October 2014

Identification and Quantification of *FXN* Antisense Transcript 1 (*FAST-1*) in Friedreich Ataxia

A Thesis Submitted for the Degree of Doctor of Philosophy by

Madhavi Sandi

Biosciences Division

Department of Life Sciences

College of Health Sciences and Life Sciences

Brunel University London, Uxbridge UB8 3PH

Declaration

I hereby declare that the research presented in this thesis is my own work, except where otherwise specified, and has not been submitted for any other degree.

Madhavi Sandi

Abstract

Friedreich ataxia (FRDA) is a lethal autosomal recessive neurodegenerative disorder caused by expanded GAA repeats in the *FXN* gene, resulting in local epigenetic changes and reduced expression of the mitochondrial protein frataxin. The disease is characterised by neurodegeneration of large sensory neurons of the dorsal root ganglia and spinocerebellar tracts. It has been recently reported that a novel frataxin antisense transcript, *FAST-1*, is overexpressed in FRDA patient derived fibroblasts. However, the lack of fundamental information about *FAST-1* gene such as size, sequence, length and its origin has hindered the understanding of its interactions with *FXN* gene. Therefore, I proposed to investigate these characteristics of *FAST-1* in a panel of FRDA cells and mouse models.

Firstly, using Northern blot hybridisation with small and large riboprobes, I identified two bands with different sizes (~500 bp and 9 kb size), representing potential *FAST-1* transcripts. Then to confirm the exact size and the location of the *FAST-1* gene, I performed 5'- and 3' RACE experiments, followed by cloning and sequencing. This analysis resulted in identification of the 5'- and 3'-ends of *FAST-1*, which mapped to nucleotide positions '-359' and '164' of the *FXN* gene, giving the total length of *FAST-1* as 523 bp size. Strikingly, the full-length 523 bp *FAST-1* transcript also corresponds to one of the Northern blotting results where I identified a band at approximately 500 bp size, indicating that the Northern blotting may have correctly identified the same full-length *FAST-1* transcript.

Subsequently, by optimising number of experimental parameters within our lab, I developed a robust qRT-PCR method to quantify *FAST-1* expression levels. Using this technique, I analysed the expression pattern of this antisense transcript in various FRDA cell

lines and mouse models. I confirmed the original finding of increased *FAST-1* levels in human FRDA fibroblasts, and further quantified *FAST-1* levels in FRDA mouse model cell lines and tissues. However, no consistently altered patterns of *FAST-1* expression were identified in relation to *FXN* expression. Therefore, either they are not directly connected, as originally reported by De Biase *et al.*, or their relationship varies between cell and tissue types.

Lastly, improved understanding of epigenetic changes in FRDA and growing evidence on long-gene regulation led me to study the 'neighbouring genes' rather than just focusing on the *FXN* gene. Therefore, I studied a region of approximately 750 kb on both sides of the *FXN* and quantified the expression levels of two genes (*PGM5* and *PIP5K1β*) on 5'- end and four genes (*TJP2, FAM189A2, APBA1* and *PTAR1*) on 3'- end of *FXN* gene in human primary fibroblasts. I found that *PGM5* and *PIP5K1β* genes, located at 5'- end of the *FXN* genes, were downregulated in FRDA fibroblasts and these findings coincide with the recent epigenetic changes identified in FRDA, where significant enrichment of gene repressive histone marks and increased DNA methylation were shown in upstream region of GAA repeats in intron 1 of the *FXN* gene. Out of four genes that were studied in the 3'- end of the *FXN* gene, only one gene (*APBA1*) was downregulated, which suggests that there are fewer repressive epigenetic marks downstream of the GAA repeat.

Acknowledgements

First and foremost I would like to express my deep and sincere gratitude to my supervisor Dr Mark Pook for accepting me as his student and giving me the opportunity to work on such an interesting research project. I would greatly appreciate for his excellent guidance, patience, motivation and providing me with an excellent atmosphere for doing research. He has always been extremely supportive and has provided me with critical help and guidance throughout my PhD. I would like to express my special thanks to my second supervisor, Dr Evegeny Makarov.

I would also like to express my special thanks to Dr Sahar Al-Mahdawi and Dr Chiranjeevi Sandi for their valuable and constructive suggestions during the planning and development of this research work. Their keen insight, high experience, and thoughtful reflection on the different aspect of research abilities are outstanding to which I aspire.

It is my pleasant duty to thank past and present members of 'Ataxia Research Group' at Brunel University London, especially Dr Sara Anjomani Virmouni, Dr Vahid Ezzatizadeh and Hassan Khonsari for their advice and assistance in keeping my research progress on schedule.

It is my pleasant honour to thank European Friedreich's Ataxia Consortium for Translational Studies (EFACTS) for providing financial support.

A huge thank you to my friends Dr Punam Bhullar, Dr Sheba Adam Zahir, Dr Maryam Ojani, Dr Gonul, Jessica, Hannah, Chetana, Yaghoub, Parisa, Savi, Becky, Ane and Hussain for being amazing colleagues; you certainly made this a more enjoyable experience.

I would also like to express my appreciation to Christine Newton and Alison Marriott for their incessant support, advice and assistance and all the other technical staff who have contributed to this work in many ways.

Finally, I would like to thank my parents, Rajeshwar and Suvarna and my brothers Shiva and Hari for providing support throughout my life. I thank you for never having anything but confidence on me and always supporting my dreams.

I would like to say a special “Thanks” to a special person in my life, my husband, Chiru, who has always been supporting, guiding, encouraging and enthusiastic about my work.

Last but not least, I would also like to say a ‘big thank you’ to both of my daughters, Sanjana and Meghana, for giving me a peaceful time for studying and writing this report.

Dedication

This thesis is dedicated to my parents, Rajeshwar and Suvarna, my husband, Chiranjeevi, and my children, Sanjana and Meghana. Every challenging work needs self-efforts as well as guidance of elders especially those who were very close to our heart. I give my deepest expression of love and appreciation for the encouragement that you gave me and the sacrifices you made during this PhD. I hope this work will be an inspiration for my children to achieve high in their lives.

Abbreviations

µg	microgram
µl	microlitre
AMV RT	avian myeloblastosis virus reverse transcriptase
AP2	Activating protein 2
APBA1	Amyloid beta A4 precursor protein-binding family A member 1
APP	Amyloid precursor protein
ATP	adenosine triphosphate
β-gal	β-galactosidase
β-ME	β-mercaptoethanol
BAC	bacterial artificial chromosome
bp	base pair
BSA	bovine serum albumin
CCDS	consensus coding sequence
ChIP	chromatin immunoprecipitation
CNS	central nervous system
CoQ	coenzyme Q
Ct	cycle threshold
CTCF	CCCTC-binding factor
cyaY	Ecoli frataxin orthologue
DEPC	diethyl pyrocarbonate
DM1	Myotonic dystrophy
DMEM	Dulbecco's modified Eagle's medium
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DNMT	DNA methyltransferases
DNMTase	DNA methyltransferases
dNTP	deoxynucleotide triphosphate
DRG	dorsal root ganglion
DTT	Dithiothreitol
E.coli	*Escherichia coli*
EDTA	ethylene diamine-tetra acetic acid
ENCODE	Encyclopedia of DNA Elements
EPO	erythropoietin
FAM189A2	Family with sequence similarity 189, member A2
FAST-1	Frataxin antisense transcript-1
FBS	fetal bovine serum
FCS	fetal calf serum
Fe-S	iron-sulphur
FISH	fluorescence in situ hybridization
FRAXA	Fragile X syndrome
FRDA	Friedreich ataxia
FXN	*Frataxin*

GAPDH	*Glyceraldehyde-3-phosphate dehydrogenase*
GSP	gene specific primer
GTP	guanosine triphosphate
HBA1	hemoglobin alpha 1
HCl	hydrochloric acid
HD	Huntington's disease
HDACi	histone deacetylase inhibitors
HMTase	histone methyltransferases
HP1	heterochromatin protein 1
HPRT	*Hypoxanthine guanine phosphoribosyl transferase*
IMS	industrial methylated spirit
ISC	iron sulphur cluster
kb	kilo base
kDa	kilodalton
kg	kilogram
LB medium	Luria-Bertani medium
MAGUK	membrane-associated guanylate kinases
mci	millicurie
MCK	muscle creatine kinase
MDB	membrane desalting buffer
mg	milligram
min	minutes
MitoQ	mitoquinone
ml	millilitre
MMLV	moloney murine leukemia virus
MOPS	3-(N-morpholino)propansulfonic acid
MPP	mitochondrial processing peptidase
mRNA	messenger RNA
NaCl	sodium chloride
NaOH	sodium hydroxide
NAT	natural antisense transcript
ng	nanogram
NGSP	nested gene specific primer
nm	nanometre
NSC	neural stem cell
NSE	neuron-specific enolase
PA	poly adenylation
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
pen-strep	penicillin and streptomycin
PEV	position effect variegation
PGM5	phosphoglucomutase 5
PIP5K1B	phosphatidylinositol-4-phosphate 5-kinase, type I, beta

PPAR	peroxisome proliferator-activated receptor
PRC2	polycomb Repressive Complex 2
PTAR1	protein prenyltransferase alpha subunit repeat containing 1
qRT-PCR	quantitative real-time RT-PCR
RACE	rapid amplification of cDNA ends
rhu-EPO	recombinant human erythropoietin
RNA	ribonucleic acid
RNA	ribonucleic acid
ROS	reactive oxygen species
rpm	revolutions per minute
rRNA	ribosomalRNA
RT-PCR	reverse transcriptase PCR
SCA	spinocerebellar ataxia
SDS	sodium dodecyl sulphate
sec	seconds
SOC	super-optimal broth with catabolite repression
SP-PCR	small-pool PCR
SRF	serum response factor
SSC	saline-sodium citrate
Taq	*Thermus aquaticus*
TBE	tris-borate-EDTA
TE	tris EDTA
TEMED	tetramethylethylenediamine
TFR	transferrin
TJP2	tight junction protein 2
TNR	trinucleotide repeat
Tris	tris(hydroxymethyl)aminomethane
TRS	trinucleotide/triplet repeats sequence
TSS	transcription start site
U	units
Ultrahyb	ultrasensitive hybridization buffer
UPM	10X Universal Primer A Mix
UTR	untranslated region
UV	ultra violet
WT	wild-type
XIST	X-inactivs specific transcript
YAC	yeast artificial chromosome
Yfh1	yeast frataxin homologue 1
ZO-2	zona occludens 2

Table of Contents

List of Figures

List of Tables

Chapter 1 - Introduction

Friedreich ataxia (FRDA) is an autosomal recessive degenerative disease that primarily affects the nervous system and the heart. Nicholaus Friedreich, who was a professor of medicine in Heidelberg in the second half of the 19th century, provided the original description as a "degenerative atrophy of the posterior columns of the spinal cord." FRDA is the most common inherited ataxia affecting approximately one in every 40,000 individuals (Delatycki *et al.* 2000). The carrier rate has been approximately at 1:90 (Pandolfo 2009). FRDA is a fatal, autosomal recessive neurodegenerative disease caused by homozygous expansion of GAA repeats within intron 1 of the *FXN* gene (Campuzano *et al.* 1996). Normal individuals have 5 to 30 GAA repeats, with a premutation range from 41 to 65 GAA repeats. However, FRDA patients have been found to contain from 66 to over 1000 GAA triplet repeats (Pandolfo 2002). The size of the expanded GAA triplet repeat sequence affects the age of onset and the magnitude of this disease. Furthermore, the expansion of the GAA repeats affects the expression of a mitochondrial protein, frataxin (Campuzano *et al.* 1997). Although the exact function of frataxin not known but it is believed to involved in iron-sulphur cluster biosynthesis (Pandolfo and Pastore 2009). Frataxin insufficiency leads to increased oxidative stress (Emond *et al.* 2000; Schulz *et al.* 2000), impaired ATP production (Lodi *et al.* 1999), and mitochondrial iron accumulation (Delatycki *et al.* 1999). The primary sites of FRDA pathology are the large sensory neurons of the dorsal root ganglia (DRG) and the dentate nucleus of the cerebellum (Koeppen *et al.* 2007; Koeppen *et al.* 2009).

Generally the first symptoms appear at childhood, but age of onset may vary from childhood to adult (Pandolfo 2002). Clinical features include progressive limb and gait ataxia, absent lower limb reflexes, extensor plantar responses, dysarthria (Harding 1981). There is also some pathological involvement of non-neuronal tissues, with cardiomyopathy

as a common secondary effect and risk of diabetes also found in 10% of FRDA patients (Schulz *et al.* 2009). The main sites of pathology include the DRG, posterior columns of the spinal cord, corticospinal tracts and cardiac muscle (Durr *et al.* 1996). Within 20 years after the first appearance of symptoms, affected individuals are confined to a wheelchair and most commonly die in early adulthood from the associated heart disease (Delatycki *et al.* 2000).

1.1 - Frataxin gene structure and function

The '*FXN*' (initially called *X25*) gene, a mutated gene responsible for FRDA, was originally mapped to chromosomal region 9q13 (Chamberlain *et al.* 1988) and is localized in the proximal long arm (Figure 1.1). More precisely, the frataxin gene is located on the long arm (q) of chromosome 9 between the Giemsa bands 13 and 21.1 from base pair 71,650,478 to base pair 71,715,093 (Ensembl 2006).

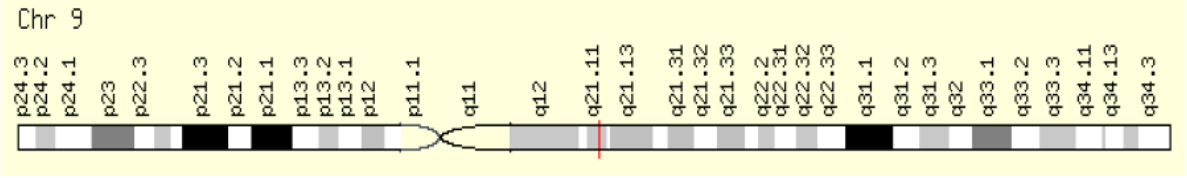

Figure 1.1 - The above image represents the location of *FXN* gene (red mark) on human chromosome 9 (Ensembl 2006).

The *FXN* gene spans 95 kb of genomic DNA and contains seven exons, 1-5a, 5b and 6 (Figure 1.2) (Campuzano *et al.* 1996). The gene is transcribed in the centromere to telomere direction. Generally the first five exons, 1 to 5a, transcribe and produce the major transcript

mRNA of 1.3kb size encoding a 210 amino acid mitochondrial protein called frataxin (Campuzano *et al.* 1996). Exon 5b is an alternatively spliced exon and exon 6 is non-coding. The *FXN* is expressed in every cell, although in varying levels in different tissues and during development (Campuzano *et al.* 1996; Koutnikova *et al.* 1997). *FXN* mRNA is predominantly expressed in tissues with a high rate of metabolism, including liver, heart, brown adipose tissue and skeletal muscles (Koutnikova *et al.* 1997). However, highest levels of *FXN* were identified in the spinal cord and somewhat lower levels in cerebellum. Frataxin expression is generally higher in mitochondria-rich cells, such as cardiomyocytes and neurons (Koutnikova *et al.* 1997).

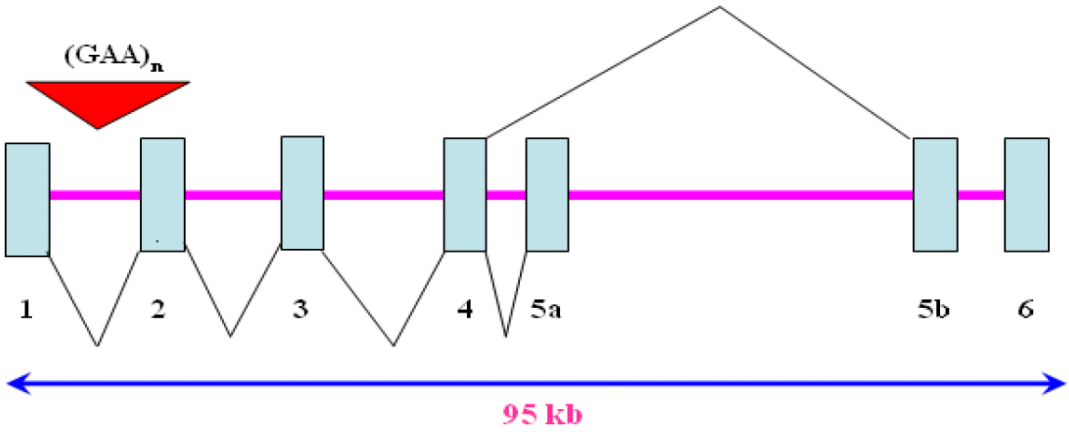

Figure 1.2 - Schematic representation of exons and splicing pattern of the *FXN* gene. The GAA repeat in intron 1 is indicated (▼). This diagram was adapted from (Cossee *et al.* 1997).

1.1.1 - Alternative splicing of the *FXN* gene

Although the exact function of frataxin is unclear at the present, it has been shown that frataxin is involved in the control of intracellular iron homeostasis. In addition, it is possible that frataxin may also have specific functions within each organism or cellular tissue. It has recently emerged that alternative splicing may have occurring in FRDA due to the identification of several *FXN* isoforms in addition to the canonical transcripts of *FXN*. Pianese *et* al. (2002) identified a new alternatively spliced form of *FXN* gene, called A1 isoform, expressed at a low level in different human tissues. The splicing occurred between positions 2 and 3 of the original serine codon (AGT), involving a frame shift with the appearance of a stop codon with in exon 5a resulting A1 isoform (Pianese *et al.* 2002). Xia *et* al. (2012) performed 5'-RACE with RNA isolated from HEK293 cells, in which a reverse primer located in exon 4 was used. This study has identified two novel tissue specific transcript variants, encoding two isoforms of frataxin that lack the mitochondrial targeting sequence and are therefore different from the canonical transcript (Xia *et al.* 2012). It is also believed that these transcripts may be generated through alternative splicing or exon skipping (Xia *et al.* 2012).

1.1.2 - Mutation in the *FXN* gene causes FRDA

The common mutation in FRDA is the large expansion of GAA trinucleotide repeat in the intron between the first and second exons of the frataxin gene. Majority of the FRDA patients (96%) are homozygous for GAA expansions (Campuzano *et al.* 1996), however, a small proportion of the patients (approximately 4%) are compound heterozygous, having

one allele with GAA repeats and the other allele with point mutations (Campuzano *et al.* 1996; Bidichandani *et al.* 1997).

1.1.3 - Instability of GAA expanded repeats

The FRDA-associated expanded alleles show both intergenerational and somatic (mitotic) instability.

Intergenerational instability

The GAA repeat sequence is unstable when transmitted from parent to offspring, with a non-pathogenic premutation of 34 repeats demonstrating expansion into pathogenic range (Montermini *et al.* 1997). Expanded alleles are equally likely to further expand or contract during maternal transmission, but most often contract during paternal transmission (Monros *et al.* 1997; Pianese *et al.* 1997).

Somatic instability

Expanded GAA triplet repeats show extensive instability in cultured cells, in the blood, central nervous system, dorsal root ganglia (DRG), spinal cord and the heart (Bidichandani *et al.* 1999; De Biase *et al.* 2007). Small-pool PCR (SP-PCR) experiments have shown that expanded GAA triplet repeats are very unstable in the peripheral leukocytes of patients (Sharma *et al.* 2002). The threshold expansion length for the initiation of somatic variability is between 26 and 44 uninterrupted GAA repeats (Sharma *et al.* 2002). Somatic instability starts after early embryonic development and continues after birth throughout life, resulting in progressive, age dependant accumulation of larger GAA triplet repeat

expansions, specifically in DRG (Figure 1.3) (De Biase *et al.* 2007). DRG degeneration is the primary cause of neurologic problems in FRDA patients.

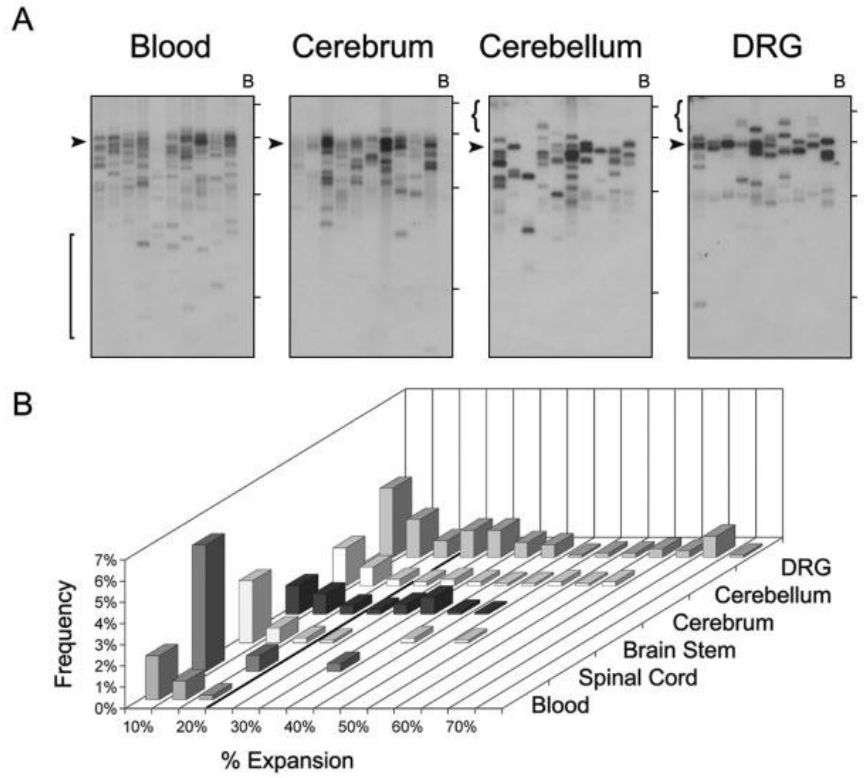

Figure 1.3 - Somatic instability in FRDA patient cells and tissues. The degree of somatic instability in DRG is more pronounced compared to other tissues (De Biase *et al.* 2007).

1.1.4 - Genotype-phenotype correlation

It is not well understood whether there is any correlation between the *FXN* gene (genotype) and the resulting physical trait or pattern of abnormalities (phenotype). However, it has been suggested that there is an inverse correlation between the size of the smaller expansion and both the age of onset and rate of disease progression (Durr *et al.* 1996; Delatycki *et al.* 1999). The two major complications of disease are cardiomyopathy and diabetes. Cardiomyopathy frequently arises in patients with large expansions in the smaller allele and is independent of the duration of this disease and diabetes does not appear to be associated with either the number of GAA repeats or the duration of disease (Durr *et al.* 1996; Delatycki *et al.* 1999; see review Santos *et al.* 2010), but develops during the later stages of disease (Filla *et al.* 1996). Other clinical manifestations, such as dysarthria, skeletal deformities, optic atrophy and hearing loss, show direct correlation to GAA expansion size (Durr *et al.* 1996; Montermini *et al.* 1997). Residual levels of frataxin vary according to the expansion and cell type. In peripheral blood leukocytes, frataxin levels in patients range from 5 to 30% of normal. The size of the smaller allele is inversely correlated to the amount of residual frataxin, providing a potential biochemical basis for the genotype-phenotype correlation with this allele (Gellera *et al.* 2007).

1.2 - Frataxin protein: structure

Frataxin is a small acidic protein highly conserved in most organisms from bacteria to mammalians and its expression is ubiquitous (Adinolfi *et al.* 2002). Frataxin in humans and Yfh1p in yeast exhibit sequence similarity and are functional homologs (Babcock *et al.* 1997; Wilson and Roof 1997). Analysis of crystal structure has revealed that mature frataxin is a compact, globular protein containing an N-terminal α-helix; seven β strands (β1-β5, β6 and β7) and a C-terminal end (Figure 1.4) (Dhe-Paganon *et al.* 2000; Musco *et al.* 2000).

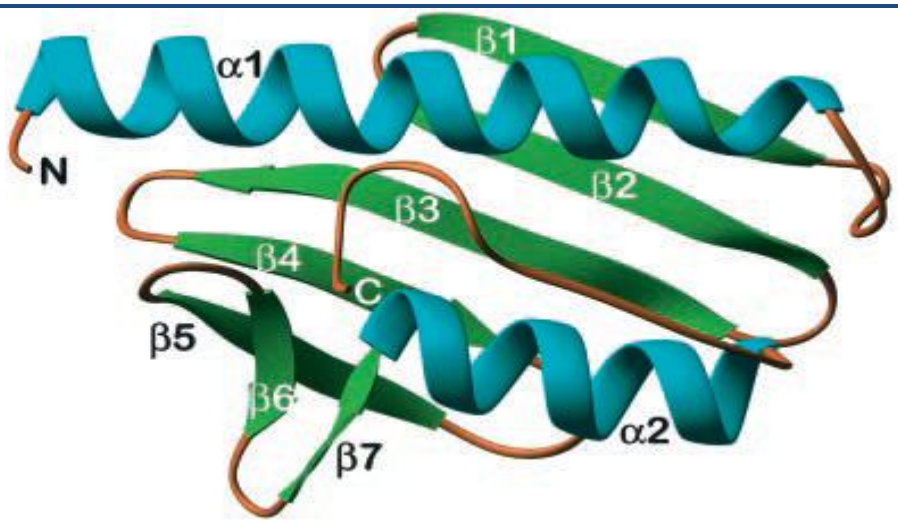

Figure 1.4 - Structure of frataxin. Ribbon diagram showing the fold of frataxin, a compact αβ sandwich, with helices coloured *turquoise* and β strands in *green* (Dhe-Paganon *et al.* 2000).

1.2.1 - Mitochondrial localization of human frataxin

Both targeting signal prediction and tissue pattern of expression initially suggested that mouse frataxin is a mitochondrial protein. To test whether human frataxin also contains a mitochondrial targeting sequence, the human sequence was tagged at the 3'end with β-galactosidase (β-gal) in a eukaryotic expression vector (Koutnikova *et al.* 1997). When transfected in HeLa cells, the fusion protein was appeared to localised in cytoplasmic granules consistent with mitochondrial localisation, while β-gal protein alone showed a diffused cytoplasmic staining. Confirmation of mitochondrial identity of the cytoplasmic granules was obtained by double indirect immune-cytofluorescence with β-gal and anti-mitochondrial protein antibodies (Koutnikova *et al.* 1997).

1.2.2 - Frataxin protein maturation

Frataxin is translated in cytoplasmic ribosomes and imported into the mitochondria as a large precursor, where the targeting sequence is proteolytically removed in a two-step process by the mitochondrial processing peptidase (MPP) enzyme to produce the mature frataxin protein (Schmucker *et al.* 2008). MPP first cleaves the precursor to give an intermediate form, followed by conversion of this product to mature form (Koutnikova *et al.* 1998; Gordon *et al.* 1999). The human frataxin is synthesized as a 210 amino acid precursor (23.1kDa MW), and processing *in vitro* may originate intermediate form of 169 amino acid size (18.8kDa MW), but only mature frataxin form of 130 amino acid size (14.3kDa MW) has functional significance *in vivo* (Condo *et al.* 2007; Schmucker *et al.* 2008).

1.2.3 - Frataxin protein function

Frataxin is a highly conserved mitochondrial protein from bacteria to humans. Genetic and biochemical studies support a role of frataxin as a multifunctional protein in different iron-dependent mitochondrial pathways (Pandolfo and Pastore 2009; Schmucker and Puccio 2010; Schmucker *et al.* 2011). Although the exact function of frataxin is still unknown, complete deficiency of frataxin in mice leads to early embryonic lethality (Cossee *et al.* 2000), indicating that frataxin is essential for early development. In eukaryotes, frataxin is necessary for normal mitochondrial function (Koutnikova *et al.* 1997; Pianese *et al.* 1997). Frataxin plays an important role in intracellular iron metabolism which includes formation of Fe-S clusters, biogenesis of heme (Yoon and Cowan 2004), iron binding/storage and iron chaperone activity (Cavadini *et al.* 2002) (Figure 1.5). Frataxin has a role in controlling cell survival as the frataxin-deficient cells are more sensitive to oxidative stress (Wong *et al.* 1999; Condo *et al.* 2007)**,** and the evidence of both apoptotic and autophagic cell death is found in frataxin-deficient animal models (Cossee *et al.* 2000; Simon *et al.* 2004). The investigation of the yeast frataxin homologue gene 1 (*Yfh1*) demonstrated that the frataxin protein was involved in iron efflux from mitochondria (Radisky *et al.* 1999). Adinolfi and colleagues, using *Escherichia coli* (*E.coli*) frataxin orthologue (CyaY), have shown that frataxin is not only act as an iron chaperone but also function as a molecular regulator to inhibit the formation of 2Fe-2S to keep the iron in a bio-available form (Adinolfi *et al.* 2009). Furthermore, accumulation of iron was consistently observed in autopsy of heart muscle (Bradley *et al.* 2000) and the dentate nucleus (Waldvogel *et al.* 1999; Koeppen *et al.* 2007) of FRDA patients. Frataxin is involved in controlling cellular oxidative stress by reducing the production of reactive oxygen species (ROS), protects the DNA against iron-induced oxidative damage (Pandolfo and Pastore 2009).

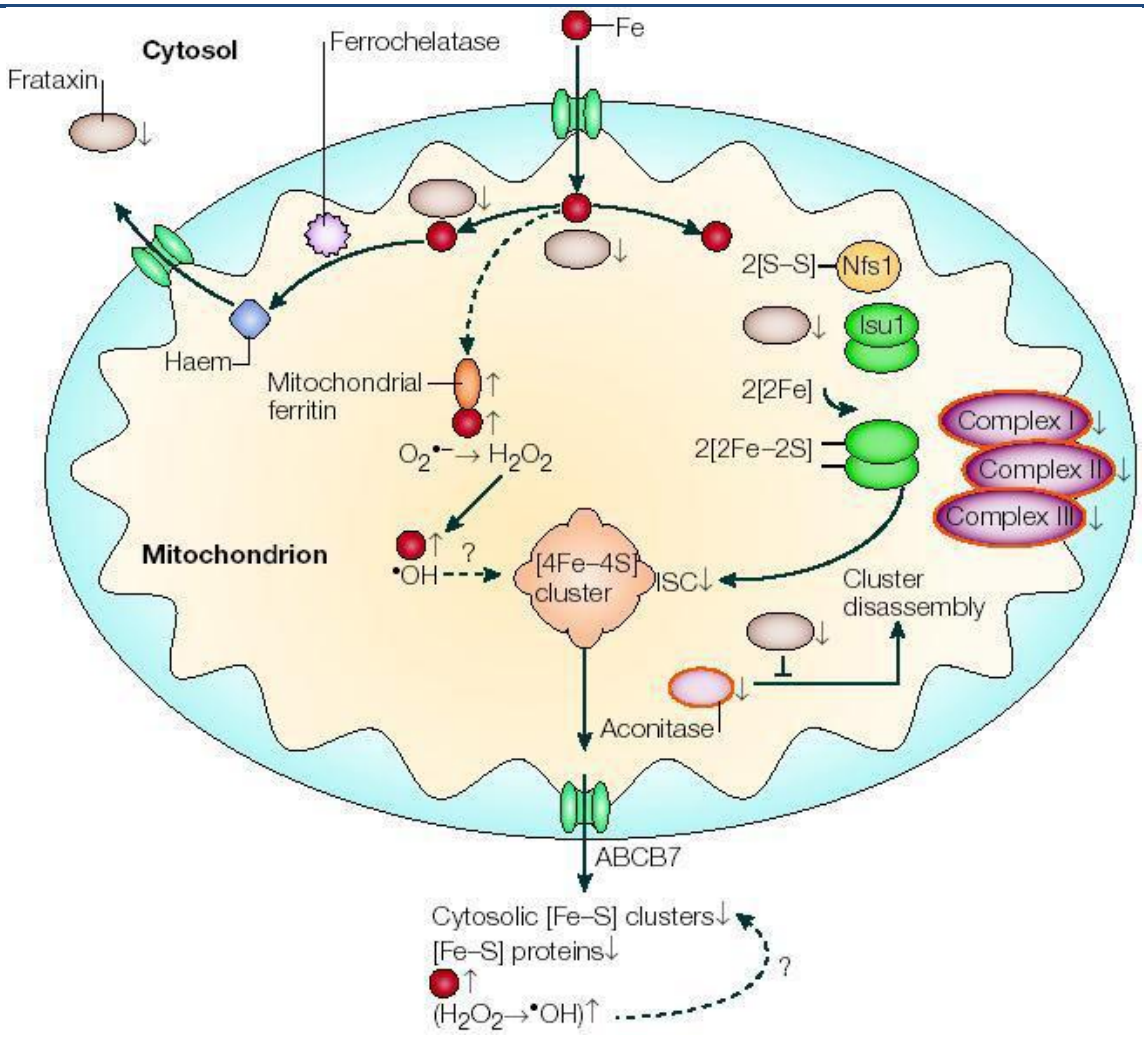

Figure 1.5 - The above image represents the alterations in mitochondrial biochemistry associated with reduced frataxin levels in FRDA (Gatchel and Zoghbi 2005).

1.3 - Physiopathological mechanisms of frataxin deficiency

The physiological consequences of frataxin deficiency are a severe disruption of iron-sulfur cluster (ISC) enzymes (Rotig *et al.* 1997; Puccio *et al.* 2001), mitochondrial iron overload coupled to cellular iron dysregulation (Babcock *et al.* 1997; Puccio *et al.* 2001) and increased sensitivity to oxidative stress (Babcock *et al.* 1997; Wong *et al.* 1999). The first evidence linking frataxin to iron metabolism was the identification of increased iron content in FRDA patient hearts (Bradley *et al.* 2000), in the dentate nucleus (Waldvogel *et al.* 1999; Koeppen *et al.* 2007) and in mitochondria of yeast strains with a deletion of the frataxin homologue (Yfh1) (Radisky *et al.* 1999). These studies suggest that frataxin plays a major role in regulating mitochondrial iron transport (Babcock *et al.* 1997). Furthermore, early biochemical studies in the endomyocardial biopsies of FRDA patients demonstrated deficiencies of the ISC-containing subunits of the mitochondrial electron transport complexes I,II and III and of iron-sulphur proteins aconitase (Rotig *et al.* 1997). Frataxin deficiency leads to primary mitochondrial and extramitochondrial ISC deficits (Rotig *et al.* 1997; Puccio *et al.* 2001). The role of oxidative stress in the pathogenesis of FRDA was discovered with the demonstration that frataxin-deficit in yeast and cultured cells from FRDA patients exhibits increased sensitivity to oxidative stress (Chantrel-Groussard *et al.* 2001; Jiralerspong *et al.* 2001). Numerous studies have demonstrated an impaired response of antioxidant enzymes in cell lines and model organisms (Wong *et al.* 1999; Chantrel-Groussard *et al.* 2001). A recent study presented data suggesting increase in nuclear and mitochondrial DNA damage in peripheral blood samples from FRDA patients (Haugen *et al.* 2010). However, it is not clear whether the increase in DNA damage is a consequence of an impaired antioxidant or directly linked to an ISC deficit as most of the damaged proteins are recognised as DNA repair proteins or ISC proteins (Rudolf *et al.* 2006).

1.4 - Molecular mechanisms of GAA repeat expansions and abnormal structures

FRDA, the most common form of ataxia, is caused by an expanded GAA repeat in intron 1 of the *FXN* gene. The implications of severely reduced *FXN* mRNA and protein have been demonstrated in FRDA patients' tissue samples and cultured cells (Campuzano *et al.* 1997). It has been shown by a number of studies that the reduced levels of *FXN* mRNA and protein are the result of inhibition of the *FXN* gene at the transcriptional level and not at the post transcriptional RNA processing level (Delatycki *et al.* 2000). Furthermore, the recent epigenetic changes occurring at the vicinity of the *FXN* gene is further implicated this hypothesis. The expanded GAA•TTC repeat sequence associated with FRDA adopts non-B DNA structures, such as triplex structures and formation of RNA.DNA hybrid (Grabczyk and Usdin 2000; Grabczyk *et al.* 2007) and sticky DNA structures (Mariappan *et al.* 1999; Sakamoto *et al.* 1999).

1.4.1 - Triplex structures

The GAA•TTC tract in FRDA is recognised as poly-purine.poly-pyrimidine (Pur.Pyr) sequence containing only purines (R) in one strand and pyrimidine (Y) on the other and it may adopt a number of unusual nucleic acid structures, including triple helices (Frank-Kamenetskii and Mirkin 1995). Triplexes in general may form R•R•Y or Y•R•Y, depending on whether the third strand is purine-rich or pyrimidine-rich, and can be formed as intermolecular structures or as folded intramolecular structures (Grabczyk and Usdin 2000). R•R•Y intramolecular triplex forms behind the RNA polymerase during transcription of a long GAA•TTC tract, trapping the polymerase and the movement of RNA polymerase along the template locally unpairs the DNA duplex and generates a wave of negative super coiling.

At transcription bubble, the polymerase covers the Y (TTC) template strand but the single stranded portion of GAA non-template strand is available to initiate triplex formation, promoting a formation of R•R•Y structures.

1.4.2 - RNA•DNA hybrid

Grabczyk *et al*. (2007) have recently reported the formation of a persistent RNA•DNA hybrid by transcription of the *FXN* gene and using a T7 RNA polymerase *in vitro*. The presence of RNA•DNA hybrids was observed using a GAA repeat number of 44 or 88 (see review Marmolino and Acquaviva 2009). Interestingly, during *in vitro* transcription of longer repeats, T7 RNA polymerase arrests in the promoter distal end of the GAA•TTC tract and an extensive RNA•DNA hybrid is tightly linked to this arrest (Grabczyk *et al.* 2007) (Figure 1.6). This has indicated that RNA•DNA hybrid formation appears to be one of the consequences of the long GAA•TTC repeats.

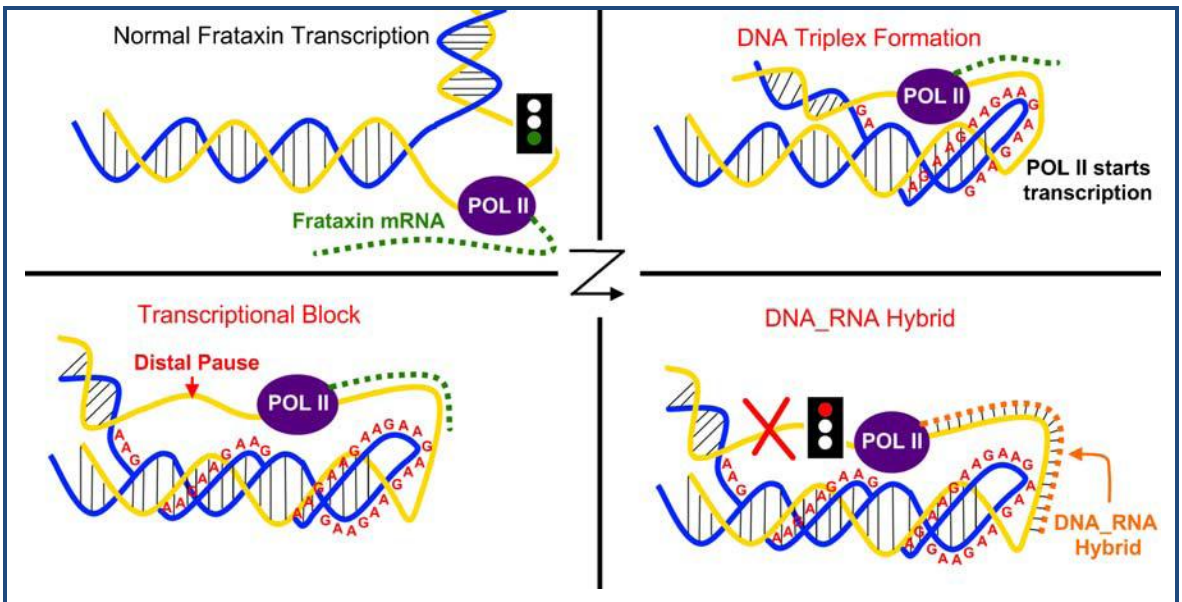

Figure 1.6 – GAA repeat mediated gene silencing through the formation of abnormal non B-DNA structures in FRDA (Grabczyk *et al.* 2007; Marmolino and Acquaviva 2009).

1.4.3 - Sticky DNA structures

Long GAA•TTC repeats from FRDA patients are reported to have highly tangled triple helical structure, called 'sticky DNA' structure. Son *et al.* (2006) have described that sticky DNA structures were formed by the association of two long GAA•TTC repeat sequences and this was a novel and unprecedented feature of the repeat (Son *et al.* 2006) (Figure 1.7). The fine structure of sticky DNA remains unknown, but it might be reminiscent of the composite triplex structure of the two distant homopurine-homopyrimidine runs (Christophe *et al.* 1985).

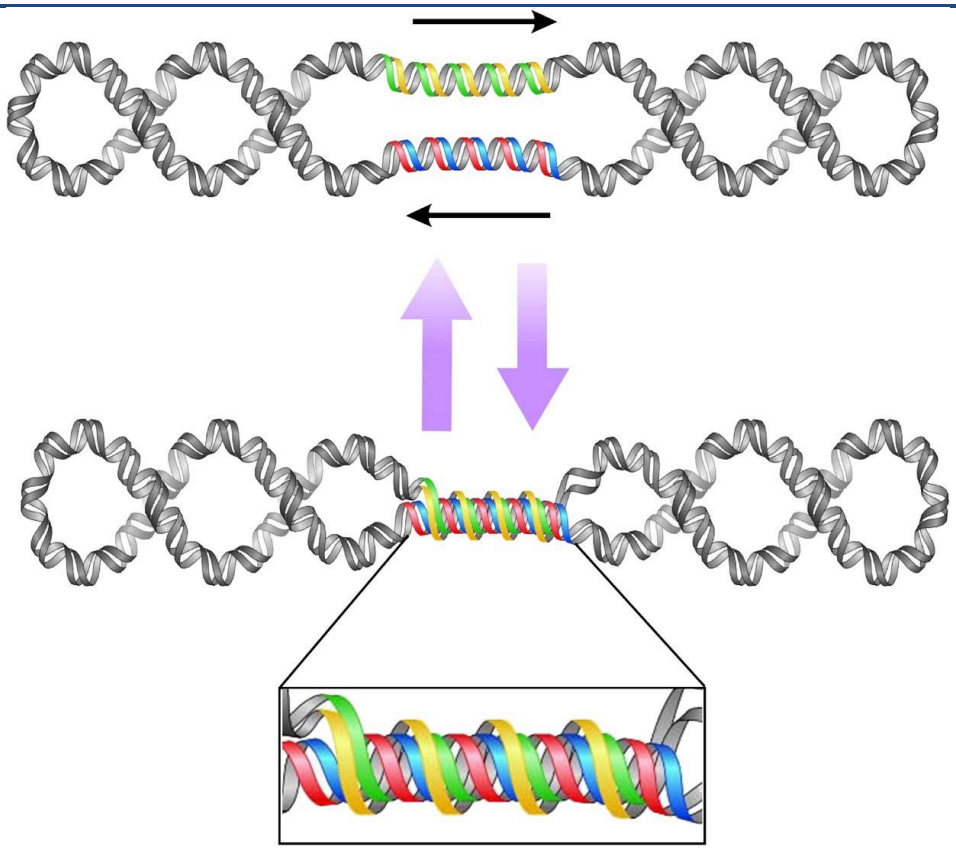

Figure 1.7 – The above model represents the formation of sticky DNA in a closed circular plasmid (Son *et al.* 2006).

1.5 - Therapeutic approaches

FRDA is a devastating neurodegenerative disorder with currently no systematic/specific therapy. Since the identification of the disease gene causing FRDA in 1996, identification of cellular function of frataxin has so far remained elusive. Therefore, it is a big challenge for researchers to develop a specific therapy for FRDA. However, significant amount of research has been initiated to tackle this life threatening disease and to develop a potential therapy for FRDA.

1.5.1 - Iron chelators

Iron chelation therapy overall aim is to maintain a 'safe' iron status at all times, ideally, to prevent iron accumulation and iron-related complications. Iron chelation is the current treatment of systemic iron overload diseases (Richardson *et al.* 2001). In FRDA, iron is in excess in mitochondria and is relatively depleted in the cytosol. Ideally, drugs that can able to redistribute the iron from overloaded regions to the depleted regions are clinically more effective rather than a total depletion of the iron from the body. In an attempt to initiate the iron chelation therapy for FRDA, desferoxamine was used but that resulted in a very poor membrane penetration and it chelates iron in the extracellular compartment and cytosol, promoting cellular iron depletion. Furthermore, desferoxamine has shown reduced Fe (II) toxicity on mitochondrial complex II in FRDA heart tissues, reduced levels of aconitase activities and frataxin (Rustin *et al.* 1999). To further identify the possible iron chelation therapy, a small phase I/II clinical trial has been carried out using orally-available iron chelator, deferiprone in individuals affected with FRDA (Boddaert *et al.* 2007). Deferiprone is shown to distribute in the CNS, crossing membranes, and remove excess iron from

mitochondria. Further, it has less tendency to cause overall iron depletion and may also redistribute iron between overloaded and depleted compartments (Marmolino 2011).

1.5.2 - Antioxidant therapy

Decreased mitochondrial respiratory chain function and increased oxidative stress have been implicated in the pathogenesis of FRDA. Therefore, there is raising interest that use of energy enhancement therapies or antioxidant treatments may be beneficial for FRDA. Among the antioxidant molecules, combination of coenzyme Q10 (CoQ_{10}) and vitamin E, Idebenone and mitoquinone (MitoQ), an analog and a derivative of CoQ_{10} respectively been widely tested (Marmolino and Acquaviva 2009). Treatment with a combination of CoQ_{10} and vitamin E improves mitochondrial function and reduced oxidative stress (Hart *et al.* 2005). Idebenone is a short-chain benzoquinone structurally related to coenzyme Q_{10}, which penetrates membranes and enters mitochondria much more efficiently than CoQ_{10}. It was first developed to treat Alzheimer's disease in Japan and therefore has extensive clinical safety. However, treatment of FRDA patients with Idebenone did not show any beneficial effect as neither the disease progression nor the neurological features of the patients did not affected as compared to placebo-treated FRDA patients (Marmolino 2011). MitoQ, a mitochondria-targeted antioxidant, was shown several hundred fold more potent than untargeted analog, idebenonoe, and it protects FRDA fibroblasts from endogenous oxidative stress (Jauslin *et al.* 2003). These results indicating that using mitochondria-targeted antioxidant compounds are more beneficial than general CoQ_{10} therapy in FRDA.

1.5.3 - HDAC inhibitors

Histone deacetylase (HDAC) inhibitors can affect transcription by increasing the acetylation of histones and transcription factors. In view of the recent epigenetic changes at *FXN* gene locus in FRDA, it has been proposed that the reversal of such epigenetic modifications could represent a useful therapeutic approach for FRDA (Figure 1.8).

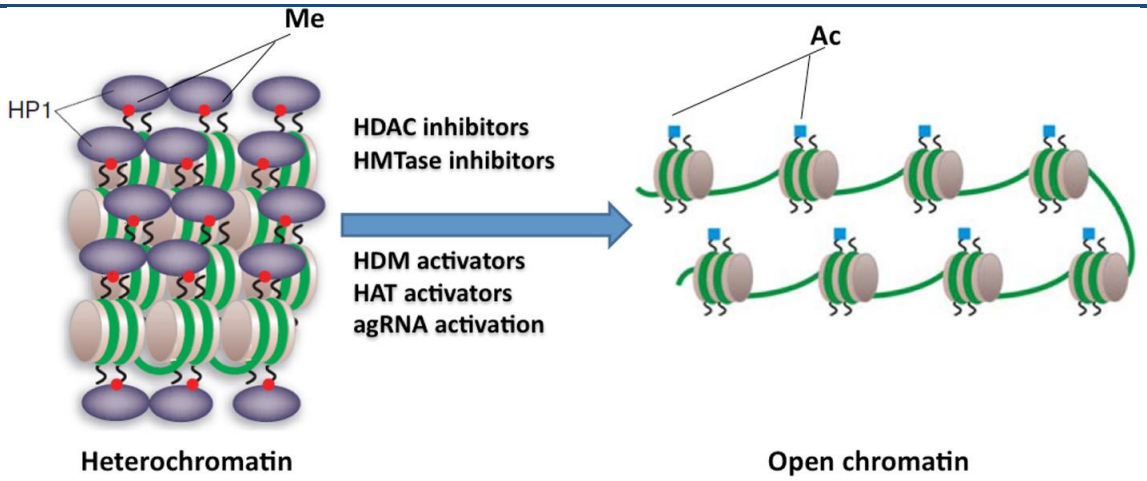

Figure 1.8 - Potential epigenetic based therapies for FRDA.

