Investigating the role of *FXN* antisense transcript 1 in Friedreich ataxia

Hajar Mikaeili

College of Health and Life Sciences

Department of Life Sciences

Division of Biosciences

December 2017

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.

Hajar Mikaeili

Abstract

Friedreich ataxia (FRDA) is a neurodegenerative disorder that is inherited in an autosomal recessive pattern. The most common FRDA mutation is hyperexpansion of a GAA triplet repeat sequence in the first intron of the affected gene, frataxin (*FXN*), resulting in decreased frataxin protein expression. The hyperexpanded GAA repeats can adopt unusual DNA structures and induce aberrant epigenetic changes leading to heterochromatin mediated gene silencing. Several epigenetic changes, including increased levels of DNA methylation, histone modifications, repressive chromatin formation and elevated levels of non-coding RNA have been reported in FRDA. It has been reported that a novel *FXN* antisense transcript (*FAST-1*), is present at higher levels in FRDA patient-derived fibroblasts and its overexpression is associated with the depletion of CTCF, a chromatin insulator protein, and heterochromatin formation involving the critical +1 nucleosome. Previously, characteristics of *FAST-1* were investigated in our lab and a full-length *FAST-1* transcript containing a poly (A) tail was identified. To investigate any possible effects of *FAST-1* on *FXN* expression, I first overexpressed this *FAST-1* transcript in three different non-FRDA cell lines and a consistent decrease of *FXN* expression was observed in each cell type compared to control cells. I also identified that *FAST-1* copy number is positively correlated with increased *FAST-1* expression, which in turn is negatively correlated with *FXN* expression in *FAST-1* overexpressing cells. Additionally, we found that *FAST-1* overexpression is associated with increased levels of DNA methylation at CpG sites U6 and U11 of the *FXN* upstream GAA repeat region, together with CTCF depletion and heterochromatin formation at the 5′UTR of the *FXN* gene. To further investigate the role of *FAST-1* in *FXN* gene silencing, I used a small hairpin RNA (shRNA) strategy to knock down *FAST-1*

expression in FRDA fibroblast cells. I found that knocking down *FAST-1* increases *FXN* expression, but not to the level of control cells. Lastly, I investigated the pattern of *FAST-1* expression and histone modifications at the *FXN* transgene in our new FRDA mouse model, designated YG8LR. The YG8LR mice showed decreased levels of *FXN* expression and H3K9ac and increased levels of *FAST-1* expression and H3K9me3. Our data suggest that since *FAST-1* is associated with *FXN* gene silencing, inhibition of *FAST-1* may be an approach for FRDA therapy.

Acknowledgments

I would like to express my deepest gratitude to my supervisor Dr. Mark Pook for his advice, guidance and trust throughout my PhD. His understanding helped me during the hardest of times.

I must also thank Dr. Sahar Al-Mahdawi for her great support, mentorship and encouragement. She made a huge contribution for my development. I am also thankful to Sahar for performing the MethylScreen study.

I would like to thank Dr. Aurelien Bayot, my colleagues Mursal Sherzai and Saba Saqlain for their help and support. I also like to thank my friend Marianne Henry for her kind help during my PhD.

Lastly, I would like to thank my husband, Mohammad, for his patience and love without which I could not have managed to get through my PhD.

Abbreviations

µg	Microgram
µl	Microliter
Ago	Argonaute
AMV RT	Avian Myeloblastosis Virus Reverse Transcriptase
ANRIL	Antisense Non-Coding RNA in the INK4 Locus
BCA	Bicinchoninic Acid
BDNF	Brain-Derived Neurotrophic Factor
bp	Base Pair
BRD	Bromodomain
BSA	Bovine Serum Albumin
cDNA	Complementary DNA
ChIP	Chromatin Immunoprecipitation
Ct	Cycle Threshold
CMV	Cytomegalovirus
CTCF	CCCTC-Binding Factor
CO2	Carbon Dioxide
DAPI	4', 6-Diamidino-2-phenylindole
dH2O	Distilled Water
DM1	Myotonic Dystrophy
DMEM	Dulbecco's Modified Eagle's Medium
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic Acid
DNMT	DNA Methyl Transferases
dNTP	Deoxy Nucleotide Triphosphate
DTT	Dithiothreitol
E.coli	*Escherichia coli*
EDTA	Ethylene Diamine-Tetra Acetic Acid
EphB2	Ephrin Receptor B2
FAST-1	Frataxin Antisense Transcript-1
FBS	Fetal Bovine Serum

Fe-S	Iron-Sulphur
FRAXA	Fragile X Syndrome
FRDA	Friedreich Ataxia
FTLD	Frontotemporal Lobar Degeneration
FXN	Frataxin
FXTAS	Fragile X-Associated Tremor and Ataxia Syndrome
GDNF	Glial-Derived Neurotrophic Factor
GFP	Green Fluorescent Protein
HBA1	Hemoglobin Alpha 1
hCD2	Heterochromatin-Sensitive Human CD2
HCl	Hydrochloric Acid
HD	Huntington's Disease
HDACi	Histone Deacetylase Inhibitors
HEK293	Human Embryonic Kidney 293 Cells
HMTase	Histone Methyl Transferases
HOTAIR	HOX Transcript Antisense RNA
HP1	Heterochromatin Protein 1
HPRT	Hypoxanthine Guanine Phosphoribosyl Transferase
ISC	Iron Sulphur Cluster
Kb	Kilo Base
Kcnq1ot1	KCNQ1 Opposite Strand Transcript 1
kDa	Kilo Dalton
kg	Kilogram
LB medium	Luria-Bertani Medium
LV	Lentiviral Vector
M	Molar
mg	Milligram
min	Minutes
ml	Millilitre
MOI	Multiplicity of Infection
MPP	Mitochondrial Processing Peptidase
mRNA	Messenger RNA

NaCl	Sodium Chloride
NaOH	Sodium Hydroxide
NAT	Natural Antisense Transcript
ng	Nanograms
NSC	Neural Stem Cell
ODN	Oligodeoxyribonucleotides
Oct-4	Octamer-Binding Transcription Factor 4
PBS	Phosphate-Buffered Saline
PCR	Polymerase Chain Reaction
pen-strep	Penicillin and Streptomycin
PEV	Position Effect Variegation
PRC	Polycomb Repressive Complex
puro	Puromycin
qRT-PCR	Quantitative Reverse Transcription PCR
RAN	Repeat Associated Non-AUG
RITS	RNA-Induced Transcriptional Gene Silencing
RNA	Ribonucleic Acid
rpm	Revolutions Per Minute
rRNA	Ribosomal RNA
sec	Seconds
shRNA	Short Hairpin RNA
SCA2	Spinocerebellar Ataxia Type 2
Taq	*Thermus Aquaticus*
TBE	Tris-Borate-EDTA
TNR	Trinucleotide Repeat
TSG	Tumour Suppressor Gene
TSS	Transcription Start Site
U	Units
UTR	Untranslated Region
WT	Wild-Type
XIST	X-Inactive Specific Transcript
YAC	Yeast Artificial Chromosome

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

1.1 Friedreich ataxia

Friedreich ataxia (FRDA, OMIM #229300), the most common inherited ataxia, is an autosomal recessive neurodegenerative disorder. The first symptoms, which are limb and gait ataxia and loss of tendon reflexes in the lower limbs, appear before the age of 25, and then patients develop dysarthria, loss of ability to walk and loss of sensation. Scoliosis is also present in most case of FRDA and is clearly progressive. There is no treatment for FRDA and patients mostly die of progressive hypertrophic cardiomyopathy in their early adulthood. FRDA is a rare genetic disorder with an estimated prevalence of one case per 50,000 individuals. It is commonly caused by hyperexpansion of a GAA triplet repeat in the first intron of both copies of the *FXN* gene (Campuzano *et al.*, 1996). Only 5% of patients are heterozygotes for a point mutation in one allele and a GAA expansion on the other allele (Pandolfo, 2002; Gellera *et al.*, 2007). Normal individuals have up to ~40 triplets, whereas in FRDA patients the number of GAA repeat can be from 70 to 1700 (Dürr *et al.*, 1996; Pandolfo, 2002). The GAA repeat size determines the level of residual frataxin and is directly correlated to an earlier age of onset and more rapid rate of disease progression. The GAA repeat expansion causes a transcriptional defect of the *FXN* gene, and reduced levels of frataxin, a 210 amino acid mitochondrial protein, results in mitochondrial dysfunction and oxidative damage (Campuzano *et al.*, 1997; Pandolfo, 2008).

1.2 *FXN* gene structure and expression

In 1988 Chamberlain *et al* carried out genetic linkage analysis on 22 families with three or more affected siblings and detected that the mutation causing FRDA occurs in chromosome 9 (Chamberlain *et al.*, 1988). Further linkage studies on about 120 affected families localised the disease locus to chromosome 9q13-9q21 (Hanauer *et*

al., 1990). In 1996 the FRDA locus was narrowed down to 9q13 and a gene called *X25* was identified in that region (Campuzano *et al.*, 1996). The gene, initially called *X25* and later changed to *FRDA*, is now known as *FXN*. Using direct complementary DNA (cDNA) selection and exon amplification, the coding sequences of *FXN* gene were identified. An inverse polymerase chain reaction (PCR) approach was used to identify the intronic sequence flanking the known coding regions and therefore, the *FXN* gene structure was determined (Campuzano *et al.*, 1996). Screening the *FXN* gene for mutations in FRDA patients revealed that an expanded trinucleotide repeats (GAA) within the first intron of *FXN* gene accounted for about 98% of FRDA cases. The *FXN* gene is composed of seven exons spanning 95 kb of genomic DNA. Exons 1 to 5a produce the major transcript, a 1.3kb mRNA that encodes a 210-amino acid protein, called frataxin. Since the exon 5b has an in-frame stop codon, alternative splicing can give a rise to a 171-amino acid coding transcript (Campuzano *et al.*, 1996). Exon 6 is entirely noncoding. Two major transcription start sites, TSS1 and TSS2, were identified at 221bp and 62bp upstream of the ORF, respectively. Greene *et al.* revealed that a region located between -221 and -121 bases upstream of the ATG translation start site plays a crucial role in regulation of the *FXN* promoter (Greene *et al.*, 2005). No TATA box element was apparent in this region and the Initiator (Inr)-like elements associated with the TSS1 did not affect the *FXN* gene expression (Kumari *et al.*, 2011). Two transcription factors, serum response factor (SRF) and transcription factor activator protein 2 (TFAP2), were found to bind directly to this region and mutagenesis of either SRF or TFAPA2 binding sites significantly decreased *FXN* promoter activity (Li *et al.*, 2010). The *FXN* gene is expressed in all cells but it has variable tissue-specific levels of expression. Generally, in tissues with high metabolic rate, the frataxin expression is high. *FXN* is

significantly expressed in heart and spinal cord, intermediately in liver, skeletal muscle, and pancreas and minimally in other tissues (Koutnikova *et al.*, 1997). In the central nervous system (CNS), the highest frataxin expression has been detected in dorsal root ganglia (DRG), much lower in cerebellum and lowest in cerebellar cortex (Delatycki *et al.*, 2000). The frataxin mRNA level in FRDA patients is varied in an inverse GAA length-dependent manner. Depending on the GAA repeat size, FRDA patients can have 13%-30% of the normal frataxin mRNA levels (Pianese *et al.*, 2004; Bidichandani *et al.*, 1998).

1.3 Frataxin protein

Frataxin (FXN) is a nuclear-encoded 210 amino acid protein which is produced in the cytosol as a precursor form. It contains N-terminus mitochondrial targeting peptides that direct the precursor to mitochondria matrix. Data obtained from transfection of mammalian cells with human *FXN* in GFP-expression localised frataxin to the mitochondria (Babcock *et al.*, 1997). Following import of the precursor into the mitochondria, mitochondrial processing peptidase (MPP) proteolytically removes the N-terminal sequence in two sequential steps and generates the 14.5 kDa functional mature frataxin encoded by amino acids 81-210 (m_{81}-FXN) (Campuzano *et al.*, 1996; Gibson *et al.*, 1996; Cavadini *et al.*, 2000; Schmucker *et al.*, 2008). Recently it has been shown that before the mitochondrial import, RNF125 as the E3 ligase enzyme ubiquitinates a significant amount of frataxin precursor and targets them to ubiquitin/proteasome degradation. Preventing frataxin degradation is considered as a potential therapeutic approach for FRDA (Rufini *et al.*, 2011; Benini *et al.*, 2017). Solution and crystal structure of mature frataxin revealed that frataxin is a compact, globular protein containing two α-helices, a middle β sheet region and a C terminal coil. Frataxin is characterized by a unique fold which allows two terminal α-helices to

frame a platform composed of five antiparallel β strands. C terminus of β5 with β6 and β7 form a smaller β sheet which interacts with the planes and confers an overall compact α-β sandwich structure (Dhe-Paganon *et al.*, 2000; Musco *et al.*, 2000; Bencze *et al.*, 2006). Alignment of human frataxin and CyaY sequences revealed that FXN C-terminal region is evolutionary highly conserved portion of the protein, which all missense mutation found so far in individuals with heterozygous FRDA affect this domain. This observation suggests that C-terminus is essential for folding and function of FXN. Mutation affecting N-terminal region of FXN causes an atypical and milder FRDA phenotype in which patients show slower disease progression (Figure 1.1) (Cossée *et al.*, 1999; De Michele *et al.*, 2000; Musco *et al.*, 2000).

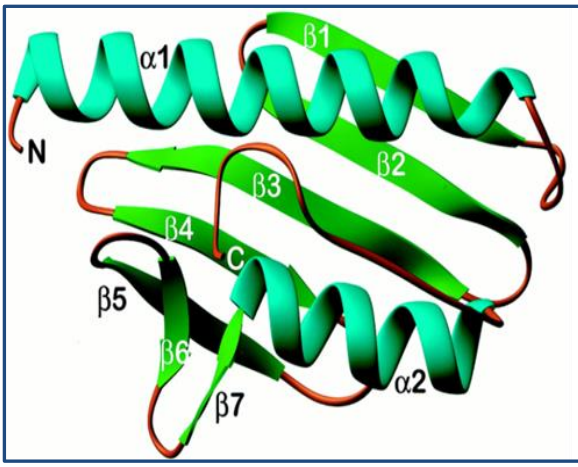

Figure 1.1: Structure of frataxin. Crystal structure of frataxin shows a compact α-β sandwich in which two α-helices pack against a 7-residue β-sheet. Strands β1–β5 form a flat antiparallel β sheet that interacts with two terminal helices. The C terminus of β5 and strands β6 and β7 form a second smaller β-sheet (Musco *et al.*, 2000).

1.3.1 Frataxin and iron homeostasis

Presence of frataxin in a wide range of species from bacteria to humans demonstrates that it is an evolutionary conserved protein. Although its exact molecular function is unclear at the present, early embryonic death in nullizygous *Fxn* mice suggests that frataxin is essential for emberyonic development (Cossée *et al.*, 2000). It is generally accepted that FXN is involved in iron homeostasis and assembly of iron-sulphur (Fe-S) clusters (Gibson *et al.*, 1996; Adinolfi *et al.*, 2009; Pastore and Puccio, 2013). Babcock *et al.* (1997) provided the first insight into the

role of FXN. They reported that deletion of Yfh1 (yeast *FXN* homologue gene) results in accumulation of mitochondrial iron and hypersensitivity to oxidative stress (Babcock *et al.*, 1997). Further studies revealed that reintroduction of Yfh1 leads to rapid export of accumulated iron from the mitochondria into the cytosol. This observation indicates a regulatory role for frataxin homologue Yfh1p in maintaining mitochondrial iron homeostasis (Radisky *et al.*, 1999). There are several lines of evidence indicating that FXN is involved in the assembly of iron-sulphur (Fe-S) clusters. It has been reported that FXN deficiency results in reduced activity of Fe-S proteins such as the Krebs cycle enzyme aconitase. Deficient activities of complexes I–III of the respiratory chain and of the aconitases (Puccio *et al.*, 2001) in FRDA mouse models and Δyfh1 yeast and FRDA patients (Rotig *et al.*, 1997) have been detected (Pandolfo and Hausmann, 2013). Puccio *et al.* (2001) generated two different tissue-specific knock-out mouse models and they reported a 77%-87% complex II deficiency and a 70-74% aconitase deficiency without noticeable mitochondrial iron accumulation in both models (Puccio *et al.*, 2001). Interestingly, it has been reported that Fxn deficient mice show 50% of Fe-S enzyme activity at 4 weeks age (Seznec *et al.*, 2004). These results propose an important but not essential role for FXN in Fe-S cluster biosynthesis.

1.3.2 Oxidative stress

Due to impaired mitochondrial iron efflux in FRDA, iron accumulates in mitochondria. This excess iron, which is primarily ferrous Fe [II], leads to increased Fenton-mediated toxic free radical (hydroxyl radicals $OH^{\bullet}$) production and consequently increased oxidative stress. $OH^{\bullet}$ is a highly toxic and can damage any biological macromolecule such as mitochondrial DNA and consequently result in cell death. Several studies were consistent with the idea that iron accumulation leads to

oxidative stress in FRDA. Measuring the concentration of 8 hydroxy 2' deoxyguanosine (8OH2'dG; a marker of oxidative DNA damage) in FRDA patients (Schulz *et al.*, 2000), serum malondialdehyde (MDA; a marker of lipid peroxidation) in FRDA mouse model (Al-Mahdawi *et al.*, 2006) and FRDA patients (Bradley *et al.*, 2004) showed an increase in the level of these oxidative parameters. However, results from other studies questioned the role of oxidative stress in FRDA (Armstrong *et al.*, 2010). Data from conditional knock-out mouse models revealed that iron accumulates in mitochondria in a time-dependent manner and it occurs after inactivation of Fe-S dependent enzymes (Puccio *et al.*, 2001). Moreover, it has been demonstrated that complete Fxn deficiency does not induce oxidative stress in neuronal tissues of an FRDA mouse model (Seznec *et al.*, 2005). Conflicting results obtained from different studies point out the need of further research on contribution of iron accumulation and oxidative damage in FRDA.

1.4 GAA repeat expansion in FRDA

Campuzano *et al.* (1997) showed that compared to normal individuals, the intron 1 in FRDA patients is larger and it contains pure GAA repeat expansion. 96% of FRDA chromosome abnormality is due to the homozygous expansion of GAA repeat within the first intron of the *FXN* gene. The remaining cases are heterozygous for GAA repeat expansion on one allele and a point mutation on the other (Campuzano *et al.*, 1996). A normal sequence carries ~6-36 triplet repeats, whereas affected individuals have approximately 70 up to 1700 triplets, most commonly 600–900 GAA triplets on both alleles of the *FXN* gene (Pandolfo, 2009; Cossee *et al.*, 1997; Bidichandani *et al.*, 1998). Investigating the GAA repeat size in normal chromosomes revealed length polymorphism in normal alleles. 83% of normal alleles carry 6-10 GAA units, which are named small normal (SN) and 17% of them contain 12-36 triplets and form

large normal group (LN) (Montermini *et al.*, 1997; Pandolfo, 1998; Monticelli *et al.*, 2004). Like all simple-sequence repeats, the basic polymorphism-generating mechanism can be accounted for size heterogeneity in normal alleles. Based on this mechanism, during DNA replication, occasional 'stuttering' of DNA polymerase can result in size heterogeneity in normal alleles (Montermini *et al.*, 1997). Due to the slippage during replication of repetitive sequences, the new strand mispairs with the template strand, creating an unpaired repeat loop. After a second round of replication this 'slip out' can give a rise to expanded repeat (Mirkin, 2006). Probably a rare or singular event caused the jump from a small normal to large normal group. Several studies have demonstrated that the long normal alleles are interrupted by hexa-nucleotide repeats (GAGGAA), which prevent these large normal alleles from further expanding and forming pathological expanded alleles. It has been directly observed that uninterrupted long normal alleles, named as permutations, can undergo large expansion into an allele containing >650 triplets (Montermini *et al.*, 1997; Patel and Isaya, 2001; Holloway *et al.*, 2011).

1.4.1 Instability of GAA expanded repeats

The pathological expanded allele size is highly variable and it shows both somatic (mitotic) and intergenerational instability. That means that the GAA repeat size continues to differ within the tissues and across the generations (Pianese *et al.*, 1997). Several studies have demonstrated that parental gender plays an important role in GAA repeat size variation. It has been observed that in maternal transmission both contraction and expansion of the GAA allele can be detected, whereas following paternal transmission there is a tendency for the GAA repeat to contract (Pianese *et al.*, 1997; De Michele *et al.*, 1998). In addition, as the parents age, the contractions in the paternal transmission and the expansions in the maternal transmission are

more evident (Kaytor *et al.*, 1997; De Michele *et al.*, 1998). It has been observed that sibs with same parental mutant chromosome show different repeat lengths. It has been documented that somatic mosaicism in FRDA patients is length-dependent, tissue-specific and age-dependent (Clark *et al.*, 2007). Mitotic GAA instability shows a strong tendency to contract. The expanded GAA alleles in peripheral leukocytes revert to normal size, whereas in the DRG, a predilection for further expansion has been detected (Sharma *et al.*, 2002). In addition, alleles containing less than 25 triplets are completely stable and when the number of triplets exceeds 44, somatic instability can be detected (Sharma *et al.*, 2002; Clark *et al.*, 2007). Comparing somatic instability in a FRDA 18-week foetus and a 24-year-old FRDA patient, both homozygous with the same E allele sizes, revealed that foetal tissue contains significantly lower levels of somatic instability. Furthermore, an experiment analysing data from 2 to 49-year-old FRDA patients and carriers has shown that mutation loads increase from 17% to 78.7%, respectively in an age dependent manner. Thus it can be concluded that somatic instability occurs mainly after early embryonic development and it develops progressively through the life (De Biase *et al.*, 2007). Several models explain the mechanisms of repeat instability and all of them support a role for replication in repeat instability (Pearson *et al.*, 2005). Forming secondary structure during replication can provide an explanation for repeat contraction and expansion. During replication of triplet repeat, the single-stranded part of lagging strand forms hairpin. Replication across hairpins results in contraction and slippage of nascent lagging strand leads to expansion (Figure 1.2) (Heidenfelder *et al.*, 2003; Mirkin, 2006). Recently, it was demonstrated that expanded GAA repeats at the *FXN* locus in FRDA obstructed the progression of replication machinery and allowed the replication forks to proceed predominantly in a 3′-5′ direction, which could promoted

repeat expansion (Gerhardt *et al.*, 2016). In addition, recent studies have shown that changes in epigenetic status surrounding repeat expansions also contribute to repeat instability (Libby *et al.*, 2008; Dion and Wilson, 2009). A comparative analysis of the epigenetic context of disease-associated trinucleotide repeats (TNRs) and random set of repeats revealed that the expansion-prone repeats differ from background TNRs in their epigenetic profiles. This data indicates that certain epigenetic status is associated with repeat expansion (Essebier *et al.*, 2016).

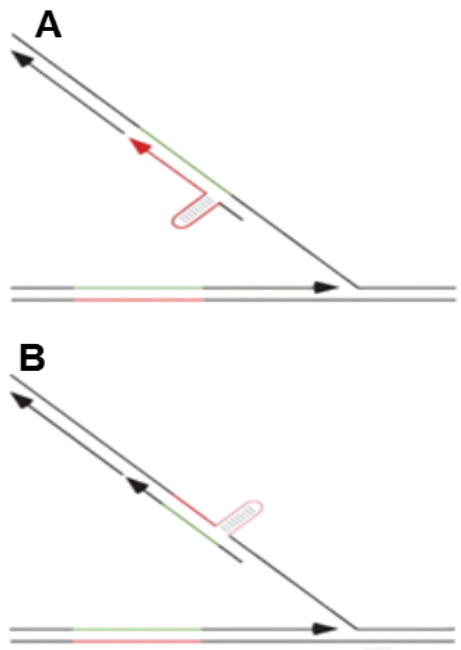

Figure 1.2: Repeat instability during replication. (A) Formation of a hairpin in the newly synthesised lagging strand causes expansion. **(B)** Formation of a hairpin in the single stranded part of lagging strand template. Skipping replication of hairpin structure leads to contraction of the triplet repeat (Mirkin, 2006).

1.5 Abnormal structures

1.5.1 Abnormal triplex structure

The exact mechanism through which the GAA repeat expansion leads to *FXN* gene silencing is not fully understood yet. Several mechanisms have been proposed to explain how the triplet expansion inhibits frataxin expression. It has been suggested that GAA repeat expansion allows the formation of non-B-DNA structures, such as triplex-based sticky DNA, R-loops, and heterochromatin formation at pathogenic *FXN* alleles, which can lead to frataxin deficiency. Bidichandani *et al.* (1998)

demonstrated that GAA expansion is associated with an unusual DNA structure and transcription deficiency in a length-dependent manner (Bidichandani *et al.*, 1998; Rajeswari, 2012). The GAA.TTC repeat is a homopurine (R) – homopyrimidine (Y) mirror repeat, which can adapt non-B-DNA conformation, including triplexes (H-DNA), sticky DNA or hairpin structures. The GAA repeat can form either R·R·Y or Y·R·Y intramolecular triplexes by fold- back mechanism. In both motifs, the purine has the central position and forms normal Watson-Crick pairs with the complementary pyrimidine strand (Figure 1.3). Binding of the third strand, containing either purine (R·R·Y) or pyrimidine (Y·R·Y), is through 'Hoogsteen' or 'reverse-Hoogsteen type' of hydrogen bonding (Sakamoto *et al.*, 1999; Grabczyk and Usdin, 2000c; Sakamoto *et al.*, 2001c; Mirkin, 2006).

