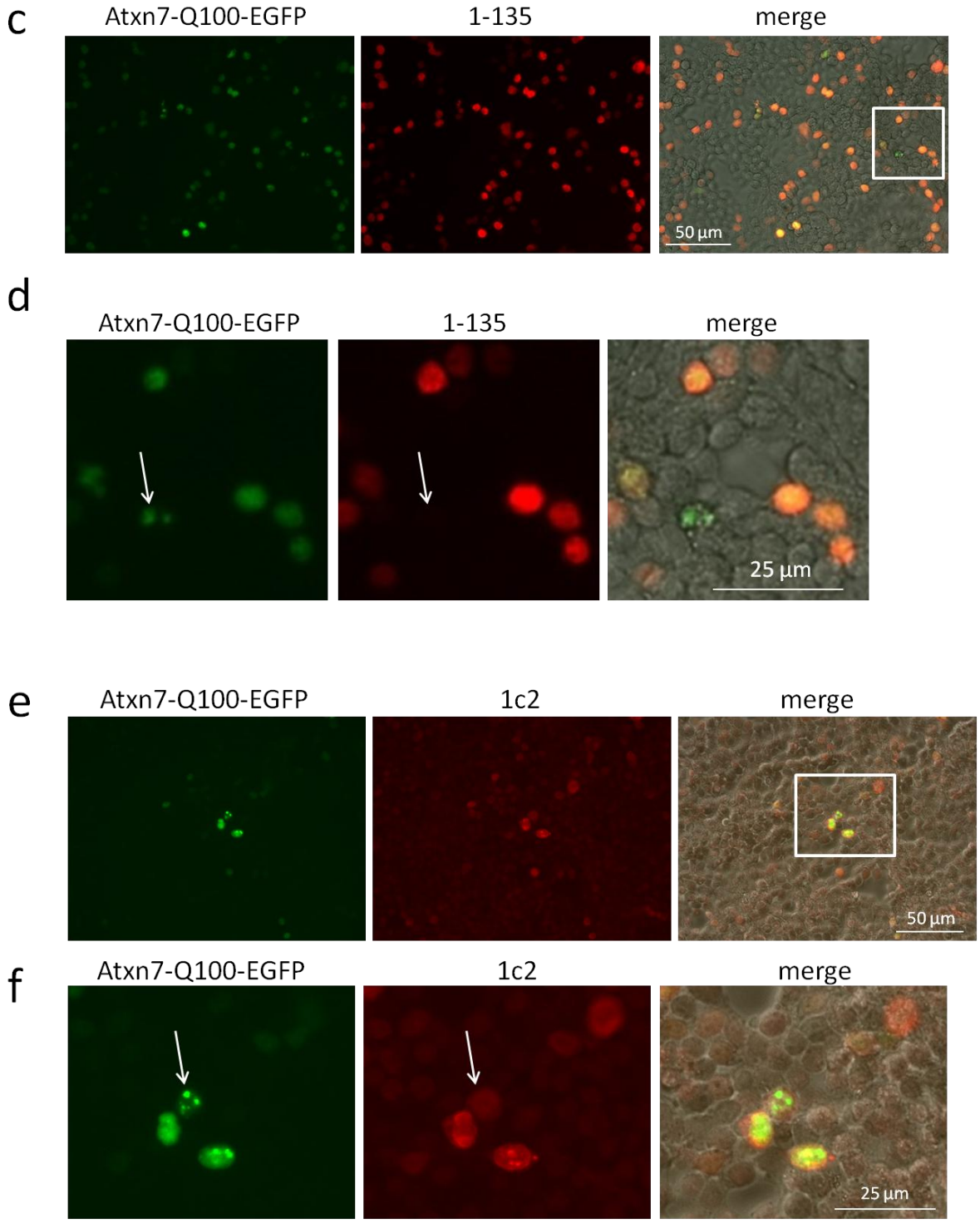


Figure 5-2 – Staining of HEK-293 cells expressing Ataxin-7-Q100-EGFP. HEK-293 cells were fixed 48 h post-transfection with pAtaxin-7-Q100-EGFP and stained with antibodies targeting Ataxin-7 **(a-d)** or the polyQ expansion **(e-f)**. All primary antibodies were used at 1:1000, with an appropriate Alexa-594-conjugated secondary antibody (red channel) at 1:1000. Transfected cells are indicated by their EGFP expression (green channel). **(b,d,f)** show enlargements of the area indicated in (a,c,e), respectively. Arrows indicate cells containing densely aggregated Ataxin-7.

5.2.1.2 Ataxin-7 protein expression in patient fibroblasts

To investigate protein expression and the presence of aggregates, fibroblasts from the Coriell SCA7 patient and the normal control line were immunostained with anti-Ataxin-7 (1-135) and anti-polyQ (1c2) antibodies. The 1-135 antibody did not reveal any signs of nuclear aggregation (Figure 5-3a and b). Although bright speckles were occasionally visible, these appeared in both cell lines (Figure 5-3b), therefore were likely an artefact of the staining procedure. This result was not unexpected given that peripheral tissues rarely exhibit NIs in patients, except in cases of very large expansions (Ansorge et al. 2004).

The anti-polyQ 1c2 antibody can detect expanded (but not normal) Ataxin-7 (Trottier et al. 1995). However, the staining pattern was not distinguishable between patient and normal fibroblasts (Figure 5-3a and b). In both lines, 1c2 staining appeared as weakly nuclear with strong perinuclear staining, strongly resembling the Golgi apparatus when compared to a similar cell line (Figure 5-3c). This could perhaps be explained by detection of Ataxin-2, which normally resides in the Golgi and typically has a polyQ region of 22 repeats (Huynh et al. 2003), which is perhaps within the detection range of 1c2. Interestingly, the Golgi accumulates and packages proteins into aggregates to then send them for secretion. Therefore, this staining could correspond to an enrichment of proteins with long polyglutamine tracts in the Golgi of these cell lines.

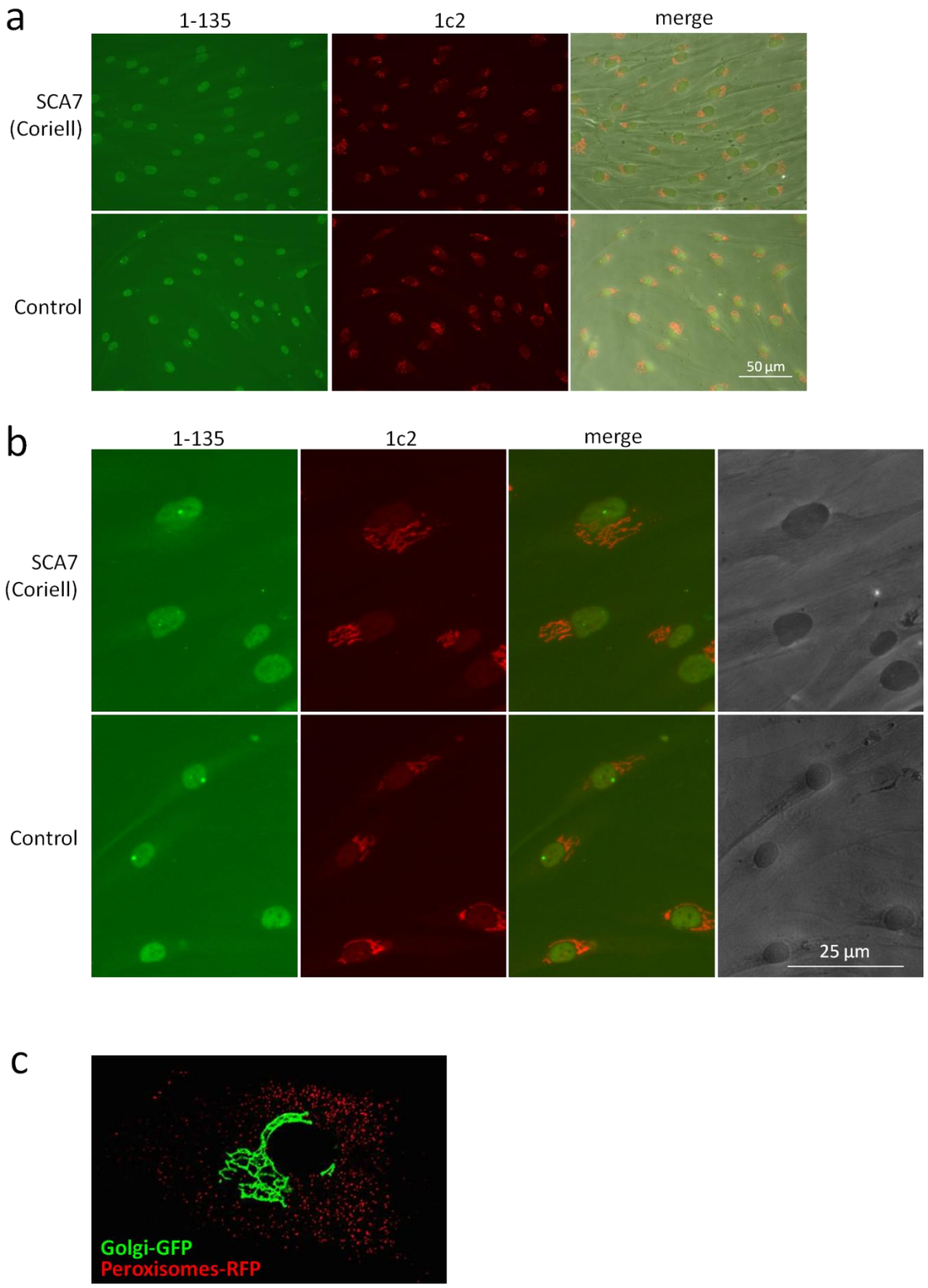


Figure 5-3 – Staining of SCA7 patient versus normal control with Ataxin-7 antibodies. (a) Fibroblasts cells from the Coriell (55Q) and control (10Q) lines were plated in 24-well plates and immunostained with the 1-135 antibody at 1:100 and 1c2 antibody at 1:500, along with Alexa-488 (green channel) and Alexa-594 (red channel) fluorescent secondary antibodies,

respectively. Images are also shown merged with the phase contrast view (merge). **(b)** Enlarged images as in (a), except green-red merged images are shown, along with phase contrast view. **(c)** Fluorescent image demonstrating the appearance of Golgi in human (neonatal) dermal fibroblasts. Source: www.invitrogen.com.

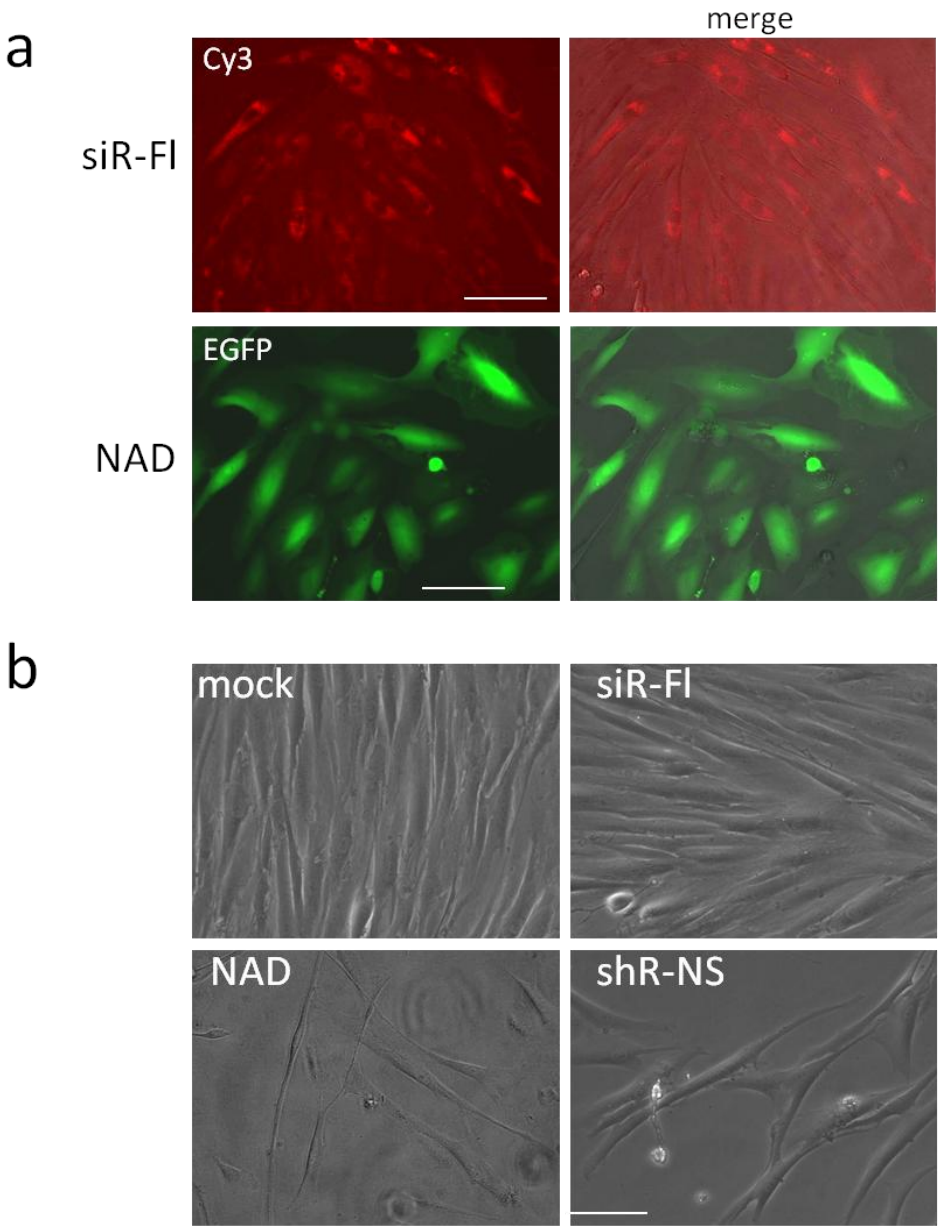
5.2.2 Silencing Ataxin-7 in patient fibroblasts

5.2.2.1 Silencing of ataxin-7 mRNA

Fibroblasts were transfected with siRNA and plasmids to assess their silencing ability against Ataxin-7 in these patient-derived cells, initially at the mRNA level. Using electroporation, transfection efficiencies were high (approximately 90%) for siRNA and moderate (~50%) for plasmids (Figure 5-4a). However, cell death was greater for fibroblasts transfected with shRNA or mirtron plasmids than siRNA-transfected cells, leading them to appear more sparse (Figure 5-4a). Even 7 d post-transfection, cells transfected with plasmids were still relatively sparse, whereas cells which received siRNA or underwent mock electroporation treatment (no nucleic acid) had grown normally and were confluent (Figure 5-4b). This toxic effect appeared to be similar regardless of which plasmid was transfected, indicating that it is likely caused by the combination of electroporation treatment with plasmid delivery, and not a plasmid- or transgene-specific effect.

Following successful transfection, silencing abilities were assessed. siR-P16, which targets both alleles of Ataxin-7 at the concentrations used (Scholefield et al., in preparation), produced a decrease of more than 80% in the total (normal plus mutant) *ataxin-7* mRNA levels, even after 7 days ($p<0.005$) (Figure 5-4c). Silencing with shRNA (shR-1) was more modest, at approximately 50% after 7 d ($p<0.05$) (Figure 5-4d). This reduced level of silencing may be largely explained by the lower transfection efficiency of plasmids compared to siRNA. However, no silencing was observed for mirtrons at this time-point (Figure 5-4e); this

was also confirmed in another patient fibroblast line. However, measured instead at 4 d post-transfection, approximately 50% silencing was detectable in two independent experiments ($p<0.005$, Figure 5-4f). A possible explanation could be that plasmid-transfected fibroblasts become stressed after an extended period of time in a sparse culture, which may affect gene expression levels.



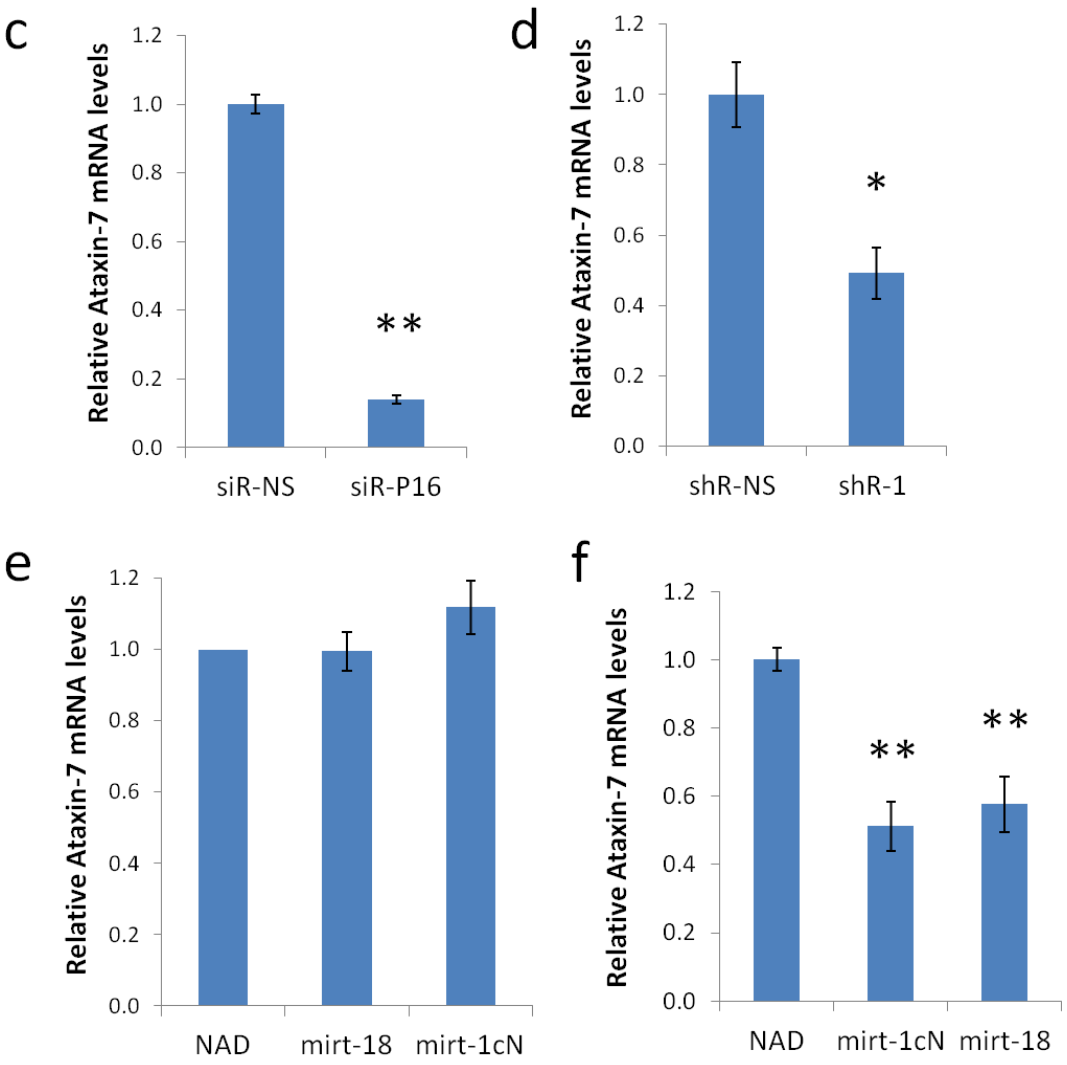


Figure 5-4 – Silencing of Ataxin-7 mRNA in patient fibroblasts. (a) Images of transfected Coriell SCA7 fibroblast cells 48 h post-electroporation. Cells were electroporated under identical conditions with a Cy3-labelled siRNA (siR-Fl) or pEGFP-NAD. Fluorescent images are also shown merged with brightfield view (merge). Scale bar = 50 µm. **(b)** Phase contrast images of electroporated cells as in (a), 7 d post-transfection. Also represented are cells electroporated in the absence of nucleic acid (mock) and shRNA-transfected cells (shR-NS). Scale bar = 50 µm. **(c)** *Ataxin-7* was detected by qPCR, carried out on RNA extracted from 1518 fibroblasts transfected with siRNA at 20 nM. **(d)** *Ataxin-7* qPCR was carried out on RNA extracted from Coriell fibroblasts 7 d post-transfection with shRNA or **(e)** mirtrons at 1 µg/ml. For (c)-(e), Ataxin-7 levels were normalised to internal control gene β-actin and N=2, except for NAD in (e), where one sample produced very high Ct values for both genes and was discarded. **(f)** *Ataxin-7* qPCR was carried out on RNA extracted from Coriell fibroblasts, 4 d post-transfection with mirtrons at 1 µg/ml. This experiment was carried out by M. Varela and TaqMan assays were used for both *ataxin-7* and GAPDH internal control. N=4 (two independent experiments). For (c)-(f), error bars represent standard error and asterisks (*/**) represent p<0.05/0.005 respectively.

5.2.2.2 Protein level silencing

Following the detection of silencing of Ataxin-7 at the mRNA level, it was then sought to verify that this leads to a reduction at the protein level. Although protein-level silencing could not be observed in standard cell lines up to 4 d post-transfection, fibroblasts were cultured for longer since they cease dividing once confluent, potentially allowing more time for Ataxin-7 turnover.

Firstly, transfected fibroblasts were immunostained with the 1-135 antibody 7 d post-transfection with siRNA or mirtrons. However, no difference in the staining intensity could be seen with siR-P16 treatment compared to a non-targeting control siRNA (siR-NS) (Figure 5-5a), despite the high efficiency of transfection (Figure 5-4a). For mirtrons, transfected cells can be discriminated by their EGFP fluorescence, yet no reduction in Ataxin-7 staining intensity was evident in these cells (Figure 5-5b). This was also carried out at 4 d post-transfection with the same observations. These results may be explained by the apparent Ataxin-7 immunostaining being an artefact due to either low levels of Ataxin-7 protein expression in these cells and/or cross-reactivity of the antibody. Alternatively, the presumed small reduction in protein levels may not visible by this method, or Ataxin-7 may be stabilised by post-transcriptional modifications in response to silencing.