HDAC inhibitors revert silent heterochromatin to an active chromatin conformation with both positive and negative effects on gene expression (Di Prospero and Fischbeck 2005; Riessland *et al.* 2006). Initial treatment of FRDA lymphoblastoid cells using a panel of commercially available HDAC inhibitors revealed that only the benzamide compound BML-210 (N1-(2-aminophenyl)- N8-phenyloctanediamide) produced a significant increase of *FXN* mRNA expression (Herman *et al.* 2006). Later, the use of HDAC inhibitors has explored in different cell systems and mouse models of FRDA (Rai *et al.* 2008; Rai *et al.* 2010; Sandi *et al.* 2011). Recently, the class III HDAC inhibitor, nicotinamide has been shown to increase frataxin expression and decrease H3K9me3 and H3K27me3 at the *FXN* gene in FRDA cells and mouse models and this compound is now in early stage clinical trials (Chan *et al.* 2013;

Libri *et al.* 2014). However, the possibility of using HDAC inhibitors to restore the transcriptional deficit at the *FXN* gene in patients should take in to consideration their positive and negative effects on both activating and inactivating gene transcription (Marmolino 2011).

1.5.4 - Recombinant human erythropoietin (rhu-EPO)

Erythropoietin (EPO) is a 34 kDa glycoprotein hormone produced by yet unidentified cells in the peritubular renal interstitium in response to decreased renal tissue oxygen delivery and controls erythropoiesis. EPO is shown to cross the blood brain barrier and therefore it would suit to use as therapeutic compound to treat brain disorders. EPO has recently received considerable attention because of the unexpected finding that it also has broad neuroprotective and cardioprotective capabilities (Bogoyevitch 2004; Li *et al.* 2004). Recombinant human EPO (rhu-EPO) has been reported to increase frataxin protein levels in different FRDA cell lines, such as lymphocytes, cardiomyocytes and cardiac fibroblasts, and P19-derived neuronal-like cells (Sturm *et al.* 2005). A first clinical trial with rhu-EPO over a period of 6 months showed stable and significant increase of frataxin expression (Boesch *et al.* 2008) suggesting that rhu-EPO is a good candidate for FRDA therapy. The effect of EPO on frataxin expression is clearly posttranscriptional. However, the precise mechanism underlying this pathway and the possible cofactors involved are unknown (Marmolino 2011).

1.5.5 - Peroxisome proliferator-activated receptor (PPAR)- γ agonists

Peroxisome proliferator-activated receptor (PPAR) proteins are nuclear fatty acid proteins and are thought to play an important role on metabolic diseases such as obesity and insulin resistance. PPAR is further divided into three subtypes: PPARα, PPARδ/β and PPARγ. PPAR-γ plays a key role in numerous cellular functions including mitochondrial biogenesis and ROS metabolism (Wu *et al.* 1999; Kelly and Scarpulla 2004). PPAR-γ agonists, such as rosiglitazone and pioglitazone, are commonly used for the treatment of diabetes mellitus (Richter *et al.* 2007). In addition, PPAR-γ agonists are known to ameliorate number of others diseases including neurodegenerative disorders, especially Alzheimer's disease (Heneka and Landreth 2007). The treatment of FRDA human fibroblasts and SKNBE (human neuroblastoma) with azelaoyl PAF (APEF), a commercially available PPAR-γ agonist, has shown significant increase in frataxin mRNA and protein levels (Marmolino *et al.* 2009).

1.5.6 - Interferon gamma

Testi and colleagues have shown that interferon gamma, a cytokine involved in multiple aspects of iron metabolism and the immune response, upregulated the frataxin expression in human cells and FRDA mouse models (Tomassini *et al.* 2012). More strikingly, interferon gamma treatment also shows significant improvement in locomotor and motor coordination performances in FRDA mice (Tomassini *et al.* 2012).

1.6 - Mouse models

The development and use of animal models of FRDA are essential requirements to understand the physiological function of frataxin and FRDA pathogenesis and the investigation of potential FRDA therapeutic strategies.

1.6.1 - Knockout mouse models

To study the mechanism of the disease, Cossee *et al.* generated a mouse model by targeted deletion of *Fxn* exon 4. Heterozygous mice were normal but the homozygous deletions caused embryonic lethality a few days after implantation, demonstrating an important role for frataxin during embryonic development (Cossee *et al.* 2000). Using a conditional gene-targeting approach, Puccio and Colleagues have generated two lines of mouse models, muscle frataxin-deficient line and a neuron/cardiac muscle frataxin-deficient line (Puccio *et al.* 2001). These two lines together reproduce important progressive pathophysiological and biochemical features of the human disease (Puccio *et al.* 2001).

1.6.2 - Knock in mouse models

With the aim to generate an animal model showing a reduction in frataxin expression similar to FRDA patients, a frataxin knock in mouse was generated by introducing 230 GAA repeats into the mouse *frataxin* gene by homologous recombination (Miranda *et al.* 2002). The homozygous mutations led to a 25% reduction in the levels of frataxin. GAA repeat knock in mice were crossed with frataxin knockout mice to obtain double heterozygous mice expressing 25-36% of wild-type frataxin levels. These mice were viable

and failed to develop FRDA associated pathology (Miranda *et al.* 2002), indicating that longer repeats might be necessary to generate a good model of FRDA in mice.

1.6.3 - FRDA YAC transgenic mouse models

As a complementary approach in the generation of FRDA mouse models, Pook and colleagues generated and characterised transgenic mice that contain the entire *FXN* gene within a human YAC clone of 370 kb (Pook *et al.* 2001).

Generation of FRDA YAC transgenic mice

In an initial effort to assess whether human frataxin could function in a mouse background and substitute for loss of endogenous murine frataxin, a human wild-type FRDA yeast artificial chromosome (YAC) (Figure 1.9) transgenic mouse line was generated and crossbred this with heterozygous *Fxn* exon 4 deletion knockout mice (Pook *et al.* 2001). The resultant homozygous *Fxn* knockout offspring were phenotypically normal, indicating that the lack of endogenous mouse frataxin was rescued by expression of functional YAC-derived human frataxin. These results demonstrate that the re-introduction of human frataxin onto a mouse null background is an effective method, paving the way for future FRDA mouse models and frataxin functional studies.

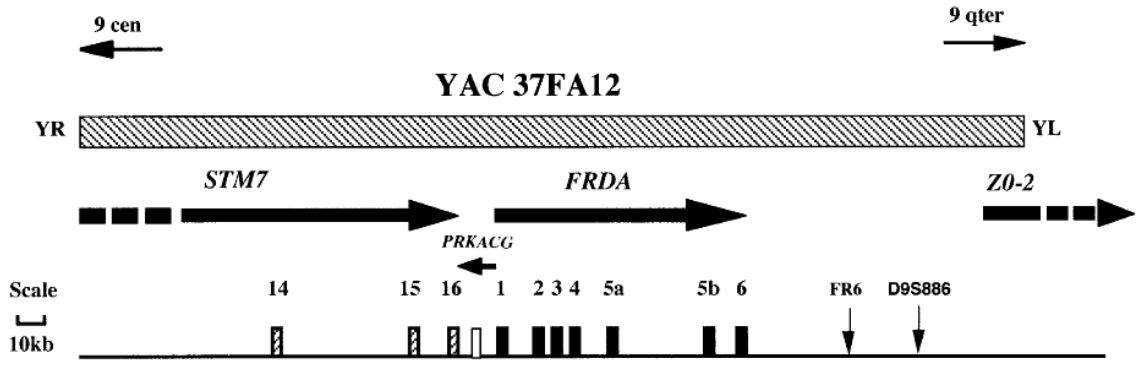

Figure 1.9 – The position of YAC 37FA12 with respect to the FRDA locus at 9q13 (Pook *et al.* 2001).

Generation of GAA containing FRDA YAC transgenic mice

Since the re-introduction of human frataxin onto a mouse null background rescued the embryonic lethality of the mouse, Pook and Colleagues have further generated two lines of mice, YG8 and YG22, that additionally contain GAA repeat expansions derived from FRDA patient DNA (Figure 1.10) (Al-Mahdawi *et al.* 2004). Both of these lines show intergenerational and somatic instability, although no large expansion was detected. However, the expanded FRDA YAC mouse models showed age related somatic instability, particularly expansions in the cerebellum, in a manner resembling to human FRDA condition. Therefore, these mice will be useful resources for further understanding of the FRDA GAA repeat expansion mechanisms within an *in vivo* mammalian system.

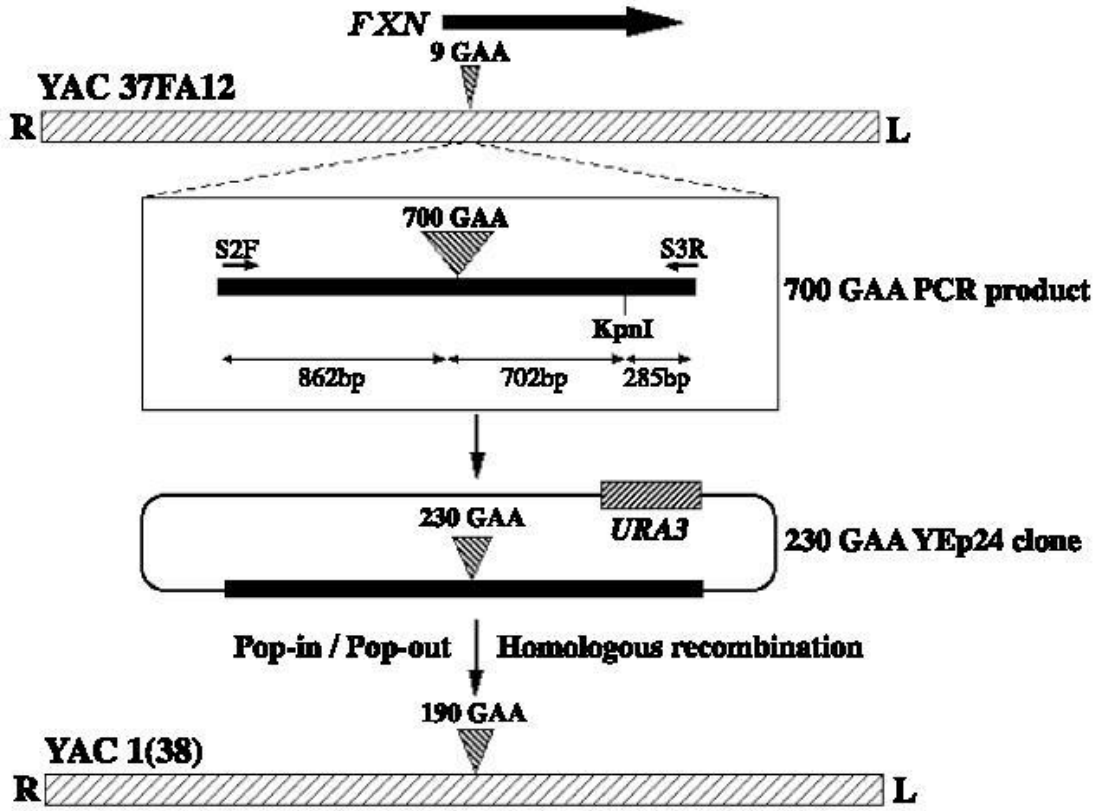

Figure 1.10 – Generation of GAA repeat expansion containing FRDA YAC transgenic mouse model by pop-in/pop-out homologous recombination strategy (Al-Mahdawi *et al.* 2004).

Both founder mice, YG8 and YG22 were shown to contain transgene sequences spanning the whole 370 kb human YAC clone, but differed in their GAA content as YG8 contained two copies of GAA repeats with 190 and 90 repeats (GAA190+90), whereas YG22 was shown to have only one copy of 190 GAA repeats (GAA190) (Table 1.1) (Al-Mahdawi *et al.* 2004). It has been further confirmed that the YG8 and YG22 rescue mice show reduced levels of human *FXN* mRNA in all tissues and decreased levels of human frataxin protein in some tissues, compared to Y47, a normal-sized GAA containing FRDA YAC transgenic mice (Al-Mahdawi *et al.* 2006; Al-Mahdawi *et al.* 2008). These mice represented the first GAA repeat expansion-based FRDA mouse models that exhibited progressive FRDA-like

pathology with GAA repeat instability and were thus of use in testing potential therapeutic strategies, particularly GAA repeat-based strategies (Al-Mahdawi *et al.* 2006).

Table 1.1 - General characterisation of *FXN* YAC transgenic mouse lines (Al-Mahdawi *et al.* 2004).

Transgenic line	YAC transgene integrity	*FXN* copy number	Founder GAA repeat length(s)	Range of GAA repeats in offspring
YG8	Complete	2	190 + 90	<9 to 223
YG22	Complete	1	190	<9 to 235

1.7 - Epigenetic changes in FRDA

Exactly how *FXN* transcriptional silencing is achieved in FRDA is not well understood, however recent evidence indicates that an epigenetic abnormality is an important underlying mechanism (Saveliev *et al.* 2003). Epigenetic mechanisms, including DNA methylation and hydroxymethylation, post-translational histone modifications, chromatin remodelling, and non-coding RNA effects, produce effects on gene expression without involving changes in the primary DNA sequence. Epigenetic mechanisms play a crucial role in silencing or activation of many genes during development. Epigenetic-based silencing has three main steps: Firstly, a decision-making process that targets specific silencing complexes to the DNA sequences to be inactivated. Secondly, a chromatin structuring process that results in efficient inhibition of RNA polymerases or other nuclear enzymes and finally, perhaps the most crucial step in epigenetic silencing is the propagation of the silent chromatin through DNA replication and mitosis to the daughter cells (Beisel and Paro 2011). Consistent with the later hypothesis, several FRDA disease-associated epigenetic changes

have been identified in the immediate vicinity of the expanded GAA repeats of the *FXN* gene. To further strengthen this hypothesis, Saveliev *et al.* (2003) showed that when expanded GAA repeats were introduced into the mouse genome, position-effect variegation (PEV) of an adjacent reporter gene was induced. PEV results when a gene normally in euchromatin is juxtaposed with heterochromatin by rearrangement or transposition. PEV is the hallmark of heterochromatin-mediated gene silencing and is thought to be caused by variable spreading of heterochromatin, which renders the gene inaccessible to the transcriptional machinery (Dillon and Festenstein 2002). Importantly, it has recently been reported that that the major *FXN* transcriptional start site, which is normally in a nucleosome-depleted region, is rendered inaccessible by altered nucleosome positioning in FRDA (Chutake *et al.* 2014).

1.7.1 - DNA methylation

The methylation of DNA is a process shared by both eukaryotic and prokaryotic cells, and it serves as an epigenetic method of modulating gene expression. Methylation plays a role in genomic stability and carcinogenesis, and it offers a target for the treatment of malignancy (Goffin and Eisenhauer 2002). The process of methylation is carried out by DNA methyltransferases (DNMT). These enzymes catalyze the covalent addition of a methyl group from a donor *S*-adenosyl methionine to the 5' position of cytosine, predominantly within the CpG dinucleotide (Robertson 2001). Initial investigations revealed increased DNA methylation of three specific CpG sites immediately upstream of the expanded GAA repeat sequence in FRDA patient derived lymphoblastoid cell lines and primary lymphocytes (Greene *et al.* 2007). Al-Mahdawi *et al*. studied FRDA patient autopsy brain, heart and cerebellum tissues, the most clinically relevant tissues in FRDA. This study revealed

significantly increased DNA methylation at the upstream region of GAA•TTC repeats (Al-Mahdawi *et al.* 2008). Similar DNA methylation changes were also identified in brain, heart and cerebellum tissues of two lines of FRDA YAC transgenic mice (YG8R and YG22R) (Al-Mahdawi *et al.* 2008).

1.7.2 - Histone modifications

The nucleosome consists of an octamer of four pairs of histone proteins, H2A, H2B, H3, and H4, surrounded by 147 bp of DNA (Kouzarides 2007). The core histones are globular except for their N-terminal ‘‘tails,’’ which are unstructured. A striking feature of histones tails is that they possess a large number of modified residues (Kouzarides 2007). Initial findings suggest that FRDA is caused by expanded GAA•TTC repeats, which trigger an abnormal structure (Sakamoto *et al.* 1999; Sakamoto *et al.* 2001). However recent studies have shown that FRDA may also be caused by increased levels of DNA methylation, decreased histone acetylation and increased histone methylation (Greene *et al.* 2007; Al-Mahdawi *et al.* 2008). Histone modifications at the *FXN* locus were first identified by the Gottesfeld lab, which reported lower levels of several acetylated H3 and H4 lysine residues, together with increased di- and trimethylation of H3K9 in the upstream GAA•TTC regions of FRDA lymphoblastoid cells (Herman *et al.* 2006). Usdin and colleagues then reported increased H3K9me2 levels within *FXN* intron 1 in FRDA lymphoblastoid cells (Greene *et al.* 2007) and our laboratory reported changes of histone modifications at the *FXN* promoter, upstream and downstream GAA regions in FRDA patients and YAC transgenic mouse tissues (Al-Mahdawi *et al.* 2008). De Biase *et al.* reported that FRDA patient fibroblasts have significantly higher levels of H3K27 trimethylation (H3K27me3) and H3K9 trimethylation

(H3K9me3) at the *FXN* 5'UTR region, coupled with elevated levels of heterochromatin proteins HP-1α and HP-1γ, compared to normal fibroblasts (De Biase *et al.* 2009). Histone modifications such as H3K4me3 and H3K79me3, both known to associate with open and active chromatin state of the gene, have been significantly reduced in the upstream and downstream GAA repeat regions of the *FXN* gene in FRDA cells (see review Sandi *et al.* 2014), indicating that there is a transcriptional elongation defect in FRDA cells. Further analysis of the epigenetic status of FRDA has revealed decreased levels of H3K4me3 in the upstream GAA repeat region of the *FXN* gene but not in the promoter region suggesting FRDA cells experiencing a more pronounced defects of the post-transcription and elongation stages of *FXN* gene expression rather than an early transcription initiation (Figure 1.11) (see review Sandi *et al.* 2014).

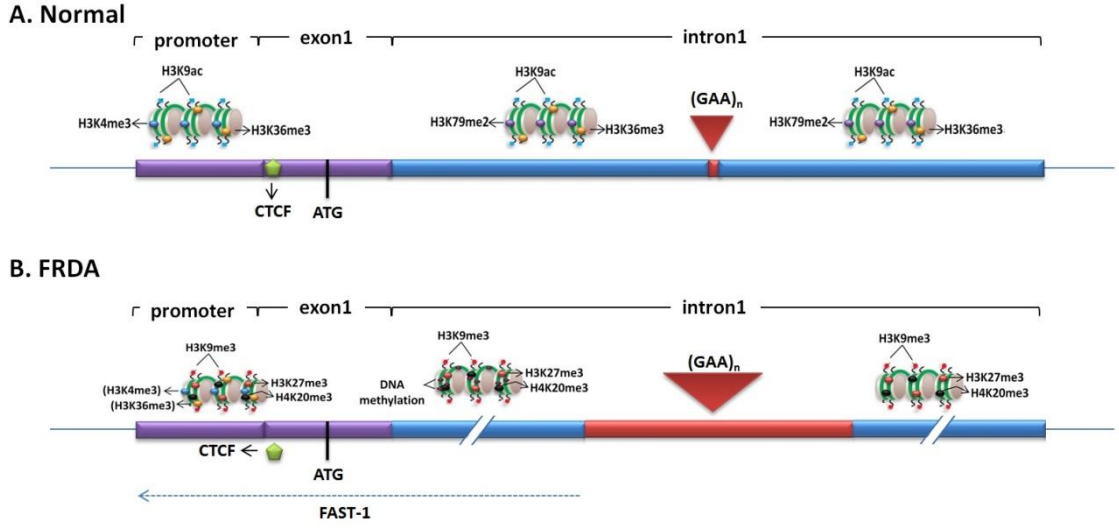

Figure 1.11 – The *FXN* chromatin organization in normal individuals and FRDA patients (Sandi *et al.* 2013).

1.7.3 - The Role of Antisense transcription and CTCF

Antisense transcription is a phenomenon where the antisense transcripts can be transcribed from the strand opposite to that of the sense transcript of either protein coding or non-protein coding genes. Recent studies have shown that antisense transcripts detected in various genes, including non-pathogenic alleles, known as natural antisense transcripts (NATs). In general, the level of antisense transcription is significantly lower than that of the coding sense transcripts. Antisense transcripts accumulate in the nucleus (Djebali *et al.* 2012). However, some antisense transcripts have been found to be associated with chromatin (Chu *et al.* 2011), and a few may occur in mitochondria and cytoplasm (Djebali *et al.* 2012). Antisense transcripts can arise from independent promoters, bidirectional promoters of divergent transcriptional units or cryptic promoters (Pelechano and Steinmetz 2013). The identification of antisense transcripts is technically challenging, as it requires strand specific approaches. Nevertheless, multiple reports have recently shown that antisense transcripts can either be involved in inhibition of the same gene from where they originate (*cis*-acting) or inhibition of genes at different locations *(trans*-acting) (Munroe and Lazar 1991; Morris *et al.* 2008; Chung *et al.* 2011). The fact that both antisense and sense transcripts originate from the same region suggests that antisense transcripts function more frequently in *cis* than other non-coding RNA that commonly function in *trans* (Guttman and Rinn 2012). Antisense transcripts or their transcription can affect all stages of gene expression including transcriptional initiation, co-transcriptional processes, and post-transcriptional processes. Antisense expression can affect transcription initiation through transcriptional interference, in which one act of transcription negatively affects a second one in cis. This occurs by promoter competition, by occlusion of binding sites due to the passage of RNA polymerase or even by chromatin or DNA modifications (Shearwin *et al.*

2005). Antisense expression can regulate transcription initiation by affecting DNA methylation, one example is the repression of the haemoglobin α1 gene (*HBA1*) in α-thalassaemia patients, where an aberrant *LUC7L* (putative RNA-binding protein Luc7-like) transcript runs antisense in to the *HBA1* locus and methylates its promoter CpG island, which silences *HBA1* expression and consequently causes disease (Tufarelli *et al.* 2003). Transcription initiation can also be controlled by affecting histone modifications, an example is mammalian X chromosome inactivation, in which the long ncRNA *XIST* spreads over one copy of the X chromosome and recruits repressive chromatin-remodelling complexes, such as Polycomb repressive complex 2(PRC2) (Pelechano and Steinmetz 2013). Antisense expression can regulate gene expression after transcription initiation by transcriptional interference that occurs co-transcriptionally, one example is the repression of the *IME4* locus (which encodes a key regulator of meiosis) in budding yeast by its antisense transcript regulator of meiosis 2 (*RME2*) (Gelfand *et al.* 2011). Antisense expression can also regulate post transcriptional life of a sense mRNA, as in the case of zinc finger E-box-binding homeobox 2 gene (*ZEB2*), which encodes a transcriptional repressor of E-cadherin, its antisense expression controls translation efficiency by preventing the processing of large 5'-intron that contain an internal ribosome entry site on the *ZEB2* mRNA. This does not change the abundance of the ZEB2 mRNA but increases its translation efficiency (Beltran *et al.* 2008).

Antisense transcripts also associated with microsatellite repeat expansion diseases such as Huntington's disease (HD) (Chung *et al.* 2011), Fragile X syndrome (FRAXA) (Ladd *et al.* 2007; Khalil *et al.* 2008) , Spinocerebellar Ataxia 7 (SCA7) (Sopher *et al.* 2011), Spinocerebellar Ataxia 8 (SCA8) (Moseley *et al.* 2006), and Myotonic dystrophy type 1 (DM1) (Cho *et al.* 2005; Yu *et al.* 2011). Therefore, the study of antisense transcription in gene

silencing machinery may provide further insight into the mechanisms of neurodegenerative disorders, including FRDA. To identify the presence of any antisense transcript in FRDA**,** De-Biase *et al*. (2009) performed strand specific reverse transcription PCR using a primer located upstream of *FXN* transcription start site 3 (TSS3), which revealed the significantly increased levels of frataxin antisense transcript 1 (*FAST-1*). In addition, it has been reported in FRDA patient derived primary fibroblasts that there is a severe depletion of CTCF (CCCTC-binding factor) in the 5' UTR region of the *FXN* gene (De Biase *et al.* 2009). The depletion of CTCF in FRDA patients also associated with increased *FAST-1* and heterochromatin formation (De Biase *et al.* 2009).

CTCF is a highly conserved 11-zinc finger (ZF) nuclear protein, originally recognised as a transcription factor that binds to avian and mammalian *MYC* promoters (Lobanenkov *et al.* 1990). CTCF is involved in a variety of transcription regulatory functions, including transcription activation, transcription repression and genomic imprinting (Phillips and Corces 2009). It has generally accepted that DNA methylation typically prevents the binding of CTCF. However, growing evidence has suggested that CTCF can also prevent the spreading of DNA methylation and thus maintains DNA methylation-free zones in the genome (Filippova *et al.* 2005; Engel *et al.* 2006). Furthermore, CTCF binding sites have been identified in the repeat expansion flanking regions of several TNR disorders, such as FRAXA (Ladd *et al.* 2007), DMI (Filippova *et al.* 2001), and SCA7 (Libby *et al.* 2008). The loss of CTCF binding at the DM1 CTG expansion is associated with the spread of heterochromatin and local CpG methylation (Cho *et al.* 2005). RT-PCR experiments indicated that *FAST-1* overlaps with the *FXN* gene, including at least portions of intron 1, exon 1, the 5'UTR, and extends further upstream into the nucleosome free region and upstream of *FXN* TSS3 (Figure 1.12) (De Biase *et al.* 2009). More recently, our laboratory has also reported increased expression

of *FAST-1* in FRDA mouse fibroblasts (Sandi *et al.* 2014) and reduced CTCF binding at the 5' UTR region of the *FXN* gene in FRDA cerebellum tissues (Al-Mahdawi *et al.* 2013). These findings support the hypothesis that CTCF prevents the spread of heterochromatin at *FXN* gene locus. However, further consideration of antisense transcription, CTCF binding, and other associated factors are needed, since they are likely to be highly relevant to the development of an epigenetic-based therapy for FRDA.

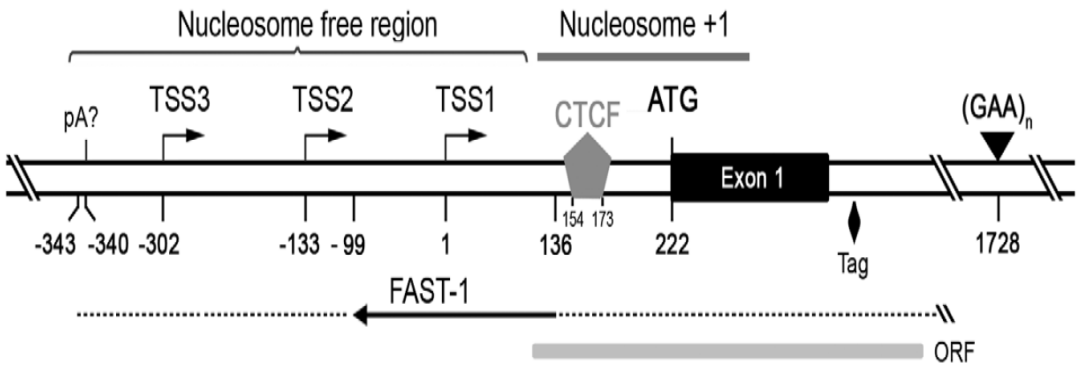

Figure 1.12 – The location and the possible extension of the *FAST-1* in FRDA (De Biase *et al.* 2009).

1.8 - Aims of the study

FRDA is a progressive neurodegenerative disorder with no effective therapy. Exactly how *FXN* transcriptional silencing is achieved in FRDA is not well understood, however recent evidence indicates that an epigenetic abnormality is an important underlying mechanism. Recently, De Biase *et al*. (2009) have reported that FRDA patients have high levels of *FAST-1* expression and increased heterochromatin protein 1 (HP1) binding at the *FXN* gene. They demonstrated the existence of *FAST-1* transcripts by RT-PCR experiments, indicating that the *FAST-1* gene overlaps with the *FXN* gene, extending from part of intron 1 to the upstream of *FXN* TSS3. Although they speculated that there might be a polyadenylation sequence located upstream of *FXN* TSS3, the exact details of the *FAST-1* gene structure remain unknown. Therefore, the first aim of this project was to further characterise the *FAST-1* gene structure, including 3'-end and 5'-ends, and to identify the total length of *FAST-1*.

Furthermore, since the initial study was limited only to the fibroblast cells, the second aim of this thesis was to develop a robust method to determine the expression levels of *FAST-1* in different cells model systems, FRDA mouse models and human FRDA patient biopsy tissues. As previously explained, there might be long-range GAA repeat-induced silencing effects on other genes adjacent to *FXN* (e.g. *PIP5K1B*). Therefore, the third aim of this project was to identify the effect of GAA repeat expansions on expression levels of flanking genes, including the *PIP5K1B* and *PGM5* genes at the 5'-end, and the *TJP2, FAM189A2, APBA1* and *PTAR1* genes at the 3'-end of the *FXN* locus (Figure 1). The specific aim was to investigate the mRNA expression levels of these *FXN* flanking genes in control and FRDA fibroblast cells.

Chapter 2 - Materials and Methods

2.1 - Solutions/reagents

General solutions

- TE buffer: 10mM Tris-HCl (pH 7.5), 1mM EDTA
- Orange G loading dye (6X): 0.35% Orange G dye, 30% sucrose
- 1X TBE: 90mM Tris, 90mM boric acid, 2mM EDTA
- 1X TAE: 40mM Tris, 20mM Acetic acid, 1mM EDTA
- Tris/glycerol homogenization buffer: 100mM Tris-HCl (pH 9.0),
- 15% glycerol (filter sterilized via 0.2µm pore)

Northern blot analysis

- 10X denaturing gel buffer (contains formaldehyde)
- 10X MOPS (morpholinopropanesulfonicacid) gel running buffer (200mM MOPS pH7.0, 80mM sodium acetate, and 10mM EDTA pH 8.0)
- ULTRAhyb (50% formamide)
- Transfer buffer (20X SSC - 3M NaCl, 0.3M trisodium citrate)
- Low stringency wash solution (2X SSC, 0.1% SDS)
- High stringency wash solution (0.1X SSC, 0.1% SDS)

DMEM medium

1X DMEM medium, 10% FCS, 2% pen-strep (5000U/ml penicillin and 5000mg/ml of streptomycin)

Fibroblast freezing medium

DMEM culture medium with 10% (v/v) DMSO

2.2 - Primers

Primers for genotyping and quantification of *FXN, FAST-1*

Various primers were used for genotyping the FRDA YAC transgenic mouse models (Table 2.1). Also, variety primers were used for the quantification of *FXN*, *FAST-1*, *FXN* flanking genes and endogenous control (Table 2.2 - Table 2.7). Primer sequences were obtained either from previous studies (as indicated) or newly designed by using primer 3 software (Rozen and Skaletsky 2000) and all primers were purchased from Sigma-Aldrich.

Table 2.1 - **Primers used for genotyping FRDA YAC transgenic mice**

Primers	Sequence (5'-3')	Size
GAA repeat (Campuzano et al. 1996)		
GAA - F	GGGATTGGTTGCCAGTGCTTAAAAGTTAG	457 bp + (GAA)n
GAA - R	GATCTAAGGACCATCATGGCCACACTTGCC	
FXN knockout (Cossee et al. 2000)		
WJ5	CTGTTTACCATGGCTGAGATCTC	
WN39 (WT specific)	CCAAGGATATAACAGACACCATT	520 bp
WC76 (KO specific)	CGCCTCCCCTACCCGGTAGAATTC	245 bp

Table 2.2 - **List of primers used for quantification of *FXN* expression**

Primers	Sequence (5'-3')	Size
FXN expression (Human Specific) (Al-Mahdawi et al. 2008)		
Fxn-h-RT-F	CAGAGGAAACGCTGGACTCT	172 bp
Fxn-h-RT-R	AGCCAGATTTGCTTGTTTGGC	
FXN expression (Human and Mouse)		
FRT I-F	TTGAAGACCTTGCAGACAAG	121 bp
RRT II-R	AGCCAGATTTGCTTGTTTGG	
Gapdh (Human) (Al-Mahdawi et al. 2008)		
Gapdh-h-F	GAAGGTGAAGGTCGGAGT	226 bp
Gapdh-h-R	GAAGATGGTGATGGGATTTC	

Gapdh (Mouse) (Al-Mahdawi et al. 2008)		
Gapdh-m-F	ACCCAGAAGACTGTGGATGG	81 bp
Gapdh-m-R	GGATGCAGGGATGATGTTCT	
HPRT (Human))		
HPRT-h-F1*	GGTGAAAAGGACCCCACGA	90 bp
HPRT-h-R1*	TCAAGGGCATATCCTACAACA	
HPRT (Mouse)		
HPRT-m-F1	ATGAAGGAGATGGGAGGCCA	80 bp
HPRT-m-R1	TCCAGCAGGTCAGCAAAGAA	

*primer sequences were obtained from Bayot *et al.* (2013).

Table 2.3 - **Primers for *FXN* sense riboprobe**

Primers	Sequence (5'-3')
FXN.f1	GGCCAAGCAGCCTCAATTTG
FXN.r1	GTAAATGCAACCGGGAGAACCAG
FXN-K-T7.f2	cGGTACCtaatacgactcactatagGGCCAAGCAGCCTCAATTTG
FXN.r2B	GcggaTCCTGTGGGGGAGCAGCTAGAGG
FXN.r3B	GcggaTCCAACCGGGAGAACCAGAGAAG

Table 2.4 - **Primers for *FXN* antisense riboprobe**

Primers	Sequence (5'-3')
FXN.f4	gcGGTACCtaatacgactcactataGGGGGAGCAGCTAGAGGTTAG
FXN.r4	GcggaTCCAAGCAGCCTCAATTTGTG
FXN.f5	GcGGTACCtaatacgactcactataGGGTGGCCCAAAGTTCCAG
FXN.r5	gcggaTCCTCTCGGGCGCCGCGCAG

FXN.f1, *FXN*.f4 and *FXN*.f5 are with *Kpn I* restriction site. *FXN*.r2B, *FXN*.r3B, *FXN*.r4 and *FXN*.r5 are with *Bam HI* restriction site. The sizes of different primer sets are listed below:

FXN.f1+*FXN*.r1 = 676 bp

FXN-K-T7.f2 + *FXN*.r2b = 238 bp

FXN-K-T7.f2 + *FXN*.r3b = 675 bp

FXN.f4+*FXN*.r4 = 231 bp

FXN.f5+*FXN*.r5 = 263 bp

Table 2.5 - **Primers for amplifying *FAST-1* transcript**

Primers	Sequence (5'-3')	Size
FAST RT*	CCAAGCAGCCTCAATTTGTG	
FAST F1*	GTGGGGGAGCAGCTAGAGG	207bp
FAST R1*	CACTTCCCAGCAAGACAGC	
N-FAST-F1	CAGCTAGAGGTTAGACCTCAG	183 bp
N-FAST-R1	CAGCAGCTCCCAAGTTCCTC	
N- FAST F2	GACCCAAGGGAGACTGCAG	88 bp
FAST R1*	CACTTCCCAGCAAGACAGC	

*primer sequences were obtained from De Biase *et al.* (2009).

Table 2.6 - **Primers for amplifying *FXN* flanking genes**

Primers	Sequence (5'-3')	Size
PGM5 expression		
PGM5-F	GGGGAGAGTTTGGAGTGAAG	128 bp
PGM5-R	GTCGATTCGGAGATCAGGAC	
PIP5K1B expression		
PIP5K-F1*	AGCAGCCTTGATGAAGAAGC	110 bp
PIP5K-R1*	GCAACCACAATTTCATCTTCTTC	
TJP2 expression		
TJP2-F	GAAAGGCTTAATCCCCAACA	153 bp
TJP2-R	AGGTCTTCCCGACTTTTCCT	
FAM189A2 expression		
FAM189A2-F1	CCAGGTGTGATCTGGTGGAC	127 bp
FAM189A2-R1	ACAGCACTACAGGGCTGAAC	
APBA1 expression		
APBA1-F	GTCCTGGGTATTGAGCAGGT	163 bp
APBA1-R	CAAGATGATCTGCCACGTCT	
PTAR1 expression		
PTAR1-F	TGCATAGATGCCCAATGTTT	152 bp
PTAR1-R	GAGATGGAGGTCTGTGGTGA	

*primer sequences were obtained from Bayot *et al.* (2013).

Table 2.7 - Primers for RACE (5'-RACE and 3'-RACE)

Primers	Sequence (5'-3')
SMARTer II A*	AAGCAGTGGTATCAACGCAGAGTACXXXXX
UPM-long *	CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGT
UPM-short*	CTAATACGACTCACTATAGGGC
3'-RACE PCR	
RACE-FAST-F1	GCCGCAGGCACTCTTCTGTGGGGGA
RACE-N-FAST-F1	GACCCAAGGGAGACTGCAGCCTGGTG
3'-RACE CDS Primer A*	AAGCAGTGGTATCAACGCAGAGTAC(T)30 V N
5'-RACE PCR	
RACE-FAST-R1	GCACCCACTTCCCAGCAAGACAGCAG
RACE-FAST-R2	GACCTCCAAGCTTTGCCTCCCTCAAG
RACE-N-FAST-R2	GACAGCAGCTCCCAAGTTCCTCCTG
RACE-N-FAST-R3	GCCACCAGGCTGCAGTCTCCCTTG
RACE-N-FAST-R6	GCCCAGGAGGCAGAGGTTGCAGTGAG
5'-RACE CDS Primer A*	(T)25V N

*primer sequences were obtained from SMARTer ™RACE cDNA amplification manual

2.3 - General techniques

Dilutions or stock solutions were prepared in deionised water (18.2 MΩ). The only exceptions were RNA experiments where RNase-free sterile water was used. Centrifugation of small samples (≤ 1.5ml) at room temperature was performed using a standard bench top micro-centrifuge (14K, Bio-Rad) and at 4°C using a refrigerated micro-centrifuge (5415R, Eppendorf). Incubations were performed either in water baths (Grant) or in a heating block (DB-2A, Techne). The pH of solutions was determined using a pH meter (Delta 340, Mettler) and pH adjustments were made by adding either concentrated HCl or NaOH. DNA samples were stored at 4°C for short time in the fridge and in the cold room for long term. RNA samples were stored at -80°C. All mouse tissues were stored at -80°C.

2.3.1 - Agarose gel electrophoresis

Agarose gel electrophoresis was used to separate DNA or RNA molecules by size. Nucleic acids are negatively charged and are moved through an agarose matrix by an electric field (electrophoresis). Shorter molecules move faster and migrate further. In our laboratory, I have made a range of 1% to 3% gels with agarose (Ultrapure electrophoresis grade, Invitrogen) for the separation in 1X TBE/TAE. The appropriate amount of agarose was weighed and boiled in standard microwave and allowed to cool on bench for few minutes. Then, ethidium bromide was added to the gel (0.2µg/ml) and swirled the melted agarose solution gently to mix with ethidium bromide followed by pouring the solution in to the casting gel tray with required comb. Once the PCR has finished, the resultant products were added with 6X orange G loading dye to a final concentration of 1X. The PCR products, along with a 1 kb plus DNA ladder (Invitrogen, Figure 2.1), were run on a gel for a period of 30min at 60V unless otherwise specified. The gels were visualized and documented by a UV transilluminator imaging system (Bio-Rad).

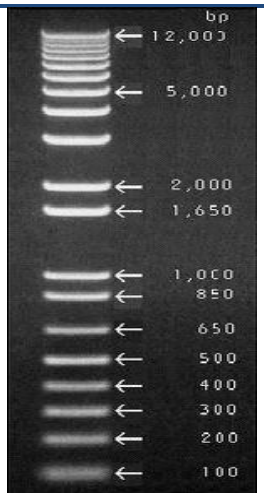

Figure 2.1 – The above image represents a typical 1kb plus DNA ladder run on 0.9% agarose gel and stained with ethidium bromide (Invitrogen).

2.3.2 - Denaturing urea polyacrylamide gel electrophoresis (Urea PAGE)

Polyacrylamide-urea gel electrophoresis provide high resolution of RNAs and are capable of resolving single stranded fragments of RNA that differ in length by as little as one nucleotide. The presence of the denaturant urea in the gel prevents the formation of secondary structures and ensures that the RNA molecules migrate through the gel as linear species. Polyacrylamide gels are run in the vertical position. For a 5% urea polyacrylamide gel of 20cm X 20cm, 2ml of 10X TBE, 9.6g of urea, 2.5ml of 40% (w/v) acrylamide (38%, w/v acrylamide; 2% w/v bisacrylamide) and water was added to make a final volume of 20 ml. The solution was stirred to dissolve with a baked stir bar. 125µl of 10% ammonium per sulphate and 16µl of TEMED (N,N,N',N'-Tetramethylethylenediamine) were added to the solution, which was then mixed briefly and immediately poured into the assembled gel plates with the help of a pipette. Cleaned gel plates were assembled with spacers and clamped by large binder clips on either side. The comb was immediately inserted and the gel was allowed to solidify. After the gel had polymerized, comb and binder clips were removed and the wells were rinsed with distilled water. Gel plates were placed in the lower buffer reservoir of the vertical electrophoresis apparatus. Plates were clipped into the apparatus by placing the metal plate over the front gel plate. 0.5X TBE buffer was added to the upper and lower reservoirs and wells were rinsed with 0.5X TBE buffer using a syringe to remove traces of urea and unpolymerised acrylamide from the wells. The gel apparatus was connected to the power supply and the gel was pre run for 20min at 700 volts. RNA samples were resuspended in formamide loading buffer. The RNA samples were heat denatured at 90°C for 1min, loaded into the bottom of the well carefully and run at 700 volts for 2hours until the bromophenol blue dye had migrated to the bottom of the gel. Gel apparatus was

disassembled and gel was visualized and documented by a UV transilluminator imaging system (Bio-Rad).

2.3.3 - DNA extraction - phenol/chloroform method

Genomic DNA extraction by the phenol/chloroform method was used for samples where greater quality was necessary. Tail samples or cell pellet were collected in eppendorf tubes and 400µl of tail digestion buffer and 10µl of proteinase K (50mg/ml) were added, followed by a brief vortex and incubation at 55°C overnight. After the incubation, samples were vortexed and 400µl of phenol (equilibrated with Tris-HCl pH 8.0) was added. Samples were mixed well by vortexing twice for 15sec and centrifuged at 14K rpm for 5min at 4°C. Then, 350µl of the supernatant was removed to a fresh eppendorf and 350µl of chloroform/isoamyl alcohol (24:1, v/v) was added. Samples were mixed briefly by inverting the eppendorf tubes and centrifuged again at 14K rpm for 10min at 4°C. Afterwards, 300µl of the resulting supernatant was removed to a fresh eppendorf and 30µl of 3M sodium-acetate (pH 5.2) was added. 800µl of absolute ethanol was then added and the samples were mixed by inverting the tube several times and centrifuged at 14K rpm for 10min at 4°C. The ethanol was drained off and the pellet washed with 1ml of 70% ethanol. The samples were again centrifuged at 14K rpm for 5min at 4°C and then ethanol was carefully drained off and the DNA pellet was air dried by inverting the eppendorf tube on paper towels for a period of approximately 10min. The DNA pellet was resuspended in 50-100µl of TE buffer and stored at 4°C.

2.3.4 - Extraction of total RNA

Total RNA was extracted from human fibroblast cells, mice fibroblasts cells, neural stem cells (NSCs), differentiated NSCs and tissues by following one of the three methods.

- Trizol® method (Invitrogen)
- NucleoSpin® RNA II method (Macherey Nagel)
- Tripure method (Roche)

Extraction of total RNA - Trizol® method (Invitrogen)

Total RNA was extracted from human fibroblast cells, mice fibroblasts cells, NSCs, differentiated NSCs and tissues using the Trizol® method following supplier guidelines. About 10^6 cells were used to extract the RNA. The Trizol® was added to the cell pellet or directly to the cells which are attached to the culture flask. The cell pellet was first washed once with PBS and collected by centrifugation at 1.5K rpm for 5min. The cell pellet was loosened by flicking the tube gently and resuspended in 1ml of Trizol®. For the extraction of RNA from tissues (20-30mg), initially 500µl of Trizol® was added to the tissue sample in an eppendorf tube and homogenised by a blue arrowhead tissue homogenizer tips, followed by adding remaining 500µl of Trizol® and then sample was mixed by inverting the tube. Samples were incubated for 10min at room temperature. Then 0.2ml of chloroform per 1ml of Trizol® was added, followed by vigorous shaking of samples for 15sec and incubation for a further 15min at room temperature. Samples were phase separated by centrifugation at 14K rpm for 15min at 4°C. The upper aqueous phase (~0.5ml) was then transferred to a freshly-labelled eppendorf tube and RNA was precipitated by adding 0.5ml of isopropyl alcohol. Samples were incubated for 10min at room temperature and centrifuged at 14K

rpm for 15min at 4°C. The supernatant was carefully removed and the RNA pellet was washed once with 1ml of 75% ethanol and centrifuged again at 8K rpm for 5min at 4°C. The supernatant was removed carefully and the RNA pellet was briefly dried for a period of 5-10min and resuspended in 30µl of RNase free water. RNA samples were stored at -80°C, until required to use.

Extraction of total RNA- NucleoSpin® RNA II method (Macherey Nagel)

Total RNA was extracted from human fibroblast cells using the NucleoSpin® RNA II total RNA isolation kit following supplier guidelines. 10^6 cells were used to extract the RNA. The cell pellet was washed once with PBS and collected by centrifugation at 1.5K rpm for 5min. Cells were lysed by adding 350µl buffer RA1 and 3.5µl β-mercaptoethanol (β-ME) to the cell pellet and vortexing vigorously. Then the cell suspension was transferred to the NucleoSpin® filter (violet ring) placed in a collection tube and centrifuged for 1min at 14K rpm. Afterwards, the NucleoSpin® filter was discarded, 350µl of 70% ethanol was added to the lysate and resuspended by pipetting up and down. Lysate was loaded to the NucleoSpin® RNA II column (light blue ring) placed in a collection tube followed by centrifugation for 30sec at 14K rpm. NucleoSpin® RNA II column was transferred to a new collection tube and 350µl membrane desalting buffer (MDB) was added to the column and centrifuged at 14K rpm for 1min. DNase reaction mixture was prepared by adding 10µl reconstituted rDNase to 90µl rDNase reaction buffer. 95µl DNase reaction mixture was then applied on to the silica membrane of the column and incubated at room temperature for 15min. Afterwards, silica membrane of NucleoSpin® RNA II column was subjected to a first wash, by adding 200µl RA2buffer and centrifuged for 30sec at 14K rpm. Later, the column was transferred to a new collection tube. During a second wash, 600µl RA3 buffer was

added to the NucleoSpin® RNA II column and centrifuged for 30sec at 14K rpm. Flow through was discarded. For a final wash, 250µl buffer RA3 was added to the NucleoSpin® RNA II column and centrifuged for 2min at 14K rpm to dry the membrane completely. Finally, RNA was eluted by adding 40µl RNase free water to NucleoSpin® RNA II column and centrifuged at 14K rpm for 1min.

Extraction of total RNA - Tripure method (Roche)

Total RNA was extracted from human fibroblast cell pellets using the Tripure isolation reagent following supplier guidelines. About 10^6 cells were used to extract the RNA. The cell pellet was first washed once with PBS and collected by centrifugation at 1.5K rpm for 5min. The cell pellet was loosened by flicking the tube gently and resuspended in 1ml of Tripure isolation reagent. Samples were then incubated for 5min at room temperature to dissociate nucleoprotein complexes. Remaining steps are same as described in RNA extraction by Trizol® method section.

2.3.5 - Complementary DNA (cDNA) synthesis

cDNA was synthesized by one of the following 3 different methods:

- Cloned AMV first-strand cDNA synthesis kit (Invitrogen)
- Superscript®III Reverse Transcriptase (Invitrogen)
- Quantitect Reverse Transcription kit (QIAGEN)

cDNA synthesis - Cloned AMV first-strand cDNA synthesis kit (Invitrogen)

cDNA was synthesized by using cloned AMV first strand cDNA synthesis kit (Invitrogen). On ice, 5µl (1µg) of RNA was added to 7µl of primer component mastermix (7µl of DEPC-water, 2µl 10mM dNTP mix and 1µl FAST RT primer or Oligo$(dT)_{20}$ primer). For Oligo(dT) cDNA synthesis, 4µl of DEPC-water, 2µl 10mM dNTP mix and 1µl Oligo$(dT)_{20}$ primer was added and for FAST RT strand specific cDNA synthesis, 4.5µl DEPC-water, 2µl 10mM dNTP mix and 0.5µl FAST RT(10µM) primer was added. RNA and primer were denatured by keeping the samples at 65°C for 5min and followed by immediately placing the samples on ice. Then the following reagents were added in order: 4µl 5X cDNA synthesis buffer, 1µl DEPC-water, 1µl 0.1M DTT, 1µl RNase OUT™ (40U/µl) and 1µl cloned AMV RT (15U/µl). The 20µl reaction mixture was gently mixed by flicking the tube and briefly centrifuged to bring all the contents to the bottom. The reaction mixture was incubated at 55°C for 60min. The reverse transcriptase reaction was terminated by keeping the samples at 85°C for 5min. The cDNA samples were used immediately or stored at -20°C.

cDNA synthesis - Superscript®III reverse transcriptase (Invitrogen)

cDNA was synthesized by using Superscript®III reverse transcriptase kit (Invitrogen). All the reactions were carried on ice. 2µl FAST RT (10µM) or Oligo dN_6 primer (random hexamer) was added to 10µl RNA (100ng/µl). RNA and primer were denatured by placing samples at 90°C for 40sec, followed by incubation at room temperature for 5min and samples were immediately placed on ice. Then the following reagents were added in order: 4µl First strand buffer, 1µl dNTP (10mM), 1.5µl DTT (0.1M), 0.5µl RNase OUT™ (40U/µl), 1µl Superscript®III RT(200U/µl). The 20µl reaction mixture was gently mixed by flicking the tube

and briefly centrifuged to bring all the contents to the bottom. The reaction mixture was incubated at 50°C for 1hour. The cDNA samples were used immediately or stored at -20°C.

cDNA synthesis - Quantitect reverse transcription kit (Qiagen):

cDNA was synthesized by using Quantitect reverse transcription kit (Qiagen). 1µg RNA sample was added to 2µl genomic DNA wipeout Buffer (7X) and made to a total volume of 14µl with DEPC-water. The samples were incubated at 42°C for 2min to remove genomic DNA contamination. Reverse transcription mastermix was prepared on ice by adding following reagents in order: 1µl Quantitect reverse transcriptase, 4µl Quantitect RT buffer (5X), 1µl FAST RT primer (10µM), 14µl above entire genomic DNA elimination reaction (which contains template RNA). The 20µl reaction mixture was gently mixed by flicking the tube and briefly centrifuged to bring all the contents to the bottom. The reaction mixture was incubated at 42°C for 30min. The cDNA samples were used immediately or stored at -20°C.