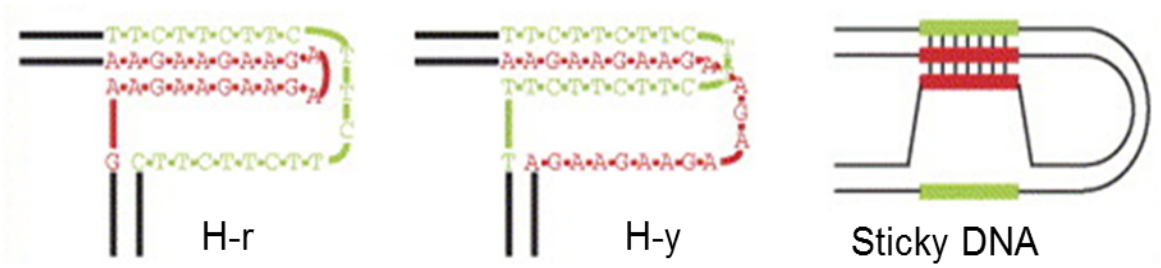

Figure 1.3: Unusual DNA structures formed by triplet repeats. Purines are red and pyrimidines are shown in green. H-r is a triplex composed of purine and pyrimidine strands which are bound together with classical Watson–Crick base pairing. The third strand contains purine stretch and runs antiparallel to the central polypurine strand. H-y the third strand comprises of polypyrimidines, which runs parallel to the central purine strand. Sticky DNA formed by binding of two long GAA repeats to each other in an intra-molecular configuration (Mirkin, 2006; Rajeswari, 2012).

1.5.2 Sticky DNA

Sticky DNA is one of the suggested unusual DNA conformations that is proposed for the GAA.TTC repeats. Sakamoto *et al.* (1999) reported that in a negatively supercoiled plasmid at natural pH, association of two long GAA.TTC repeats in the R·R·Y configurations leads to the formation of sticky DNA. They proposed a strand exchange model in which two R·R·Y triplexes exchange their Y strands to form an extremely stable hybrid triplex (Figure 1.3) (Sakamoto *et al.*, 1999). It has been reported that sticky DNA structure directly binds to RNA polymerase, sequesters it and thus, strongly impairs *in vitro* transcription in both a *cis* and *trans* manner (Sakamoto *et al.*, 2001c). Sakamoto *et al.* (2001) showed that introducing interruptions into GAA repeats abolishes the formation of sticky DNA and therefore reduces its negative effect on GAA transcription (Sakamoto *et al.*, 2001a).

1.5.3 R-loops

Several studies have provided evidence that the unusual DNA structure within the GAA repeat impedes transcription of the *FXN* mRNA. It has been proposed that triplex formation may prevent binding of transacting factors (TF) to transcription-control sequences and abolish transcriptional enhancement (Kohwi and Kohwi-Shigematsu, 1991), or alternatively it may trap the RNA polymerase and cause a pause site (Sakamoto *et al.*, 2001c). Grabczyk and Usdin proposed a model to explain the role of DNA triplex in impeding the GAA repeat transcription. Their model suggests that transient formation of DNA triplexes behind of an advancing RNA polymerase within a long GAA.TCC tract leads to R-loop (RNA.DNA hybrid) formation. R-loop formation negatively affects *FXN* transcription elongation and consequently results in reduced frataxin mRNA and protein levels (Figure 1.4) (Grabczyk and Usdin, 2000a; Grabczyk and Usdin, 2000c). It has been

demonstrated that RNA.DNA hybrid formation is a fundamental feature of GAA repeat transcription (Grabczyk *et al.*, 2007). Groh *et al.* (2014) observed R-loop formation over the expanded GAA region. Its level was in correlation with the GAA expansion length and it was co-localised with the H3K9me2 chromatin marker (Groh *et al.*, 2014). Several studies have demonstrated that elimination of RNA.DNA hybrids by RNase-H, which specifically disrupts RNA.DNA hybrids, significantly improves mRNA elongation (Huertas and Aguilera, 2003; Groh *et al.*, 2014). Moreover, observing this hybrid formation even in the 'pre-mutation' size alleles, suggests that RNA.DNA hybrid formation contributes to GAA repeat instability (Grabczyk *et al.*, 2007).

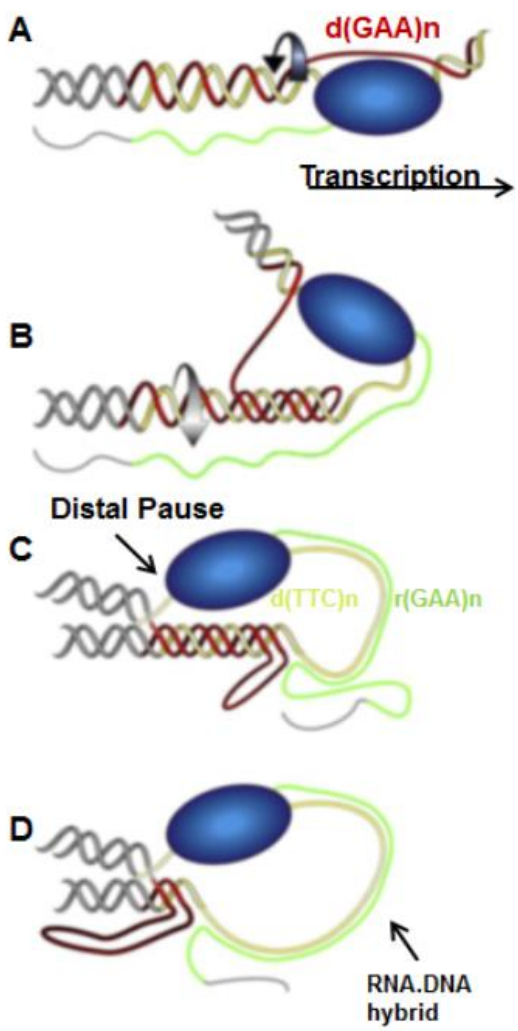

Figure 1.4: An intramolecular triplex impedes transcription and gives rise to RNA.DNA hybrid. The purine (GAA or R) strand of the repeat is red, the pyrimidine (TTC or Y) strand is yellow, the flanking DNA is gray and the nascent transcript is green.

A) Polymerase advancing within a long GAA.TCC repeat unzips the DNA helix and the DNA behind it becomes negatively supercoiled. Negative supercoiling promotes triplex formation. RNA polymerase covers the Y (TCC) template strand and the non-template (GAA) strand is available to fold back and become the third strand.

B) As the helix rotates to form the triplex, the negative supercoils relax and consequently the triplex formation spreads over the region. Following the involvement of the non-template GAA strand in the formation of the triplex, part of the template strand (yellow) becomes temporarily single stranded.

C) Since the RNA polymerase has trouble to deal with the triplex/duplex junction, it pauses in the distal end of the GAA.TCC repeats. Paused RNA polymerase generates truncated transcript at the 3′ end of the structure. The nascent transcript (green) and template strand (yellow) bind together and form a RNA.DNA hybrid.

D) A stable RNA.DNA hybrid formed (Grabczyk *et al.*, 2007).

1.6 Epigenetic changes in FRDA

The word “epigenetic” literally means “on top of or in addition to genetics”. Based on new scientific findings, epigenetics is defined as ‘‘the study of changes in gene function that are mitotically and/or meiotically heritable and that do not involve a change in the DNA sequence.’’ (Dupont *et al.*, 2009; Kanherkar *et al.*, 2014). Major epigenetic mechanisms, include DNA methylation, histone modification and ncRNA-mediated gene silencing. Disruption of these epigenetic mechanisms can alter gene expression by chromatin remodeling. Accumulative evidence suggests that mutation and non-B-DNA structure could account for epigenetic changes and heterochromatin formation, thereby impeding gene transcription (Chen *et al.*, 1995; Al-Mahdawi *et al.*, 2008). Several epigenetic changes and abnormal heterochromatinisation have been identified in the immediate vicinity of the expanded GAA repeats of the *FXN* gene (Sandi *et al.*, 2014). The first evidence for epigenetic involvement in FRDA came from a study by Saveliev and colleagues. In eukaryotes, DNA wraps around the histone proteins to form a nucleoprotein complex called "chromatin". Heterochromatin is a highly condensed and transcriptionally inaccessible type of chromatin, which contains large blocks of repetitive DNA sequences. Heterochromatin is invasive and has a tendency to affect the expression of nearby genes (Yandim *et al.*, 2013). Saveliev *et al.* showed that in transgenic mice, juxtapositioning of GAA expanded repeats with heterochromatin-sensitive human CD2 (hCD2) transgene alters expression by position effect variegation (PEV) (Saveliev *et al.*, 2003). They showed that irrespective of the chromosomal location of the hCD2 transgene, inclusion of the GAA repeats can supress hCD2 expression. In addition, they also reported that the overexpression of a powerful modifier of PEV, heterochromatin protein (HP1), in CD2 GAA transgenic mouse lines significantly

decreases CD2 expression, whereas in the absence of GAA, HP1 overexpression did not affect hCD2 expression (Saveliev *et al.*, 2003). PEV occurs when a gene that was previously located in a euchromatic region, is positioned proximal to heterochromatic region. The relocated gene exhibits varied expression because of the change in its position (Elgin and Reuter, 2013). PEV is a hallmark of heterochromatin mediated gene silencing. Recent studies strongly suggest that epigenetic changes including DNA methylation, histone modifications and repressive chromatin formation are involved in the *FXN* gene silencing in FRDA (Elgin and Grewal, 2003; Herman *et al.*, 2006).

1.6.1 DNA methylation

DNA methylation is a silencing epigenetic mark that involves the addition of a methyl group to the C5 position of cytosine in the context of CpG dinucleotides. Gene silencing is the predominant consequence of DNA methylation. Therefore, aberrant DNA methylation is a feature of a number of important human diseases, such as cancer in which DNA methylation in the promoter region of key tumour suppressor genes promotes oncogenic progression (Pook, 2012). Recently, a large amount of evidence has emerged to show that DNA methylation has an important role in repeat expansion diseases. Disease-associated DNA methylation has also been identified in other trinucleotide repeat expansion diseases, such as myotonic dystrophy type 1 (DM1) (López Castel *et al.*, 2010), fragile X syndrome (FRAXA) and FRDA. The expansion of CGG repeats to over 200 triplets within the 5′UTR of the *FMR1* gene leads to hypermethylation of the CpG sites along the *FMR1* promoter region, transcriptional silencing and development of fragile X syndrome (Chandler *et al.*, 2003; Naumann *et al.*, 2009). Pathological expansion of of a G_4C_2 repeat in *C9orf72* is the most common genetic cause of frontotemporal lobar degeneration (FTLD) and

amyotrophic lateral sclerosis (ALS). A recent study showed that the increase of C*9orf72* repeat length correlates with earlier disease onset, increased methylation states of 5′ CpG islands and reduced promoter activity (Gijselinck *et al.*, 2016). In FRDA patients, similar patterns have been reported. The *FXN* gene has three putative CpG islands (Figure 1.5) (http://cpgislands.usc.edu/) and several studies have investigated the role of DNA methylation in FRDA. Studying DNA methylation status at the *FXN* locus has revealed an overall increase in the methylation of CpG island upstream of GAA repeat in FRDA derived lymphoblastoid cells (Greene *et al.*, 2007). Al-Mahdawi and colleagues also reported an increase in DNA methylation upstream of GAA repeats (UP) in two lines of FRDA YAC transgenic mice (YG8 and YG22) and clinically important FRDA tissue such as heart, brain and cerebellum (Al-Mahdawi *et al.*, 2008). A similar pattern was demonstrated in two large scale studies confirming the previous observation about hypermethylation in the UP region of the *FXN* gene in FRDA individuals. These studies reported a positive correlation between triplet expansion size, disease severity and the level of upstream methylation (Evans-Galea *et al.*, 2012) and a strong negative correlation between methylation level and the age of disease onset (Castaldo *et al.*, 2008; Sandi *et al.*, 2013b; Yandim *et al.*, 2013). Lorcinz and colleagues demonstrated that intragenic DNA methylation may also reduce transcription elongation efficiency by inducing heterochromatin formation, suggesting that intronic DNA methylation at the upstream region of the GAA repeat is involved in *FXN* gene silencing (Lorincz *et al.*, 2004).

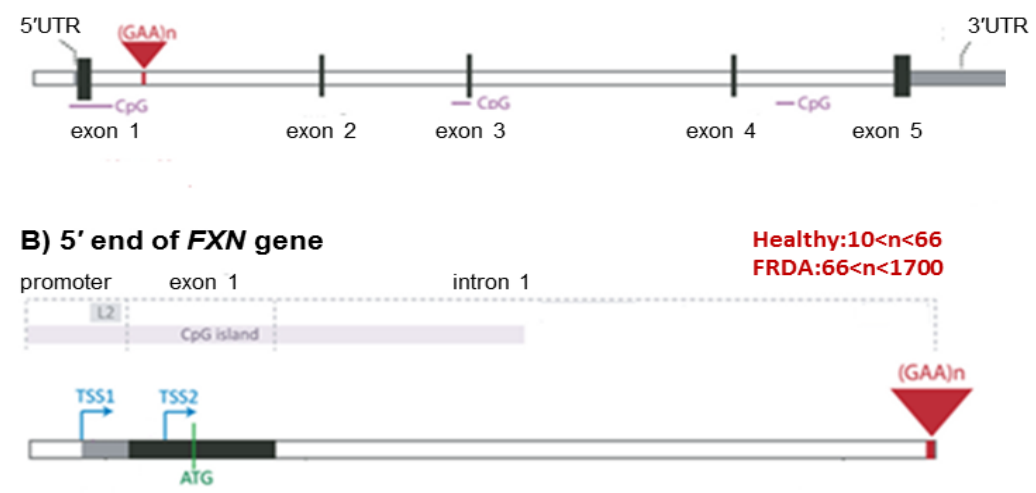

Figure 1.5: The human Frataxin (*FXN*) gene and its regulatory components. **(A)** Full body of the *FXN*. *FXN* gene consists of five exons and four introns. It has three putative CpG islands. *FXN*'s first intron contains a number of (GAA)n triplet repeats. Normal alleles contain 10–66 triplets while disease-causing expanded *FXN* alleles have higher numbers. **(B)** Regulatory elements at the 5′ end of *FXN*. The first CpG island is located upstream of the GAA repeats and covers promoter, first exon and the beginning of first intron (Yandim *et al.*, 2013).

1.6.2 Histone modifications

Eukaryotic DNA is tightly packed into nucleosomes and a nucleosome is composed of DNA coiled around an octamer, which is made up of two copies of each histone protein- H2A, H2B, H3, and H4. There are different types of histone modifications, which can affect gene expression by determining accessibility of the DNA for transcription. The simplest and most extensively studied classes of histone modifications are the acetylation and methylation of lysine (K) residues. Acetylation of lysine is a hallmark of transcriptionally active chromatin. This is probably due to the presence of acetyl groups (CH_3CO-) which can weaken the binding between the histones and negatively-charged DNA backbone and allow the chromatin structure to

open up (Allfrey *et al.*, 1964; Halleck and Gurley, 1981; Yandim *et al.*, 2013). Histone acetylation also results in the recruitment of chromatin remodeling factors that lead to more permissive and transcriptionally competent chromatin structure. Such factors are recruited through their acetyl-lysine recognition domains, known as bromodomains (BRD) (Swygert and Peterson, 2014; Ferri *et al.*, 2016). Hyperacetylation of histone H3 and H4 are known as active marks, whereas methylation of H3K9 and H3K27 are known to be a hallmark of heterochromatin formation and subsequent gene silencing (Karmodiya *et al.*, 2012; Wang *et al.*, 2008). It is commonly recognized that H3K9 and H3K27 methylation are important in heterochromatin formation. It is well established that a conserved histone methyltransferase (HMTase), SUV39H1/2, selectively methylates lysine 9 of histone H3 (Rea *et al.*, 2000) and subsequently generates a binding site for HP1 proteins, a family of heterochromatic adaptor molecules (Lachner *et al.*, 2001). In turn, HP1 recruits more HMTases and creates a positive feedback loop, resulting in heterochromatin formation (Figure 1.6). On the other hand, H3K27 methylation is catalysed by polycomb repressive complex 2 (PRC2). Previous studies suggested that both markers recruit distinct protein machineries. However, recent studies showed that these marks frequently colocalise in the genome and they cooperate in gene silencing maintenance (Boros *et al.*, 2014; Saksouk *et al.*, 2015; van Kruijsbergen *et al.*, 2015). Hypoacetylation of histone H3 and H4, with increased levels of H3K9 di- and trimethylation in the upstream of GAA repeat, were first reported histone modifications within the *FXN* locus (Herman *et al.*, 2006). Investigating FRDA patient brain, cerebellum and heart tissues and YAC transgenic mice revealed significant levels of H3 and H4 histone deacetylation and increased H3K9 di- and tri-methylation at the promoter, downstream and upstream of the GAA

repeats (Al-Mahdawi *et al.*, 2008). Furthermore, repressive marks such as H3K9me3 and H3K27me3 and HP1, were highly enriched in 5′UTR of the *FXN* gene (De Biase *et al.*, 2009). Kim *et al. (*2011) reported significantly decreased levels of H3K4me3 at a region immediately upstream of the GAA repeats. Numerous studies suggest that H3K4me3 modification is often positively correlated with transcription initiation and more accessible chromatin structure for transcription factors. In addition, it has been shown that H3K4me3 facilitates recruitment of transcription post-initiation factors and enhances the efficiency of transcript elongation. This finding indicates a defect in *FXN* transcription elongation rather than a defective transcriptional initiation in FRDA (Sims *et al.*, 2007; Kim *et al.*, 2011). H3K36me3 and H3K79me2 are known as a hallmark of transcription elongation. Decreased methylation of H3K79 and H3K36 in FRDA cell lines supports the idea that frataxin deficiency is a result of defective *FXN* transcription elongation (Kim *et al.*, 2011; Sandi *et al.*, 2013b; Sandi *et al.*, 2014). However recent studies have demonstrated that repressive chromatin extends from the GAA repeat towards the *FXN* promoter and encompasses the *FXN-TSS* and renders it transcriptionally non-permissive, and thereby causes a severe deficiency of transcriptional initiation (Chutake *et al.*, 2014a). Indeed, the transcriptional deficiency upstream of the expanded GAA-TR significantly correlates with repeat length (Chutake *et al.*, 2014c). Taken together, recent studies demonstrate that the expanded GAA repeats impair expression predominantly at the initiation stage of transcription, but also block RNAPII elongation.

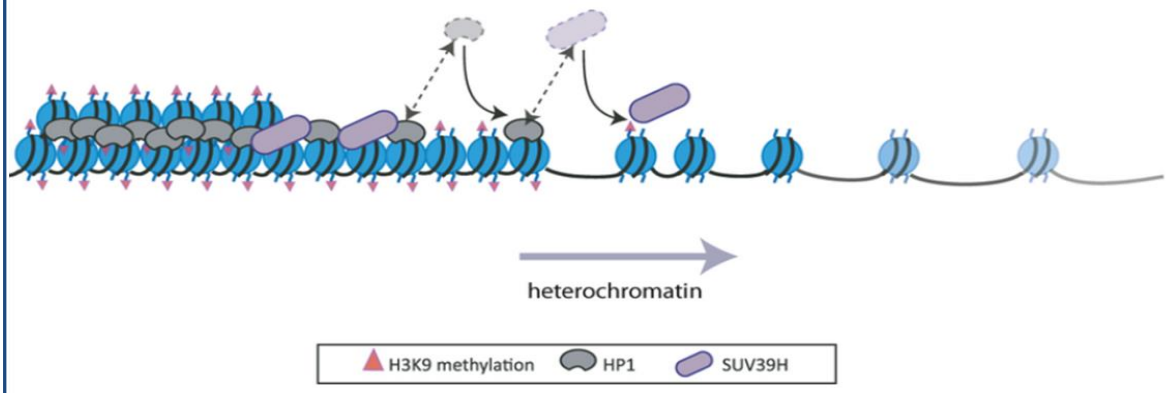

Figure 1.6: Spreading of heterochromatin via HP1 and SUV39H. There is an increased level of H3K9me3 and HP1 , at the 5′UTR of *FXN* gene. Here, the histone methyltransferase Drosophila suppressor of variegation 3-9 (human homologue) (SUV39H) catalyses H3K9 trimethylation, which is subsequently recognised by HP1. While HP1 molecules self-dimerise and create a more compact chromatin, they also recruit more SUV39H to the locus to exacerbate heterochromatin's spreading (Yandim *et al.,* 2013).

1.6.3 CTCF

CCCTC-binding factor (CTCF) is an 82-kDa evolutionarily conserved and ubiquitously expressed 11-zinc-finger DNA-binding protein. 100% evolutionary conservation of the entire CTCF 11-Zn-finger region indicates that these domains are involved in a highly conserved biological function (Filippova *et al.*, 1996). CTCF is a multifunction protein, initially characterised as a transcriptional repressor of the chicken *c-myc* gene (Lobanenkov *et al.*, 1990; Klenova *et al.*, 1993; Kim *et al.*, 2015). Later, CTCF was found to also function as a chromatin insulator. Indeed, it is the only insulator protein known in vertebrates. Insulators are DNA-protein complex that protect genes from inappropriate signals emanating from their surrounding environments (West *et al.*, 2002). Insulators can function as enhancer-blockers and block enhancer-promoter interactions or act as barriers against spread of heterochromatin into a neighbouring domain (Yang and Corces, 2011). There are

several recent studies proposing that insulators may mediate three-dimensional looping of genomic regions with the primary goal of organising the eukaryotic genome into epigenetically heritable states (Yang and Corces, 2011; Herold *et al.*, 2012). The insulator activity of CTCF was first identified at the 5′ end of the chicken ß-globin locus. Bell *et al.* (1999) showed that CTCF was able to interfere with enhancer-promoter communication in a directional manner and separate the locus from neighbouring heterochromatin (Bell *et al.*, 1999). The insulator function of CTCF was also reported at the imprinted *Igf2* (insulin-like growth factor 2)–*H19* locus. It was shown that CTCF can bind to the unmethylated imprinting control region (ICR), and block the access of *Igf2* to the distal enhancer (Figure 1.7) (Bartolomei *et al.*, 1991; Bell and Felsenfeld, 2000).

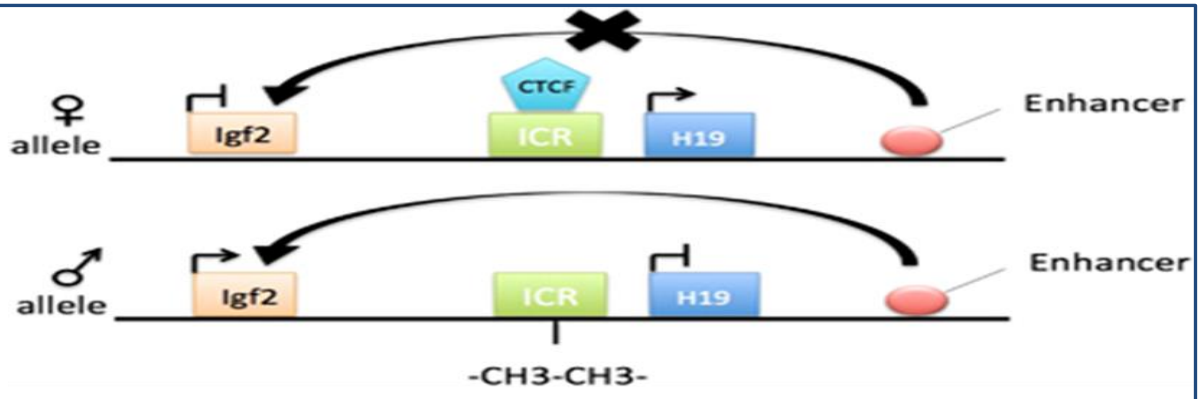

Figure 1.7: The neighbouring *Igf2* and *H19* genes are reciprocally imprinted. *Igf2* is paternally expressed and *H19* is maternally expressed. These two genes are separated from each other by a differentially methylated region (DMR) ICR. The ICR contains multiple CTCF-binding sites that function as an insulator. CTCF fails to bind to the methylated paternal ICR and thus the enhancer is able to induce *Igf2* gene expression. On the other hand, CTCF binding to unmethylated maternal ICR, blocks the access of *Igf2* to the enhancer sheared with *H19,* resulting in *H19* gene activation and *Igf2* inactivation (Kim *et al.,* 2015).

Some studies have suggested that CTCF may act as a barrier at heterochromatin boundaries and protect genes against PEV. Genome-wide studies of the localisation of CTCF in three cell types revealed that CTCF-binding sites are significantly enriched at the boundaries between the repressive H3K27me3 and active H2AK5ac domains, indicating that CTCF may be involved in the chromatin barrier function (Cuddapah *et al.*, 2009). A recent study by Weth *et al.* (2014) has demonstrated that upon CTCF depletion, H3K27me3 spreads into the nearby active domain and thereby causes gene repression. They also showed that CTCF removes H3K27me3, a hallmark of repressed domain, and actively converts repressed chromatin into active chromatin (Weth *et al.*, 2014).