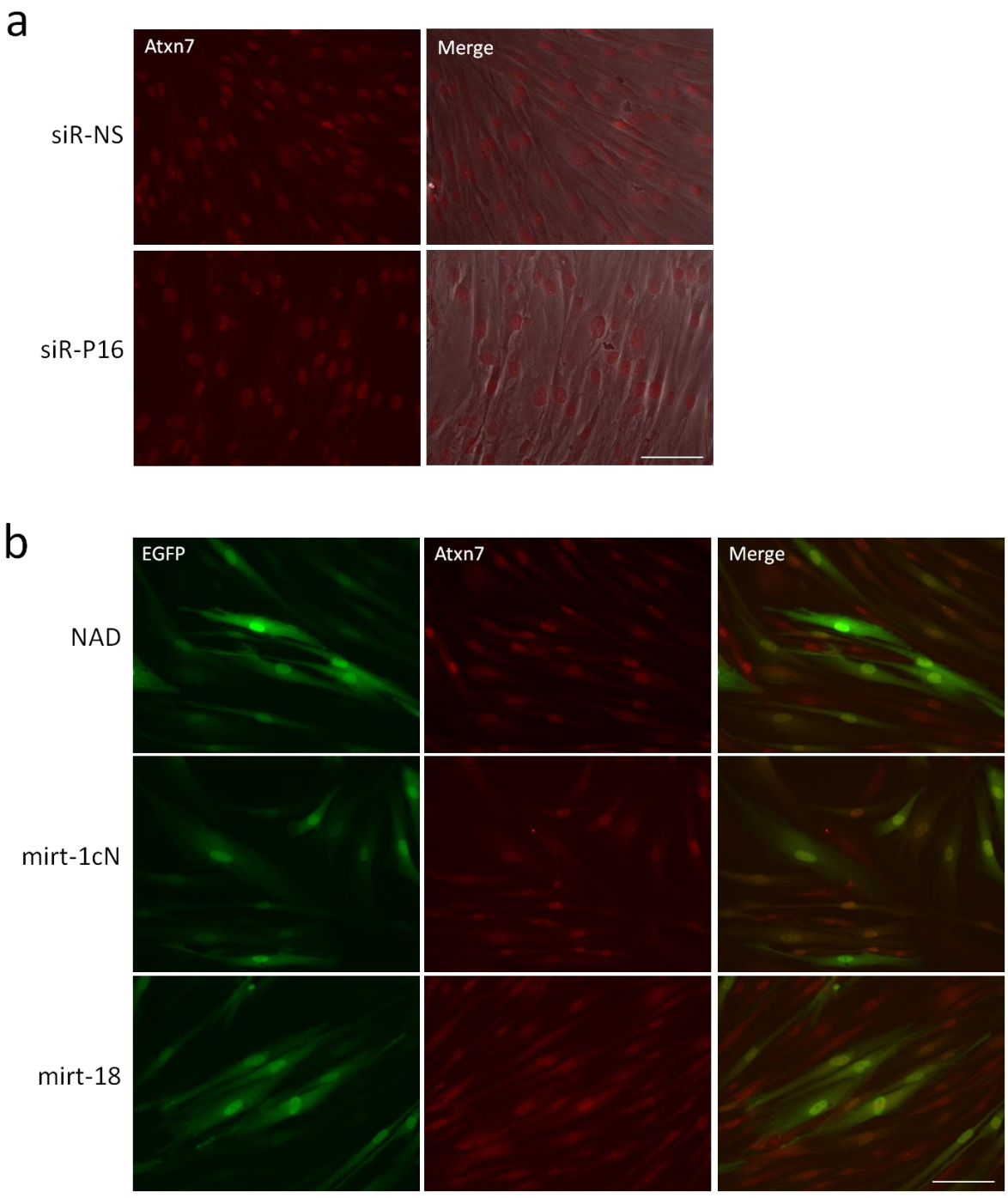


Figure 5-5 – Immunostaining of SCA7 patient fibroblasts after siRNA or mirtron transfection. Fibroblasts from the Coriell line were transfected with **(a)** siRNA at 20 nM or **(b)** mirtrons at 1 µg/ml. Cells were fixed after 7 d and stained with the Ataxin-7 1-135 antibody at 1:100, with an Alexa-594 fluorescent secondary antibody (1:500). In (a), images are also shown merged with phase contrast view and in (b) with the green channel, indicating EGFP-expressing (transfected) cells. Scale bar = 50 µm.

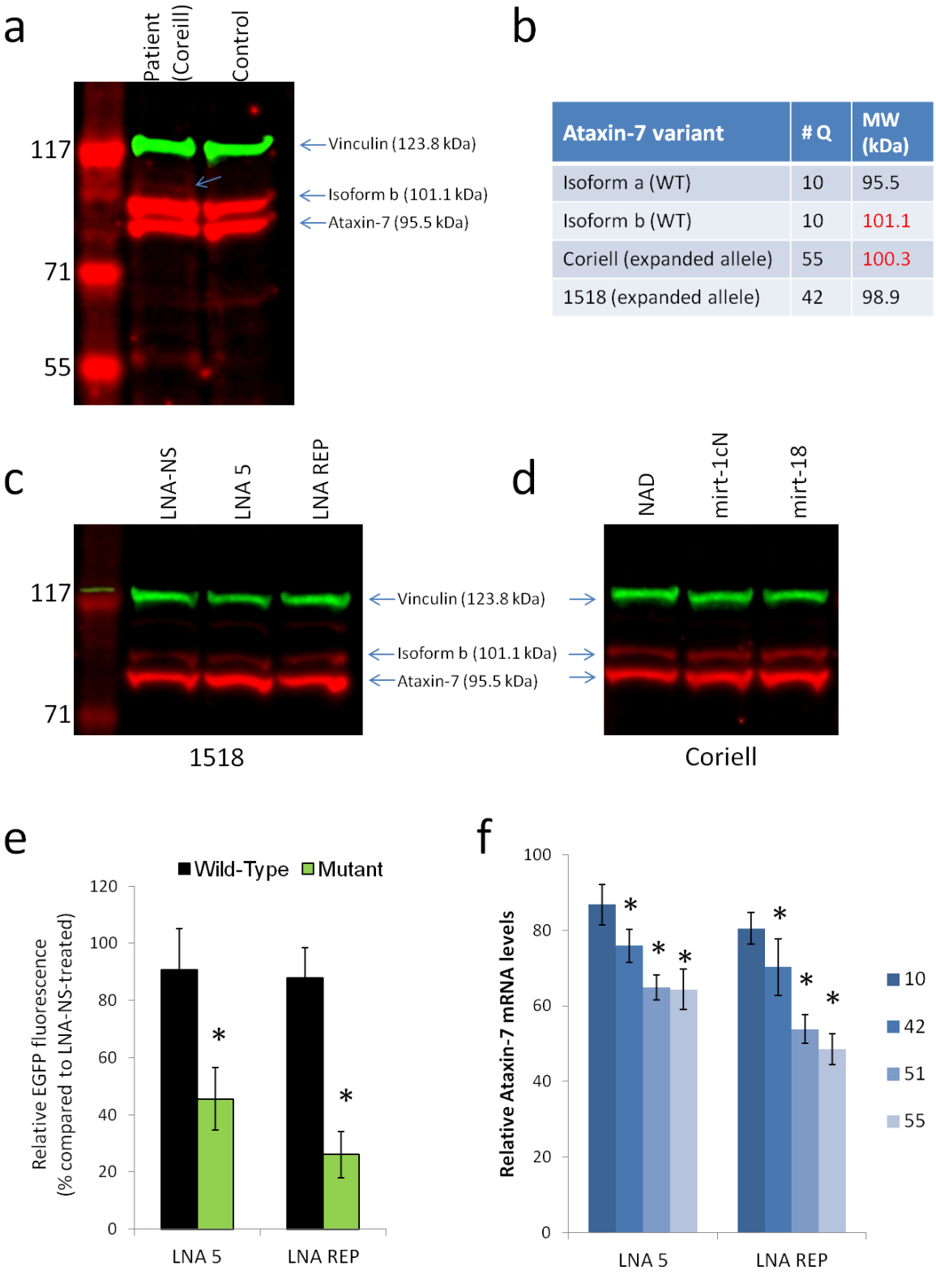
Western blotting was also used to investigate silencing on Ataxin-7 protein. This may provide some advantages over immunostaining, since proteins are concentrated, which may overcome problems involving low expression levels, and separated according to molecular weight, aiding their correct identification. To investigate the detection of endogenous Ataxin-7 in fibroblasts by Western blot, protein was first extracted from untreated Coriell SCA7 and normal fibroblast lines. Following electrophoresis, the 1-135 antibody was used for protein detection. For the Coriell line, a band was clearly visible around 95 kDa, as expected for normal Ataxin-7, and also a band around 100 kDa, corresponding to the size of the mutant protein (Figure 5-6a). However, this band was also observed for normal control fibroblasts, which do not possess an expanded Ataxin-7 allele. This band was also found in HD patient-derived fibroblasts (not shown). A band of this size was not distinct in (human) SH-SY5Y cells or (mouse) N2A cells (Figure 3-13 and Figure 3-15), indicating a cell-line specific effect. Although it may represent non-specific binding to an unrelated protein, an alternative explanation could be the expression of Ataxin-7 isoform b, which has a molecular weight of 101 kDa (for a non-expanded allele) (Figure 5-6b). Isoform b is a splice variant of Ataxin-7 and is enriched in the brain, but can also be detected in non-neuronal human tissues (Einum et al. 2003). Above this band, a faint band was seen only for the patient cells, which could correspond to the mutant variant of isoform b (Figure 5-6a, arrow). Exposing the same blot with the 1c2 antibody revealed no bands (data not shown), despite its strong (largely non-specific) immunostaining results seen earlier (Figure 5-3). This indicates that any non-specific proteins to which it binds had run off the gel or simply that this antibody was not sufficiently optimised in this procedure. Alternatively, the mutant protein may be trapped in the stacking gel. This is difficult to eliminate with the available antibodies, as they did not bind to highly aggregated Ataxin-7 protein (Figure 5-2b and d), though this could be investigated

using a standard protein stain. However, it is unlikely that 100% of the protein should be aggregated, since even when over-expressed, some remains soluble (Figure 5-2).

To further investigate, protein was then extracted from an alternative patient cell line (1518). In this line the mutant protein has a shorter expansion and hence a predicted molecular weight of 98.9 kDa (Figure 5-6b), which should be distinguishable from the presumed isoform b band. However, no additional band was visible at the expected size for the mutant protein (Figure 5-6c). In this experiment, the fibroblasts were transfected with two different antisense oligonucleotides which had been shown to achieve selective silencing of mutant over WT transfected Ataxin-7 (M. Varela, Figure 5-6e). These contained locked nucleic acid (LNA) residues, which can increase the affinity of oligonucleotides for their target mRNA (Veedu and Wengel 2009), and were targeted either to the CAG expansion (LNA REP) or the expansion plus its flanking sequence (LNA 5) (M. Varela, in preparation). Silencing of endogenous Ataxin-7 mRNA in patient fibroblasts was also demonstrated (M. Varela, Figure 5-6f). Efficacy correlated with CAG-repeat length, and in the Coriell line, 40% (LNA5) and 50% (LNA-REP) silencing of total Ataxin-7 mRNA was achieved after 3 d. For the Western blot, protein was harvested 7 d post-transfection. However, there was no visible reduction in intensity of any band compared to the control (Figure 5-6c). Similarly, fibroblasts from the Coriell line were transfected with mirtrons (which target both wild-type and mutant Ataxin-7), but again, after 7 days of mirtron expression, no reduction in Ataxin-7 protein was apparent (Figure 5-6d).

Together, these results indicate that the mutant Ataxin-7 protein may not be detectable in fibroblasts with the methods used, and that protein-level silencing cannot be observed. Ataxin-7 protein expression was lower in this cell line than N2A cells when compared directly (not shown), which may make subtle differences in protein levels more difficult to

distinguish. Alternatively, it could be that none of the immunostaining or Western bands represented either mutant or wild-type Ataxin-7. This is unlikely given the affinity shown to transfected Ataxin-7, and the correct size of the band in the Western blot. Mass spectrometry could be used to confirm the identity of the Western bands. It is possible that only N-terminal cleavage products of mutant Ataxin-7 are present, as commonly detected in patient material and mouse models (Einum et al. 2001; Yvert et al. 2000; Helmlinger et al. 2004a; Garden et al. 2002). However, the Ataxin-7 antibodies used here were raised against Ataxin-7 N-terminal fragments (amino acids 1-15 and 1-135) (Ström et al. 2002) and the smallest cleavage product contains amino acids 1-266 (Young et al. 2007). Therefore, these antibodies should be able to detect such fragments if present.



Ataxin-7 variant	# Q	MW (kDa)
Isoform a (WT)	10	95.5
Isoform b (WT)	10	101.1
Coriell (expanded allele)	55	100.3
1518 (expanded allele)	42	98.9

Figure 5-6 – Ataxin-7 protein detection in SCA7 patient fibroblasts by Western blotting. (a) Western blot for protein extracted from SCA7 patient fibroblasts (Coriell) and control/normal fibroblasts, probed with anti-Ataxin-7 antibody 1-135 (labelled Atxn7) and loading control anti-vinculin (vin). 25 µg total protein was loaded. **(b)** Table detailing the number of glutamine residues (#Q) and molecular weight of Ataxin-7 proteins in normal and

SCA7 patient (Coriell and 1518) alleles, including Ataxin-7 isoform b, a splice variant of Ataxin-7. Red text indicates the presumed co-incident bands in (a). **(c)** Western blot carried out as in (a), except protein was extracted from 1518 fibroblasts 7 d post-transfection with LNA-oligonucleotides. **(d)** Western blot carried out as in (a), except protein was extracted from Coriell SCA7 fibroblasts, 7 d post-transfection with pEGFP-NAD, mirt-18 or mirt-1cN. **(e)** Relative fluorescence measurement in protein extracted from HEK-293 cells was used to indicate the expression of Ataxin-7-Q10-DsRED (wild-type) or Ataxin-7-Q100-EGFP (mutant) when co-transfected with LNAs. Values represent the DsRED/EGFP fluorescence relative to LNA-NS (non-specific control) treated cells. Protein was extracted 3 d post-transfection. N=3 and error bars represent standard deviation (Varela et al, in preparation). **(f)** The expression of endogenous *ataxin-7* mRNA was measured 3 d after antisense oligonucleotide treatments in control, 1518, 1519 and Coriell SCA7 patient fibroblasts (repeat sizes 10, 42, 51 and 55, respectively). Ataxin-7 levels were measured with a TaqMan assay and normalised to internal control GAPDH (also measured by TaqMan assay). Values are given as percentages of relative to LNA-NS (non-specific control) for each cell line. N=3 and error bars represent standard error. For (e) and (f), asterisks (*) indicate $p<0.05$ compared to LNA-NS-treated cells.

5.2.3 Transcriptional rescue of genes dysregulated in SCA7

Although protein aggregates were not detected in the patient-derived fibroblasts, alternative measures of phenotype could be highly informative for assessing the efficacy of novel therapies. It is well established that many genes demonstrate dysregulated expression in polyQ disorders, including SCA7. The expression of many genes was altered in a mouse model of SCA7 (Chou et al. 2010). Importantly, some of these changes were found to be mirrored in SCA7 patient fibroblasts; specifically, the level of DNAJA1 (DnaJ homolog subfamily A member 1) mRNA was increased and UCHL1 (Ubiquitin carboxy-terminal hydrolase L1) decreased in patient cells compared to controls, each by approximately a factor of two (Scholefield et al, in preparation). Both of these genes have functional significance in SCA7 pathogenesis, being involved in protein folding and the Ubiquitin-proteasome system (UPS), respectively. Further, it was shown that their expression could be rescued to some extent when Ataxin-7 was silenced using siRNA (Scholefield et al, in preparation, and Figure 5-7a). This effect was observed at 7 d post transfection but not at 4 d, indicating that it is a downstream effect of the reduction of Ataxin-7.

In fibroblasts transfected with shR-1 (where Ataxin-7 mRNA levels were reduced by approximately 50%), DNAJA1 was repressed slightly but not with statistical significance, and UCHL1 levels were not significantly altered (Figure 5-7b). Similarly, in mirtron-transfected cells, the restorations in gene expression of DNAJA1 and UCHL1 were not seen (Figure 5-7c). Indeed, mirt-1cN led to an increase in DNAJA1 levels, contrary to the earlier observation with siRNA.

It was previously noted that at lower siRNA concentrations (where Ataxin-7 mRNA is reduced by 50% or less), DNAJA1 mRNA levels were not significantly altered, whereas UCHL1 mRNA still demonstrated an increase (J Scholefield et al., in preparation). This may explain

the lack of restoration of DNAJA1 in the case of shRNA and mirtrons, neither of which reduce Ataxin-7 mRNA by more than 50% 7 d post-transfection. For mirtrons, the silencing appeared to be transient, which may not be sufficient to rescue the downstream effects of mutant Ataxin-7.

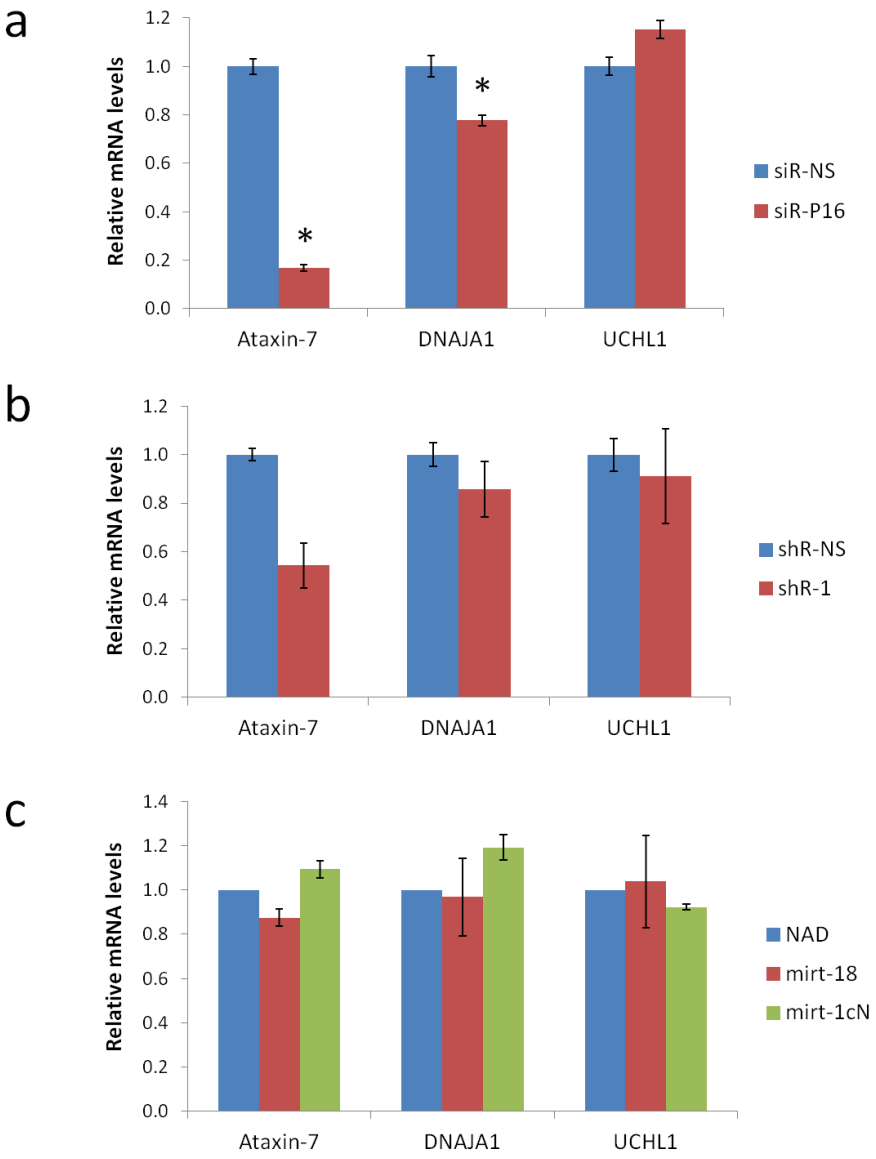


Figure 5-7 – Effects of Ataxin-7 silencing on gene expression. Expression levels of *ataxin-7*, DNAJA1 and UCHL1 were measured by qPCR in Coriell patient fibroblasts, 7 d post-transfection with **(a)** siRNA at 20 nM, **(b)** shRNA at 1 µg/ml or **(c)** mirtrons at 1 µg/ml. Expression levels were normalised to internal control gene β-actin, and values obtained for siR-NS, shR-NS and NAD were normalised to 1. N=2 except for NAD, where one biological replicate produced high Ct values for all genes tested and was discarded. Error bars represent standard error. For (a)-(c), asterisks (*) represent $p<0.05$ (1-tailed t-test).

5.3 Lentiviruses as a vector for mirtrons

A novel SCA7 patient-derived neuronal cell model was produced and characterised by J. Scholefield and L. Watson and was available for the testing of novel therapies (see section 0). Post-mitotic neurons are difficult to transfect through typical lipid-based or electroporation methods, but can be transduced using viral vectors capable of infecting non-dividing cells, such as lentiviruses (Lundberg et al. 2008; Linterman et al. 2011) or magnetic nanobeads (Fallini et al., 2010). Certain recombinant Adeno-Associated viruses (AAV) can also target neuronal cells. However, AAV can package expression cassettes no larger than 5 kb, and the viral titres obtained become lower as this limit is approached (Wu et al. 2010). Lentiviruses have a larger packaging capacity of 8-10 kb (Sinn et al. 2005), which would allow greater flexibility for the later testing of K&R constructs for Ataxin-7 (including promoter, the 3 kb Ataxin-7 cDNA and a fluorescent tag along with a mirtron or shRNA). Further, lentiviruses expressing shRNAs targeting Ataxin-7 were previously constructed (J. Scholefield, unpublished data), therefore lentiviruses expressing mirtrons were also designed.

5.3.1 Production of lentiviruses to silence Ataxin-7

Lentiviruses were made using the HIV7 backbone as described in sections 2.2.8 and 2.3.5. The shRNA-expressing virus was constructed by inserting a U6-promoter-driven cassette in the pLKO.1 cloning vector, containing either a non-specific shRNA (shR-NS) or shR-P16 (cloning carried out by J. Scholefield, Figure 5-8a). The vector additionally contained a CMV-driven EGFP gene to allow visualisation of viral infection. In order to produce mirtron-expressing viruses, the pLKO.1 EGFP gene was replaced with one version of the EGFP-mirtron gene, containing either the double-mirtron construct 1cN-18 or two NAD introns (NAD-NAD) (Chapter 3; Figure 5-8b). Including mirt-18 allowed the virus to potentially be

used against mouse Ataxin-7 but this construct was spliced more efficiently than that containing mirt-18 alone (Figure 3-10).

Lentiviruses were packaged in HEK-293 T cells by co-transfection of each modified pLKO.1 plasmid with the complementary packaging vectors. Packaging of the virus leads to the appearance of lysed cells (Figure 5-8c). Virus particles were harvested by collection of media after 48 and 72 h, filtered to remove cells and cell debris, then each concentrated to 100 µl.

During the packaging process, EGFP expression appeared brighter for the mirtron viruses compared to the shRNA viruses (Figure 5-8c), despite identical conditions during transfection. Splicing of introns is known to increase transgene expression in some cases (Brinster et al. 1988), therefore the spliced EGFP from the mirtron viruses may conceivably produce a higher level of expression than the intron-less EGFP from the shRNA viruses. However, mirtron viruses produced a larger viral pellet following ultracentrifugation compared to shRNA viruses, indicating that more lentiviral particles were produced. This indicated that packaging was more efficient for the mirtron viruses. This may be explained by the size of the cDNA being inserted and the packaging limits of lentiviruses. Although the cDNA being inserted was not close to this limit, the genome was indeed smaller for the mirtron constructs: the U6-shRNA cassette added approximately 360 bp, whereas the two mirtrons together added 2 x 72 bp, leading to an overall difference of just over 200 bp. This could perhaps have led to differences in packaging efficiency.

10 ml aliquots of filtered media containing the NAD-NAD and mirtron virus were not ultracentrifuged and these were used in some experiments (“un-concentrated”). The major limitation of this is that a larger volume has to be added to cells which may compromise the maximum possible infection which can be achieved, and the media may not be compatible

with all cell types such as neuronal cultures. However, these were favoured for use in SH-SY5Y and HEK-293 cells and appeared to demonstrate better stability during storage.