2.3.6 - Determination of RNA/DNA quantity and purity

RNA/DNA concentration and purity was determined by using NanoDrop™ 2000c spectrophotometer (NanoDrop, Thermo Scientific). The absorption (A) of ultra violet light (UV-light) at 260nm by RNA/DNA was used to determine the concentration and the quality by using $A_{260/280}$ ratio.

2.3.7 - DNase I treatment of RNA by using DNase I

DNase I treatment of RNA (1µg) was performed by adding 2µl of RNA (500ng/µl) to 8µl of mastermix (1µl 10X DNase I buffer, 1µl DNase I enzyme, 6µl water) and incubated at room temperature for 15min. DNase I enzyme was inactivated by adding 1µl of 25mM EDTA to the reaction mixture and incubated at 65°C for 10min. The DNase I treated RNA samples were used immediately or stored at -80°C.

2.3.8 - Polymerase chain reaction (PCR) - standard or gradient PCR

Polymerase chain reactions (PCR) were performed to amplify the target template using forward and reverse primers in a Bio-Rad tetrad PCR machine. The right concentrations and precise annealing temperatures of each of the primer set used in this study were assessed by gradient PCR. In the gradient PCR, the annealing temperature was set to a range of ±5°C to the annealing temperature calculated by the following formula:

T_M=4(G+C) + 2(A+T)°C

T_A= T_M - 5°C (where T_M-melting temperature, T_A-annealing temperature)

2.3.9 - Purification of PCR products from agarose gels

Purification of PCR products from agarose gels by following one of the two methods.

i. Wizard® SV Gel and PCR Clean-Up System kit (Promega)

ii. Gene clean method

iii. Nucleotrap gel extraction (Clontech)

Purification of PCR products from agarose gels by Wizard® SV Gel and PCR Clean-Up System kit (Promega)

This method involves the use of agarose gel electrophoresis and a Wizard® SV Gel and PCR Clean-Up System kit (Promega) to purify the PCR products. 1% agarose gel was prepared with a wide-toothed comb to form large wells. PCR products (~70µl) were separated on a 1% agarose TBE mini-gel for approximately 1hr at 70V in 1x TBE buffer, leaving an empty well between each sample loaded. Using a scalpel blade and under high wavelength UV light (362nm), the bands were excised from the gel and transferred to eppendorf tubes. The weight of the excised gel block was determined in milligrams (mg) and of membrane binding solution (4.5M guanidine isothiocyanate, 0.5M potassium acetate pH 5.0) was added at a ratio of 10µl of solution per 10mg of agarose gel slice. The agarose blocks were completely dissolved by incubating at 55°C in a water bath. The dissolved gel mixture was transferred to the SV minicolumn assembly (SV minicolumn placed in a collection tube) and incubated for 1min at room temperature and centrifuged at 14K rpm for 1min. The SV minicolumn was removed from the spin column assembly and the solution in the collection tube was discarded, then the SV minicolumn was returned to the collection Tube. The SV minicolumn was washed by adding 700µl of membrane wash solution (10mM potassium acetate, pH 5.0, 80% ethanol, 16.7µM EDTA pH 8.0) and centrifuged at 14K rpm for 1min. The wash step was repeated by adding 500µl of membrane wash solution to the SV minicolumn assembly and centrifuged at 14K rpm for 5min. The collection tube was emptied and centrifuged the column assembly for 1min at 14K rpm to allow removal of any residual ethanol. The SV minicolumn was then transferred to a clean eppendorf tube and 50µl of nuclease-free water was directly applied to the centre of the column without

touching the membrane followed by incubation at room temperature for 1min and then centrifugation at 14K rpm for 1min. The SV minicolumn was discarded and the eluted DNA was stored at 4°C.

Purification of PCR products from agarose gels by the Gene clean method

This method involves the use of agarose gel electrophoresis and a Gene clean kit to purify PCR products. 1% agarose gel was prepared with a wide-toothed comb to form large wells. PCR products (~50µl) were separated on a 1% agarose TBE mini gel for approximately 1hour at 70V in 1x TBE buffer, leaving an empty well between each sample loaded. Using a scalpel blade and under high wavelength UV light (362nm), the bands were excised from the gel and transferred to eppendorf tubes. The weight of the excised gel block was determined in milligrams (mg). 1/2 volume of TBE modifier and 4.5 volumes of NaI (sodium iodide) were added to a given volume of agarose gel slice. The agarose blocks were completely dissolved by incubating at 55°C in a dry bath for 10min. Glass milk was resuspended by vortexing for 1min. After incubation, 5µl glass milk was added to the tubes, mixed and incubated on ice for 30min to allow binding of the DNA to the silica matrix. The tubes were centrifuged for approximately 5sec at high speed. The supernatant was removed and the pellet was resuspended in 200µl new wash solution. The sample was then centrifuged for 15sec in the centrifuge and the supernatant was discarded. The wash procedure was repeated for two more times. The eppendorf tube was centrifuged for a few seconds to remove any residual supernatant. 10µl elution buffer was added to the tubes, incubated at 55°C for 5min and centrifuged at high speed for 30sec. The supernatant containing the eluted DNA was removed carefully and placed in a new labelled eppendorf tubes. In this first elution step, approximately 80% of the bound DNA was eluted. A second elution was done by adding

10µl elution buffer to the pellet, followed by incubating at 55°C for 5min and centrifuged at high speed for 5min. A second elution was resulted in an additional 10-20% recovery of eluted DNA.

Purification of PCR products from agarose gels by NucleoTrap gel extraction

5µl of PCR product was electrophoresed on a 1.2% agarose TAE mini gel for approximately 50min at 70V in 1X TAE buffer. Using a scalpel blade and under high wavelength UV light (362nm), the bands were excised from the gel and transferred to eppendorf tubes. The weight of the excised gel block was determined in milligrams (mg). For every 100 mg of agarose gel slice 300µl of buffer NT1 was added. Nucleotrap suspension was vortexed thoroughly until the beads are completely resuspended. 10µl of Nucleotrap suspension was added to the gel slice and incubated at 50°C in a water bath for 15min, vortexing for every 2-3min during the incubation period. The sample was then centrifuged at 10K rpm for 30sec at room temperature and the supernatant was discarded. 500µl of buffer NT2 was added to the pellet and vortex briefly. The tube was centrifuged at 10K rpm for 30sec at room temperature and the supernatant was discarded completely. Then the pellet was washed with 500µl of buffer NT3 and the sample was vortexed followed by centrifugation at 10K rpm for 30sec at room temperature. The supernatant was discarded completely and this wash step was repeated twice. The pellet was air dried by inverting the tubes on paper towels for 10min. Then 20µl of elution buffer NE was added to the pellet, which was resuspended by vortexing. The resuspended sample was incubated at room temperature for 10min and centrifuged at 10K rpm for 30sec. The supernatant containing the eluted DNA was removed carefully and placed in a new labelled eppendorf tubes

2.3.10 - Precipitation of RNA with ethanol

Precipitation of RNA with ethanol is the standard method to recover RNA from aqueous solutions. 0.1 volume of 3M sodium acetate and 2-3 volumes of ethanol were added to the RNA solution and mixed by repeated inversion, followed by incubation at -20°C for 1hour. RNA was recovered by centrifugation at 14K rpm for 30min at 4°C. The supernatant was discarded carefully without disturbing the pellet. The Pellet was washed with 1ml of ice-cold 70% ethanol and centrifuged at 14K rpm for 10min at 4°C. Supernatant was discarded carefully and vacuum centrifuged (Savant speed vac plus) by leaving the eppendorf tube lid open for 20min. The RNA pellet was stored at -80°C.

2.3.11 - Radio-labelling of RNA probe (Riboprobe) by in vitro transcription

Strand specific single-stranded RNA probes are not only easier to make but also generally yield stronger signals in hybridization reactions than DNA probes. The DNA template for transcribing RNA probes was generated by PCR with primers whose 5' ends encodes T7 promoter for T7 bacteriophage-encoded DNA-dependent RNA polymerases. Following purification of the PCR products from the gel (section 2.3.9 -), these double stranded DNA templates were radiolabelled by *in vitro* transcription reactions.

Radio-labelling of riboprobe by [α-^{32}P] GTP

For radio-labelling, the following components were added in the order: 8.45µl water, 2.5µl 1 mg/ml BSA, 0.8µl $MgCl_2$, 3.75µl 10X low GTP mix (5mM rATP, 5mM rCTP, 5mM rUTP, 0.5mM GTP), 2.5µl 10X TRO buffer, 1µl RNase out, 3µl DNA template, 2µl [α-^{32}P] GTP

(10mCi/ml) and 1µl T7 polymerase (~10U). The reagents were mixed by gently tapping the eppendorf tubes, centrifuged the tube for 2sec to transfer all the solution to the bottom of the tube and incubated at 37°C for 2hours.

2.3.12 - Purification of radiolabelled riboprobe by SPIN-pureTM G-50 columns

SPIN-pure™ G-50 columns were used to remove the unincorporated dNTPs (deoxy ribonucleotides) from the radiolabelled riboprobe. A SPIN-pure™ G-50 column was packed with sephadex G-50, which is a pre-hydrated gel. The column was tapped gently so that the dry gel settles to the bottom of the spin column. The top column cap was removed and 0.75ml of water was added, followed by replacing the cap and vortexed vigorously for 10sec. Air bubbles were removed sharply by tapping the bottom of the column. The top cap and the bottom cap of the column were removed and the column was placed in an eppendorf tube. Then, the column was centrifuged at 5K rpm for 2min to remove the interstitial fluid. Column was then transferred to a new eppendorf tube followed by adding 25µl of in vitro transcription reaction mixture on to the gel in the column and centrifuged at 5K rpm for 2min. The sample is purified by the retention of low molecular weight contaminants in the matrix while the larger molecules of interest were exchanged into the buffer and eluted into the eppendorf tube.

2.3.13 - Northern blotting

Northern blotting is a technique for size fractionating of RNA in a gel, followed by transfer and immobilization on a solid support (membrane) in such a manner that the relative positions of the RNA molecules are maintained. The resulting Northern blot is then hybridized with a radiolabelled riboprobe complementary to the RNA of interest. Signal generated from detection of the probe can be used to determine the size and abundance of the target RNA. Northern blotting was performed by using Ambion® NorthernMax® kit.

Preparation of gel

Agarose gel was prepared in a ventilating hood as 10X denaturing gel buffer contains formaldehyde. 1.25g of agarose was dissolved in 112.5ml DEPC (diethylpyrocarbonate) water, and 12.5ml 10X denaturing gel buffer was added for every 125ml of gel. The solutions were heated and mixed well by swirling gently and care was taken to avoid forming bubbles in the gel solution. The gel was poured to about 0.6cm in thickness in a horizontal gel tray with a thick teeth comb. After the gel solidified at room temperature, the comb was removed carefully. The gel tray was placed into the electrophoresis chamber and the gel was covered by about 1cm with 1X MOPS gel running buffer.

Preparation of RNA samples

Ethidium bromide was added to the formamide loading dye to a final concentration of 10µg/ml for direct visualisation of RNA after electrophoresis. Formamide loading dye with ethidium bromide was then added to the RNA pellet and also to the RNA millennium marker™. The sample was vortexed gently and incubated at 65°C for 15min in a dry heating block to denature RNA secondary structure and was then immediately placed on ice.

Electrophoresis

RNA samples (~17µg) and RNA millennium marker™ (5µg) were loaded into the wells of the gel and run at 80 volts for 3hours. Then the gel was examined under UV light and photographed by placing the transparent ruler next to the gel in a UV transilluminator imaging system (Bio-Rad) so that the positions of size markers can later be transferred to the film.

Transfer of RNA from gels to the nylon membrane

The denatured RNA was transferred from the agarose gel to the surface of a positively charged nylon membrane (Amersham Hybond™-N+ membrane) by a capillary action in an upward direction. Before transferring of RNA to the membrane, the following blotting materials were prepared: The unused gel above the wells was removed using a scalpel, Amersham Hybond™-N+ membrane was cut to the same size as the gel (15cm X 13.2cm), eight pieces of Whatman 3MM filter paper were cut to the size slightly larger than gel, white paper towels were folded in half to create a stack of 5cm in height, one piece of Whatman 3MM filter paper was cut large enough to cover the area of the gel and to reach across into the transfer buffer reservoir (to act as a wick). A supporting tray was placed in a blotting tray and 500ml of transfer buffer was added to the blotting tray. The long Whatman 3MM filter paper was soaked in the transfer buffer and placed over the support tray so that the ends remain in the transfer buffer. It was ensured that there were no bubbles trapped between support tray and filter paper. The gel was placed on the filter paper, Amersham Hybond™-N+ membrane soaked in transfer buffer was placed on the gel and was ensured that no air bubbles were trapped in between. Four pieces of Whatman 3MM filter papers soaked in transfer buffer were placed on the membrane followed by four pieces of dry

Whatman 3MM filter papers. A 5cm stack of white paper towels were placed on top followed by a glass plate. The whole blotting apparatus was covered with saran wrap to prevent evaporation of transfer buffer. A weight was placed on top of the stack (e.g. a litre bottle of solution) to assure an even contact of all the stack components. The time required for transferring of RNA from gel to the membrane was 2hours. After the transfer process had completed, the transfer set up was disassembled. The membrane was placed in a UV cross linker (UVP) to crosslink the RNA immediately.

Prehybridization and hybridization

ULTRAhyb (Ambion) was preheated to 68°C and the bottle was swirled to dissolve any precipitated material. The cross linked membrane was transferred to the hybridization tube. 20ml of preheated (68°C) ULTRAhyb (Ambion) was added to the hybridization tube and the membrane was prehybridized by placing the tube in a roller bottle type hybridization oven (Grant Boekel) at 68°C for 30min.

Adding probe to the prehybridized blot

25µl of radiolabelled riboprobe (2.3.11) was mixed with 1ml of ULTRAhyb and was immediately transferred to the hybridization tube. The two solutions were mixed well and hybridized overnight (20 hours) at 68°C. After the incubation, ULTRAhyb was removed and disposed according to the radioactive safety guidelines.

Washing and exposure to film

Low stringency wash: Low stringency wash solution (2X SSC, 0.1% SDS) was heated at 50°C for 2min to dissolve the SDS. The hybridized membrane was washed two times with 40ml low stringency wash solution by agitating at room temperature for 5min in a roller bottle type hybridization oven. After the Low stringency wash, the blot was removed from the hybridization tube, wrapped in a plastic saran wrap and exposed to X-ray film (Fujifilm) for 5 hours at -80°C for autoradiography. After developing the X-ray film (Fujifilm), the hybridized membrane was subjected to high stringency wash.

High stringency wash: High stringency wash solution (0.1XSSC, 0.1% SDS) was heated at 50°C for 2min to dissolve the SDS. The hybridized membrane was washed two times with 40ml high stringency wash solution by agitating at 68°C for 15min in a roller bottle type hybridization oven. After the high stringency wash, the blot was removed from the hybridization tube, wrapped in a plastic saran wrap and exposed to X-ray film (Fujifilm) for 4 days at -80°C for autoradiography.

2.3.14 - Quantitative real-time RT PCR (qRT-PCR)

Quantitative real-time reverse transcriptase PCR (qRT-PCR) was performed using SYBR® green master mix (Applied Biosystems) in a real time PCR machine (ABI Prism 7900HT, Applied Biosystems). qRT-PCR reactions were carried out in 96-well plates (Microamp, Applied Biosystems) in triplicates. A final volume of 20µl mastermix was prepared containing 10µl of 2x SYBR® green mastermix, 0.5-1.0µl of 2.5-5pmol of optimized respective forward and reverse primers, 5µl of (5x diluted) cDNA and distilled water to make

the final volume up to 20µl. Samples were minimized to light exposure. Target and endogenous mastermixes were prepared separately and added to the plate with a repetitive electronic pipette (Rainin) followed by adding the cDNA. Then the plate was sealed with real time plate sealers (Microamp, Applied Biosystems) and the contents were mixed gently by shaking the plate. The plate was then briefly (1min) centrifuged at 1K rpm to bring all the contents to the bottom of the well.

The cycling conditions varied according to the application and were optimized to amplify the different targets with similar efficiencies. Following each real time PCR reaction a dissociation curve run was performed by increasing the temperatures gradually from 60°C to 95°C. Relative quantification values were determined by $2^{-\Delta\Delta Ct}$ method using SDS 2.4 software (Applied Biosystems).

2.3.15 - Cloning and Sequencing

Cloning was performed using TOPO TA® cloning method. Cloning is the process where the desired PCR products were cloned in to a vector and then the recombinant vector was transformed in to the competent *E.Coli.* 1µl pCR™4-TOPO® vector, 1µl salt solution (1.2 M NaCl; 0.06 M MgCl) and 4µl gel purified desired PCR product were added to the tube and incubated at room temperature for 30min. After incubation, the TOPO® Cloning reaction was placed on ice in order to transform competent cells.

Transforming competent cells:

3µl TOPO® Cloning reaction was added to one vial (50µl) of DH5α™comptent cells and incubated on ice for 30min. The cells were subjected to heat shock treatment for 30sec at 42^0C in a water bath (Grant) without shaking. After heat shock treatment, the tube was immediately placed on ice for 2 min. Then 250µl of S.O.C. (super optimal broth with catabolite repression) medium was added to the tube and placed in a horizontal shaker to rotate the tube at 200rpm for 1 hour at 37^0C. After shaking, 50 µl of transformation solution was spread on prewarmed agar plates containing 50µg/ml of ampicillin and 50µg/ml of kanamycin, and incubated overnight at 37^0C. For analysing positive clones, 12 colonies were picked and cultured overnight in 5ml of LB medium containing 50µg/ml of ampicillin and 50µg/ml of kanamycin.

Plasmid DNA purification using the QIAprep spin mini kit

The 5ml bacterial cultures were harvested in 15ml universal tubes at 3K rpm for 5min at room temperature. The supernatant was removed by inverting the centrifuge tube until all medium had been drained and then the bacterial cell pellet was dried by placing the inverted tubes on paper towels for 5min. The bacterial cell pellet was vortexed for few seconds then resuspended in 250µl Buffer P1 buffer (50mMTris-HCl, pH 8.0, 10mM EDTA, 100µg/ml RNase A) by vortexing until no cell clumps remained and then transferred to the eppendorf tube. After, 250µl buffer P2 (200mM NaOH, 1% SDS) was added to the tube and mixed gently by inverting the tube 4-6 times, until the solution became viscous and slightly clear. Subsequently, 350µl buffer N3 (4.2 M Gu-HCl, 0.9 m potassium acetate ph 4.8) was added to the tube, and mixed immediately and thoroughly by inverting the tube 4-6 times.

Then the tube was centrifuged at 13K rpm for 10min at room temperature. The supernatant was applied to the QIAprep spin column by pipetting, and centrifuged for 60sec. The flow through was discarded and the QIAprep spin column was washed by adding 500µl buffer PB (5M Gu-HCl, 30% isopropanol), and centrifuged for 60sec. Then the flow through was discarded and the QIAprep spin column was washed again by adding 750µl buffer PE (10mm Tris-HCl, ph 7.5, 80% ethanol), and centrifuged for 60sec. The flow through was discarded, and centrifuged for an additional 1min to remove the residual buffer. The QIAprep spin column was placed in a clean 1.5ml eppendorf tube and the DNA was eluted by adding 50µl buffer EB (10mM Tris-Cl, pH 8.5) to the centre of the QIAprep spin column and centrifuged for 1min.

Restriction enzyme digestion of recombinant plasmids with *Eco*RI

The recombinant plasmids for inserts were analysed by restriction enzyme digestion with EcoRI. 1µl EcoRI enzyme, 2µl 10X EcoRI buffer, 12µl sterile water was added to the 5µl of plasmid DNA and incubated at 37°C in a water bath for 2hours. 4µl EcoRI digested plasmid DNA and 1µl undigested plasmid were analysed together on a 1% agarose gel.

Sequencing

For sequencing of recombinant plasmids by next generation DNA sequencing method, 30µl recombinant plasmids DNA (100ng/µl) in eppendorf tube were sent to Beckman Coulter Genomics.

2.3.16 - RACE (Rapid amplification of complementary DNA ends)

5’- and 3’- RACE was performed using SMARTer™ RACE cDNA amplification kit (Clontech) and it involves two steps:

I) Generating RACE-ready cDNA

II) Rapid amplification of cDNA ends

SMARTer IIA oligonucleotide, 5'-CDS Primer A, 3'-CDS Primer A, SMART Scribe™ Reverse Transcriptase (100 U/µl) was provided by the kit.

i) Generating RACE-ready cDNA

In two separate eppendorf tubes, 5'-RACE-Ready cDNA mix (1.0-2.75µl RNA and 1.0µl 5'-CDS Primer A) and 3’- RACE-Ready cDNA mix (1.0-3.75µl RNA and 1.0µl 3'-CDS Primer A) were prepared. For the control RACE experiment, 1µl control mouse heart total RNA (1µg/µl) was used. Sterile water was added to the eppendorf tubes for a final volume of 3.75µl for 5’-RACE and 4.75µl for 3’-RACE. The contents in the eppendorf tube were mixed gently and centrifuged briefly. Then the eppendorf tubes were incubated at 72°C for 3min, and the tubes were cooled to 42°C for 2min in a thermocycler. After cooling, the tubes were centrifuged briefly for 10sec at 14K rpm to collect the contents at the bottom of the tube. For the 5'-RACE cDNA synthesis reaction, 1µl SMARTer IIA oligonucleotide was added to the eppendorf tube and the contents were mixed well by vortexing and centrifuged briefly. Following mastermix was prepared at room temperature for 5’- & 3’-RACE-Ready cDNA synthesis reactions: 2.0µl 5X First-Strand Buffer, 1.0µl DTT (20 mM), 1.0µl dNTP Mix (10 mM), 0.25µl RNase Inhibitor (40 U/µl) and 1.0µl SMART Scribe™ Reverse Transcriptase (100 U). 5.25µl Master Mix was then added to the denatured RNA for 5'-RACE-

Ready cDNA synthesis and 3'- RACE-Ready cDNA synthesis to a total volume of 10µl. Then the contents of the tubes were mixed by gentle vortex and centrifuged briefly to collect the contents at the bottom of the tube. The eppendorf tubes were incubated at 42°C for 90min followed by 70°C for 10min incubation in a hot-lid thermal cycler. Then the first strand cDNA reaction product was diluted by adding 100µl Tricine-EDTA Buffer.

ii) Rapid Amplification of cDNA Ends (RACE)

This procedure describes the 5'-RACE and 3'-RACE PCR reactions that generate the 5'- and 3'- cDNA fragments. It was recommended to perform positive control for 5'- and 3'-RACE using the TFR primers, UPM, and control RACE-Ready cDNAs. For performing RACE PCR reaction, gene specific primers were designed with 50-70% GC content, 23-28 nucleotides long and with a T_m >70^0C which enables the use of touch down PCR. This also avoids non-specific complementarities to the 3'-end of universal primer mix and specific to the gene of interest. For 5'-RACE PCR, an antisense gene specific reverse primer, and for 3'-RACE PCR sense gene specific forward primer was designed. RACE PCR reactions were performed with the Advantage®2 Polymerase Mix (Table 2.8 & Table 2.9). Following master mix was prepared for 5'- and 3'-RACE PCR reactions by adding 34.5µl PCR-grade water, 5.0µl 10X Advantage 2 PCR buffer, 1.0µl dNTP mix (10 mM) and 1.0µl 50X Advantage®2 polymerase mix in an eppendorf tube and the contents were mixed by vortexing (without introducing bubbles) and centrifuged briefly.

Table 2.8 - 5'-RACE PCR reaction set up

Component	Tube No. & Description				
	1 5'-RACE Sample	2 5'-TFR* (+ Control)	3 GSP1 + GSP2† (+ Control)	4 UPM only (– Control)	5 GSP1 only (– Control)
5'-RACE-Ready cDNA (experimental)	2.5 µl	2.5 µl	2.5 µl	2.5 µl	2.5 µl
UPM (10X)	5 µl	5 µl	–	5 µl	–
GSP1 (10 µM)	1 µl	–	1 µl	–	1 µl
GSP2 (10 µM)	–	–	1 µl	–	–
Control 5'-RACE TFR Primer (10 µM)	–	1 µl	–	–	–
H_2O	–	–	4 µl	1 µl	5 µl
Master Mix	41.5 µl	41.5 µl	41.5 µl	41.5 µl	41.5 µl
Final Volume	50 µl	50 µl	50 µl	50 µl	50 µl

* Skip this reaction if your RNA is nonmouse.
† Skip this reaction if your GSPs will not create overlapping RACE fragments.

Table 2.9 - 3'-RACE PCR reaction set up

Component	Tube No. & Description				
	1 3'-RACE Sample	2 3'-TFR* (+ Control)	3 GSP1 + GSP2† (+ Control)	4 UPM only (– Control)	5 GSP1 only (– Control)
3'-RACE-Ready cDNA (experimental)	2.5 µl	2.5 µl	2.5 µl	2.5 µl	2.5 µl
UPM (10X)	5 µl	5 µl	–	5 µl	–
GSP1 (10 µM)	–	–	1 µl	–	–
GSP2 (10 µM)	1 µl	–	1 µl	–	1 µl
Control 3'-RACE TFR Primer (10 µM)	–	1 µl	–	–	–
H_2O	–	–	4 µl	1 µl	5 µl
Master Mix	41.5 µl	41.5 µl	41.5 µl	41.5 µl	41.5 µl
Final Volume	50 µl	50 µl	50 µl	50 µl	50 µl

* Skip this reaction if your RNA is nonmouse.
† Skip this reaction if your GSPs will not create overlapping RACE fragments.

Following touchdown PCR programme was performed in a thermal cycler

- 5 cycles: 94°C 30 sec
  72°C 3 min

- 5 cycles: 94°C 30 sec
  70°C 30 sec
  72°C 3 min

- 25 cycles: 94°C 30 sec
  68°C 30 sec
  72°C 3 min

To visualize the results, 5µl of PCR products were separated in 1.2% agarose TAE mini-gels along with a 1kb+ DNA ladder (Invitrogen) at 75V for ~30min and the remaining 45µl of each PCR reaction was stored at 4°C until the control experiment has satisfactorily worked.

Nested RACE PCR

The primary RACE PCR product is resulted in smear PCR products and to increase the sensitivity of the PCR we have performed nested RACE PCR. Primary RACE PCR product was diluted in tricine-EDTA buffer (10X, 50X, 100X). For nested RACE PCR reaction, 34.5µl PCR-grade water, 5.0µl 10X Advantage®2 PCR Buffer, 1.0µl dNTP Mix (10 mM) and 1.0µl 50X Advantage 2 Polymerase Mix, 1µl nested universal primer A (10µM), 1µl nested gene specific primer (10µM) and 2.5µl of diluted Primary RACE PCR product were added in a PCR tube, the contents were centrifuged briefly and PCR in a thermal cycler.

- 1 cycle 94°C 30 sec
- 20 cycles: 94°C 30 sec
  68°C 30 sec
  72°C 3 min
- 1 cycle: 72°C 10 min

2.4 - Establishment of primary fibroblasts

None of the currently available cell models reproduce all of the essential molecular and cellular disease mechanisms that are known to occur in FRDA patients. Therefore, there is still a great need for the development of additional cell models. I have established primary fibroblasts from Y47R, YG8R, YG22R and YG8sR mouse kidney tissues.

2.4.1 - Culture medium of fibroblasts

Preparation

The culture medium for fibroblasts (DMEM) was prepared by mixing the aliquots of FCS (10% v/v) and pen-strep (1-2% v/v) on 500ml of 1X DMEM medium. The contents were mixed well and filter sterilized with a 0.22µm pore size filter unit (Nalgene) and stored at 4°C until required to use.

2.4.2 - Culture procedure of fibroblasts

The culture procedure of fibroblasts in our lab was originally based on Gomes-Pereira and Monckton (2004) and it is recommended for the establishment of cell cultures from large mouse organs like lung, liver, kidney and heart

Procedure

The procedure to establish the fibroblasts have recently been described (Sandi *et al.* 2014) briefly the kidneys were aseptically collected from the selected mice. The tissues were then chopped into small pieces followed by incubation with trypsin-EDTA (0.25%) for 60-90

min. Primary cultures were grown in DMEM medium with 10% FCS and 1% pen-strep in 5% CO_2 at 37°C.

2.4.3 - Subculture of fibroblasts

Prior to the subculture, the culture medium, PBS and trypsin-EDTA solutions were prewarmed at 37°C in a water bath. The old medium was removed carefully from the flask and washed once with sterile PBS. 1ml (for every 25cm^2 culture flask) of prewarmed trypsin-EDTA was added and gently rock the flask to ensure that the entire monolayer of cells was covered with the trypsin solution. The flask was then incubated at 37°C until the cells begin to detach, usually 3-5 min. Care should be taken to avoid leaving the cells exposed to the trypsin solution longer than necessary. After observing the cells under microscope to ensure that all the cells have detached and floating, a small amount of serum-containing medium (usually 5ml) was added to inactivate the trypsin. Depending on the cell quantity and growth rate of the cells subculture was performed between the ranges of 1:5 to 1:10 ratio.

2.4.4 - Cell quantification and viability (Trypan blue exclusion assay)

The trypan blue exclusion test is used to determine the number of viable cells present in a cell suspension. It is based on the principle that live cells possess intact cell membranes that exclude certain dyes, such as trypan blue, Eosin, or propidium, whereas dead cells do not. To perform the trypan blue assay, the cells were first digested with trypsin-EDTA and cell suspension was made with DMEM medium. 1:1 dilution of 0.4%

trypan blue solution and cell suspension (10µl each) was made and cell number was counted using a Countess® automated cell counter (Invitrogen).

2.4.5 - Mycoplasma PCR testing

Mycoplasma PCR was carried out using Mycosensor™ PCR Assay kit (Stratagene) and the procedure was performed according to the manufacturer's recommendations. The primer set in the kit recognises most Mycoplasma infections using as little as 100 µl of cell culture supernatant. Briefly, 100µl of supernatant from test cell culture flask was carefully transferred into 1.5ml eppendorf tube and the lid was tightly closed followed by incubation at 95°C for 5 min. The sample was spun quickly (2-3 sec) and 10µl of strataclean resin was added. The suspension was mixed by gently flicking the tube and a quick centrifugation (5-10 sec) was performed. The clear supernatant (test sample) was transferred to a fresh tube and sample was placed on ice. Mycoplasma PCR mastermix was prepared by adding the following reagents in order; 33µl dH_20, 5µl 10x PCR buffer (Qiagen), 1µl Q buffer (Qiagen), 2.5µl 10mM dNTP mix, 1µl primer mix, 2µl *Taq* DNA polymerase, 5µl internal control and 5µl test sample. The cycling conditions of Mycoplasma PCR were as described for "PCR using *Taq* DNA polymerase" in Table 2.10. To visualize the results, 10-15µl of PCR products were separated in 1-2% agarose TBE mini-gels along with a 1kb+ DNA ladder (Invitrogen) at 75V for ~30min.

Table 2.10 - Mycoplasma PCR cycling parameters

Cycle(s)	Temperature	Duration			
		PCR using *Taq* DNA polymerase[a]	PCR using *Taq* DNA polymerase with dUTP/UNG decontamination	PCR using the Brilliant QPCR master mix	PCR using the Brilliant QPCR master mix with dUTP/UNG decontamination
1	37°C	—	10 minutes	—	10 minutes
	94°C	—	10 minutes	10 minutes	10 minutes
35	94°C	30 seconds	30 seconds	30 seconds	30 seconds
	55°C	1 minute	1 minute	1 minute	1 minute
	72°C	1 minute	1 minute	1 minute	1 minute

2.4.6 - Cryopreservation and regeneration of fibroblasts

The aim of cryopreservation is to enable stocks of cells to be stored to prevent the need to have all cell lines in culture at all times. Cells, to be frozen, were first harvested by trypsin-EDTA digestion and resuspended in 1ml of DMEM medium with 10% (v/v) DMSO in cryotube vials. Cells were mixed with the freezing medium at room temperature to a final concentration of viable cells in a range of 10^6 to 10^7 per ml. The cryotubes were then placed in a cooling container and transferred to -80°C to reduce the temperature 1°C per minute. Once the temperature is down to -50°C or approximately 3 hours after placing the cooling container at -80°C, transfer the cryotubes into liquid nitrogen for long-term storage..

In contrast to the cryopreservation, the cells should be thawed as quickly as possible. The cryotube with desired cell lines was carefully removed from the liquid nitrogen tank and immediately placed at 37°C water bath. The tube is intermittently shaken until the cell freezing solution was completely thawed. To prevent any possible contamination, the cryotube was rinsed with 70% ethanol before it was transferred to the laminar air flow cabinet. Once the cells were completely thawed, the cell suspension was added to a 15 ml

centrifuge tube containing 10 ml of prewarmed culture medium. Cells were then split in between two culture flasks and growth was observed daily.

2.5 - Statistical analysis

Statistical analyses such as descriptive measurements and graphical visualization were done using Microsoft excel 2007 software. Statistical values comparing two sample groups were determined using the student's t test and a statistical significance level of 5% was chosen to determine if the mean values differed significantly or not.

Chapter 3 - Identification of *FAST-1* by Northern Blotting

3.1 - Introduction

Although the exact mechanism by which the GAA repeat expansion leads to decreased frataxin expression is not known, evidence suggests that GAA repeat expansions can produce heterochromatin-mediated gene silencing effects (Saveliev *et al.* 2003). Several FRDA disease-related epigenetic changes have been identified in the immediate vicinity of the expanded GAA repeats of the *FXN* gene. The potential role of epigenetic mechanisms in FRDA disease was first suggested by the finding that long GAA repeats could suppress the expression of a nearby heterochromatin-sensitive cell surface reporter gene in transgenic mice by position effect variegation (Saveliev *et al.* 2003). Antisense transcription is considered as one of the epigenetic mechanisms that may affect *FXN* gene expression without involving changes in the primary DNA sequence. Antisense RNA has been ascribed to roles in several molecular mechanisms, including the regulation of gene expression. Recent studies have shown that antisense transcripts can be detected in various genes, including the non-pathogenic alleles, known as natural antisense transcripts (NATs).

It has been recently reported that *FAST-1* is transcribed from the *FXN* antisense strand, the opposite strand to the protein coding strand or sense strand. Bidichandani and colleagues have reported significantly increased levels of a frataxin antisense transcript 1 (*FAST1*) in FRDA fibroblast cells, associated with depletion of CTCF binding at the 5' UTR region of the *FXN* gene (De Biase *et al.* 2009). This finding has suggested possible involvement of these factors in the heterochromatin and *FXN* gene silencing processes of FRDA disease. FRDA is a disorder that currently has no effective therapy. However, since FRDA is associated with several epigenetic changes that result in a partial deficiency of frataxin mRNA and protein, reversing the epigenetic changes to upregulate frataxin

expression may prove to be an effective therapy. As a part of epigenetic therapy, and since very little is known about the *FAST-1*, it is important to identify and characterise the *FAST-1* to unravel any underlying mechanism that lead to the *FXN* gene silencing in FRDA. Therefore, I aimed to identify the *FAST-1* initially by Northern blot analysis. Northern blot analysis is one of the most reliable and widely used standard methods for validating and quantitating mRNAs and small RNAs. It not only detects the presence of a transcript but also indicates size and relative comparison of transcript abundance on a single membrane. Numerous reports have described various improvements of Northern blotting, which propelled this technique as the method of choice to study gene expression changes. It also reveals information about RNA identity and size allowing a deeper understanding of gene expression levels.

3.2 - Preparation of *FAST-1* radiolabelled riboprobe

In order to identify a *FAST-1* transcript by Northern blot analysis, I used three primary fibroblast cell lines derived from unaffected individuals (H.Normal, GM04503, GM07492) and two fibroblast cell lines from FRDA patients (GM03816, GM03665). Genomic DNA was extracted from all these cell lines by the phenol-chloroform method and a gradient PCR was performed using *FXN*.f1 and *FXN*.r1 primers, which amplify a 676 bp DNA fragment between exon 1 and intron 1 of the *FXN* gene (Figure 3.1). Gradient PCR revealed that the DNA fragment was efficiently amplified at 53°C and therefore all subsequent PCRs were done at this annealing temperature only.

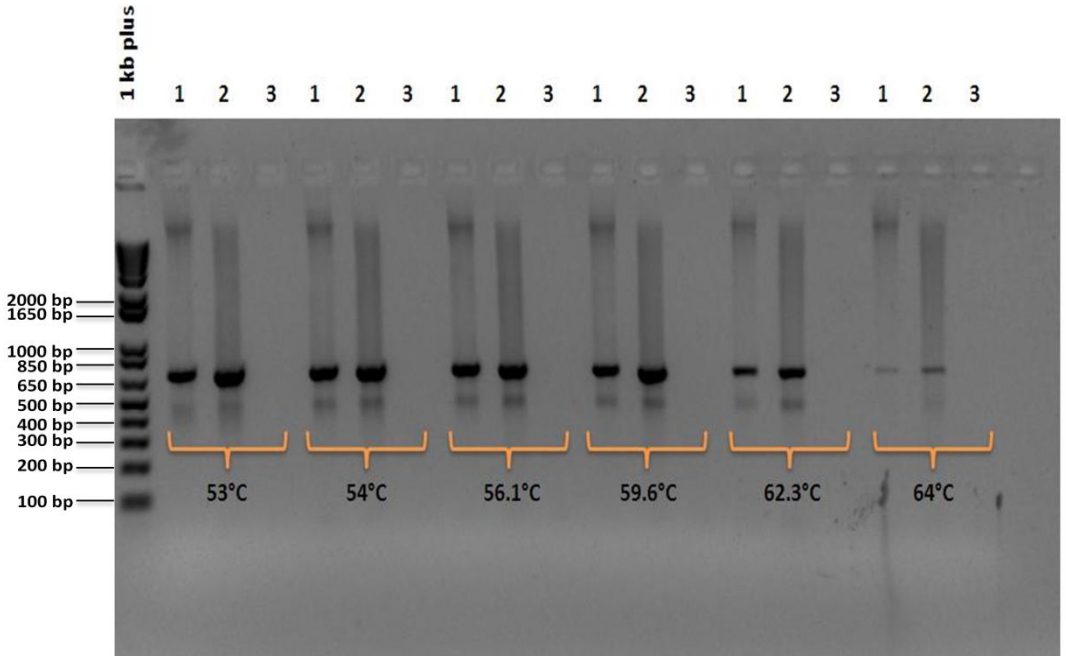

Figure 3.1 - Gradient PCR amplification with *FXN*.f1 and *FXN*.r1 primers to generate a DNA fragment with 676 bp size between exon 1 and intron 1 of the *FXN* gene. Lanes 1, 2 and 3 are human genomic DNA, HeLa cell genomic DNA and water, respectively.

Afterwards, the resultant PCR products were gel purified using the Wizard® SV Gel and PCR Clean-Up System kit (Promega). The gel purified PCR products were then diluted 5X and 2µl of diluted PCR products were used for nested PCR amplifications. A long DNA riboprobe with 675 bp size was amplified using *FXN*-K-T7.f2 and *FXN*.r2B primers and a short DNA riboprobe with 238 bp size was amplified using *FXN*-K-T7.f2 and *FXN*.r3B primers (Figure 3.2). The *FXN*-K-T7.f2 primer contains a T7 promoter overhang, which is essential for the RNA polymerase function during the *in vitro* transcription process.

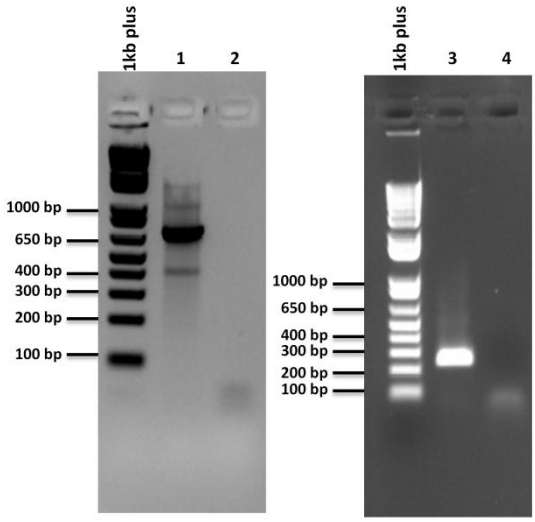

Figure 3.2 – The above image represents the amplification of long and short riboprobes. Lanes 1 & 3 - long DNA riboprobe (675 bp) and short DNA riboprobe (238 bp), respectively; 2 & 4 - water control.

The riboprobes were then gel purified and radiolabelled with [α-P^{32}] GTP by *in vitro* transcription. Afterwards, the riboprobes were purified using the SPIN-pure^{TM} G-50 columns and analysed by running them on a 5% urea PAGE gel. The gel was then exposed to an X-ray film (FujiFilm) for 1 hour at -80°C for autoradiography followed by developing the X-ray film using an automated developer (Agfa) (Figure 3.3).

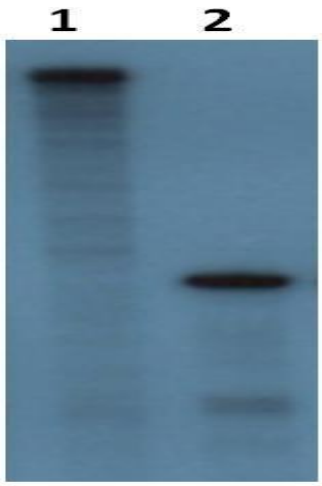

Figure 3.3 – X-ray film image developed by exposing the radiolabelled riboprobes which analysed on a 5% urea PAGE gel. Lane 1 - long riboprobe (675 bp) and lane 2- short riboprobe (238 bp).

3.3 - Northern blotting

For Northern blotting analysis, I have isolated total RNA from three primary fibroblast cell lines from unaffected individuals (H.Normal, GM04503, GM07492) and two from FRDA patients (GM03816, GM03665). The RNA was eluted in 60µl of nuclease-free water by NucleoSpin®RNAII-total RNA isolation kit according to the manufacturer's recommendations. The concentration of RNA was measured by Nanodrop and then precipitated for concentrating the RNA. The RNA samples (17µg) and RNA millennium marker™ (5µg) were mixed with formamide loading dye and ethidium bromide followed by denaturation of the samples by incubating at 65°C for 15min. Afterwards, the samples were loaded on a 1% denaturing agarose gel and run for 3 hours at 80 volts. The gel was then examined under UV light and imaged by placing the transparent ruler next to the gel in a UV transilluminator imaging system. The sharp and clear 28S and 18S ribosomal bands in the gel indicate that the RNA was intact and good quality (Figure 3.4). The denatured RNA was then

transferred to Hybond^{TM}-N+ membrane followed by placing the membrane in a UV crosslinker.

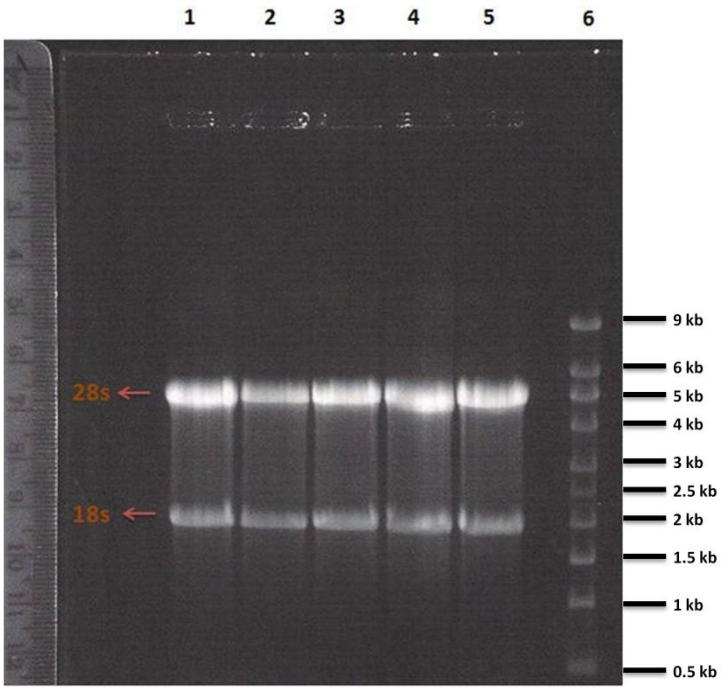

Figure 3.4 –Agarose gel electrophoresis of total RNA extracted from control and FRDA primary fibroblasts. The 18S and 28S ribosomal RNA bands are clearly visible in the intact RNA sample. Lanes: 1-3 control, 4-5 FRDA, and 6-RNA millennium marker.

3.4 - Results

3.4.1 - Hybridisation of Northern blot with *FAST-1* radiolabelled short riboprobe

Northern blotting was performed by hybridizing the radiolabelled short riboprobe (238 bp size) to the cross-linked Northern blot (Hybond^{TM}-N+ membrane). After the low stringency wash, the blot was exposed to an X-ray film for 5 hours at -80°C followed by developing the X-ray film using developer equipment (Figure 3.6).

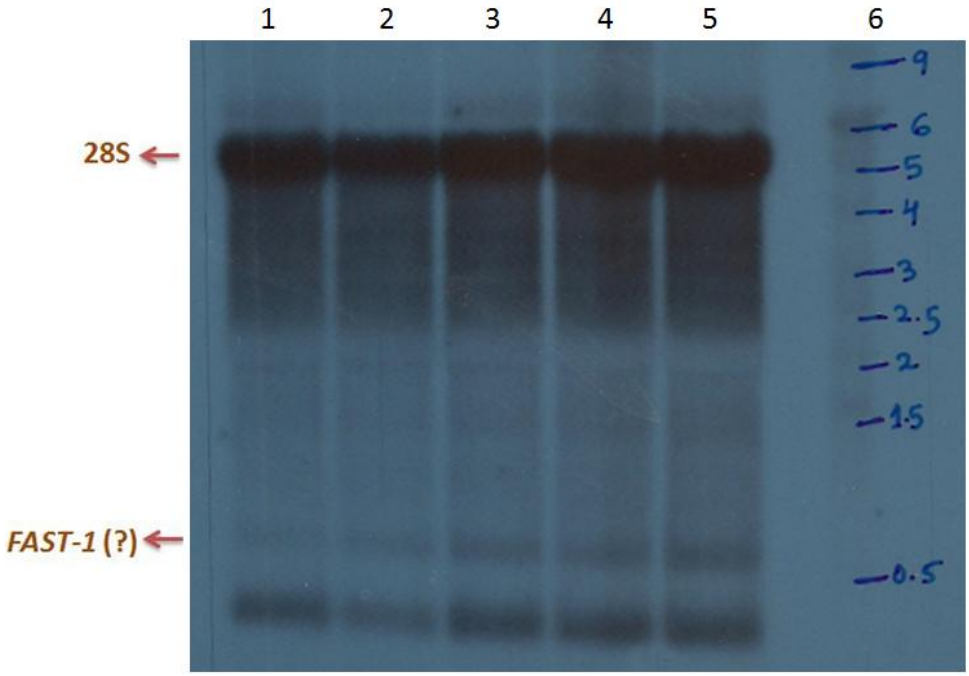

Figure 3.5 – X-ray film image developed by exposing to the Northern blot hybridized with radiolabelled short riboprobe (238 bp) after low stringency wash. Lanes: 1-3 control, 4-5 FRDA, and 6- RNA millennium marker.

The Northern blot analysis after the low stringency wash using the short riboprobe generated two distinct bands at different locations. One clear band was seen around 5kb size, which may represent non-specific hybridisation to the 28S rRNA. A second band was observed just above 500 bp which may represent *FAST-1* RNA.

After the low stringency wash, the blot was subjected to high stringency wash and then exposed to an X-ray film for 4 days and at -80°C followed by development of the X-ray film using developer equipment (Figure 3.6). The Northern blot analysis after the high stringency wash generated two distinct bands at different locations. One clear band was seen around 5kb size, which may represent non-specific hybridisation to the 28S rRNA. A second faint band was observed above 9 kb which may represent *FAST-1* RNA. The band

above 500 bp size which was observed after the low stringency wash, disappeared after the high stringency wash.