CTCF plays a role in several TNR diseases. Results from genome-wide mapping studies and studying TNR diseases such as myotonic dystrophy (DM1), spinocerebellar ataxia type 2 (SCA2) and type 7 (SCA7), and Huntington's disease (HD), show that CTCF binding sites are present at one or both sides of repeat expansion regions (Filippova *et al.*, 2001; Libby *et al.*, 2008). The role of CTCF in establishing the local chromatin structure at repeats was first shown at the *DM1* locus, where two CTCF binding sites flank *DM1* CTG repeats and restrict repressive chromatin structure to repeat region. Loss of CTCF binding in congenital DM1 was associated with the spread of heterochromatin into nearby regions (Cho *et al.*, 2005). In FRDA, a single CTCF binding site has been identified in the 5′ untranslated region (5′UTR) region of the *FXN* gene. De Biase *et al*. (2009) reported that the GAA repeat expansion in FRDA is associated with severe depletion of CTCF in the 5′UTR of the *FXN* gene. However, since other genes were not found to be affected, it appears that CTCF depletion is not a generalised defect in FRDA (De Biase *et al.*, 2009). Studying FRDA cerebellum also demonstrated that CTCF occupancy decreased to

the value of only 20% compared to unaffected cerebellum controls (Al-Mahdawi *et al.*, 2013). It has also been reported that CTCF depletion in FRDA is associated with higher levels of an antisense transcript and heterochromatin formation at the *FXN* locus (De Biase *et al.*, 2009).

1.6.4 The role of antisense transcription in FRDA

Antisense transcription refers to the transcription from the opposite strand of a protein coding strand (sense strand) and the production of a single stranded RNA, known as an antisense transcript. This antisense transcript may have a partial or a complete overlap with the sense transcript. Antisense transcripts are widespread in eukaryotic genomes. It has been estimated that at least 20% of human genes have an antisense partner (Chen *et al.*, 2004). Similar to sense transcription, antisense transcription is driven by a promoter that can be an independent unidirectional promoter or a divergent promoter (bidirectional promoter) (Pelechano and Steinmetz, 2013). Various models have been suggested to explain how antisense transcripts are involved in regulation of gene expression. In many organisms, antisense transcripts have emerged as key regulators of gene expression in an epigenetic manner. A classic example is mammalian X chromosome inactivation, in which *XIST* silences one of the two X chromosomes in females. The action of *XIST* is negatively regulated in *cis* by its antisense transcript, *TSIX*. *TSIX* transcription induces histone modifications and alters chromatin conformation in the *XIST* promoter region. It has been shown that loss of *TSIX* transcription abolishes heterochromatin formation at the *XIST* promoter and is associated with reduced CpG methylation and aberrant histone modification in the 5′ region of *Xist* gene (Ohhata *et al.*, 2008).

Antisense RNA can also affect expression of its sense partner by DNA methylation and chromatin modifications at the promoter region. Yu *et al*. (2008) reported that in

leukaemia, expression of *p15*, a well-known tumour suppressor gene (TSG), is negatively-regulated in *cis* and in *trans* by its antisense transcript, *p15AS*. *p15AS* transcription leads to dicer-independent heterochromatin formation through an increase in H3K9 dimethylation and a decrease in dimethylation of H3K4 in cells (Yu *et al.*, 2008). Antisense RNA transcribed at the *p21* locus, another TSG, represses *p21* gene expression. The *p21* antisense RNA acts as an effector molecule, which directs epigenetic regulatory complexes and H3K27me3 enrichment to the sense promoter region (Morris *et al.*, 2008; Faghihi and Wahlestedt, 2009). Important evidence of the relationship between antisense RNA and DNA methylation was provided by Tufarelli *et al*. (2003). They showed that heavily methylated CpG islands and the silencing of the haemoglobin *α1* gene (*HBA2*) in patients with a class of α-thalassemia are correlated with the transcription of an antisense transcript. An aberrant *LUC7L* transcript which is transcribed from the opposite strand to *HBA2* gene, extends into the CpG-island region of *HBA2* gene and causes hypermethylation and silencing of the *HBA2* gene (Tufarelli *et al.*, 2003; Morris, 2012; Pelechano and Steinmetz, 2013). Antisense transcripts with a potential pathogenic impact have been detected in several repeat associated diseases. At the Huntington's disease (HD) locus, a natural antisense transcript has been detected that regulates the level of *HTT* expression (Chung *et al.*, 2011). It has been proposed that an antisense transcript spanning the CGG repeat region of *FMR1*, might contribute to the pathogenesis of fragile X syndrome (FXS) and fragile X-associated tremor and ataxia syndrome (FXTAS) (Ladd *et al.*, 2007). Recent studies show that the *FMRA1* antisense transcript, *ASFMR1*, supports a non-canonical type of protein translation called repeat associated non-AUG (RAN) translation, which derives the production of toxic proteins from CGG repeats (Krans *et al.*, 2016). In addition,

studies suggest that antisense transcription in spinocerebellar ataxia type 8 (SCA8) (Moseley *et al.*, 2006), amyotrophic lateral sclerosis/frontotemporal dementia (ALS/FTD) (Zu *et al.*, 2013) and DM1 (Cho *et al.*, 2005) is a fundamental pathologic feature of these repeat-associated diseases. The expanded allele in congenital DM1 is associated with an antisense transcript emanating from the adjacent *SIX5* regulatory region, the loss of CTCF binding and the propagation of heterochromatin to the surrounding regions (Cho *et al.*, 2005; Gudde *et al.*, 2017).

Literature supporting the notion that antisense transcripts are involved in heterochromatin formation and the regulation of their partner mRNAs expression inspired De Biase *et al*. (2009) to investigate the presence of any antisense transcript in FRDA. They performed a strand-specific reverse transcription PCR with a primer located upstream of *FXN* transcription start site 3 (TSS3) and discovered an antisense transcript overlapping with the CTCF binding site (Figure 1.8). This novel antisense transcript named *FAST1* (*FXN* Antisense Transcript – 1) was produced at significantly higher levels in FRDA fibroblasts. Sequencing and BLAST analysis of the *FAST-1* transcript confirmed that it was originated from the *FXN* locus. Higher levels of *FAST-1* were associated with the severe CTCF depletion and coincidentally heterochromatin formation in the 5′ untranslated sequence (5′UTR) of the *FXN* gene (De Biase *et al.*, 2009). A comparison of CTCF occupancy at the 5′UTR of fibroblast cell lines was carried out between two normal individuals versus two FRDA patients, revealing a four-fold reduction in CTCF in the 5′UTR of the FRDA samples. In addition, they examined CTCF occupancy at 2 other loci on chromosome 9q in FRDA and normal cell lines and showed that FRDA cells do not have a generalized defect of CTCF binding. The +1 nucleosome of the *FXN* gene includes the CTCF binding site in the 5′UTR. It is followed by a long nucleosome free region, which is

composed of the three reported TSS of the *FXN* gene (Figure 1.8) (De Biase *et al.*, 2009). De Biase and colleagues observed that knocking down CTCF by siRNA in wild type fibroblasts resulted in high levels of *FAST1* and reduced *FXN* transcription. Moreover, they detected significant H3K9me3, H3K27me3 and HP1 enrichment in FRDA fibroblasts, suggesting repressive heterochromatin formation involving the +1 nucleosome and ensuing transcriptional silencing of the *FXN* gene (De Biase *et al.*, 2009).

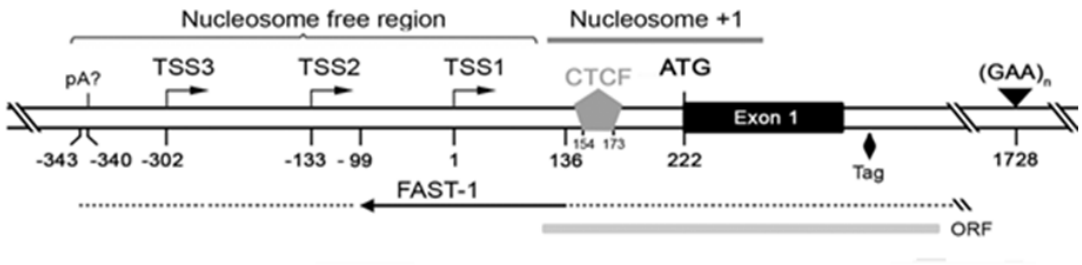

Figure 1.8: The *FAST-1*, an antisense transcript overlapping with the CTCF binding site. Schematic representation of the relevant portion of the *FXN* gene shows that the CTCF binding site in the 5′UTR is preceded by a nucleosome free region and *FAST-1* is overlapping with *FXN* gene. The solid line represents the part of *FAST-1* that has been amplified using one round of PCR, dotted line shows extended transcript that has been detected via nested PCR. The grey bar represents open reading frame (De Biase *et al.*, 2009).

In order to determine the exact size and location of *FAST-1*, characteristics of *FAST-1* were investigated in our lab and a full length *FAST-1* transcript with a total length of 523bp in size, was identified. Mapping the 3′ and 5′ ends of *FAST-1* transcript onto the genome precisely localised them to nucleotides -359 and 164 of the *FXN* gene, respectively. A poly (A) signal was also identified in the *FAST-1* sequence at nucleotide positions of -283 to -288 (Sandi and Pook, 2015). A tag corresponding to *FAST-1* was identified in a study of the human antisense transcriptome (position

70840691 on the "-" strand of chromosome 9 in the HCT116 cell line), which maps close to the start of intron 1 of the *FXN* gene (He *et al.*, 2008) and encompasses the promoter region (Figure 1.8). This potential putative *FAST-1* comprises a 462 base pair open reading frame, which encodes a putative 154 amino acids. The open reading frame is intact in humans and chimpanzees but it does not contain a Kozak consensus sequence (De Biase *et al.*, 2009). The part of *FAST-1* that has been amplified and used in this work is shown as a solid line in (Figure 1.9).

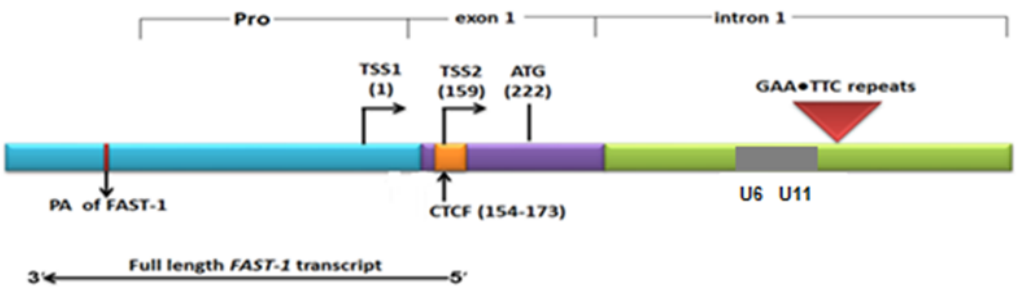

Figure 1.9: Schematic diagram of the *FXN* 5′UTR, exon 1 and intron 1 regions. It is showing the GAA repeats and differentially methylated region in FRDA. The numbers represent the CpG sites, which are numbered according to (Sandi *et al.*, 2013a). The 5′ end of *FXN* gene showing the region corresponding to the full length *FAST-1* transcript. The 5′ end of the *FAST-1* occupies the CTCF binding site in the *FXN* 5′UTR. TSS1- Transcription start site 1, TSS2- Transcription start site 2, ATG- Translation start site, PA- Polyadenylation, Pro- Promoter region (Sandi and Pook, 2015).

1.7 Mouse models

1.7.1 Knock-out and knock in mouse models

Cossée and colleagues generated an *Fxn* knock-out mouse model by deleting the most conserved and functionally important sequence of the *Fxn* gene, exon 4. Homozygous frataxin deletion resulted in the complete loss of frataxin function and early embryonic lethality (Cossée *et al.*, 2000). Early embryonic lethality indicates the

need of residual frataxin levels for survival until after birth. Puccio and colleagues then developed two viable lines of conditional knock-out mouse models, in which *Fxn* was deleted in specific tissues. They generated a muscle frataxin-deficient line and a neuron/cardiac muscle frataxin-deficient line, which together could mimic several pathophysiological and biochemical characteristics of FRDA (Puccio *et al.*, 2001). Since GAA repeat expansion is the main cause of FRDA and none of the knock-out mouse models had GAA repeat expansions, Miranda *et al.,* inserted or 'knocked in' $(GAA)_{230}$ within the first intron of the *Fxn* gene. To achieve further reduction in frataxin levels, the knock-in (KI) mice were crossed with knock-out (KO) mice and KIKO mice $Fxn^{-/230GAA}$ were generated. KIKO mice showed 25-30% frataxin expression, but these mice did not initially show defects in motor coordination, iron metabolism, response to iron binding or GAA repeat instability (Miranda *et al.*, 2002). However, neurobehavioral deficits have recently been discovered in KIKO mice (McMackin *et al.*, 2017).

1.7.2 FRDA YAC transgenic mouse models

To investigate whether human frataxin can stay functional in a mouse environment and rescue the embryonic lethal phenotype of homozygous *Fxn* knock-out mice, a human wild-type FRDA yeast artificial chromosome (YAC) transgenic mouse line was crossed twice with heterozygous *Fxn* exon 4 deletion knock-out mice. The result was a phenotypically normal mouse line with no endogenous mouse frataxin that only expressed YAC-derived human frataxin (Pook *et al.*, 2001). To study the relation of GAA repeat expansion to FRDA, Al-Mahdawi and colleagues introduced GAA repeat expansions into human YACs and generated two lines of transgenic mice, YG8 and YG22. Both of these mouse lines, showed intergenerational and age-related somatic instability of the GAA expansion (Al-Mahdawi *et al.*, 2004) reduced

levels of frataxin mRNA and protein expression and decreased aconitase activity (Al-Mahdawi *et al.*, 2006).

1.8 Therapies

Currently, there is no effective cure for FRDA, however different therapeutic approaches have been investigated to either treat the symptoms or address the cause of the disease (Figure 1.10). Although the exact molecular mechanisms underlying FRDA are not fully understood, unusual DNA structures and repressive chromatin formation as a result of GAA repeat expansion were proposed to play a role. It has been reported that the presence of complementary oligodeoxyribonucleotides (ODN) during transcription inhibits particular types of triplex formation *in vitro* and consequently this increases the yield of the full-length RNA. However, due to the high concentration required and the potential degradation, using unmodified ODNs *in vivo* is not practical (Grabczyk and Usdin, 2000a). Other molecules such as polyamides can disrupt formation of FRDA triplexes and sticky DNA by locking GAA repeat into double-stranded B-DNA structure and subsequently increase frataxin mRNA and protein levels in FRDA lymphoblast cells. However, microarray analysis of polyamide effects on the global gene expression has revealed that the polyamides significantly changed the expression of a small number of genes (Burnett *et al.*, 2006). In a recent study, Corey and colleagues developed two distinct synthetic agents, an anti-GAA duplex RNA and a single-stranded locked nucleic acid LAN. These two synthetic nucleic acids employ same strategy to restore curative FXN levels in FRDA cells. They target GAA repeats and thereby, prevent R-loops and resultant repressive chromatin formation in FRDA cells. This study demonstrated that a simple binding to the mutant transcript is sufficient to re-activate frataxin expression in FRDA cells. No changes in GAA repeat-containing genes were

reported in this study (Li *et al.*, 2016). Another novel therapeutic strategy is delivering a functional *FXN* gene to the main site of pathology. It has been reported that delivering functional *FXN* into heart cells of a mouse model with complete deletion of frataxin in cardiac and skeletal muscle, can completely prevent the development of cardiac disease and even reverse it when cardiomyopathy is already established (Perdomini *et al.*, 2014).

FRDA is caused by a partial deficiency of FXN protein. Therefore, by delivering a functional protein into the main site of pathology exogenously, the FRDA phenotype can be rescued. Payne and colleagues used a cell-penetrant peptide, TAT, to deliver human FXN to mitochondria in FRDA fibroblast cells and conditional knock-out mice. They reported reduced sensitivity to oxidant stress in FRDA fibroblast cells and increased cardiac function, aconitase activity, growth velocity and lifespan in conditional *Fxn* knock-out mice after treatment with TAT-FXN fusion protein (Vyas *et al.*, 2012). Since GAA repeat expansions induce epigenetic changes and heterochromatin formation, resulting in transcriptional silencing of the *FXN* gene, the potential reversibility of epigenetic changes has made them attractive targets for FRDA treatment. In particular, histone deacetylase inhibitors (HDACi) have been used to boost frataxin expression. It has been reported that treating KIKI mice with a pimelic *o*-aminobenzamide HDACi (compound 106) restores normal frataxin levels and increases acetylation at lysine residues in histones H3 and H4 near the GAA repeat (Rai *et al.*, 2008). Also, long term application of HDACi compound 109 in YG8R FRDA mice produced increased frataxin protein in brain and improved motor coordination and locomotor activity (Chutake *et al.*, 2016; Sandi *et al.*, 2011). A phase Ib clinical trial with HDACi compound 109/RG2833 has been successfully completed (Soragni *et al.*, 2014; Bürk, 2017). Nicotinamide (vitamin B3) is another

HDACi that has been reported to upregulate *FXN* expression in different FRDA models (Chan *et al.*, 2013). A clinical study revealed that nicotinamide can significantly upregulate frataxin protein and attenuate heterochromatin formation in FRDA patients, however, no significant clinical improvement was reported (Libri *et al.*, 2014). Furthermore, compound 5 (C5) is a small molecule with HDACi activity, which has been shown to restore FRDA histone acetylation to wild type levels and increase the *FXN* expression in FRDA patient-derived primary lymphocytes. The C5 molecule was identified by using an FRDA genomic DNA reporter cell model to screen and pre-select potential therapeutic compounds. Using this type of cell model system can considerably increase the efficiency and speed of primary compound screening (Lufino *et al.*, 2013). To justify further development of such candidate therapeutic compounds, FRDA mouse models can be used to validate the initial results obtained from *in vitro* studies (Al-Mahdawi *et al.*, 2008).

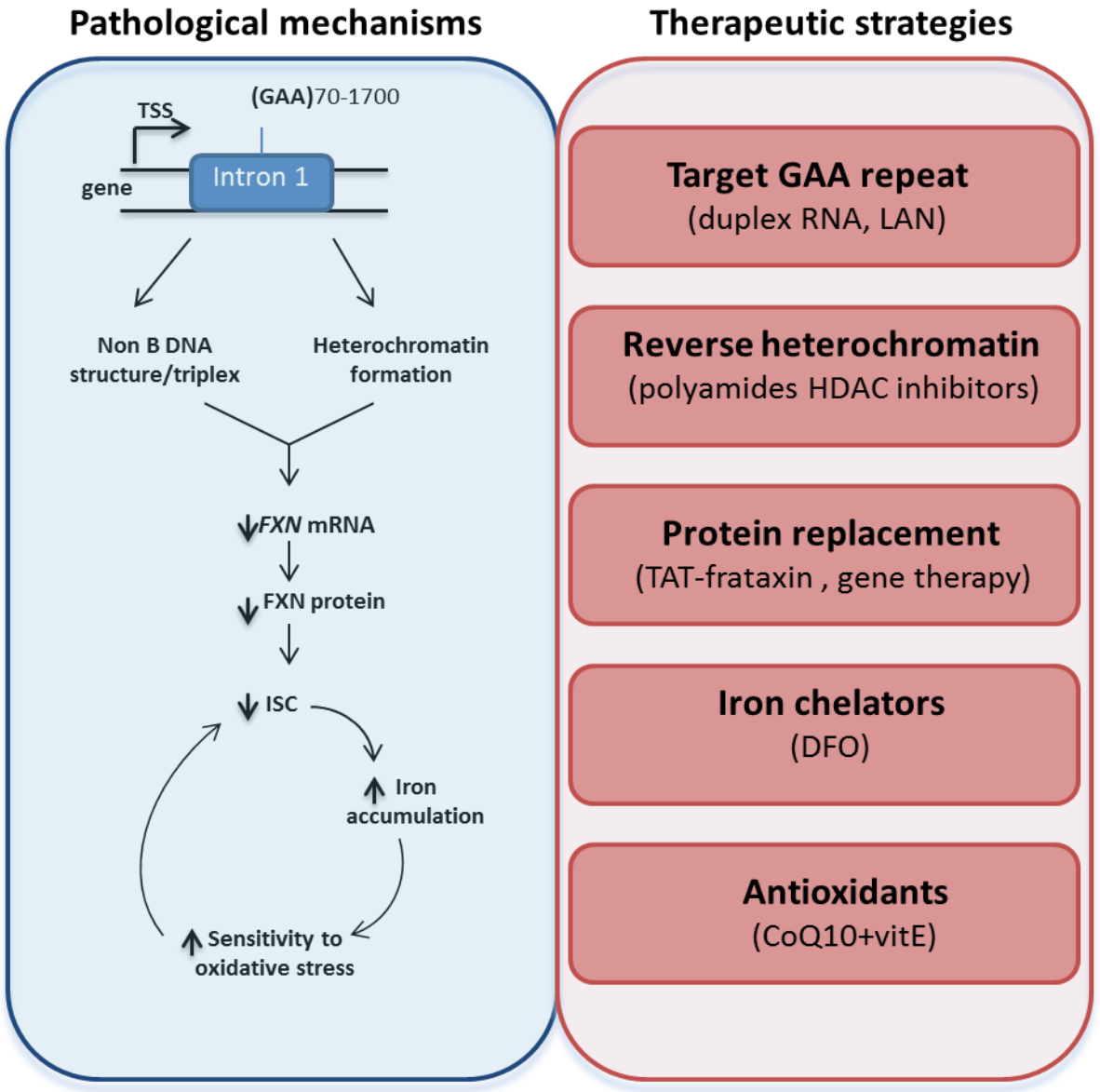

Figure 1.10: Pathological mechanisms and therapeutic strategies in FRDA (Schmucker and Puccio, 2010).

1.9 Aims of study

The exact molecular mechanisms causing FRDA are not yet well understood. However, accumulating evidence suggests that epigenetic alterations are involved in *FXN* gene silencing. In this research project, I aimed to study any possible effects that *FAST-1* might have on the *FXN* gene expression. Initially, I established stable HEK293, HeLa and fibroblast cell lines that overexpress *FAST-1*. Then, we

examined the *FXN* mRNA and protein expression, DNA methylation and histone modification profiles of the *FAST-1* overexpressing cells, as described in Chapter 3.

To investigate whether knocking down *FAST-1* could improve *FXN* expression in FRDA cells, as the next step of our study, I used the shRNA system to knock down *FAST-1* in FRDA cells, as described in Chapter 4.

Recently, the Ataxia Research Group at Brunel University London established a new YG8-derived FRDA mouse model, YG8LR, which was shown to have an average of 420 GAA repeats. YG8LR mice show decreased levels of *FXN* mRNA and frataxin protein in various tissues compared to YG8sR and control mice (Saba Saqlain, PhD thesis). To further characterise our new mouse model, I analysed the expression levels of *FAST-1* and histone modifications in YG8LR mice, as described in Chapter 5.

Chapter 2: Material and methods

2.1 General cell culture maintenance

Cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) high glucose, GlutaMAX™ supplement (Gibco) with addition of 10% foetal bovine serum (Gibco) and 2% penicillin/streptomycin 5000U/ml (Gibco). The complete media was filtered through a 0.22μm pore size filter unit (Nalgene). The prepared medium was stored at 4°C. Cells were grown at 37°C in 5% CO_2 humid incubator.

2.1.1 Thawing of cryopreserved cells

Vials were taken out from liquid nitrogen and transported in dry ice. To release any residual liquid nitrogen, a soaked tissue in 70% alcohol was wrapped around the vials and lids were loosened. To rapidly thaw the cells, the vials were put in a 37°C water bath for 1-2 minutes. Cells were transferred into a 15ml conical tube containing 10ml pre-warmed complete culture medium and centrifuged for 5 minutes at 200xg. The supernatant was discarded and the pellet was resuspended in 5ml medium and transferred into either 25cm^2 or 75cm^2 flasks containing complete culture medium. Cells were incubated at 37°C, 5% CO_2 with 92-95% humidity.

2.1.2 Cell freezing medium

Freezing medium was comprised of 10% (v/v) DMSO and 90% DMEM culture medium containing penicillin/streptomycin and FBS.

2.1.3 Passage of cultured cells

When cells reached 80-90% confluency they were passaged into fresh growth medium. First, DMEM, trypsin and PBS were warmed in a 37°C water bath. After washing cells with PBS, approximately 1ml of 0.25% Trypsin/EDTA was added and the flask was incubated in a humidified 5% CO_2 incubator at 37°C for 5 minutes to allow cells to detach from the flask. Trypsin/EDTA was inactivated by adding 5ml of

DMEM. Cells were pelleted at 200xg for 5 minutes. The supernatant was aspirated off and cells were recovered in DMEM and were subcultured at a 1:2 or 1:4 split ratio.

2.1.4 Mycoplasma PCR test

Mycoplasma PCR was carried out using MycoSensor PCR Assay Kit (Agilent technologies). 100μl of cell culture supernatant was heated at 95° for 5 minutes and centrifuged for 1 minute. StrataClean resin was resuspended by vortexing for 30 seconds and 10μl of it was added to heated supernatant and mixed by flicking the tube. To pellet the resin, the mixture was centrifuged for 1 minute at 13K rpm (microfuge) and 20μl of supernatant was transferred to a PCR tube on ice. The supernatant was used as the PCR template. A premaster mixture was prepared for 4 reactions by adding the following solutions:

- 54μl of PCR grade water
- 100μl of REDtaq
- 8μl of primer mix
- 16μl of internal control template

45μl of mastermix was transferred into each PCR tube and 5μl of test samples (HEK293 and HeLa) were added to the corresponding tubes. 5μl of sterile water and 5μl of positive control template were added to negative control and positive control reactions respectively. PCR amplification was performed with the following cycling conditions:

Table 2.1: Mycoplasma test PCR program

Step	Temperature	Time	Cycles
Initial denaturation	94°C	30sec	1
Denaturation	94°C	30sec	
Annealing	55°C	1min	35
Elongation	72°C	1min	
Extension	72°C	10mins	1

To visualize the results, 10μl of PCR product was run on a 1% agarose TBE mini-gel along with a 1kb plus DNA size marker (Invitrogen) at 75V for ~30 minutes (Figure 2.1). The following table 2.2 was used to interpret the results.