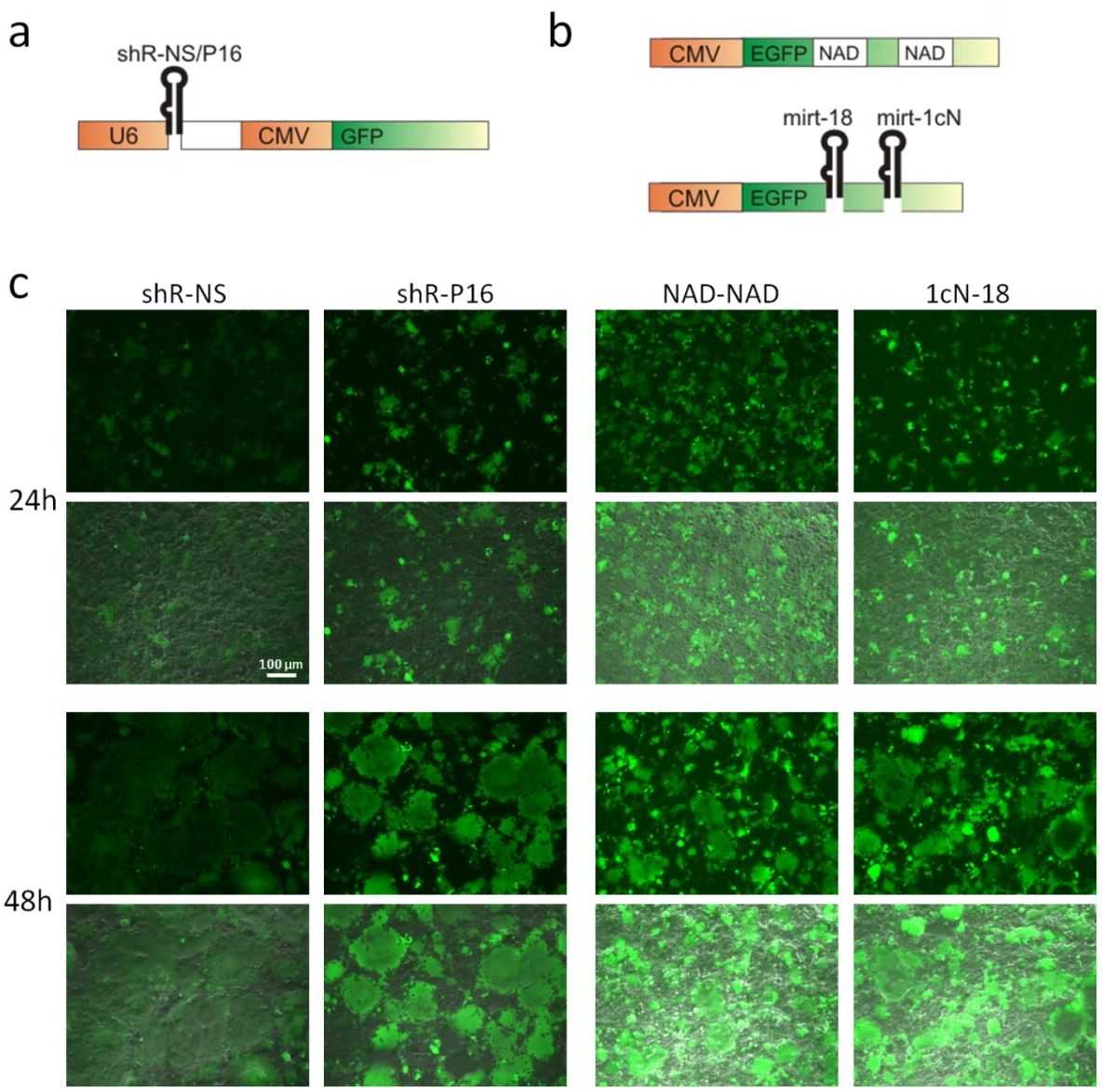


Figure 5-8 – shRNA and mirtron lentivirus design and production. **(a)** Schematic diagram showing the design of shRNA-expressing lentiviruses. Promoters are shown in orange boxes. **(b)** Schematic diagram showing the design of the NAD-containing (top) and mirtron-containing (bottom) lentiviruses. **(c)** Representative fluorescent images of HEK-293 T cells during packaging of each virus, showing increasing areas of cell lysis between 24 and 48 h post-transfection. EGFP images (top row) are also shown merged with phase contrast images.

5.3.2 Estimation of lentiviral titre

HEK-293 cells were infected with different dilutions of the shRNA and mirtron lentiviruses in order to test their infectivity and estimate their titre. For the shRNA viruses, the rate of infection was comparable for shR-NS and shR-P16 (Figure 5-9a). A more significant difference was observed for the mirtron viruses (Figure 5-9b). A rough estimation of titre was obtained through counting the proportion of fluorescent cells and inferring the total number of active viral particles (transducing units, TU) per ml of virus added to the well (Table 5-1). The shRNA lentiviruses each had a titre in the region of 1-2 x 10^6 TU/ml, and the un-concentrated mirtron lentiviruses had a titre in the region of 2-6 x 10^4 TU/ml (equivalent to an approximate 1:100 dilution of the concentrated viruses).

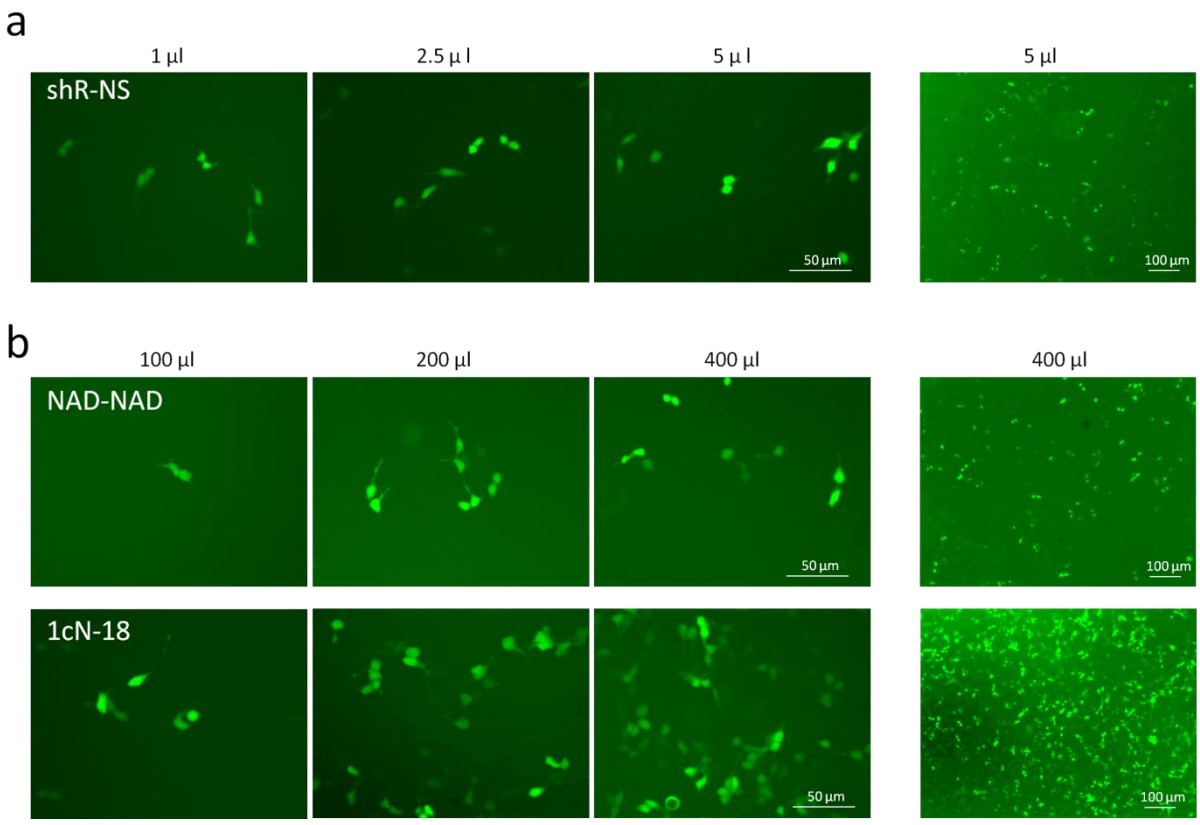


Figure 5-9 – Lentiviral infections in HEK-293 cells. EGFP fluorescent images of HEK-293 cells infected with varying volumes of virus as indicated, 48 h post-infection (at confluency). To indicate that these images are representative of the infection rate across each well, cells are also shown at a lower magnification for the highest volume in each case. **(a)** Cells were infected with shRNA viruses which both achieved similar rates of infection (shR-NS shown here). **(b)** Cells were infected with un-concentrated NAD-NAD and 1cN-18 mirtron viruses which demonstrated noticeably different titres.

Table 5-1 – Lentiviral titre estimation. Calculations were based upon the higher magnification images in Figure 5-9 (and comparable images for HIV-shR-P16). The number of cells expressing EGFP was counted (EGFP cells) and divided by the approximate total number of cells in the image (500) (EFGP/total). This value was multiplied by the total approximate number of cells in each well (1×10^5) to obtain the total number of transducing units (TU). This was divided by the viral volume added to each well (Viral vol (ml)) to obtain the titre in TU/ml, which was then averaged for each virus. Calculations were excluded (excl) where proportion of EGFP-expressing cells was less than 0.01 to minimise inaccuracy.

	NAD-NAD			**1cN-18**			**shR-NS**			**shR-P16**		
Viral vol (ml)	0.4	0.2	0.1	0.4	0.2	0.1	0.005	0.0025	0.001	0.005	0.0025	0.001
EGFP cells	17	12	2	75	35	10	14	14	8	16	8	2
EGFP/total	0.034	0.024	0.004	0.15	0.07	0.02	0.028	0.028	0.016	0.032	0.016	0.004
TU	6800	4800	excl	30000	14000	4000	5600	5600	3200	6400	3200	excl
TU/ml	1.7×10^4	2.4×10^4		7.5×10^4	7.0×10^4	4.0×10^4	1.1×10^6	2.2×10^6	3.2×10^6	1.3×10^6	1.3×10^6	
Mean TU/ml	2.1×10^4			6.2×10^4			2.2×10^6			1.3×10^6		

5.3.3 Expression of shRNA and mirtrons from lentiviruses

In order to confirm that the lentiviruses expressed the shRNA and mirtrons, the production of the pre-miRNA hairpin species was investigated. Using RNA extracted from infected HEK-293 cells, quantitative RT-PCR was carried out using custom-designed assays specific for shR-P16, mirt-18 or mirt-1cN. Unfortunately, the shR-P16 hairpin could not be detected, even for the positive control (RNA extracted from cells efficiently transfected with shR-P16 plasmid; data not shown). Both mirt-18 and mirt-1cN, on the other hand, were successfully detected in cells transfected with the 1cN-18 plasmid and cells infected with the 1cN-18 virus (but not in cells infected with the NAD-NAD virus, Figure 5-10a). Compared to cells transfected with the viral EGFP-1cN-18 plasmid in the same experiment, viral mirt-18 levels were approximately 230-fold lower and mirt-1cN 110-fold lower (Figure 5-10b). The difference between viral and plasmid expression levels is likely to be explained by the low infection rate (approximately 1-5%, presumably single-copy) compared to plasmid transfection (90%, multiple copies). In addition, some lentiviruses may have been packaged post-splicing. To avoid this, an improved design could incorporate the EGFP-mirt cassette on the opposite strand (Jonsson et al. 1995). The difference between mirt-18 and mirt-1cN is more unexpected, since each transcript should produce similar levels of both mirtrons. It was proposed in Chapter 3 (section 3.3.2) that the mirt-1cN hairpin is processed into mature miRNA more rapidly than mirt-18. It is possible that at the low levels expressed through viral infection, the mirt-18 hairpin can be more efficiently processed, causing less saturation of the RNAi pathway. Alternatively, the discrepancy may be explained through unequal sensitivities of the assays or inaccuracy at high Ct values.

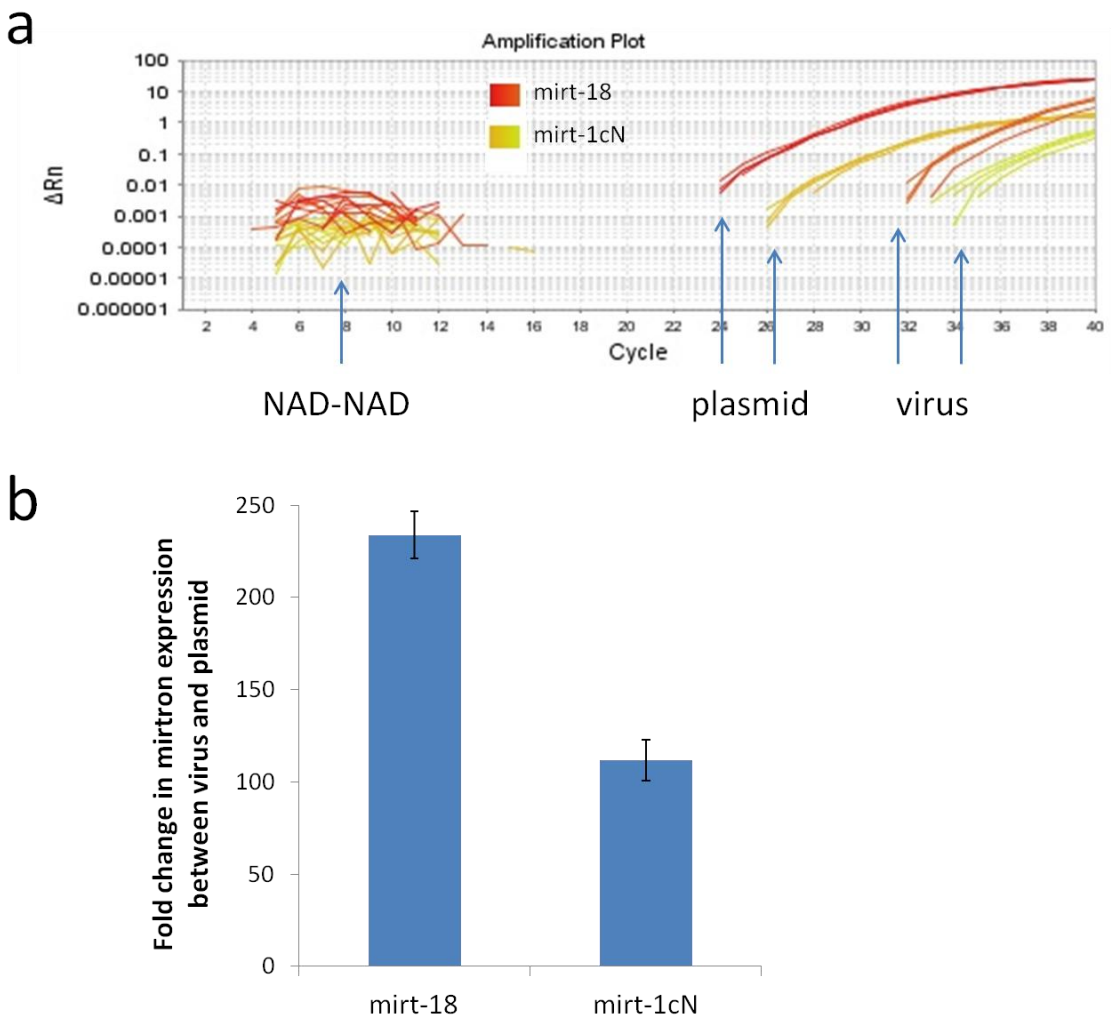


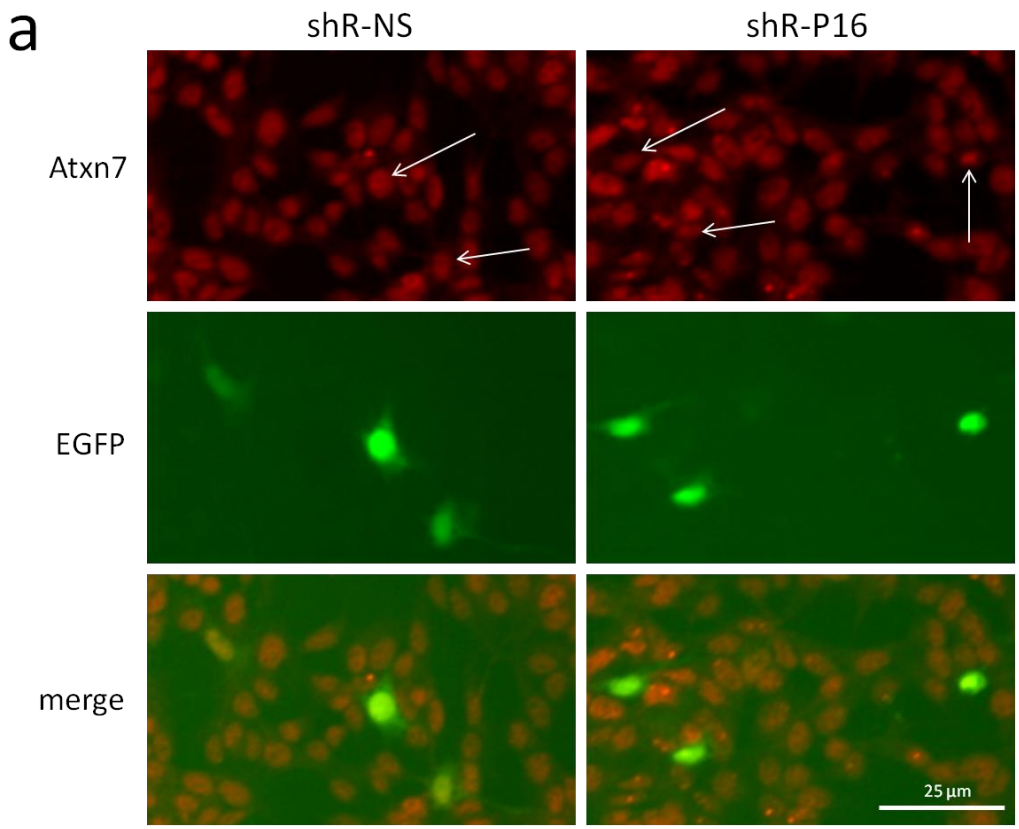
Figure 5-10 – Mirtron expression in HEK-293 cells infected with lentivirus compared to plasmid-transfected cells. RNA was extracted from HEK-293 cells 3 d post-infection with 5 µl mirrtron lentivirus in 48-well plates or 3 d post-transfection with the equivalent pLKO.1 plasmid. qPCR was used to detect mirtron pre-miRNA hairpin expression using custom-designed TaqMan assays for small-RNA. **(a)** qPCR amplification plot for mirtron hairpins. The key indicates the mirtron being assayed. Labels indicate traces representing the NAD-NAD virus, 1cN-18 plasmid and 1cN-18 virus. The traces represent all biological (2) and technical (2) replicates. **(b)** Mirtron levels were normalised to GAPDH internal control using the delta-delta Ct method. The fold-change in expression of each mirtron between virus-infected and plasmid-transfected cells is presented. GAPDH expression was also measured with a TaqMan assay. N=2. Error bars represent standard error.

5.3.4 Silencing in a standard neuronal cell line

5.3.4.1 Immunostaining to detect lentiviral-directed silencing of Ataxin-7

Following the confirmation of mirtron pre-miRNA expression from lentiviruses, it was then possible to seek proof that the lentiviruses can silence Ataxin-7. A human neuronal cell line (SH-SY5Y) was initially chosen for this purpose. However, the low infection rates obtained, particularly for the shRNA viruses, could limit the detection of silencing using RNA or protein lysates. Immunostaining was therefore initially used to asses Ataxin-7 protein silencing, because it allows individual infected cells to be identified by their EGFP expression.

SH-SY5Y cells were first infected with the shRNA viruses and immunostained with the 1-135 anti-Ataxin-7 antibody 48 h after infection (Figure 5-11a). Some cells appeared more weakly stained than uninfected cells; however, the frequency of this observation was not greater for shR-P16 compared to shR-NS. SH-SY5Y cells were then infected with the mirtron viruses and immunostained after a longer period of infection, 96 h (Figure 5-11b). However, Ataxin-7 protein silencing was again not strongly evident. Protein-level silencing was also difficult to detect with plasmid-delivered mirtrons, as shown in Chapter 3, likely to be limited by the low protein levels and insufficient sensitivity of the assays used. Therefore, seeking evidence of silencing at the mRNA level was attempted.


a
shR-NS
shR-P16
Atxn7
EGFP
merge
25 µm

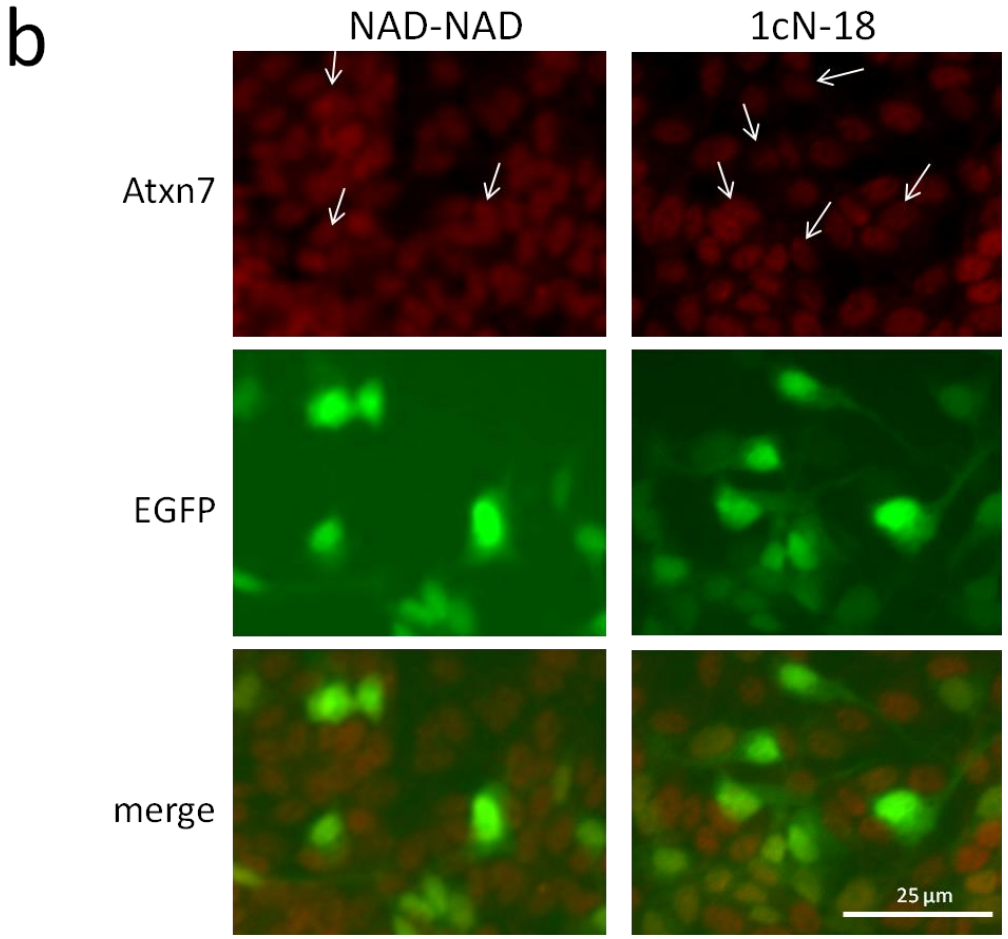

b
NAD-NAD
1cN-18
Atxn7
EGFP
merge
25 µm

Figure 5-11 – Immunostaining of Ataxin-7 in SH-SY5Y cells infected with lentiviruses. Representative fluorescent images of cells immunostained for Ataxin-7 protein with primary antibody 1-135 (1:100) and an Alexa-594-conjugated secondary antibody (1:500, red channel, "Atxn7"). EGFP fluorescence indicates infected cells. EGFP and merged images were enhanced for clarity. **(a)** SH-SY5Y cells were plated in 24-well plates and infected with 3 µl of each shRNA virus. Cells were fixed and stained 48 h later. **(b)** SH-SY5Y cells were plated in 24-well plates and infected with 1 ml un-concentrated mirtron virus. Cells were fixed and stained 96 h later.