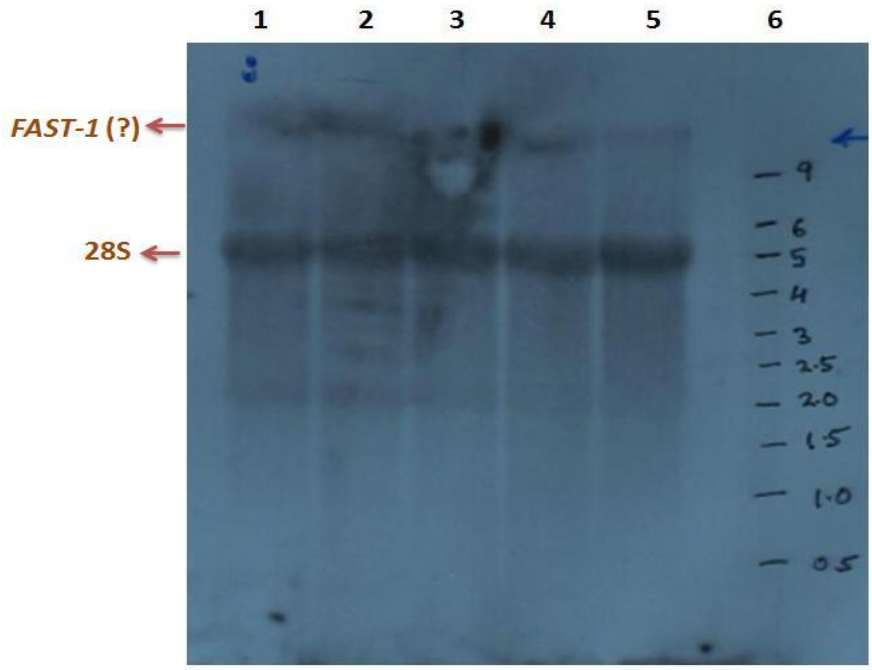

Figure 3.6 – X-ray film image developed by exposing to the Northern blot hybridized with radiolabelled short riboprobe (238 bp) after low and high stringency wash. Lanes: 1-3 control, 4-5 FRDA, and 6- RNA millennium marker.

3.4.2 - Hybridisation of Northern blot with *FAST-1* radiolabelled long riboprobe

Since the bands on the film were faint, I have decided to hybridise the same membrane with the radiolabelled long riboprobe (675 bp). To perform this, the previously hybridised radiolabelled short riboprobe on the Northern blot was stripped by washing with 20ml of 0.1% SDS for 30min at 68°C followed by hybridisation with the radiolabelled long riboprobe. After low and high stringency washes, the blot was exposed to an X-ray film for 4 days at -80°C, followed by development of the X-ray film using developer equipment (Figure 3.7).

The Northern blot analysis using the long riboprobe generated two distinct bands at different locations, consistence with the results obtained using short riboprobe. However, the bands that appear above 9kb size were stronger with the long riboprobe than with the short probe (Figure 3.7).

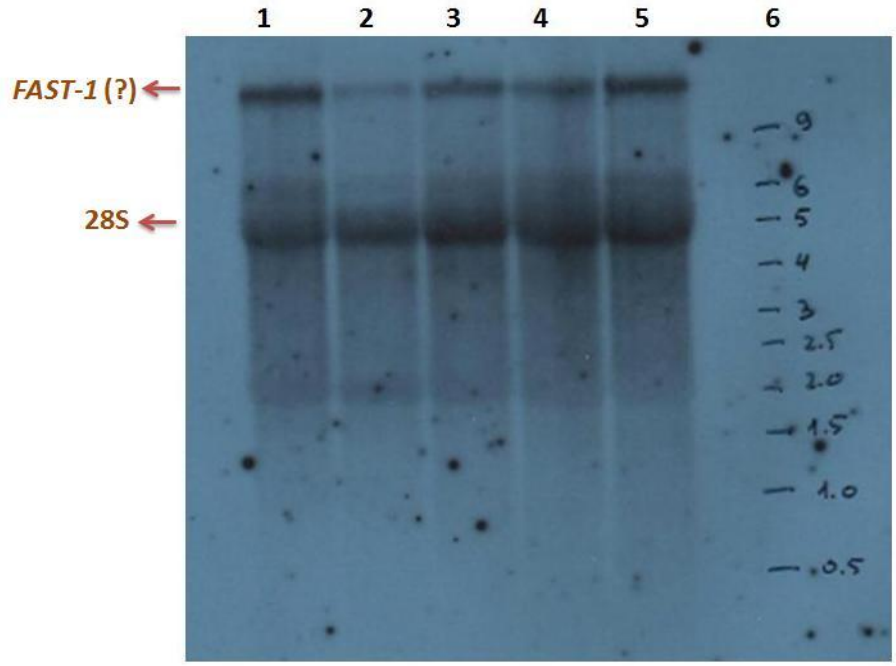

Figure 3.7 – X-ray film image developed by exposing to the Northern blot hybridized with radiolabelled long riboprobe (675 bp). Lanes: 1-3 control, 4-5 FRDA, and 6- RNA millennium marker.

3.5 - Discussion

Antisense transcripts have been implicated in heterochromatin formation via H3K9me3 and subsequent recruitment of HP1 (Grewal and Elgin 2007; Iida *et al.* 2008). In recent years, attention has been drawn to the fact that a significant fraction of the transcriptome comprises RNAs containing sequences that are complementary to other endogenous RNAs. These natural antisense transcripts (NATs) can have protein-coding properties but mainly represent non-coding RNAs (ncRNAs) (reviewed in Beiter *et al.* 2009). NATs have been implicated with diverse regulatory functions at various levels, including imprinting, X-inactivation, RNA processing, RNA export, and transcriptional regulation (reviewed in Beiter *et al.* 2009). The discovery of *FAST-1* is therefore not surprising in itself. Furthermore, the detection of higher levels of *FAST-1* in the region showing heterochromatin formation in FRDA patients offers a plausible mechanistic basis for the epigenetic abnormality involving in *FXN* gene silencing.

In order to eventually investigate the potential expression of *FAST-1* in human FRDA YAC transgenic mouse models, I have first tried to reproduce the results of De Biase *et al.* (2009) by investigating *FAST-1* expression in human FRDA fibroblast cells. As a different approach, I have initially performed Northern blotting experiments to detect *FAST-1* expression. The Northern blotting was performed by hybridising the Northern blot with sense radiolabelled riboprobe to detect antisense RNA of *FXN*, *FAST-1*. During the Northern blot hybridisation using a short radiolabelled riboprobe, after a low stringency wash, we have observed a band with a size of over 500 bp. And after the high stringency wash, the band corresponding to this size was disappeared but resulted in a new band size appeared at above 9kb size. Furthermore, the Northern blotting hybridisation with a long

radiolabelled riboprobe, after low and high stringency washes, produced a thick band over 9kb in size. Since the Northern blotting experiment resulted in identification of two bands with different sizes (one band with over 500 bp size and another band with above 9kb size), it became hard to speculate which band might correctly correspond to the antisense RNA of *FXN*, the *FAST-1*. Therefore, to further confirm the presence of *FAST-1*, and to identify the 5'- and 3'- ends, and total length of *FAST-1*, I have performed RACE (rapid amplification of cDNA ends) experiments and this will be explained in the next chapter.

Chapter 4 - Identification of *FAST-1* by Rapid Amplification of cDNA Ends

4.1 - Introduction

A growing number of natural antisense RNA transcripts of different eukaryotic genes have been detected over the last few years and the widespread existence of natural antisense RNA in eukaryotes is now well accepted (Rosok and Sioud 2004). Although the function of most of these antisense RNAs are largely not known, growing evidence has suggested their involvement in gene regulation. The initial discovery of pervasive antisense transcription was faced with justifiable scepticism (Johnson *et al.* 2005). Therefore, precise identification of such antisense transcripts is technically challenging issue, as it requires identification of strand-specific approaches (see review Pelechano and Steinmetz 2013). Similarly, in regards to the *FAST-1*, it is also important to determine the exact size and the location to resolve whether it is derived from the mRNA or the genome. I therefore decided to perform 5' and 3' rapid amplification of cDNA ends (RACE) to identify the full length *FAST-1*.

RACE is a powerful tool to quickly obtain full-length cDNA when the sequence is only partially known. Starting with cellular mRNA, PCR is used to amplify regions between the known parts of the sequence and nonspecific tags at the ends of the cDNA. RACE results in the production of a cDNA copy of the RNA sequence of interest, produced through reverse transcription and subsequent PCR amplification of the cDNA copies. In classic RACE, homopolymer tailing of the cDNA is used to append a linker sequence to the cDNA terminus. Extension of unknown regions of the cDNA is then achieved through PCR using a gene-specific primer and a primer that can bind and prime DNA synthesis from the linker sequence. By performing RACE analysis of *FAST-1*, followed by cloning and sequencing, I

have identified a full-length *FAST-1* transcript and determined its location with relation to the *FXN* gene.

4.2 - Methodology

To obtain a full length sequence of *FAST-1*, 5'- and 3'- RACE experiments were performed using SMARTer™ RACE cDNA amplification kit (Clontech). The details of this technique are described below and the entire protocol is adopted from Clontech (www.clonetech.com).

Smart Technology: The cornerstone of SMARTer™ RACE cDNA amplification kit is SMART technology, which provides a mechanism for generating full-length cDNAs by reverse transcription, made possible by the joint action of a SMARTer IIA oligonucleotide and SMARTScribe reverse transcriptase (a variant of MMLV RT). The SMARTScribe RT, upon reaching the end of an RNA template, exhibits terminal transferase activity, adding 3-5 residues to the 3'- end of the first strand cDNA. The SMARTer oligo contains a terminal stretch of modified bases that anneal to the extended cDNA tail, allowing the oligo to serve as a template for the RT. SMARTScribe RT switches templates from the mRNA molecule to the SMARTer oligo, generating a complete cDNA copy of the original RNA with the additional SMARTer sequence at the end. Since the template switching activity of the RT occurs only when the enzyme reaches the end of the RNA template, the SMARTer sequence is typically only incorporated into full-length, first strand cDNAs (Figure 4.1). Following reverse transcription, SMART technology allows the first strand cDNA to be used directly in 5'- and 3'- RACE PCR reactions. Incorporation of universal primer binding sites in a single-step during first strand cDNA synthesis eliminates the need for tedious second-strand synthesis and adaptor ligation. The only requirement of SMARTer™ RACE cDNA

amplification kit is to know at least 23-28 nucleotides of sequence information to design gene specific primers (GSPs) for the 5'- and 3'- RACE reactions.

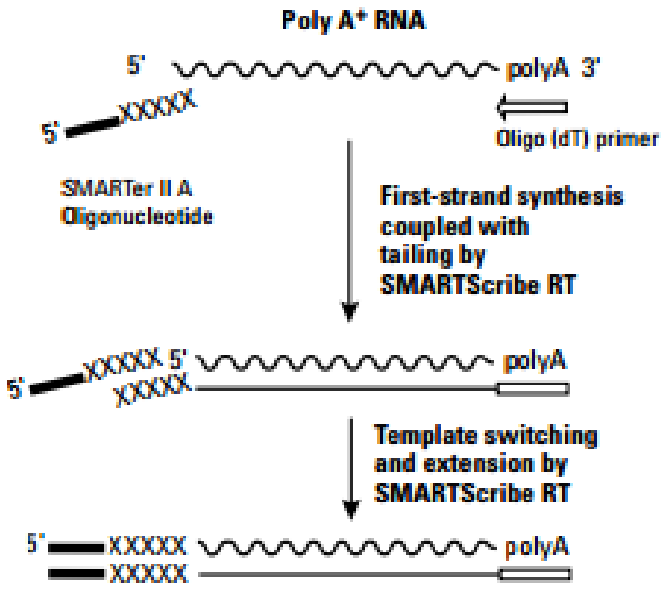

Figure 4.1 – Mechanism of SMARTer cDNA synthesis. First strand synthesis is primed using a modified oligo (dT) primer. After SMARTScribe reverse transcriptase (RT) reaches the end of the mRNA template, it adds several non-template residues. The SMARTer II A oligonucleotide anneals to the tail of the cDNA and serves as an extended template for SMARTScribe RT (Clonetech).

Primer design: Ideal GSPs should be 23-28 nucleotides long, with a GC content of 50-70%, and T_m (melting temperature) greater than 70°C. They should not be complementary to the 3'-end of the universal primer mix (UPM) and specific to gene of interest. For experiments, an antisense primer should be designed for 5'-RACE PCR and a sense primer for the 3'-RACE PCR. By designing the primers that give a 100-200 bp overlap in the RACE products, it is possible to use the primers together as a positive control for the PCR reactions. Furthermore, nested PCR may be necessary where the level of background or nonspecific amplification in the 5'- or 3'- RACE reaction is too high with a single GSP (Figure 4.2).

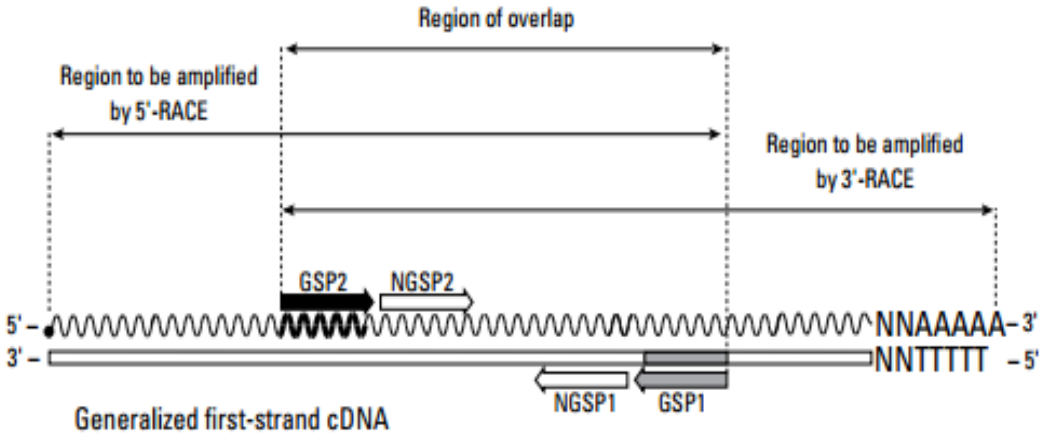

Figure 4.2 – Relationship of GSPs to the cDNA template. This diagram shows a generalised first-strand DNA template and the GSPs are designed to produce overlapping RACE products. This overlap permits the use of the primers together in a control PCR reaction (Clonetech).

4.3 - Identification of 5'- and 3'- ends of *FAST-1* by 5'- and 3'- RACE experiments

FAST-1 was initially identified in human fibroblasts. However, the size and location of this transcript have not yet been reported. Therefore, to identify a full length *FAST-1* transcript, I performed RACE analysis and subsequent cloning and sequencing. The details of this complex procedure and critical approaches along with our findings are discussed in subsequent sections below.

4.3.1 - Positive control RACE cDNA synthesis using the mouse heart total RNA

Before proceeding with the *FAST-1* RACE experiment, I performed a positive control cDNA synthesis using the mouse heart total RNA provided with the SMARTer™ RACE cDNA amplification kit. A schematic representation of 5' RACE and 3' RACE work-flow can be seen in Figure 4.3. The positive control 5'-RACE ready cDNA was generated by using mouse heart

total RNA, 5'-RACE CDS primer A and SMARTer II A oligonucleotide. 3'-RACE ready cDNA was generating by using 3'-RACE ready cDNA, mouse heart total RNA and 3'-RACE CDS primer A. The cDNA was synthesised as mentioned in materials and methods section (section 2.3.16). Positive control RACE PCR experiment was performed using the RACE ready cDNAs generated from the control mouse heart total RNA, which amplified the ends of transferrin receptor (TFR) cDNA. Control 5'-RACE ready cDNA was amplified by using 5'-RACE TFR and UPM (Universal primer mix A) primers and control 3'-RACE ready cDNA was amplified by using 3'-RACE TFR and UPM primers, employing a touchdown PCR programme. Simultaneously, an internal control RACE PCR experiment was also performed with the 5'- and 3'-RACE ready cDNAs using both 5'-RACE TFR and 3'-RACE TFR primers. The analysis of RACE PCR products in a 1.2% agarose gel indicated the presence of ladder of faint bands and one thick band at approximately 2.1kb size in 5' RACE PCR (Figure 4.4). However, the 3' RACE-PCR was amplified one thick band at approximately 3.1kb size along with few other faint bands (Figure 4.4). As expected, the internal control PCR for 5'- and 3'-RACE cDNAs resulted in an amplification of overlapping region of 380 bp size products (Figure 4.4). These results indicate successful execution of the RACE experiment using our laboratory conditions and therefore I proceeded to the next steps of the process.

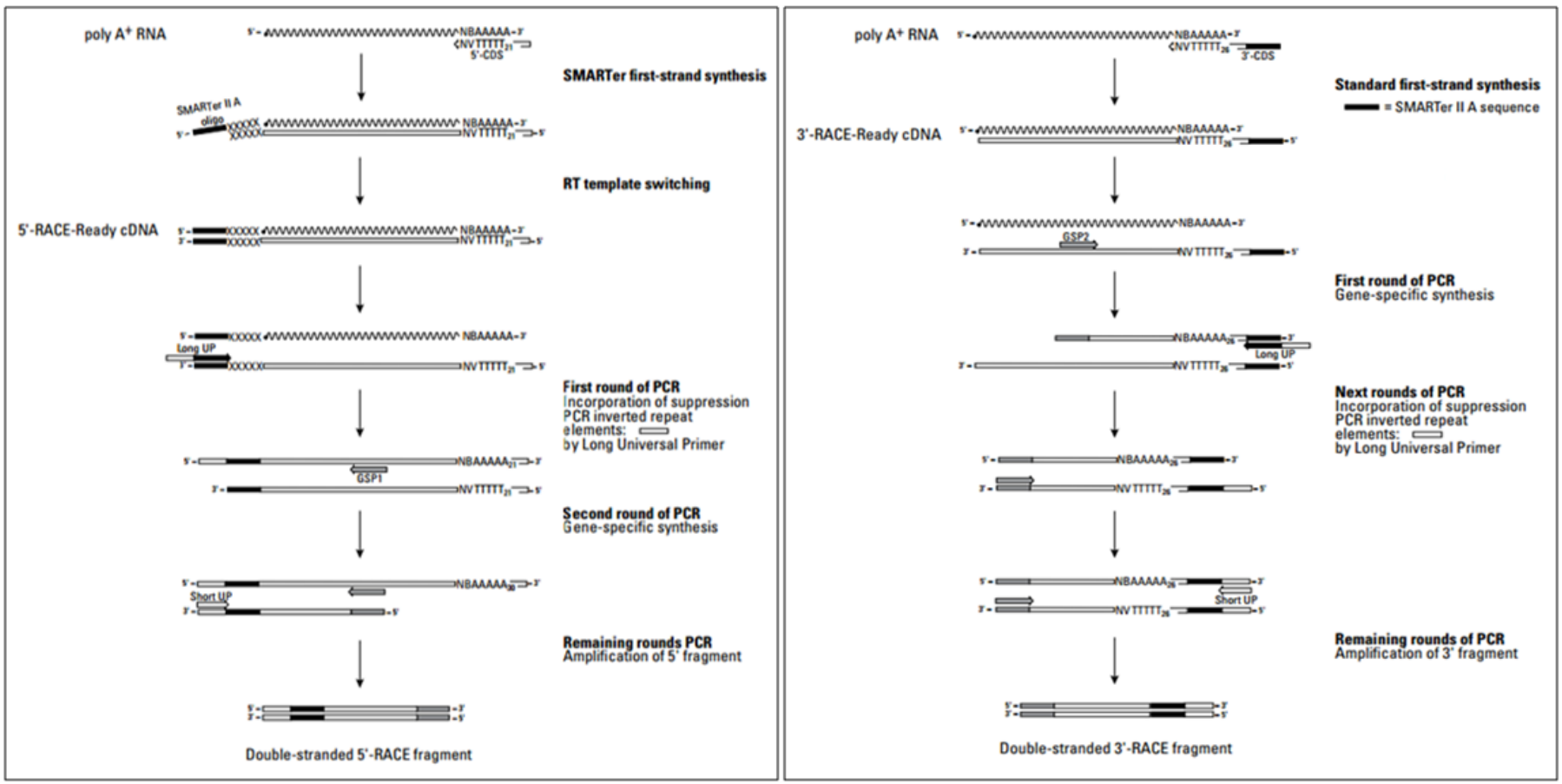

Figure 4.3 – The above image represents the flow chart of 5’-RACE and 3’-RACE work-flow (Clonetech).

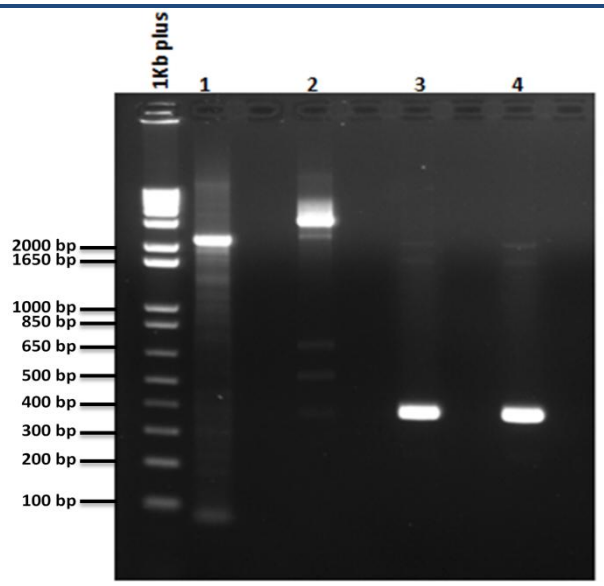

Figure 4.4 – Positive control RACE PCR experiment performed using the RACE ready cDNAs generated from the control mouse heart total RNA. Lane 1: 5'-RACE PCR generated 2.1 kb sized product. Lane 2: 3'-RACE PCR amplified a band with the size of approximately 3.1 kb. Lane 3: Internal control for 5'-RACE cDNA amplified an overlapping region of 380 bp. Lane 4: Internal control for 3'-RACE cDNA amplified 380 bp overlapping region.

4.3.2 - 5'- and 3'- *FAST-1* RACE cDNA synthesis

The *FAST-1* RACE experiment was performed using a human FRDA fibroblast cell line (GM04078). The procedure was performed exactly the same as positive control RACE PCR with the exception that RNA of the human FRDA cell line was used instead of mouse heart total RNA. For amplification of 5'-and 3'- *FAST-1* RACE cDNA, primers were designed in antisense and sense orientations, respectively. The SMARTer oligo sequence used to synthesise the cDNA was 5'–AAGCAGTGGTATCAACGCAGAGTACXXXXX–3', where 'X' represents the undisclosed base in the proprietary SMARTer oligo sequence. For primary and nested PCR amplifications, *FAST-1* specific primers were designed for both 5'- and 3'-RACE experiments. In addition, the primary PCR primers were designed to amplify the

overlapping region between the primers therefore this could act as a positive control. The full list of primary and nested PCR primers and their locations is listed below (Figure 4.5).

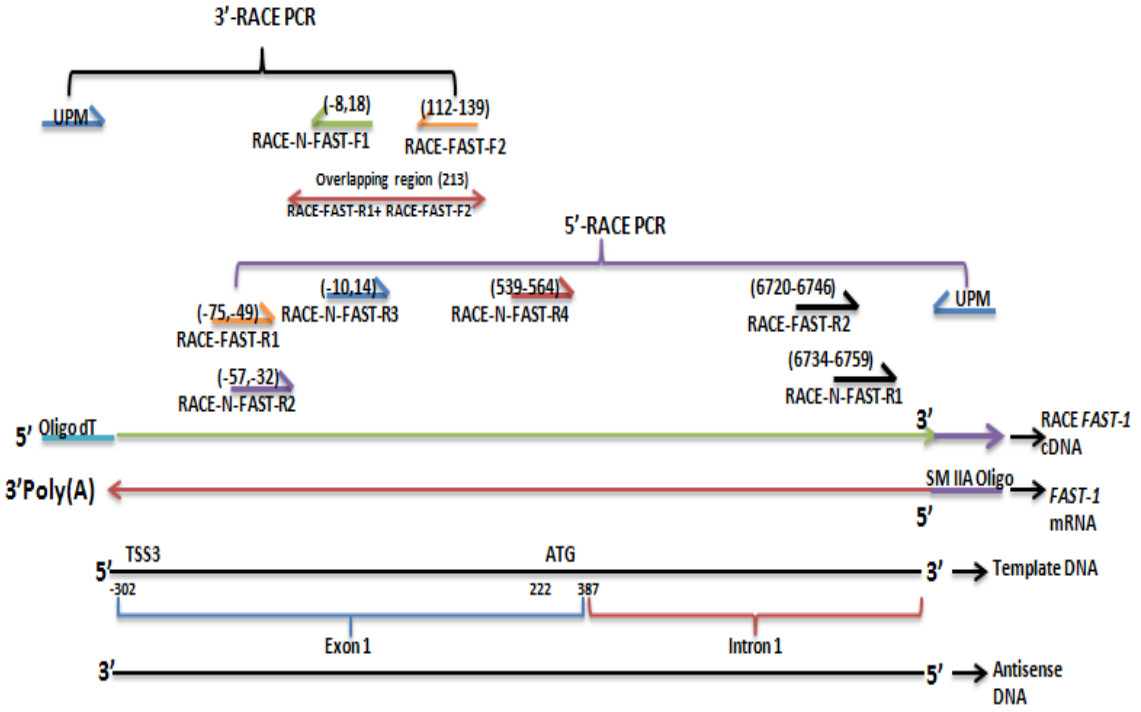

Figure 4.5 – Schematic representation of the positions of 5'-RACE and 3'-RACE PCR primers, numbered in relation to the main *FXN* transcription start site (TSS1) = +1. Primers RACE-FAST-F2, RACE-N-FAST-F1, RACE-FAST-R1, RACE-N-FAST-R2, RACE-N-FAST-R3 are designed from 5'-UTR region of *FXN* gene, RACE-N-FAST-R4, RACE-N-FAST-R1, RACE-FAST-R2 primers are designed from intron 1 region of *FXN* gene.

4.3.3 - 5'- and 3'- *FAST-1* RACE primary PCR

5'- and 3' *FAST-1* RACE cDNA samples were amplified using RACE-FAST-R1 and RACE-FAST-F2 primers, respectively, with the UPM primer common to both reactions. Positive control RACE PCR reactions were performed by amplifying both, 5'- and 3'- *FAST-1* RACE cDNAs using RACE-FAST-F2 and RACE-FAST-R1, which should amplify an overlapping region of 213 bp. Negative control RACE PCR reactions were also performed by using only either *FAST-1* specific primers or UPM primer. The PCR products were then run on a gel to verify

the band sizes. The 5' and 3' *FAST-1* RACE PCR reactions produced ladders of bands with different sizes. However, the positive control PCR reactions, which use both 5' and 3' *FAST-1* specific primers, generated only one single band with 213 bp size, representing the overlapping region of the PCR (Figure 4.6). Furthermore, the negative RACE PCR reactions performed using the UPM primer show smeared bands in the gel for both 5' and 3' RACE reactions. However, the gene-specific primer negative RACE PCRs showed no bands and therefore represented good negative controls (Figure 4.6).

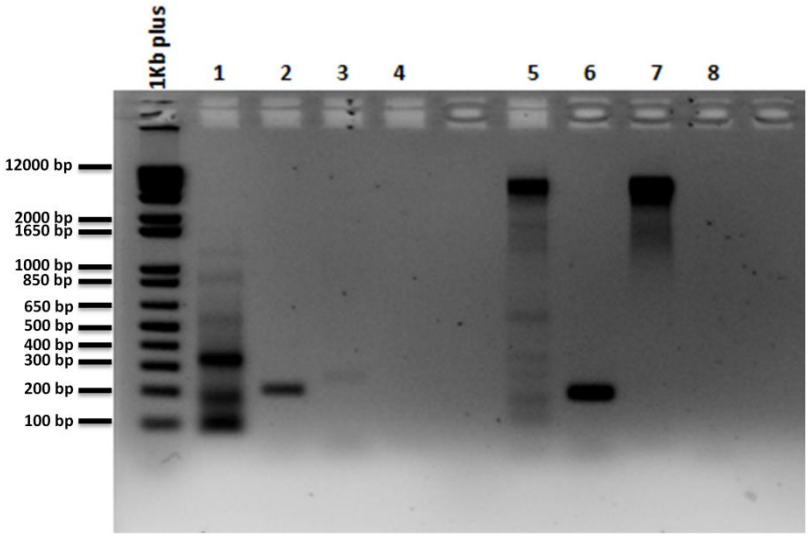

Figure 4.6 – Primary 3'- and 5'- *FAST-1* RACE PCR amplification. Lanes 1 & 5: 5'- and 3'- RACE primary PCR, Lanes 2 & 6: Positive control for 5'- and 3'- *FAST-1* RACE PCR reactions, Lanes 3 & 7: Negative control for 5'- and 3'- *FAST-1* RACE PCR reaction amplified using UMP primer only, Lanes 4 & 8: Negative control for 5'- and 3'- *FAST-1* RACE PCR reactions amplified by using only gene specific primer.

4.3.4 - Nested 3'- *FAST-1* RACE PCR

Since the primary 5'- and 3'- *FAST-1* RACE PCR reactions produced smeared bands, I then performed a "nested" PCR reaction using the NUP primer (nested universal primer A, supplied with the SMARTer RACE cDNA amplification kit) and a nested gene specific primer. For nested PCR amplification, the primary PCR product was diluted 50X with tricine EDTA buffer buffer. To identify the correct annealing temperatures, a gradient nested PCR was performed at three different temperatures, 68°C, 70°C and 72°C, using RACE-N-FAST-F1 and NUP primers. After analysing the PCR products in agarose gels, I found that 68°C and 70°C produced one band of approximately 500 bp in size, whereas 72°C did not produced any bands (Figure 4.8). Therefore, I performed all of the nested 3' *FAST-1* RACE PCRs at 68°C and the resultant PCR products were cut from agarose gels, followed by purification using the Nucleotrap gel extraction kit (section 2.3.9)

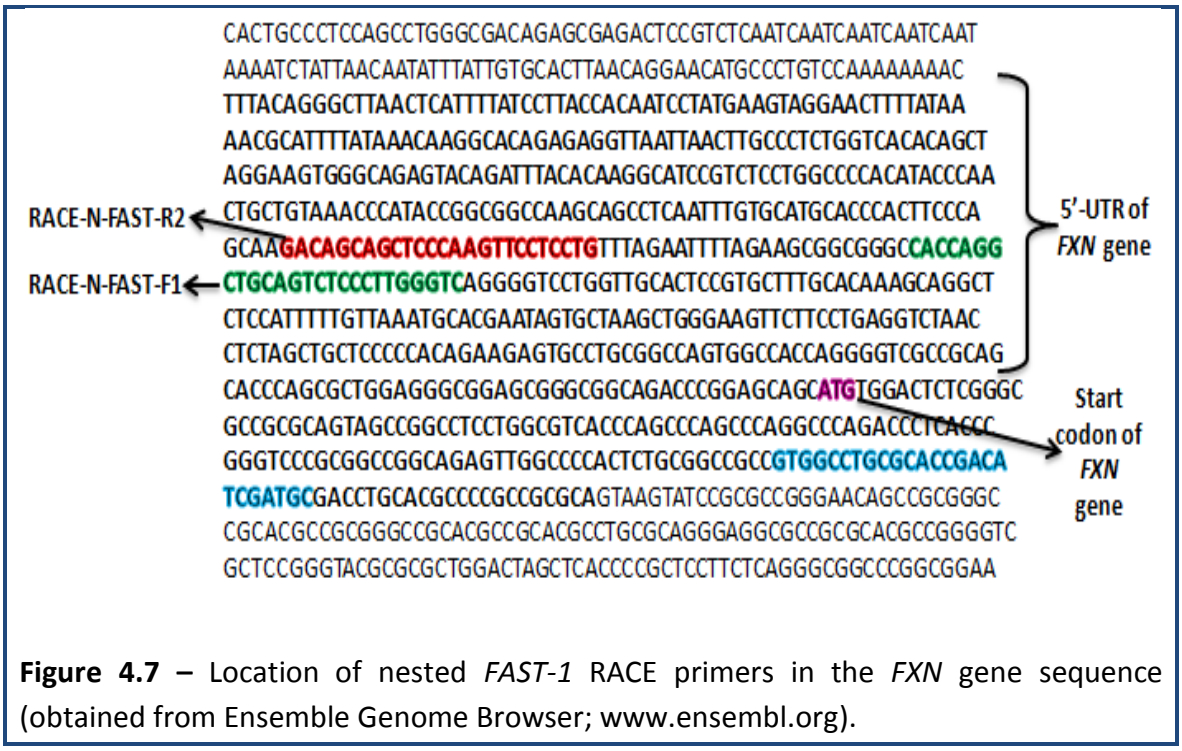

Figure 4.7 – Location of nested *FAST-1* RACE primers in the *FXN* gene sequence (obtained from Ensemble Genome Browser; www.ensembl.org).

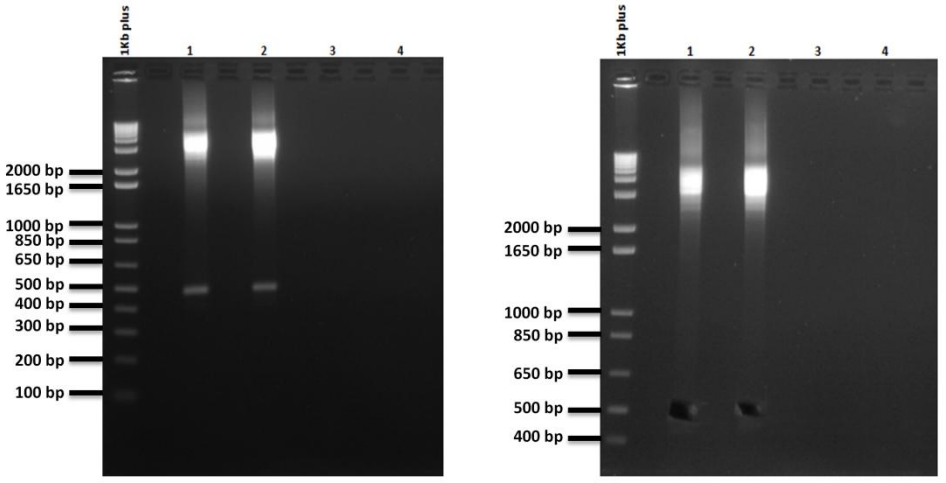

Figure 4.8 – Pre- and post-excision ethidium bromide-stained agarose gel showing the 3' *FAST-1* RACE nested PCR products. Lanes 1 -3: Nested PCR at 68°C, 70°C and 72°C, respectively. Lane 4: water control. The PCR products were then cut out and gel purified for cloning (right side image).

4.3.5 - Nested 5'- *FAST-1* RACE PCR

Nested 5'- *FAST-1* RACE PCR was performed initially with 50X diluted 5'- *FAST-1* RACE primary PCR product using RACE- FAST-R2, located in intron 1 region of the *FXN* gene (Figure 4.6), and NUP primers at 68°C annealing temperature. This PCR did not produce any band and therefore a series of dilutions (10X, 25X, 50X and 100X) were made and the PCR was performed again. Although a very faint band appeared with a 10X dilution, none of the other dilutions produced any bands (Figure 4.9). Therefore, to improve the efficiency of the PCR and to amplify the specific fragments, I designed more primers to perform nested 5'- *FAST-1* RACE PCR. RACE-N-FAST-R2 and RACE-N-FAST-R3 primers were designed from 5'-UTR region of the *FXN* gene, whereas RACE-N-FAST-R4, RACE-N-FAST-R1 and RACE-FAST-R2 primers were designed from intron 1 region of the *FXN* gene (Figure 4.5).

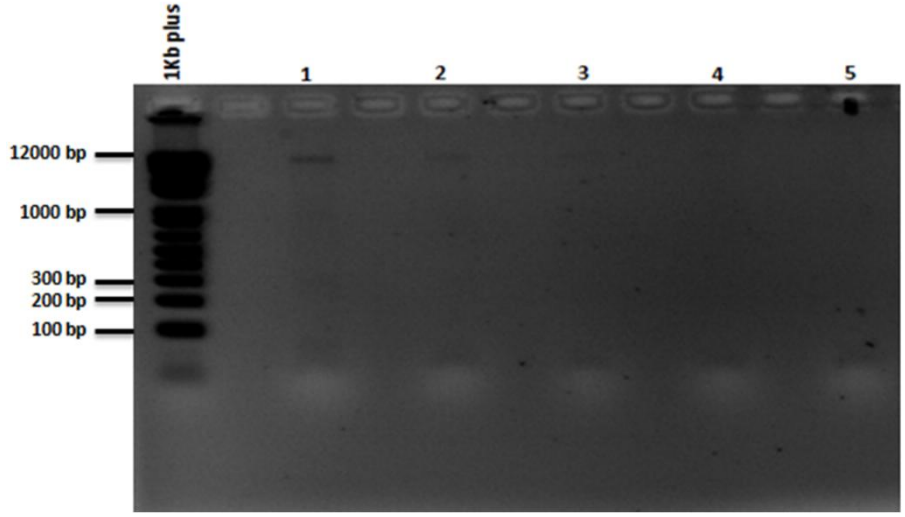

Figure 4.9 – Ethidium bromide-stained agarose gel showing the 5' *FAST-1* RACE nested PCR products. Lanes 1 -4: Nested PCR with 10X, 25X, 50X and 100X dilutions, respectively. 5: water control.

A series of 5' *FAST-1* RACE nested PCRs was performed by combining the NUP primer with one of the newly designed gene specific primers, such as RACE-FAST-R2, RACE-N-FAST-R1, RACE-N-FAST-R2, RACE-N-FAST-R3 and RACE-N-FAST-R4 at 68°C annealing temperature. After analysing the PCR products on agarose gels, I found that the RACE-N-FAST-R2 and NUP primer combination resulted in one band corresponding to a size between 200-300 bp. Similarly, RACE-N-FAST-R3 and NUP primers generated a band between 100-200 bp in size and no products were amplified with the other PCR primers, such as RACE-FAST-R2, RACE-N-FAST-R1 and RACE-N-FAST-R4 (Figure 4.10). The two nested 5'- *FAST-1* RACE PCR products, which were amplified by RACE-N-FAST-R2 and NUP primers and RACE-N-FAST-R3 and NUP primers, were cut from the gel followed by purification of the DNA by the Nucleotrap gel extraction kit (section 2.3.9).

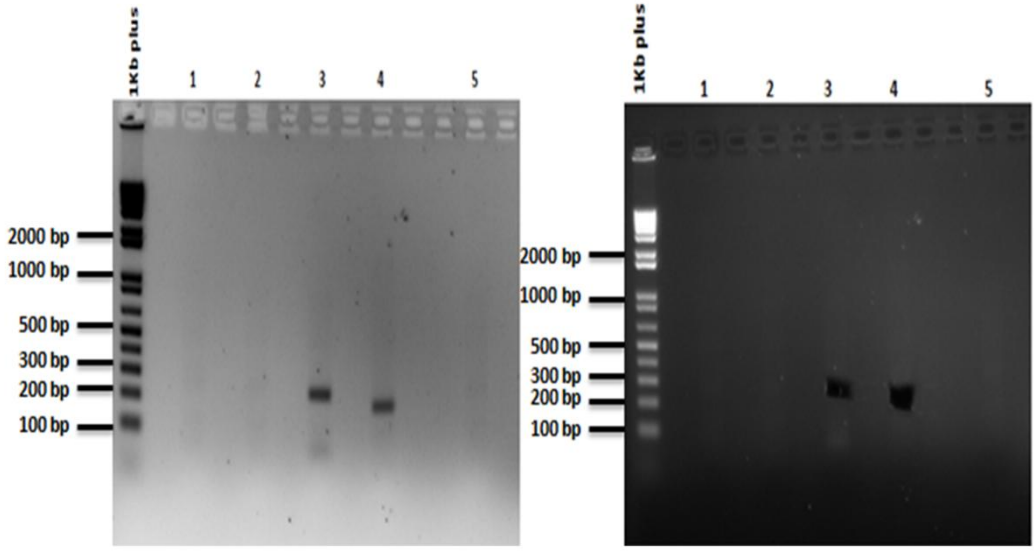

Figure 4.10 – Ethidium bromide-stained agarose gel showing the 5' *FAST-1* RACE nested PCR products. Lanes 1 -5: Nested PCR using a combination of NUP primer and one of the four gene specific primers, such as RACE-FAST-R2, RACE-N-FAST-R1, RACE-N-FAST-R2, RACE-N-FAST-R3 and RACE-N-FAST-R4, respectively. The PCR products were then cut and gel purified for cloning (right side image).

4.3.6 - Cloning of 3'- and 5'- RACE PCR products

After gel extraction, a nested 3'- *FAST-1* RACE PCR product with a size between 400 bp and 500 bp which was amplified by RACE-N-FAST-F1 and NUP primers, and two nested 5'- *FAST-1* RACE PCR products, a 200-300 bp product (amplified by RACE-N-FAST-R2 and NUP primers) and a second 100 bp-200 bp product (amplified by RACE-N-FAST-R3 and NUP primers), were cloned into pCR™4-TOPO® vector by the TOPO TA® cloning method (section 2.3.15). To extract plasmids containing the inserted gene of interest, 14 positive clones of 3' *FAST-1* RACE PCR and 12 clones of 5' *FAST-1* RACE PCR were picked and cultured overnight, followed by extraction of the plasmid DNA.

4.3.7 - Restriction enzyme digestion of 3'- *FAST-1* RACE recombinant plasmids with *Eco*RI

Prior to sending the samples for sequencing and to confirm the presence of correct sized inserts (PCR products), I digested 14 recombinant plasmids isolated from nested 3' *FAST-1* RACE clones by *Eco*RI restriction enzyme for 2 hours at 37°C. The *Eco*RI digested 3'-*FAST-1* RACE recombinant plasmids yielded correctly-sized products between 400 bp and 500 bp (which was amplified by RACE-N-FAST-F1 and NUP primers), when resolved on an agarose gel (Figure 4.11). Out of 14 recombinant plasmids, 8 plasmids (plasmids numbered 1, 2, 4 5, 8, 11, 12 and 13) were sequenced in both orientations using M13 forward (−20), and M13 reverse primers using the next generation DNA sequencing method (performed by Beckman Coulter Genomics).

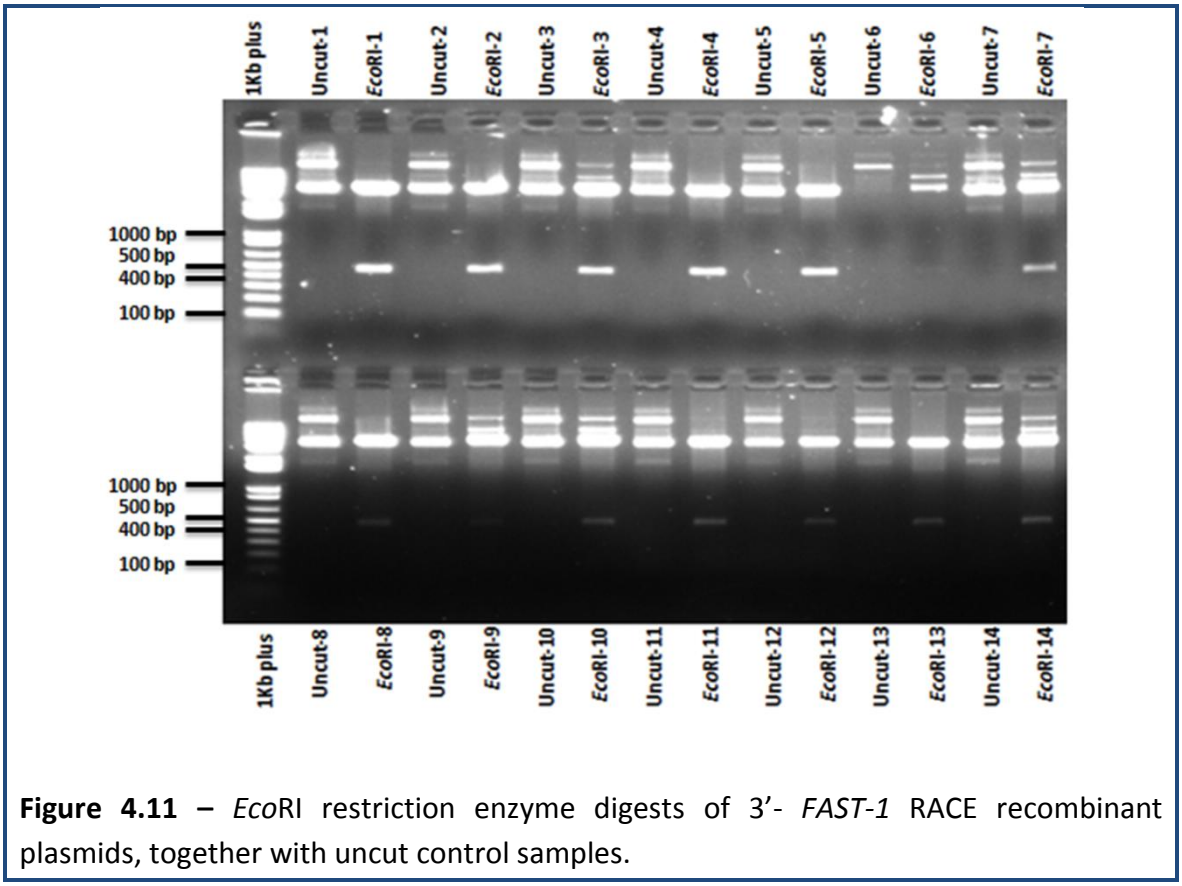

Figure 4.11 – *Eco*RI restriction enzyme digests of 3'- *FAST-1* RACE recombinant plasmids, together with uncut control samples.

4.3.8 - Restriction enzyme digestion of 5'- *FAST-1* RACE recombinant plasmids with *Eco*RI

The first set of twelve 5'- *FAST-1* RACE recombinant plasmids, which were generated by cloning a PCR product of 200 bp and 300 bp size, amplified by RACE-N-FAST-R2 and NUP primers, and a second set of twelve 5'- *FAST-1* RACE recombinant plasmids, which were generated by cloning a PCR product of 100 bp and 200 bp size, amplified by RACE-N-FAST-R3 and NUP primers, were digested with restriction enzyme *Eco*RI. After digestion, the products were run on an agarose gel and this confirmed that all of the plasmids yielded the correctly-sized products (Figure 4.12 & Figure **4.13**).

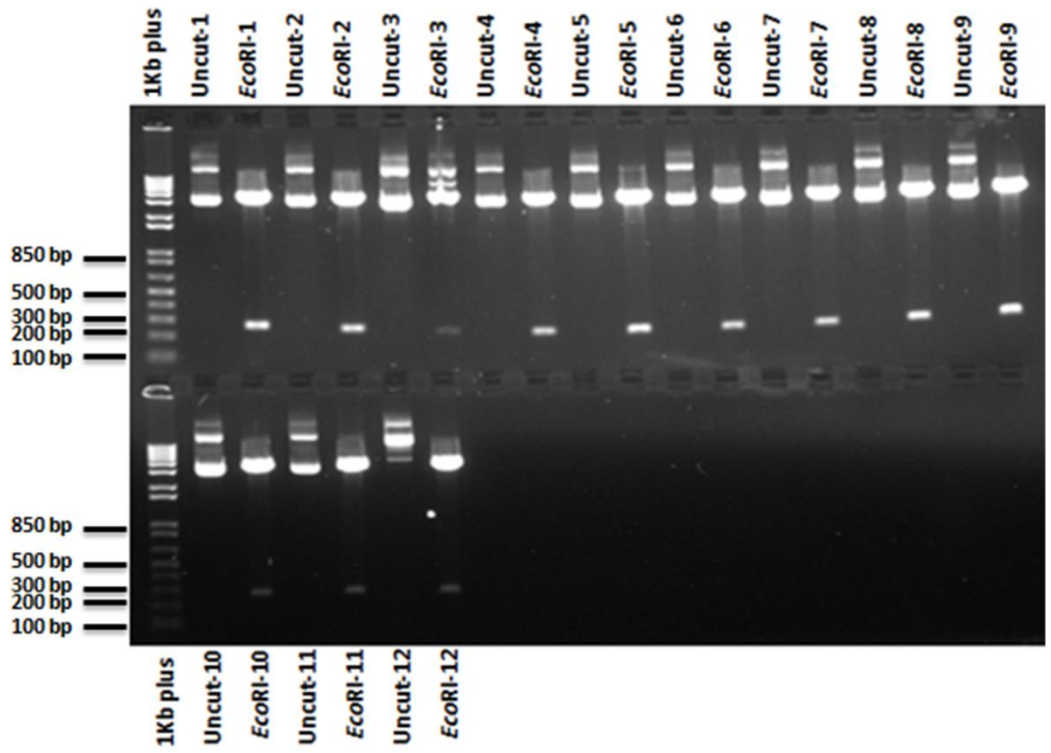

Figure 4.12 – *Eco*RI restriction enzyme digests of first set of 5'- *FAST-1* RACE recombinant plasmids (generated by cloning a PCR product of 200-300 bp size amplified with RACE-N-FAST-R2 and NUP primers), together with uncut control samples.