Table 2.2: Mycoplasma test PCR results

PCR template	PCR product(s)	Result
Cell culture extract +	315bp and 500bp	Mycoplasma infection
internal control template	315bp only	Heavy Mycoplasma infection
	500bp only	No Mycoplasma infection
Positive control +	None or 315bp only	Failed PCR reaction
internal control template	315bp and 500bp	Expected control result
Negative control +	500bp	Expected control result
internal control template	315bp and 500bp	Contaminated reagents

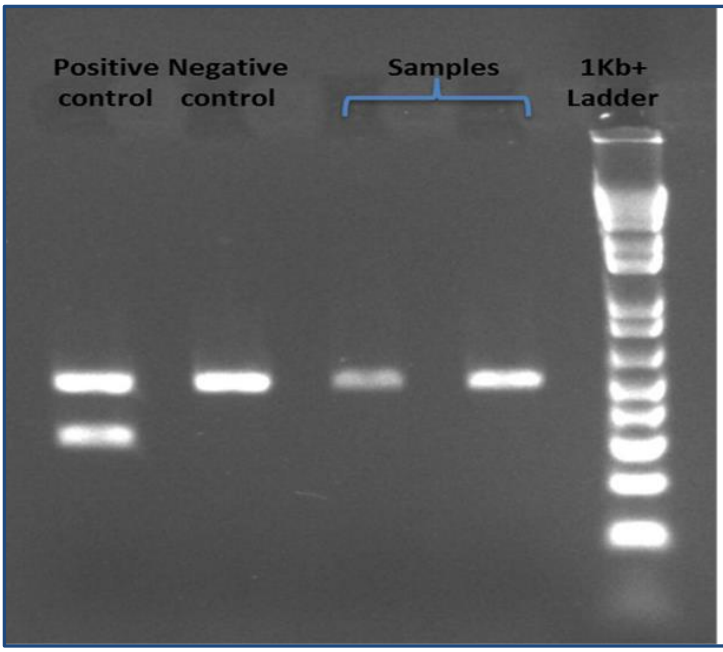

Figure 2.1: Mycoplasma PCR reaction products on a standard 1% agarose gel. The Mycoplasma PCR yields amplification of the 315bp PCR product from the most common species of *Mycoplasma* and the 500bp internal control product. The expected results for the controls are two bands at 315bp and 500bp for the positive control and one band at 500bp for the negative control.

2.2 Assay standards and kit

REDTaq® ReadyMix™ PCR Reaction Mix	MERCK
Taq DNA Polymerase PCR Master Mix	QIAGEN
Fast SYBR™ Green Master Mix	Applied Biosystems
Power SYBR® Green Master Mix	Applied Biosystems
Plasmid miniprep/ maxiprep/ megaprep kit	QIAGEN
Cloned AMV First-Strand cDNA Synthesis Kit	MERCK
DNaseI treatment RNase free kit	Thermo Scientific™
Custom Plus TaqMan Copy Number Assays	Thermo Scientific™
BCA Protein Assay Kit	Thermo Scientific™
GENECLEAN® III Kit	MP Biomedicals
Neon™ transfection system	Thermo Scientific™
Dipstick Assay Kit for Frataxin Quantity	ABCAM
Chromatrap ChIP QPCR Kit	Chromatrap
Custom Lentiviral shRNA	MERCK
Aconitase Assay Kit	Cambridge Biosciences Ltd

2.3 Primers

Primer sequences were obtained either from previous studies (as indicated) or newly designed and all primers were purchased from Sigma-Aldrich.

Table 2.3: Primers used for genotyping

Primer name	Sequence (5'-3')	Amplicon size
GAA repeat (Campuzano *et al.* 1996)		
GAA – F	GGGATTGGTTGCCAGTGCTTAAAAGTTAG	457bp+
GAA - R	ATCTAAGGACCATCATGGCCACACTTGCC	(GAA)n
***FXN* knock-out (Cossee et al. 2000)**		
WJ5	CTGTTTACCATGGCTGAGATCTC	520bp
WN39 (WT specific)	CCAAGGATATAACAGACACCATT	245bp
WC76 (KO specific)	CGCCTCCCCTACCCGGTAGAATTC	

Table 2.4: Primers used for quantification of *FXN* expression

Primer name	Sequence (5'-3')	Amplicon size
***FXN* expression (Human Specific) (Al-Mahdawi *et al.*, 2008)**		
Fxn-h-RT-F	CAGAGGAAACGCTGGACTCT	172bp
Fxn-h-RT-R	AGCCAGATTTGCTTGTTTGGC	
***FXN* expression (Human and Mouse)**		
FRT I-F	TTGAAGACCTTGCAGACAAG	121bp
RRT II-R	AGCCAGATTTGCTTGTTTGG	
***HPRT* (Human) (Bayot *et al.*, 2013)**		
HPRT-h-F1	GGTGAAAAGGACCCCACGA	90bp
HPRT-h-R1	TCAAGGGCATATCCTACAACA	
***HPRT* (Mouse) (Bayot *et al.*, 2013)**		
HPRT-m-F1	ATGAAGGAGATGGGAGGCCA	80bp
HPRT-m-R1	TCCAGCAGGTCAGCAAAGAA	

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

Primer name	Sequence (5'-3')	Amplicon size
***FAST-1* cDNA synthesis (De Biase *et al.*, 2009)**		
FAST RT	CCAAGCAGCCTCAATTTGTG	
PCR for *FAST-1* cDNA (De Biase *et al.*, 2009)		
FAST F1	GTGGGGGAGCAGCTAGAGG	
FAST R1	CACTTCCCAGCAAGACAGC	207bp
***FAST-1* expression**		
N- FAST F2	GACCCAAGGGAGACTGCAG	88bp
FAST R1	CACTTCCCAGCAAGACAGC	

Table 2.6: Primers used for ChIP qPCR analysis

Primer name	Sequence (5'-3')	Amplicon size
***FXN* – 5'UTR (De Biase *et al.*, 2009)**		
FXN-ChIP-R	CGAGAGTCCACATGCTGCTCC	131bp
FXN-ChIP-F	TCCTGAGGTCTAACCTCTAGCTGC	
***FXN* – promoter(Herman *et al.*, 2006)**		
FXN-pro - F	CCCCACATACCCAACTGCTG	147bp
FXN-pro - R	GCCCGCCGCTTCTAAAATTC	
H19 – CTCF positive(Burke *et al.*, 2005)		
h-H19 - F	CCCATCTTGCTGACCTCAC	
h-H19 - R	AGACCTGGGACGTTTCTGTG	165bp
***H19*neg – CTCF negative (Burke *et al.*, 2005)**		
h-H19neg - F	CATCATGGTGTCCTCACAGG	
h-H19neg – R	AGCTCTAAGGGAGGCTCCAG	161bp

2.4 General techniques

2.4.1 Agarose gel electrophoresis

Agarose gel electrophoresis was used to separate DNA fragments on the basis of their size. 0.8%-2% agarose gels were prepared in 1×TBE. The agarose was completely dissolved in TBE by using a standard microwave and allowed to cool for 5 minutes. Ethidium bromide was added to a final concentration of 0.2µg/ml. The gel was poured in a gel tank and allowed to set. 6×Orange G loading dye was added to samples to get a final concentration of 1x electrophoresis buffer (1xTBE) was added on top of the gel. Samples along with 1kb plus ladder (Figure 2.2) were run at 50V until to get a good band separation. Gels were visualized and documented by using a UVP GDAS (Gel Documentation Analysis System) 1200 under the ultraviolet light.

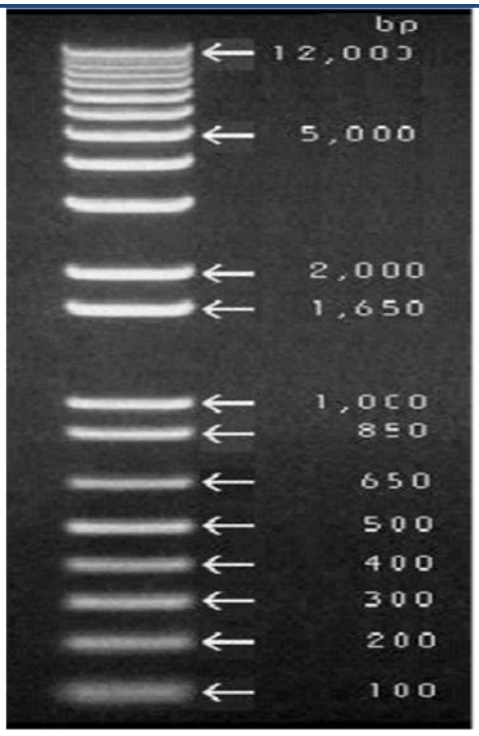

Figure 2.2 : 1kb plus DNA ladder

2.4.2 DNA extraction by the phenol/chloroform method

This method was used to extract genomic DNA from samples where greater DNA quality was necessary. Tissue samples or cell pellets were collected in Eppendorf tubes and 400µl of digestion buffer (100mM Tris (pH8.0), 5mM EDTA, 200mM NaCl, 0.2% (w/v) SDS) and 10µl of proteinase K (50mg/ml) were added, followed by a brief vortex and incubation at 55°C overnight. After digestion, by adding required sterile water the initial volume was adjusted to 200µl. An equal ratio of phenol was added to the mixture (200µl) and it was vortexed for 30 seconds. The tube was centrifuged at 14k rpm (microfuge) for 5 minutes at 4°C. The upper aqueous phase was removed and transferred to a fresh Eppendorf tube and 1 volume of chloroform/isoamyl alcohol (24:1, v/v) was added. The samples were mixed by vortexing for 30 seconds and centrifuged at 14k rpm for 10 minutes at 4°C. The upper layer was transferred to a fresh Eppendorf tube. 1/10 volume of 3M sodium acetate (pH 5.2) and 2 volumes of absolute ethanol were added. The mixture was incubated at -80°C overnight. To precipitate DNA, the tube was centrifuged 14k rpm for 30 minutes at 4°C and the DNA pellet was washed with 70% ethanol. After centrifuging for 8 minutes, the ethanol was drained off carefully and the DNA pellet was air-dried for 5 minutes. The DNA pellet was resuspended in 50µl of TE buffer.

2.4.3 Extraction of total RNA using the Trizol® method (Invitrogen)

Total RNA was extracted from approximately 10^6 cells. Medium was first aspirated off and cells were rinsed with 10ml of ice cold PBS. Trizol® was added directly to the cells and it was incubated for 5 minutes at room temperature. Cell lysate was passed through a pipet up and down several times. For the extraction of RNA from tissues, initially 30mg of a tissue was homogenized in 1ml of Trizol® in an Eppendorf tube. Samples were mixed and incubated for 10 minutes at room temperature. 0.2ml of

chloroform per 1ml of Trizol® was added, followed by vigorous shaking of samples for 15 seconds and incubation for a further 15 minutes at room temperature. Samples were phase separated by centrifugation at 14K rpm (microfuge) for 15 minutes 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 10 minutes at room temperature and centrifuged at 14K rpm for 15 minutes at 4°C. The supernatant was carefully removed and the RNA pellet was washed once with 1ml of 75% ethanol (made with DEPC-water) and centrifuged again at 8K rpm for 5 minutes at 4°C. The supernatant was removed carefully and the RNA pellet was briefly dried for 5-10 minutes and resuspended in 20µl of DEPC-water. RNA samples were stored at -80°C.

2.4.4 Checking RNA integrity

Isolation of intact RNA is essential for quantitative gene expression analysis. To check RNA integrity an aliquot of the RNA samples along with a DNA ladder were run on a 1.5% agarose gel. Intact total RNA had sharp, clear 28S and 18S rRNA bands (eukaryotic samples) (Figure 2.3).

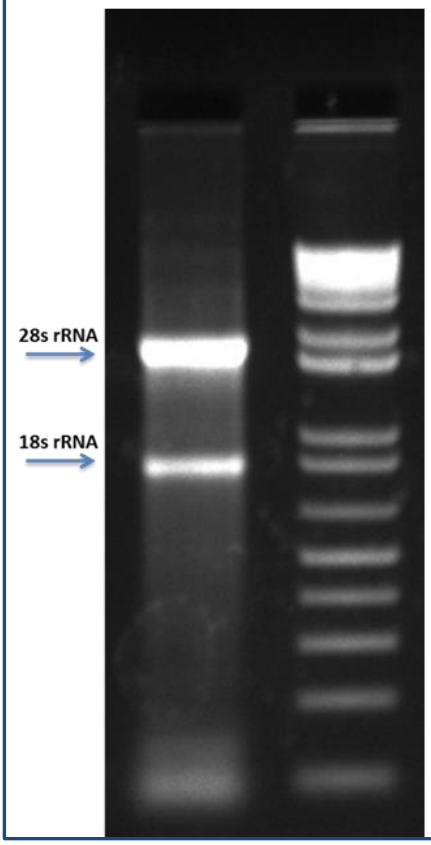

Figure 2.3: Intact RNA. Intact total RNA was run beside a Marker on a 1.5% agarose gel. The 18S and 28S ribosomal RNA bands are clearly visible in the intact RNA sample.

2.4.4 DNase I treatment of RNA

Deoxyribonuclease I, Amplification Grade (DNase I, Amp Grade, ThermoScientific) was used to eliminate DNA contamination from RNA samples. To remove any genomic DNA following reagents were added to 1µg of total RNA:

- 1µl 10×DNase I reaction buffer
- 2µl DNase I, Amp Grade, (1U/µl)
- DPEC-treated water to 10µl

The reaction was incubated for 15 minutes at room temperature. DNaseI reaction was inactivated by adding 1µl of 25mM EDTA solution to the reaction mixture and heating it for 10 minutes at 65°C. The RNA samples were used immediately for reverse transcription or stored at -80°C.

2.4.5 Complementary DNA (cDNA) synthesis

Complementary DNA (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.

For Oligo(dT) cDNA synthesis, 5µl (1µg) of RNA was added to 4µl of DEPC-water, 2µl 10mM dNTP mix and 1µl Oligo(dT)20 primer. For FAST RT strand specific cDNA synthesis, 5µl (1µg) of RNA was added to 4.5µl DEPC-water, 2µl 10mM dNTP mix and 0.5µl FAST RT (10µM) primer. All the preparation steps were carried out on ice. The mixtures were heated at 65°C for 5 minutes and then placed on ice immediately after. To bring all the contents of the tubes to the bottom, samples were centrifuged briefly and the following reagents were added in order: 4µl 5xcDNA 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 and briefly centrifuged. The reaction mixture was incubated at 55°C for 60 minutes. The reaction was terminated by heating the samples at 85°C for 5 minutes. The cDNA samples were used immediately or stored at -20°C (Sandi and Pook, 2015).

2.4.6 Determination of RNA/DNA concentration 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) was measured at 260nm to determine the DNA/RNA concentration and the quality was determined by using $A_{260/280}$ ratio.

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

Polymerase chain reactions (PCR) were performed using a Bio-Rad tetrad PCR machine. The right concentrations and precise annealing temperatures of each of the primer set 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:

TM = 4(G+C) + 2(A+T) °C

TA = TM - 5°C (where TM-melting temperature, TA-annealing temperature)

2.4.8 GAA PCR

1µl of extracted DNA was used as a template for GAA PCR together with 24µl of the reaction mixture. The GAA PCR primers (Table 2.7), master mix components (Table 2.8) and conditions are shown in (Table 2.9).

Table 2.7: GAA PCR primers

Primer name	Sequence (5'-3')	Amplicon size
GAA Forward	GGGATTGGTTGCCAGTGCTTAAAAGTTAG	
GAA Reverse	GATCTAAGGACCATCATGGCCACACTTGCC	421bp

Table 2.8: GAA PCR master mix components

Component	Volume(µl)/Reaction
2×Qiagen Master Mix (Qiagen) containing MgCl2, Taq DNA polymerase and dNTPs)	12.5µl
Q buffer	5µl
50µM GAA reverse primer	1µl
50µM GAA forward primer	1µl
dH_2O	to 25µl

Table 2.9: GAA PCR program

Step	Temperature	Time	Cycles
Initial denaturation	94°C	2mins	1
Denaturation **Annealing** **Elongation**	94°C 60°C 68°C	10sec 30sec 45sec	10
Denaturation **Annealing** **Elongation**	94°C 60°C 68°C	10sec 30sec 1min, with 20 sec increments	20
Extension	68°C	10mins	1

To visualise the PCR products, a 1% agarose gel with 50ml 1×TBE (89mM Tris-borate, 2mM EDTA, pH 8.3) buffer was prepared (section 2.4.1).10µl of samples along with $1kb^{+}$ DNA ladder (Invitrogen) were run at 75V for 30-60 minutes. Three controls were used. Control samples with known genotype were used as positive and negative controls. A blank control containing no DNA was used as a contamination control. No PCR amplification was expected in negative and blank controls (Figure 2.4).

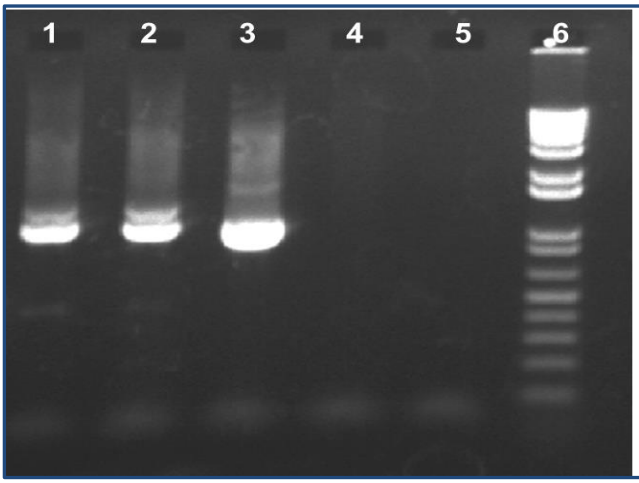

Figure 2.4: Agarose gel electrophoresis of GAA PCR reaction products from FRDA transgenic mouse tissues. Lanes 1 and 2 are samples. Lanes 3, 4 and 5 are positive, negative and blank (dH_2O) controls respectively. A $1Kb^{+}$ DNA ladder was used for size evaluation.

2.4.9 *Fxn* KO PCR

To determine the knockout genotype of *Fxn* (wild-type, heterozygous, rescue) the KO PCR was performed (Figure 2.5). 1µl of extracted DNA was used as a template for PCR together with 24µl of the reaction mixture. The KO PCR primers (Table 2.10), master mix components (Table 2.11) and conditions are specified in (Table 2.12).

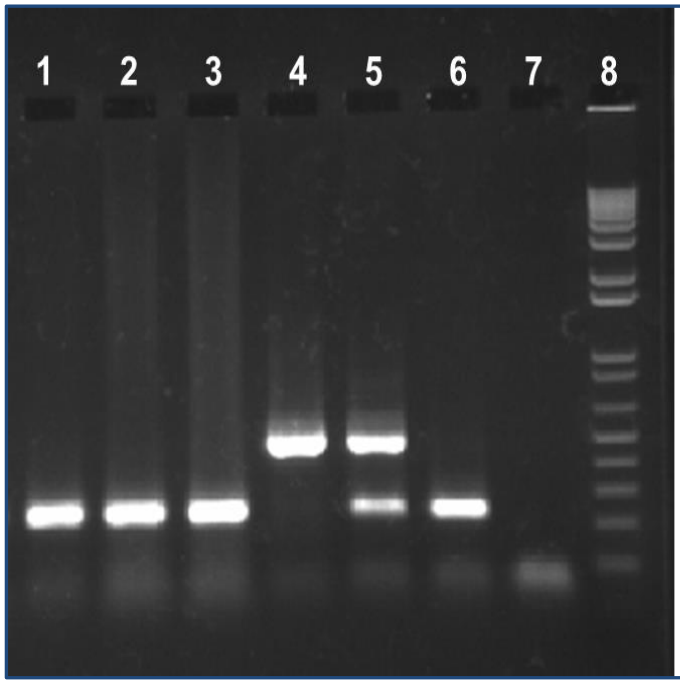

Figure 2.5: Agarose gel electrophoresis of *Fxn* KO PCR reaction products from FRDA transgenic mouse tissues. Lanes 1 to 3 are test samples. Lanes 4, 5, 6 and 7 are wild-type, heterozygous, rescue and blank (dH_2O) controls respectively. A $1Kb^+$ DNA ladder was used for size evaluation.

Table 2.10: *Fxn* KO PCR primers

Primer name	Sequence (5'-3')	Amplicon size
WJ5	CTGTTTACCATGGCTGAGATCTC	Wild-type allele (WJ5-WN39) = 520bp
WN39(WT specific)	CCAAGGATATAACAGACACCATT	
WC76 (KO specific)	CGCCTCCCCTACCCGGTAGAATTC	Mutant allele (WJ5-WC76) = 245bp

Table 2.11: *Fxn* KO PCR master mix components

Component	Volume(µl)/Reaction
REDTaq sigma master mix containing MgCl2, Taq DNA polymerase and dNTPs)	12.5µl
50µM WJ5	0.5µl
50µM WN39	0.5µl
50µM WC76	0.5µl
dH_2O	to 25µl

Table 2.12: *Fxn* KO PCR program

Step	Temperature	Time	Cycles
Initial denaturation	94°C	1min	1
Denaturation **Annealing** **Elongation**	94°C 54°C 72°C	20sec 20sec 20sec	30
Extension	72°C	10mins	1

2.4.10 Quantitative real-time PCR (qPCR)

Quantitative real-time PCR (qPCR) amplification was performed using SYBR Green (Applied Biosystems) in a QuantStudio 7 Flex Real-Time PCR instrument (Applied Biosystems). Reactions were carried out in 96 well plates (MicroAmp, Applied Biosystems), in triplicates, and in a final volume of 10µl containing optimized concentration of forward and reverse primers and gDNA/cDNA. Distilled water was added to non-template control reactions instead of template DNA. Mastermix was prepared on ice while trying to minimize light exposure. The reaction plate was sealed with MicroAmp® Optical Adhesive Film (Applied Biosystems) and centrifuged briefly at 1k rpm. The cycling conditions were selected according to the master mix

used and application. Alongside each real-time PCR run, a dissociation curve was performed by gradually increasing the temperature from 60°C to 95°C. Relative quantification values were determined by the $2^{-\Delta\Delta Ct}$ method.

2.4.11 Protein extraction

10^7 cells were collected in a 1.5ml Eppendorf tube and homogenised in 300 of RIPA (RadioImmunoPrecipitation Assay) buffer (Sigma-Aldrich) and 12µl of 25xprotease inhibitor (ThermoScientific). Samples were centrifuged at 14k rpm for 30 minutes at 4°C and the clear supernatant was collected in a fresh Eppendorf tube. A 50µl aliquot of the protein lysate was taken for BCA assay and DTT was added to the rest at a final concentration of 1mM. The protein lysates were stored at -80°C.

2.4.12 Determination of protein concentration using the BCA protein assay

The protein concentration was measured using a Pierce® BCA Protein Assay Kit (Thermo Scientific) following the manufacturer's instructions. This method combines the reduction of Cu^{2+} to Cu^{1+} by protein in an alkaline medium with highly sensitive and selective colorimetric detection of the cuprous cation ($Cu^{1+)}$ using bicinchoninic acid (BCA). The purple-coloured product of this assay has a strong absorbance at 562nm. To construct a standard curve, the contents of one bovine serum albumin (BSA) ampoule (2mg/ml) was diluted into several Eppendorf tubes as shown in (Table 2.13).

Table 2.13: Preparation of BSA standards for BCA analysis

Tube	BSA volume	dH_2O volume	Final volume	Final concentration
A	30 µl	10	20	1500µg/ml
B	20 µl	20	20	1000µg/ml
C	20 µl of A	20	40	750µg/ml
D	20 µl of B	20	20	500µg/ml
E	20 µl of D	20	20	250µg/ml
F	20 µl of E	20	30	125µg/ml
G	10 µl of F	40	50	25µg/ml
H	0	40	40	0

To perform BCA assay, protein lysate was diluted 1:10 with dH_2O.To prepare working buffer, 50 parts of BCA Reagent A were mixed with 1 part of BCA Reagent B (50:1, Reagent A: B). 10 µl of BSA standards and diluted protein samples were added to 200 µl of working buffer (A+B). The BCA reactions were incubated for 30 minutes at 37°C.

To measure A_{562}, the Protein BCA application was selected from NanoDrop™ 2000c home page. A standard curve was prepared by plotting A_{562} measurement for each BSA standard against its concentration. The standard curve was then used to determine the protein concentration of each study sample.

2.4.13 Dipstick assay

The level of frataxin protein was measured by lateral flow immunoassay with the Frataxin Protein Quantity Dipstick Assay Kit (MitoSciences) according to the manufacturer's instructions.

First 2µg of protein in 25µl of extraction buffer (buffer A) was mixed with 25µl of 2xblocking buffer (buffer B) and was added to individual wells on a 96-well plate (provided with the kit) with gold-conjugated monoclonal antibody at the bottom of each well. The samples were incubated for 5 minutes, allowing the gold-conjugate to

hydrate. The mixture was then resuspended gently using a pipette and dipsticks were inserted into the well. The samples were allowed to wick up onto the dipstick until the entire sample volume was absorbed (15-20 minutes). The dipsticks were washed for 20 minutes with 30µl of washing buffer (buffer C) in an empty well of the microplate. The dipstick was air-dried for approximately 20 minutes and the signal intensity was measured with a MS-1000 Immunochromatographic Reader (MitoSciences) (Figure 2.6).