5.3.4.2 Improvement of lentiviral infection in SH-SY5Y cells

Since the proportion of infected cells would be a limiting factor for the detection of Ataxin-7 silencing in cell lysates, and the total viral volume available was limited, the addition of polybrene (hexadimethrine bromide) was tested to improve infection (Toyoshima and Vogt 1969). With polybrene at concentrations up to 8 µg/ml, the proportion of infected cells for both shRNA and mirtron viruses increased by approximately a factor of 2 compared to no polybrene (Figure 5-12a and b). At 12 µg/ml and above it became toxic, particularly after 72 h (Figure 5-12b), consistent with published findings (Seitz et al. 1998). 8 µg/ml polybrene was therefore used in further experiments.

5.3.4.3 Silencing of Ataxin-7 mRNA with lentiviruses

SH-SH5Y cells were harvested 4 d after infection with shRNA lentiviruses, when approximately 30-40% of cells were infected (Figure 5-13a). qPCR indicated that the shR-P16 lentivirus achieved a small (10%) but significant ($p<0.05$) reduction in Ataxin-7 mRNA, compared to the non-specific control. (Figure 5-13b). A higher infection rate (50-60%) was observed for mirtron viruses (Figure 5-13a), yet no silencing was observed (Figure 5-13b). An equivalent siRNA (siR-P16) was used as a positive control, and achieved approximately 90% transfection efficiency and 30% silencing at the same time-point (Figure 5-13b). It is encouraging that silencing was observed for the shRNA lentivirus, though the experiment

should be repeated to be certain of this effect. The silencing may be limited by the relatively low infection rate combined with a lower copy number per cell compared to plasmid or siRNA transfection. Further, at low levels shR-P16 demonstrates allele-specificity towards a SCA7-linked SNP (Scholefield et al. 2009). Poor cell health may also have been an issue, particularly for the 6 d time-point used for mirtron viruses.

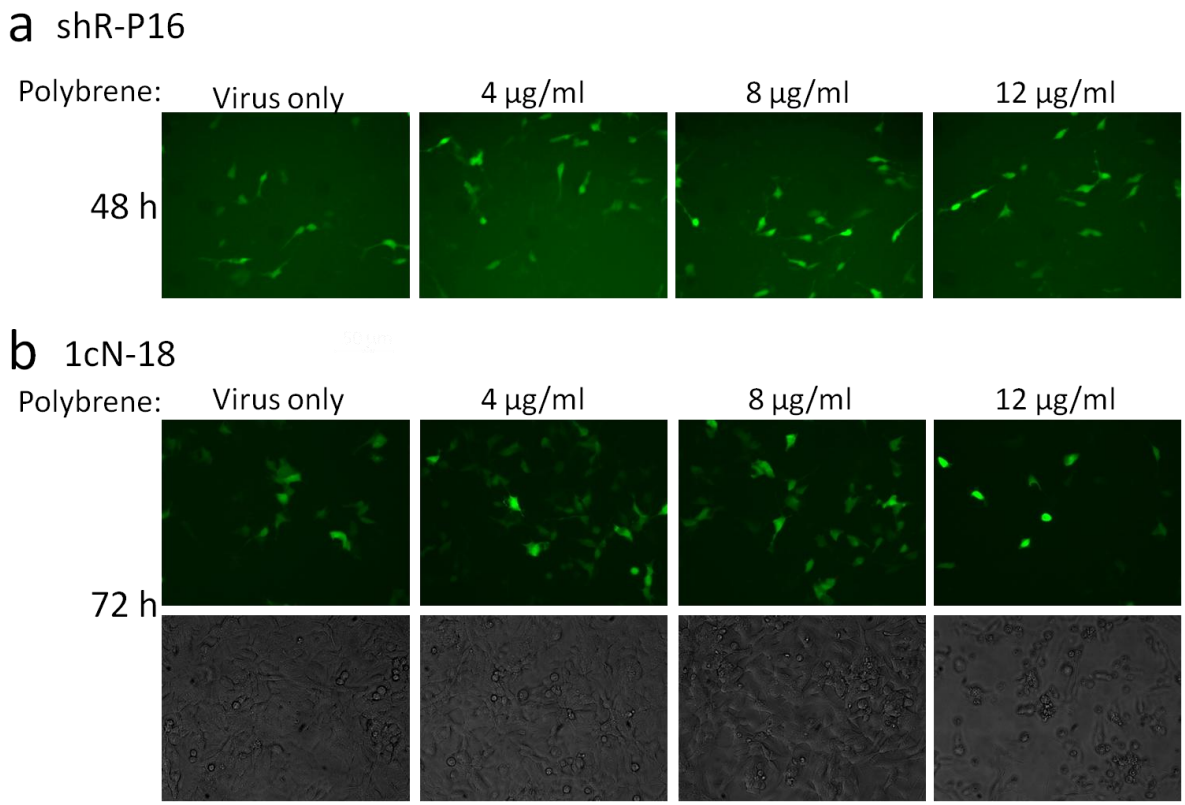


Figure 5-12 – Lentiviral infection was improved with polybrene. SH-SY5Y cells were plated in 24-well plates and polybrene was added concurrently with the lentivirus. Media was changed 24 h later. **(a)** Representative fluorescent images of cells infected with 3 µl shR-P16 virus with polybrene at the stated concentrations, 48 h post-infection. **(b)** Representative images of cells infected with 1 µl 1cN-18 virus with polybrene at the stated concentrations, 72 h post-infection. Corresponding brightfield images are also shown to visualise cell death.

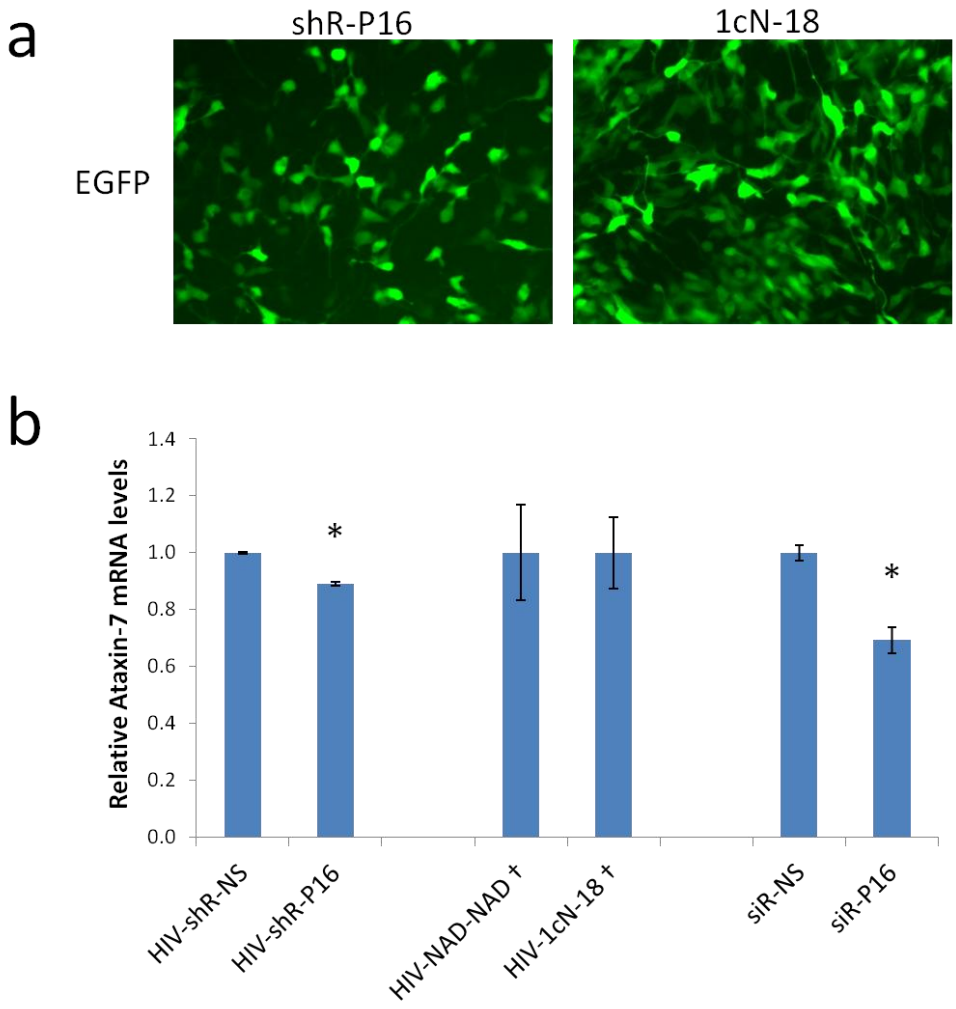


Figure 5-13 – Silencing of endogenous Ataxin-7 mRNA with lentiviruses in SH-SY5Y cells. SH-SY5Y cells were plated in 48-well plates and infected with lentiviruses: 10 µl shRNA lentivirus with polybrene at 8 µg/ml or 0.5 ml un-concentrated mirtron lentivirus. **(a)** Representative fluorescent images of EGFP expression indicating infected cells, 4 d post-infection with HIV-shR-P16 and 6 d post-infection with HIV-1cN-18 (†). **(b)** Ataxin-7 mRNA levels were measured by qPCR and normalised to β-actin, in RNA extracted from SH-SY5Y cells in (a). The value for each non-specific control (HIV-shR-NS, NAD-NAD) was normalised to 1. qPCR was also carried out on SH-SY5Y cells 96 h post-transfection with an siRNA equivalent to shR-P16 (siR-P16) at 20 nM, and the value for siR-NS was normalised to 1. N=2, error bars represent standard error and asterisks (*) represent $p<0.05$ compared to the appropriate non-specific control.

5.4 Patient-derived neurons

Patient fibroblasts can be reprogrammed into induced pluripotent stem cells (iPSCs) through the delivery of key reprogramming genes (Takahashi et al. 2007). iPSCs may then be differentiated into neuronal cultures through the provision of appropriate small molecule stimuli and culture conditions (Dimos et al. 2008; Ebert et al. 2009; Dottori and Pera 2008). This was recently attained for SCA7 patient cells (Luo et al. 2012; Watson 2012), (Scholefield et al., unpublished data). This creates novel human cell models for SCA7, naturally containing an Ataxin-7 expansion but of a more relevant cell type than fibroblasts. Such models for other polyQ disorders displayed more disease-relevant phenotypes than fibroblasts (Koch et al. 2011), and hence may provide a useful tool for analysis of SCA7 pathogenesis and novel treatments. Patient-derived neuronal cultures were produced from two SCA7 patient fibroblast cell lines and probed for phenotypic features, as discussed below. The lentiviruses produced in the previous section were then used to transduce these cultures in an attempt to silence Ataxin-7.

5.4.1 Neuron production and characterisation.

5.4.1.1 Production

The work in this section was carried out by J. Scholefield and L. Watson. iPSCs were generated from patient fibroblast cell lines and their pluripotent character was confirmed, described in section 2.3.2. An adapted protocol for neuronal differentiation (Dottori and Pera 2008) was then used to derive neuronal cells from two SCA7 patients as well as controls. This protocol produces mixed neuronal and glial populations. One method for differentiation of Purkinje cells (the most severely affected neurons in SCA7) from stem cells had been published (Tao et al. 2010), but the chosen method was selected for its simplicity.

This was deemed a reasonable compromise given that SCA7 affects a number of neuronal and glial cells, as discussed in Chapter 1.

5.4.1.2 Characterisation

The work in this section was carried out by L. Watson and J. Scholefield (Watson 2012). Firstly, it was verified that no significant contractions or expansions of the CAG tract occurred during the iPSC and neuronal differentiation procedures. iPSC-derived neuronal cells were confirmed to have neuronal character by their positive immunostaining for the marker βIII-tubulin. There was no detectable difference in the differentiation efficiency of SCA7 iPSCs compared to those derived from control fibroblasts, suggesting that neuronal development is not hindered by CAG expansions of this size in Ataxin-7, as also seen in HD (Koch et al. 2011; Camnasio et al. 2012).

An approximate two-fold increase in *ataxin-7* expression was detected by qPCR in SCA7 iPSC-derived neurons compared to controls, after one week of differentiation. This is perhaps attributable to epigenetic regulation via CTCF, involving the SCAANT1 antisense transcript (Sopher et al. 2011). CAG expansion in *ataxin-7* was previously correlated to reduced SCAANT1 and de-repression of an alternative *ataxin-7* promoter, significantly increasing the levels of *ataxin-7*. However, immunostaining indicated no differences in Ataxin-7 expression at the protein level, nor differences in localisation, between SCA7 and control iPSC lines. No NIs were detected. Hence, the iPSC-derived neurons appeared to represent an early, pre-NI stage of disease, which was independent of the disease stage of the patient from which they were derived.

Dysregulation of certain transcripts is often the earliest detectable change in the pathogenesis of polyQ diseases. The expression of 15 disease-relevant genes, previously found to be altered in patients or animal models, was investigated in the iPSC-derived

neurons. 12 of these showed significant changes in expression levels (or noticeable trends through differentiation) compared to wild-type iPSC-derived neurons. Notably, this included the two genes which were altered in fibroblasts – DNAJA1 and UCHL1, showing similar but less significant changes in the iPSC-derived neurons.

5.4.2 Silencing of Ataxin-7 in neurons

Neuronal cultures derived from 1519 iPSCs were treated with the concentrated lentiviruses produced in section 5.3. This patient was heterozygous for the SNP towards which shR-P16 was directed, which should allow the shRNA expressed from the virus to silence the mutant allele. Viral volumes were selected based upon prior infections in order to produce comparable levels of expression. Polybrene was also added at 8 µg/ml to enhance infection. 48 h post-infection, no fluorescence was visible in any of the virus-infected cultures. 4 d post-transfection, fluorescence was visible in some cells, but none of the neuronal cells emanating from the neurosphere. Infected cells generally had a fibroblast-like morphology, although some were more filamentous (Figure 5-14). It is likely that they were an artefact of the differentiation procedure; cells which survive the process but do not form part of the neurosphere and do not become fully differentiated. However, they may have some neuronal character, as evidenced by their weak staining with the neuronal marker βIII-tubulin (Watson 2012). During the subsequent days, the number of EGFP-expressing cells increased as these cell populations divided, but no EGFP expression was ever observed in the neurosphere extensions.

Lentivirus-mediated silencing activity was assessed in these cells by immunostaining with anti-Ataxin-7 antibody 1-15, 7 d post-infection (Figure 5-14). However, there was no strong correlation between EGFP expression and Ataxin-7 staining intensity, indicating that protein-level silencing could not be observed. The low infection rate would limit the analysis of

silencing in cell lysates. Furthermore, the heterogeneous nature of the cell populations (containing cells at many stages of differentiation) may lead to varying levels of Ataxin-7 within the population, complicating any comparisons made between each well.

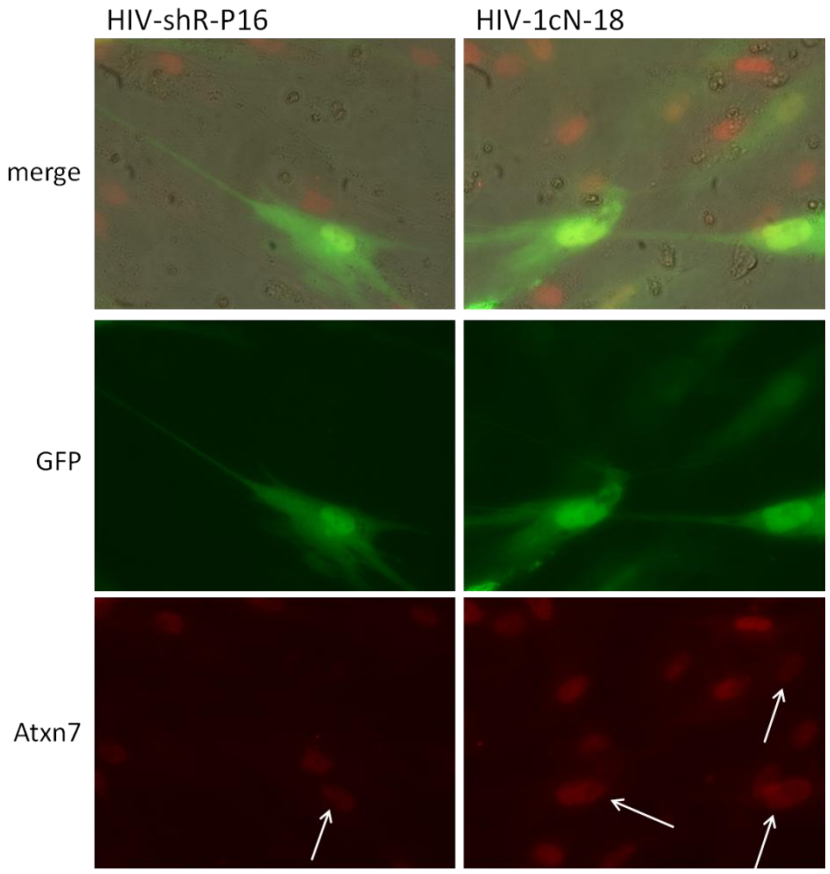


Figure 5-14 – Ataxin-7 immunostaining in patient-derived cultures infected with lentiviruses. Plates containing neuronal cultures derived from the 1519 fibroblast line were infected with lentiviruses (20 µl shR-NS/P16, 10 µl NAD-NAD or 5 µl 1cN-18; all with polybrene at 8 µg/ml). They were fixed 7 d later and stained with the anti-Ataxin-7 1-15 antibody (1:100) with an Alexa-594-conjugated secondary antibody (1:500) (red channel, Atxn7). Only fibroblast-like cells were infected. Representative fluorescent images of shR-P16 and 1cN-18 infected cells are shown, indicating that there were no differences in Ataxin-7 immunostaining between infected and un-infected cells. Images are also shown merged with the phase contrast view. Arrows indicate cells expressing viral EGFP.

5.5 Discussion

5.5.1 Patient fibroblasts as a model

Dermal fibroblast cultures are derived from simple skin biopsies from patients. They provide a relatively accessible disease cell model which is straightforward to culture and can be transfected efficiently with siRNA and moderately well with plasmids. Furthermore, certain disease-related phenotypes can be detected in fibroblasts derived from patients of other polyQ disorders, such as proteasome impairment and nuclear RNA aggregation in HD (Seo et al. 2004; de Mezer et al. 2011). In the case of SCA7, significant changes to the expression of disease-relevant genes were detected (Scholefield et al, in preparation). These changes could be corrected to some extent with siRNA targeting Ataxin-7. Fibroblasts therefore provide a valuable SCA7 cell model in which to test the efficacy of novel therapies and to indicate at multiple levels (mRNA, protein and downstream effects) whether cells naturally carrying the mutation can be corrected.

Protein aggregates, a key feature of affected cells in SCA7 and other polyQ disorders, could not be detected by immunostaining in fibroblasts, and no aggregated material was apparent by Western blotting. It may be that Ataxin-7 protein expression in this cell type is so low that aggregates do not form (or are readily processed by the cell). Similarly, many transcriptional-level alterations detected in mouse models are presumably specific to affected cell types and were not detected in patient-derived fibroblasts (Scholefield et al, in preparation). The derivation of fibroblast lines from patients with larger expansions may be more likely to present with protein aggregation (Ansorge et al. 2004).

Silencing of Ataxin-7 was detected at the mRNA level in patient fibroblast cells. Impressively, mirtrons achieved more than 50% silencing of Ataxin-7 mRNA 4 d post-transfection. Given

the plasmid transfection efficiency of approximately 50%, this indicates a potent silencing effect. However, results were variable and this effect was not detected at 7 d post-transfection. It is possible that prolonged cell stress following the electroporation transfection procedure caused alterations to gene expression. No protein-level silencing could be detected, even with high-efficiency siRNA and LNA transfection, reflecting the observations in Chapter 3. It appears likely that the methods employed were not sensitive enough to visualise any change in endogenous Ataxin-7 protein levels on the timescale used. It is also possible that the protein detected is not Ataxin-7 or a compensatory mechanism may post-transcriptionally stabilise Ataxin-7, which remain to be explored.

5.5.2 Lentiviruses

Lentiviruses are retroviruses which are capable of infecting non-dividing cells such as neuronal cells and are showing promise in clinical trials for gene therapy. Lentiviruses designed to target Ataxin-7 with shRNA or mirtrons were produced.

Obtaining high viral titres proved difficult, particularly for shRNA-containing viruses, despite a repetition of the packaging procedure for each virus. This limited the number of experiments which could be carried out, and inhibited the potential for *in vivo* application. This could be circumvented through commercial production, but at a substantial cost. Further, it is well known that freeze-thawing can severely impact the infective capacity of viruses, which may also have impacted these lentiviruses, given that the aliquot size must be chosen before titre has been established.

Nevertheless, it was shown that the lentiviruses could infect HEK-293 cells and human neuronal SH-SY5Y cells and that mirtron hairpins were expressed in HEK-293 cells. Surprisingly, silencing of endogenous Ataxin-7 in SH-SY5Y cells at the mRNA level was

detectable only for the shRNA-expressing virus, and was minimal. This may explain why no protein-level silencing was detected in infected cells. At low levels, shR-P16 has an allele-specific silencing effect towards a disease-linked SNP (Scholefield et al. 2009), which may explain its minimal efficacy in this case, as was also seen with plasmid transfection (data not shown). However, this does not explain the lack of silencing activity of mirtrons, which may alternatively be accountable to insufficient infection or expression and/or poor cell health and requires repeating. The expression of the virally-delivered mirtrons should also be confirmed in this cell line.

5.5.3 Patient-derived neurons as a model

Neuronal cultures were produced through reprogramming of patient-derived dermal fibroblast cultures via iPS cells. This provided a human cell model for SCA7, naturally containing an Ataxin-7 expansion but of a more relevant cell type than those directly available from patients, hence a useful tool for analysis of SCA7 pathogenesis and novel treatments.

The neuronal cultures were highly problematic to transduce. It proved difficult to infect the distinctly neuronal cells of the population with the shRNA- and mirtron-expressing lentiviruses. More success may be achieved by transducing cells at the iPSC stage, which can then be selected to differentiate neuronal cultures with the transgene integrated into the genome (Kamata et al. 2010; Chatterjee et al. 2011). Alternatively, other lentivirus pseudotypes or AAV may be developed to express mirtrons.

iPSC-derived neurons are more difficult to maintain in culture compared to fibroblasts. The differentiation protocol produced heterogeneous populations containing cells at many stages of differentiation, making the analysis of gene expression in cell lysates difficult. It

would be interesting to investigate the effect of the mutation in more homogeneous Purkinje cell cultures, a population which is severely affected in patients.

Expression of *ataxin-7* was approximately two-fold greater in SCA7 patient iPSC-derived neurons than controls. However, immunostaining indicated no differences in Ataxin-7 protein expression and localisation between SCA7 and control iPSC lines. No NIs were detected. Hence, the iPSC-derived neurons appeared to represent an early, pre-NI stage of disease. In order to produce a detectable phenotype, excitation stimuli could potentially be used to attempt to trigger the aggregation of mutant protein, as observed in SCA3 neuronal cultures (Koch et al. 2011).

Dysregulation of 12 disease-relevant genes was found in the patient iPSC-derived neurons compared to wild-type equivalents. Notably, this included the two genes which were altered in fibroblasts – DNAJA1 and UCHL1, showing similar but less significant changes. It is feasible that the extensive transcriptional and cellular reorganisation occurring during reprogramming and differentiation alters the levels of Ataxin-7 and the cellular response to its mutation.