Out of twelve 5' *FAST-1* RACE recombinant plasmids, 4 plasmids from each set (4, 5, 6 and 7) were sequenced in both orientations using M13 forward (−20) and M13 reverse primers using the next generation DNA sequencing method (performed by Beckman Coulter Genomics).

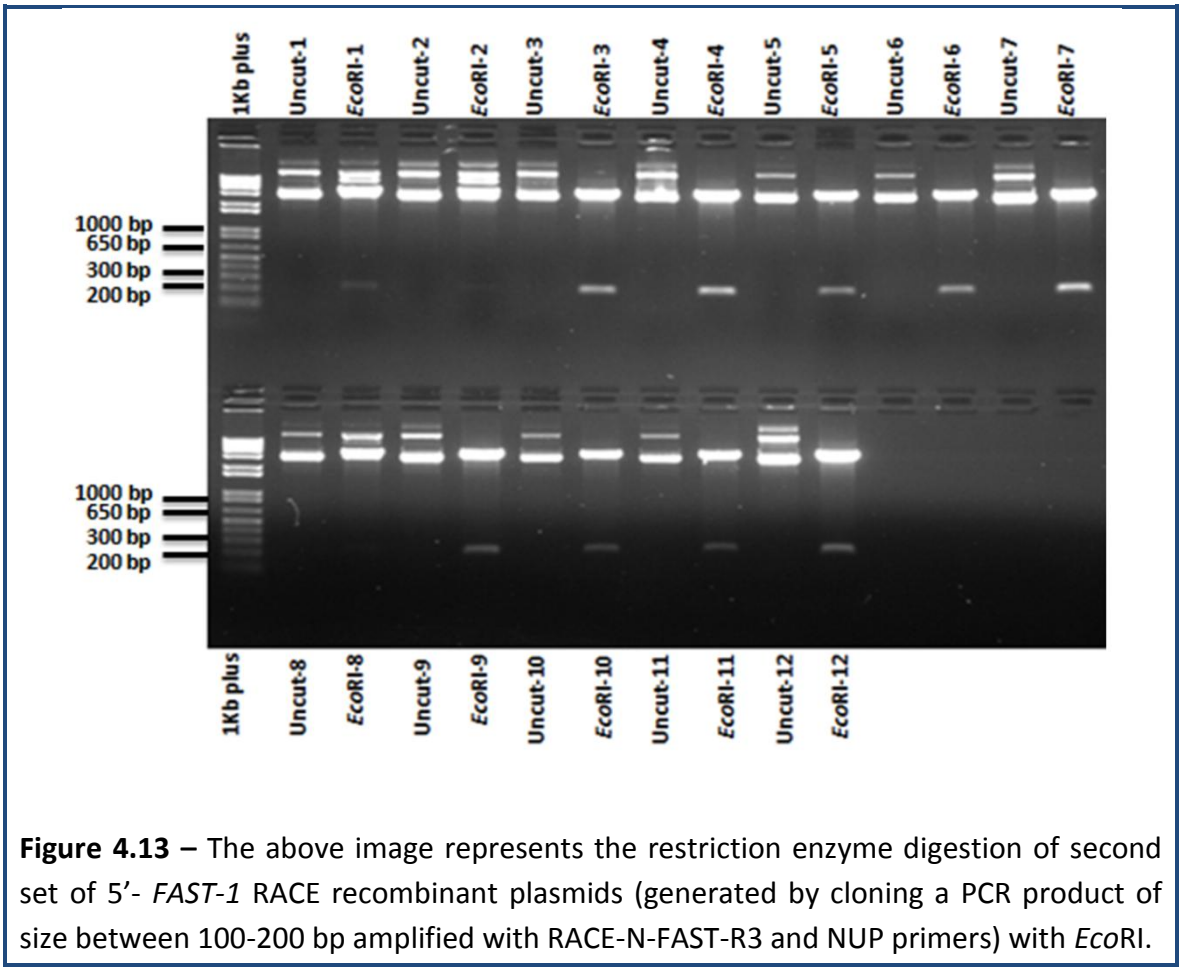

Figure 4.13 – The above image represents the restriction enzyme digestion of second set of 5'- *FAST-1* RACE recombinant plasmids (generated by cloning a PCR product of size between 100-200 bp amplified with RACE-N-FAST-R3 and NUP primers) with *Eco*RI.

4.3.9 - 3'- *FAST-1* RACE recombinant plasmid sequencing

Eight 3'- *FAST-1* RACE recombinant plasmids were sequenced in both orientations using two primers, M13 forward (−20), and M13 reverse. 3'- *FAST-1* RACE recombinant plasmids contained an insert which was amplified by RACE-N-FAST-F1 and NUP primers in a nested 3'- *FAST-1* RACE PCR. After reading the sequence and comparing with the reference

gene sequencing from Ensembl, I confirmed that the cloning and sequencing had been successful. However, the sequencing of 3'-*FAST-1* RACE recombinant plasmid performed in the orientation of NUP primer failed due to the presence of the poly T tail sequence.

The 3' *FAST-1* RACE sequence analysis revealed that the 3'-end of the *FAST-1* transcript was located at nucleotide position -359 bp (in relation to the main *FXN* TSS1 = +1) and the polyadenylation signal was identified at nucleotide position -341 bp to -346 bp. Furthermore, all eight 3' *FAST-1* RACE PCR clone sequences contained an additional 'T' nucleotide at position -329 bp on *FXN* gene locus compared to the genomic DNA sequence (Ensembl) (Figure 4.14). This could be due to a possible nucleotide polymorphism.

Figure 4.14 - 3'-*FAST-1* RACE recombinant plasmid (clone-1) sequencing performed in the orientation of the RACE-N-FAST-F1 primer.

4.3.10 - 5'- *FAST-1* RACE recombinant plasmid sequencing

The two sets of 5'- *FAST-1* RACE recombinant plasmids, which were generated by cloning two nested 5'- *FAST-1* RACE PCR products, a 200 bp-300 bp product (amplified by RACE-N-FAST-R2 and NUP primers) and a second 100 bp-200 bp product (amplified by RACE-N-FAST-R3 and NUP primers), were sequenced (4 clones from each set) in both orientations using M13 Forward (−20), and M13 Reverse primers.

After analysing the sequence, the 5'- *FAST-1* RACE experiment, which was performed to identify the 5'-end of the *FAST-1*, was deemed successful in both orientations. The sequence analysis revealed that the 5'-end of the *FAST-1* was located at nucleotide position 164 bp of the *FXN* gene locus (Figure 4.15 & Figure 4.16). From the results of 5' - and 3'- *FAST-1* RACE experiments, the 5'- and 3'- ends of *FAST-1* were mapped on genomic DNA sequence from nucleotide positions -359 to 164 of the *FXN* gene. Therefore, the total length of the full-length *FAST-1* transcript was found to be 523 bp in size.

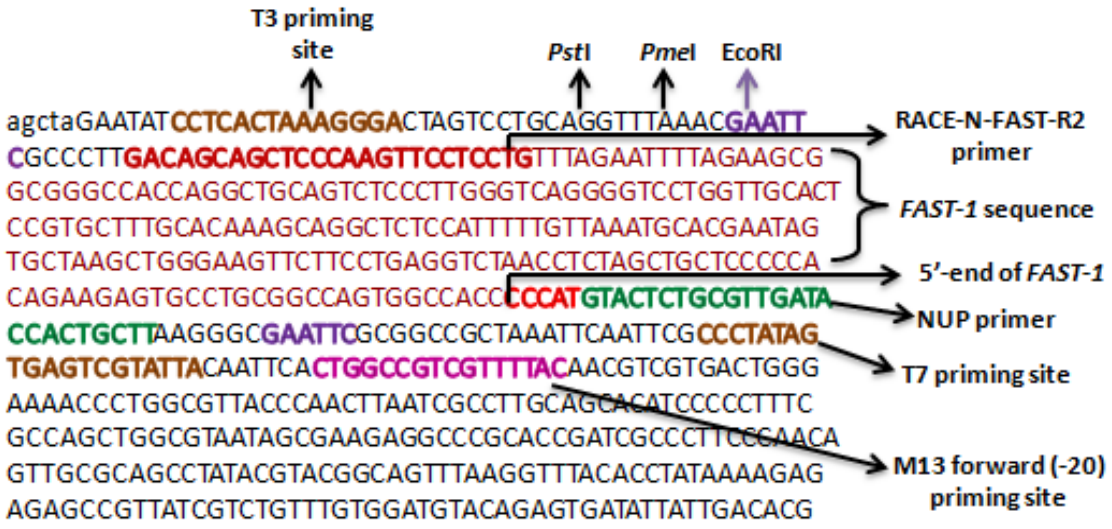

Figure 4.15 – Sequencing of a 5'-*FAST-1* RACE recombinant plasmid (generated by cloning a PCR product of 200 bp-300 bp size and amplified with RACE-N-FAST-R2 and NUP primers), performed in the orientation of RACE-N-FAST-R2 primer.

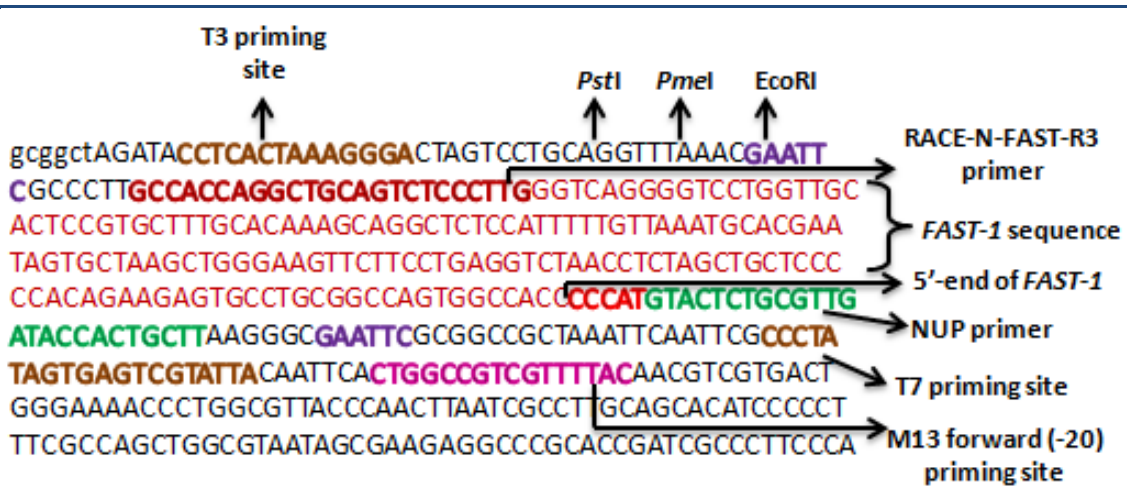

Figure 4.16 – Sequencing of a 5'-*FAST-1* RACE recombinant plasmid (generated by cloning a PCR product of size between 100-200 bp and amplified with RACE-N-FAST-R3 and NUP primers), performed in the orientation of RACE-N-FAST-R3 primer.

4.4 - Identification of Full-length a *FAST-1* transcript

After confirmation of the 5'- and 3'- ends of *FAST-1* by 5'- and 3'- RACE experiments, the total length of the *FAST-1* was calculated to be 523 bp in size. To confirm this further, I amplified the full length *FAST-1* by designing the primers at the end of 5'- and 3'-ends of *FAST-1* sequence (Figure 4.17) followed by cloning and sequencing.

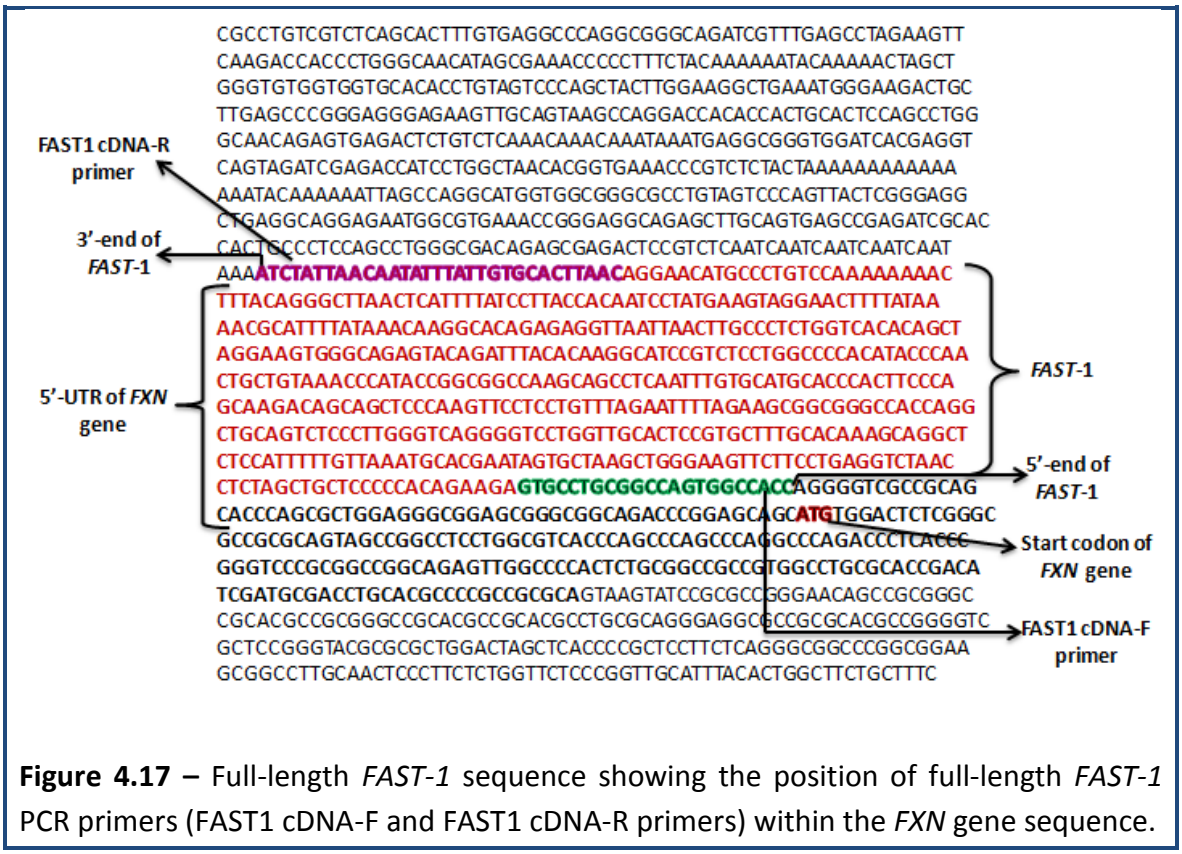

Figure 4.17 – Full-length *FAST-1* sequence showing the position of full-length *FAST-1* PCR primers (FAST1 cDNA-F and FAST1 cDNA-R primers) within the *FXN* gene sequence.

4.4.1 - Amplification of a full-length *FAST-1* sequence

The full-length *FAST-1* sequence was amplified with FAST1 cDNA-F (forward primer) and FAST1 cDNA-R (reverse primer). I used a human *FXN* BAC DNA sample (#clone 586E2), together with a non-*FXN* negative control BAC DNA sample, a non-FRDA control fibroblast cDNA sample and FRDA fibroblast cDNA sample, performing the PCR amplification at 54°C

annealing temperature. The resultant PCR products were run on a 1% agarose gel, revealing a single band of approximately 500-600 bp in size from each of the *FXN* BAC DNA, non-FRDA control fibroblast cDNA and FRDA fibroblast cDNA samples, but no band was identified in the non-*FXN* negative control BAC DNA sample. The single band from the *FXN* BAC DNA was then cut from the gel, followed by purifying the product using the Nucleotrap gel extraction kit (Figure 4.18).

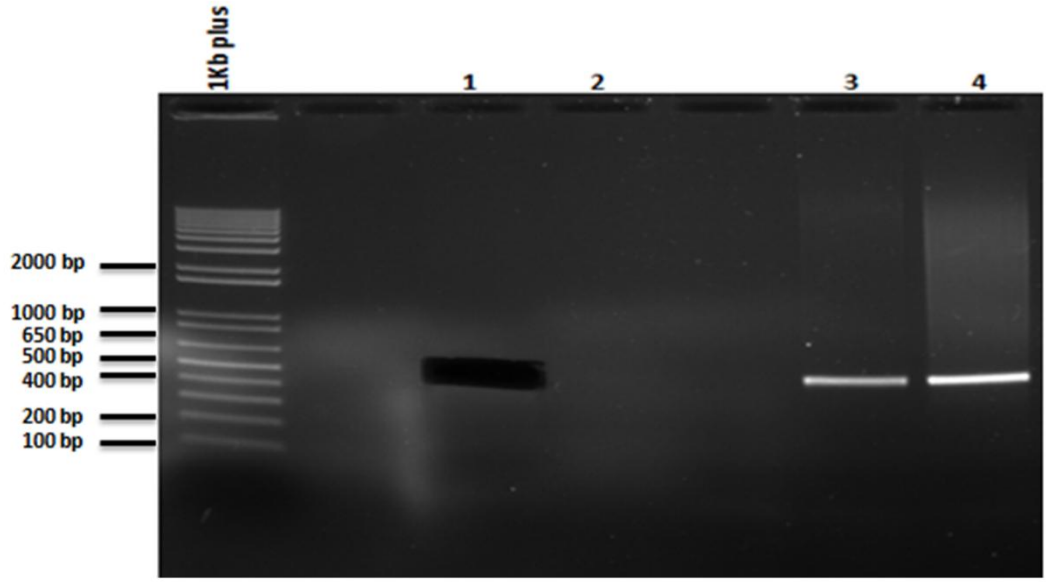

Figure 4.18 – Amplification of a full-length *FAST-1* PCR product with FAST1 cDNA F and FAST1 cDNA R primers from: 1) BAC DNA (586E2), 2) negative BAC DNA, 3) control fibroblast cDNA and 4) FRDA fibroblast cDNA.

Before cloning, 2µl of purified full length *FAST-1* PCR product was analysed on a 1.2% agarose gel to confirm the correct sized product (Figure 4.19).

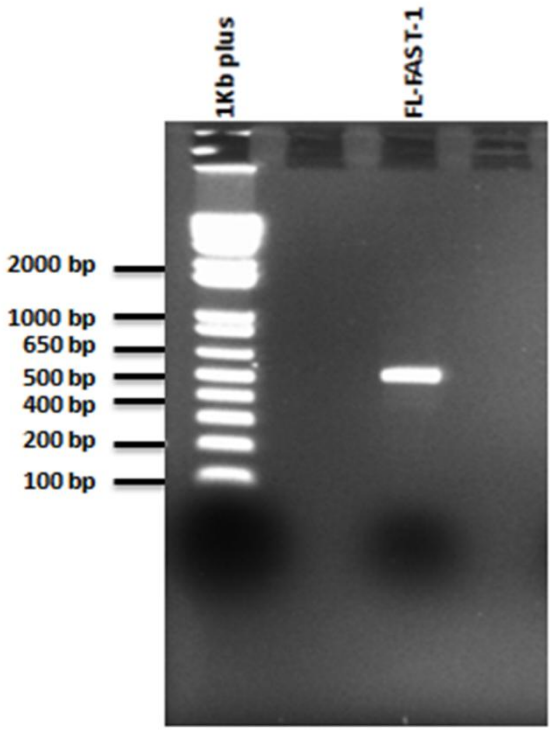

Figure 4.19 – A gel purified full length *FAST-1* PCR product analysed on a 1.2% agarose gel for size confirmation

4.4.2 - Cloning of a full-length *FAST-1* sequence

The gel purified full length *FAST-1* PCR product was cloned into pCR™4-TOPO® vector by TOPO TA® cloning, followed by transferring in to the DH5α™ comptent cells as previously described (section 2.3.15). To analyse positive clones, out of 95 colonies grown on agar plate, 12 colonies were picked and cultured overnight in 5ml of LB medium containing 50µg/ml of ampicillin and 50µg/ml of kanamycin, followed by plasmid DNA extraction using the QIAprep spin miniprep kit.

4.4.3 - Restriction enzyme digestion of full length *FAST-1* recombinant plasmids with *Eco*RI

To identify accurately-sized inserts prior to sequencing, the recombinant plasmids isolated from 12 full-length *FAST-1* clones were subjected to enzymatic digestion with *Eco*RI. This clearly indicated the presence of correctly-sized inserts in all 12 clones (Figure 4.20). Four, out of 12 full-length *FAST-1* recombinant plasmids (1, 2, 3 & 4) were sequenced by the next generation DNA sequencing method (performed by Beckman Coulter Genomics).

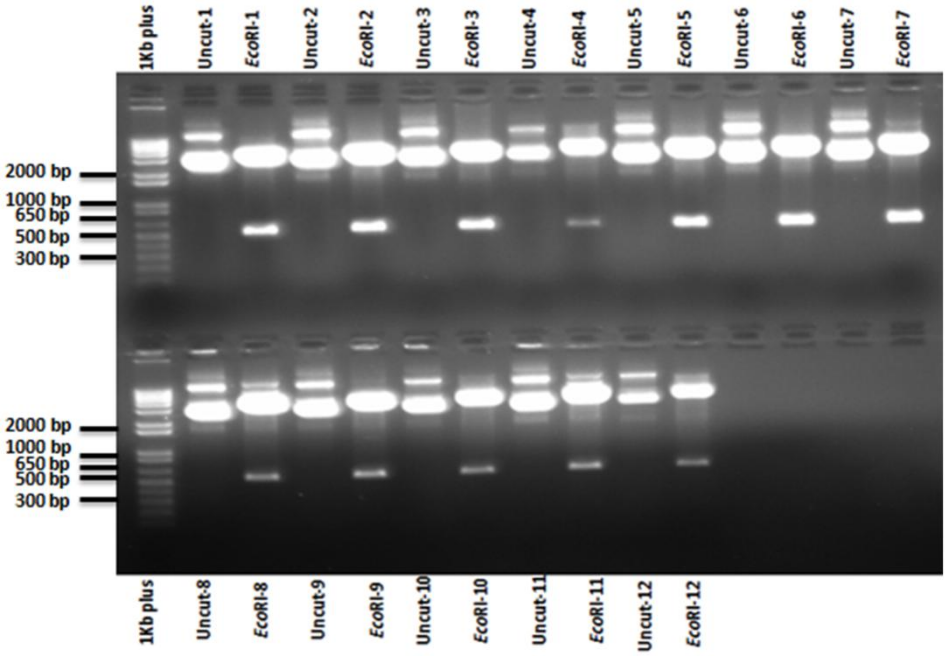

Figure 4.20 – *Eco*RI restriction enzyme digests of full length *FAST-1* recombinant plasmids, together with uncut controls.

4.4.4 - Sequencing of full length *FAST-1* recombinant plasmids

Four full-length *FAST-1* recombinant plasmids were sequenced in both orientations using two primers, M13 forward (−20) and M13 reverse. Sequence analysis confirmed the full-length *FAST-1* clone to be 523 bp in size, 164 bp to -359 bp region (Figure 4.21 & Figure 4.22). The full length *FAST-1* transcript occupies the region of CTCF, two *FXN* transcription factors, serum response factor (SRF) & activating protein 2 (AP2), and extends into the poly A region of the *FXN* gene.

Figure 4.21 – Full-length *FAST-1* recombinant plasmid (clone 1), sequenced in the orientation of FAST1 cDNA F primer.

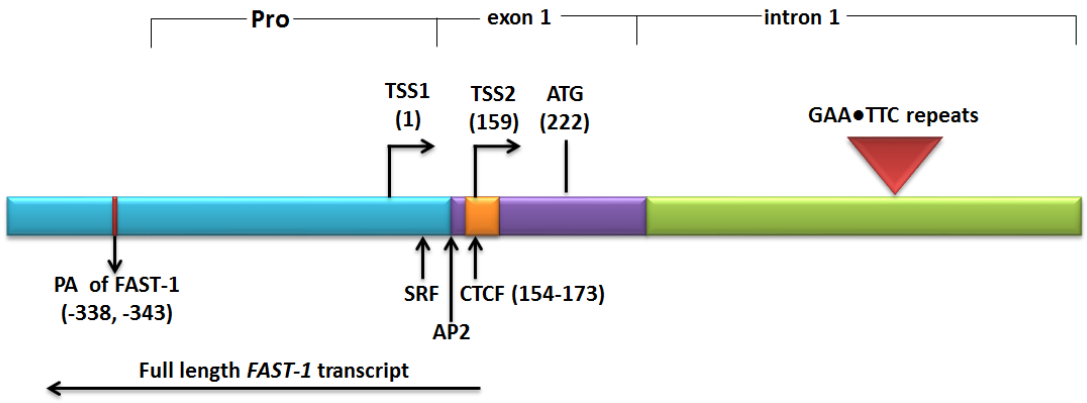

Figure 4.22 – The 5' end of the *FXN* gene showing the location of the full length *FAST-1* with 523 bp long. It also had a polyadenylation site located between -338 and -343. TSS1 - Transcription start site 1, TSS2- Transcription start site 2, ATG-Translation start site, CTCF-CCCTC binding factor, SRF-Serum response factor, AP2-Activating protein 2, PA-Poly adenylation, Pro-promoter region.

4.5 - Discussion

Although antisense transcription has been known to occur in prokaryotes for many years, the widespread occurrence of antisense transcripts in humans and mice has only recently been documented. The prevalence of lncRNAs and NATs has been reported in a variety of organisms. While a consensus has yet to be reached on their global importance, an increasing number of examples have been shown to be functional, regulating gene expression at the transcriptional and post-transcriptional level. Similarly, the mechanism of *FXN* gene silencing in FRDA patients is also not fully understood. However, the leading hypothesis, based on the recent therapeutic approaches, is the occurrence of epigenetic changes around the GAA repeats. Additionally, the *FXN* gene has also been reported to contain an antisense transcript, *FAST-1* (De Biase *et al.* 2009), which may be involved in *FXN* gene silencing by yet an unknown mechanism.

Since there was very limited information available about the *FAST-1*, including the lack of key information, such as length and location, I first investigated Northern blot analysis in an attempt to identify the *FAST-1* size. However, this Northern blotting technique was only partially successful, revealing two different sized bands and uncertainty if either band accurately corresponds to *FAST-1*. Therefore, in an effort to further characterise *FAST-1*, I used RACE technology. The RACE experiment was initiated by speculating that the *FAST-1* RNA may have a poly(A) tail. The 3'- and 5'-RACE CDS primer A in the RACE PCR kit contains a poly(T) tail and these primers should only anneal to any *FAST-1* RNA if it contains a poly(A) tail. This process then initiates the reverse transcription and thus produces 3'- and 5' *FAST-1* RACE cDNA. Based on my results, I confirmed that the *FAST-1* transcript possesses a poly(A) tail. Due to the sensitivity and the complex nature of the RACE experiment, I

optimised several approaches to make sure that this technique is reproducible. One of these parameters is the annealing temperature of the RACE PCR. Secondly, the concentration of the cDNA is also considered essential.

The primary *FAST-1* RACE PCR using gene specific primers resulted in smeared bands and or multiple ladder bands in both 3'- and 5'- *FAST-1* RACE PCRs. However, simultaneous amplification of positive controls produced one band at 213 bp size, indicating efficiency of the RACE PCR. Since the primary PCR failed to produce a single specific band, which could have been due to low abundance of the *FAST-1* transcript, I then performed nested PCR. The nested PCRs of 3'- and 5'- *FAST-1* RACE experiments produced single bands. These PCR products were then gel purified and cloned for the sequencing analysis. After cloning and sequencing the relatively abundant fragment obtained from the PCR of 3'- and 5'-*FAST-1* RACE, I was able to detect the expected *FAST-1* hybrid sequence. From the sequencing results of 5'- and 3'- *FAST-1* RACE recombinant plasmids, the 3'- and 5'- ends of *FAST-1* were mapped on genomic DNA sequence at nucleotide positions of -359 and 164 of the *FXN* gene and the total length of *FAST-1* was found to be 523 bp in size. A poly(A) signal was also identified in the *FAST-1* sequence at nucleotide positions of -341 to -346. In addition, the full length *FAST-1* transcript (523 bp) also corresponds to one of the Northern blotting technique results where I identified a band at approximately 500 bp size, indicating that the Northern blotting may also have correctly identified the same full-length *FAST-1* transcript. Furthermore, due to the presence of a poly(A) tail in the *FAST-1* transcript, it could be more intriguing to conclude that *FAST-1* RNA might be an mRNA species.

In conclusion, by performing 3'- and 5'-RACE experiments, I have identified a full

length *FAST-1* transcript, which starts from the 5'-UTR region of the *FXN* gene. Importantly, the 5'-end of the *FAST-1* occupies the recently reported CTCF binding site in the *FXN* 5'UTR region (De Biase *et al.* 2009; Al-Mahdawi *et al.* 2013), suggesting the possible regulation of *FAST-1* expression by CTCF binding and/or involvement of *FAST-1* in epigenetic changes that occur at the start of the *FXN* gene.

Chapter 5 - Quantification of *FAST-1* by qRT-PCR

5.1 - Introduction

My previous results investigating *FAST-1* transcripts by Northern blot analysis and RACE experiments supported the existence of *FAST-1* at the *FXN* locus. However, to develop a more robust technique for the identification and quantification of *FAST-1* levels, I have used quantitative real-time PCR (qRT-PCR) approach. For *FAST-1* mRNA quantification, strand-specific cDNA, synthesised by a FAST RT primer, was used. Whereas for *FXN*, GAPDH/Gapdh and HPRT/Hprt quantifications, cDNA synthesised by oligo$(dT)_{20}$ primer was used. *FAST-1* mRNA levels were quantified in human fibroblasts, *FRDA* YAC transgenic mouse cells (fibroblasts and NSCs) and tissues (brain, cerebellum and liver), FRDA BAC transgenic mouse tissues (brain, cerebellum and liver), genetically modified mice (Bmi-1 & HP1γ) tissues (brain, cerebellum and liver), diazoxide treated FRDA YAC transgenic mouse tissues (brain and liver) and histone methyltransferase (HMTase) inhibitor-treated mouse fibroblasts.

5.2 - Identification and quantification of *FAST-1* in human fibroblasts

Total RNA was isolated from three normal (GM04503, GM07492, GM08399) and eight FRDA fibroblast cells (GM03816, GM04078, GM03665, FA1, FA2, FA3, FA4, FA5) by the Trizol method. The quality and concentration of the RNA samples were checked by NanoDrop™ 2000c spectrophotometry and agarose gel electrophoresis, respectively (Figure 5.1). To remove genomic DNA contamination, RNA samples were then subjected to DNase I treatment using amplification grade DNase I (Invitrogen™). Untreated RNA and DNase I-treated RNA was amplified using FAST F1 and FAST R1 primers and the PCR products were

analysed on a 2% agarose gel. Untreated RNA amplified the correctly-sized PCR products (207 bp), whereas DNase I treated RNA samples did not show any product. This indicates that the DNase I treatment worked successfully; hence no genomic DNA contamination was present in the final DNase-treated RNA samples (Figure 5.2).

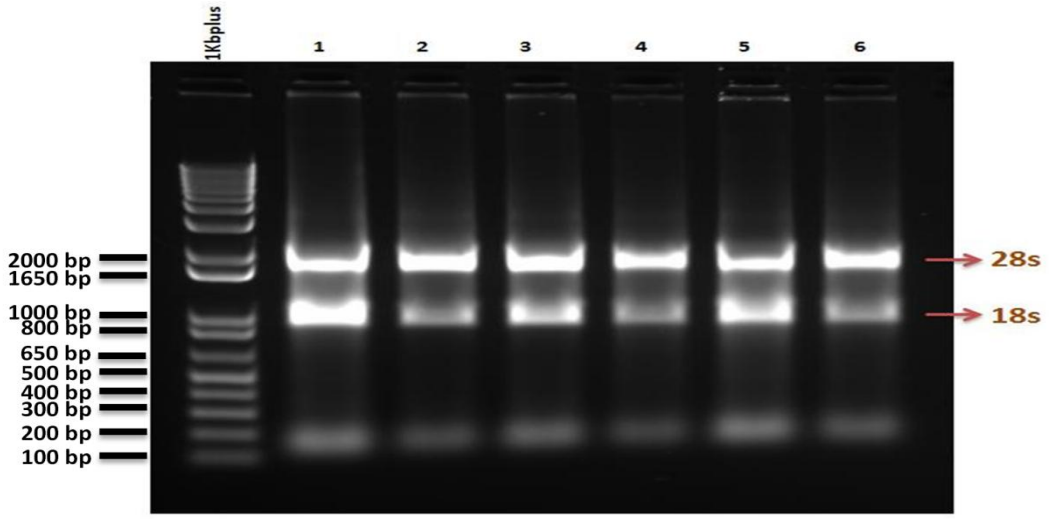

Figure 5.1 – The above typical image represents the RNA analysed on a 1% agarose gel. Lanes: 1-3 control and 4-6 FRDA samples.

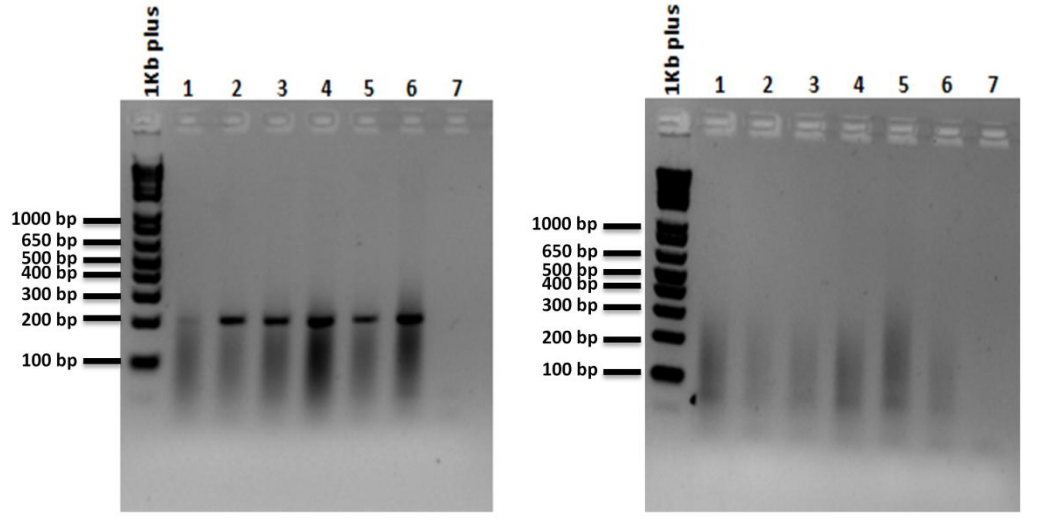

Figure 5.2 – The *FAST-1* primary PCR of untreated RNA (A) and DNase l-treated RNA (B). Lanes: 1-3 control, 4-6 FRDA, and 7-water.

Afterwards, the RNA was converted into cDNA by following the method described in De Biase *et al.* (2009). Briefly, a strand specific cDNA was synthesized from 1µg of DNase I treated RNA by Quantitect reverse transcription kit (Qiagen) by using a FAST RT primer. The resultant cDNA was then amplified using FAST F1 and FAST R1 primers, which should amplify a PCR product of 207 bp in size. However, none of the samples produced an expected band when the PCR products were analysed on a 2% agarose gel (Figure 5.3). However, a nested PCR using 2µl of 5X diluted primary PCR products (amplified by FAST F1 and FAST R1 primers) with N-FAST-F1 and N-FAST-R1 primers resulted in PCR products of the correct size (183 bp) (Figure 5.3).

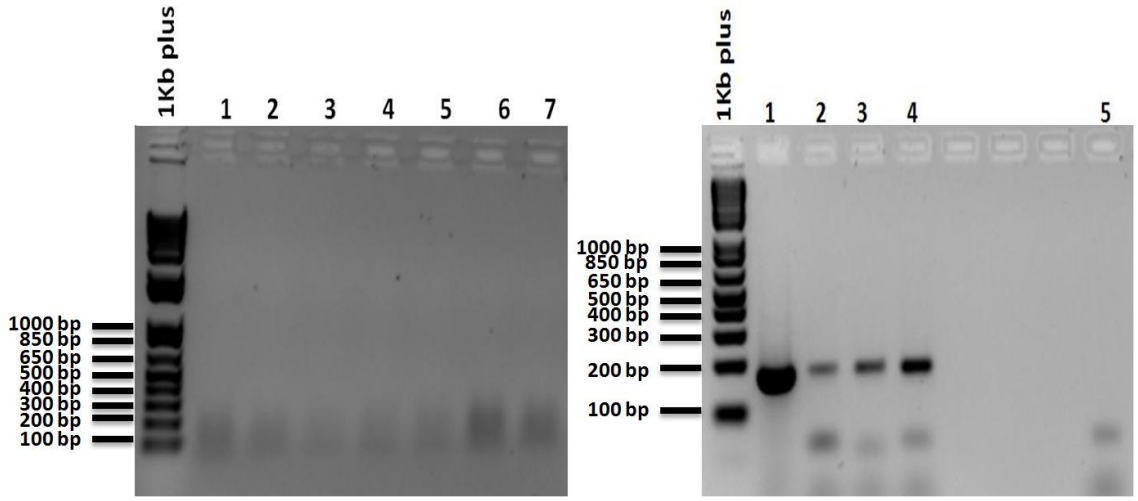

Figure 5.3 – Primary PCR (A) and nested PCR (B) of *FAST-1* using FAST F1 and FAST R1, and N-FAST-F1 and N-FAST-R1 primers, respectively. PCR products were run on a 2% agarose gel. A) Lanes: 1-3 control, 4-6 FRDA, and 7-water. B) Lanes: 1-2 control, 3-4 FRDA, and 5-water.

Although I have used exactly the same steps and reagents as described in De Biase *et al.* (2009) paper to amplify the *FAST-1*, I did not managed to amplify *FAST-1* by the primary PCR. Therefore, I optimised the technique using other approaches. Firstly, the RNA was extracted by the Trizol® method from the control and FRDA primary fibroblasts. The quality and concentration of the RNA were checked and any genomic DNA contamination was eliminated from the RNA samples by DNase I treatment. DNase I treated RNA samples were then used to synthesise the cDNA by cloned AMV first-strand cDNA synthesis using FAST RT primer. To confirm that *FAST-1* amplification was not from endogenous priming, cDNA synthesis was performed without FAST RT primer. After the cDNA synthesis, a primary PCR was performed using 2µl of cDNA (100ng/µl) with the N-FAST F2 and FAST R1 primer set which amplify a PCR product of 88 bp size and the resultant PCR products were run on a 2% agarose gel. The results indicated that only cDNA synthesised by the FAST RT primer amplified the *FAST-1* transcript, whereas the cDNA synthesised without the FAST RT primer resulted in no products confirming that the *FAST-1* amplification is not produced from endogenous priming (Figure 5.4).

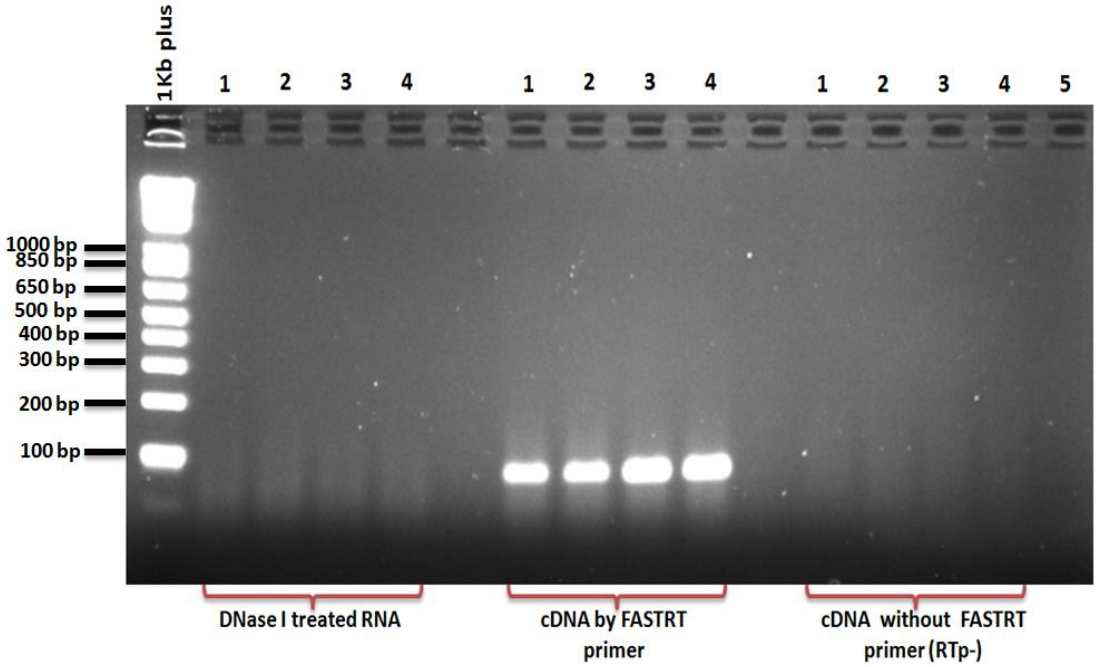

Figure 5.4 – PCR amplification of DNase I-treated RNA, cDNA prepared by FAST RT primer and cDNA prepared without FAST RT primer. Lanes: 1-2 control, 3-4 FRDA, and 5-water.

To quantify the *FAST-1* and *FXN* mRNA levels in human primary fibroblasts of normal and FRDA cells, qRT-PCR measurements were performed as described (section 2.17). To account for possible differences in gene expression efficiency or RNA amounts, the Ct values obtained for *FAST-1* and *FXN* were normalized to *GAPDH* and *HPRT* (endogenous control). For quantification of *FAST-1* (an antisense RNA), cDNA synthesised by strand specific FAST RT primer was used, whereas for *FXN*, *GAPDH* and *HPRT* quantifications, cDNA synthesised by oligo$(dT)_{20}$ primer was used. Since *FAST-1* has low levels of expression and to reduce the Ct values of *FXN*, *GAPDH* and *HPRT* closer to the *FAST-1* values, different dilutions of cDNA were used for qRT-PCR (1:5 dilution for *FAST-1*, 1:10 for *FXN* and 1:100 dilution for *GAPDH* and *HPRT*). Each PCR-reaction was performed in triplicate on a MicroAmp optical 96-well PCR plate (Applied Biosystems) and repeated at least twice. The passage numbers of all

primary fibroblasts were closely matched throughout all mRNA quantification experiments to avoid any possible cell culture variability.

QRT-PCR analysis of human fibroblasts clearly indicated increased levels of *FAST-1* and reduced levels of *FXN* expression in FRDA fibroblasts. *FAST-1* expression was increased by 41% in FRDA fibroblasts ($p<0.01$) (Figure 5.5), consistent with previous reports (De Biase *et al.* (2009). *FXN* levels were reduced by 77% ($p<0.001$) in fibroblasts from FRDA patients compared to fibroblasts derived from unaffected individuals (Figure 5.5).

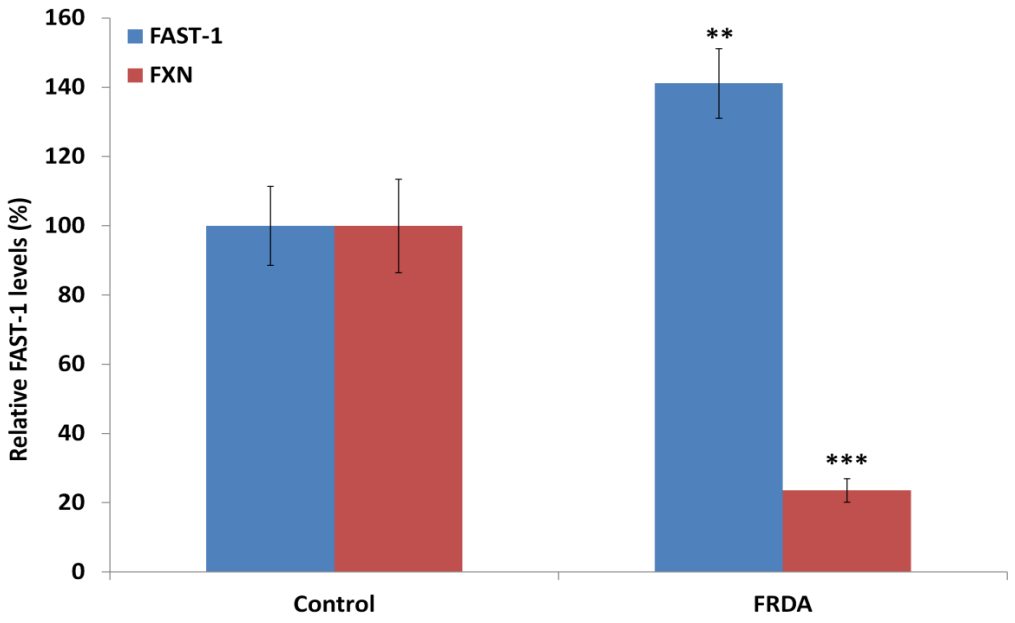

Figure 5.5 – Relative *FXN* and *FAST-1* mRNA levels in unaffected and FRDA patient derived fibroblasts (n=3-8), **p<0.01, ***p<0.001, error bars ± 1 SEM.

5.3 - Quantification of *FAST-1* levels in FRDA YAC transgenic mouse cells

To investigate the FRDA molecular disease mechanisms and therapy, our group has previously established three human *FXN* YAC transgenic mouse models: Y47R, containing normal-sized $(GAA)_9$ repeats, YG8R which initially contained expanded GAA repeats of 90 and 190 units, but which have subsequently been bred to now contain expanded GAA repeats of 120 and 220 units and YG22R, whose GAA repeats expanded from 190 units to 170 - 260 units, respectively (Al-Mahdawi *et al.* 2004). YG8sR (YG8R with a small GAA band) line was recently generated by crossbreeding YG8R with Y47R containing a contraction of GAA repeat sequence to a single band with 120 GAA repeat units.

To understand the mechanism underlying *FXN* gene silencing with regards to the *FAST-1*, we have used primary fibroblast cells and NSCs derived from Y47R, YG8R, YG22R and YG8sR mouse models. *FAST-1* and *FXN* mRNA expression levels in mouse fibroblasts and NSCs were quantified by qRT-PCR. The Y47R values were used as a calibrator for the relative YG8R, YG22 and YG8sR values. QRT-PCR analysis was performed using *Gapdh* and *Hprt* as endogenous controls and experiments were carried out in triplicate.

QRT-PCR analysis of mouse fibroblasts identified increased levels of *FAST-1* and reduced levels of *FXN* expression in YG8R and YG22R fibroblasts compared to the Y47R fibroblasts. These findings are in agreement with our pervious results using human fibroblasts. However, YG8sR fibroblasts showed reduced levels of both *FXN* and *FAST-1* compared to Y47R. In YG8R and YG22R fibroblasts, *FAST-1* expression levels were increased by 80% ($p<0.05$) and 220% ($p<0.01$), respectively, whereas *FXN* levels were reduced by 39% ($p<0.001$) and 27% (ns), respectively, compared to Y47R fibroblasts. However, in YG8sR fibroblasts both *FAST-1* and *FXN* levels were reduced by 31% ($p<0.05$) and 88% ($p<0.001$),

respectively (Figure 5.6).

The levels of *FAST-1* expression in YG8sR and YG8R NSCs were reduced by 43% ($p<0.05$) and 37% (ns), respectively, and *FXN* levels were decreased by 84% ($p<0.001$) and 42% ($p<0.01$), respectively, compared to Y47R NSCs. In YG22R NSCs, *FAST-1* expression was increased non-significantly by 40% and *FXN* expression was decreased by 35% ($p<0.05$) (Figure 5.6). Since there is differential expression of *FAST-1* in different cell lines it is speculated that *FAST-1* may be influenced by cell type-specific effects.