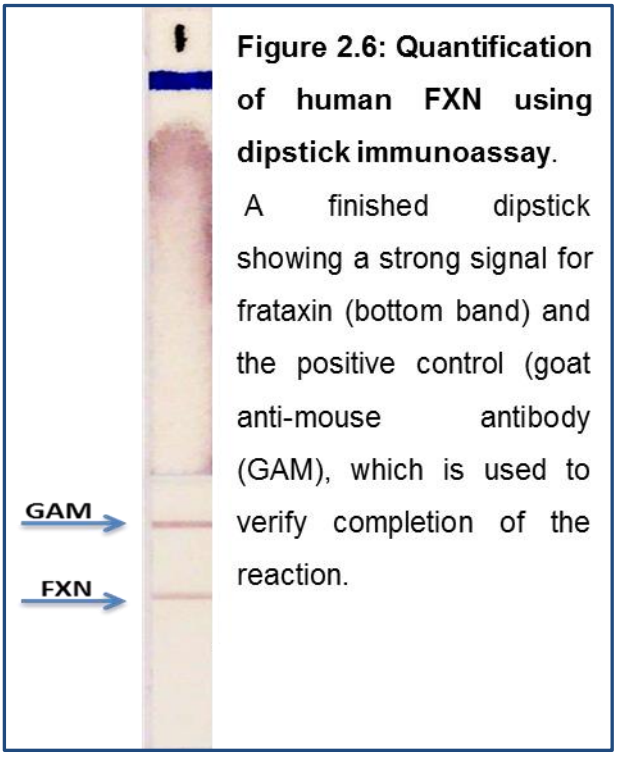

Figure 2.6: Quantification of human FXN using dipstick immunoassay. A finished dipstick showing a strong signal for frataxin (bottom band) and the positive control (goat anti-mouse antibody (GAM), which is used to verify completion of the reaction.

2.4.14 Chromatin immunoprecipitation (ChIP) analysis

This procedure was performed by using ChIP qPCR kit (Chromatrap) with an acetylated H3 (Lys9) (Upstate, 7-352), trimethyl-H3 (Lys9) (Upstate, 07-442) and CTCF (Upstate, 7-729), antibody on formaldehyde cross-linked samples. Initially, 4 million HeLa cells were cultured in a T75 flask. Once the flask became confluent, cells were washed with warm PBS and fixed in basic cell culture media containing 1% formaldehyde for 10 minutes at room temperature on a rocking platform. To

quench the reaction, the fixation solution was removed and 4ml of 0.65M glycine was added and cells were incubated for 5 minutes at room temperature with gentle agitation on a rocking platform. The cells were collected by scraping in ice cold PBS containing 1μl of proteinase inhibitor cocktail and pelleted by centrifugation at 3500×g for 5 minutes at 4°C. The cell pellets were resuspended in 800μl Hypotonic Buffer (provided by the kit) and incubated at 4°C for 10 minutes. The hypotonic slurries were pelleted by centrifugation at 5000xg for 5 minutes at 4°C and they were resuspended in 500μl of pre-warmed to 40°C Lysis buffer (provided by the kit). Samples were then incubated at 4°C for 10 minutes. Bioruptor® Pico sonication System (Diagenode) was used to achieve the desired length of DNA fragments (100bp-500bp). Chromatin samples from HeLa cells were sheared for 5 cycles of 30″ on/ 30″ off, using 0.1ml Bioruptor® Microtubes. The cell debris was removed by centrifugation for 10 minutes at 14K rpm (microfuge) 4°C, and the DNA collected as the supernatant phase. 1μl of proteinase inhibitor cocktail was added to the samples. To assess the quality of the DNA shearing, a 25μl aliquot of sonicated and unsonicated chromatin were reverse-cross-linked by adding 5μl of 1M $NaHCO_3$ and 5μl of 5M NaCl in a final volume of 50μl. Samples were then incubated at 65°C for 2 hours. 1μl of Proteinase K (provided with the kit) was added and incubated for 1 hour at 37°C. The reaction was stopped by adding 2μl of Proteinase K stop solution (provided with the kit). 5μl of the reverse-cross-linked chromatin samples were separated on a 1% agarose 1xTBE mini-gel. To have a correct size assessment, all chromatin samples were reverse-cross-linked before agarose gel electrophoresis (Figure 2.7). DNA fractionation was considered acceptable when the majority of the DNA was ≤500bp. The chromatin concentration was measured at 260nm. The reading was multiplied by 2 to account for the dilution during the reverse cross-

linking. The chromatin concentration was used to determine the volume of chromatin needed for the immunoprecipitation (IP) step.

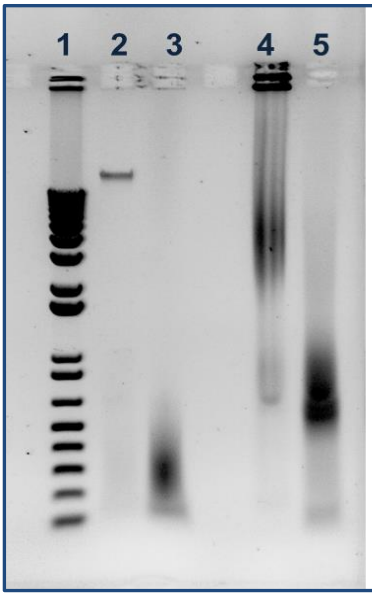

Figure 2.7: Comparison of shearing assessment with 1% agarose gel using reverse-cross-linked versus cross-linked chromatin.
To avoid any misleading result for the size assessment, only reverse-cross-linked samples should be used.4 million HeLa cells were fixed and were sheared for 5 cycles of 30″ on/ 30″ off. Prior to analysis on a 1% agarose gel, samples 2 and 3 were reverse-cross-linked. Samples 4 and 5 were loaded directly on the gel without any treatment.
1-1Kb+ ladder. 2- Reverse-cross-linked, unsonicated chromatin. 3- Reverse-cross-linked, sonicated chromatin. 4- Cross-linked, unsonicated chromatin. 5- Cross-linked, sonicated chromatin.

To minimise non-specific background interference from suspended solids, the sonicated chromatin was spun down at 16000×g for 10 minutes at 4°C.

A molar ratio 1:2 of chromatin to antibody was used for each IP. 2μg of sheared chromatin, 4μg of antibodies of interest, 5μl of Wash Buffer 1 (provided with the kit) and 1μl of PIC were added to a final volume of 40μl. For every antibody IP, the equivalent amount of chromatin in a final volume of 20μl was set aside and labelled as input. Spin Columns were prepared by adding 600μl of the Column Conditioning Buffer (provided with the kit). The conditioning step was repeated twice. Once the columns were ready, IP slurries were transferred to corresponding Chromatrap® spin column and incubated for 1 hour at 4°C on a rocking platform with gentle agitation. Each column was washed with the provided wash buffers five times. The chromatin was subsequently eluted by adding 50μl of elution buffer to each column, followed by incubation at RT, for 15minutes. The samples were centrifuged for 30

seconds at top speed. The cross-links were reversed and DNA was used for downstream processing, PCR and qPCR.

Real-time PCR was performed in triplicate using a QuantStudio 7 Flex Real-Time PCR instrument (Applied Biosystems). Each qPCR reaction was carried out in triplicates in a 10µl reaction volume by using 2µl of the eluted immunoprecipitated DNA, 5µl of Fast SYBR Green Master Mix (Applied Biosystems) and 12.5pmol of each of the respective forward and reverses primers (Table 2.6). Relative quantification values were normalised to input and minus antibody samples, and finally determined in relation to a control region.

Chapter 3: Overexpressing *FAST-1* in non-FRDA cell lines

3.1 Introduction

In order to determine if *FAST-1* has any effect on *FXN* gene expression, *FAST-1* was overexpressed in 3 different non-FRDA cell lines, HeLa, HEK293 and fibroblast, each with normal *FXN* expression. Previously, full-length *FAST-1* was isolated and characterised in our lab. To allow overexpression of *FAST-1*, the full-length *FAST-1* transcript was first cloned into a mammalian expression vector with a strong cytomegalovirus (CMV) promoter to drive high levels of *FAST-1* transcription. The three non-FRDA human cell lines were stably transfected with the recombinant plasmid and increased expression of the *FAST-1* was confirmed by isolating total RNA and measuring *FAST-1* expression levels in the transfected cell lines by quantitative real-time reverse transcription PCR (qRT-PCR). After confirming *FAST-1* overexpression, *FXN* mRNA expression changes were studied in the *FAST-1* overexpressing cell lines by qRT-PCR. In addition, frataxin protein expression levels were also measured using the frataxin lateral flow dipstick immunoassay.

FAST-1 expression levels in stably transfected cell lines were not only affected by the promoter driving *FAST-1* expression, but also by the *FAST-1* copy number in the genome. Therefore, *FAST-1* copy numbers were determined by quantitative TaqMan qPCR. Several aspects of epigenetics are strongly linked to non-coding RNAs. To investigate a potential explanation for the *FAST-1* overexpression effect on the *FXN* gene, epigenetic changes were studied at the *FXN* locus of the *FAST-1* overexpressing cells. Several previous studies have demonstrated that specific histone modifications are strongly associated with *FXN* gene silencing. Therefore, chromatin immunoprecipitation (ChIP) was carried out using antibodies against the characteristic heterochromatin marker trimethylated histone 3 lysine 9 (H3K9me3) and euchromatin marker acetylated histone 3 lysine 9

(H3K9ac). It has been reported that there is a negative association between *FAST-1* expression and CTCF occupancy level at the *FXN* locus. In FRDA, higher levels of *FAST-1* were associated with decreased CTCF binding at the 5′UTR of the *FXN* gene. On the other hand, global CTCF knockdown in normal cells was associated with increased level of *FAST-1*. In fact, after knocking down CTCF, *FAST-1* levels in normal cells showed a corresponding doubling and were equalized to the level seen in FRDA (De Biase *et al.*, 2009). To examine the effect of *FAST-1* overexpression on CTCF occupancy, ChIP with anti-CTCF was performed on *FAST-1* overexpressing HeLa cells. To complete the epigenetic studies on *FAST-1* overexpressing cells, the DNA methylation profiles of *FAST-1* overexpressing fibroblast cells were investigated. In this chapter, three non-FRDA *FAST-1* stably-overexpressing cell lines were generated and the corresponding effects on *FXN* expression and epigenetic changes were studied in those cell lines.

3.2 Materials and methods

3.2.1 Cloning

To overexpress *FAST-1* in HEK293, HeLa and fibroblast cell lines, *FAST-1* was cloned into a mammalian expression vector named pcDNA3 in which expression is driven by the human CMV promoter and it also contains a Geneticin® selectable marker. Previously in our laboratory, a full-length *FAST-1* fragment was cloned into pCR™4-TOPO® vector by TOPO TA® cloning and then it was transformed into DH5α competent cells.

3.2.1.1 Growth and maintenance of *E. coli*

For long-term storage, bacterial cultures were stored in 15% (v/v) glycerol at -80°C. *Escherichia coli* (*E. coli*) were grown in liquid LB media at 37°C with agitation at 250 rpm. *E. coli* transformed with plasmid was grown overnight in 5ml LB medium containing 50µg/ml of ampicillin at 37°C with agitation at 250 rpm. Transformed cells were spread onto agar plates containing 50µg/ml of the relevant selective antibiotic. The plates were incubated overnight at 37°C, after which colonies were picked using sterile 20µl pipette tips and grown overnight in 5ml liquid LB cultures. The plasmids were purified by using QIAprep spin mini kit.

3.2.1.2 Plasmid DNA purification using the QIAprep spin mini kit

5ml bacterial cultures were harvested by centrifuging at 4K rpm (microfuge) for 5 minutes at room temperature. The supernatant was discarded completely. Following vortexing, the pellet was resuspended in 250µl of P1 buffer (50mMTris-HCl, pH 8.0, 10mM EDTA, 100µg/ml RNase A). The plasmid DNA was denatured by adding 250µl of buffer P2 (200mM NaOH, 1% SDS). To neutralize the solution and re-anneal the denatured DNA, 350µl of buffer N3 (4.2 M Gu-HCl, 0.9 m potassium

acetate pH 4.8) was added and mixed immediately. After centrifuging for 10 minutes at 13K rpm in a microfuge, the supernatant was transferred to QIAprep spin column. To remove trace nuclease activity, 500µl of PB buffer (5M GU-HCl, 30% isopropanol) was used to wash the column and followed by 60 seconds centrifuging. The column was washed again with 750µl of PE (10mm Tris-HCl, pH 7.5, 80% ethanol) and centrifuged for 60 seconds. To remove residual wash buffer, the empty column was centrifuged for an additional 60 seconds. To elute the DNA, 50µl of pre-warmed buffer EB (10mM Tris-Cl, pH 8.5) was added to the column and left to stand for 5 minutes. Following centrifuging for 5 minutes, the purified plasmid DNA was collected in 1.5ml Eppendorf tube.

3.2.1.3 Restriction enzyme digestion

1µg of plasmid DNA was digested in a final volume of 20µl containing 1x buffer (supplied by the manufacturer) and the restriction enzyme. The volume of enzyme used varied depending of the concentration of the enzyme stock but did not exceed 10% (v/v) of the total reaction volume. Reactions were carried out as the manufacturer's instructions at the recommended temperature and duration. DNA digestion was verified by agarose gel electrophoresis.

3.2.1.4 Purification of DNA fragments from agarose gels by using the GeanClean III Kit

0.8% agarose mini-gels were prepared with a wide-toothed comb. Linearized vector and digested recombinant plasmid containing *FAST1* (~20µl) were run for approximately 2.5 hours at 75V in 1×TBE buffer. The desired bands were excised from the gel by a scalpel blade under 320nm wavelength UV light. After measuring the weight of excised bands, an approximate volume of gel slice was determined

(100mg=100μl). 1/2 volume of TBE modifier and 4.5 volumes of NaI were added to the gel slices. Following 5 minutes incubation at 55°C in a water bath the gel slices were completely dissolved. According to the final volume of the mix, calculated volume of glassmilk was added to each sample and the samples were then incubated for 5 minutes at 55°C in a water bath with gentle inversion every 1 minute to resuspend the glassmilk. Following centrifuging for 5 seconds at 13K rpm, the pellets were washed 3 times by adding 250μl of 'new wash' solution. The pellets were air-dried and then resuspended in 10μl of elution buffer. Following centrifuging at high speed for 30 seconds, the purified DNA samples were collected and their concentration and purity was determined by a NanoDrop™ 2000c spectrophotometer. DNA concentration was determined by reading the absorbance at 260nm wavelength (A_{260}). To assess the quality of DNA, the ratio of the absorbance at 260nm and 280nm ($A_{260/280}$) was measured. A ratio of ~1.8 is accepted as pure DNA. 1μl of gel purified plasmid and *FAST-1* fragments were run on a 1% agarose gel (Figure 3.1).

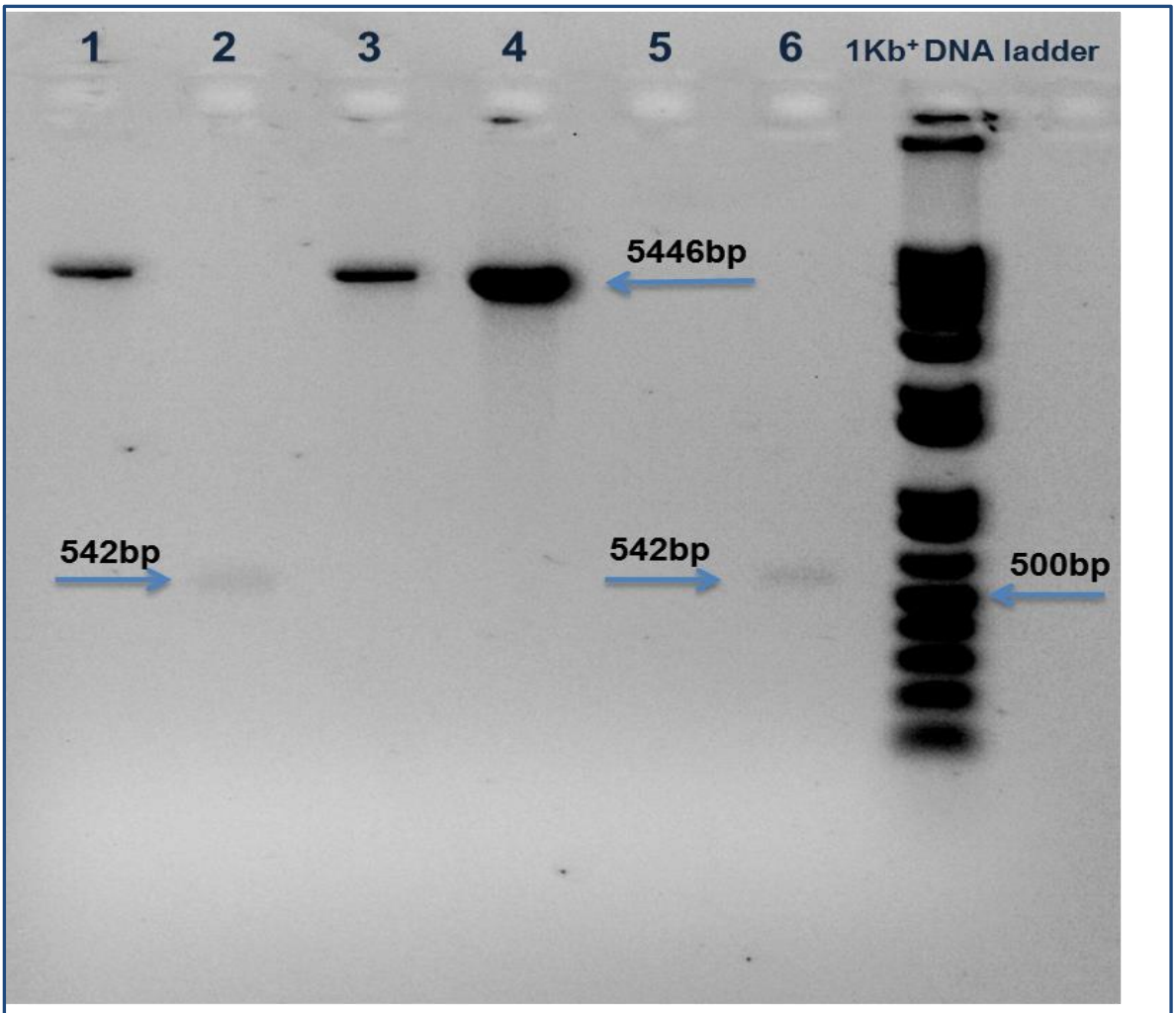

Figure 3.1: Gel purified pcDNA3 and *FAST-1*. 1µl of each gel purified samples was loaded on a 1% agarose gel. Lanes 1, 3 and 4: The gel purified pcDNA3 which appeared on the gel at 5446bp. Lanes 2 and 6: The gel purified *FAST-1* fragment with 542bp size.

3.2.1.5 Dephosphorylating expression vectors

To prevent recircularization of plasmid vectors, 2µg of the gel purified linearized plasmid was dephosphorylated in a final volume of 30µl containing 1µl of CIP (calf intestinal alkaline phosphatase) and 3µl of 10xdephosphorylation buffer. The reaction was carried out for 30 minutes at 37°C in water bath. The reaction was stopped by adding 5mM EDTA and heating the tube at 85°C for 15 minutes. The DNA was extracted by the phenol/chloroform method.

3.2.1.6 DNA ligation

A 1:2 molar ratio of vector to insert was used to clone the *FAST-1* fragment into 20ng of vector. 1μl of T4 DNA ligase and 2μl of 5×ligase buffers were added to a final volume of 10μl. Ligation reactions were incubated for 18 hours at 14°C.

3.2.1.7 Transforming competent cells

50μl of DH5α competent cells were transferred into pre-chilled Eppendorf tube (s). 5μl of DNA was added to the cells and mixed gently. Following incubation on ice for 30 minutes, the cells were subjected to a heat-shock at 42°C in a water bath, for exactly 20 seconds. The tube was then placed on ice for 2 minutes. 250μl of pre-warmed (37°C) SOC (super optimal broth with catabolite repression) was added to cells. To let the transformed cells express the antibiotic resistance, cells were shaken at 220 rpm, 37°C, for 1 hour. 100μl of SOC containing transformed cells was spread on an agar plate containing 50mg/ml of a relevant antibiotic. The plate was incubated in a 37°C incubator overnight.

3.2.1.8 Colony PCR screen

To identify plasmids containing the insert DNA, colony PCR was performed. Individual transformants were picked by pipette tips and dipped into 10μl sterile water. The tips were streaked on fresh agar plates containing the relevant antibiotic. To release the plasmid DNA from the cell, sterile water containing a transformant was heated to 98°C for 10 minutes. The mixture was used as a template for amplification reactions. Three primers were designed to determine the presence and correct molecular size of the insert (Table 3.1).

Table 3.1: List of primers used for colony screen

Primers	Sequences (5'-3')	Amplicon size
Insert positive		
pcDNA3-F	AGACCCAAGCTTGGTACCGA	623bp
pcDNA3-R	GCCAGTGTGATGGATATCTG	
Right orientation		
pcDNA3-F	AGACCCAAGCTTGGTACCGA	149bp
Insert-R	TAGTGCTAAGCTGGGAAGTTC	

The pcDNA3-F/ pcDNA3-R primers were designed to produce an amplicon of 623bp only if the insert is present in the plasmid construct. In another PCR amplification, pcDNA3-F was paired with an insert specific primer (Insert-R) to produce an amplicon of 149bp only if the insert is in the correct orientation. 1μl of 5mM forward primer, 1μl of 5mM reverse primer, 6.25μl of REDtaq ReadyMix PCR Reaction Mix (2×) (Sigma), 3.75μl of PCR-grade water and 1μl of heated transformed cells were added in a PCR tube. The PCR tube was centrifuged briefly and PCR was performed in a thermal cycler (Table 3.2).

Table 3.2: Screen PCR program

Step	Temperature	Time	Cycles
Initial denaturation	94°C	1min	1
Denaturation	94°C	20sec	
Annealing	55°C	30sec	30
Elongation	72°C	1min and 30sec	
Extension	72°C	1min	1

To visualize the results, PCR products were run on a 1% agarose gels along with a 1kb plus DNA size marker (Invitrogen) at 75V for ~45 minutes. Transformants that

produced both 623bp and 149bp amplicons were considered as positive clones which contain the *FAST-1* insert in the correct orientation (Figure 3.2). To perform diagnostic restriction enzyme digestion, 5 positive clones were picked and grown in 5ml LB and 5μl ampicillin and shaken at 220 rpm in a 37°C incubator overnight.

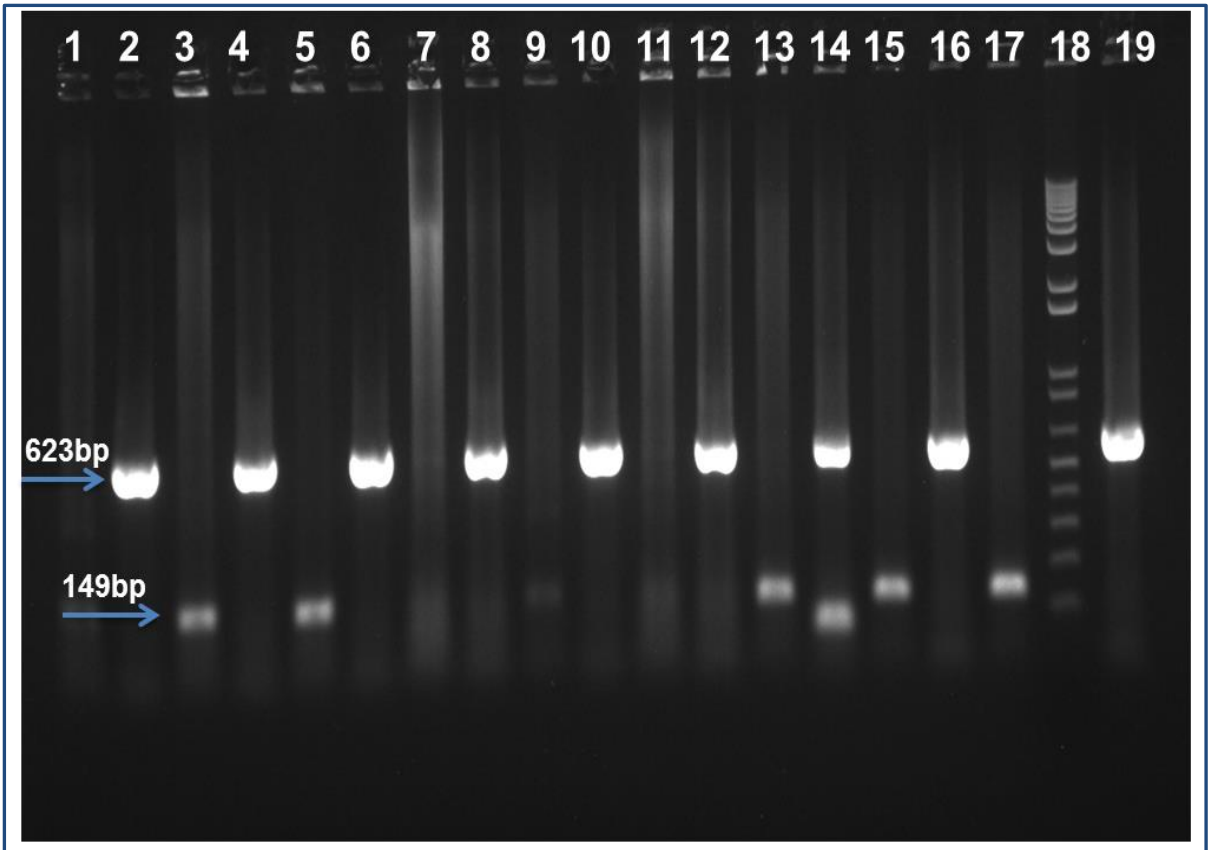

Figure 3.2: Colony PCR screen. The arrow pointing to 623bp amplicon size shows that all colonies contained the desired insert DNA. Lanes 1, 3, 5, 9, 13, 15, and 17: The existence of an amplicon of 149bp indicated that the desired DNA insert was in the correct orientation. Lane 18: 1kb plus DNA ladder.