The patient from which the 1519 neuronal cultures were developed had an expansion size of 51 repeats and hence adult-onset SCA7. Large expansions lead to much more rapid SCA7 disease onset (Yoo et al. 2003; Ansorge et al. 2004) and increased severity of some phenotypes in stem cell-derived HD neuronal cultures (Conforti et al. 2013). Therefore, the derivation of neuronal cultures from SCA7 patients with very long CAG tracts may be more likely to present with protein aggregation and phenotypic changes in the short timescale of *in vitro* neuronal differentiation.

5.5.4 Conclusions

Silencing of Ataxin-7 was elicited by mirtrons in patient-derived dermal fibroblast cells at the mRNA level. As in Chapter 3, no endogenous protein-level silencing could be detected. Lentiviruses were made which could produce mirtron expression in infected cells. These were tested in patient-derived neurons, but the distinctly neuronal cells of the population could not be infected. More success may be achieved by transducing cells at the iPSC stage, before differentiation.

Fibroblast cultures provide a relatively accessible disease cell model which can be transfected at high efficiency to test the efficacy of novel gene therapies. Although certain disease phenotypes can be detected in patient fibroblasts, SCA7 largely does not affect non-neuronal cells in patients and dermal fibroblasts therefore do not closely represent the disease state. Notably, protein aggregates could not be detected by immunostaining in fibroblasts. Neuronal cultures reprogrammed from patient fibroblast cultures theoretically provide a better model for SCA7. However, the heterogeneous populations of differentiated and undifferentiated cells make phenotypic analysis difficult. No aggregation of Ataxin-7 was detected, but 12 genes had altered transcript levels in patient-derived neuronal cultures compared to controls.

Before further development, it would be necessary to confirm the ability of the silencing technologies here to reduce endogenous Ataxin-7 protein. It is promising, however, that some phenotypic improvement can be reproducibly detected in response to siRNA treatment in terms of transcriptional changes.

6 Discussion

6.1 Overview

The aim of this project was to investigate the use of mirtrons and other intronic miRNA in novel gene knockdown-replacement therapies, focussing on SCA7 as a model. This consisted of three main areas of investigation:

- Optimisation of gene silencing for Ataxin-7 by mimicking mirtrons and miRNA clusters.
- Modification of an Ataxin-7 cDNA sequence to confer RNAi resistance, and incorporation of mirtrons into this sequence for co-delivery in transfected cell models.
- Investigating phenotypic improvements in SCA7 patient-derived cell models.

Gene therapy is advancing in clinical trials, holding great potential for treatment of otherwise incurable diseases. In particular, dominant mutations may be targeted by RNA interference for allele-specific silencing. This has the potential to be more specific than mediating downstream effects of the mutation, hence safer and more effective.

Combined gene knockdown–replacement (K&R) strategies, where a gene is silenced in a non-allele-specific manner while providing an RNAi-resistant replacement copy, may be the only feasible approach in many diseases where mutations are not easily amenable to allele-specific silencing, but where non-allele-specific silencing is potentially deleterious (Rossmiller et al. 2012). This includes the polyQ (CAG expansion) diseases. SCA7 is a rare polyQ disease, but presents a useful model. SCA7 is caused by an expanded CAG repeat in the gene

encoding Ataxin-7. As in other polyQ disorders, SCA7 pathogenesis is complex, and is likely to involve both gain- and loss-of-function.

The functions of Ataxin-7 are still emerging, but it appears that effects of its silencing on development can be severe (Yanicostas et al. 2012), hinting at the potential importance of maintaining its levels during therapy. Silencing Ataxin-7 in an unaffected model organism beyond the embryonic stage could reveal the functional roles of Ataxin-7 in adults.

There has been little data published on silencing Ataxin-7. However, importantly, transcriptional-level silencing of expanded Ataxin-7 which occurred in the *R7E* SCA7 model mouse did not reverse disease progression (Helmlinger et al. 2004a). Therefore it is important to determine whether early silencing of Ataxin-7 can lead to phenotypic improvements as in other polyQ diseases. Transgene overexpression may have contributed to the irreversibility observed in this model, and the disease symptoms may have progressed more quickly if mutant expression was sustained. Nonetheless, this indicates that therapies for SCA7 may be most beneficial with the earliest possible intervention.

6.2 Mirtron-based RNAi

RNAi therapies have been shown to direct specific gene silencing in clinical trials through the delivery of siRNA molecules (Davis et al. 2010; Leachman et al. 2010; DeVincenzo et al. 2008, 2010; Strumberg et al. 2012). Stable expression constructs in the form of short hairpin RNAs (shRNA) or artificial microRNAs (miRNAs) are preferable in their potential for providing long-term therapy.

Synthetic shRNA and miRNA are transcribed as stem-loop transcripts and then processed by the microprocessor into pre-miRNAs. Numerous studies have shown the advantages of mimicking endogenous miRNA hairpins over the use of shRNA, in terms of potency as well as

safety (Maczuga et al. 2012). Many natural miRNAs are found within introns (intronic miRNAs). Mirtrons are a special class of intronic miRNA where the ends of the short pre-miRNA stem-loops are defined by the splice sites (Ruby et al. 2007a). Therefore, the typical Drosha processing step is not involved in generating the mature product. Intronic miRNA, including mirtrons, are particularly attractive for K&R applications (see section 6.4).

Mirtrons can be adapted to successfully silence a gene of interest *in vitro* when placed as introns within a reporter gene (Seow et al. 2012; Sibley et al. 2012). A potential benefit of using mirtrons in RNAi-based gene therapy is the strict definition of the pre-miRNA though the splicing reaction, perhaps minimising off-target effects by producing a well-defined seed region compared to miRNA and shRNA processed by the microprocessor. Further, the components of the microprocessor can be sequestered into nuclear RNA aggregates, reducing their activity levels, which has been shown in one nucleotide expansion disorder so far (Sellier et al. 2013).

6.2.1 Artificial mirtrons can silence Ataxin-7

In this thesis, an algorithm was used to select appropriate target sites within Ataxin-7, against which a number of mirtrons were designed. They were screened for splicing capability and for silencing of a matched dual Luciferase construct. Although a wide variety of results were seen across the set, several had silencing activity. Two were selected as the most promising mirtrons. It was demonstrated by qPCR that these mirtrons can elicit silencing of the endogenous *ataxin-7* transcript in HEK-293 cells and both human and mouse neuroblastoma cell lines (SH-SY5Y and Neuro2A) and patient-derived fibroblasts.

Although mirtrons could reduce aggregation of mutant Ataxin-7 in transfected cells, it was not possible to detect endogenous protein-level knockdown in any cell line. This was also the

case for siRNAs which give a greater measurable effect against mRNA. This indicates that the extent of silencing and the turnover of Ataxin-7 protein do not give an observable change in 3-7 days. It is possible that a compensatory mechanism may up-regulate Ataxin-7 levels, perhaps through inhibition of its turnover. Indeed, post-translational modifications of Ataxin-7 cleavage products at Lysine 257 can influence their stability (Mookerjee et al. 2009; Janer et al. 2010). This could pose a serious issue for silencing therapies and would be an interesting avenue to investigate.

Contrary to shRNA treatment, mirtron-directed silencing was reduced at later time-points in SH-SY5Y cells and patient-derived fibroblasts. These apparent transient effects, along with the poor silencing ability of mirt-18 in SH-SY5Y, require further investigation.

6.2.2 Mirtron biogenesis

Deep sequencing demonstrated that mature miRNAs were produced from the artificial mirtrons. A lower total number of reads were detected for mirt-18 than mirt-1cN, yet achieved a similar level of silencing, further indicating that the mature species derived from mirtrons can be potent even at low levels (Seow et al. 2012). The ends of the mature species defined by Dicer were more variable than those defined by splicing. 5′-arm mirtrons were therefore more accurately processed (and also had more favourable strand bias) using the miR-1224 backbone. The range of target sites available in a given gene may necessitate the use of 3′-arm mirtrons in some cases, therefore their design may require further optimisation or the use of alternative mirtron backbones. It would be interesting to directly compare the mature products arising from mirtrons to those from their shRNA-equivalent to determine the relative accuracy of mirtron processing.

By making mutations to the 5′ splice site of mirt-18 and mirt-1cN, it was shown that the silencing activity of these mirtrons was largely dependent upon splicing. However, for both "un-spliceable" mirtrons a small but non-significant level of silencing of full-length Ataxin-7, but not Luciferase targets, was detected, and for mirt-18 a small number of mature species were detected by deep sequencing. It may be that when splicing does not occur, the hairpin-containing transcript can become available to processing by Drosha and either the canonical miRNA pathway or the simtron pathway (Havens et al. 2012).

6.2.3 Artificial mirtron design

Mirtron design involves a more restricted choice of target site compared to other hairpin sh/miRNAs, to ensure their recognition by splicesome factors as well as the downstream factors of the RNAi pathway. Weakening the hairpin structure for mirt-18 reduced its silencing capability despite the increase in splicing. It is clear that a careful balance must be achieved between these factors when designing artificial mirtrons.

The identity of the first five bases of the mirtron and the number and size of interruptions to the poly-pyrimidine tract appeared very important for splicing ability. Poly-pyrimidine tract interruptions were minimised through appropriate choice of target site and (for 5′ arm mirtrons) modification of the passenger strand. However, the limited choice of mirtron target sites does not necessarily make this possible. Inclusion of intronic pentanucleotide motifs could be investigated to improve splicing (Lim and Burge 2001).

It was demonstrated that multiple mirtrons could be delivered within a single transcript. However, the splicing efficiency of both mirtrons was compromised, though their silencing activity was not measurably reduced. The secondary structure of the transcript perhaps becomes more complex in this setting.

Off-target effects should also be investigated. This would be particularly relevant in tissues or cells representing the target tissues, i.e. brain and retinal cells. The safety profile of artificial mirtrons compared to other sh/miRNAs has not yet been investigated *in vivo*. Theoretically, the safety of mirtrons should be comparable to other miRNA mimics.

Overall, the main disadvantage of the application of artificial mirtrons is the difficulty of the design process. Their potential benefits, which remain to be experimentally demonstrated, warrant further investigation.

6.3 Other intronic miRNAs

No intronic miRNA mimics have yet been tested in therapeutic K&R applications. However, mimicking the three miRNAs in the naturally-occurring miR-106b cluster, and placing this 800 bp intron within a replacement gene is a possible option (Aagaard et al. 2008). This system may be advantageous over mirtrons in avoiding splicing-related constraints, allowing a less restricted choice of miRNA target sites, and hence potentially greater silencing. In addition, three hairpins are incorporated into a single intron, which may be more straightforward to apply to a multi-targeting K&R therapy. Further, this construct has shown silencing efficacy *in vivo* (Pihlmann et al. 2012). However, unlike mirtrons, each hairpin is reliant upon processing by Drosha.

Artificial mirtrons and a shRNA hairpin targeting Ataxin-7 were each expressed as miRNA mimics within the miR-106b cluster. Some were effective against Luciferase targets and mCherry-tagged full-length mutant Ataxin-7. However, for one position in the cluster, all three miRNAs tested were unable to produce substantial silencing. Interestingly, different hairpin positions were previously problematic (Aagaard et al. 2008; Pihlmann et al. 2012), which is perhaps a limitation of this technology.

Splicing of the exons of this construct was confirmed. However, it will be crucial to test whether splicing also occurs in the context of a different gene for application to combination therapy. Once a successful cluster combination is constructed, their molecular manipulation is then more complex compared to mirtrons. Mirtrons, at less than 100 bp, can be cloned into specific splice sites within a chosen gene cheaply by using appropriately designed long primers. Nonetheless, these constructs warrant further investigation.

6.4 Mirtron-based K&R

Mirtrons were developed into a potential combination therapy for SCA7, whereby mirtrons silencing all endogenous Ataxin-7 were co-delivered as introns within an RNAi-resistant wild-type gene which serves as a replacement.

6.4.1 Advantages of mirtron-based K&R therapy

A number of studies have proven the principle of concurrent gene silencing and replacement and shown safety and therapeutic benefits *in vivo* (Rossmiller et al. 2012). However, evidence for advantages over silencing-only therapy are currently limited, since many studies were carried out on a knockout background meaning that silencing alone was not beneficial.

It is desirable to deliver both miRNA and replacement gene in a single construct to ensure cells receive both aspects of the therapy (Greenwald et al. 2012). To achieve this, mirtrons and other intronic miRNAs could be nested within a replacement gene. Mirtrons may be particularly useful for this purpose due to their minimal size, which is desirable in limited-capacity delivery vectors such as lentiviruses and Adeno-associated viruses. Also the potential advantages attributed to their processing mechanism are also relevant, as detailed above.

6.4.2 Silent modifications in Ataxin-7 convey resistance to silencing by mirtrons

The redundancy within the genetic code was exploited to design RNAi-resistant versions of the Ataxin-7 gene which were delivered alongside mirt-18 and 1cN and substantially resistant to repression. For clinical applications it would be valuable to confirm the protein sequence and function.

6.4.3 Mirtrons and a replacement gene can be co-delivered under a single promoter

Mirtrons were co-delivered with an RNAi-resistant Ataxin-7 gene under a single promoter by cloning an EGFP-mirt cassette as a tag after the ORF of mirtron-resistant Ataxin-7. Both the expression of replacement Ataxin-7 and successful splicing of the mirtrons were indicated by EGFP fluorescence and RT-PCR. The mirtrons encoded by these constructs were produced at low levels yet they showed significant silencing activity against Luciferase constructs. This demonstrated that mirtrons can be delivered as introns under the same promoter as a replacement gene and effectively repress a target, supporting their potential for development into single-construct K&R systems.

For clinical development it would also be necessary to remove the fluorescent tag and also to produce a construct of minimal size for ease of packaging into a suitable delivery system.

6.4.4 Mirtrons can be effectively spliced from some sites in Ataxin-7

The splicing of mirtrons from a host transcript is context-dependent. This led to difficulties in finding sites within Ataxin-7 from which mirtrons could be spliced efficiently, though a number of sites were tested. Computational analysis of splice sites proved a useful tool, but

may be limited by the lack of consideration of RNA secondary structure. Sites 8 and 10, despite scoring well, did not produce good splicing efficiency.

For one site (9), mirt-1cN was spliced at approximately 50% efficiency, and 70% silencing of a Luciferase target was achieved. High-throughput sequencing confirmed that the mature mirt-1cN species was produced, with strand selection comparable to pEGFP-mirt-1cN, but slightly lower accuracy of processing. Its lower splicing efficiency may lead to some mirtron-containing transcripts being available for other miRNA processing pathways. Before further development, it will be important to confirm mirtron splicing from this construct in a neuronal cell line, since splice sites may be selectively activated or repressed according to cell type (Keppetipola et al. 2012). The ability of this construct to silence endogenous Ataxin-7 requires confirmation. Protein-level silencing and the functional capacity of the replacement Ataxin-7 gene also need to be established.

6.4.5 Further development of mirtron-based K&R

It appears that a combination of mirtron-based secondary structure, poorly understood splice site recognition and distance from the promoter can contribute to the context-dependence encountered in achieving efficient mirtron splicing from within Ataxin-7. This problem highlights the need for better understanding and prediction of mirtron splicing efficiencies, which could limit development of this strategy for other diseases. However, more thorough computational prediction may improve this situation and the constantly decreasing cost of gene synthesis allows much easier screening of different intron locations.

Inserting introns into UTRs was investigated as an alternative strategy to co-deliver mirtrons with a transgene, without the need to identify efficient splice sites in the cDNA. Once a mirtron is known to be efficiently spliced from EGFP, it may be transferred to a UTR along

with a stretch of each flanking exon to ensure a good context for splicing. The 5' UTR was chosen to minimise the chances of triggering nonsense-mediated decay and this also avoids the apparent detrimental effects of placing an intron several thousand bp from the transcriptional start site.

The NAD intron was cloned along with 8 bp of each flanking exon from EGFP and tested for splicing by RT-PCR. However, no splicing appeared to occur. Perhaps the 16 bp total flanking sequence was not sufficient to recreate the well-recognised context in the EGFP plasmid or alternatively the intron may not be a sufficient distance from the end of the promoter to allow the efficient binding of spliceosomal factors. This approach must also avoid introducing alternative start codons which would disrupt transgene expression.

Tailed-mirtron mimics are another potential option. They are dependent upon splicing for their biogenesis, but a tail on either the 5' or 3' end of the hairpin requires digestion by an exonuclease other than Drosha (Babiarz et al. 2008; Flynt et al. 2010). Tailed-mirtrons could be constructed either by mimicking endogenous tailed-mirtrons or by adding an intronic tail to artificial mirtrons. This could have the advantage of allowing one of the splice sites to be less subject to secondary structure. Endogenous tailed mirtron tails can be short (<100 nt), hence still allowing construction of a K&R system of close to minimal size. This remains a promising option which could be further investigated.

6.5 Patient-derived cells as models for SCA7

The use of standard cell lines can only indicate the effects of novel therapeutics on the expression of either mutant transgenic or wild-type endogenous genes. Therefore, the therapeutic benefits of novel therapies may be better explored in cells which naturally carry

the disease-causing mutation and show measurable phenotypes compared to controls. Two such cell models were investigated.

6.5.1 Patient fibroblasts as a model

Dermal fibroblast cultures, derived from relatively non-invasive patient skin biopsies, are straightforward to culture and can be transfected with siRNA and plasmids. They have detectable phenotypic changes in some other polyQ disorders (Seo et al. 2004; de Mezer et al. 2011). SCA7 patient-derived fibroblasts were investigated for some phenotypic features and for correction through RNAi.

Protein aggregates, a key feature of affected cells in SCA7 and other polyQ disorders, could not be detected by immunostaining in patient-derived fibroblasts, and no aggregated material was apparent by Western blotting. It may be that Ataxin-7 protein expression in this cell type is so low that aggregates do not form (or are readily degraded). Fibroblasts derived from patients with larger expansions may be more likely to present with protein aggregation (Ansorge et al. 2004). Depending on the expansion size, the accumulation of aggregations in patients takes a number of years, which is impossible to reproduce with cellular models.

Significant changes to the expression of disease-relevant genes were recently detected in SCA7 patient-derived fibroblasts (Scholefield et al, in preparation), representing some of the early changes in SCA7 pathogenesis. These changes were reproducibly partially corrected with siRNA specifically targeting Ataxin-7. Therefore, this potentially provides a measurable phenotype against which to test the efficacy of novel therapies.

However, it is not known whether this phenotypic rescue reflects the potential for improvement in animal models or patients, and as such it does not necessarily provide more information than the Ataxin-7 silencing. These effects were not achieved by shRNA or

mirtrons, which may be attributable to insufficient Ataxin-7 silencing. Other disease-related phenotypes could be investigated, such as proteasome impairment and nuclear RNA aggregation, as detected in fibroblasts derived from HD patients (Seo et al. 2004; de Mezer et al. 2011). However, aside from aggregation, non-neuronal SCA7 patient cells may be unlikely to exhibit pathogenic features, even with very large expansions (Ansorge et al. 2004).

Patient-derived fibroblasts are a simple and valuable model for testing novel SCA7 therapies. More thorough investigation of their phenotypes may reveal further advantages.

6.5.2 Patient-derived neurons as a model

Neuronal cultures were differentiated from patient-derived fibroblasts via iPS cells, providing an additional SCA7 human cell model naturally containing an Ataxin-7 expansion. Being of a more disease-relevant cell type, this model is potentially favourable over the use of fibroblasts, hence a useful tool for analysis of SCA7 pathogenesis and novel treatments. However, along with some promising features, there are a significant number of difficulties.

A relatively simple differentiation procedure was chosen. Nonetheless, iPS-derived neurons are more difficult to maintain in culture compared to fibroblasts. Heterogeneous populations were produced, containing cells of mixed neuronal types and glia, at various stages of differentiation. The variability of the neurosphere precursors made comparisons between cultures difficult. These factors may cause differences in expression of Ataxin-7 and other genes, complicating the analysis of gene expression in cell lysates. The neuronal cultures were highly problematic to transduce. It proved difficult to infect the distinctly neuronal cells of the population with the shRNA- and mirtron-expressing lentiviruses. More success may be achieved by transducing cells at the iPS stage, which can then be selected to differentiate neuronal cultures with the transgene integrated into the genome.

Expression of *ataxin-7* was higher in SCA7 patient iPSC-derived neurons than controls. However, immunostaining did not indicate corresponding differences in Ataxin-7 protein expression, nor any sign of aggregation. Hence, the iPSC-derived neurons appeared to represent an early, pre-NI stage of disease. In accordance with this, dysregulation of 12 disease-relevant genes was found in patient iPS-derived neurons compared to wild-type equivalents (Watson 2012). Most of these changes were increasingly significant through longer periods of differentiation. Notably, this included the two genes which were altered in fibroblasts – DNAJA1 and UCHL1, showing similar but less significant changes. It is feasible that the extensive transcriptional and cellular reorganisation occurring during reprogramming alters the levels of Ataxin-7 and the cellular response to its mutation.

The patients from which neuronal cultures were produced had expansion sizes of up to 51 repeats and the cultures were tested for aggregates up to six weeks post-differentiation. In comparison, SCA7 knock-in mice with 100 repeats did not display symptoms or tissue pathology until 8-9 months of age (although the presence of protein aggregation was not reported) (Chen et al. 2012). Even in mice with 266 repeats, microaggregates and large inclusions were not detected until eight and ten weeks of age, respectively (Yoo et al. 2003). Therefore, producing neuronal cultures with very long CAG tracts may be more likely to present with protein aggregation and phenotypic changes in the relatively short timescale of *in vitro* neuronal differentiation. Alternatively, in order to induce a detectable phenotype, excitation stimuli could potentially be used to attempt to trigger the aggregation of mutant protein, as observed in SCA3 neuronal cultures (Koch et al. 2011). It would also be interesting to investigate the phenotype of patient-derived Purkinje cell cultures, which are severely affected in patients.

6.5.3 SCA7 mouse models

Following success *in vitro*, novel therapies require testing in a representative animal model before advancement to the clinic. This is the only way to represent the complex cell-cell relationships which are highly relevant in the pathogenesis of SCA7.

The retinal phenotype makes SCA7 a good choice as a model of polyQ diseases. The eye is an accessible tissue which is amenable to gene therapy via local injection, allowing the testing of therapeutic efficacy in animal models (and eventually patients). Indeed, gene therapies for diseases of the retina have progressed to clinical trials. Furthermore, photoreceptor function can be measured non-invasively by electroretinogram (ERG), allowing disease progression and functional effects of therapeutic intervention to be monitored over time.

Although SCA7 disease progression has been studied extensively in a number of mouse models, few potential therapies have been tested. Knock-in models are thought to represent the disease most accurately, since over-expression of transgenes can exacerbate the phenotype of the disease, potentially biasing the outcome of therapeutic interventions.

Moreover, mouse models must be carefully designed or selected dependent upon the therapies being tested and compared. For example, to validate the allele-specificity of therapies directed toward SNPs, the mouse must possess the SNP on its mutant allele and the other allele(s) should be expressed at equivalent levels. This may represent a limitation in the development of such therapies. Non-allele-specific therapies would have to be directed against both the human and mouse genes in order to be tested in most mouse models.

Testing K&R therapies in mice would require the replacement (human) gene to rescue the function of the mouse protein. In the case of SCA7, this is not likely to present a problem

because the *ataxin-7* gene is highly conserved (McMahon et al. 2005; Yanicostas et al. 2012). Silencing Ataxin-7 in a normal mouse could reveal functional roles which are essential to maintain, further informing the development of the most appropriate novel gene therapy.

6.6 Viral delivery of mirtrons

Expression constructs require a vehicle for delivery. The most promising vectors for delivery to neuronal cells are Adeno-associated viruses (AAV), lentiviruses and nanoparticles (Auricchio et al. 2001; Cai et al. 2010; Hauswirth et al. 2008).