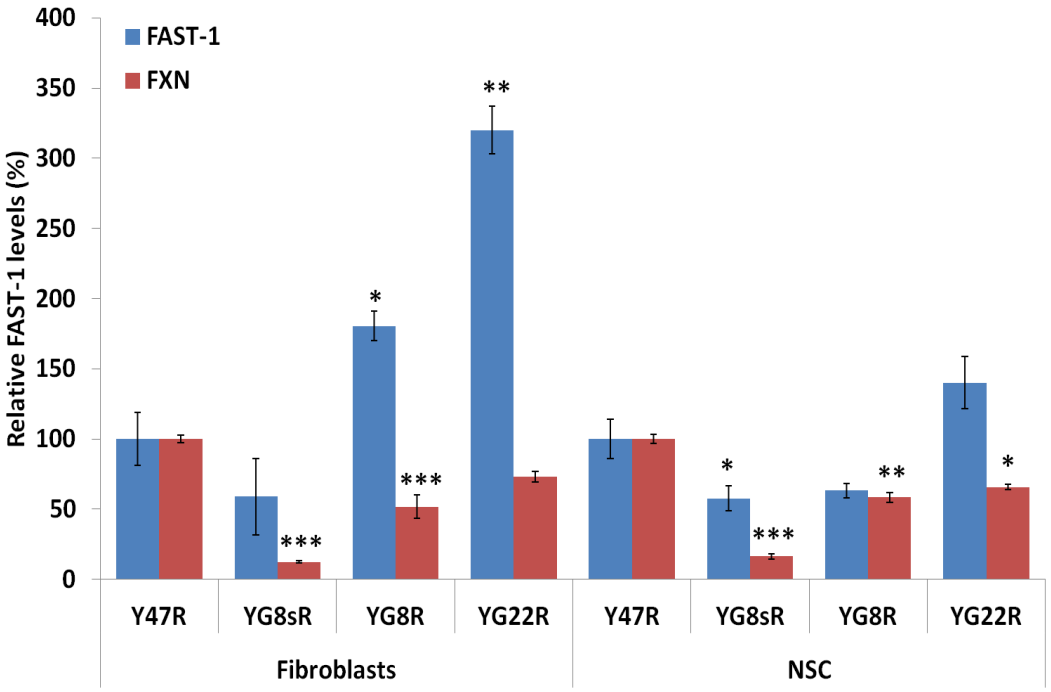

Figure 5.6 – Relative *FXN* and *FAST-1* mRNA levels in fibroblasts derived from Y47R, YG8sR, YG8R and YG22R mouse models (n=2), *p<0.05, **p<0.01, ***p<0.001, error bars ± 1 SEM.

5.4 - Quantification of *FAST-1* levels in FRDA YAC transgenic mouse tissues

To examine the levels of *FAST-1* and *FXN* expression in FRDA YAC transgenic mouse (Y47R, YG8R and YG8sR) tissues - brain, cerebellum (neuronal tissues) and liver (non-neuronal tissue), I performed qRT-PCR analysis using *Gapdh* and *Hprt* as endogenous controls. All values were normalized to Y47R expression and experiments were carried out in triplicate. The relative levels of mRNA expression in brain, cerebellum and liver tissues were then calibrated by calculating the means of the RQ values, Y47R brain values were set arbitrarily as 100%.

My findings revealed that *FAST-1* levels in YG8R mice were increased by 154% in brain ($p<0.05$) and 223% in cerebellum ($p<0.01$), and 30% in liver (ns), compared to Y47R mice (Figure 5.7). In contrast, *FAST-1* levels were decreased by 82% in brain ($p<0.001$), 75% in cerebellum ($p<0.001$) and 50% in liver (ns) in YG8sR mice. Similarly, *FXN* expression levels in YG8sR mice were decreased by 74% in brain ($p<0.001$), 82% in cerebellum ($p<0.001$) and 88% in liver ($p<0.01$). *FXN* levels in YG8R mice were reduced by 40% in brain ($p<0.05$) and 50% in liver (ns) compared to Y47R brain tissues (Figure 5.7). However, in cerebellum a non-significant increase of 27% in *FXN* expression was observed in YG8R mice. In general, I have noticed that there is a trend in YG8sR mice for both *FAST-1* and *FXN* expression to be reduced, however the YG8R data looks very similar to the human fibroblasts data. Due to the differential expression of *FAST-1* in different tissues it is speculated that *FAST-1* may be influenced by tissue-specific effects.

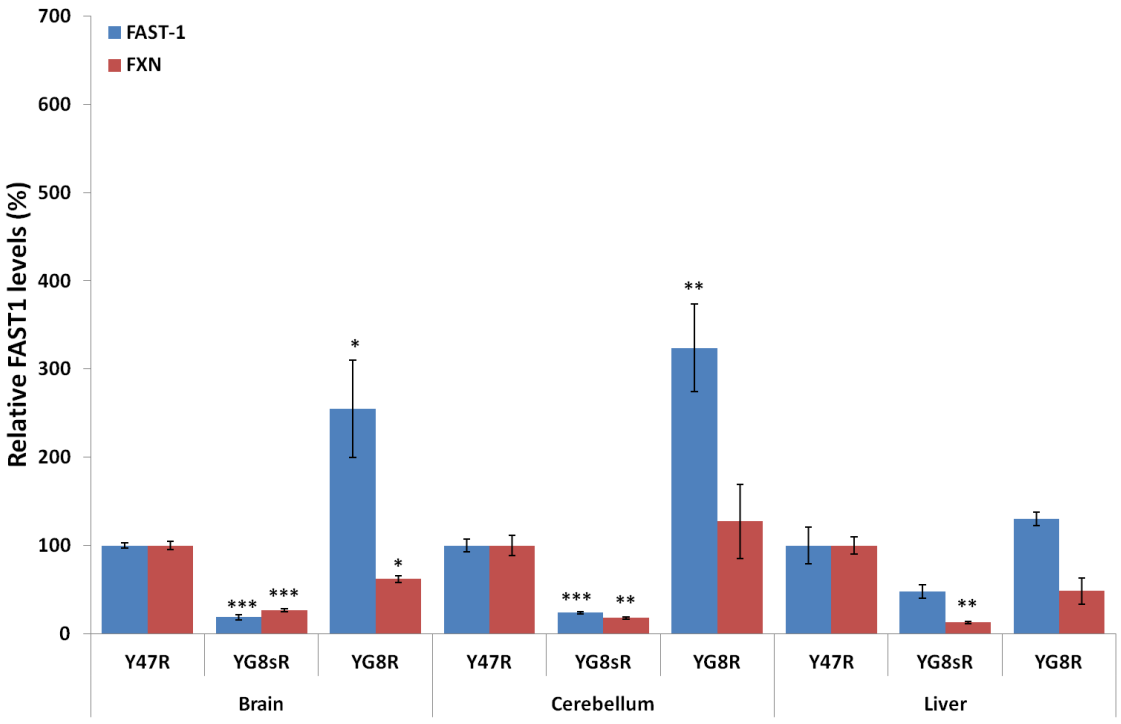

Figure 5.7 – Relative *FXN* and *FAST-1* mRNA levels in brain, cerebellum and liver tissues of Y47R, YG8sR and YG8R mouse models (n=2), *p<0.05, **p<0.01, ***p<0.001, error bars ± 1 SEM.

5.5 - Quantification of *FAST-1* levels in FRDA BAC transgenic mouse tissues

Sarsero *et al*. (2004) have generated a *FXN* bacterial artificial chromosome (BAC) transgenic mouse model by inserting a 188 kb BAC clone (pBAC265) containing exons 1-5b of the normal human FRDA locus. The *Fxn* knockout mice embryonic lethality was successfully rescued by the human FRDA orthologue (Sarsero *et al.* 2004). However, rather surprisingly, the rescued mice did not exhibit any phenotypic or behavioural differences when compared to wild type mice. Subsequently, similar to the generation of expanded FRDA YAC transgenic mouse model, a *FXN* BAC transgenic mouse model has also been generated by inserting large GAA expansion mutation (>500 GAA repeats), but this mouse, which retains a selectable marker gene and has an interrupted GAA repeat sequence, failed

to exhibit any FRDA-like phenotype (J. Sarsero, personal communication).

In addition to the identification of *FAST-1* expression levels in our FRDA YAC transgenic mouse models, we were interested to know the *FAST-1* levels in FRDA BAC transgenic mouse tissues. Therefore, I quantified *FAST-1* and *FXN* expression levels in brain, cerebellum and liver tissues of *FXN* BAC transgenic control mouse (9 GAA repeats) and FRDA BAC transgenic mouse with expanded GAA repeats (500 GAA repeats) by qRT-PCR. All *FAST-1* and *FXN* values were normalized to *FXN* BAC transgenic control mouse brain. QRT-PCR analysis indicated that the *FAST-1* expression levels in GAA expansion BAC transgenic mice were decreased by 76% in brain ($p<0.01$) and 88% in liver ($p<0.01$) as compared to the control mice. However, in cerebellum tissues the *FAST-1* levels were increased to 731% ($p<0.001$) in FRDA mice. Similarly, *FXN* levels were decreased by 55% in brain ($p<0.001$), 74% in cerebellum ($p<0.001$) and 94% in liver ($p<0.01$) (Figure 5.8) compared to the levels in control mice. These results show that there are no consistent correlations, except in cerebellum tissues, between *FAST-1* and *FXN* expression levels in brain and liver tissue types, at least in this mouse model.

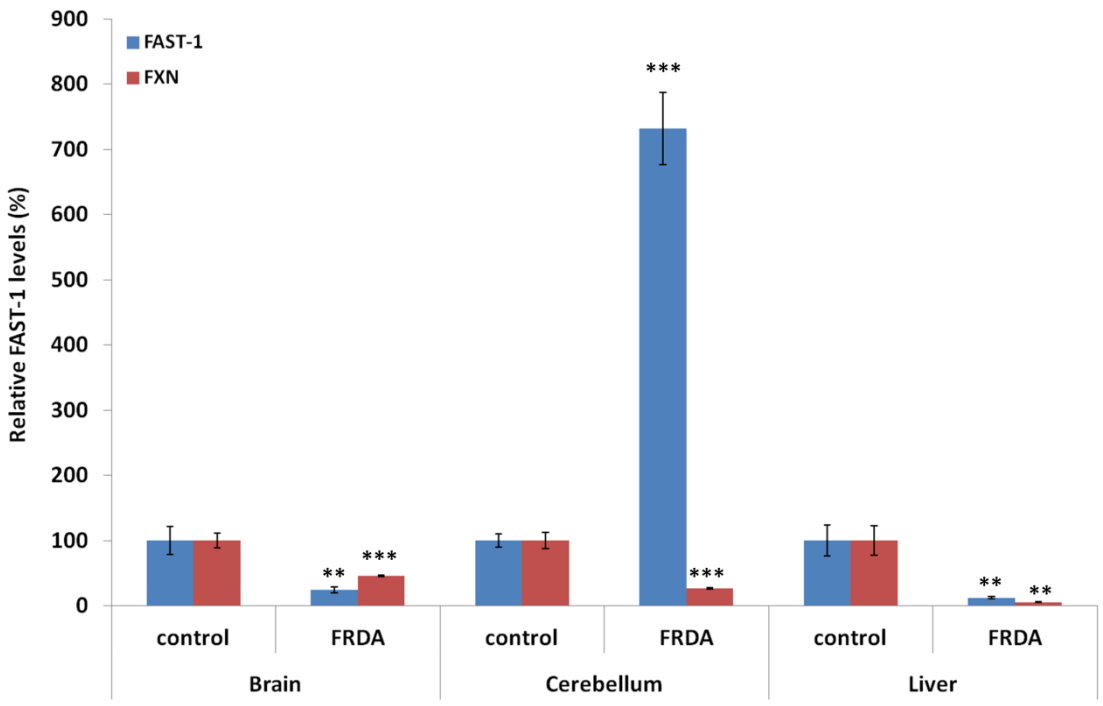

Figure 5.8 – Relative *FXN* and *FAST-1* mRNA levels in brain, cerebellum and liver tissues of control and FRDA BAC transgenic mouse models (n=2) **p<0.01, ***p<0.001, error bars ± 1 SEM.

5.6 - Quantification of *FAST-1* levels in double genetically modified mice

To understand the FRDA disease mechanism and to find an effective therapy, we have generated a FRDA YAC transgenic mouse model with expanded repeats (up to 270 GAA repeats). One of the characteristic features of this mouse model is that it spontaneously exhibits somatic and intergenerational GAA repeat instability (Al-Mahdawi *et al.* 2004). Eventually, we have cross bred our FRDA YAC transgenic mice with number of mismatch repair (MMR) knockout mice to understand the underlying mechanism of GAA repeat instability (Bourn *et al.* 2012; Ezzatizadeh *et al.* 2012; Ezzatizadeh *et al.* 2014). Similarly, to further characterise the GAA repeat dynamics and *FXN* gene silencing, we have recently crossed GAA repeat expansion-containing *FXN* transgenic mice with Bmi-1 and HP1γ

heterozygous knockout mice. The details of *FAST-1* and *FXN* expression analysis in these double genetically modified mouse models are discussed below.

5.6.1 - Quantification of *FAST-1* levels in Bmi-1 mouse tissues

Bmi-1 is a member of Polycomb family of transcription repressors, has been implicated in controlling cell senescence by repression of p19arf/MDM2/p53 pathway (Jacobs *et al.* 1999; Park *et al.* 2004; Molofsky *et al.* 2005). In the central nervous system (CNS), Bmi-1 protein is localised within NSCs, as well as mature neurons and astrocytes, indicating its critical role in brain development and maintenance (Hayry *et al.* 2008; Chatoo *et al.* 2009). It has recently been reported that Bmi-1 is required for maintaining endogenous antioxidant defences in the brain, and its absence subsequently causes premature brain degeneration (Cao *et al.* 2012). Since there are number of functions implicated to the Bmi-1 protein, we were interested to investigate the role of this protein in FRDA disease progression. Therefore, we have bred a line of our FRDA YAC transgenic mice (YG8sR) with the Bmi-1 heterozygous knockout mice to produce double genetically modified mice with both a *FXN* transgene and a Bmi-1 heterozygous knockout allele.

FAST-1 and *FXN* expression levels were quantified in Bmi-1 WT and Bmi-1 heterozygous knockout mouse tissues (brain, cerebellum and liver) by qRT-PCR. All values were normalised to the Bmi-1 WT brain. QRT-PCR analysis of Bmi-1 heterozygous knockout mouse tissues showed that *FAST-1* levels were non-significantly decreased in FRDA Bmi-1 heterozygous knockout mouse brain and non-significantly increased in cerebellum and liver tissues compared to Bmi-1 WT mice (Figure 5.9). In contrast, *FXN* expression levels did not change in any of the three tissues studied, brain, cerebellum and liver. Since there is no

correlation between the WT and heterozygous allele of the Bmi-1 gene on *FXN* (and *FAST-1*) expression, it would be most unlikely that the Bmi-1 had an effect on *FAST-1* gene expression.

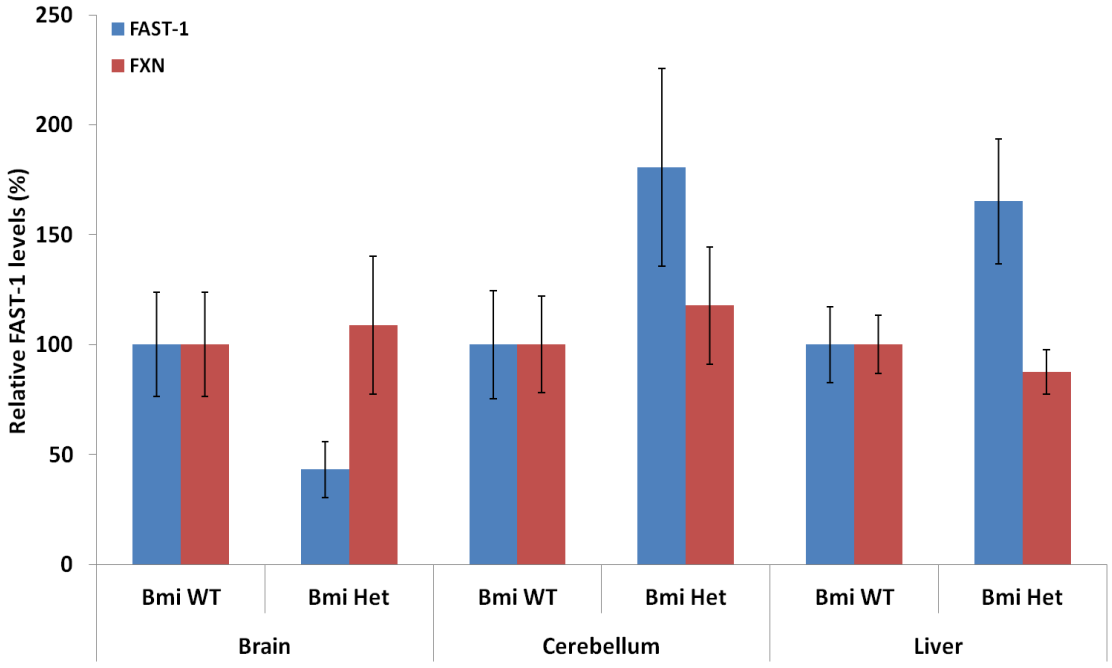

Figure 5.9 – Relative *FXN* and *FAST-1* mRNA levels in brain, cerebellum and liver tissues of Bmi-1 wild type and Bmi-1 heterozygous transgenic mouse models (n=3) error bars ± 1 SEM.

5.6.2 - Quantification of *FAST-1* levels in HP1γ heterozygous tissues

Heterochromatin protein 1 (HP1) is a small non-histone protein and highly conserved across the evolution, that is typically found in heterochromatin. Mammalian cells contain three HP1 proteins, HP1α, HP1β and HP1γ. Heterochromatin, recruited by HP1, is a repressive chromatin state that is characterized by densely packed DNA and low transcriptional activity. Heterochromatin-induced gene silencing is important for mediating

developmental transitions, and in addition, it has more global functions in ensuring chromosome segregation and genomic integrity (Hahn *et al.* 2010). FRDA has recently been recognised, by a number of remarkable findings, as an epigenetic disorder (see reviews Sandi *et al.* 2013; Yandim *et al.* 2013; Sandi *et al.* 2014). Also, it is strongly believed that gene repressive marks, especially H3K9me3, were significantly enriched at *FXN* gene locus (Herman *et al.* 2006; Al-Mahdawi *et al.* 2008; De Biase *et al.* 2009; Kim *et al.* 2011). Importantly, nucleosomes containing H3K9me3 is known to recruit HP1, which then binds to H3K9 methyltransferases. In support of this hypothesis, FRDA fibroblasts have shown significant enrichment of H3K9me3 coupled with both subunits of HP1 (HP1α and HP1γ) (De Biase *et al.* 2009) at the vicinity of the *FXN* gene locus that may perhaps increased the heterochromatin in FRDA cells that led to the *FXN* gene silencing. Therefore, one of the possible suggestions is that decreased HP1 levels may partially rescue the *FXN* gene from being switched off. Previous epigenetic studies on FRDA have provided enough evidence of the GAA repeat expansion-induced transcriptional inhibition of *FXN*. However, no studies have focused directly on elucidating the role of HP1 in FRDA. Therefore, we first set out to investigate the role of this protein *in vivo* using our FRDA YAC transgenic mouse model. We then bred HP1γ heterozygous mice with expanded FRDA YAC transgenic mouse model (YG8sR) to generate double transgenic mouse model with *FXN* and HP1γ. These double transgenic mouse models were in 50% B6 and 50% 129 backgrounds.

The expression levels of *FAST-1* and *FXN* in transgenic FRDA HP1γ WT and transgenic FRDA HP1γ heterozygous mouse tissues (brain, cerebellum and liver) were quantified by qRT-PCR. The HP1γ WT brain values were used as a calibrator to express the *FAST-1* and *FXN* values relatively. After analysing these results we have found that *FAST-1* levels in FRDA

HP1γ heterozygous mouse were increased by 50% in brain (ns), decreased by 37% in cerebellum (ns) and no change in liver compared to FRDA HP1γ WT mouse (Figure 5.10). No change was seen in *FXN* levels of brain and non-significantly decreased in cerebellum, but in liver tissues the *FXN* expression was decreased by 30% and reached a statistical significance ($p<0.05$) in FRDA HP1γ heterozygous mouse compared to FRDA HP1γ WT mouse brain (Figure 5.10).

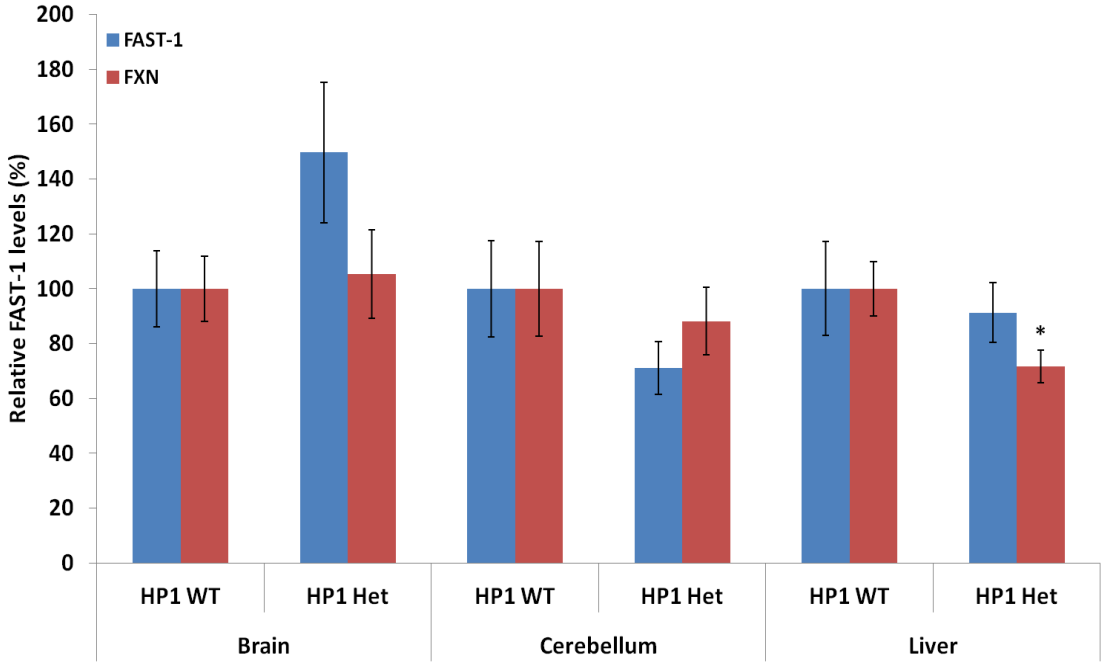

Figure 5.10 – Relative *FXN* and *FAST-1* mRNA levels in brain, cerebellum and liver tissues of HP1 wild type FRDA mice and HP1 heterozygous FRDA transgenic mouse models (n=3) *$p<0.05$, error bars ± 1 SEM.

5.7 - Quantification of *FAST-1* levels in diazoxide-treated FRDA YAC transgenic mouse tissues

Diazoxide is a mitochondrial ATP-sensitive potassium (mitoK_{ATP}) channel activator that is thought to act by binding to a specific conformation of the sulfonylurea receptor (SUR) component of this heteromeric channel (Bryan and Aguilar-Bryan 1999). It is currently approved for oral use in humans (e.g. Proglycem) as a vasodilator to treat hypertension and as an inhibitor of insulin secretion to treat hypoglycemia due to insulinoma or congenital hyperinsulinaemia. However, research data from diabetic rat models have shown that diazoxide may actually improve insulin secretion, prevent pancreatic β-cell apoptosis and improve pancreatic β-cell functionality during the chronic hyperglycaemia associated with diabetes (Kwon *et al.* 2006; Huang *et al.* 2007; Ma *et al.* 2007). Diazoxide has also been used in humans to improve β-cell function in subjects with type 2 diabetes (Guldstrand *et al.* 2002) and to improve glycemic control in subjects with type 1 diabetes (Radtke *et al.* 2010). Therefore, diazoxide may prove beneficial to the subset of FRDA patients who suffer with diabetes. Other reports have described cardioprotective effects of diazoxide in mice with ischaemia/reperfusion injury (Suzuki *et al.* 2003; Ahmad *et al.* 2006), which may have an impact on the cardiac pathology seen in FRDA.

We evaluated the long-term effects of diazoxide (3mg/kg) treatment of YG8sR mice over a period of 3 months. After treatment, mice were culled and tissues were collected by snap frozen in liquid nitrogen. *FAST-1* and *FXN* expression levels in brain and liver tissues were assessed by qRT-PCR, comparing diazoxide-treated mice with untreated mice. We had previously identified that changes in behavioral tests were differentially exhibited in male and female mice. Therefore, we compared expression levels in both male and female mice

together or male or female mice alone. This analysis revealed no change in *FAST-1* and *FXN* expression levels in brain tissues of diazoxide-treated mice when both male and female values were taken together or when female values were considered alone (Figure 5.11). However, when we analysed the data for male mice, *FAST-1* expression was increased by 1.8-fold ($p<0.001$) without any change in *FXN* gene levels (Figure 5.11). In liver tissue, the *FAST-1* and *FXN* expression levels were significantly reduced in diazoxide-treated liver samples when both male and female values were taken together ($p<0.05$ for both) or when female values were considered alone ($p<0.01$ for both) (Figure 5.12). Similar to brain tissue, liver tissue of diazoxide-treated male mice have shown a significant increase in *FAST-1* ($p<0.001$) but no change in *FXN* levels as compared to untreated liver sample (Figure 5.12). These results indicated that long-term treatment of diazoxide in YG8sR had no effect on *FXN* gene expression levels, but increased *FAST-1* expression in male mice in both tissues studied (brain and liver). We did not notice any gender specific *FAST-1* expression in any of our previous analysis. Therefore, we conclude that this could be just a gender specific effect of the drug that may lead an increase in *FAST-1* expression.

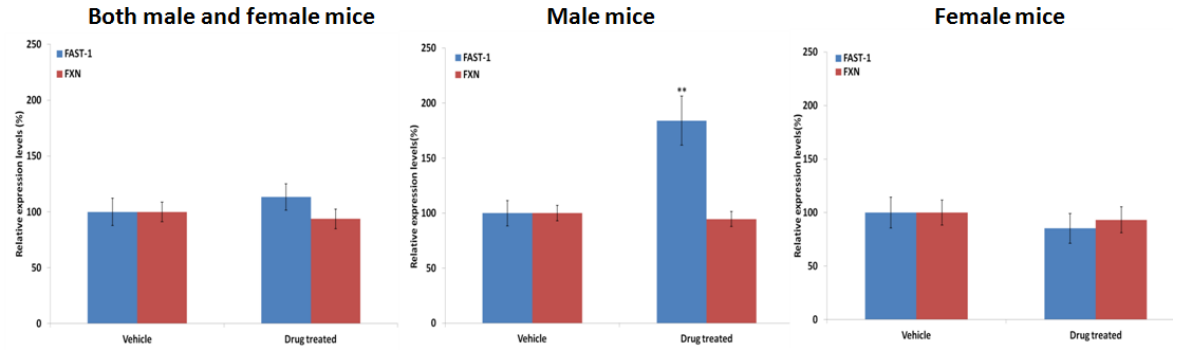

Figure 5.11 – Relative *FXN* and *FAST-1* mRNA levels in brain tissues of diazoxide treated YG8sR mice. The analysis of both male and female mice together or female mice alone revealed no change in expression *FXN* or *FAST-1* levels. However, the analysis of male mice alone has shown significant increases in *FAST-1* levels in drug treated mice (n=4). **p<0.01, error bars ± 1 SEM).

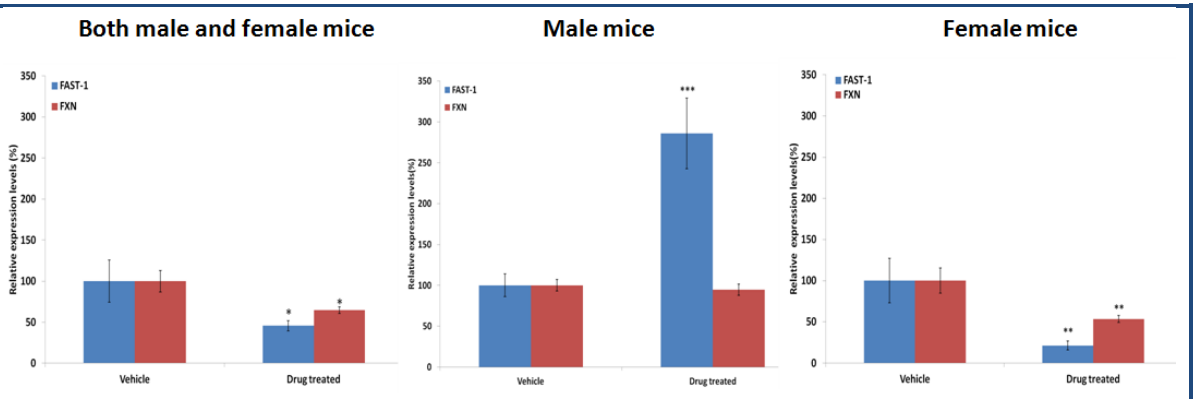

Figure 5.12 – Relative *FXN* and *FAST-1* mRNA levels in liver tissues of diazoxide treated YG8sR mice. The analysis of both male and female mice together or female mice alone has revealed significant reduction in both *FXN* and *FAST-1* expression levels. However, the analysis of male mice alone has shown only significant increases in *FAST-1* levels in drug treated mice and no change in *FXN* levels (n=4) (*p<0.05, **p<0.01, ***p<0.001 error bars ± 1 SEM).

5.8 - Quantification of *FAST-1* levels in HMTase inhibitor-treated mouse cells

The finding that FRDA has an epigenetic aetiology has encouraged the investigation of epigenetic-based therapies, in particular the use of histone deacetylase (HDAC) inhibitors, to reverse *FXN* gene silencing (See review Sandi *et al.* 2013). Our group has previously investigated the long-term safety and efficacy of three 2-aminobenzamide HDAC inhibitors, designated 106, 136 and 109, in YG8R YAC transgenic mice (Sandi *et al.* 2011). Compound 109 has subsequently been taken forward by Repligen Corp. to a phase I clinical trial as RG2833 (Soragni *et al.* 2014), and new compounds are currently being investigated. As an additional epigenetic-based therapeutic approach for FRDA therapy, we recently proposed to investigate the use of HMTase inhibitors. HMTase enzymes are divided into lysine-specific and arginine-specific groups. They have great specificity and usually modify only one particular histone residue. Since FRDA is associated with significant increases in repressive histone marks at the *FXN* locus, such as hypermethylation of H3K27 and H3K9, it would be of great interest to consider the use of compounds that could reverse this epigenetic state to activate *FXN* gene expression. Some encouraging preliminary studies of FRDA cells and mice have shown that the class III HDAC inhibitor, nicotinamide, can decrease levels of H3K9 and H3K27 trimethylation at the *FXN* locus to upregulate *FXN* gene expression (Chan *et al.* 2013). We have now investigated a series of HMTase inhibitors in our laboratory and I have investigated *FAST-1* and *FXN* expression levels following treatment of mouse primary fibroblast cells.

In our preliminary experiments, we used BIX01294 (G9a inhibitor) at 1nM and 1µM concentration for 48 hours. After the treatment, cells were collected and RNA was extracted by the Trizol method, followed by cDNA synthesis. Treatment of BIX01294 in Y47R mouse

primary fibroblast cells resulted in no change in *FAST-1* expression although a non-significant decrease (29%) in cells treated with 1µM of BIX01294. However, *FXN* levels were increased to 196% ($p<0.05$) and 138% ($p<0.05$) with 1nM and 1µM of the drug, respectively as compared to untreated sample Figure 5.13). In YG8R mouse cells, a dose dependent decrease in *FAST-1* expression was noted. These levels were reduced by 25% ($p<0.01$) and 41% ($p<0.01$) with 1nM and 1µM of the drug, respectively as compared to untreated sample Figure 5.13). However, reduced *FAST-1* concentrations were not correlated with increases in *FXN* expression as at the same concentrations, *FXN* expression was also reduced by 25% ($p<0.01$) and 19% ($p<0.01$), respectively compared to the untreated control cells (Figure 5.13). Similarly, treatment with BIX01294 in YG8sR also resulted in significant reduction of *FAST-1* levels at both concentrations used, but more pronounced at 1nM (Figure 5.13). *FXN* expression levels were reduced in both the drug-treated samples (Figure 5.13).

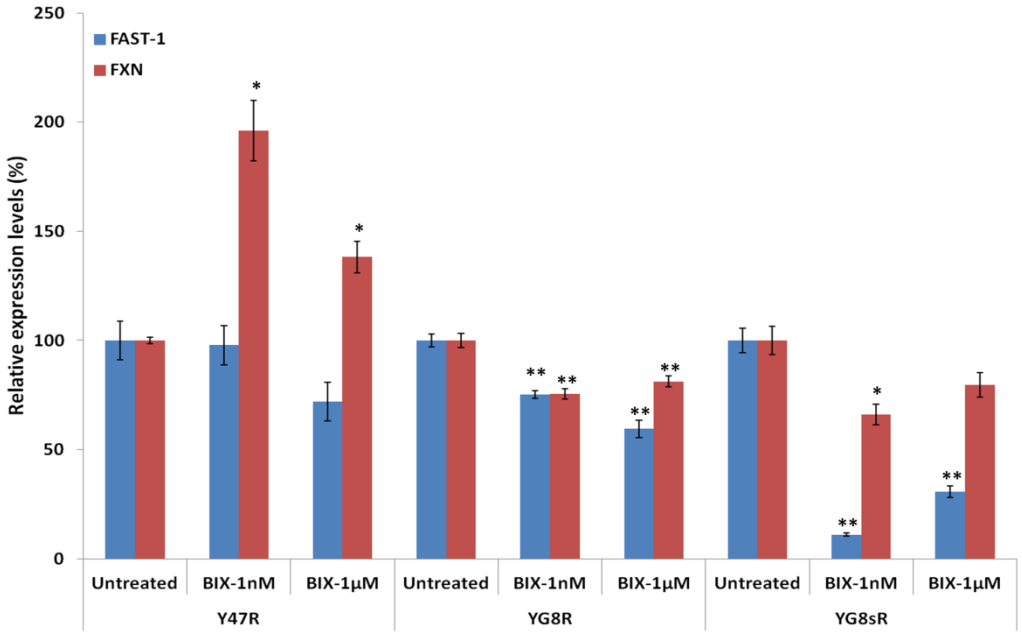

Figure 5.13 – Relative *FXN* and *FAST-1* mRNA levels in Y47R, YG8sR and YG8R mouse primary fibroblasts following treatment with BIX01294 at 1nM and 1µM concentrations for 48 hours (*$p<0.05$, **$p<0.01$, ***$p<0.001$ error bars ± 1 SEM).

5.9 - Discussion

Antisense transcripts have been generally implicated in heterochromatin formation via association with H3K9me3 and subsequent recruitment of HP1 (Grewal and Elgin 2007; Grewal and Jia 2007; Iida *et al.* 2008). In recent years, attention has been drawn to the fact that a significant fraction of the transcriptome comprises RNAs containing sequences that are complementary to other endogenous RNAs. Natural antisense transcripts (NATs) are RNA molecules that are transcribed from the opposite DNA strand compared with other transcripts and overlap in part with sense RNA (reviewed in Beiter *et al.* 2009). It is very interesting to note that most of the antisense transcription expression is not always associated with the expression of the sense gene, which suggests the use of alternative transcriptional regulatory elements (He *et al.* 2008). These natural antisense transcripts (NATs) can have protein-coding properties but mainly represent non-coding RNAs (ncRNAs). NATs have been implicated with diverse regulatory functions at various levels, including imprinting, X-inactivation, RNA processing, RNA export, and transcriptional regulation (reviewed in Beiter *et al.* 2009). Therefore, the discovery of *FAST-1* is therefore not surprising in itself. Furthermore, the detection of higher levels of *FAST-1* in the region showing heterochromatin formation in FRDA patients offers a plausible mechanistic basis for the epigenetic abnormality involving in *FXN* gene silencing.

To investigate the potential expression of *FAST-1* in human FRDA YAC transgenic mouse models, I first attempted to reproduce the results of De Biase *et al.* (2009) by investigating *FAST-1* expression in human FRDA fibroblast cells. This required establishing a robust qRT-PCR-based method for detecting and quantifying human *FAST-1* expression within our lab. As discussed in De Biase *et al*. paper, a strand specific cDNA synthesized by

quantitect reverse transcription kit (QIAGEN) using FAST RT primer was not successful in *FAST-1* amplification through primary PCR, but amplified in nested PCR. Therefore, to amplify *FAST-1* in a primary PCR, I optimised the cDNA synthesis using different approaches and found that strand specific cDNA synthesized by cloned AMV first-strand cDNA synthesis kit using FAST RT primer was a good template for *FAST-1* amplification, whereas for endogenous control, cDNA synthesised by an oligo$(dT)_{20}$ primer was used. QRT-PCR analysis of human fibroblasts clearly indicated increased levels of *FAST-1* and reduced levels of *FXN* expression compared to fibroblasts derived from unaffected individuals. These results are in agreement with the previous findings of De Biase *et al.* (2009). Therefore, we were interested to investigate *FAST-1* and *FXN* expression levels in FRDA YAC transgenic mouse (Y47R, YG8sR, YG8R and YG22R) cells and tissues. Initially I quantified *FAST-1* and *FXN* expression levels in FRDA YAC transgenic mouse fibroblasts, which clearly showed increased levels of *FAST-1* and reduced levels of *FXN* expression in YG8R and YG22R fibroblasts compared to the Y47R fibroblasts. These findings are in agreement with our pervious results using human fibroblasts. However, although YG8sR mice are derived from a founder YG8R mouse, the gene expression pattern for both *FAST-1* and *FXN* is differed from YG8R mice fibroblasts. Even though, the ratio of *FAST-1* and *FXN* is maintained similar to YG8R mice cells, the overall expression of *FAST-1* and *FXN* were dramatically reduced in YG8sR mice cells. Since the interactions between *FXN* (sense) and *FAST-1* (antisense) transcripts are not known, it is paramount important to hypothesize such interactions to unravel any possible mechanisms underlying this phenomenon. Further investigations of *FAST-1* in NSCs displayed a trend towards reduced *FAST-1* expression in the YG8R and YG8sR cells, indicating that the *FAST-1* expression may be cell type dependent.

I then extended my investigations of *FAST-1* into two different mouse models, YAC FRDA and BAC FRDA mouse models to see if there is any difference in expression levels in different tissues. My findings revealed that *FAST-1* expression is only increased in brain and cerebellum tissues of YG8R mice, similar to my initial findings with YG8R fibroblast cells. *FAST-1* levels were significantly increased in cerebellum tissues in the BAC FRDA mouse model, indicating the *FAST-1* expression may be dependent on other factors in addition to the length of GAA repeats and *FXN* expression levels. To further characterise the GAA repeat dynamics and *FXN* gene silencing, we recently developed two double genetically modified mouse models in our laboratory, crossing YG8sR mice with either Bmi-1 or HP1γ knockout mouse models. Quantification of *FAST-1* and *FXN* expression levels in these double genetically modified mice revealed no significant differences between WT and heterozygous Bmi-1 or HP1γ knockout mice when the comparison was made within the same tissue. This suggests that neither Bmi-1 nor HP1γ significantly affect the regulation of *FAST-1* expression.

Having developed a robust qRT-PCR based technique to detect *FAST-1* expression levels, I then utilised this technique to find changes in *FAST-1* following treatment with various drugs. Quantification of *FAST-1* in diazoxide-treated mouse brain and liver tissues revealed a gender specific effect on *FAST-1* expression. Therefore, gender specific variations in *FAST-1* expression may be due to a differential effect of the drug. More strikingly, HMTase treated cells showed dose dependent reduction in *FAST-1* expression. However, although *FAST-1* levels were significantly downregulated, there were no corresponding increases of *FXN* levels, rather I noticed non-significant reductions in *FXN* expression.

In summary, I have developed a robust qRT-PCR based method to detect and quantify the *FAST-1* expression levels. I also analysed the expression pattern of this antisense transcript in various corresponding FRDA cell lines and mouse models. The original report of the existence of the *FAST-1* is further confirmed in our three different systems, including human cell lines, mouse cell lines and mouse tissues. However, no consistent and distinct patterns of *FAST-1* expression were identified in relation to *FXN* expression. Therefore, either they are not directly connected, as originally reported (De Biase *et al* 2009) or their relationship varies between cell and tissue types. Further studies, such as knockdown or over expression of *FAST-1* with subsequent measurement of *FXN* levels, may be necessary to identify any solid correlation between *FAST-1* and *FXN*. In addition, the identification of any tissue-specific and cell type-specific expression patterns of *FAST-1* might guide future in-depth study to find a mechanistic role of this novel antisense transcript in FRDA disease progression.

Chapter 6 - Quantification of mRNA levels of *FXN* flanking genes by qRT-PCR

6.1 - Introduction

An unsolved problem in genetics is the finding that gene silencing can 'spread' along a chromosome. Transcriptional control is a major mechanism for regulating gene expression. The complex machinery required to affect this control is still emerging from functional and evolutionary analysis of genomic architecture. In almost all organisms, the precise regulation of gene transcription is controlled by a complex system of interactions between transcription factors, histone modifications and modifying enzymes and chromatin conformation. The requirement for complex spatio-temporal regulation is most obvious during the development and differentiation. More strikingly, the spatio-temporal regulation requires a precise gene switching which controls the generation of many different cell types, at the right time and the right place, from a single fertilised cell (van Heyningen and Bickmore 2013). This complex procedure is tightly regulated by number of *cis*- and *trans*-acting elements. DNA damage in any of these regions has the capacity to disrupt this highly regulated process, although it can be argued that functional DNA repair mechanism remove any damaged bases. However, growing evidence has suggested that the efficiency and specificity with which cells repair such regions may affect transcriptional regulation. Long-range silencing seems to be a dynamic process and this involves local diffusion of histone-modifying enzymes from source binding sites to low-affinity sites nearby. To address the complex nature of this silencing procedure, Talbert and Henikoff have proposed number of models, based on various contexts and organisms, to show that multiple mechanisms have evolved that silence genes at distance (Talbert and Henikoff 2006).

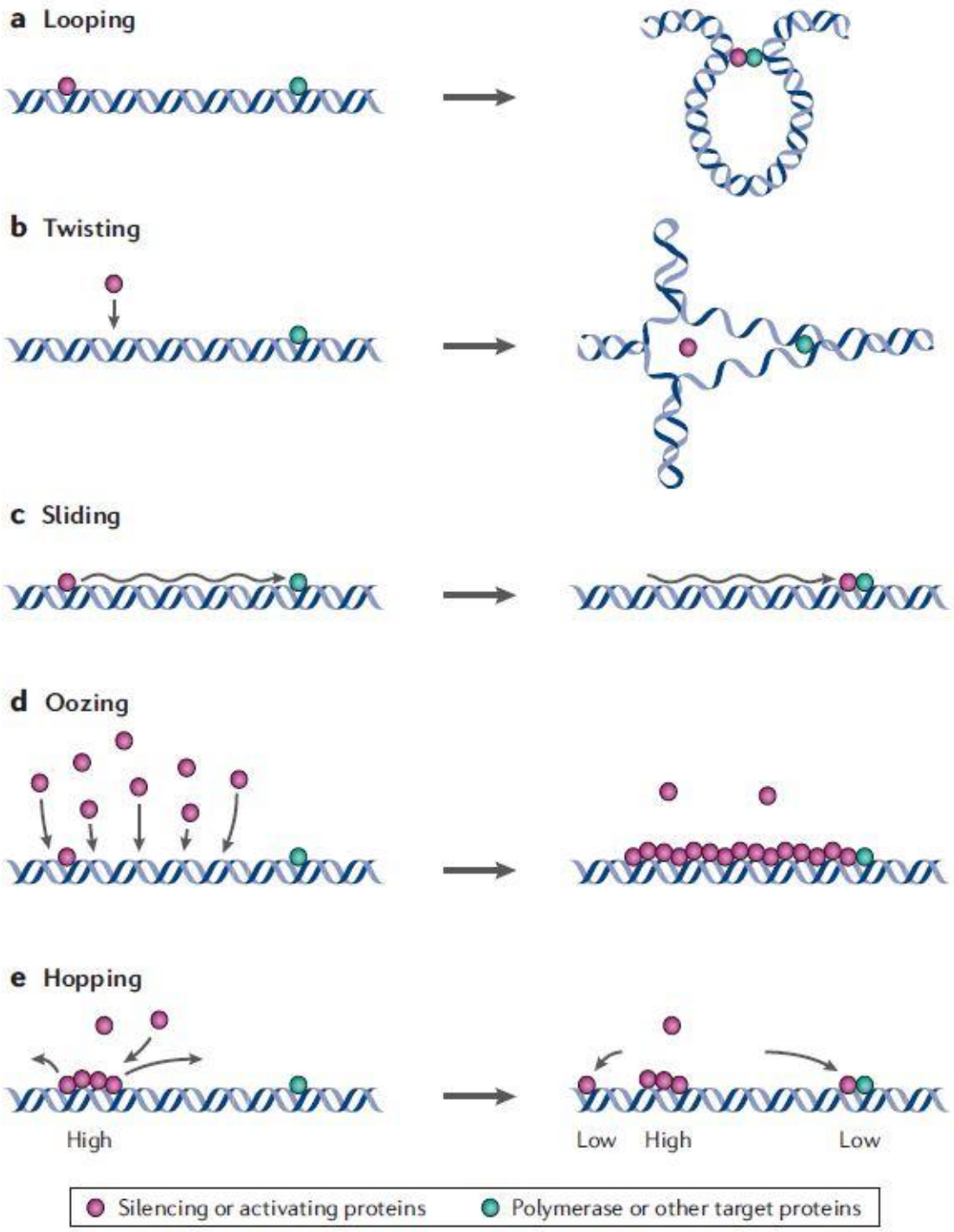

Figure 6.1 – Classes of models for action at a distance. [a] Looping (contact between distant sites), [b] Twisting (propagation of a conformational change in the DNA), [c] Sliding (tracking along a chromosome) [d] Oozing (binding of one protein facilitates adjacent binding of the next, and so on) and [e] Hopping (active release of a histone-modifying enzyme from a source site to bind nearby sites of low affinity) (adapted from (Talbert and Henikoff 2006).

There is currently very limited information available on the regulation of *FXN* gene in humans. The *FXN* promoter, which lacks a TATA sequence, is not controlled by the Inr/DPE-like elements found in the vicinity of the TSS (Greene *et al.* 2005). A recent study, which covered 21.3 kb region upstream of the exon 1 of the *FXN* gene (includes *PRKACG* gene), has identified eight highly conserved non-coding regions (Puspasari *et al.* 2011). Deletion of one of these regions has significantly reduced *FXN* gene expression indicating that sequences in these conserved regions, most probably, act as enhancers or binding sites for transcription factors that activate gene expression. As genetic alterations affecting gene expression at a distance, while keeping transcriptional unit intact, are currently emerging as contributing factors in several human genetic diseases. In support of this hypothesis, it has been recently shown that *PIP5K1β* (phosphatidylinositol-4-phosphate 5-kinase, type I, beta) gene expression was altered in FRDA patient derived fibroblasts and lymphocytes compared to healthy individuals (Bayot *et al.* 2013).

The *PIP5K1β* gene is located ~26 kb away from the 5′ end of the *FXN* gene. Although it is currently unknown the mechanism underlying reduced expression of *PIP5K1β* gene, it is believed that the GAA repeat may have a position effect variegation or *cis*-effect that might have in turn reduced this gene expression. This finding has encouraged me to identify different gene expression levels flanking the 3'- and 5' ends of the *FXN* gene. As distinct mechanisms may be consistent with effects of GAA triplet repeats on transcription of genes over distances of hundreds of kilobases, it will be of particular interest to investigate the molecular basis for this long-range gene-silencing effect (Talbert and Henikoff 2006; Deng and Blobel 2010). Therefore, in order to account for any possible explanation of the effect of GAA repeat expansion on *FXN* and its neighbouring genes, I have chosen approximately 750

kb size on both sides of the human *FXN* gene. This region encompass, in addition to *PIP5K1B* gene, *PGM5* gene at the 5'-end of *FXN* gene locus, and *TJP2, FAM189A2, APBA1* and *PTAR1* genes which flank the *FXN* locus at the 3'-end (Figure 6.2). I have quantified mRNA expression levels of these genes in a panel of human control and FRDA fibroblast cell lines. This analysis has, indeed, shown reduced expression levels of 3 out of 6 genes studied. The *PRKACG* transcript, which transcribes in the opposite direction of *FXN*, was undetectable in our experimental conditions, despite the use of three different primer pairs, indicating very low levels of gene expression in primary fibroblast cells. Therefore, this has to be, in future, investigated in other relevant tissues where this gene expression is expected to be abundant, such as testis.

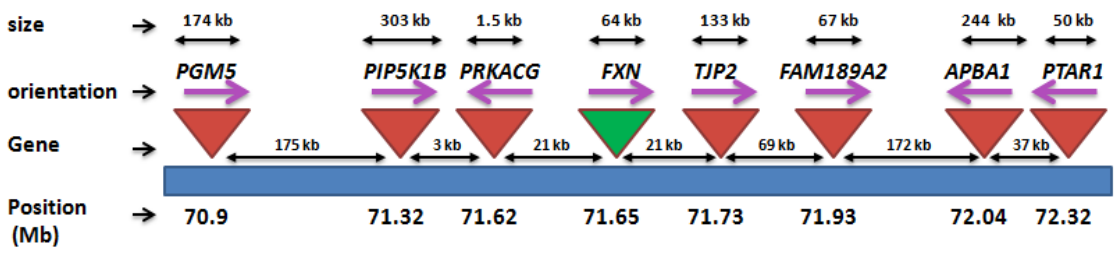

Figure 6.2 – The map of chromosome 9 which covers approximately 750 kb on both sides of human *FXN* gene (green triangle) and its close neighbouring genes (red triangles). Gene transcription orientation (pink arrows), gene sizes and their positions were represented.