3.2.1.9 Diagnostic restriction enzyme digestion

The plasmid DNA was isolated from the overnight bacterial cell cultures using QIAprep spin Miniprep kit, and then 2ul of the DNA were incubated with *Eco*RI for 2 hours at 37°C. To verify that the plasmid backbone and insert are of the expected

size, the digested plasmid was then run on a 1% agarose gel along with a 1kb plus DNA ladder (Figure 3.3).

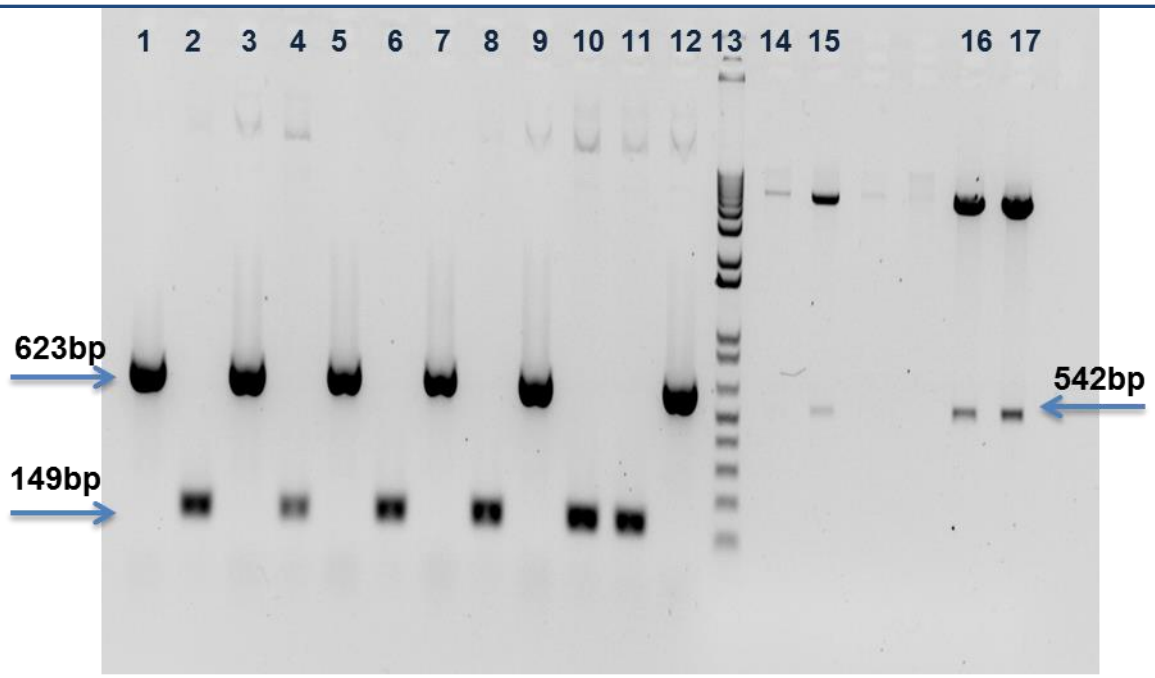

Figure 3.3: Colony PCR screen and diagnostic restriction digest. The colonies which passed the colony PCR screen were picked and DNA was extracted from them. The PCR with same primers (pcDNA3F/pcDNA3R. pcDNA3F/InsertR) was performed to check the amplicon size from pure plasmid DNA. Lanes 1 to 12: Amplicons resulting from 6 positive colonies. The presence of the 623bp and 149bp amplicons showed that the plasmid construct contained insert in the correct orientation. Lane 13: 1kb plus DNA ladder. Lanes 14, 15, 16, 17: Extracted plasmids DNA were digested with *Eco*RI and 10μl of the digestion reactions were run on a 1% agarose gel. Released fragment which appeared on the gel at 542bp, demonstrated that plasmid construct contained the right size insert.

3.2.2. Transfections

3.2.2.1 Lipofectamine

Plasmid DNA containing the *FAST-1* insert and the pEGFP-C1 plasmid were prepared by using the EndoFree Plasmid Maxi Kit. To measure the transfection efficiency, pEGFP-C1 plasmid was co-transfected with test plasmid (pcDNA3/*FAST-1*) into a separate well of HEK293 and HeLa cells. Mock-transfection was performed for both cell lines. Appropriate numbers of HEK293 and HeLa cells were seeded in a 24-well plate to obtain 80% confluency on the day of transfection. Lipofectamine 3000 (Life Technologies) was used as a DNA carrier for the transfection. 2μl of the Lipofectamine 3000 was diluted in 50μl of Opti-MEM I Reduced Serum Medium (Life Technologies) and 1.6μg of DNA was mixed with 50μl of Opti-MEM and 2μl of P3000 Reagent. The DNA-liposome complex was prepared by adding diluted DNA into diluted Lipofectamine and incubating for 5 minutes at room temperature. 50μl of DNA-liposome mixture was added to each well containing cells. The cells were incubated at 37°C in a 5% CO_2 incubator with 92-95% humidity for 5 hours. To reduce toxicity of Lipofectamine, after 5 hours the transfection media was removed and replaced with complete medium and incubated overnight at 37°C in a 5% CO_2 incubator. The following day, cells were trypsinised and transferred into 5ml complete medium and seeded in a 6-well plate in 1/5, 1/20 and 1/50 dilutions. Cells were incubated overnight at 37°C in a 5% CO_2 incubator with 92-95% humidity. The following day, the media was aspirated and replaced with fresh DMEM containing the appropriate concentration of G418 antibiotic. The selection media was replaced every 3-4 days. According to the antibiotic kill assay, HeLa and HEK293 cells were selectable with 200μg/ml and 400μg/ml of G418, respectively.

3.2.2.2 Electroporation

To measure the transfection efficiency, pEGFP-C1 plasmid was co-transfected with test plasmid (pcDNA3/*FAST-1*) into a separate well of fibroblast cells. One to two days prior to electroporation, cells were seeded in flasks with fresh growth medium to reach 70-90% confluency on the day of transfection. On the day of transfection, stable transfections of fibroblast cells were performed by using Neon™ transfection system. Cells were washed twice with PBS (-Mg+, Ca_2+) and cell counts were performed after trypsinization. 10^6 cells were transferred into a 1.5ml microcentrifuge tube and cells were centrifuged at 400xg for 5 minutes at room temperature. The cell pellet was resuspended in 140µl of buffer R (provided with the kit) and 5µg of DNA was added to the mixture. The cells were picked up with the Neon™ pipette by using Neon™ tips and exposed to an electric current (1200V, 20ms, 2 pulses) in the Neon™ cuvette filled with 3ml of Buffer E (provided with the kit). After electroporation, samples were diluted 10-fold in 900µl medium. 100µl of the diluted samples were added to 1900µl pre-warmed culture medium. Cells were then left to grow in media without antibiotics for 24 hours. The following day, the *GFP* expression was visualised under a fluorescent microscope (Figure 3.4) and media was aspirated and replaced with fresh DMEM containing appropriate concentration of G418 antibiotic. The selection media was replaced every 3-4 days. According to the antibiotic kill assay, transfected fibroblast cells were selectable with 100µg/ml of G418.

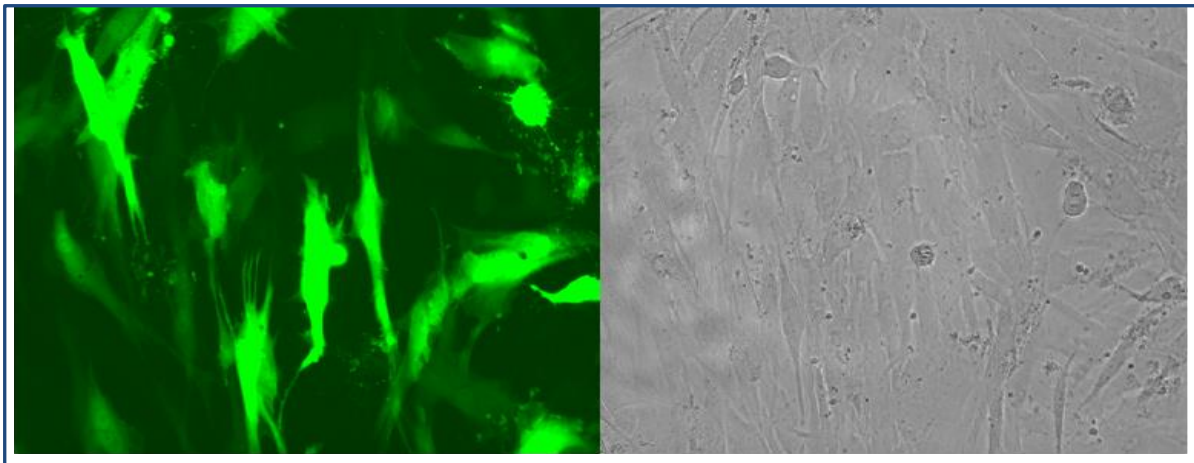

Figure 3.4: The expression of GFP (green fluorescent protein) in fibroblast cells co-transfected with pcDNA3/*FAST-1* and pEGFP-C1 plasmids under a fluorescent microscope. Cells were transfected by Neon™ transfection system and visualised by the Floid Cell Imaging Station fluorescent microscope, 24 hours post transfection.

3.2.3 Antibiotic kill curves

To determine the optimal antibiotic concentration for selecting stable colonies, antibiotic kill curves were determined. The appropriate numbers of cells were seeded in a 24-well plate to reach up to 20% confluency at the day of transduction. Cells were incubated overnight at 37°C, 5% CO_2 with 92-95% humidity. The next day medium was replaced with complete medium containing antibiotic G418 ranging from 0-700µg/ml for HeLa and fibroblast cells and 0-1200µg/ml for HEK293 cells. The experiment was performed in triplicate for each cell line. The media containing selection antibiotic was replaced every 4 days. After 2 weeks, cells were trypsinised and suspended in DMEM. To determine the lowest concentration that causes complete cell death, cell viability was measured in different concentrations of G418. 20µl of 0.4% (w/v) Trypan blue solution was added to 20µl of cell suspension and allowed to stand for 5 minutes. A small amount of Trypan blue-cell suspension mixture was transferred to a chamber port of Countess™ Cell Counting Chamber Slide (Invitrogen) and live and dead cells were counted by a Countess™ Automated

Cell Counter (Invitrogen). The lowest concentration which killed 100% of the cells was chosen as the optimal concentration.

3.2.4 Overexpressing *FAST-1* in non-FRDA cell lines

To construct a *FAST-1* delivery vector, full length *FAST-1* transcript was inserted into the multiple cloning site of expression vector pcDNA3. The pcDNA3 plasmid is a mammalian expression vector with the CMV promoter and ampicillin resistance in bacteria and neomycin resistance in mammalian cells. *FAST-1*-pCRTM4-TOPO® and pcDNA3-containing *E. coli* bacteria were cultured overnight at 37°C in LB medium containing 50µg/ml ampicillin. Plasmids were purified using the QIAprep spin mini kit (section 3.2.1.2). To release the *FAST-1* insert from the pCRTM4-TOPO® vector and to linearise pcDNA3, both plasmids were digested by *Eco*RI and run on 0.8% agarose gel. Desired bands were cut out and purified using a GeanClean III Kit (section 3.2.1.4). The pcDNA3 vector was dephosphorylated and ligated with the insert (*FAST-1*). DH5αTM competent cells were transformed with ligation product and plated overnight on LB/AMP. Colonies were screened via PCR (section 3.2.1.8) and the presence of the insert was confirmed by restriction enzyme digestion (section 3.2.1.9). Once the presence and correct orientation of the *FAST-1* insert was confirmed, human HeLa, HEK293 and fibroblast cell lines were stably transfected with pcDNA3/*FAST-1* (section 3.2.2). A GFP-tag plasmid was co-transfected to check the transfection efficiency. The transfected cells were ells were put into G418 selection. In order to obtain a monoclonal cell population, limiting dilution method was performed. Two days after transfection, transfected HeLa and HEK293 cells were harvested and diluted 1/100 and 1/50 respectively. Cells were cultured in selective media for 5 days to expand. The culture dishes were examined under an inverted microscope. The well-isolated clones were found under the

microscope and a circle was drawn around each clone on the bottom of the culture dish. Then, cells were washed with PBS and picked by placing a drop of trypsin on each cell clone. The resistant single cell colonies were picked and further grown. Since fibroblast cells need secreted factors from their neighbouring cells to grow, they did not undergo this step. 12 stable *FAST-1* overexpressing HeLa and HEK293 cell lines and 3 stable *FAST-1* overexpressing fibroblast cell lines were developed in selective media containing 200μg/ml, 400μg/ml and 100μg/ml of G418, respectively.

3.2.5 Quantification of *FAST-1* and *FXN* mRNA levels in *FAST-1* overexpressing cells

Total RNA was isolated from *FAST-1* overexpressing cells using the Trizol method. Once RNA samples were prepared, agarose gel electrophoresis was used to visually assess the quality of RNA samples. RNA concentration was measured by NanoDropTM 2000c spectrophotometry. To remove any residual gDNA, RNA samples were subjected to DNase I treatment (Invitrogen). DNase I-treated RNA was amplified using FAST F1 and FAST R1 primers (for *FAST-1* cDNA) and PCR products were run on an agarose gel. Amplification of DNase I-treated RNA samples indicated that there was no gDNA contamination in the RNA samples. To synthesise cDNA, the cloned AMV first-strand cDNA synthesis kit was used as per the manufacturer's instructions. Oligo (dT) and *FAST-1*-specific primer (FAST RT) were used for cDNA synthesis. To ensure that *FAST-1* amplification was not from self-priming, FAST RT primer was omitted from the cDNA synthesis mixture. In order to check if the cDNA synthesis had worked, the *FAST-1* cDNA was amplified with the N-FAST F2 and FAST R1 primer set to produce an 88bp product. The oligo-dT-primed *FXN* and *HPRT* cDNAs were amplified using FXN-h-F/R and HPRT-h-F/R primer sets respectively. The PCR products were run on a 2% agarose gel. In

samples without FAST RT primer, no PCR product was detected on the agarose gel, indicating that *FAST-1* amplification was not a result of endogenous priming. The conditions and the PCR master mix components are shown in (Table 3.3).

Table 3.3*: FAST-1* cDNA PCR program

Step	Temperature	Time	Cycles	Components	Volume(µl)/Reaction
Initial denaturation	95°C	30sec	1	**Qiagen MasterMix (2X)**	12.5
Denaturation	95°C	30sec		**Q buffer**	5
Annealing	60°C	1min	40	**N-FAST F2 primer(10µM)**	1
Elongation	70°C	1min		**FAST-R1 primer(10µM)**	1
Extension	72°C	10mins	1	**dH_2o**	3.5

2µl of the diluted synthesised cDNA was used as a template for qRT-PCR together with 8µl of the reaction mixture. The qRT-PCR experiment was run on standard program using a QuantStudio 7 Flex Real-Time PCR instrument (Applied Biosystems). The qPCR master mix components are shown in (Table 3.4).

Table 3.4: *FAST-1* qPCR components

Components ***FAST-1***	Components ***FXN***	Components ***HPRT***	Volume per reaction (µl)
Power SYBR® Green	***Power SYBR® Green***	***Power SYBR® Green***	*5*
N-FAST F2 primer(50µM)	**FXN-h-F primer(50µM)**	**HPRT-h-F primer(50µM)**	0.05
FAST-R1 primer(50µM)	**FXN-h-R primer(50µM)**	**HPRT-h-R primer(50µM)**	0.05
dH_2O	dH_2O	dH_2O	2.9
cDNA (1:5)	cDNA (1:10)	cDNA (1:100)	2

3.2.6 TaqMan ® copy number assay

TaqMan® copy number assays were run simultaneously with a TaqMan® copy number reference assay in a duplex real-time PCR. The copy number assay amplifies the target genomic sequence of interest and the reference assay amplifies a sequence that is known to exist in two copies in a diploid genome. A TaqMan® copy number reference assay (Applied Biosystems), RNase P forward and reverse primers and a TaqMan TAMRA™probe, and two custom designed TaqMan® copy number assays (Applied Biosystems), containing *FAST-1*-detecting forward and reverse primers and a TaqMan MGB probe, were used in this experiment. The master mix listed in (Table 3.5) was prepared.

Table 3.5: Reaction mixture components and volumes

Reaction mixture component	Volume per well (µl) 96-well plate
TaqMan® Universal Master Mix II (Applied Biosystems)	10.0
TaqMan® Copy Number Assay, 20× working stock	1.0
TaqMan® Copy Number Reference Assay, 20×	1.0
Nuclease-free water	4.0
Total Volume	16.0

The following samples were run on each plate using four replicates for each gDNA sample:

- **Samples:** gDNA samples in which the copy number of the target is unknown.
- **No Template Controls (NTC):** A sample that does not contain a DNA template, used for the detection of contamination and background fluorescence.

- **Calibrator sample:** A DNA sample with a known copy number for the target of interest.

16µl of the reaction mixture was added into each well of a MicroAmp® *Fast* Optical 96-well Reaction Plate (Applied Biosystems). The diluted gDNA samples were vortexed and spun briefly before adding to the plate. 4µl of gDNA (5ng/µl) was added to the wells containing the reaction mixture. The reaction mixture and gDNA were mixed by pipetting up and down for several times. The reaction plate was sealed with MicroAmp® Optical Adhesive Film (Applied Biosystems) and centrifuged briefly. The reaction plate was loaded into a QuantStudio 7 Flex Real-Time PCR instrument (Applied Biosystems) and a standard program was run. The number of *FAST-1* copies in each genomic DNA sample was determined by relative quantitation (RQ) normalized to the known copy number of the reference sequence using the comparative Ct (ΔΔCt) method and the Ct data were subsequently compared to a calibrator sample (non-*FAST-1* transfected cells) known to have two copies of *FAST-1* sequence, analysed by Applied Biosystems CopyCaller™ Software v.2.0. The predicted copy number values were calculated by rounding up the actual copy number values. TaqMan ® Copy Number Assay requires two target specific primers and a probe which anneals to its complementary sequence between forward and reverse primers. TaqMan probes are covalently joined to a reporter dye at their 5' ends and a quencher dye at their 3′ ends. When the quencher is in the close proximity to the reporter, no increase in fluorescent signal from the reporter dye can be detected. During each round of PCR, *Taq* DNA polymerase, which has 5' exonuclease activity, cuts the probe and physically separates the reporter from the quencher dye. This creates a permanent increase in fluorescence of the reporter dye. QuantStudio 7 Flex Real-Time PCR instrument (Applied Biosystems) can

monitor and record this increase in fluorescence and generate an amplification plot (Figure 3.5).

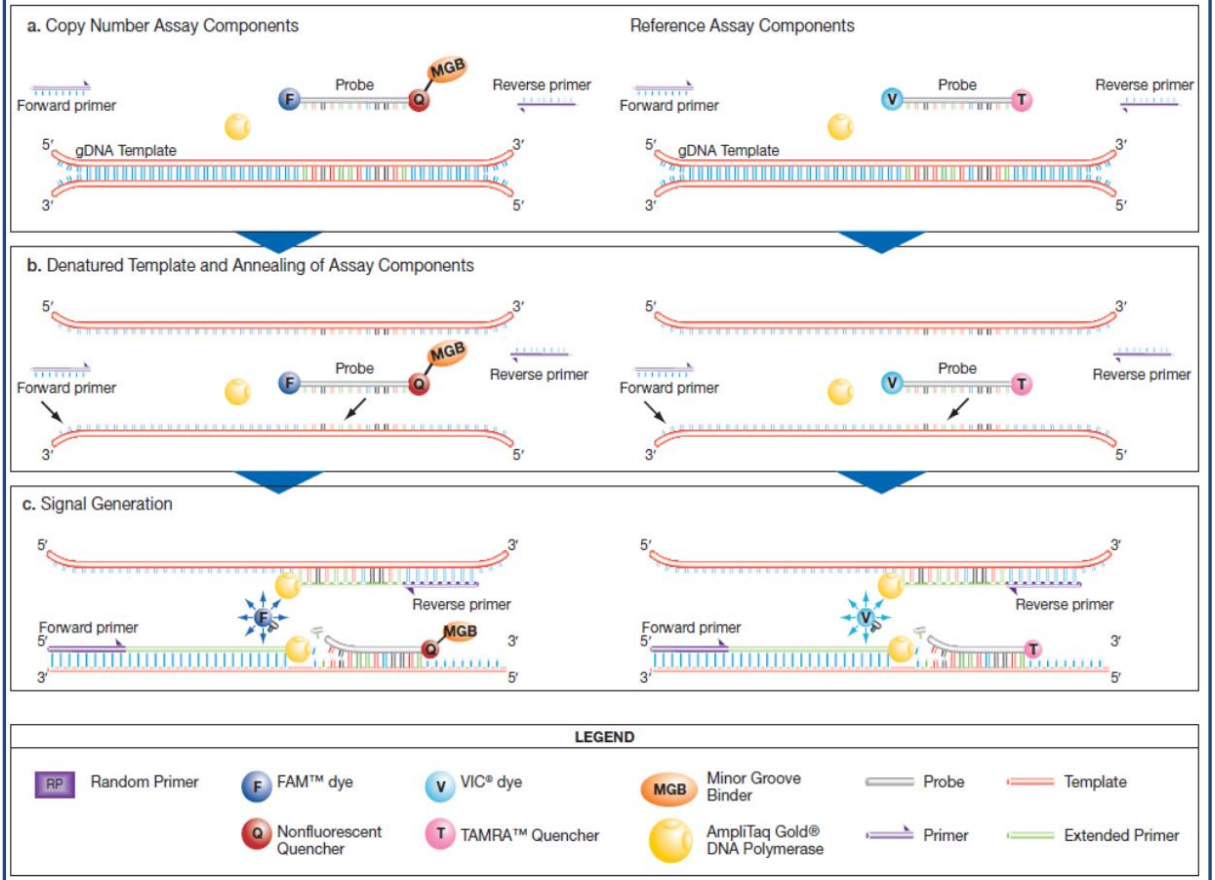

Figure 3.5: PCR and detection of target and reference gDNA sequences in a duplex reaction (Applied Biosystems, 2010).

3.2.7 MethylScreen assay

DNA methylation analysis was performed using the previously described MethylScreen assay (Al-Mahdawi et al., 2013). This technique uses combined restriction digestion of DNA with a methylation-sensitive restriction enzyme (MSRE) and methylation-dependent restriction enzyme (MDRE) as both single and double digests (Figure 3.6). The MethylScreen assay was used to analyse the *FXN* U6 and U11 CpG regions (Figure 3.7). Genomic DNA was extracted by the phenol/chloroform method and 0.5µg of each sample was used in each of four tubes,

for four different restriction enzyme digestions: (1) a MSRE, (2) a MDRE, (3) both MSRE and MDRE (double digest, DD), and (4) neither MSRE or MDRE (mock control). The digestion reaction components are shown in (Table 3.6 and 3.7).

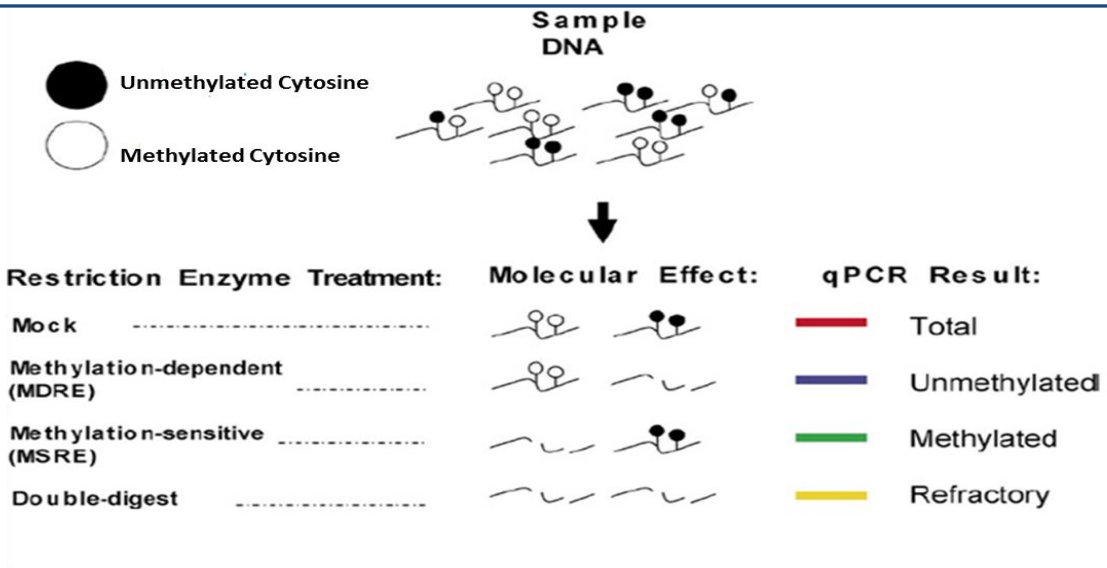

Figure 3.6: Overview of the MethylScreen technology procedure performed at the *FXN* CpG6 (U6) and CpG11 (U11) sites. The SYBR Green kinetic reaction profiles are presented as follows: red represents the mock-treated reaction, quantifying the total molecular population under study; blue reflects the methylation-dependent restriction enzyme (MDRE) reactions, quantifying the unmethylated molecular population; green represents the methylation-sensitive restriction enzyme (MSRE)-treated portion and allows quantification of the densely methylated fraction. The fraction of the population not amenable to analysis results from using the double-digest as template (gold) (Holemon *et al.*, 2007).