Lentiviruses designed to target Ataxin-7 with shRNA or mirtrons were produced, which successfully infected HEK-293 cells and human neuronal SH-SY5Y cells. Mirtron hairpins were expressed from the mirtron-expressing virus, which is encouraging for the further development of mirtrons as therapeutics in whole organisms. Surprisingly, silencing of endogenous Ataxin-7 in SH-SY5Y cells at the mRNA level was detectable only for the shRNA-expressing virus, and was minimal. The lack of silencing efficacy of mirtrons expressed from viruses may be accountable to splicing of viral RNA before packaging. This requires further investigation, through RT-PCR for splicing, comparison of mirtron expression relative to EGFP and selecting for infected cells to produce populations which all express the viral genes.

Obtaining high viral titres proved difficult, particularly for shRNA-containing viruses, despite a repetition of the packaging procedure for each virus. This limited the number of experiments which could be carried out, and inhibited the potential for *in vivo* application. This could be circumvented through commercial production, but at a substantial cost.

6.7 Further research

Key areas of research required for further development of this approach as a therapy for SCA7. Firstly, protein-level silencing could not be detected, even using siRNA, seven days after transfection. Although in fibroblasts some presumed downstream effects were observed, it is important to determine that Ataxin-7 can be reduced at the protein level. It may be that compartmentalisation or post-translational modification of endogenous Ataxin-7 may lead to high protein stability. Towards this aim, the half-life of the endogenous protein could be determined by pulse-chase analysis or cycloheximide treatment.

Secondly, the silencing effect of mirtrons appeared weakened over time in a number of cell lines, where shRNA-directed silencing was maintained. The biogenesis of mirtrons over time requires further investigation, for example by qPCR or Northern blotting, to ensure that this technology can be used to provide long-term therapeutic benefit.

Thirdly, for use in a K&R therapy, the functionality of the replacement gene should be determined. Although the function of Ataxin-7 is not well established, some assays could be used to verify that the modified gene is functional. For example, co-localisation/co-immunoprecipitation assays can be used to demonstrate interactions with known binding partners such as CRX and α-tubulin (La Spada et al. 2001; Nakamura et al. 2012). The composition and histone acetylation activity of STAGA/TFTC complexes could also be investigated (Helmlinger et al. 2004c; Palhan et al. 2005). Promoter-reporter assays can be used to demonstrate that the replacement gene functions correctly in Ataxin-7-dependent transcriptional activity (La Spada et al. 2001; Ström et al. 2005; Young et al. 2007).

Once these issues are resolved, it will be important to investigate the efficacy and safety of mirtrons *in vivo*, which will require the production of an effective delivery mechanism.

Ideally, a K&R construct which is more efficiently spliced should be developed. This could be investigated by inserting mirtrons into a different transgene to confirm that efficient splicing can be achieved. The use of tailed-mirtrons could also be investigated to increase the splicing potential of mirtron-based K&R therapies,

6.8 Concluding remarks

The work in this thesis focussed on the development of artificial mirtrons for novel knockdown-and-replacement therapies for a polyQ disorder, SCA7. SCA7 provides a unique opportunity as a model for polyQ diseases due to its retinal phenotype. The K&R approach has numerous advantages in polyQ diseases as well as other dominant conditions caused by mutations which are difficult to target. Intronic miRNAs may be ideal for this application. In particular, mirtrons allow construction of K&R therapies of minimal size and are not dependent on the microprocessor. Synthetic mirtrons were successfully developed and shown to silence Ataxin-7 mRNA in numerous cell lines. Mirtron-based K&R appears to be a promising strategy but requires more extensive investigation and optimisation, particularly related to splicing. Before further development, it would be necessary to confirm the ability of mirtrons to reduce endogenous Ataxin-7 protein and assess their safety and efficacy in animal models.

Fibroblast cultures provide a relatively accessible disease cell model which can be transfected at high efficiency to test the efficacy of novel therapies. It is promisingthat siRNA treatment appears to induce downstream effects of Ataxin-7 silencing, which may represent phenotypic improvement. However, Ataxin-7 protein aggregates could not be detected. Neuronal cultures reprogrammed from patient fibroblast cultures theoretically provide a better cell model for SCA7 and several genes had altered transcript levels. However, the

heterogeneous populations of differentiated and undifferentiated cells make phenotypic analysis difficult.

Overall, this work suggests that mirtron-based gene K&R is achievable and warrants further investigation. This approach potentially offers novel mutation-independent therapies applicable to many dominant conditions and may present a number of benefits over other miRNA- and shRNA-based approaches.

7 References

Aagaard LA, Zhang J, von Eije KJ, Li H, Saetrom P, Amarzguioui M, Rossi JJ. 2008. Engineering and optimization of the miR-106b cluster for ectopic expression of multiplexed anti-HIV RNAs. *Gene Ther* **15**: 1536–1549.

Abou-Sleymane G, Chalmel F, Helmlinger D, Lardenois A, Thibault C, Weber C, Mérienne 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.

Ajayi A, Yu X, Lindberg S, Langel U, Ström A-L. 2012. Expanded ataxin-7 cause toxicity by inducing ROS production from NADPH oxidase complexes in a stable inducible Spinocerebellar ataxia type 7 (SCA7) model. *BMC Neurosci* **13**: 86.

Alves S, Nascimento-Ferreira I, Auregan G, Hassig R, Dufour N, Brouillet E, Pedroso de Lima MC, Hantraye P, Pereira de Almeida L, Deglon N. 2008. Allele-Specific RNA Silencing of Mutant Ataxin-3 Mediates Neuroprotection in a Rat Model of Machado-Joseph Disease. *PLoS ONE* **3**: e3341.

Andersen PM, Al-Chalabi A. 2011. Clinical genetics of amyotrophic lateral sclerosis: what do we really know? *Nat Rev Neurol* **7**: 603–615.

Ando Y, Maida Y, Morinaga A, Burroughs AM, Kimura R, Chiba J, Suzuki H, Masutomi K, Hayashizaki Y. 2011. Two-step cleavage of hairpin RNA with 5’ overhangs by human DICER. *BMC Mol Biol* **12**: 6.

Ansorge O, Giunti P, Michalik A, Van Broeckhoven C, Harding B, Wood N, Scaravilli F. 2004. Ataxin-7 aggregation and ubiquitination in infantile SCA7 with 180 CAG repeats. *Ann Neurol* **56**: 448–452.

Auricchio A, Kobinger G, Anand V, Hildinger M, O’Connor E, Maguire AM, Wilson JM, Bennett J. 2001. Exchange of surface proteins impacts on viral vector cellular specificity and transduction characteristics: the retina as a model. *Hum Mol Genet* **10**: 3075–3081.

Babiarz JE, Ruby JG, Wang Y, Bartel DP, Blelloch R. 2008. Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs. *Genes Dev* **22**: 2773–2785.

Bader AG, Brown D, Winkler M. 2010. The Promise of MicroRNA Replacement Therapy. *Cancer Res* **70**: 7027–7030.

Baek D, Villén J, Shin C, Camargo FD, Gygi SP, Bartel DP. 2008. The impact of microRNAs on protein output. *Nature* **455**: 64–71.

Banno H, Katsuno M, Suzuki K, Tanaka F, Sobue G. 2012. Pathogenesis and molecular targeted therapy of spinal and bulbar muscular atrophy (SBMA). *Cell Tissue Res* **349**: 313–320.

Becher MW, Rubinsztein DC, Leggo J, Wagster MV, Stine OC, Ranen NG, Franz ML, Abbott MH, Sherr M, MacMillan JC, et al. 1997. Dentatorubral and pallidoluysian atrophy (DRPLA). Clinical and neuropathological findings in genetically confirmed North American and European pedigrees. *Mov Disord* **12**: 519–530.

Berezikov E, Chung W-J, Willis J, Cuppen E, Lai EC. 2007. Mammalian mirtron genes. *Mol Cell* **28**: 328–336.

Bogerd HP, Karnowski HW, Cai X, Shin J, Pohlers M, Cullen BR. 2010. A mammalian herpesvirus uses noncanonical expression and processing mechanisms to generate viral MicroRNAs. *Mol Cell* **37**: 135–142.

Bohula EA, Salisbury AJ, Sohail M, Playford MP, Riedemann J, Southern EM, Macaulay VM. 2003. The efficacy of small interfering RNAs targeted to the type 1 insulin-like growth factor receptor (IGF1R) is influenced by secondary structure in the IGF1R transcript. *J Biol Chem* **278**: 15991–15997.

Borel F, van Logtenstein R, Koornneef A, Maczuga P, Ritsema T, Petry H, van Deventer SJ, Jansen PL, Konstantinova P. 2011. In vivo knock-down of multidrug resistance transporters ABCC1 and ABCC2 by AAV-delivered shRNAs and by artificial miRNAs. *J RNAi Gene Silencing* **7**: 434–442.

Boudreau RL, Martins I, Davidson BL. 2009a. Artificial microRNAs as siRNA shuttles: improved safety as compared to shRNAs in vitro and in vivo. *Mol Ther* **17**: 169–175.

Boudreau RL, McBride JL, Martins I, Shen S, Xing Y, Carter BJ, Davidson BL. 2009b. Nonallele-specific silencing of mutant and wild-type huntingtin demonstrates therapeutic efficacy in Huntington's disease mice. *Mol Ther* **17**: 1053–1063.

Bowman AB, Yoo S-Y, Dantuma NP, Zoghbi HY. 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**: 679–691.

Bridge AJ, Pebernard S, Ducraux A, Nicoulaz A-L, Iggo R. 2003. Induction of an interferon response by RNAi vectors in mammalian cells. *Nat Genet* **34**: 263–264.

Brinster RL, Allen JM, Behringer RR, Gelinas RE, Palmiter RD. 1988. Introns increase transcriptional efficiency in transgenic mice. *Proc Natl Acad Sci USA* **85**: 836–840.

Browne SE, Beal MF. 2006. Oxidative damage in Huntington's disease pathogenesis. *Antioxid Redox Signal* **8**: 2061–2073.

Brunak S, Engelbrecht J, Knudsen S. 1991. Prediction of human mRNA donor and acceptor sites from the DNA sequence. *J Mol Biol* **220**: 49–65.

Bryer A, Krause A, Bill P, Davids V, Bryant D, Butler J, Heckmann J, Ramesar R, Greenberg J. 2003. The hereditary adult-onset ataxias in South Africa. *J Neurol Sci* **216**: 47–54.

Buratti E, Baralle FE. 2004. Influence of RNA secondary structure on the pre-mRNA splicing process. *Mol Cell Biol* **24**: 10505–10514.

Burns DM, D'Ambrogio A, Nottrott S, Richter JD. 2011. CPEB and two poly(A) polymerases control miR-122 stability and p53 mRNA translation. *Nature* **473**: 105–108.

Bustin SA. 2010. Why the need for qPCR publication guidelines?--The case for MIQE. *Methods* **50**: 217–226.

Cai X, Conley SM, Nash Z, Fliesler SJ, Cooper MJ, Naash MI. 2010. Gene delivery to mitotic and postmitotic photoreceptors via compacted DNA nanoparticles results in improved phenotype in a mouse model of retinitis pigmentosa. *FASEB J* **24**: 1178–1191.

Camnasio S, Delli Carri A, Lombardo A, Grad I, Mariotti C, Castucci A, Rozell B, Riso PL, Castiglioni V, Zuccato C, et al. 2012. The first reported generation of several induced pluripotent stem cell lines from homozygous and heterozygous Huntington's disease patients demonstrates mutation related enhanced lysosomal activity. *Neurobiol Dis* **46**: 41–51.

Cancel G, Duyckaerts C, Holmberg M, Zander C, Yvert G, Lebre AS, 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**: 2519–2530.

Cashman SM, Binkley EA, Kumar-Singh R. 2005. Towards mutation-independent silencing of genes involved in retinal degeneration by RNA interference. *Gene Ther* **12**: 1223–1228.

Chatterjee P, Cheung Y, Liew C. 2011. Transfecting and nucleofecting human induced pluripotent stem cells. *J Vis Exp*.

Cheloufi S, Dos Santos CO, Chong MMW, Hannon GJ. 2010. A dicer-independent miRNA biogenesis pathway that requires Ago catalysis. *Nature* **465**: 584–589.

Chen S, Peng G-H, Wang X, Smith AC, Grote SK, Sopher BL, La Spada AR. 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.

Chen YC, Gatchel JR, Lewis RW, Mao C-A, Grant PA, Zoghbi HY, Dent SYR. 2012. Gcn5 loss-of-function accelerates cerebellar and retinal degeneration in a SCA7 mouse model. *Hum Mol Genet* **21**: 394–405.

Chi SW, Zang JB, Mele A, Darnell RB. 2009. Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. *Nature* **460**: 479–486.

Chiang HR, Schoenfeld LW, Ruby JG, Auyeung VC, Spies N, Baek D, Johnston WK, Russ C, Luo S, Babiarz JE, et al. 2010. Mammalian microRNAs: experimental evaluation of novel and previously annotated genes. *Genes Dev* **24**: 992–1009.

Chou A-H, Chen C-Y, Chen S-Y, Chen W-J, Chen Y-L, Weng Y-S, Wang H-L. 2010. Polyglutamine-expanded ataxin-7 causes cerebellar dysfunction by inducing transcriptional dysregulation. *Neurochem Int* **56**: 329–339.

Christine CW, Starr PA, Larson PS, Eberling JL, Jagust WJ, Hawkins RA, VanBrocklin HF, Wright JF, Bankiewicz KS, Aminoff MJ. 2009. Safety and tolerability of putaminal AADC gene therapy for Parkinson disease. *Neurology* **73**: 1662–1669.

Chung K-H, Hart CC, Al-Bassam S, Avery A, Taylor J, Patel PD, Vojtek AB, Turner DL. 2006. Polycistronic RNA polymerase II expression vectors for RNA interference based on BIC/miR-155. *Nucleic Acids Res* **34**: e53.

Cifuentes D, Xue H, Taylor DW, Patnode H, Mishima Y, Cheloufi S, Ma E, Mane S, Hannon GJ, Lawson ND, et al. 2010. A novel miRNA processing pathway independent of Dicer requires Argonaute2 catalytic activity. *Science* **328**: 1694–1698.

Comyn O, Lee E, MacLaren RE. 2010. Induced pluripotent stem cell therapies for retinal disease. *Curr Opin Neurol* **23**: 4–9.

Conforti P, Camnasio S, Mutti C, Valenza M, Thompson M, Fossale E, Zeitlin S, MacDonald ME, Zuccato C, Cattaneo E. 2013. Lack of huntingtin promotes neural stem cells differentiation into glial cells while neurons expressing huntingtin with expanded polyglutamine tracts undergo cell death. *Neurobiol Dis* **50**: 160–170.

Crabb TL, Lam BJ, Hertel KJ. 2010. Retention of spliceosomal components along ligated exons ensures efficient removal of multiple introns. *RNA* **16**: 1786–1796.

Croce CM. 2009. Causes and consequences of microRNA dysregulation in cancer. *Nat Rev Genet* **10**: 704–714.

Cullen BR. 2009. Viral and cellular messenger RNA targets of viral microRNAs. *Nature* **457**: 421–425.

Curtis HJ, Sibley CR, Wood MJA. 2012. Mirtrons, an emerging class of atypical miRNA. *Wiley Interdiscip Rev RNA* **3**: 617–632.

Custer SK, Garden GA, Gill N, Rueb U, Libby RT, Schultz C, Guyenet SJ, Deller T, Westrum LE, Sopher BL, et al. 2006. Bergmann glia expression of polyglutamine-expanded ataxin-7 produces neurodegeneration by impairing glutamate transport. *Nat Neurosci* **9**: 1302–1311.

Czech B, Malone CD, Zhou R, Stark A, Schlingeheyde C, Dus M, Perrimon N, Kellis M, Wohlschlegel JA, Sachidanandam R, et al. 2008. An endogenous small interfering RNA pathway in Drosophila. *Nature* **453**: 798–802.

Darty K, Denise A, Ponty Y. 2009. VARNA: Interactive drawing and editing of the RNA secondary structure. *Bioinformatics* **25**: 1974–1975.

David G, Abbas N, Stevanin G, Dürr 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.

Davis ME, Zuckerman JE, Choi CHJ, Seligson D, Tolcher A, Alabi CA, Yen Y, Heidel JD, Ribas A. 2010. Evidence of RNAi in humans from systemically administered siRNA via targeted nanoparticles. *Nature* **464**: 1067–1070.

DeVincenzo J, Cehelsky JE, Alvarez R, Elbashir S, Harborth J, Toudjarska I, Nechev L, Murugaiah V, Van Vliet A, Vaishnaw AK, et al. 2008. Evaluation of the safety, tolerability and pharmacokinetics of ALN-RSV01, a novel RNAi antiviral therapeutic directed against respiratory syncytial virus (RSV). *Antiviral Res* **77**: 225–231.

DeVincenzo J, Lambkin-Williams R, Wilkinson T, Cehelsky J, Nochur S, Walsh E, Meyers R, Gollob J, Vaishnaw A. 2010. A randomized, double-blind, placebo-controlled study of an RNAi-based therapy directed against respiratory syncytial virus. *Proc Natl Acad Sci USA* **107**: 8800–8805.

DiFiglia M, Sena-Esteves M, Chase K, Sapp E, Pfister E, Sass M, Yoder J, Reeves P, Pandey RK, Rajeev KG, et al. 2007. Therapeutic silencing of mutant huntingtin with siRNA attenuates striatal and cortical neuropathology and behavioral deficits. *Proc Natl Acad Sci USA* **104**: 17204–17209.

Dimos JT, Rodolfa KT, Niakan KK, Weisenthal LM, Mitsumoto H, Chung W, Croft GF, 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* **321**: 1218–1221.

Doench JG, Sharp PA. 2004. Specificity of microRNA target selection in translational repression. *Genes Dev* **18**: 504–511.

Dogan RI, Getoor L, Wilbur WJ, Mount SM. 2007. Features generated for computational splice-site prediction correspond to functional elements. *BMC Bioinformatics* **8**: 410.

Dottori M, Pera MF. 2008. Neural differentiation of human embryonic stem cells. *Methods Mol Biol* **438**: 19–30.

Dykxhoorn DM, Novina CD, Sharp PA. 2003. Killing the messenger: short RNAs that silence gene expression. *Nat Rev Mol Cell Biol* **4**: 457–467.

Ebert AD, Yu J, Rose FF Jr, Mattis VB, Lorson CL, Thomson JA, Svendsen CN. 2009. Induced pluripotent stem cells from a spinal muscular atrophy patient. *Nature* **457**: 277–280.

Ebert MS, Neilson JR, Sharp PA. 2007. MicroRNA sponges: competitive inhibitors of small RNAs in mammalian cells. *Nat Methods* **4**: 721–726.

Eden E, Brunak S. 2004. Analysis and recognition of 5’ UTR intron splice sites in human pre-mRNA. *Nucleic Acids Res* **32**: 1131–1142.

Einum DD, Clark AM, Townsend JJ, Ptacek LJ, Fu Y-H. 2003. A novel central nervous system-enriched spinocerebellar ataxia type 7 gene product. *Arch Neurol* **60**: 97–103.

Einum DD, Townsend JJ, Ptácek LJ, Fu YH. 2001. Ataxin-7 expression analysis in controls and spinocerebellar ataxia type 7 patients. *Neurogenetics* **3**: 83–90.

Eisenberg I, Alexander MS, Kunkel LM. 2009. miRNAS in normal and diseased skeletal muscle. *J Cell Mol Med* **13**: 2–11.

Elbashir SM, Harborth J, Lendeckel W, Yalcin A, Weber K, Tuschl T. 2001. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **411**: 494–498.

Ely A, Naidoo T, Mufamadi S, Crowther C, Arbuthnot P. 2008. Expressed anti-HBV primary microRNA shuttles inhibit viral replication efficiently in vitro and in vivo. *Mol Ther* **16**: 1105–1112.

Ender C, Krek A, Friedländer MR, Beitzinger M, Weinmann L, Chen W, Pfeffer S, Rajewsky N, Meister G. 2008. A human snoRNA with microRNA-like functions. *Mol Cell* **32**: 519–528.

Evans SJ, Douglas I, Rawlins MD, Wexler NS, Tabrizi SJ, Smeeth L. 2013. Prevalence of adult Huntington's disease in the UK based on diagnoses recorded in general practice records. *J Neurol Neurosurg Psychiatr*.

Feng Z, Gao F. 2012. Stem cell challenges in the treatment of neurodegenerative disease. *CNS Neurosci Ther* **18**: 142–148.

Fernandez-Valverde SL, Taft RJ, Mattick JS. 2010. Dynamic isomiR regulation in Drosophila development. *RNA* **16**: 1881–1888.

Feyeux M, Bourgois-Rocha F, Redfern A, Giles P, Lefort N, Aubert S, Bonnefond C, Bugi A, Ruiz M, Deglon N, et al. 2012. Early transcriptional changes linked to naturally occurring Huntington's disease mutations in neural derivatives of human embryonic stem cells. *Hum Mol Genet* **21**: 3883–3895.

Flynt AS, Greimann JC, Chung W-J, Lima CD, Lai EC. 2010. MicroRNA biogenesis via splicing and exosome-mediated trimming in Drosophila. *Mol Cell* **38**: 900–907.

Furrer SA, Mohanachandran MS, Waldherr SM, Chang C, Damian VA, Sopher BL, Garden GA, La Spada AR. 2011. Spinocerebellar ataxia type 7 cerebellar disease requires the coordinated action of mutant ataxin-7 in neurons and glia, and displays non-cell-autonomous bergmann glia degeneration. *J Neurosci* **31**: 16269–16278.

Garden GA, Libby RT, Fu Y-H, Kinoshita Y, Huang J, Possin DE, Smith AC, Martinez RA, Fine GC, Grote SK, 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**: 4897–4905.

Gatchel JR, Watase K, Thaller C, Carson JP, Jafar-Nejad P, Shaw C, Zu T, Orr HT, Zoghbi HY. 2008. The insulin-like growth factor pathway is altered in spinocerebellar ataxia type 1 and type 7. *Proc Natl Acad Sci USA* **105**: 1291–1296.

Giering JC, Grimm D, Storm TA, Kay MA. 2008. Expression of shRNA from a tissue-specific pol II promoter is an effective and safe RNAi therapeutic. *Mol Ther* **16**: 1630–1636.

Ginn SL, Alexander IE, Edelstein ML, Abedi MR, Wixon J. 2013. Gene therapy clinical trials worldwide to 2012 - an update. *J Gene Med* **15**: 65–77.

Giunti P, Stevanin G, Worth PF, David G, Brice A, Wood NW. 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**: 1594–1603.

Goguel V, Wang Y, Rosbash M. 1993. Short artificial hairpins sequester splicing signals and inhibit yeast pre-mRNA splicing. *Mol Cell Biol* **13**: 6841–6848.

Gouw LG, Castañeda MA, McKenna CK, Digre KB, Pulst SM, Perlman S, Lee MS, Gomez C, Fischbeck K, Gagnon D, et al. 1998. Analysis of the dynamic mutation in the SCA7 gene shows marked parental effects on CAG repeat transmission. *Hum Mol Genet* **7**: 525–532.

Greenberg J, Solomon GAE, Vorster AA, Heckmann J, Bryer A. 2006. Origin of the SCA7 gene mutation in South Africa: implications for molecular diagnostics. *Clin Genet* **70**: 415–417.

Greenwald DL, Cashman SM, Kumar-Singh R. 2012. Mutation-independent rescue of a novel mouse model of Retinitis Pigmentosa. *Gene Ther* **20**: 425–434.