6.2 - Quantification of mRNA levels of *FXN* flanking genes by qRT-PCR

6.2.1 - *PGM5* gene

PGM5, also known as aciculin, was identified as a 60 kDa cytoskeletal protein (Belkin and Burridge 1994) and it is a major dystrophin-associated protein. It consists of 11 exons and spans approximately 174 kb size on chromosome 9q13. *PGM5* is located at around 482 kb apart from the 5'end of the *FXN* gene. *PGM5* is involved in inter conversion of glucose-1-phosphate and glucose-6-phosphate and also the role of *PGM5* is essential in carbohydrate formation from glucose-6-phosphate (Edwards *et al.* 1995). In addition, *PGM5* is also found to interact with the cytoskeletal proteins dystrophin and utrophin, possibly by playing a role in cytoskeletal organisation and function (Belkin and Burridge 1995). As the major symptoms of FRDA are motor coordination impairments, gene expression regulation in the skeletal muscle tissue would be crucial for motor function control in FRDA patients.

To determine the mRNA expression levels of *PGM5*, I used 2 control and 7 FRDA patient derived fibroblasts. One of the control fibroblast values was arbitrarily set at 100%. For this analysis, HPRT gene expression levels were used as an endogenous control. Quantification of mRNA expression levels by qRT-PCR revealed significantly reduced expression of both genes studied, *PGM5* and *FXN* in FRDA fibroblasts. *PGM5* expression was reduced to 48% ($p<0.05$) and *FXN* expression was reduced to 18% ($p<0.001$) in FRDA fibroblasts as compared to control fibroblasts (Figure 6.3).

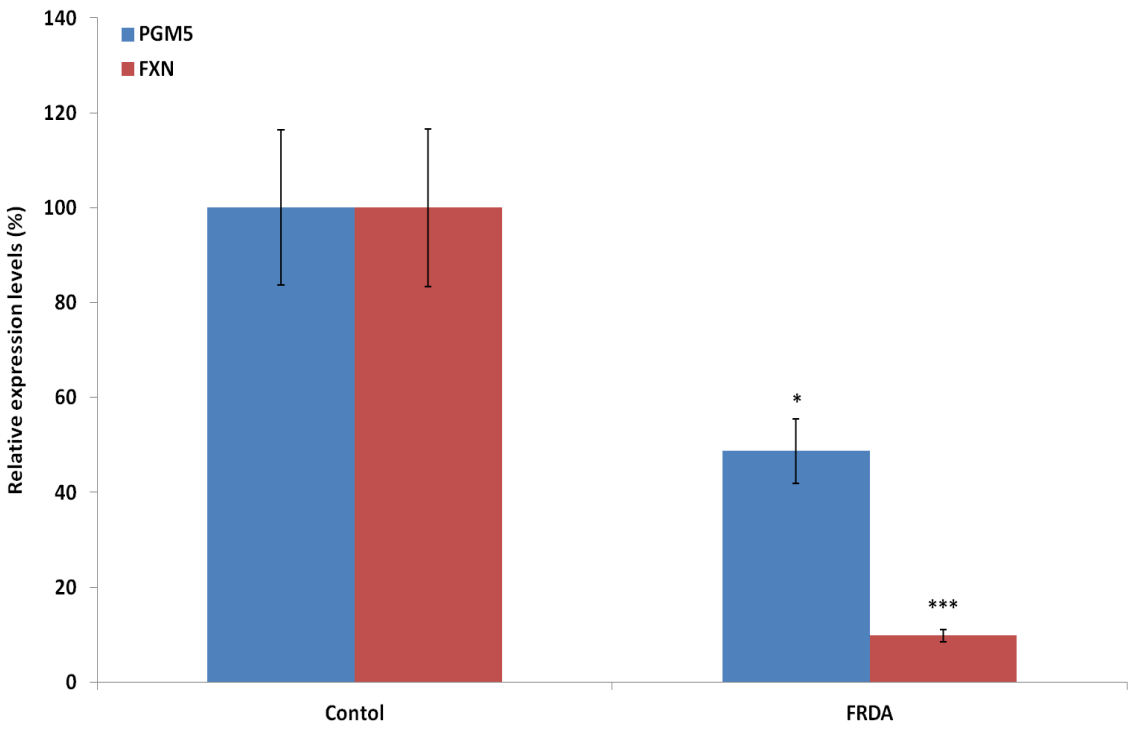

Figure 6.3 – Relative *PGM5* and *FXN* mRNA levels in unaffected and FRDA patient derived fibroblasts (n=2-7), *p<0.05, ***p<0.001 error bars ± 1 SEM).

6.2.2 - *PIP5K1β* gene

PIP5K1β (phosphatidylinositol-4-phosphate 5-kinase type-1 beta) gene consists of 18 exons and spans more than 300 kb on chromosome 9q13. It is located 26 kb away from 5'-end of the *FXN* gene. *PIP5K1β* belongs to the PIP5K1 family of enzymes that phosphorylate phosphatidylinositol 4-phosphate [PI(4)P] to generate the important signalling molecule phosphatidylinositol(4,5)bisphosphate [PI(4,5)P2] (van den Bout and Divecha 2009). Based on the linkage analysis, *PIP5K1β* was initially considered as a candidate gene for FRDA (Carvajal *et al.* 1995; Carvajal *et al.* 1996; Pook *et al.* 1997). However, after the identification of *FXN* gene as a main driver for FRDA disease progression, this hypothesis was eventually discarded. Recently, Rustin and colleagues have demonstrated that *FXN* and *PIP5K1β* expression levels were concomitantly reduced in FRDA patient fibroblasts and blood cells (Bayot *et al.* 2013). Although it is unclear whether the GAA repeat expansion had any effect on *PIP5K1β* gene expression, however, growing evidence on the existence of *FAST-1* and spreading of heterochromatin in the upstream region of the GAA repeats supports the hypothesis that the expansion of GAA repeat may have contribute to the silencing of this gene. *PIP5K1β* depletion in FRDA cells is associated with cytoskeleton anomalies, including destabilisation of actin network and delayed cell spreading (Bayot *et al.* 2013).

Since I proposed to investigate the neighbouring genes on either side of the *FXN* locus, I also included *PIP5K1β* to see whether I could reproduce the results of Bayot *et al*. In addition to the commercially available three FRDA cell lines, I also used same 5 FRDA fibroblast cell lines that Bayot *et al* used on their paper. Using the qRT-PCR method, I quantified the mRNA expression levels of *PIP5K1β* and *FXN* gene expression levels in 3 control and 8 FRDA cell lines. Although I found variable levels of *PIP5K1β* and *FXN* levels in

individual cell lines (data not shown), the mean levels of *PIP5K1β* were significantly reduced by 50% ($p<0.01$) and *FXN* levels were reduced by 88% ($p<0.001$) in FRDA fibroblasts compared to control cell lines (Figure 6.4), in agreement with previous results.

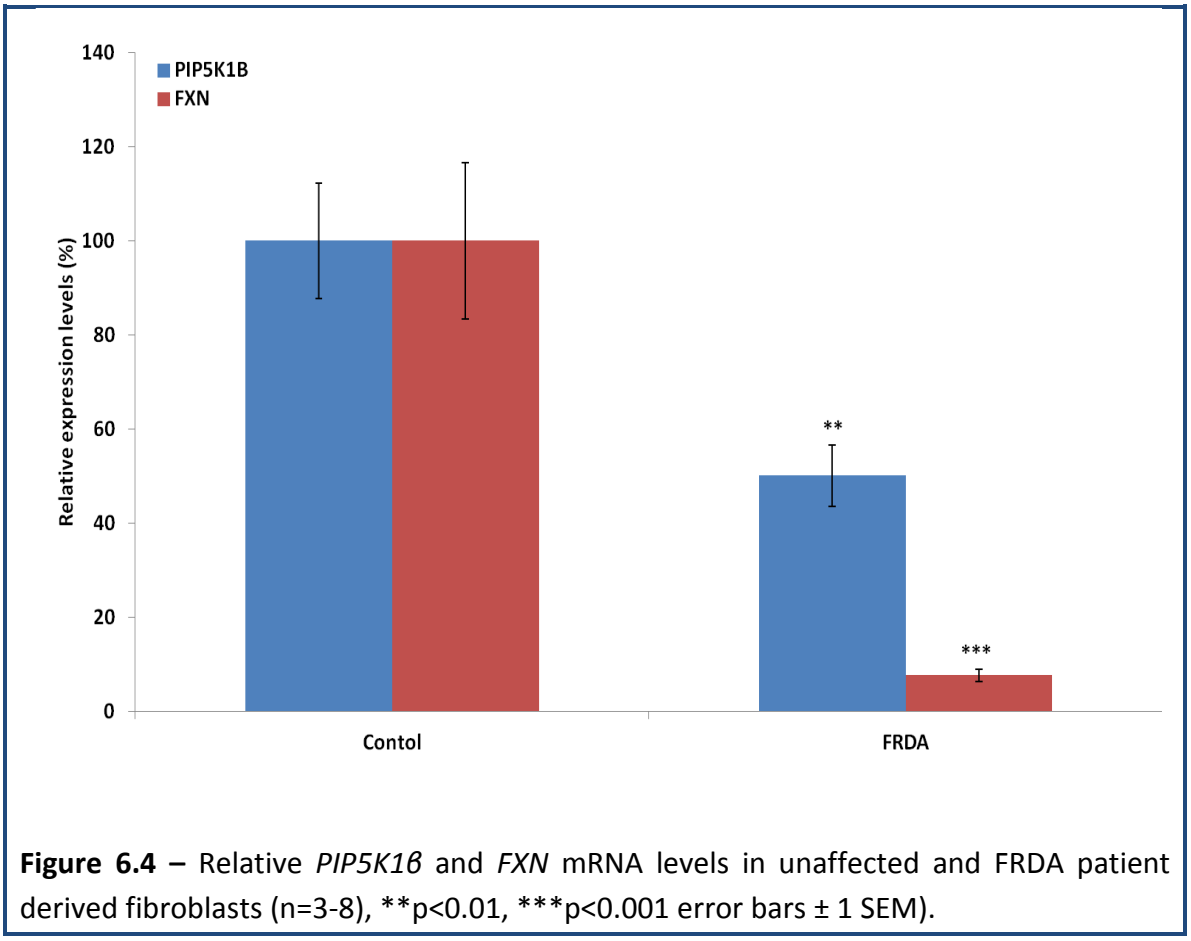

Figure 6.4 – Relative *PIP5K1β* and *FXN* mRNA levels in unaffected and FRDA patient derived fibroblasts (n=3-8), **p<0.01, ***p<0.001 error bars ± 1 SEM).

6.2.3 - *TJP2* gene

Tight junction protein 2 (*TJP2*) gene, encodes a zonula occluden (ZO-2) protein which is a member of the membrane-associated granulate kinase homolog (MAGUK) family. ZO-2 is a 160 KD tight junction scaffold protein that belongs to the MAGUK family and is a multidomain molecule that binds to a variety of cell signalling proteins, to the actin cytoskeleton, and to gap, tight, and adherens junction proteins. In sparse cultures, ZO-2 is present at the nucleus and associates with molecules active in gene transcription and pre-mRNA processing. ZO-2 inhibits the Wnt signalling pathway, reduces cell proliferation, and promotes apoptosis and its absence, mutation, or over expression is present in various human diseases, including deafness and cancer (Gonzalez-Mariscal *et al.* 2012). Importantly, *TJP2* plays an important role in cytosolic component of several classes of cell-cell junctions (Fanning *et al.* 2012). Through interaction with cytoskeletal proteins and integral membrane proteins, members of the TJP family have an important role in the localisation of components of these paracellular structures. It has been recently reported that mutations in *TJP2* cause progressive cholestatic liver disease (Sambrotta *et al.* 2014).

Quantification of *TJP2* gene, 21 kb away from 3'end of the *FXN* gene, mRNA levels by qRT-PCR revealed a non-significant reduction in FRDA cells although *FXN* gene levels were reduced to 12% ($p<0.001$) to the levels of control cells (Figure 6.5). This result indicating that the transcription machinery may perhaps less affected at 3' end of the *FXN* gene.

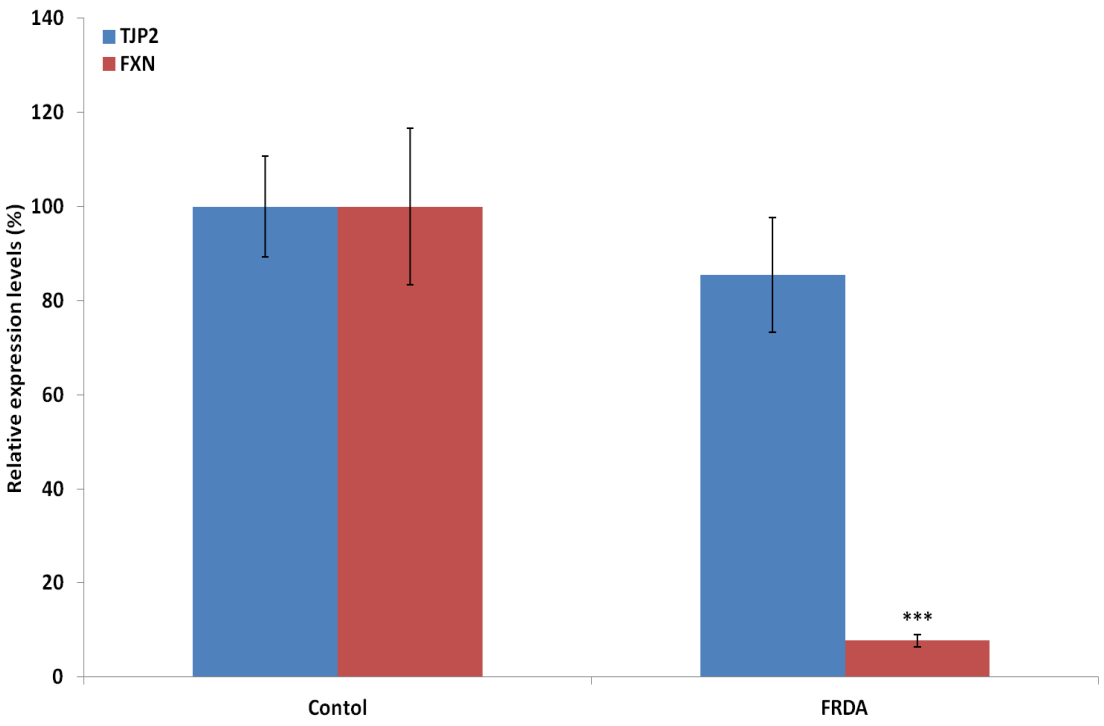

Figure 6.5 – Relative *TJP2* and *FXN* mRNA levels in unaffected and FRDA patient derived fibroblasts (n=3-8), ***p<0.001 error bars ± 1 SEM).

6.2.4 - *FAM189A2* gene

FAM189A2 gene (Homo sapiens family with sequence similarity 189, member A2) functions are largely unknown. However, the Human Protein Atlas shows a wide expression pattern with strong staining in epididymus and moderate expression in testicular cell types. The gene encodes a 450 aa membrane protein with 87% identity in mouse. Similar to *TJP2* gene, *FAM189A2* is also located on the 3' end of the *FXN* gene with approximately 225 kb distance. The quantification of mRNA levels of *FAM189A2* gene in human fibroblasts revealed no difference between FRDA and control cells (Figure 6.6).

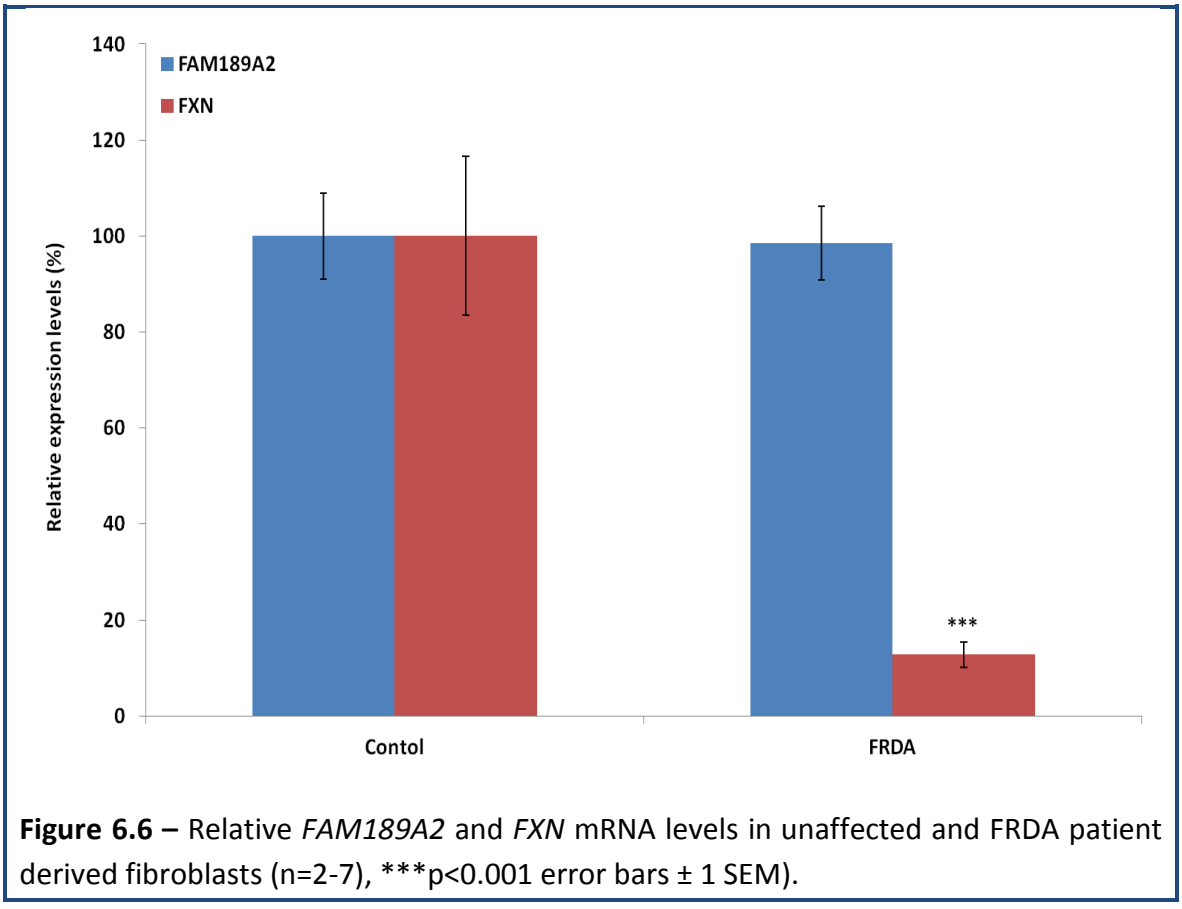

Figure 6.6 – Relative *FAM189A2* and *FXN* mRNA levels in unaffected and FRDA patient derived fibroblasts (n=2-7), ***p<0.001 error bars ± 1 SEM).

6.2.5 - *APBA1* gene

Amyloid beta A4 precursor protein-binding family A member 1 (*APBA1*) is a protein that in humans is encoded by the *APBA1* gene. *APBA1* spans over 244 kb on chromosome 9 and is composed of 13 exons and has multiple transcription start sites (Chai *et al.* 2012). It is approximately 462 kb away from the 3' end of the *FXN* gene. *APBA1* belongs to X11 protein family. The X11 protein family are multidomain proteins composed of conserved phosphotyrosine-binding (PTB) domain (Rogelj *et al.* 2006). They are involved in formation of multiprotein complexes and two of the family members, X11α and X11β, are expressed primarily in neurones. In humans, *APBA1* gene also encodes X11α protein. X11α is a brain specific multi-modular protein that interacts with the Alzheimer's disease amyloid precursor protein (APP). Aggregation of amyloid-β peptide (Aβ), an APP cleavage product, is believed to be central to the pathogenesis of Alzheimer's disease (Chai *et al.* 2012).

Since I found some interesting results with *PGM5* and PI5K1B gene, I then extended my investigations to quantify the *APBA1* gene levels. I used same cell lines as before and experiments were repeated at least twice. The analysis of *APBA1* mRNA quantifications revealed a significant reduction in *APBA1* levels in FRDA fibroblasts. The *APBA1* levels were reduced by 50% ($p<0.01$) and *FXN* levels were reduced by 82% ($p<0.001$) in FRDA compared to control fibroblasts (Figure 6.7). These results lead to my speculation that there is perhaps either a synergistic or added effect of GAA repeats and *FXN* gene silencing contributing to *APBA1* gene regulation.

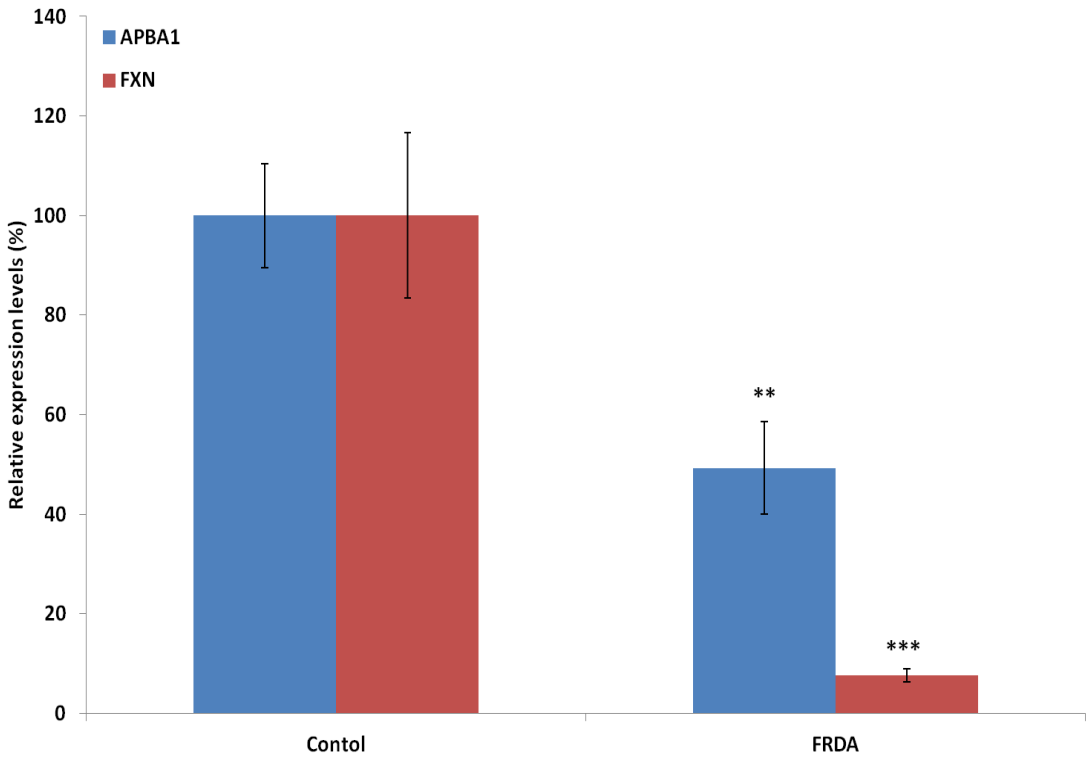

Figure 6.7 – Relative *APBA1* and *FXN* mRNA levels in unaffected and FRDA patient derived fibroblasts (n=3-8), **p<0.01, ***p<0.001 error bars ± 1 SEM).

6.2.6 - *PTAR1* gene

PTAR1 (prenyltransferase alpha subunit repeat containing 1) is located on the chromosome 9 at 9q21.12 position. It is approximately 50 kb in size and contains 8 exons. *PTAR1* is about 743 kb away from 3' end of the *FXN* gene. The exact function of *PTAR1* is unknown and therefore it is now included in the human collaborative consensus coding sequence (CCDs) group project. The key goal of this project is to identify and accurately annotate all protein-coding genes. The quantification of *PTAR1* mRNA by qRT-PCR revealed a 23% (ns) reduction in *PTAR1* levels in FRDA fibroblast cells (Figure 6.8).

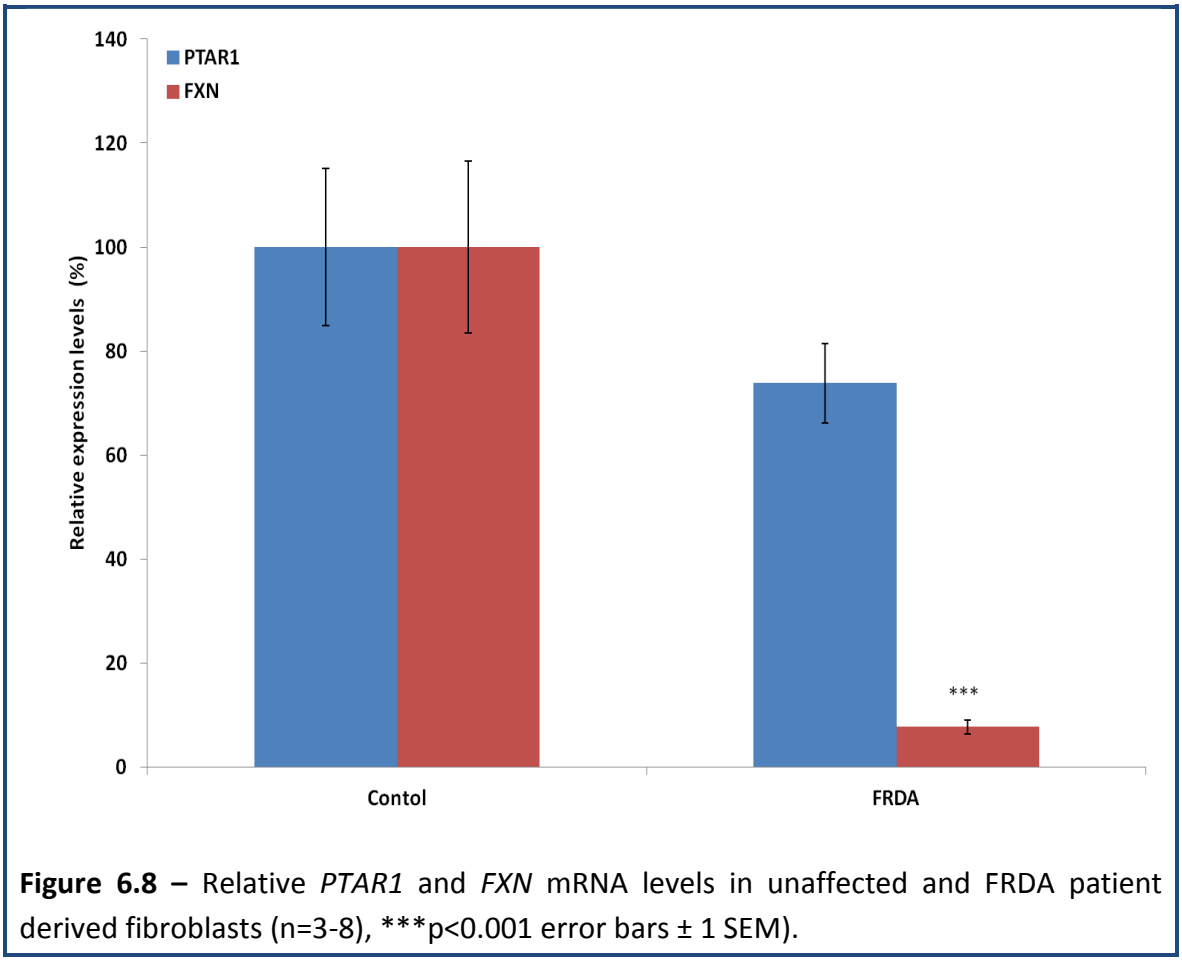

Figure 6.8 – Relative *PTAR1* and *FXN* mRNA levels in unaffected and FRDA patient derived fibroblasts (n=3-8), ***p<0.001 error bars ± 1 SEM).

6.3 - Discussion

Increased importance of epigenetic changes in FRDA and growing evidence on long-range gene regulation led me to study the adjacent genes rather than just focusing on *FXN* gene. Therefore, I chose a region of approximately 750 kb on both sides of the *FXN* gene that covered three genes on 5' end and four genes on 3' end of the *FXN* gene. Strikingly, I found that two genes (I did not managed to amplify *PRKACG*) on 5' end of the *FXN* gene (*PGM5* and *PIP5K1β*) have significantly downregulated in FRDA fibroblasts. These results are in agreement with the recent epigenetic changes identified in FRDA, where upstream of the GAA repeats in intron 1 of the *FXN* gene have shown significant enrichment of gene repressive histone marks and increased DNA methylation (Greene *et al.* 2007; Al-Mahdawi *et al.* 2008; Kim *et al.* 2011; Evans-Galea *et al.* 2012). Therefore, it is more likely that these repressive epigenetic marks may have an effect on its neighbouring genes. Moreover, the repressive epigenetic marks are less pronounced on the downstream of the GAA repeats. This could perhaps support our findings that only one gene (*APBA1*) out of 4 genes studied were downregulated in the 3' end of the *FXN* gene. Since there is no effect on the expression of two proximal genes of 3' end of *FXN* gene, it is more unlikely that the reduced *APBA1* expression is due to the GAA repeat expansion but there are likely to be other factors at play.

The underlying mechanisms that explain whether the downregulation of the *FXN* neighbouring genes are due to solely acquiring gene repressive signals from a heterochromatinized *FXN* promoter gene or just a general downregulation of genes is yet to investigated. However, it has been recently reported that depletion of CTCF, an 11 zinc

finger highly conserved transcription factor, in the 5' UTR region of expanded *FXN* alleles may initiate spreading of repressive chromatin from GAA expansion by increasing the expression of *FAST-1* (De Biase *et al.* 2009). Systematic genome-wide mapping by ChIP analysis has shown that CTCF binds to tens of thousands of genomic sites, often distal to transcription start sites of genes, but also at promoter and enhancer sequences (Barski *et al.* 2007; Holwerda and de Laat 2013). Although it is too early to predict a mechanistic link between *FXN* gene silencing, due to expanded GAA repeats, and the concurrent reduction of its neighbouring genes, but it might be an involvement of one or more of the above mentioned epigenetic changes are at play. Therefore, it would be paramount important to identify such interactions by using transgene or reporter assay techniques.

One of the key features of the transcription factory, an active gene transcription unit that is clustered in a discrete site within the eukaryotic nucleus, is that disparate transcription units can share the same factory at a given frequency, a concept initially referred as 'aggregation neighbouring active genes' (Cook 2002). Using a combination of RNA FISH/RNAPII Ser-5 co-localisation and 3C (chromosome conformation capture) assays, further revealed that co-localisation of 'neighbouring genes' can share the same factories at higher than expected frequencies (Osborne *et al.* 2004). Similarly, evidence has recently been confirmed that one transcription factor can regulate multiple genes. Therefore, based on the above observations, one possible speculation is that maybe the genes on 5' end of the *FXN* gene share the same transcription factors and downregulation of one gene may concomitantly affect other genes. It is also important to note that the co-localisation of active loci is not limited to genes on the same chromosome (*cis*) but can be expanded to the whole genome (*trans*), although *trans*-co-localisation occurs with reduced frequency

(Schoenfelder *et al.* 2010; Wei *et al.* 2013). Very recently, a patient with 46,XY gonadal disorder of sex development has been shown to have a 454 kb duplication within 9q21.11, involving the five genes *PIP5K1β*, *PRKACG*, *FXN*, *TJP2* and *FAM189A2* (Norling *et al.* 2013), indicating a cumulative contribution of this set of genes together to disease progression.

In conclusion, I have identified that *PGM5* and *PIP5K1β* genes, located at 5' end of the *FXN* genes, were downregulated, although the potential implications of these genes in FRDA disease progression/pathogenesis requires further investigation. Since I have analysed the *FXN* flanking gene signature only in fibroblasts, it would also be important to investigate the similar changes in disease relevant tissues, ideally neuronal and cardiac tissues, from FRDA patients. The implications of the *APBA1* gene in Alzheimer's disease are well documented and therefore a further link now needs to be established for FRDA. Nevertheless, our data shed new light on the consequences of the *FXN* gene expression in patient cells and might provide unexpected responses to some of the unanswered questions raised in FRDA.

Chapter 7 - General Discussion and Conclusions

Friedreich ataxia is an autosomal recessive neurodegenerative disease affecting the central and peripheral nervous system for which currently there is no therapy. The most common mutation causing the disease is a large trinucleotide GAA repeat expansion in the first intron of *FXN* gene encoding 'frataxin', a mitochondrial protein. The subsequent deficiency of frataxin protein leads to neurological disability, increased risk of diabetes mellitus, cardiomyopathy and premature death. FRDA is the commonest hereditary ataxia, accounting for approximately half of all inherited ataxias. Although the mechanism by which the GAA repeat expansion induces silencing of *FXN* gene is currently not known, recent in vitro and in vivo investigations has been put forward that FRDA is associated with non-B DNA conformation and/or heterochromatin-mediated *FXN* gene silencing. In support to the later hypothesis, differential DNA methylation of CpG islands and histone modifications such as increased methylation and decreased acetylation have been reported in FRDA.

In recent years, attention has been drawn to the fact that a significant fraction of the transcriptome comprises RNAs containing sequences that are complementary to other endogenous RNAs. These natural antisense transcripts (NATs) can have protein-coding properties but mainly represent non-coding RNAs (ncRNAs). Antisense RNA transcripts, which play a role in gene regulation, have previously been associated with number of microsatellite repeat expansion diseases such as Huntington disease (HD) (Chung *et al.* 2011), FRAXA (Ladd *et al.* 2007; Khalil *et al.* 2008), SCA7 (Sopher *et al.* 2011), SCA8 (Moseley *et al.* 2006) and DM1 (Cho *et al.* 2005; Yu *et al.* 2011). The discovery of *FAST-1* is therefore not surprising in itself. In general, the level of antisense transcription is significantly lower than that of the coding sense transcripts. Nevertheless, multiple reports have recently shown that antisense transcripts can either be involved in inhibition of the same gene from

where they originate (*cis*-acting) or inhibition of genes at different locations *(trans*-acting) (Carmichael 2003; Chen *et al.* 2004; Wang *et al.* 2005). Therefore, the study of antisense transcription in gene silencing machinery may provide further insight into mechanisms of neurodegenerative disorders, including FRDA. The exact mechanism of how increased *FAST-1* expression leads to *FXN* gene silencing is unknown, but it has been suggested that depletion of CTCF in FRDA eventually results in disease associated epigenetic changes for the transcriptional repression of the *FXN* gene. In support of this hypothesis, loss of CTCF binding site in DM1 CTG expansion is associated with the spread of heterochromatin and DNA methylation (Cho *et al.* 2005). However, how CTCF is depleted in FRDA and how this leads to *FXN* gene silencing is still unclear, there are number of reports that have been identified which potentially contribute to the further understanding of *FAST-1.*

With thousands of documented ncRNAs, pervasive transcription has been described in virtually all eukaryotic organisms (Berretta and Morillon 2009; Mercer *et al.* 2009). The increased number of genome wide analyses, including recently published ENCODE results (Bernstein *et al.* 2012; Harrow *et al.* 2012), have indicated that the genome is surprisingly complex. It is also believed that due to this complex nature of the genome there is a shift in gene organization from linear to modular model, in which it is possible for a sequence to be transcribed into a range of sense and antisense, coding and non-coding transcripts (see review Mercer *et al.* 2009). Although majority of the ncRNA functions are not known, steady growth in understanding these large unknown areas of the genome put forward some functional characteristics. These investigations have revealed that most of the ncRNAs regulate transcription via chromatin modulation (Nagano and Fraser 2011; Kugel and Goodrich 2012; Rinn and Chang 2012; Geisler and Coller 2013). For instance, HOTAIR (HOX transcript antisense RNA) is associated with Polycomb repressive complex 2 (PRC2) and

Lysine-specific demethylase 1 (LSD1). PRC2 is known to recruit the repressive histone mark, H3K27me3, and LSD1 is responsible for removal of active H3K4me2. Based on these findings it could be more intriguing to speculate that *FAST-1* may perhaps somehow adopt epigenetic modulation in FRDA that leads to *FXN* gene silencing.

Transcription of the genome is tightly controlled by number of factors. As soon as nascent pre-mRNAs are generated, they undergo post-transcriptional modifications and are spliced into one of many potential isoforms. However, antisense RNA and pre-mRNA have been suggested to form higher order RNA:RNA duplexes, which in turn inhibits the splicing due to spliceosome inaccessibility (Krystal *et al.* 1990; Munroe and Lazar 1991). A similar mechanism may be applied to *FAST-1* (Figure 7.1) and this needs further investigations. The emerging understanding of the functional characteristics of NATs has also revealed that large number of NATs are also associated with spliceosomes and regulate/influence the alternative splicing (Munroe and Lazar 1991). In addition, antisense transcripts have also been found to regulate transcription factors of a gene to control its expression (Bumgarner *et al.* 2009; Bumgarner *et al.* 2012). For example, the human dihydrofolate reductase (*DHFR)* gene encodes a lncRNA, which binds to the promoter and the transcription factor IIB, forming a complex. Through these interactions, the lncRNA disassembles the complex, which in turn leads to suppression of *DHFR* expression (Martianov *et al.* 2007). It has recently been identified that two transcription factors, SRF and TFAP2, as well as EGR3-like sequence, that work together to regulate expression of the *FXN* gene (Li *et al.* 2010). Since the locations of these transcription factors fall within the vicinity of the recently identified *FAST-1* location, it is more tempting to speculate that *FAST-1* may have an effect on these transcription factors that in turn contribute to the silencing of *FXN* gene expression (Figure 7.1).

Since there was very limited information available about the *FAST-1* (De Biase *et al.* 2009) and the lack of key information, such as length and the location, to understand its role in *FXN* gene silencing, I utilised the Northern blotting technique. In an attempt to identify the *FAST-1* size, Northern blot hybridisation with two different sized radiolabelled riboprobes (238 bp and 675 bp), resulted in the identification of two different sized bands, one at ~500 bp and the other one at over 9 kb size. Subsequently, using RACE experiment coupled with cloning and sequencing, I identified the 5'- and 3'- ends of *FAST-1* and plotted the full length *FAST-1* as 523 bp size. These results support my initial findings with Northern blot (~500 bp size band), thus confirming that 523 bp sequence may only be one potential *FAST-1* transcript in FRDA. However, we cannot rule out the possibility that alternative splicing or different transcription start sites may generate different sized transcripts and hence I observed two bands in Northern blotting. To support this hypothesis, exonic splicing enhancer regulatory elements within the context of bidirectional coding sequences have shown alternative splicing of an antisense mRNA of TRα gene which resulted in two different transcripts with altered sizes (Salato *et al.* 2010). Furthermore, I have also identified that the 5'-end of the *FAST-1* occupies the recently identified CTCF binding site at 5' UTR region of *FXN* gene suggesting the possible involvement of *FAST-1* in CTCF disposition and epigenetic changes that occurring at the vicinity of the *FXN* gene (De Biase *et al.* 2009; Al-Mahdawi *et al.* 2013), as discussed earlier.

To further improve the efficacy of *FAST-1* identification and subsequent quantifications, I also developed a qRT-PCR based method. Due to the sensitivity and the complex nature of this experiment I have optimised several approaches to make sure that this technique is reproducible. Using the qRT-PCR method I have first identified that the *FAST-1* levels were significantly increased in FRDA fibroblasts, consistent with previous

results identified by Sanjay and colleagues (De Biase *et al.* 2009). Extended analysis of *FAST-1* in FRDA YAC mouse model cell lines and tissues have revealed increased *FAST-1* and reduced *FXN* gene expression levels in primary fibroblasts, and brain, cerebellum and liver tissues of YG8R mice. However, in other mice I found no direct correlation, except in few occasions, between *FAST-1* and *FXN* gene. Increased *FAST-1* expression in one particular cell line or tissue did not show any decrease in *FXN* expression. Similar results were also identified with other mouse models, such FRDA BAC transgenic mouse model and double transgenic FRDA YAC mouse models (Bmi-1 het KO and HP1γ het KO). Therefore, in order to identify the effect of *FAST-1* in *FXN* expression, it is paramount important to identify the interactions between these two genes. Also, since the full sequence of *FAST-1* is identified, it would be more beneficial to construct an expression plasmid clone to overexpress *FAST-1* and identify the subsequent consequences on *FXN* gene expression. Similarly you could knock down expression of *FAST-1* and look at the potential effects of this on *FXN* expression. We cannot rule out the possibility that, since the *FAST-1* expression is very low compared to *FXN*, it might not have any direct effect on *FXN* gene but may be affecting other *FXN* gene regulating factors such as controlling the transcription factors of the gene. Antisense expression has been shown to regulate transcription initiation by affecting DNA methylation, which usually leads to their long-term repression (Lister *et al.* 2009). R-loops (RNA:DNA hybrids) are induced at GAA repeats following *in vitro* transcription and in bacteria (Grabczyk *et al.* 2007; Reddy *et al.* 2011), indicating that GAA repeats in FRDA patients are associated with increased number of R-loops. Importantly, increased numbers of R-loops in FRDA are associated with *FXN* gene silencing (Groh *et al.* 2014), and also R-loops are shown to increase antisense transcription (Skourti-Stathaki *et al.* 2014). These

findings can help to pave new paths/mechanisms for precise understanding of the series events occurring at the *FXN* gene locus and subsequent gene silencing (Figure 7.1).

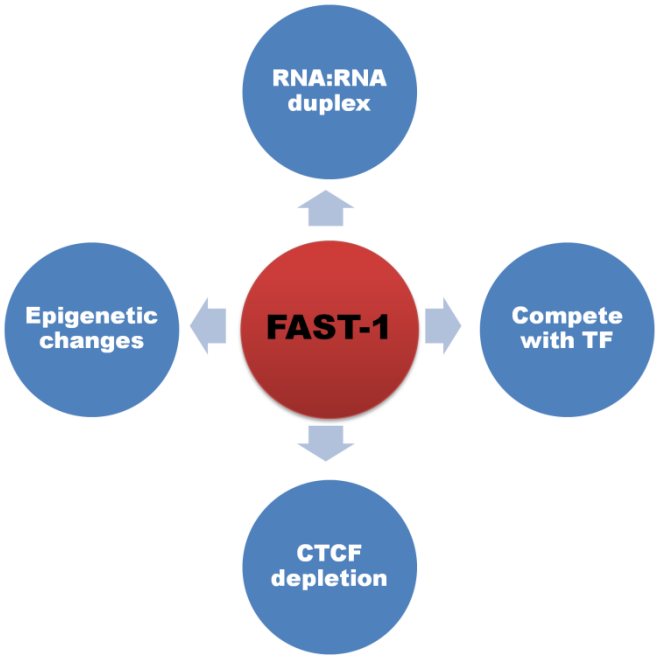

Figure 7.1 – Putative *FAST-1* mechanisms by which *FXN* gene silencing is achieved in FRDA. It has been well documented that antisense transcripts form higher order RNA:RNA duplexes with nascent pre-mRNA of the sense strand. In addition, number of other mechanisms also proposed which includes, antisense transcripts are known to compete with the transcription factors (TF) thereby reduce gene expression. Depletion of CTCF followed by increased gene repressive histone marks and elevated levels of DNA methylation is another hallmark of the antisense transcripts, like *FAST-1*, which in turn maintains long-range gene repression.

Finally, the analysis of *FXN* flanking genes has revealed that there could be, potentially, an effect of GAA repeat mediated gene silencing on the 5'-end of the *FXN* gene. However, the potential implications of these genes in FRDA disease progression and pathogenesis require further investigation. For example, *PIP5K1β* depletion in FRDA cells is associated with cytoskeleton anomalies, including destabilisation of actin network and

delayed cell spreading (Bayot *et al.* 2013). Similarly, identification of functions of other genes in FRDA cells may further strengthen our understanding of FRDA. Remarkably, the expression of some antisense transcripts is linked to the activity of neighbouring genes (Xu *et al.* 2011; Sigova *et al.* 2013). Therefore, it would also important to investigate the effect of *FAST-1* gene in its neighbouring genes and this can be achieved by knockdown of *FAST-1* expression.

In conclusion, using a combination of Northern blotting technique and RACE PCR, I have identified full length *FAST-1* transcript. To hasten the quantification of *FAST-1*, a robust qRT-PCR based method has been established and validated. I have also further identified that *FAST-1* expression levels were increased in human and mouse fibroblasts that correlated with reduced *FXN* expression. In addition, YG8R mice tissues have also shown consistently increased *FAST-1* together with low levels of *FXN* expression, further confirming the hypothesis that *FAST-1* may have direct or indirect role in *FXN* gene repression by one of many mechanisms discussed. Importantly, inhibitions of NATs by single-stranded oligonucleotides or siRNAs can transiently upregulate locus-specific gene expression (Modarresi *et al.* 2012). Therefore, inhibition of *FAST-1* may prove to be an effective target for FRDA therapy since, by yet unknown mechanism, *FAST-1* is associated with *FXN* gene silencing. Furthermore, antigene RNA (agRNA), a small duplex RNAs of 19 bp in length that target gene promoters, can target both sense and antisense stands, and coding and non-coding transcripts (Watts *et al.* 2010). Therefore, the use of agRNA to activate *FXN* gene by targeting *FXN* promoter or the *FAST-1* transcript may prove useful, since both mechanisms may be involved in reversing *FXN* gene silencing and thus ameliorating FRDA disease.

References

Adinolfi S., Iannuzzi C., Prischi F., Pastore C., Iametti S., Martin S. R., Bonomi F. and Pastore A. (2009). "Bacterial frataxin CyaY is the gatekeeper of iron-sulfur cluster formation catalyzed by IscS." Nat Struct Mol Biol **16**(4): 390-396.

Adinolfi S., Trifuoggi M., Politou A. S., Martin S. and Pastore A. (2002). "A structural approach to understanding the iron-binding properties of phylogenetically different frataxins." Hum Mol Genet **11**(16): 1865-1877.

Ahmad N., Wang Y., Haider K. H., Wang B., Pasha Z., Uzun O. and Ashraf M. (2006). "Cardiac protection by mitoKATP channels is dependent on Akt translocation from cytosol to mitochondria during late preconditioning." Am J Physiol Heart Circ Physiol **290**(6): H2402-2408.

Al-Mahdawi S., Pinto R. M., Ismail O., Varshney D., Lymperi S., Sandi C., Trabzuni D. and Pook M. (2008). "The Friedreich ataxia GAA repeat expansion mutation induces comparable epigenetic changes in human and transgenic mouse brain and heart tissues." Hum Mol Genet **17**(5): 735-746.

Al-Mahdawi S., Pinto R. M., Ruddle P., Carroll C., Webster Z. and Pook M. (2004). "GAA repeat instability in Friedreich ataxia YAC transgenic mice." Genomics **84**(2): 301-310.

Al-Mahdawi S., Pinto R. M., Varshney D., Lawrence L., Lowrie M. B., Hughes S., Webster Z., Blake J., Cooper J. M., King R.*, et al.* (2006). "GAA repeat expansion mutation mouse models of Friedreich ataxia exhibit oxidative stress leading to progressive neuronal and cardiac pathology." Genomics **88**(5): 580-590.

Al-Mahdawi S., Sandi C., Mouro Pinto R. and Pook M. A. (2013). "Friedreich Ataxia Patient Tissues Exhibit Increased 5-Hydroxymethylcytosine Modification and Decreased CTCF Binding at the FXN Locus." PLoS One **8**(9): e74956.

Babcock M., de Silva D., Oaks R., Davis-Kaplan S., Jiralerspong S., Montermini L., Pandolfo M. and Kaplan J. (1997). "Regulation of mitochondrial iron accumulation by Yfh1p, a putative homolog of frataxin." Science **276**(5319): 1709-1712.

Barski A., Cuddapah S., Cui K., Roh T. Y., Schones D. E., Wang Z., Wei G., Chepelev I. and Zhao K. (2007). "High-resolution profiling of histone methylations in the human genome." Cell **129**(4): 823-837.

Bayot A., Reichman S., Lebon S., Csaba Z., Aubry L., Sterkers G., Husson I., Rak M. and Rustin P. (2013). "Cis-silencing of PIP5K1B evidenced in Friedreich's ataxia patient cells results in cytoskeleton anomalies." Hum Mol Genet **22**(14): 2894-2904.

Beisel C. and Paro R. (2011). "Silencing chromatin: comparing modes and mechanisms." Nat Rev Genet **12**(2): 123-135.

Beiter T., Reich E., Williams R. and Simon P. (2009). "Antisense transcription: A critical look in both directions." Cellular and Molecular Life Sciences **66**(1): 94-112.

Belkin A. M. and Burridge K. (1994). "Expression and localization of the phosphoglucomutase-related cytoskeletal protein, aciculin, in skeletal muscle." J Cell Sci **107 (Pt 7)**: 1993-2003.

Belkin A. M. and Burridge K. (1995). "Localization of utrophin and aciculin at sites of cell-matrix and cell-cell adhesion in cultured cells." Exp Cell Res **221**(1): 132-140.

Beltran M., Puig I., Pena C., Garcia J. M., Alvarez A. B., Pena R., Bonilla F. and de Herreros A. G. (2008). "A natural antisense transcript regulates Zeb2/Sip1 gene expression during Snail1-induced epithelial-mesenchymal transition." Genes Dev **22**(6): 756-769.