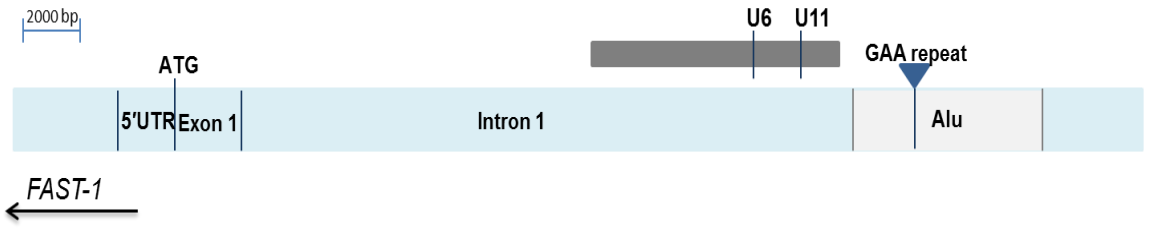

Figure 3.7: The *FXN* region of MethylScreen analysis. Schematic diagram of 2.5kb of the *FXN* exon 1 and intron 1 region showing the GAA repeat, CpG6 (U6) and CpG11(U11) sites and the region corresponding to *FAST-1* transcript (Al-Mahdawi *et al.*, 2013).

Table 3.6: U11 MethylScreen digestion reactions components

Components	Volume	*Aji*I	McrBC	Double Digest	Mock
0.5µg DNA	5µl	✓	✓	✓	✓
***Aji*I (5U/µl)**	2µl	✓		✓	
McrBC (10U/µl)	1µl		✓	✓	
10 X TANGO™	10µl	✓	✓	✓	✓
GTP (100mM)	1µl		✓	✓	
50% Glycerol	3µl				✓

Each reaction was made up to 100ul with dH_2O

Table 3.7: U6 MethylScreen digestion reactions components

Components	Volume	*Hpy*188III	McrBC	Double Digest	Mock
0.5µg DNA	5µl	✓	✓	✓	✓
***Hpy*188III (10U/µl)**	1µl	✓		✓	
McrBC (10U/µl)	1µl		✓	✓	
10 X TANGO™	10µl	✓	✓	✓	✓
GTP (100mM)	1µl		✓	✓	
50% Glycerol	3µl				✓

Each reaction was made up to 100ul with dH_2O

For site U11, *Aji*I (Fermentas) and McrBC (Fermentas) were used respectively as the MSRE and MDRE, respectively. For site U6, Hpy188III (New England Biolabs) and McrBC (Fermentas) were used respectively as the MSRE and MDRE respectively. The digestion reactions were carried out for 16 hours at 37°C in water bath. The reactions were stopped and enzymes were deactivated by heating the tubes at 65°C for 20 minutes. Samples were then placed on ice and briefly centrifuged to collect contents at the bottom of the tubes. To check the efficiency of digestion reactions,

5µl of each reaction was run on a 1% agarose gel. 5µl of DNA was used as a template for qPCR together with 15µl of the reaction mixture. Real-time PCR was performed in triplicate using a QuantStudio 7 Flex Real-Time PCR instrument (Applied Biosystems). Each qPCR reaction was carried out in triplicates in a 20µl reaction volume by using 500ng of the digested DNA, 10µl of Fast SYBR Green (Applied Biosystems) and 1µl of 5µM each of the respective forward and reverse primers (see Table 3.8 for primers used). The MSRE enzymes selectively digested DNA at unmethylated cytosine residues and left the fully methylated DNA intact. Therefore, in MSRE digested samples, the amplified products were generated from the fraction of fully methylated DNA. In contrast, MDRE enzymes cut DNA at methylated cytosine residues. Because methylated DNA was fully digested by MDRE, the amplified products in qPCR represented the fraction of unmethylated DNA. In double digested tubes, where DNA was digested by both MSRE and MDRE, full digestion and no PCR amplification were expected. In principle, no digestion should occur in mock control and the lowest Ct values belonged to the undigested controls (Figure 3.8). Therefore, Fmet and Rmet primers produced an amplicon of 196bp in the U11 mock control and a 160bp product was expected in the U6 mock control. PCR quantification was carried out using the ΔCt method (values were calculated as $2^{\Delta Ct}$ (mock-digest) with the mock value set at 100%) and RQ Manager Software (Applied Biosystems). MethylScreen DNA methylation values were then calculated as follows: Densely methylated DNA (DM)=(MSRE-DD)/(100-DD)×100; unmethylated DNA (UM)=(MDRE-DD)/(100-DD)×100; intermediately methylated DNA (IM)=100–(DM+UM).

Table 3.8: Primer used in MethylScreen studies

Primer name	Sequence (5'-3')	Amplicon size
Fmet	GATCCGTCTGGGCAAAGGCCAG	
Rmet	ATCCCAAAGTTTCTTCAAACACAATG	196bp
U6 F	GAAGATGCCAAGGAAGTGGTAG	
U6 R	GAGCAACACAAATATGGCTTGG	160bp

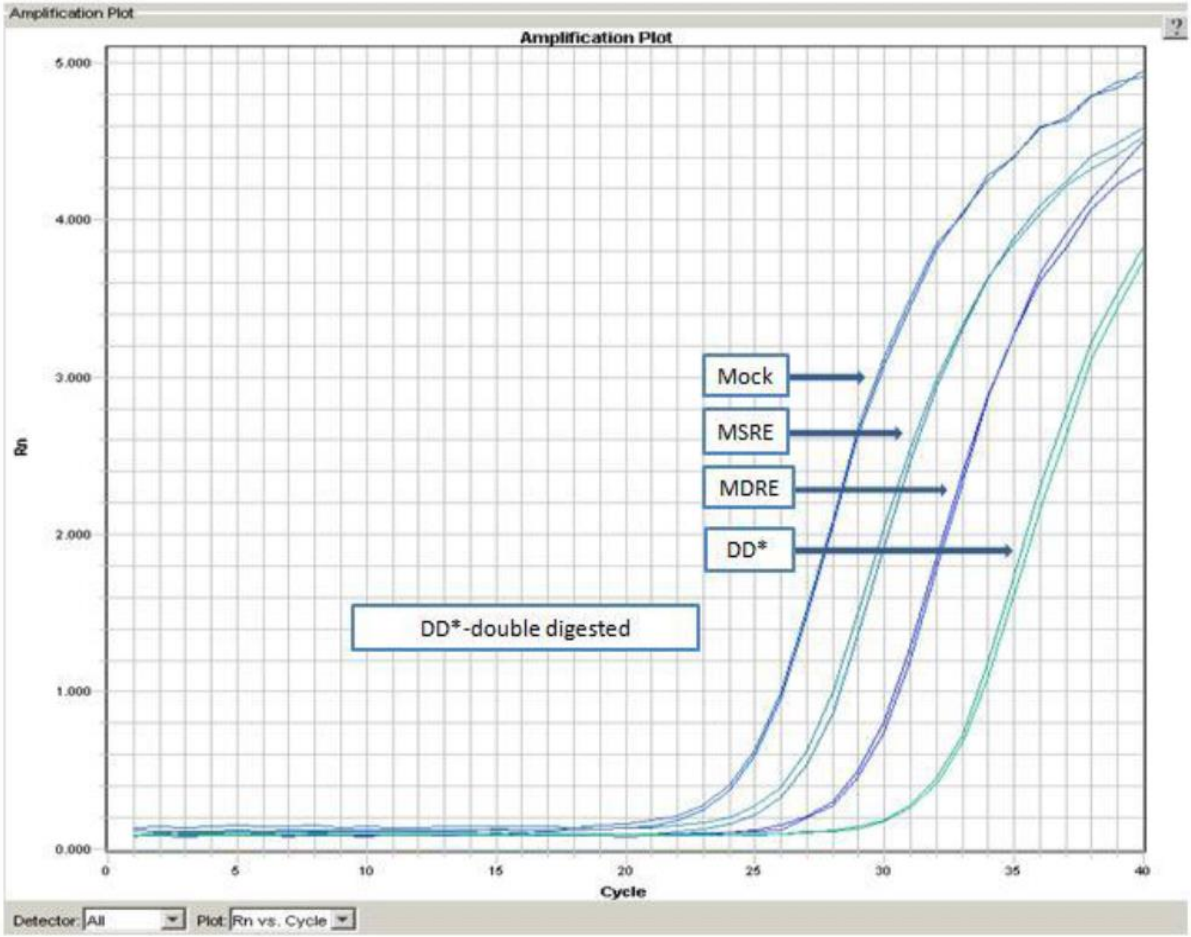

Figure 3.8: A typical qPCR amplification plot of MethylScreen. Mock controls have earlier Ct values compare to double digested samples (DD). The difference in Ct values between the double digest and the mock reactions represents the analytical window of the assay.

3.3 Results

3.3.1 Quantification of *FAST-1* levels in *FAST-1* overexpressing cells

As described in section 3.2.4, twelve stable *FAST-1* overexpressing HeLa cell lines, twelve stable *FAST-1* overexpressing HEK293 cell lines and three stable *FAST-1* overexpressing fibroblast cell lines were developed. To confirm *FAST-1* overexpression in *FAST-1* transfected cells, qRT-PCR measurements were performed as previously described in section 3.2.5, using primers designed to detect human-specific frataxin antisense transcript, *FAST-1*. Total RNA was isolated from stably overexpressing *FAST-1* cells, and treated with DNase I and converted into cDNA. The resultant cDNA was used as a template in qRT-PCR reactions. To quantify *FAST-1* expression the cDNA synthesised by FAST RT primer was diluted 1:5 and used in qRT-PCR reactions; whereas for *FXN* and *HPRT* expression cDNA synthesised by oligo (dT) primer was diluted 1:10 and 1:100 respectively and used as qRT-PCR template. All the Ct values obtained for *FAST-1* and *FXN* were normalised to the endogenous control gene *HPRT*, then calibrated to normal cells before transfection and the mean of control was set to 100%. Analysis of qRT-PCR measurement confirmed a very significant increase of *FAST-1* expression in transfected HeLa cells (396%, $P<0.001$), transfected HEK293 cells (660%, $P<0.05$) and transfected fibroblast cells (332%, $P<0.01$) compared with non-transfected control cells (Figure 3.9).

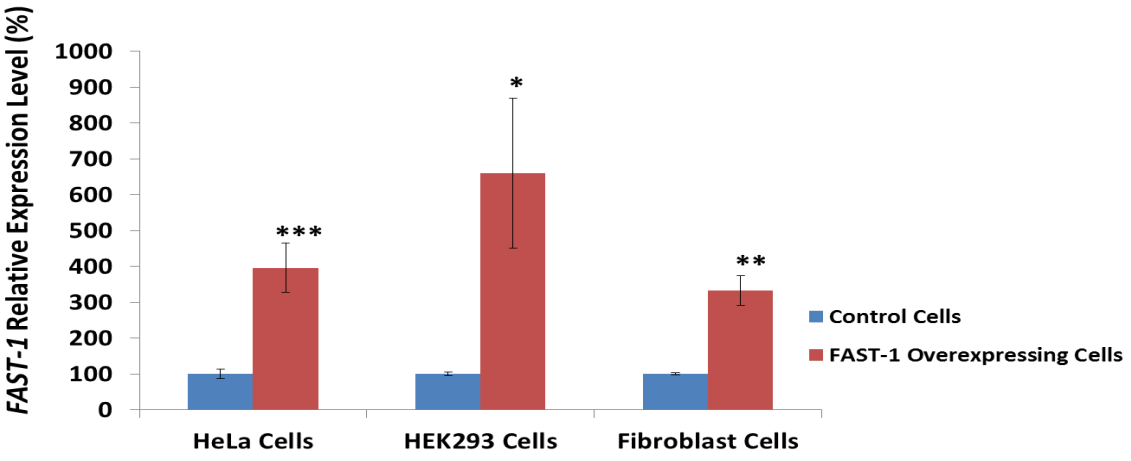

Figure 3.9: Relative *FAST-1* expression in *FAST-1* overexpressing cells. *FAST-1* overexpression was confirmed in *FAST-1* transfected cells compared to non-transfected control cells. Data were normalized to the mean *FAST-1* level of non-transfected control cells taken as 100%. (n=12, *FAST-1* overexpressing HeLa and HEK293 clones), (n=3, *FAST-1* overexpressing fibroblast clones), *P* values were generated by t-test. *p<0.05, **p<0.01, ***p<0.001, error bars=SEM.

3.3.2 Quantification of *FXN* mRNA levels in *FAST-1* overexpressing cells

To assess the effect of *FAST-1* overexpression on *FXN* mRNA expression, qRT-PCR measurements were performed as previously described in section 3.2.5 using primers designed to detect human *FXN* cDNA. QRT-PCR analysis of *FAST-1* overexpressing cells revealed a very significant reduction of *FXN* mRNA in the *FAST-1* overexpressing HeLa cells (45%, *P*<0.05), HEK293 cells (47%, *P*<0.001) and fibroblast cells (42%, *P*<0.001) compared with non-transfected control cells. To examine if the transfection reagents or the transfection process itself has any effects on the *FXN* expression in target cells, *FXN* mRNA levels were measured in empty vector-transfected samples. There was no significant difference detected between the empty vector-transfected and non-transfected control cells (Figure 3.10).

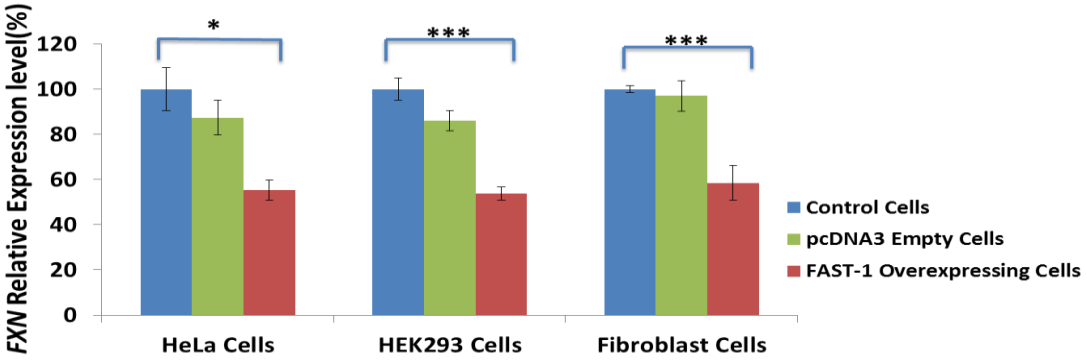

Figure 3.10: Relative *FXN mRNA* expression in *FAST-1* overexpressing cells. *FXN* mRNA expression levels were reduced in *FAST-1* overexpressing cells compared to non-transfected cells. No significant difference detected between the empty vector-transfected and non-transfected control cells. Data were normalized to the mean *FXN* level of non-transfected control cells taken as 100%. (n=12, *FAST-1* overexpressing HeLa and HEK293 clones), (n=3, *FAST-1* overexpressing fibroblast clones), *P* values were generated by t-test. *p<0.05, ***p<0.001, error bars=SEM.

3.3.3 Determination of *FAST-1* copy number in *FAST-1* overexpressing cells

To determine the copy number of *FAST-1* integrated into the genome, ~20 ng samples of genomic DNA from *FAST-1* overexpressing HeLa cell lines were analysed by TaqMan qPCR copy number assay. To assess the reliability of each copy number call, the results were analysed by Applied Biosystems CopyCaller™ Software v.2.0. Eight out of twelve *FAST-1* overexpressing HeLa cell lines had acceptable confidence values and absolute zero score (<2.65) to pass the test. As previously described in section 3.2.6, duplexed TaqMan reactions were applied to accurately quantitate the level of *FAST-1* in *FAST-1* overexpressing cells relative to the endogenous reference gene RNase P. RNase P was known to have two copies in a diploid genome. Non-transfected cells known to have two copies of *FAST-1* were used as the calibrator samples. Results from the TaqMan assay indicated that

there was a significant positive correlation between *FAST-1* copy number and *FAST-1* expression (R =0.7, *P*<0.05) (Figure 3.11).

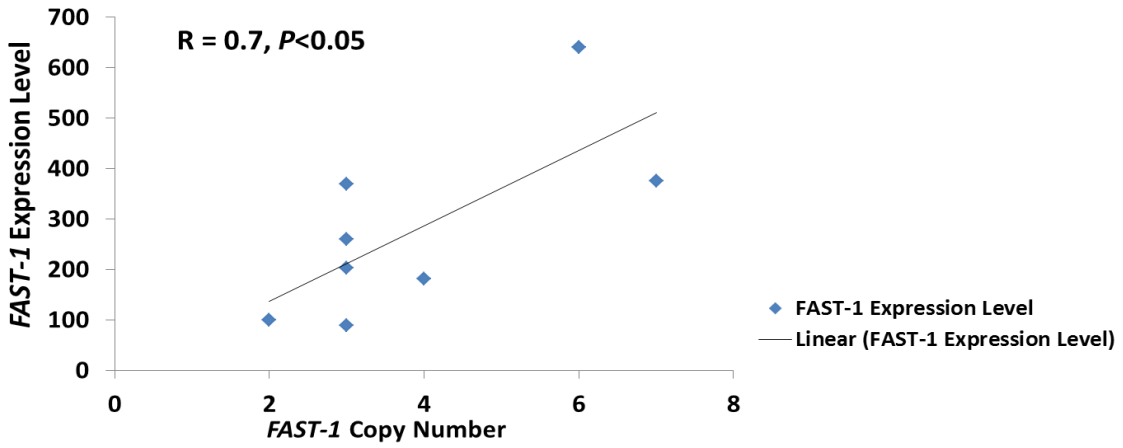

Figure 3.11: TaqMan copy number assay to determine the *FAST-1* copy number in *FAST-1* overexpressing clones. There was a significant positive correlation between *FAST-1* copy number and *FAST-1* expression (R =0.7, *P*<0.05) in *FAST-1* overexpressing HeLa cells. (n=8, *FAST-1* overexpressing HeLa clones).

Scatter plot analysis was conducted between *FXN* and *FAST-1* expression for the *FAST-1* overexpressing HeLa cells, revealing a strong negative correlation between *FAST-1* expression and *FXN* expression (R = -0.58, *P*<0.05) (Figure 3.12).

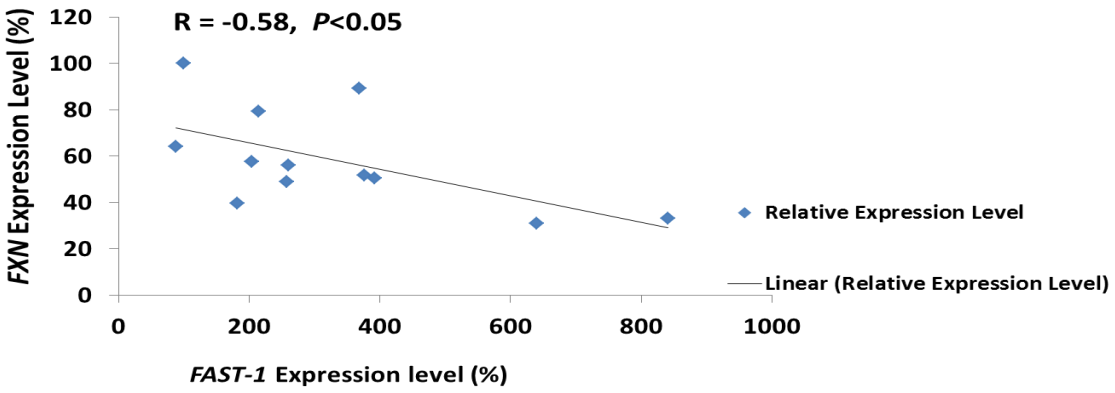

Figure 3.12: Scatter plot of *FXN* mRNA expression level versus *FAST-1* expression in *FAST-1* overexpressing HeLa cells. There was a significant negative correlation between *FAST-1* expression and *FXN* expression (R =-0.58, *P*<0.05) in *FAST-1* overexpressing HeLa cells. (n=12, *FAST-1* overexpressing HeLa clones).

3.3.4 Quantification of FXN protein levels in *FAST-1* overexpressing cells

To determine the levels of human frataxin protein expression in the *FAST-1* overexpressing cells, the Frataxin Protein Quantity Dipstick lateral flow immunoassay assay was performed as described in section 2.4.13. Non-transfected cells were used as the controls in this experiment. All frataxin expression levels obtained from the *FAST-1* overexpressing cells were calibrated to non-transfected control cells and the mean of control was set to 100%. The experiment was performed in triplicates. Results from these experiments demonstrated that frataxin protein expression was significantly decreased by 40% (*P*<0.01) in *FAST-1* overexpressing HeLa cells, 55% (*P*<0.05) in *FAST-1* overexpressing HEK293 cells and 58% (*P*<0.001) in *FAST-1* overexpressing fibroblast cells compared with non-transfected control cells (Figure 3.13).

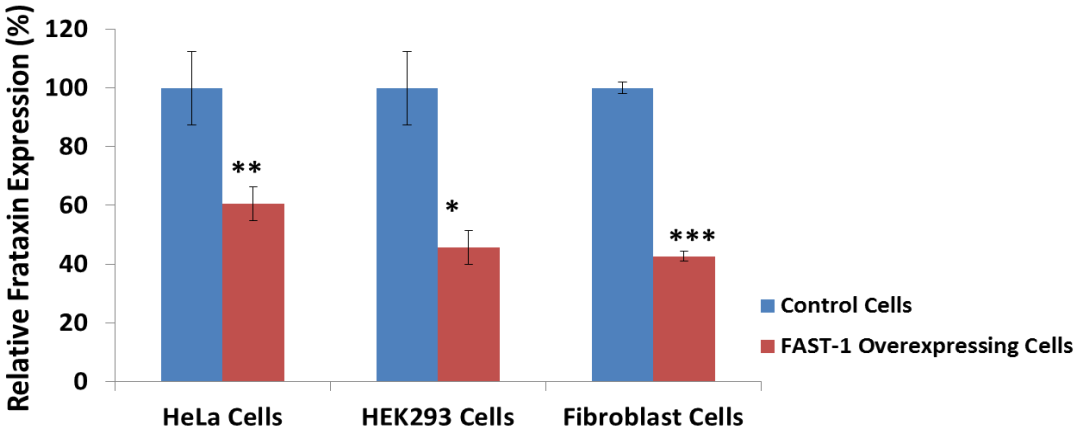

Figure 3.13: Relative frataxin protein expression in *FAST-1* over expressing cells. Frataxin protein expression levels were reduced in *FAST-1* overexpressing cells compared to non-transfected cells. Data were normalized to the mean *FXN* level of non-transfected control cells taken as 100%. (n=12, *FAST-1* overexpressing HeLa and HEK293 clones), (n=3, *FAST-1* overexpressing fibroblast clones), *P* values were generated by t-test. *p<0.05 ,**p<0.01, ***p<0.001, error bars=SEM.

3.3.5 Investigation of histone modifications in *FAST-1* overexpressing HeLa cells

Measuring frataxin mRNA and protein expression before and after *FAST-1* overexpression in 3 different cell lines, HeLa, HEK293 and fibroblast, demonstrated that *FAST-1* overexpression exerted a negative impact on the *FXN* gene expression. To investigate potential explanation for the effects of *FAST-1* overexpression on *FXN* gene expression, specific histone modifications were studied in *FAST-1* overexpressing HeLa cells. Three different *FAST-1* overexpressing HeLa colonies, which showed the lowest level of *FXN* expression, empty vector-transfected cells and a non-transfected HeLa cells were utilized in this experiment. Acetylated and trimethylated histone H3K9 modifications were investigated by ChIP analysis at the

FXN 5′UTR region. The ChIP procedure was performed as previously described in section 2.4.14. Real-time PCR was performed in triplicate in a final volume of 10µl containing 12.5pmol of each of the respective forward and reverse primers (see Table 2.6 for primers used). The ChIP results were normalized to input and -Ab values were taken into account. ChIP assay results showed significantly decreased acetylation of H3K9 at the *FXN* 5′UTR region in *FAST-1* overexpressing HeLa cells compared to non-transfected control cells. There were no statistically significant differences in H3K9 acetylation levels between the empty vector-transfected and non-transfected control cells (Figure 3.14).

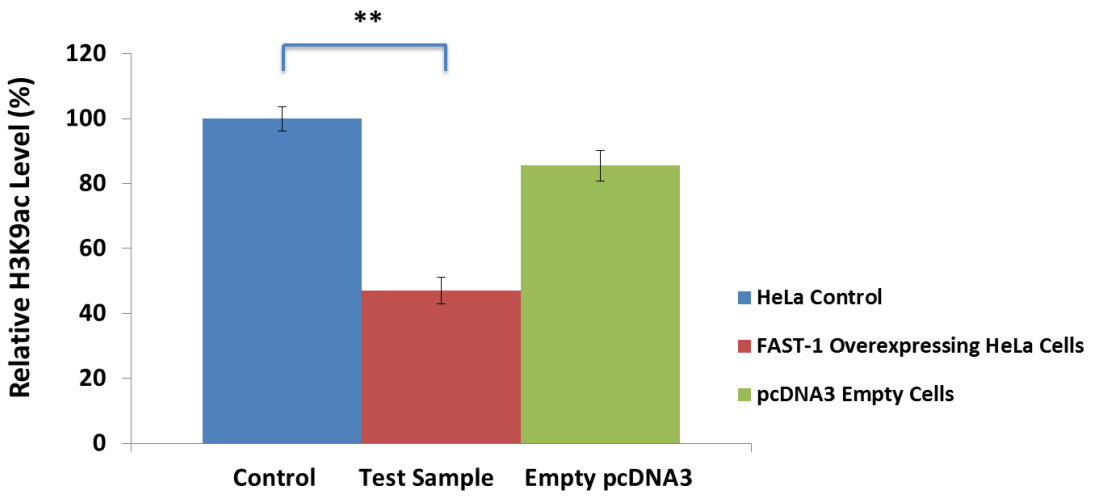

Figure 3.14: ChIP with anti-H3K9ac in *FAST-1* overexpressing HeLa cells. Results showed a significant decrease of H3K9 acetylation at the 5′UTR of the *FXN* gene in *FAST-1* overexpressing HeLa cells. No significant differences were detected between empty transfected-vector and non-transfected Hela control cells. The data were from three independent chromatin preparations, with each experiment done in triplicate. (n=3, *FAST-1* overexpressing HeLa clones), *P* values were generated by t-test. **p<0.01, error bars=SEM.