Grimm D. 2011. The dose can make the poison: lessons learned from adverse in vivo toxicities caused by RNAi overexpression. *Silence* **2**: 8.

Grimm D, Streetz KL, Jopling CL, Storm TA, Pandey K, Davis CR, Marion P, Salazar F, Kay MA. 2006. Fatality in mice due to oversaturation of cellular microRNA/short hairpin RNA pathways. *Nature* **441**: 537–541.

Grondin R, Kaytor MD, Ai Y, Nelson PT, Thakker DR, Heisel J, Weatherspoon MR, Blum JL, Burright EN, Zhang Z, et al. 2012. Six-month partial suppression of Huntingtin is well tolerated in the adult rhesus striatum. *Brain* **135**: 1197–1209.

Hacein-Bey-Abina S, Von Kalle C, Schmidt M, McCormack MP, Wulffraat N, Leboulch P, Lim A, Osborne CS, Pawliuk R, Morillon E, et al. 2003. LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. *Science* **302**: 415–419.

Hafner M, Landthaler M, Burger L, Khorshid M, Hausser J, Berninger P, Rothballer A, Ascano M Jr, Jungkamp A-C, Munschauer M, et al. 2010. Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* **141**: 129–141.

Haley B, Zamore PD. 2004. Kinetic analysis of the RNAi enzyme complex. *Nat Struct Mol Biol* **11**: 599–606.

Han J, Lee Y, Yeom K-H, Nam J-W, Heo I, Rhee J-K, Sohn SY, Cho Y, Zhang B-T, Kim VN. 2006. Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. *Cell* **125**: 887–901.

Han J, Pedersen JS, Kwon SC, Belair CD, Kim Y-K, Yeom K-H, Yang W-Y, Haussler D, Blelloch R, Kim VN. 2009. Posttranscriptional crossregulation between Drosha and DGCR8. *Cell* **136**: 75–84.

Hannon GJ, Rossi JJ. 2004. Unlocking the potential of the human genome with RNA interference. *Nature* **431**: 371–378.

Haussecker D, Huang Y, Lau A, Parameswaran P, Fire AZ, Kay MA. 2010. Human tRNA-derived small RNAs in the global regulation of RNA silencing. *RNA* **16**: 673–695.

Hauswirth WW, Aleman TS, Kaushal S, Cideciyan AV, Schwartz SB, Wang L, Conlon TJ, Boye SL, Flotte TR, Byrne BJ, et al. 2008. Treatment of leber congenital amaurosis due to RPE65 mutations by ocular subretinal injection of adeno-associated virus gene vector: short-term results of a phase I trial. *Hum Gene Ther* **19**: 979–990.

Havens MA, Reich AA, Duelli DM, Hastings ML. 2012. Biogenesis of mammalian microRNAs by a non-canonical processing pathway. *Nucleic Acids Res* **40**: 4626–4640.

Helmlinger D, Abou-Sleymane G, Yvert G, Rousseau S, Weber C, Trottier Y, Mandel J-L, Devys D. 2004a. Disease progression despite early loss of polyglutamine protein expression in SCA7 mouse model. *J Neurosci* **24**: 1881–1887.

Helmlinger D, Bonnet J, Mandel J-L, Trottier Y, Devys D. 2004b. Hsp70 and Hsp40 chaperones do not modulate retinal phenotype in SCA7 mice. *J Biol Chem* **279**: 55969–55977.

Helmlinger D, Hardy S, Abou-Sleymane G, Eberlin A, Bowman AB, Gansmüller A, Picaud S, Zoghbi HY, Trottier Y, Tora L, et al. 2006. Glutamine-expanded ataxin-7 alters TFTC/STAGA recruitment and chromatin structure leading to photoreceptor dysfunction. *PLoS Biol* **4**: e67.

Helmlinger D, Hardy S, Sasorith S, Klein F, Robert F, Weber C, Miguet L, Potier N, Van-Dorsselaer A, Wurtz J-M, et al. 2004c. Ataxin-7 is a subunit of GCN5 histone acetyltransferase-containing complexes. *Hum Mol Genet* **13**: 1257–1265.

Helmlinger D, Yvert G, Picaud S, Merienne K, Sahel J, Mandel J-L, Devys D. 2002. Progressive retinal degeneration and dysfunction in R6 Huntington's disease mice. *Hum Mol Genet* **11**: 3351–3359.

Hersheson J, Haworth A, Houlden H. 2012. The inherited ataxias: genetic heterogeneity, mutation databases, and future directions in research and clinical diagnostics. *Hum Mutat* **33**: 1324–1332.

Hjelm BE, Rosenberg JB, Szelinger S, Sue LI, Beach TG, Huentelman MJ, Craig DW. 2011. Induction of pluripotent stem cells from autopsy donor-derived somatic cells. *Neurosci Lett* **502**: 219–224.

Holmberg M, Duyckaerts C, Dürr A, Cancel G, Gourfinkel-An I, Damier P, Faucheux B, Trottier Y, Hirsch EC, Agid Y, et al. 1998. Spinocerebellar ataxia type 7 (SCA7): a neurodegenerative disorder with neuronal intranuclear inclusions. *Hum Mol Genet* **7**: 913–918.

Hong X, Scofield DG, Lynch M. 2006. Intron size, abundance, and distribution within untranslated regions of genes. *Mol Biol Evol* **23**: 2392–2404.

Hu J, Liu J, Corey DR. 2010. Allele-selective inhibition of huntingtin expression by switching to an miRNA-like RNAi mechanism. *Chem Biol* **17**: 1183–1188.

Hu J, Matsui M, Gagnon KT, Schwartz JC, Gabillet S, Arar K, Wu J, Bezprozvanny I, Corey DR. 2009. Allele-specific silencing of mutant huntingtin and ataxin-3 genes by targeting expanded CAG repeats in mRNAs. *Nat Biotechnol* **27**: 478–484.

Huang CC, Faber PW, Persichetti F, Mittal V, Vonsattel JP, MacDonald ME, Gusella JF. 1998. Amyloid formation by mutant huntingtin: threshold, progressivity and recruitment of normal polyglutamine proteins. *Somat Cell Mol Genet* **24**: 217–233.

Huang C-C, Sugino K, Shima Y, Guo C, Bai S, Mensh BD, Nelson SB, Hantman AW. 2013. Convergence of pontine and proprioceptive streams onto multimodal cerebellar granule cells. *Elife* **2**: e00400.

Huynh DP, Yang H-T, Vakharia H, Nguyen D, Pulst SM. 2003. Expansion of the polyQ repeat in ataxin-2 alters its Golgi localization, disrupts the Golgi complex and causes cell death. *Hum Mol Genet* **12**: 1485–1496.

Jacobson SG, Cideciyan AV, Ratnakaram R, Heon E, Schwartz SB, Roman AJ, Peden MC, Aleman TS, Boye SL, Sumaroka A, et al. 2012. Gene therapy for leber congenital amaurosis caused by RPE65 mutations: safety and efficacy in 15 children and adults followed up to 3 years. *Arch Ophthalmol* **130**: 9–24.

Janer A, Martin E, Muriel M-P, Latouche M, Fujigasaki H, Ruberg M, Brice A, Trottier Y, Sittler A. 2006. PML clastosomes prevent nuclear accumulation of mutant ataxin-7 and other polyglutamine proteins. *J Cell Biol* **174**: 65–76.

Janer A, Werner A, Takahashi-Fujigasaki J, Daret A, Fujigasaki H, Takada K, Duyckaerts C, Brice A, Dejean A, Sittler A. 2010. SUMOylation attenuates the aggregation propensity and cellular toxicity of the polyglutamine expanded ataxin-7. *Hum Mol Genet* **19**: 181–195.

Jeon CJ, Strettoi E, Masland RH. 1998. The major cell populations of the mouse retina. *J Neurosci* **18**: 8936–8946.

John B, Enright AJ, Aravin A, Tuschl T, Sander C, Marks DS. 2004. Human MicroRNA targets. *PLoS Biol* **2**: e363.

Jones D, Li Y, He Y, Xu Z, Chen H, Min W. 2012. Mirtron MicroRNA-1236 Inhibits VEGFR-3 Signaling During Inflammatory Lymphangiogenesis. *Arteriosclerosis, Thrombosis, and Vascular Biology* **32**: 633–642.

Jonsson JJ, Habel DE, McIvor RS. 1995. Retrovirus-mediated transduction of an engineered intron-containing purine nucleoside phosphorylase gene. *Hum Gene Ther* **6**: 611–623.

Kahle JJ, Gulbahce N, Shaw CA, Lim J, Hill DE, Barabási A-L, Zoghbi HY. 2011. Comparison of an expanded ataxia interactome with patient medical records reveals a relationship between macular degeneration and ataxia. *Hum Mol Genet* **20**: 510–527.

Kamata M, Liu S, Liang M, Nagaoka Y, Chen ISY. 2010. Generation of human induced pluripotent stem cells bearing an anti-HIV transgene by a lentiviral vector carrying an internal murine leukemia virus promoter. *Hum Gene Ther* **21**: 1555–1567.

Katoh T, Sakaguchi Y, Miyauchi K, Suzuki T, Kashiwabara S-I, Baba T, Suzuki T. 2009. Selective stabilization of mammalian microRNAs by 3' adenylation mediated by the cytoplasmic poly(A) polymerase GLD-2. *Genes Dev* **23**: 433–438.

Keppetipola N, Sharma S, Li Q, Black DL. 2012. Neuronal regulation of pre-mRNA splicing by polypyrimidine tract binding proteins, PTBP1 and PTBP2. *Crit Rev Biochem Mol Biol* **47**: 360–378.

Khvorova A, Reynolds A, Jayasena SD. 2003. Functional siRNAs and miRNAs exhibit strand bias. *Cell* **115**: 209–216.

Kiang A-S, Palfi A, Ader M, Kenna PF, Millington-Ward S, Clark G, Kennan A, O'reilly M, Tam LCT, Aherne A, et al. 2005. Toward a gene therapy for dominant disease: validation of an RNA interference-based mutation-independent approach. *Mol Ther* **12**: 555–561.

Kim J, Inoue K, Ishii J, Vanti WB, Voronov SV, Murchison E, Hannon G, Abeliovich A. 2007. A MicroRNA feedback circuit in midbrain dopamine neurons. *Science* **317**: 1220–1224.

Kim S-J, Kim T, Hong S, Rhim H, Kim IY, Kang S. 2003. Oxidative stimuli affect polyglutamine aggregation and cell death in human mutant ataxin-1-expressing cells. *Neurosci Lett* **348**: 21–24.

Kim Y-K, Kim VN. 2007. Processing of intronic microRNAs. *EMBO J* **26**: 775–783.

Kimchi-Sarfaty C, Oh JM, Kim I-W, Sauna ZE, Calcagno AM, Ambudkar SV, Gottesman MM. 2007. A "silent" polymorphism in the MDR1 gene changes substrate specificity. *Science* **315**: 525–528.

Kizilyaprak C, Spehner D, Devys D, Schultz P. 2011. The linker histone H1C contributes to the SCA7 nuclear phenotype. *Nucleus* **2**: 444–454.

Klionsky DJ, Emr SD. 2000. Autophagy as a regulated pathway of cellular degradation. *Science* **290**: 1717–1721.

Koch P, Breuer P, Peitz M, Jungverdorben J, Kesavan J, Poppe D, Doerr J, Ladewig J, Mertens J, Tüting T, et al. 2011. Excitation-induced ataxin-3 aggregation in neurons from patients with Machado-Joseph disease. *Nature* **480**: 543–546.

Koeppen AH. 2005. The pathogenesis of spinocerebellar ataxia. *Cerebellum* **4**: 62–73.

Köhler A, Schneider M, Cabal GG, Nehrbass U, Hurt E. 2008. Yeast Ataxin-7 links histone deubiquitination with gene gating and mRNA export. *Nat Cell Biol* **10**: 707–715.

Kordasiewicz HB, Stanek LM, Wancewicz EV, Mazur C, McAlonis MM, Pytel KA, Artates JW, Weiss A, Cheng SH, Shihabuddin LS, et al. 2012. Sustained therapeutic reversal of Huntington's disease by transient repression of huntingtin synthesis. *Neuron* **74**: 1031–1044.

Kozomara A, Griffiths-Jones S. 2011. miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res* **39**: D152–157.

Krämer A, Keller W, Appel B, Lührmann R. 1984. The 5’ terminus of the RNA moiety of U1 small nuclear ribonucleoprotein particles is required for the splicing of messenger RNA precursors. *Cell* **38**: 299–307.

Kubodera T, Yamada H, Anzai M, Ohira S, Yokota S, Hirai Y, Mochizuki H, Shimada T, Mitani T, Mizusawa H, et al. 2011. In vivo application of an RNAi strategy for the selective suppression of a mutant allele. *Hum Gene Ther* **22**: 27–34.

Kubodera T, Yokota T, Ishikawa K, Mizusawa H. 2005. New RNAi strategy for selective suppression of a mutant allele in polyglutamine disease. *Oligonucleotides* **15**: 298–302.

Ladewig E, Okamura K, Flynt AS, Westholm JO, Lai EC. 2012. Discovery of hundreds of mirtrons in mouse and human small RNA data. *Genome Res* **22**: 1634–1645.

Leachman SA, Hickerson RP, Schwartz ME, Bullough EE, Hutcherson SL, Boucher KM, Hansen CD, Eliason MJ, Srivatsa GS, Kornbrust DJ, et al. 2010. First-in-human mutation-targeted siRNA phase Ib trial of an inherited skin disorder. *Mol Ther* **18**: 442–446.

Lebbink RJ, Lowe M, Chan T, Khine H, Wang X, McManus MT. 2011. Polymerase II promoter strength determines efficacy of microRNA adapted shRNAs. *PLoS ONE* **6**: e26213.

Lee I, Ajay SS, Yook JI, Kim HS, Hong SH, Kim NH, Dhanasekaran SM, Chinnaiyan AM, Athey BD. 2009. New class of microRNA targets containing simultaneous 5’-UTR and 3’-UTR interaction sites. *Genome Res* **19**: 1175–1183.

Lee Y, Ahn C, Han J, Choi H, Kim J, Yim J, Lee J, Provost P, Rådmark O, Kim S, et al. 2003. The nuclear RNase III Drosha initiates microRNA processing. *Nature* **425**: 415–419.

Li F, Mahato RI. 2009. Bipartite vectors for co-expression of a growth factor cDNA and short hairpin RNA against an apoptotic gene. *J Gene Med* **11**: 764–771.

Li N, You X, Chen T, Mackowiak SD, Friedländer MR, Weigt M, Du H, Gogol-Döring A, Chang Z, Dieterich C, et al. 2013. Global profiling of miRNAs and the hairpin precursors: insights into miRNA processing and novel miRNA discovery. *Nucleic Acids Res* **41**: 3619–3634.

Liang FQ, Dejneka NS, Cohen DR, Krasnoperova NV, Lem J, Maguire AM, Dudus L, Fisher KJ, Bennett J. 2001. AAV-mediated delivery of ciliary neurotrophic factor prolongs photoreceptor survival in the rhodopsin knockout mouse. *Mol Ther* **3**: 241–248.

Lim LP, Burge CB. 2001. A computational analysis of sequence features involved in recognition of short introns. *Proc Natl Acad Sci USA* **98**: 11193–11198.

Lim LP, Lau NC, Garrett-Engele P, Grimson A, Schelter JM, Castle J, Bartel DP, Linsley PS, Johnson JM. 2005. Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature* **433**: 769–773.

Lindenberg KS, Yvert G, Müller K, Landwehrmeyer GB. 2000. Expression analysis of ataxin-7 mRNA and protein in human brain: evidence for a widespread distribution and focal protein accumulation. *Brain Pathol* **10**: 385–394.

Linterman KS, Palmer DN, Kay GW, Barry LA, Mitchell NL, McFarlane RG, Black MA, Sands MS, Hughes SM. 2011. Lentiviral-Mediated Gene Transfer to the Sheep Brain: Implications for Gene Therapy in Batten Disease. *Hum Gene Ther* **22**: 1011–1020.

Lipinski DM, Singh MS, MacLaren RE. 2011. Assessment of cone survival in response to CNTF, GDNF, and VEGF165b in a novel ex vivo model of end-stage retinitis pigmentosa. *Invest Ophthalmol Vis Sci* **52**: 7340–7346.

Liu J, Carmell MA, Rivas FV, Marsden CG, Thomson JM, Song J-J, Hammond SM, Joshua-Tor L, Hannon GJ. 2004. Argonaute2 is the catalytic engine of mammalian RNAi. *Science* **305**: 1437–1441.

Lorenz R, Bernhart SH, Honer zu Siederdissen C, Tafer H, Flamm C, Stadler PF, Hofacker IL. 2011. ViennaRNA Package 2.0. *Algorithms Mol Biol* **6**: 26.

Lu S, Cullen BR. 2004. Adenovirus VA1 noncoding RNA can inhibit small interfering RNA and MicroRNA biogenesis. *J Virol* **78**: 12868–12876.

Lund E, Guttinger S, Calado A, Dahlberg JE, Kutay U. 2004. Nuclear export of microRNA precursors. *Science* **303**: 95–98.

Lundberg C, Björklund T, Carlsson T, Jakobsson J, Hantraye P, Déglon N, Kirik D. 2008. Applications of lentiviral vectors for biology and gene therapy of neurological disorders. *Curr Gene Ther* **8**: 461–473.

Luo Y, Fan Y, Zhou B, Xu Z, Chen Y, Sun X. 2012. Generation of induced pluripotent stem cells from skin fibroblasts of a patient with olivopontocerebellar atrophy. *Tohoku J Exp Med* **226**: 151–159.

Maczuga P, Koornneef A, Borel F, Petry H, van Deventer S, Ritsema T, Konstantinova P. 2012. Optimization and comparison of knockdown efficacy between polymerase II expressed shRNA and artificial miRNA targeting luciferase and Apolipoprotein B100. *BMC Biotechnol* **12**: 42.

Marashi S-A, Eslahchi C, Pezeshk H, Sadeghi M. 2006. Impact of RNA structure on the prediction of donor and acceptor splice sites. *BMC Bioinformatics* **7**: 297.

Marks WJ Jr, Bartus RT, Siffert J, Davis CS, Lozano A, Boulis N, Vitek J, Stacy M, Turner D, Verhagen L, et al. 2010. Gene delivery of AAV2-neurturin for Parkinson's disease: a double-blind, randomised, controlled trial. *Lancet Neurol* **9**: 1164–1172.

Martin JJ, Van Regemorter N, Krols L, Brucher JM, de Barsy T, Szliwowski H, Evrard P, Ceuterick C, Tassignon MJ, Smet-Dieleman H. 1994. On an autosomal dominant form of retinal-cerebellar degeneration: an autopsy study of five patients in one family. *Acta Neuropathol* **88**: 277–286.

Mauger C, Del-Favero J, Ceuterick C, Lübke U, van Broeckhoven C, 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.

McBride JL, Boudreau RL, Harper SQ, Staber PD, Monteys AM, Martins I, Gilmore BL, Burstein H, Peluso RW, Polisky B, et al. 2008. Artificial miRNAs mitigate shRNA-mediated toxicity in the brain: implications for the therapeutic development of RNAi. *Proc Natl Acad Sci USA* **105**: 5868–5873.

McCullough SD, Xu X, Dent SYR, Bekiranov S, Roeder RG, Grant PA. 2012. Reelin is a target of polyglutamine expanded ataxin-7 in human spinocerebellar ataxia type 7 (SCA7) astrocytes. *Proc Natl Acad Sci USA* **109**: 21319–21324.

McMahon SJ, Pray-Grant MG, Schieltz D, Yates JR, Grant PA. 2005. Polyglutamine-expanded spinocerebellar ataxia-7 protein disrupts normal SAGA and SLIK histone acetyltransferase activity. *PNAS* **102**: 8478–8482.

De Mezer M, Wojciechowska M, Napierala M, Sobczak K, Krzyzosiak WJ. 2011. Mutant CAG repeats of Huntingtin transcript fold into hairpins, form nuclear foci and are targets for RNA interference. *Nucleic Acids Res* **39**: 3852–3863.

Millington-Ward S, Chadderton N, O'Reilly M, Palfi A, Goldmann T, Kilty C, Humphries M, Wolfrum U, Bennett J, Humphries P, et al. 2011. Suppression and replacement gene therapy for autosomal dominant disease in a murine model of dominant retinitis pigmentosa. *Mol Ther* **19**: 642–649.

Montell C, Fisher EF, Caruthers MH, Berk AJ. 1984. Control of adenovirus E1B mRNA synthesis by a shift in the activities of RNA splice sites. *Mol Cell Biol* **4**: 966–972.

Montes M, Becerra S, Sánchez-Álvarez M, Suñé C. 2012. Functional coupling of transcription and splicing. *Gene* **501**: 104–117.

Mookerjee S, Papanikolaou T, Guyenet SJ, Sampath V, Lin A, Vitelli C, DeGiacomo F, Sopher BL, Chen SF, La Spada AR, et al. 2009. Posttranslational modification of ataxin-7 at lysine 257 prevents autophagy-mediated turnover of an N-terminal caspase-7 cleavage fragment. *J Neurosci* **29**: 15134–15144.

Morlando M, Ballarino M, Gromak N, Pagano F, Bozzoni I, Proudfoot NJ. 2008. Primary microRNA transcripts are processed co-transcriptionally. *Nat Struct Mol Biol* **15**: 902–909.

Mount SM. 1982. A catalogue of splice junction sequences. *Nucleic Acids Res* **10**: 459–472.

Mulders SAM, van den Broek WJAA, Wheeler TM, Croes HJE, van Kuik-Romeijn P, de Kimpe SJ, Furling D, Platenburg GJ, Gourdon G, Thornton CA, et al. 2009. Triplet-repeat oligonucleotide-mediated reversal of RNA toxicity in myotonic dystrophy. *Proc Natl Acad Sci USA* **106**: 13915–13920.

Murphy D, Dancis B, Brown JR. 2008. The evolution of core proteins involved in microRNA biogenesis. *BMC Evol Biol* **8**: 92.

Naidoo J, Young D. 2012. Gene regulation systems for gene therapy applications in the central nervous system. *Neurol Res Int* **2012**: 595410.

Nakamura Y, Tagawa K, Oka T, Sasabe T, Ito H, Shiwaku H, La Spada AR, Okazawa H. 2012. Ataxin-7 associates with microtubules and stabilizes the cytoskeletal network. *Hum Mol Genet* **21**: 1099–1110.

Nguyen QD, Schachar RA, Nduaka CI, Sperling M, Basile AS, Klamerus KJ, Chi-Burris K, Yan E, Paggiarino DA, Rosenblatt I, et al. 2012. Dose-ranging evaluation of intravitreal siRNA PF-04523655 for diabetic macular edema (the DEGAS study). *Invest Ophthalmol Vis Sci* **53**: 7666–7674.

Nielsen TT, Marion I van, Hasholt L, Lundberg C. 2009. Neuron-specific RNA interference using lentiviral vectors. *J Gene Med* **11**: 559–569.

Nieto-Bona MP, Garcia-Segura LM, Torres-Aleman I. 1993. Orthograde transport and release of insulin-like growth factor I from the inferior olive to the cerebellum. *J Neurosci Res* **36**: 520–527.

Noma S, Ohya-Shimada W, Kanai M, Ueda K, Nakamura T, Funakoshi H. 2012. Overexpression of HGF attenuates the degeneration of Purkinje cells and Bergmann glia in a knockin mouse model of spinocerebellar ataxia type 7. *Neurosci Res* **73**: 115–121.

Nott A, Meislin SH, Moore MJ. 2003. A quantitative analysis of intron effects on mammalian gene expression. *RNA* **9**: 607–617.