Bernstein B. E., Birney E., Dunham I., Green E. D., Gunter C. and Snyder M. (2012). "An integrated encyclopedia of DNA elements in the human genome." Nature **489**(7414): 57-74.

Berretta J. and Morillon A. (2009). "Pervasive transcription constitutes a new level of eukaryotic genome regulation." EMBO Rep **10**(9): 973-982.

Bidichandani S. I., Ashizawa T. and Patel P. I. (1997). "Atypical Friedreich ataxia caused by compound heterozygosity for a novel missense mutation and the GAA triplet-repeat expansion." Am J Hum Genet **60**(5): 1251-1256.

Bidichandani S. I., Purandare S. M., Taylor E. E., Gumin G., Machkhas H., Harati Y., Gibbs R. A., Ashizawa T. and Patel P. I. (1999). "Somatic sequence variation at the Friedreich ataxia locus includes complete contraction of the expanded GAA triplet repeat, significant length variation in serially passaged lymphoblasts and enhanced mutagenesis in the flanking sequence." Hum Mol Genet **8**(13): 2425-2436.

Boddaert N., Le Quan Sang K. H., Rotig A., Leroy-Willig A., Gallet S., Brunelle F., Sidi D., Thalabard J. C., Munnich A. and Cabantchik Z. I. (2007). "Selective iron chelation in Friedreich ataxia: biologic and clinical implications." Blood **110**(1): 401-408.

Boesch S., Sturm B., Hering S., Scheiber-Mojdehkar B., Steinkellner H., Goldenberg H. and Poewe W. (2008). "Neurological effects of recombinant human erythropoietin in Friedreich's ataxia: a clinical pilot trial." Mov Disord **23**(13): 1940-1944.

Bogoyevitch M. A. (2004). "An update on the cardiac effects of erythropoietin cardioprotection by erythropoietin and the lessons learnt from studies in neuroprotection." Cardiovasc Res **63**(2): 208-216.

Bourn R. L., De Biase I., Pinto R. M., Sandi C., Al-Mahdawi S., Pook M. A. and Bidichandani S. I. (2012). "Pms2 suppresses large expansions of the (GAA.TTC)n sequence in neuronal tissues." PLoS One **7**(10): e47085.

Bradley J. L., Blake J. C., Chamberlain S., Thomas P. K., Cooper J. M. and Schapira A. H. (2000). "Clinical, biochemical and molecular genetic correlations in Friedreich's ataxia." Hum Mol Genet **9**(2): 275-282.

Bryan J. and Aguilar-Bryan L. (1999). "Sulfonylurea receptors: ABC transporters that regulate ATP-sensitive K(+) channels." Biochim Biophys Acta **1461**(2): 285-303.

Bumgarner S. L., Dowell R. D., Grisafi P., Gifford D. K. and Fink G. R. (2009). "Toggle involving cis-interfering noncoding RNAs controls variegated gene expression in yeast." Proc Natl Acad Sci U S A **106**(43): 18321-18326.

Bumgarner S. L., Neuert G., Voight B. F., Symbor-Nagrabska A., Grisafi P., van Oudenaarden A. and Fink G. R. (2012). "Single-cell analysis reveals that noncoding RNAs contribute to clonal heterogeneity by modulating transcription factor recruitment." Mol Cell **45**(4): 470-482.

Campuzano V., Montermini L., Lutz Y., Cova L., Hindelang C., Jiralerspong S., Trottier Y., Kish S. J., Faucheux B., Trouillas P.*, et al.* (1997). "Frataxin is reduced in Friedreich ataxia patients and is associated with mitochondrial membranes." Hum Mol Genet **6**(11): 1771-1780.

Campuzano V., Montermini L., Molto M. D., Pianese L., Cossee M., Cavalcanti F., Monros E., Rodius F., Duclos F., Monticelli A.*, et al.* (1996). "Friedreich's ataxia: autosomal recessive disease caused by an intronic GAA triplet repeat expansion." Science **271**(5254): 1423-1427.

Cao G., Gu M., Zhu M., Gao J., Yin Y., Marshall C., Xiao M., Ding J. and Miao D. (2012). "Bmi-1 absence causes premature brain degeneration." PLoS ONE **7**(2): e32015.

Carmichael G. G. (2003). "Antisense starts making more sense." Nat Biotech **21**(4): 371-372.

Carvajal J. J., Pook M. A., dos Santos M., Doudney K., Hillermann R., Minogue S., Williamson R., Hsuan J. J. and Chamberlain S. (1996). "The Friedreich's ataxia gene encodes a novel phosphatidylinositol-4- phosphate 5-kinase." Nat Genet **14**(2): 157-162.

Carvajal J. J., Pook M. A., Doudney K., Hillermann R., Wilkes D., al-Mahdawi S., Williamson R. and Chamberlain S. (1995). "Friedreich's ataxia: a defect in signal transduction?" Hum Mol Genet **4**(8): 1411-1419.

Cavadini P., O'Neill H. A., Benada O. and Isaya G. (2002). "Assembly and iron-binding properties of human frataxin, the protein deficient in Friedreich ataxia." Hum Mol Genet **11**(3): 217-227.

Chai K. H., McLoughlin D. M., Chan T. F., Chan H. Y. and Lau K. F. (2012). "Genomic organization and promoter cloning of the human X11alpha gene APBA1." DNA Cell Biol **31**(5): 651-659.

Chamberlain S., Shaw J., Rowland A., Wallis J., South S., Nakamura Y., von Gabain A., Farrall M. and Williamson R. (1988). "Mapping of mutation causing Friedreich's ataxia to human chromosome 9." Nature **334**(6179): 248-250.

Chan P. K., Torres R., Yandim C., Law P. P., Khadayate S., Mauri M., Grosan C., Chapman-Rothe N., Giunti P., Pook M.*, et al.* (2013). "Heterochromatinization induced by GAA-

repeat hyperexpansion in Friedreich's ataxia can be reduced upon HDAC inhibition by vitamin B3." Hum Mol Genet **22**(13): 2662-2675.

Chantrel-Groussard K., Geromel V., Puccio H., Koenig M., Munnich A., Rotig A. and Rustin P. (2001). "Disabled early recruitment of antioxidant defenses in Friedreich's ataxia." Hum Mol Genet **10**(19): 2061-2067.

Chatoo W., Abdouh M., David J., Champagne M. P., Ferreira J., Rodier F. and Bernier G. (2009). "The polycomb group gene Bmi1 regulates antioxidant defenses in neurons by repressing p53 pro-oxidant activity." J Neurosci **29**(2): 529-542.

Chen J., Sun M., Kent W. J., Huang X., Xie H., Wang W., Zhou G., Shi R. Z. and Rowley J. D. (2004). "Over 20% of human transcripts might form sense-antisense pairs." Nucleic Acids Res **32**(16): 4812-4820.

Cho D. H., Thienes C. P., Mahoney S. E., Analau E., Filippova G. N. and Tapscott S. J. (2005). "Antisense transcription and heterochromatin at the DM1 CTG repeats are constrained by CTCF." Mol Cell **20**(3): 483-489.

Christophe D., Cabrer B., Bacolla A., Targovnik H., Pohl V. and Vassart G. (1985). "An unusually long poly(purine)-poly(pyrimidine) sequence is located upstream from the human thyroglobulin gene." Nucleic Acids Res **13**(14): 5127-5144.

Chu C., Qu K., Zhong F. L., Artandi S. E. and Chang H. Y. (2011). "Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions." Mol Cell **44**(4): 667-678.

Chung D. W., Rudnicki D. D., Yu L. and Margolis R. L. (2011). "A natural antisense transcript at the Huntington's disease repeat locus regulates HTT expression." Hum Mol Genet **20**(17): 3467-3477.

Chutake Y. K., Costello W. N., Lam C. and Bidichandani S. I. (2014). "Altered nucleosome positioning at the transcription start site and deficient transcriptional initiation in friedreich ataxia." J Biol Chem **289**(22): 15194-15202.

Condo I., Ventura N., Malisan F., Rufini A., Tomassini B. and Testi R. (2007). "In vivo maturation of human frataxin." Hum Mol Genet **16**(13): 1534-1540.

Cook P. R. (2002). "Predicting three-dimensional genome structure from transcriptional activity." Nat Genet **32**(3): 347-352.

Cossee M., Puccio H., Gansmuller A., Koutnikova H., Dierich A., LeMeur M., Fischbeck K., Dolle P. and Koenig M. (2000). "Inactivation of the Friedreich ataxia mouse gene leads to early embryonic lethality without iron accumulation." Hum Mol Genet **9**(8): 1219-1226.

Cossee M., Schmitt M., Campuzano V., Reutenauer L., Moutou C., Mandel J. L. and Koenig M. (1997). "Evolution of the Friedreich's ataxia trinucleotide repeat expansion: founder effect and premutations." Proc Natl Acad Sci U S A **94**(14): 7452-7457.

De Biase I., Chutake Y. K., Rindler P. M. and Bidichandani S. I. (2009). "Epigenetic Silencing in Friedreich Ataxia Is Associated with Depletion of CTCF (CCCTC-Binding Factor) and Antisense Transcription." PLoS ONE **4**(11): e7914.

De Biase I., Rasmussen A., Endres D., Al-Mahdawi S., Monticelli A., Cocozza S., Pook M. and Bidichandani S. I. (2007). "Progressive GAA expansions in dorsal root ganglia of Friedreich's ataxia patients." Ann Neurol **61**(1): 55-60.

Delatycki M. B., Camakaris J., Brooks H., Evans-Whipp T., Thorburn D. R., Williamson R. and Forrest S. M. (1999). "Direct evidence that mitochondrial iron accumulation occurs in Friedreich ataxia." Ann Neurol **45**(5): 673-675.

Delatycki M. B., Paris D. B., Gardner R. J., Nicholson G. A., Nassif N., Storey E., MacMillan J. C., Collins V., Williamson R. and Forrest S. M. (1999). "Clinical and genetic study of Friedreich ataxia in an Australian population." Am J Med Genet **87**(2): 168-174.

Delatycki M. B., Williamson R. and Forrest S. M. (2000). "Friedreich ataxia: an overview." J Med Genet **37**(1): 1-8.

Deng W. and Blobel G. A. (2010). "Do chromatin loops provide epigenetic gene expression states?" Curr Opin Genet Dev **20**(5): 548-554.

Dhe-Paganon S., Shigeta R., Chi Y. I., Ristow M. and Shoelson S. E. (2000). "Crystal structure of human frataxin." J Biol Chem **275**(40): 30753-30756.

Di Prospero N. A. and Fischbeck K. H. (2005). "Therapeutics development for triplet repeat expansion diseases." Nat Rev Genet **6**(10): 756-765.

Dillon N. and Festenstein R. (2002). "Unravelling heterochromatin: competition between positive and negative factors regulates accessibility." Trends Genet **18**(5): 252-258.

Djebali S., Davis C. A., Merkel A., Dobin A., Lassmann T., Mortazavi A., Tanzer A., Lagarde J., Lin W., Schlesinger F*., et al.* (2012). "Landscape of transcription in human cells." Nature **489**(7414): 101-108.

Durr A., Cossee M., Agid Y., Campuzano V., Mignard C., Penet C., Mandel J. L., Brice A. and Koenig M. (1996). "Clinical and genetic abnormalities in patients with Friedreich's ataxia." N Engl J Med **335**(16): 1169-1175.

Edwards Y. H., Putt W., Fox M. and Ives J. H. (1995). "A novel human phosphoglucomutase (PGM5) maps to the centromeric region of chromosome 9." Genomics **30**(2): 350-353.

Emond M., Lepage G., Vanasse M. and Pandolfo M. (2000). "Increased levels of plasma malondialdehyde in Friedreich ataxia." Neurology **55**(11): 1752-1753.

Engel N., Thorvaldsen J. L. and Bartolomei M. S. (2006). "CTCF binding sites promote transcription initiation and prevent DNA methylation on the maternal allele at the imprinted H19/Igf2 locus." Hum Mol Genet **15**(19): 2945-2954.

Ensembl (2006). "Human Contig View, accessed October 2006 URL<http://www.ensembl.org/Homo_sapiens/Search/Results?species=Homo_sapiens;idx=;q=frataxin>."

Evans-Galea M. V., Carrodus N., Rowley S. M., Corben L. A., Tai G., Saffery R., Galati J. C., Wong N. C., Craig J. M., Lynch D. R., *et al.* (2012). "FXN methylation predicts expression and clinical outcome in Friedreich ataxia." Ann Neurol **71**(4): 487-497.

Ezzatizadeh V., Pinto R. M., Sandi C., Sandi M., Al-Mahdawi S., Te Riele H. and Pook M. A. (2012). "The mismatch repair system protects against intergenerational GAA repeat instability in a Friedreich ataxia mouse model." Neurobiol Dis **46**(1): 165-171.

Ezzatizadeh V., Sandi C., Sandi M., Anjomani-Virmouni S., Al-Mahdawi S. and Pook M. A. (2014). "MutLalpha heterodimers modify the molecular phenotype of Friedreich ataxia." PLoS ONE **9**(6): e100523.

Fanning A. S., Van Itallie C. M. and Anderson J. M. (2012). "Zonula occludens-1 and -2 regulate apical cell structure and the zonula adherens cytoskeleton in polarized epithelia." Mol Biol Cell **23**(4): 577-590.

Filippova G. N., Cheng M. K., Moore J. M., Truong J. P., Hu Y. J., Nguyen D. K., Tsuchiya K. D. and Disteche C. M. (2005). "Boundaries between chromosomal domains of X inactivation and escape bind CTCF and lack CpG methylation during early development." Dev Cell **8**(1): 31-42.

Filippova G. N., Thienes C. P., Penn B. H., Cho D. H., Hu Y. J., Moore J. M., Klesert T. R., Lobanenkov V. V. and Tapscott S. J. (2001). "CTCF-binding sites flank CTG/CAG repeats and form a methylation-sensitive insulator at the DM1 locus." Nat Genet **28**(4): 335-343.

Filla A., De Michele G., Cavalcanti F., Pianese L., Monticelli A., Campanella G. and Cocozza S. (1996). "The relationship between trinucleotide (GAA) repeat length and clinical features in Friedreich ataxia." Am J Hum Genet **59**(3): 554-560.

Frank-Kamenetskii M. D. and Mirkin S. M. (1995). "Triplex DNA structures." Annu Rev Biochem **64**: 65-95.

Gatchel J. R. and Zoghbi H. Y. (2005). "Diseases of unstable repeat expansion: mechanisms and common principles." Nat Rev Genet **6**(10): 743-755.

Geisler S. and Coller J. (2013). "RNA in unexpected places: long non-coding RNA functions in diverse cellular contexts." Nat Rev Mol Cell Biol **14**(11): 699-712.

Gelfand B., Mead J., Bruning A., Apostolopoulos N., Tadigotla V., Nagaraj V., Sengupta A. M. and Vershon A. K. (2011). "Regulated antisense transcription controls expression of cell-type-specific genes in yeast." Mol Cell Biol **31**(8): 1701-1709.

Gellera C., Castellotti B., Mariotti C., Mineri R., Seveso V., Didonato S. and Taroni F. (2007). "Frataxin gene point mutations in Italian Friedreich ataxia patients." Neurogenetics **8**(4): 289-299.

Goffin J. and Eisenhauer E. (2002). "DNA methyltransferase inhibitors-state of the art." Ann Oncol **13**(11): 1699-1716.

Gonzalez-Mariscal L., Bautista P., Lechuga S. and Quiros M. (2012). "ZO-2, a tight junction scaffold protein involved in the regulation of cell proliferation and apoptosis." Ann N Y Acad Sci **1257**: 133-141.

Gordon D. M., Shi Q., Dancis A. and Pain D. (1999). "Maturation of frataxin within mammalian and yeast mitochondria: one-step processing by matrix processing peptidase." Hum Mol Genet **8**(12): 2255-2262.

Grabczyk E., Mancuso M. and Sammarco M. C. (2007). "A persistent RNA.DNA hybrid formed by transcription of the Friedreich ataxia triplet repeat in live bacteria, and by T7 RNAP in vitro." Nucleic Acids Res **35**(16): 5351-5359.

Grabczyk E. and Usdin K. (2000). "The GAA*TTC triplet repeat expanded in Friedreich's ataxia impedes transcription elongation by T7 RNA polymerase in a length and supercoil dependent manner." Nucleic Acids Res **28**(14): 2815-2822.

Greene E., Entezam A., Kumari D. and Usdin K. (2005). "Ancient repeated DNA elements and the regulation of the human frataxin promoter." Genomics **85**(2): 221-230.

Greene E., Mahishi L., Entezam A., Kumari D. and Usdin K. (2007). "Repeat-induced epigenetic changes in intron 1 of the frataxin gene and its consequences in Friedreich ataxia." Nucleic Acids Res **35**(10): 3383-3390.

Grewal S. I. and Elgin S. C. (2007). "Transcription and RNA interference in the formation of heterochromatin." Nature **447**(7143): 399-406.

Grewal S. I. and Jia S. (2007). "Heterochromatin revisited." Nat Rev Genet **8**(1): 35-46.

Groh M., Lufino M. M., Wade-Martins R. and Gromak N. (2014). "R-loops associated with triplet repeat expansions promote gene silencing in Friedreich ataxia and fragile X syndrome." PLoS Genet **10**(5): e1004318.

Guldstrand M., Grill V., Bjorklund A., Lins P. E. and Adamson U. (2002). "Improved beta cell function after short-term treatment with diazoxide in obese subjects with type 2 diabetes." Diabetes Metab **28**(6 Pt 1): 448-456.

Guttman M. and Rinn J. L. (2012). "Modular regulatory principles of large non-coding RNAs." Nature **482**(7385): 339-346.

Hahn M., Dambacher S. and Schotta G. (2010). "Heterochromatin dysregulation in human diseases." J Appl Physiol (1985) **109**(1): 232-242.

Harding A. E. (1981). "Friedreich's ataxia: a clinical and genetic study of 90 families with an analysis of early diagnostic criteria and intrafamilial clustering of clinical features." Brain **104**(3): 589-620.

Harrow J., Frankish A., Gonzalez J. M., Tapanari E., Diekhans M., Kokocinski F., Aken B. L., Barrell D., Zadissa A., Searle S.*, et al.* (2012). "GENCODE: The reference human genome annotation for The ENCODE Project." Genome Res **22**(9): 1760-1774.

Hart P. E., Lodi R., Rajagopalan B., Bradley J. L., Crilley J. G., Turner C., Blamire A. M., Manners D., Styles P., Schapira A. H.*, et al.* (2005). "Antioxidant treatment of patients with Friedreich ataxia: four-year follow-up." Arch Neurol **62**(4): 621-626.

Haugen A. C., Di Prospero N. A., Parker J. S., Fannin R. D., Chou J., Meyer J. N., Halweg C., Collins J. B., Durr A., Fischbeck K.*, et al.* (2010). "Altered gene expression and DNA damage in peripheral blood cells from Friedreich's ataxia patients: cellular model of pathology." PLoS Genet **6**(1): e1000812.

Hayry V., Tynninen O., Haapasalo H. K., Wolfer J., Paulus W., Hasselblatt M., Sariola H., Paetau A., Sarna S., Niemela M.*, et al.* (2008). "Stem cell protein BMI-1 is an independent marker for poor prognosis in oligodendroglial tumours." Neuropathol Appl Neurobiol **34**(5): 555-563.

He Y., Vogelstein B., Velculescu V. E., Papadopoulos N. and Kinzler K. W. (2008). "The antisense transcriptomes of human cells." Science **322**(5909): 1855-1857.

Heneka M. T. and Landreth G. E. (2007). "PPARs in the brain." Biochim Biophys Acta **1771**(8): 1031-1045.

Herman D., Jenssen K., Burnett R., Soragni E., Perlman S. L. and Gottesfeld J. M. (2006). "Histone deacetylase inhibitors reverse gene silencing in Friedreich's ataxia." Nat Chem Biol **2**(10): 551-558.

Holwerda S. J. B. and de Laat W. (2013). "CTCF: the protein, the binding partners, the binding sites and their chromatin loops." Philosophical Transactions of the Royal Society B: Biological Sciences **368**(1620).

Huang Q., Bu S., Yu Y., Guo Z., Ghatnekar G., Bu M., Yang L., Lu B., Feng Z., Liu S.*, et al.* (2007). "Diazoxide Prevents Diabetes through Inhibiting Pancreatic β-Cells from Apoptosis via Bcl-2/Bax Rate and p38-β Mitogen-Activated Protein Kinase." Endocrinology **148**(1): 81-91.

Iida T., Nakayama J. and Moazed D. (2008). "siRNA-mediated heterochromatin establishment requires HP1 and is associated with antisense transcription." Mol Cell **31**(2): 178-189.

Jacobs J. J., Kieboom K., Marino S., DePinho R. A. and van Lohuizen M. (1999). "The oncogene and Polycomb-group gene bmi-1 regulates cell proliferation and senescence through the ink4a locus." Nature **397**(6715): 164-168.

Jauslin M. L., Meier T., Smith R. A. and Murphy M. P. (2003). "Mitochondria-targeted antioxidants protect Friedreich Ataxia fibroblasts from endogenous oxidative stress more effectively than untargeted antioxidants." Faseb J **17**(13): 1972-1974.

Jiralerspong S., Ge B., Hudson T. J. and Pandolfo M. (2001). "Manganese superoxide dismutase induction by iron is impaired in Friedreich ataxia cells." FEBS Lett **509**(1): 101-105.

Johnson J. M., Edwards S., Shoemaker D. and Schadt E. E. (2005). "Dark matter in the genome: evidence of widespread transcription detected by microarray tiling experiments." Trends Genet **21**(2): 93-102.

Kelly D. P. and Scarpulla R. C. (2004). "Transcriptional regulatory circuits controlling mitochondrial biogenesis and function." Genes Dev **18**(4): 357-368.

Khalil A. M., Faghihi M. A., Modarresi F., Brothers S. P. and Wahlestedt C. (2008). "A novel RNA transcript with antiapoptotic function is silenced in fragile X syndrome." PLoS One **3**(1): e1486.

Kim E., Napierala M. and Dent S. Y. (2011). "Hyperexpansion of GAA repeats affects post-initiation steps of FXN transcription in Friedreich's ataxia." Nucleic Acids Res **39**(19): 8366-8377.

Koeppen A. H., Michael S. C., Knutson M. D., Haile D. J., Qian J., Levi S., Santambrogio P., Garrick M. D. and Lamarche J. B. (2007). "The dentate nucleus in Friedreich's ataxia: the role of iron-responsive proteins." Acta Neuropathol **114**(2): 163-173.

Koeppen A. H., Morral J. A., Davis A. N., Qian J., Petrocine S. V., Knutson M. D., Gibson W. M., Cusack M. J. and Li D. (2009). "The dorsal root ganglion in Friedreich's ataxia." Acta Neuropathol **118**(6): 763-776.

Koutnikova H., Campuzano V., Foury F., Dolle P., Cazzalini O. and Koenig M. (1997). "Studies of human, mouse and yeast homologues indicate a mitochondrial function for frataxin." Nat Genet **16**(4): 345-351.

Koutnikova H., Campuzano V. and Koenig M. (1998). "Maturation of wild-type and mutated frataxin by the mitochondrial processing peptidase." Hum Mol Genet **7**(9): 1485-1489.

Kouzarides T. (2007). "Chromatin modifications and their function." Cell **128**(4): 693-705.

Krystal G. W., Armstrong B. C. and Battey J. F. (1990). "N-myc mRNA forms an RNA-RNA duplex with endogenous antisense transcripts." Mol Cell Biol **10**(8): 4180-4191.

Kugel J. F. and Goodrich J. A. (2012). "Non-coding RNAs: key regulators of mammalian transcription." Trends Biochem Sci **37**(4): 144-151.

Kwon G., Marshall C. A., Liu H., Pappan K. L., Remedi M. S. and McDaniel M. L. (2006). "Glucose-stimulated DNA synthesis through mammalian target of rapamycin (mTOR)

is regulated by KATP channels: effects on cell cycle progression in rodent islets." J Biol Chem **281**(6): 3261-3267.

Ladd P. D., Smith L. E., Rabaia N. A., Moore J. M., Georges S. A., Hansen R. S., Hagerman R. J., Tassone F., Tapscott S. J. and Filippova G. N. (2007). "An antisense transcript spanning the CGG repeat region of FMR1 is upregulated in premutation carriers but silenced in full mutation individuals." Hum Mol Genet **16**(24): 3174-3187.

Li K., Singh A., Crooks D. R., Dai X., Cong Z., Pan L., Ha D. and Rouault T. A. (2010). "Expression of Human Frataxin Is Regulated by Transcription Factors SRF and TFAP2." PLoS ONE **5**(8): e12286.

Li W., Maeda Y., Yuan R. R., Elkabes S., Cook S. and Dowling P. (2004). "Beneficial effect of erythropoietin on experimental allergic encephalomyelitis." Ann Neurol **56**(6): 767-777.

Libby R. T., Hagerman K. A., Pineda V. V., Lau R., Cho D. H., Baccam S. L., Axford M. M., Cleary J. D., Moore J. M., Sopher B. L*., et al.* (2008). "CTCF cis-regulates trinucleotide repeat instability in an epigenetic manner: a novel basis for mutational hot spot determination." PLoS Genet **4**(11): e1000257.

Libri V., Yandim C., Athanasopoulos S., Loyse N., Natisvili T., Law P. P., Chan P. K., Mohammad T., Mauri M., Tam K. T*., et al.* (2014). "Epigenetic and neurological effects and safety of high-dose nicotinamide in patients with Friedreich's ataxia: an exploratory, open-label, dose-escalation study." Lancet **384**(9942): 504-513.

Lister R., Pelizzola M., Dowen R. H., Hawkins R. D., Hon G., Tonti-Filippini J., Nery J. R., Lee L., Ye Z., Ngo Q. M*., et al.* (2009). "Human DNA methylomes at base resolution show widespread epigenomic differences." Nature **462**(7271): 315-322.

Lobanenkov V. V., Nicolas R. H., Adler V. V., Paterson H., Klenova E. M., Polotskaja A. V. and Goodwin G. H. (1990). "A novel sequence-specific DNA binding protein which interacts with three regularly spaced direct repeats of the CCCTC-motif in the 5'-flanking sequence of the chicken c-myc gene." Oncogene **5**(12): 1743-1753.

Lodi R., Cooper J. M., Bradley J. L., Manners D., Styles P., Taylor D. J. and Schapira A. H. (1999). "Deficit of in vivo mitochondrial ATP production in patients with Friedreich ataxia." Proc Natl Acad Sci U S A **96**(20): 11492-11495.

Ma Z., Portwood N., Brodin D., Grill V. and Bjorklund A. (2007). "Effects of diazoxide on gene expression in rat pancreatic islets are largely linked to elevated glucose and potentially serve to enhance beta-cell sensitivity." Diabetes **56**(4): 1095-1106.

Mariappan S. V., Catasti P., Silks L. A., 3rd, Bradbury E. M. and Gupta G. (1999). "The high-resolution structure of the triplex formed by the GAA/TTC triplet repeat associated with Friedreich's ataxia." J Mol Biol **285**(5): 2035-2052.

Marmolino D. (2011). "Friedreich's ataxia: past, present and future." Brain Res Rev **67**(1-2): 311-330.

Marmolino D. and Acquaviva F. (2009). "Friedreich's Ataxia: from the (GAA)n repeat mediated silencing to new promising molecules for therapy." Cerebellum **8**(3): 245-259.

Marmolino D., Acquaviva F., Pinelli M., Monticelli A., Castaldo I., Filla A. and Cocozza S. (2009). "PPAR-gamma agonist Azelaoyl PAF increases frataxin protein and mRNA expression: new implications for the Friedreich's ataxia therapy." Cerebellum **8**(2): 98-103.

Martianov I., Ramadass A., Serra Barros A., Chow N. and Akoulitchev A. (2007). "Repression of the human dihydrofolate reductase gene by a non-coding interfering transcript." Nature **445**(7128): 666-670.

Mercer T. R., Dinger M. E. and Mattick J. S. (2009). "Long non-coding RNAs: insights into functions." Nat Rev Genet **10**(3): 155-159.

Miranda C. J., Santos M. M., Ohshima K., Smith J., Li L., Bunting M., Cossee M., Koenig M., Sequeiros J., Kaplan J.*, et al.* (2002). "Frataxin knockin mouse." FEBS Lett **512**(1-3): 291-297.

Modarresi F., Faghihi M. A., Lopez-Toledano M. A., Fatemi R. P., Magistri M., Brothers S. P., van der Brug M. P. and Wahlestedt C. (2012). "Inhibition of natural antisense transcripts in vivo results in gene-specific transcriptional upregulation." Nat Biotechnol **30**(5): 453-459.

Molofsky A. V., He S., Bydon M., Morrison S. J. and Pardal R. (2005). "Bmi-1 promotes neural stem cell self-renewal and neural development but not mouse growth and survival by repressing the p16Ink4a and p19Arf senescence pathways." Genes Dev **19**(12): 1432-1437.

Monros E., Molto M. D., Martinez F., Canizares J., Blanca J., Vilchez J. J., Prieto F., de Frutos R. and Palau F. (1997). "Phenotype correlation and intergenerational dynamics of the Friedreich ataxia GAA trinucleotide repeat." Am J Hum Genet **61**(1): 101-110.

Montermini L., Andermann E., Labuda M., Richter A., Pandolfo M., Cavalcanti F., Pianese L., Iodice L., Farina G., Monticelli A.*, et al.* (1997). "The Friedreich ataxia GAA triplet repeat: premutation and normal alleles." Hum Mol Genet **6**(8): 1261-1266.

Montermini L., Richter A., Morgan K., Justice C. M., Julien D., Castellotti B., Mercier J., Poirier J., Capozzoli F., Bouchard J. P.*, et al.* (1997). "Phenotypic variability in Friedreich ataxia: role of the associated GAA triplet repeat expansion." Ann Neurol **41**(5): 675-682.

Morris A. R., Mukherjee N. and Keene J. D. (2008). "Ribonomic analysis of human Pum1 reveals cis-trans conservation across species despite evolution of diverse mRNA target sets." Mol Cell Biol **28**(12): 4093-4103.

Moseley M. L., Zu T., Ikeda Y., Gao W., Mosemiller A. K., Daughters R. S., Chen G., Weatherspoon M. R., Clark H. B., Ebner T. J.*, et al.* (2006). "Bidirectional expression

of CUG and CAG expansion transcripts and intranuclear polyglutamine inclusions in spinocerebellar ataxia type 8." Nat Genet **38**(7): 758-769.

Munroe S. H. and Lazar M. A. (1991). "Inhibition of c-erbA mRNA splicing by a naturally occurring antisense RNA." J Biol Chem **266**(33): 22083-22086.

Musco G., Stier G., Kolmerer B., Adinolfi S., Martin S., Frenkiel T., Gibson T. and Pastore A. (2000). "Towards a structural understanding of Friedreich's ataxia: the solution structure of frataxin." Structure **8**(7): 695-707.

Nagano T. and Fraser P. (2011). "No-nonsense functions for long noncoding RNAs." Cell **145**(2): 178-181.

Norling A., Linden Hirschberg A., Iwarsson E., Persson B., Wedell A. and Barbaro M. (2013). "Novel candidate genes for 46,XY gonadal dysgenesis identified by a customized 1 M array-CGH platform." Eur J Med Genet **56**(12): 661-668.

Osborne C. S., Chakalova L., Brown K. E., Carter D., Horton A., Debrand E., Goyenechea B., Mitchell J. A., Lopes S., Reik W*., et al.* (2004). "Active genes dynamically colocalize to shared sites of ongoing transcription." Nat Genet **36**(10): 1065-1071.

Pandolfo M. (2002). "The molecular basis of Friedreich ataxia." Adv Exp Med Biol **516**: 99-118.

Pandolfo M. (2009). "Friedreich ataxia: the clinical picture." J Neurol **256 Suppl 1**: 3-8.

Pandolfo M. and Pastore A. (2009). "The pathogenesis of Friedreich ataxia and the structure and function of frataxin." J Neurol **256 Suppl 1**: 9-17.

Park I. K., Morrison S. J. and Clarke M. F. (2004). "Bmi1, stem cells, and senescence regulation." J Clin Invest **113**(2): 175-179.

Pelechano V. and Steinmetz L. M. (2013). "Gene regulation by antisense transcription." Nat Rev Genet **14**(12): 880-893.

Phillips J. E. and Corces V. G. (2009). "CTCF: master weaver of the genome." Cell **137**(7): 1194-1211.

Pianese L., Cavalcanti F., De Michele G., Filla A., Campanella G., Calabrese O., Castaldo I., Monticelli A. and Cocozza S. (1997). "The effect of parental gender on the GAA dynamic mutation in the FRDA gene." Am J Hum Genet **60**(2): 460-463.

Pianese L., Tammaro A., Turano M., De Biase I., Monticelli A. and Cocozza S. (2002). "Identification of a novel transcript of X25, the human gene involved in Friedreich ataxia." Neurosci Lett **320**(3): 137-140.

Pook M. A., Al-Mahdawi S., Carroll C. J., Cossee M., Puccio H., Lawrence L., Clark P., Lowrie M. B., Bradley J. L., Cooper J. M*., et al.* (2001). "Rescue of the Friedreich's ataxia knockout mouse by human YAC transgenesis." Neurogenetics **3**(4): 185-193.

Pook M. A., Carvajal J. J., Doudney K., Hillermann R. and Chamberlain S. (1997). "Exon-intron structure of a 2.7-kb transcript of the STM7 gene with phosphatidylinositol-4-phosphate 5-kinase activity." Genomics **42**(1): 170-172.

Puccio H., Simon D., Cossee M., Criqui-Filipe P., Tiziano F., Melki J., Hindelang C., Matyas R., Rustin P. and Koenig M. (2001). "Mouse models for Friedreich ataxia exhibit cardiomyopathy, sensory nerve defect and Fe-S enzyme deficiency followed by intramitochondrial iron deposits." Nat Genet **27**(2): 181-186.

Puspasari N., Rowley S. M., Gordon L., Lockhart P. J., Ioannou P. A., Delatycki M. B. and Sarsero J. P. (2011). "Long range regulation of human FXN gene expression." PLoS ONE **6**(7): e22001.

Radisky D. C., Babcock M. C. and Kaplan J. (1999). "The yeast frataxin homologue mediates mitochondrial iron efflux. Evidence for a mitochondrial iron cycle." J Biol Chem **274**(8): 4497-4499.

Radtke M. A., Nermoen I., Kollind M., Skeie S., Sorheim J. I., Svartberg J., Hals I., Moen T., Dorflinger G. H. and Grill V. (2010). "Six months of diazoxide treatment at bedtime in newly diagnosed subjects with type 1 diabetes does not influence parameters of {beta}-cell function and autoimmunity but improves glycemic control." Diabetes Care **33**(3): 589-594.

Rai M., Soragni E., Chou C. J., Barnes G., Jones S., Rusche J. R., Gottesfeld J. M. and Pandolfo M. (2010). "Two new pimelic diphenylamide HDAC inhibitors induce sustained frataxin upregulation in cells from Friedreich's ataxia patients and in a mouse model." PLoS One **5**(1): e8825.

Rai M., Soragni E., Jenssen K., Burnett R., Herman D., Coppola G., Geschwind D. H., Gottesfeld J. M. and Pandolfo M. (2008). "HDAC inhibitors correct frataxin deficiency in a Friedreich ataxia mouse model." PLoS ONE **3**(4): e1958.

Reddy K., Tam M., Bowater R. P., Barber M., Tomlinson M., Nichol Edamura K., Wang Y. H. and Pearson C. E. (2011). "Determinants of R-loop formation at convergent bidirectionally transcribed trinucleotide repeats." Nucleic Acids Res **39**(5): 1749-1762.

Richardson D. R., Mouralian C., Ponka P. and Becker E. (2001). "Development of potential iron chelators for the treatment of Friedreich's ataxia: ligands that mobilize mitochondrial iron." Biochim Biophys Acta **1536**(2-3): 133-140.

Richter B., Bandeira-Echtler E., Bergerhoff K., Clar C. and Ebrahim S. H. (2007). "Rosiglitazone for type 2 diabetes mellitus." Cochrane Database Syst Rev(3): CD006063.

Riessland M., Brichta L., Hahnen E. and Wirth B. (2006). "The benzamide M344, a novel histone deacetylase inhibitor, significantly increases SMN2 RNA/protein levels in spinal muscular atrophy cells." Hum Genet **120**(1): 101-110.

Rinn J. L. and Chang H. Y. (2012). "Genome regulation by long noncoding RNAs." Annu Rev Biochem **81**: 145-166.

Robertson K. D. (2001). "DNA methylation, methyltransferases, and cancer." Oncogene **20**(24): 3139-3155.

Rogelj B., Mitchell J. C., Miller C. C. and McLoughlin D. M. (2006). "The X11/Mint family of adaptor proteins." Brain Res Rev **52**(2): 305-315.

Rosok O. and Sioud M. (2004). "Systematic identification of sense-antisense transcripts in mammalian cells." Nat Biotechnol **22**(1): 104-108.

Rotig A., de Lonlay P., Chretien D., Foury F., Koenig M., Sidi D., Munnich A. and Rustin P. (1997). "Aconitase and mitochondrial iron-sulphur protein deficiency in Friedreich ataxia." Nat Genet **17**(2): 215-217.

Rudolf J., Makrantoni V., Ingledew W. J., Stark M. J. and White M. F. (2006). "The DNA repair helicases XPD and FancJ have essential iron-sulfur domains." Mol Cell **23**(6): 801-808.

Rustin P., Munnich A. and Rotig A. (1999). "Quinone analogs prevent enzymes targeted in Friedreich ataxia from iron-induced injury in vitro." Biofactors **9**(2-4): 247-251.

Sakamoto N., Chastain P. D., Parniewski P., Ohshima K., Pandolfo M., Griffith J. D. and Wells R. D. (1999). "Sticky DNA: self-association properties of long GAA.TTC repeats in R.R.Y triplex structures from Friedreich's ataxia." Mol Cell **3**(4): 465-475.

Sakamoto N., Larson J. E., Iyer R. R., Montermini L., Pandolfo M. and Wells R. D. (2001). "GGA*TCC-interrupted triplets in long GAA*TTC repeats inhibit the formation of triplex and sticky DNA structures, alleviate transcription inhibition, and reduce genetic instabilities." J Biol Chem **276**(29): 27178-27187.

Salato V. K., Rediske N. W., Zhang C., Hastings M. L. and Munroe S. H. (2010). "An exonic splicing enhancer within a bidirectional coding sequence regulates alternative splicing of an antisense mRNA." RNA Biol **7**(2): 179-190.

Sambrotta M., Strautnieks S., Papouli E., Rushton P., Clark B. E., Parry D. A., Logan C. V., Newbury L. J., Kamath B. M., Ling S*., et al.* (2014). "Mutations in TJP2 cause progressive cholestatic liver disease." Nat Genet **46**(4): 326-328.

Sandi C., Al-Mahdawi S. and Pook M. A. (2013). "Epigenetics in Friedreich's Ataxia: Challenges and Opportunities for Therapy." Genet Res Int **2013**: 852080.

Sandi C., Pinto R. M., Al-Mahdawi S., Ezzatizadeh V., Barnes G., Jones S., Rusche J. R., Gottesfeld J. M. and Pook M. A. (2011). "Prolonged treatment with pimelic o-aminobenzamide HDAC inhibitors ameliorates the disease phenotype of a Friedreich ataxia mouse model." Neurobiol Dis **42**(3): 496-505.

Sandi C., Sandi M., Anjomani-Virmouni S., Al-Mahdawi S. and Pook M. A. (2014). "Epigenetic-based therapies for Friedreich ataxia." Frontiers in Genetics **5**.

Sandi C., Sandi M., Jassal H., Ezzatizadeh V., Anjomani-Virmouni S., Al-Mahdawi S. and Pook M. A. (2014). "Generation and Characterisation of Friedreich Ataxia YG8R Mouse Fibroblast and Neural Stem Cell Models." PLoS One **9**(2): e89488.

Santos R., Lefevre S., Sliwa D., Seguin A., Camadro J. M. and Lesuisse E. (2010). "Friedreich ataxia: molecular mechanisms, redox considerations, and therapeutic opportunities." Antioxid Redox Signal **13**(5): 651-690.

Sarsero J. P., Li L., Holloway T. P., Voullaire L., Gazeas S., Fowler K. J., Kirby D. M., Thorburn D. R., Galle A., Cheema S*., et al.* (2004). "Human BAC-mediated rescue of the Friedreich ataxia knockout mutation in transgenic mice." Mamm Genome **15**(5): 370-382.

Saveliev A., Everett C., Sharpe T., Webster Z. and Festenstein R. (2003). "DNA triplet repeats mediate heterochromatin-protein-1-sensitive variegated gene silencing." Nature **422**(6934): 909-913.

Schmucker S., Argentini M., Carelle-Calmels N., Martelli A. and Puccio H. (2008). "The in vivo mitochondrial two-step maturation of human frataxin." Hum Mol Genet **17**(22): 3521-3531.

Schmucker S., Martelli A., Colin F., Page A., Wattenhofer-Donze M., Reutenauer L. and Puccio H. (2011). "Mammalian frataxin: an essential function for cellular viability through an interaction with a preformed ISCU/NFS1/ISD11 iron-sulfur assembly complex." PLoS One **6**(1): e16199.

Schmucker S. and Puccio H. (2010). "Understanding the molecular mechanisms of Friedreich's ataxia to develop therapeutic approaches." Hum Mol Genet **19**(R1): R103-110.

Schoenfelder S., Sexton T., Chakalova L., Cope N. F., Horton A., Andrews S., Kurukuti S., Mitchell J. A., Umlauf D., Dimitrova D. S*., et al.* (2010). "Preferential associations between co-regulated genes reveal a transcriptional interactome in erythroid cells." Nat Genet **42**(1): 53-61.

Schulz J. B., Boesch S., Burk K., Durr A., Giunti P., Mariotti C., Pousset F., Schols L., Vankan P. and Pandolfo M. (2009). "Diagnosis and treatment of Friedreich ataxia: a European perspective." Nat Rev Neurol **5**(4): 222-234.

Schulz J. B., Dehmer T., Schols L., Mende H., Hardt C., Vorgerd M., Burk K., Matson W., Dichgans J., Beal M. F*., et al.* (2000). "Oxidative stress in patients with Friedreich ataxia." Neurology **55**(11): 1719-1721.

Sharma R., Bhatti S., Gomez M., Clark R. M., Murray C., Ashizawa T. and Bidichandani S. I. (2002). "The GAA triplet-repeat sequence in Friedreich ataxia shows a high level of somatic instability in vivo, with a significant predilection for large contractions." Hum Mol Genet **11**(18): 2175-2187.

Shearwin K. E., Callen B. P. and Egan J. B. (2005). "Transcriptional interference--a crash course." Trends Genet **21**(6): 339-345.

Sigova A. A., Mullen A. C., Molinie B., Gupta S., Orlando D. A., Guenther M. G., Almada A. E., Lin C., Sharp P. A., Giallourakis C. C.*, et al.* (2013). "Divergent transcription of long noncoding RNA/mRNA gene pairs in embryonic stem cells." Proc Natl Acad Sci U S A **110**(8): 2876-2881.

Simon D., Seznec H., Gansmuller A., Carelle N., Weber P., Metzger D., Rustin P., Koenig M. and Puccio H. (2004). "Friedreich ataxia mouse models with progressive cerebellar and sensory ataxia reveal autophagic neurodegeneration in dorsal root ganglia." J Neurosci **24**(8): 1987-1995.

Skourti-Stathaki K., Kamieniarz-Gdula K. and Proudfoot N. J. (2014). "R-loops induce repressive chromatin marks over mammalian gene terminators." Nature.

Son L. S., Bacolla A. and Wells R. D. (2006). "Sticky DNA: in vivo formation in E. coli and in vitro association of long GAA*TTC tracts to generate two independent supercoiled domains." J Mol Biol **360**(2): 267-284.

Sopher B. L., Ladd P. D., Pineda V. V., Libby R. T., Sunkin S. M., Hurley J. B., Thienes C. P., Gaasterland T., Filippova G. N. and La Spada A. R. (2011). "CTCF regulates ataxin-7 expression through promotion of a convergently transcribed, antisense noncoding RNA." Neuron **70**(6): 1071-1084.

Soragni E., Miao W., Iudicello M., Jacoby D., De Mercanti S., Clerico M., Longo F., Piga A., Ku S., Campau E.*, et al.* (2014). "Epigenetic therapy for Friedreich ataxia." Ann Neurol **76**(4): 489-508.

Sturm B., Stupphann D., Kaun C., Boesch S., Schranzhofer M., Wojta J., Goldenberg H. and Scheiber-Mojdehkar B. (2005). "Recombinant human erythropoietin: effects on frataxin expression in vitro." Eur J Clin Invest **35**(11): 711-717.

Suzuki M., Saito T., Sato T., Tamagawa M., Miki T., Seino S. and Nakaya H. (2003). "Cardioprotective effect of diazoxide is mediated by activation of sarcolemmal but not mitochondrial ATP-sensitive potassium channels in mice." Circulation **107**(5): 682-685.

Talbert P. B. and Henikoff S. (2006). "Spreading of silent chromatin: inaction at a distance." Nat Rev Genet **7**(10): 793-803.

Tomassini B., Arcuri G., Fortuni S., Sandi C., Ezzatizadeh V., Casali C., Condo I., Malisan F., Al-Mahdawi S., Pook M.*, et al.* (2012). "Interferon gamma upregulates frataxin and corrects the functional deficits in a Friedreich ataxia model." Hum Mol Genet **21**(13): 2855-2861.

Tufarelli C., Stanley J. A., Garrick D., Sharpe J. A., Ayyub H., Wood W. G. and Higgs D. R. (2003). "Transcription of antisense RNA leading to gene silencing and methylation as a novel cause of human genetic disease." Nat Genet **34**(2): 157-165.

van den Bout I. and Divecha N. (2009). "PIP5K-driven PtdIns(4,5)P2 synthesis: regulation and cellular functions." J Cell Sci **122**(Pt 21): 3837-3850.

van Heyningen V. and Bickmore W. (2013). "Regulation from a distance: long-range control of gene expression in development and disease." Philos Trans R Soc Lond B Biol Sci **368**(1620): 20120372.

Waldvogel D., van Gelderen P. and Hallett M. (1999). "Increased iron in the dentate nucleus of patients with Friedrich's ataxia." Ann Neurol **46**(1): 123-125.

Wang X. J., Gaasterland T. and Chua N. H. (2005). "Genome-wide prediction and identification of cis-natural antisense transcripts in Arabidopsis thaliana." Genome Biol **6**(4): R30.

Watts J. K., Yu D., Charisse K., Montaillier C., Potier P., Manoharan M. and Corey D. R. (2010). "Effect of chemical modifications on modulation of gene expression by duplex antigene RNAs that are complementary to non-coding transcripts at gene promoters." Nucleic Acids Res **38**(15): 5242-5259.

Wei Z., Huang D., Gao F., Chang W. H., An W., Coetzee G. A., Wang K. and Lu W. (2013). "Biological implications and regulatory mechanisms of long-range chromosomal interactions." J Biol Chem **288**(31): 22369-22377.

Wilson R. B. and Roof D. M. (1997). "Respiratory deficiency due to loss of mitochondrial DNA in yeast lacking the frataxin homologue." Nat Genet **16**(4): 352-357.

Wong A., Yang J., Cavadini P., Gellera C., Lonnerdal B., Taroni F. and Cortopassi G. (1999). "The Friedreich's ataxia mutation confers cellular sensitivity to oxidant stress which is rescued by chelators of iron and calcium and inhibitors of apoptosis." Hum Mol Genet **8**(3): 425-430.

Wu Z., Puigserver P., Andersson U., Zhang C., Adelmant G., Mootha V., Troy A., Cinti S., Lowell B., Scarpulla R. C*., et al.* (1999). "Mechanisms controlling mitochondrial biogenesis and respiration through the thermogenic coactivator PGC-1." Cell **98**(1): 115-124.

Xia H., Cao Y., Dai X., Marelja Z., Zhou D., Mo R., Al-Mahdawi S., Pook M. A., Leimkühler S., Rouault T. A*., et al.* (2012). "Novel Frataxin Isoforms May Contribute to the Pathological Mechanism of Friedreich Ataxia." PLoS ONE **7**(10): e47847.

Xu Z., Wei W., Gagneur J., Clauder-Munster S., Smolik M., Huber W. and Steinmetz L. M. (2011). "Antisense expression increases gene expression variability and locus interdependency." Mol Syst Biol **7**: 468.

Yandim C., Natisvili T. and Festenstein R. (2013). "Gene regulation and epigenetics in Friedreich's ataxia." J Neurochem **126 Suppl 1**: 21-42.

Yoon T. and Cowan J. A. (2004). "Frataxin-mediated iron delivery to ferrochelatase in the final step of heme biosynthesis." J Biol Chem **279**(25): 25943-25946.

Yu Z., Teng X. and Bonini N. M. (2011). "Triplet repeat-derived siRNAs enhance RNA-mediated toxicity in a Drosophila model for myotonic dystrophy." PLoS Genet **7**(3): e1001340.