O'Neill B, Millington-Ward S, O'Reilly M, Tuohy G, Kiang AS, Kenna PF, Humphries P, Farrar GJ. 2000. Ribozyme-based therapeutic approaches for autosomal dominant retinitis pigmentosa. *Invest Ophthalmol Vis Sci* **41**: 2863–2869.

O'Reilly M, Palfi A, Chadderton N, Millington-Ward S, Ader M, Cronin T, Tuohy T, Auricchio A, Hildinger M, Tivnan A, et al. 2007. RNA interference-mediated suppression and replacement of human rhodopsin in vivo. *Am J Hum Genet* **81**: 127–135.

Okamura K, Hagen JW, Duan H, Tyler DM, Lai EC. 2007. The mirtron pathway generates microRNA-class regulatory RNAs in Drosophila. *Cell* **130**: 89–100.

Packer AN, Xing Y, Harper SQ, Jones L, Davidson BL. 2008. The bifunctional microRNA miR-9/miR-9* regulates REST and CoREST and is downregulated in Huntington's disease. *J Neurosci* **28**: 14341–14346.

Palfi A, Ader M, Kiang A-S, Millington-Ward S, Clark G, O'Reilly M, McMahon HP, Kenna PF, Humphries P, Farrar GJ. 2006. RNAi-based suppression and replacement of rds-peripherin in retinal organotypic culture. *Hum Mutat* **27**: 260–268.

Palhan VB, Chen S, Peng G-H, Tjernberg A, Gamper AM, Fan Y, Chait BT, La Spada AR, Roeder RG. 2005. Polyglutamine-expanded ataxin-7 inhibits STAGA histone acetyltransferase activity to produce retinal degeneration. *Proc Natl Acad Sci USA* **102**: 8472–8477.

Palmiter RD, Sandgren EP, Avarbock MR, Allen DD, Brinster RL. 1991. Heterologous introns can enhance expression of transgenes in mice. *Proc Natl Acad Sci USA* **88**: 478–482.

Pandya-Jones A. 2011. Pre-mRNA splicing during transcription in the mammalian system. *Wiley Interdiscip Rev RNA* **2**: 700–717.

Park I-H, Zhao R, West JA, Yabuuchi A, Huo H, Ince TA, Lerou PH, Lensch MW, Daley GQ. 2008. Reprogramming of human somatic cells to pluripotency with defined factors. *Nature* **451**: 141–146.

Park J-E, Heo I, Tian Y, Simanshu DK, Chang H, Jee D, Patel DJ, Kim VN. 2011. Dicer recognizes the 5' end of RNA for efficient and accurate processing. *Nature* **475**: 201–205.

Paulson HL. 2009. The spinocerebellar ataxias. *J Neuroophthalmol* **29**: 227–237.

Pearson CE, Nichol Edamura K, Cleary JD. 2005. Repeat instability: mechanisms of dynamic mutations. *Nat Rev Genet* **6**: 729–742.

Pertea M, Lin X, Salzberg SL. 2001. GeneSplicer: a new computational method for splice site prediction. *Nucleic Acids Res* **29**: 1185–1190.

Petrs-Silva H, Yasumura D, Matthes MT, Lavail MM, Lewin AS, Hauswirth WW. 2012. Suppression of rds Expression by siRNA and Gene Replacement Strategies for Gene Therapy Using rAAV Vector. *Adv Exp Med Biol* **723**: 215–223.

Pfister EL, Kennington L, Straubhaar J, Wagh S, Liu W, DiFiglia M, Landwehrmeyer B, Vonsattel J-P, Zamore PD, Aronin N. 2009. Five siRNAs targeting three SNPs may provide therapy for three-quarters of Huntington's disease patients. *Curr Biol* **19**: 774–778.

Pihlmann M, Askou AL, Aagaard L, Bruun GH, Svalgaard JD, Holm-Nielsen MH, Dagnaes-Hansen F, Bek T, Mikkelsen JG, Jensen TG, et al. 2012. Adeno-associated virus-delivered polycistronic microRNA-clusters for knockdown of vascular endothelial growth factor in vivo. *J Gene Med* **14**: 328–338.

Rand TA, Petersen S, Du F, Wang X. 2005. Argonaute2 cleaves the anti-guide strand of siRNA during RISC activation. *Cell* **123**: 621–629.

Raper SE, Chirmule N, Lee FS, Wivel NA, Bagg A, Gao G, Wilson JM, Batshaw ML. 2003. Fatal systemic inflammatory response syndrome in a ornithine transcarbamylase deficient patient following adenoviral gene transfer. *Mol Genet Metab* **80**: 148–158.

Reynolds A, Leake D, Boese Q, Scaringe S, Marshall WS, Khvorova A. 2004. Rational siRNA design for RNA interference. *Nat Biotechnol* **22**: 326–330.

Rossmiller B, Mao H, Lewin AS. 2012. Gene therapy in animal models of autosomal dominant retinitis pigmentosa. *Mol Vis* **18**: 2479–2496.

Ruan K, Fang X, Ouyang G. 2009. MicroRNAs: novel regulators in the hallmarks of human cancer. *Cancer Lett* **285**: 116–126.

Ruby JG, Jan CH, Bartel DP. 2007a. Intronic microRNA precursors that bypass Drosha processing. *Nature* **448**: 83–86.

Ruby JG, Stark A, Johnston WK, Kellis M, Bartel DP, Lai EC. 2007b. Evolution, biogenesis, expression, and target predictions of a substantially expanded set of Drosophila microRNAs. *Genome Res* **17**: 1850–1864.

Samakoglu S, Lisowski L, Budak-Alpdogan T, Usachenko Y, Acuto S, Di Marzo R, Maggio A, Zhu P, Tisdale JF, Rivière I, et al. 2006. A genetic strategy to treat sickle cell anemia by coregulating globin transgene expression and RNA interference. *Nat Biotechnol* **24**: 89–94.

Sari Y. 2011. Huntington's Disease: From Mutant Huntingtin Protein to Neurotrophic Factor Therapy. *Int J Biomed Sci* **7**: 89–100.

Schallenberg M, Charalambous P, Thanos S. 2012. GM-CSF protects rat photoreceptors from death by activating the SRC-dependent signalling and elevating anti-apoptotic factors and neurotrophins. *Graefes Arch Clin Exp Ophthalmol* **250**: 699–712.

Schamberger A, Sarkadi B, Orban TI. 2012. Human mirtrons can express functional microRNAs simultaneously from both arms in a flanking exon-independent manner. *RNA Biol* **9**: 1177–1185.

Scherr M, Battmer K, Ganser A, Eder M. 2003. Modulation of gene expression by lentiviral-mediated delivery of small interfering RNA. *Cell Cycle* **2**: 251–257.

Scholefield J, Greenberg LJ, Weinberg MS, Arbuthnot PB, Abdelgany A, Wood MJA. 2009. Design of RNAi hairpins for mutation-specific silencing of ataxin-7 and correction of a SCA7 phenotype. *PLoS ONE* **4**: e7232.

Scholefield J, Wood MJA. 2010. Therapeutic gene silencing strategies for polyglutamine disorders. *Trends Genet* **26**: 29–38.

Seitz B, Baktanian E, Gordon EM, Anderson WF, LaBree L, McDonnell PJ. 1998. Retroviral vector-mediated gene transfer into keratocytes: in vitro effects of polybrene and protamine sulfate. *Graefes Arch Clin Exp Ophthalmol* **236**: 602–612.

Sellier C, Freyermuth F, Tabet R, Tran T, He F, Ruffenach F, Alunni V, Moine H, Thibault C, Page A, et al. 2013. Sequestration of DROSHA and DGCR8 by Expanded CGG RNA Repeats Alters MicroRNA Processing in Fragile X-Associated Tremor/Ataxia Syndrome. *Cell Rep* **3**: 869–880.

Senzer N, Barve M, Kuhn J, Melnyk A, Beitsch P, Lazar M, Lifshitz S, Magee M, Oh J, Mill SW, et al. 2012. Phase I trial of "bi-shRNAi(furin)/GMCSF DNA/autologous tumor cell" vaccine (FANG) in advanced cancer. *Mol Ther* **20**: 679–686.

Seo H, Sonntag K-C, Isacson O. 2004. Generalized brain and skin proteasome inhibition in Huntington's disease. *Ann Neurol* **56**: 319–328.

Seow Y, Sibley CR, Wood MJA. 2012. Artificial mirtron-mediated gene knockdown: functional DMPK silencing in mammalian cells. *RNA* **18**: 1328–1337.

Shah JK, Garner HR, White MA, Shames DS, Minna JD. 2007. sIR: siRNA Information Resource, a web-based tool for siRNA sequence design and analysis and an open access siRNA database. *BMC Bioinformatics* **8**: 178.

Sibley CR, Seow Y, Curtis H, Weinberg MS, Wood MJA. 2012. Silencing of Parkinson's disease-associated genes with artificial mirtron mimics of miR-1224. *Nucleic Acids Res* **40**: 9863–9875.

Sibley CR, Seow Y, Saayman S, Dijkstra KK, El Andaloussi S, Weinberg MS, Wood MJA. 2011. The biogenesis and characterization of mammalian microRNAs of mirtron origin. *Nucleic Acids Research* **40**: 438–448.

Sibley CR, Seow Y, Wood MJA. 2010. Novel RNA-based strategies for therapeutic gene silencing. *Mol Ther* **18**: 466–476.

Sieving PA, Caruso RC, Tao W, Coleman HR, Thompson DJS, Fullmer KR, Bush RA. 2006. Ciliary neurotrophic factor (CNTF) for human retinal degeneration: Phase I trial of CNTF delivered by encapsulated cell intraocular implants. *PNAS* **103**: 3896–3901.

Sinn PL, Sauter SL, McCray PB Jr. 2005. Gene therapy progress and prospects: development of improved lentiviral and retroviral vectors--design, biosafety, and production. *Gene Ther* **12**: 1089–1098.

Smibert P, Bejarano F, Wang D, Garaulet DL, Yang J-S, Martin R, Bortolamiol-Becet D, Robine N, Hiesinger PR, Lai EC. 2011. A Drosophila genetic screen yields allelic series of core microRNA biogenesis factors and reveals post-developmental roles for microRNAs. *RNA* **17**: 1997–2010.

Sopher BL, Ladd PD, Pineda VV, Libby RT, Sunkin SM, Hurley JB, Thienes CP, Gaasterland T, Filippova GN, La Spada AR. 2011. CTCF regulates ataxin-7 expression through promotion of a convergently transcribed, antisense noncoding RNA. *Neuron* **70**: 1071–1084.

Soutschek J, Akinc A, Bramlage B, Charisse K, Constien R, Donoghue M, Elbashir S, Geick A, Hadwiger P, Harborth J, et al. 2004. Therapeutic silencing of an endogenous gene by systemic administration of modified siRNAs. *Nature* **432**: 173–178.

La Spada AR, Fu YH, Sopher BL, Libby RT, Wang X, Li LY, Einum DD, Huang J, Possin DE, Smith AC, et al. 2001. Polyglutamine-expanded ataxin-7 antagonizes CRX function and induces cone-rod dystrophy in a mouse model of SCA7. *Neuron* **31**: 913–927.

La Spada AR, Wilson EM, Lubahn DB, Harding AE, Fischbeck KH. 1991. Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. *Nature* **352**: 77–79.

Stegmeier F, Hu G, Rickles RJ, Hannon GJ, Elledge SJ. 2005. A lentiviral microRNA-based system for single-copy polymerase II-regulated RNA interference in mammalian cells. *Proc Natl Acad Sci USA* **102**: 13212–13217.

Ström A-L, Forsgren L, Holmberg M. 2005. A role for both wild-type and expanded ataxin-7 in transcriptional regulation. *Neurobiol Dis* **20**: 646–655.

Ström A-L, Jonasson J, Hart P, Brännström T, Forsgren L, Holmberg M. 2002. Cloning and expression analysis of the murine homolog of the spinocerebellar ataxia type 7 (SCA7) gene. *Gene* **285**: 91–99.

Strumberg D, Schultheis B, Traugott U, Vank C, Santel A, Keil O, Giese K, Kaufmann J, Drevs J. 2012. Phase I clinical development of Atu027, a siRNA formulation targeting PKN3 in patients with advanced solid tumors. *Int J Clin Pharmacol Ther* **50**: 76–78.

Takahashi J, Fujigasaki H, Zander C, El Hachimi KH, Stevanin G, Dürr A, Lebre A-S, Yvert G, Trottier Y, de Thé H, et al. 2002. Two populations of neuronal intranuclear inclusions in SCA7 differ in size and promyelocytic leukaemia protein content. *Brain* **125**: 1534–1543.

Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, Yamanaka S. 2007. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**: 861–872.

Takahashi-Fujigasaki J, Breidert T, Fujigasaki H, Duyckaerts C, Camonis JH, Brice A, Lebre A-S. 2011. Amyloid precursor-like protein 2 cleavage contributes to neuronal intranuclear inclusions and cytotoxicity in spinocerebellar ataxia-7 (SCA7). *Neurobiol Dis* **41**: 33–42.

Tam OH, Aravin AA, Stein P, Girard A, Murchison EP, Cheloufi S, Hodges E, Anger M, Sachidanandam R, Schultz RM, et al. 2008. Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. *Nature* **453**: 534–538.

Tao O, Shimazaki T, Okada Y, Naka H, Kohda K, Yuzaki M, Mizusawa H, Okano H. 2010. Efficient generation of mature cerebellar Purkinje cells from mouse embryonic stem cells. *J Neurosci Res* **88**: 234–247.

Taylor J, Grote SK, Xia J, Vandelft M, Graczyk J, Ellerby LM, La Spada AR, Truant R. 2006. Ataxin-7 can export from the nucleus via a conserved exportin-dependent signal. *J Biol Chem* **281**: 2730–2739.

Taylor S, Wakem M, Dijkman G, Alsarraj M, Nguyen M. 2010. A practical approach to RT-qPCR-Publishing data that conform to the MIQE guidelines. *Methods* **50**: S1–5.

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

Toyoshima K, Vogt PK. 1969. Enhancement and inhibition of avian sarcoma viruses by polycations and polyanions. *Virology* **38**: 414–426.

Trottier Y, Lutz Y, Stevanin G, Imbert G, Devys D, Cancel G, Saudou F, Weber C, David G, Tora L. 1995. Polyglutamine expansion as a pathological epitope in Huntington's disease and four dominant cerebellar ataxias. *Nature* **378**: 403–406.

Unwalla HJ, Li HT, Bahner I, Li MR, Kohn D, Rossi JJ. 2006. Novel Pol II fusion promoter directs human immunodeficiency virus type 1-inducible coexpression of a short hairpin RNA and protein. *J Virol* **80**: 1863–1873.

Veedu RN, Wengel J. 2009. Locked nucleic acid as a novel class of therapeutic agents. *RNA Biol* **6**: 321–323.

Wang H-L, Yeh T-H, Chou A-H, Kuo Y-L, Luo L-J, He C-Y, Huang P-C, Li AH. 2006. Polyglutamine-expanded ataxin-7 activates mitochondrial apoptotic pathway of cerebellar neurons by upregulating Bax and downregulating Bcl-x(L). *Cell Signal* **18**: 541–552.

Wang W-X, Rajeev BW, Stromberg AJ, Ren N, Tang G, Huang Q, Rigoutsos I, Nelson PT. 2008. The expression of microRNA miR-107 decreases early in Alzheimer's disease and may accelerate disease progression through regulation of beta-site amyloid precursor protein-cleaving enzyme 1. *J Neurosci* **28**: 1213–1223.

Watanabe T, Totoki Y, Toyoda A, Kaneda M, Kuramochi-Miyagawa S, Obata Y, Chiba H, Kohara Y, Kono T, Nakano T, et al. 2008. Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. *Nature* **453**: 539–543.

Watson LM. 2012. Novel cell models for the study of Spinocerebellar Ataxia type 7 pathogenesis and therapy in a South African patient cohort. PhD Thesis, University of Cape Town, South Africa.

Wei C, Salichos L, Wittgrove CM, Rokas A, Patton JG. 2012. Transcriptome-wide analysis of small RNA expression in early zebrafish development. *RNA* **18**: 915–929.

Wellington CL, Ellerby LM, Hackam AS, Margolis RL, Trifiro MA, Singaraja R, McCutcheon K, Salvesen GS, Propp SS, Bromm M, et al. 1998. Caspase cleavage of gene products associated with triplet expansion disorders generates truncated fragments containing the polyglutamine tract. *J Biol Chem* **273**: 9158–9167.

Wen R, Song Y, Kjellstrom S, Tanikawa A, Liu Y, Li Y, Zhao L, Bush RA, Laties AM, Sieving PA. 2006. Regulation of rod phototransduction machinery by ciliary neurotrophic factor. *J Neurosci* **26**: 13523–13530.

Wheeler TM, Leger AJ, Pandey SK, MacLeod AR, Nakamori M, Cheng SH, Wentworth BM, Bennett CF, Thornton CA. 2012. Targeting nuclear RNA for in vivo correction of myotonic dystrophy. *Nature* **488**: 111–115.

Wheeler TM, Sobczak K, Lueck JD, Osborne RJ, Lin X, Dirksen RT, Thornton CA. 2009. Reversal of RNA dominance by displacement of protein sequestered on triplet repeat RNA. *Science* **325**: 336–339.

Whitehead KA, Langer R, Anderson DG. 2009. Knocking down barriers: advances in siRNA delivery. *Nat Rev Drug Discov* **8**: 129–138.

Whitney A, Lim M, Kanabar D, 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**: 140–143.

Wu H, Ye C, Ramirez D, Manjunath N. 2009a. Alternative processing of primary microRNA transcripts by Drosha generates 5' end variation of mature microRNA. *PLoS ONE* **4**: e7566.

Wu H, Ye C, Ramirez D, Manjunath N. 2009b. Alternative Processing of Primary microRNA Transcripts by Drosha Generates 5' End Variation of Mature microRNA. *PLoS One* **4**: e1020.

Wu X, Ashe J, Bushara KO. 2011. Role of olivocerebellar system in timing without awareness. *Proc Natl Acad Sci U S A* **108**: 13818–13822.

Wu Z, Yang H, Colosi P. 2010. Effect of genome size on AAV vector packaging. *Mol Ther* **18**: 80–86.

Xia H, Mao Q, Eliason SL, Harper SQ, Martins IH, Orr HT, Paulson HL, Yang L, Kotin RM, Davidson BL. 2004. RNAi suppresses polyglutamine-induced neurodegeneration in a model of spinocerebellar ataxia. *Nat Med* **10**: 816–820.

Xia XG, Zhou HX, Zhou S, Yu YP, Wu R, Xu ZS. 2005. An RNAi strategy for treatment of amyotrophic lateral sclerosis caused by mutant Cu,Zn superoxide dismutase. *J Neurochem* **92**: 362–367.

Xie J, Xie Q, Zhang H, Ameres SL, Hung J-H, Su Q, He R, Mu X, Seher Ahmed S, Park S, et al. 2011. MicroRNA-regulated, Systemically Delivered rAAV9: A Step Closer to CNS-restricted Transgene Expression. *Mol Ther* **19**: 526–535.

Xu X-M, Yoo M-H, Carlson BA, Gladyshev VN, Hatfield DL. 2009. Simultaneous knockdown of the expression of two genes using multiple shRNAs and subsequent knock-in of their expression. *Nat Protoc* **4**: 1338–1348.

Yamamoto A, Lucas JJ, Hen R. 2000. Reversal of neuropathology and motor dysfunction in a conditional model of Huntington's disease. *Cell* **101**: 57–66.

Yang J-S, Maurin T, Robine N, Rasmussen KD, Jeffrey KL, Chandwani R, Papapetrou EP, Sadelain M, O'Carroll D, Lai EC. 2010. Conserved vertebrate mir-451 provides a platform for Dicer-independent, Ago2-mediated microRNA biogenesis. *Proc Natl Acad Sci USA* **107**: 15163–15168.

Yanicostas C, Barbieri E, Hibi M, Brice A, Stevanin G, Soussi-Yanicostas N. 2012. Requirement for Zebrafish Ataxin-7 in Differentiation of Photoreceptors and Cerebellar Neurons. *PLoS One* **7**: e50705.

Yefimova MG, Messaddeq N, Karam A, Jacquard C, Weber C, Jonet L, Wolfrum U, Jeanny J-C, Trottier Y. 2010. Polyglutamine toxicity induces rod photoreceptor division, morphological transformation or death in spinocerebellar ataxia 7 mouse retina. *Neurobiol Dis* **40**: 311–324.

Yewdell JW, Lacsina JR, Rechsteiner MC, Nicchitta CV. 2011. Out with the old, in with the new? Comparing methods for measuring protein degradation. *Cell Biol Int* **35**: 457–462.

Yoo SY, Pennesi ME, Weeber EJ, Xu B, Atkinson R, Chen S, Armstrong DL, Wu SM, Sweatt JD, Zoghbi HY. 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.

Young JE, Gouw L, Propp S, Sopher BL, Taylor J, Lin A, Hermel E, Logvinova A, Chen SF, Chen S, et al. 2007. Proteolytic cleavage of ataxin-7 by caspase-7 modulates cellular toxicity and transcriptional dysregulation. *J Biol Chem* **282**: 30150–30160.

Yu J, Vodyanik MA, Smuga-Otto K, Antosiewicz-Bourget J, Frane JL, Tian S, Nie J, Jonsdottir GA, Ruotti V, Stewart R, et al. 2007. Induced pluripotent stem cell lines derived from human somatic cells. *Science* **318**: 1917–1920.

Yu X, Ajayi A, Boga NR, Ström A-L. 2012. Differential Degradation of Full-length and Cleaved Ataxin-7 Fragments in a Novel Stable Inducible SCA7 Model. *J Mol Neurosci* **47**: 219–233.

Yvert G, Lindenberg KS, Devys D, Helmlinger D, Landwehrmeyer GB, Mandel JL. 2001. SCA7 mouse models show selective stabilization of mutant ataxin-7 and similar cellular responses in different neuronal cell types. *Hum Mol Genet* **10**: 1679–1692.

Yvert G, Lindenberg KS, Picaud S, Landwehrmeyer GB, Sahel JA, Mandel JL. 2000. Expanded polyglutamines induce neurodegeneration and trans-neuronal alterations in cerebellum and retina of SCA7 transgenic mice. *Hum Mol Genet* **9**: 2491–2506.

Zamore PD, Tuschl T, Sharp PA, Bartel DP. 2000. RNAi: Double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell* **101**: 25–33.

Zander C, Takahashi J, El Hachimi KH, Fujigasaki H, Albanese V, Lebre AS, Stevanin G, Duyckaerts C, 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**: 2569–2579.

Zeng Y, Cullen BR. 2004. Structural requirements for pre-microRNA binding and nuclear export by Exportin 5. *Nucleic Acids Res* **32**: 4776–4785.

Zeng Y, Yi R, Cullen BR. 2003. MicroRNAs and small interfering RNAs can inhibit mRNA expression by similar mechanisms. *Proc Natl Acad Sci USA* **100**: 9779–9784.

Zhao Y, Lang G, Ito S, Bonnet J, Metzger E, Sawatsubashi S, Suzuki E, Le Guezennec X, Stunnenberg HG, Krasnov A, et al. 2008. A TFTC/STAGA module mediates histone H2A and H2B deubiquitination, coactivates nuclear receptors, and counteracts heterochromatin silencing. *Mol Cell* **29**: 92–101.

Zhuang Y, Weiner AM. 1986. A compensatory base change in U1 snRNA suppresses a 5' splice site mutation. *Cell* **46**: 827–835.

Zuker M, Stiegler P. 1981. Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. *Nucleic Acids Res* **9**: 133–148.