In addition, significant enrichment of H3K9me3 was detected at the 5′UTR of the *FXN* gene (Figure 3.15). H3K9me3 enrichment was determined by subtracting the respective minus antibody samples from the H3K9me3-precipitated samples, followed by normalisation to the input sample.

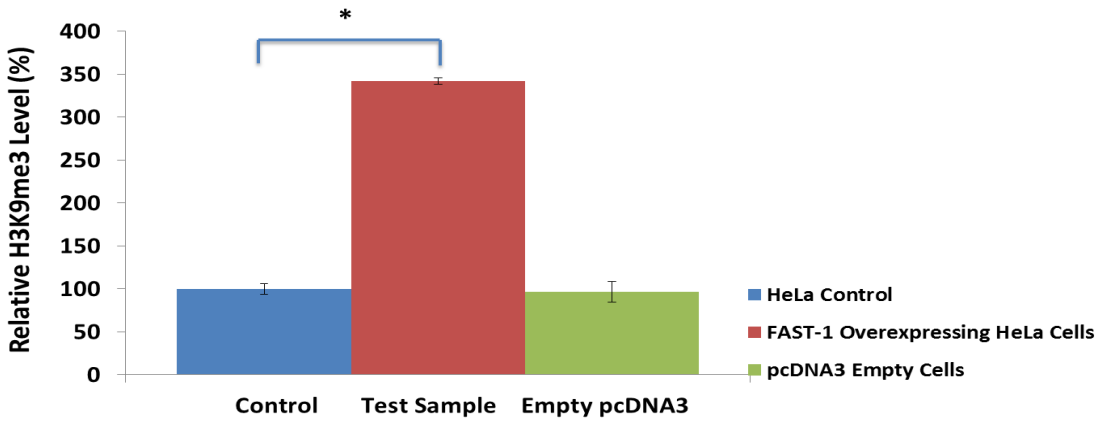

Figure 3.15: ChIP with anti-H3K9me3 in *FAST-1* overexpressing HeLa cells. Results showed enrichment of H3K9me3 at the 5′UTR of the *FXN* gene in *FAST-1* overexpressing HeLa cells. No significant differences were detected between empty transfected-vector and non-transfected Hela control cells. The data were from three independent chromatin preparations, with each experiment done in triplicate. (n=3, *FAST-1* overexpressing HeLa clones), *P* value was generated by t-test. *p<0.5, error bars=SEM.

3.3.6 Determination of CTCF levels in *FAST-1* overexpressing HeLa cells

It has previously been reported that CTCF depletion results in increased *FAST-1* transcription. To elucidate the relationship between CTCF and *FAST-1* in *FAST-1* overexpressing HeLa cells, ChIP experiments were performed with an anti-CTCF antibody. CTCF ChIP analysis revealed that CTCF occupancy was decreased sharply at the 5′UTR of the *FXN* gene. In addition, comparison of CTCF occupancy in empty transfected-vector and non-transfected HeLa control cells did not show a similar reduction (Figure 3.16).

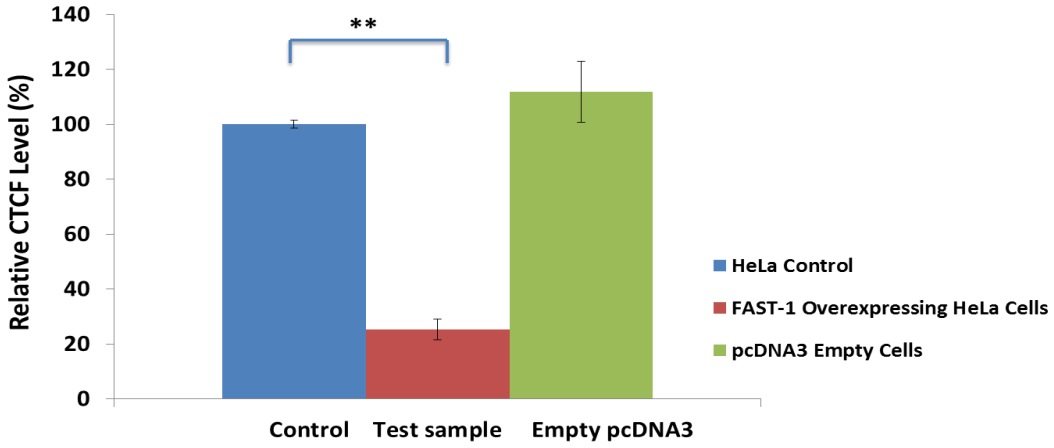

Figure 3.16: ChIP with anti-CTCF in *FAST-1* overexpressing HeLa cells. Results showed significantly less occupancy of CTCF in *FAST-1* overexpressing HeLa cells. No significant differences were detected between empty transfected-vector and non-transfected Hela control cells. The data were from three independent chromatin preparations, with each experiment done in triplicate. (n=3, *FAST-1* overexpressing HeLa clones), *P* value was generated by t-test. **p<0.01, error bars=SEM.

3.3.7 Investigation of DNA methylation in *FAST-1* overexpressing fibroblast cells

It has been reported that CpG U6 and CpG U11 show elevated levels of DNA methylation in FRDA patients compared to controls (Greene *et al.*, 2007; Al-Mahdawi *et al.*, 2013). Therefore, MethylScreen technology was used to determine the level of DNA methylation at CpG U6 and CpG U11 in two *FAST-1* overexpressing fibroblasts cell lines (FAST-1B, FAST-1C), together with FRDA fibroblasts and unaffected control fibroblasts. Results from MethylScreen assays revealed a non-significant small increase in DNA methylation at CpG site U6, with DM values increasing from 14% in unaffected control fibroblasts to 18% and 23% in FAST-1 B and FAST-1 C, respectively (Figure 3.17).

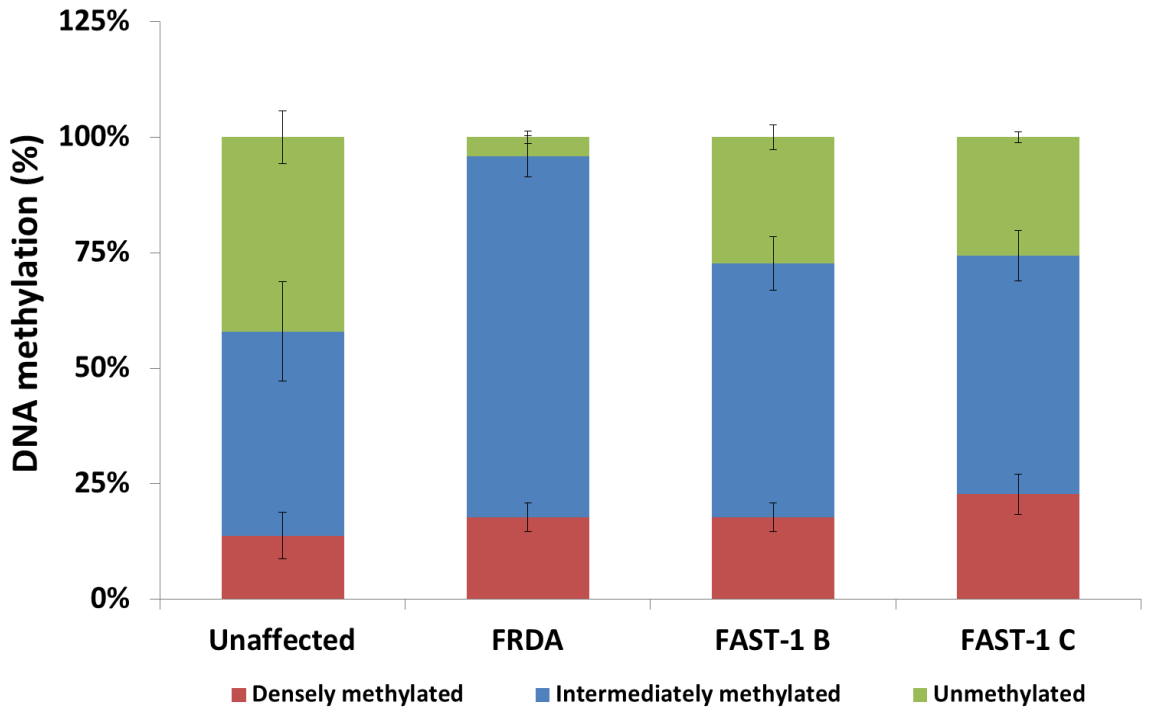

Figure 3.17: DNA methylation levels in *FAST-1* overexpressing fibroblasts analysed by MethylScreen technique. MethylScreen analysis of CpG site U6 in the *FXN* upstream GAA repeat region of *FAST-1* overexpressing cells. UM=unmethylated, IM=intermediately methylated, DM=densely methylated. The experiment was repeated twice for each cell type and each reaction was carried out in duplicate. Error bars=SEM.

At CpG site U11, the proportion of densely methylated templates in *FAST-1* overexpressing fibroblasts significantly increased from 8% in unaffected controls to 79% ($P<0.01$) and 76% ($P<0.001$) in FAST-1 B and FAST-1 C, respectively, similar to the levels found in FRDA fibroblasts (Figure 3.18).

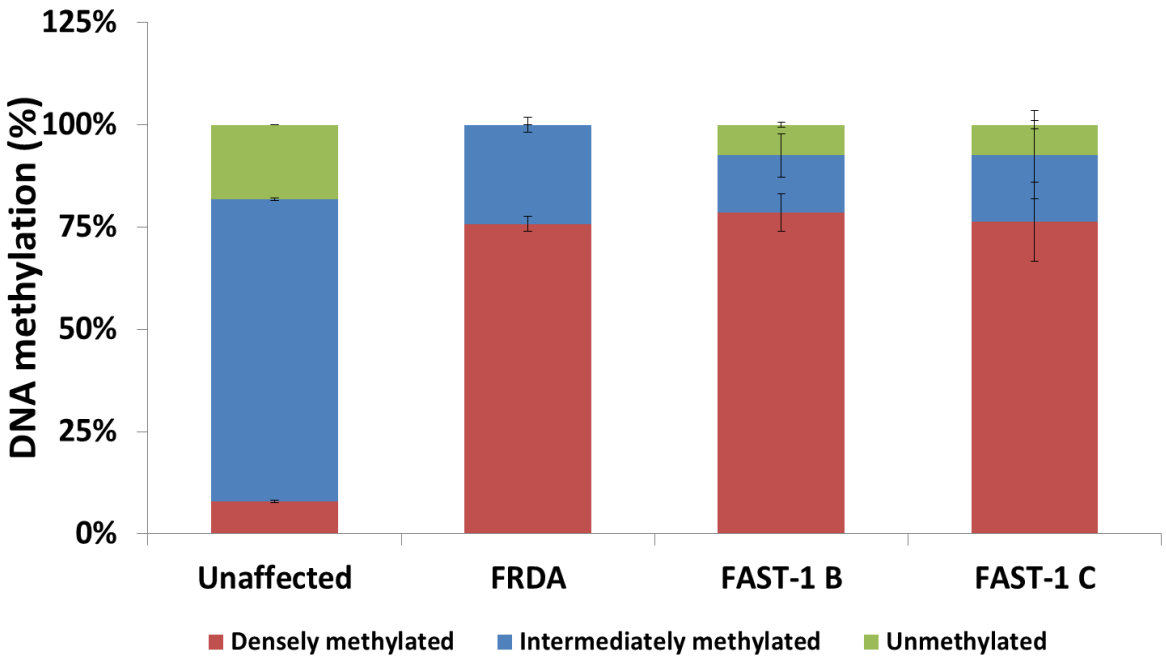

Figure 3.18: DNA methylation levels in *FAST-1* overexpressing fibroblasts analysed by MethylScreen technique. MethylScreen analysis of CpG site U11 in the *FXN* upstream GAA repeat region of *FAST-1* overexpressing fibroblasts. UM=unmethylated, IM=intermediately methylated, DM=densely methylated. The experiment was repeated twice for each cell type and each reaction was carried out in duplicate. Error bars=SEM.

3.4 Discussion

It is well established that the GAA repeat expansion results in decreased transcription of the *FXN* gene and the protein it codes for. Despite extensive investigation, the exact FRDA disease mechanisms remain unknown. However, it has been reported that the *FAST-1* expression levels were two-fold higher in FRDA patient-derived fibroblasts. Therefore, to investigate potential *FAST-1* effects on *FXN* expression, we stably overexpressed *FAST-1* in HeLa, HEK293 and fibroblast cells. QRT-PCR analysis confirmed *FAST-1* overexpression in all three cell lines. Analysis of frataxin mRNA expression revealed a significant reduction in the levels of the *FXN* transcripts in *FAST-1* overexpressing cells. Measuring frataxin protein expression levels by lateral flow immunoassay demonstrated that *FAST-1* overexpressing cells also express significantly less frataxin protein compared to control cells. In addition, no significant *FXN* gene and protein expression changes were detected in empty vector control cells. These results indicate that *FAST-1* can act *in trans* in a similar manner to the *cis*-acting *FAST-1* overexpression that has previously been identified in FRDA fibroblasts. Additionally, we identified a significant positive correlation between *FAST-1* copy number and *FAST-1* expression level ($R=0.7$, $P<0.05$) and a negative correlation between *FAST-1* copy number and *FXN* expression level ($R=-0.58$, $P<0.05$). This suggests a concentration-dependent molecular mechanism of action for *FAST-1*, rather than a potential threshold effect. To investigate the potential mechanisms for the *FAST-1* overexpression effects, two major epigenetic modifications, DNA methylation at the CpG sites U6 and U11 and histone modifications at the promoter region of the *FXN* gene in *FAST-1* overexpressing cells, were studied. *FAST-1* overexpressing cells showed reduced *FXN* expression levels and exhibited similar epigenetic changes to those observed in FRDA cells.

FAST-1 overexpressing HeLa cells showed reduced occupancy of CTCF, a decreased level of H3K9ac and an increased level of H3K9me3 at the 5′UTR of the *FXN* gene. Several studies suggest that antisense RNAs might be key regulators in gene expression. There are various potential ways by which antisense RNA involvement could cause heterochromatin formation and epigenetic silencing of the partner gene. Studies suggest that aberrantly expressed antisense RNA mostly act at the RNA level to epigenetically silent their corresponding sense genes (Whitehead *et al.*, 2009; Kanduri, 2011). Antisense RNAs can act as scaffold or directly recruit chromatin modifying proteins to chromatin. Antisense RNAs such as *Xist*, *Kcnq1ot1*, which is expressed from intron 10 of the imprinted gene *Kcnq1*, and *HOTAIR* can recruit DNA and chromatin modifying proteins, such as the PRC2 complex, which has H3K27 methyltransferase activity and catalyses heterochromatin formation and gene repression (Zhao *et al.*, 2008; Khalil *et al.*, 2009; Marchese and Huarte, 2014). The enrichment of H3K9me3, H3K27me3 and HP1 at the 5′UTR of the *FXN* gene have previously been observed and we also detected H3K9me3 enrichment and reduced H3K9ac and frataxin expression following *FAST-1* overexpression in non-FRDA HeLa cells (Al-Mahdawi *et al.*, 2008; De Biase *et al.*, 2009; Kim *et al.*, 2011). Therefore, it is plausible to speculate that *FAST-1* may induce RNA-dependent histone modification and promote the heterochromatinisation of the *FXN* gene.

In addition, it has been reported that the *FAST-1* transcript may form a double-stranded RNA with the *FXN* sense transcript and the resulting RNA duplex may interact with argonaute proteins, Ago1 and Ago2, which are key players in RNA-induced transcriptional gene silencing (RITS). Strikingly, knocking down Ago1 and Ago2 in FRDA cells reversed the repressive chromatin at the 5′UTR of the *FXN* gene (Bidichandani *et al.*, 2011). Moreover, some studies suggest that antisense

transcripts can recruit Ago proteins during gene silencing (Chu *et al.*, 2010; Schwartz *et al.*, 2014). Based on these findings and our own observations, it can be suggested that *FAST-1* overexpression may trigger an RNA interference (RNAi) pathway, causing the establishment of heterochromatin in the 5′UTR of the *FXN* gene. However, some studies argue against a role for RNAi system in FRDA, as the authors reported that they could not detect small interfering RNAs (siRNAs) matching the *FXN* locus (Punga and Bühler, 2010). Presently, none of the above possibilities can be ruled out and further studies are required. The effects of stably transfected *FAST-1* overexpression on CTCF and H3K9ac and H3K9me3 at the *FXN* 5′UTR locus indicate a potential causative role for *FAST-1* in FRDA, as opposed to the increased *FAST-1* expression that is observed in FRDA merely being the passive result of loss of CTCF binding at the *FXN* 5′UTR. Thus, it is plausible that increased *FAST-1* RNA molecules either directly displace CTCF from the *FXN* 5′UTR, or trigger heterochromatin formation in the *FXN* 5′UTR, which consequently results in CTCF depletion and reduced *FXN* gene expression.

Chapter 4: Knocking down *FAST-1* expression in FRDA fibroblasts

4.1 Introduction

Natural antisense transcripts (NATs) are prevalent in the human genome. Although they generally have very low levels of expression, they have been shown to be involved in epigenetic silencing of their associated sense gene. Therefore, inhibition of NATs may lead to upregulation of the sense genes. It has been reported that suppression of antisense transcripts to *p21 (Morris et al., 2008)*, *BDNF* (Brain-Derived Neurotrophic Factor), *GDNF* (Glial-derived Neurotrophic Factor), *EphB2* (Ephrin receptor B2) (Modarresi *et al.*, 2012b) and Oct-4 (octamer-binding transcription factor 4) (Hawkins and Morris, 2010) upregulates the corresponding mRNAs. We hypothesized that knocking down *FAST-1* as an epigenetic modulator may increase *FXN* expression, therefore, *FAST-1* was stably knocked down by short hairpin RNA (shRNA) in FRDA fibroblast cells and the subsequent consequences on *FXN* expression were investigated.

4.2 Material and methods

4.2.1 Transduction of fibroblast cells

Fibroblasts from healthy subjects (GM07492) were obtained from Coriell Cell Repository. The FRDA fibroblasts (FA1) were obtained from Dr Alexandra Durr (ICM University Hospital Pitié-Salpêtrière, Paris, France). Expanded GAA alleles for the FRDA patient were 416 & 59 triplets (FA1). To knock down *FAST-1*, Sigma custom cloning MISSION shRNA team designed a short hairpin RNA, cloned it into a pLKO.1-puro-CMV-tGFP vector and provided us with ready to use lentiviral particles for transduction. To ensure that the introduction of lentiviral particles and shRNA did not affect gene expression and that the observed outcomes were due to *FAST-1* silencing, a scrambled shRNA (pLKO.1-puro Non-Target shRNA Control Transduction Particles) was used. Human FRDA cells and non-FRDA fibroblast cells were grown in 6-well plates to reach 80% confluency. On the day of transduction, cells were rinsed with warm PBS and replaced with 500µl of fresh medium. To increase the lentiviral transduction efficiency, polybrene (8µg/ml) was mixed with *FAST-1* and scrambled LV particles and added to the cells on a plate. To increase the contact surface between viral particles and cells, a minimum volume (500µl) of medium was added. Cells were incubated at 37°C and 5% CO_2 for 6 hours and then the supernatant was replaced with 1ml fresh medium and incubated at 37°C and 5% CO_2 for 72 hours to allow time for the expression of the reporter and puromycin resistance genes. After 72 hours, the selective antibiotic, puromycin, was added to the media at a final concentration of 0.2µg/µl. The selection media was replaced every 3 days and the final cell harvest for RNA analysis was performed at 2 weeks post-transduction. A lower concentration of puromycin was continued throughout the duration of the cell culture.

4.2.2 Determination of MOI for the transduction of the FRDA fibroblasts

$5x10^4$ human FRDA fibroblasts that were growing exponentially were seeded into 6-well plates and incubated for 18-20 hours at 37°C and 5% CO_2. Triplicate wells of a 6-well plate were used for each range of multiplicity of infection (MOI) and a control. Cells were washed with PBS and 500µl of fresh medium containing 8µg/ml of polybrene was added to each well, then an appropriate volume of the LV *FAST-1*-GFP ($5.7X10^7$ TU/ml) (25µl-100µl) was added to each well to achieve a MOI of 5, 10, 15 and 20. Cells were incubated at 37°C and 5% CO_2 with viral supernatant for 6 hours and then the supernatant was replaced with 1ml fresh medium and incubated at 37°C and 5% CO_2 for 72 hours to allow time for the expression of the reporter and puromycin resistance genes. After 72 hours, puromycin antibiotic was added to the media at a final concentration of 0.2µg/µl.

Following transduction of human FRDA fibroblasts, live cells were counted using trypan blue (Fisher Scientific). ImagestreamX was used to measure the number of the GFP-positive cells. In this method, cells were trypsinised and pelleted by centrifugation at 5000xg for 5 minutes at 4°C. The pellet was washed with ice-cold PBS and fixed with 4% paraformaldehyde for 3 minutes on ice. Fixed cells were washed with PBS and the pellet was resuspended in 100µl of Accumax and 1µl Draq 5 was added to the cells. Cells were put through imaging flow cytometry using the ImagestreamX system (Amnis Inc.). Images of approximately 10,000 cells were captured on channel 1 for bright field (BF); on channel 2 for GFP representing the green GFP expression; and on channel 5 for the far-red DRAQ5™ staining of the nuclear region of each cell. Following excitation with a 488nm laser at a power setting of 70 mW, all images were captured using a 40xobjective (Figure 4.1).

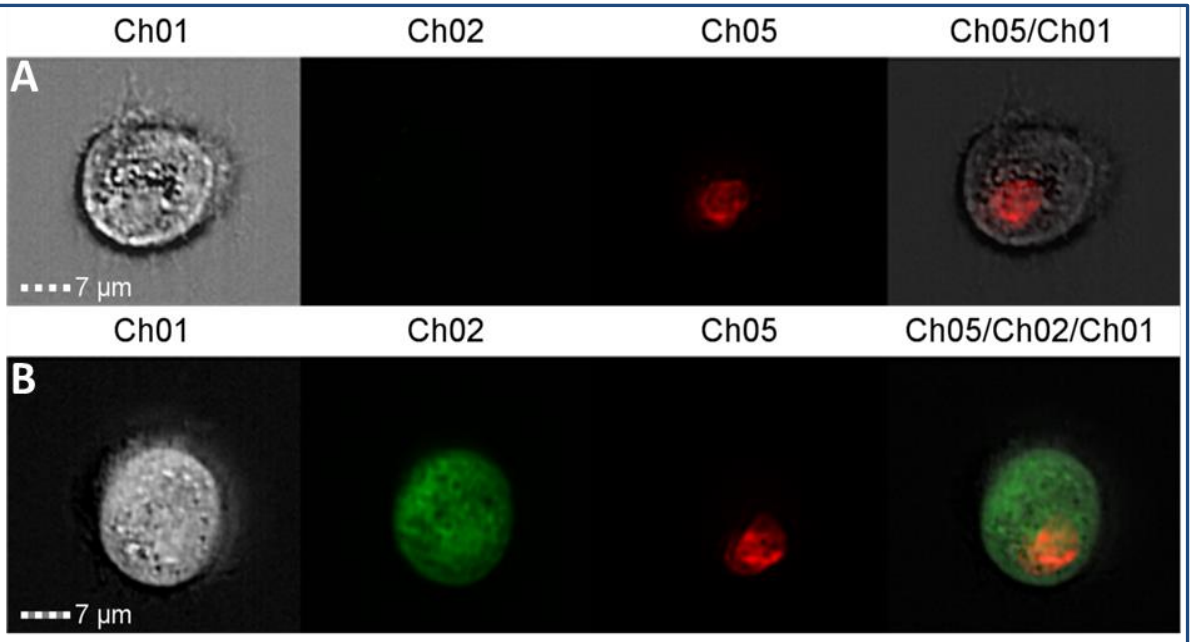

Figure 4.1: Flow cytometric imaging of FRDA fibroblasts. Channel 1 (ch01) was used for bright field ; channel 2 (ch02) for GFP representing the green GFP expression; and channel 5 (ch05) for the DRAQ5™ nuclear far-red DNA stain. **A)** A GFP-negative FRDA fibroblast cell. **B)** A GFP-positive FRDA fibroblast. FRDA fibroblasts were infected with FAST-1-GFP LV at MOI 10. Green fluorescing fibroblasts were visualised 72 hours post transduction using live cells under light microscopy and by image flow cytometry using an ImageStreamX system (40x magnification).

4.2.3 Aconitase activity

The Aconitase Assay Kit (Cayman, 705502) was used to assess the aconitase activity of transduced FRDA fibroblasts. Cells were washed with cold PBS, scraped off the flask in ice cold PBS and centrifuged at 800xg for 10 minutes at 4°C. Pellets were resuspended in 50µl of cell lytic buffer and incubated on ice for 15 minutes. Lysates were centrifuged at top speed for 10 minutes at 4°C and used immediately. Samples were diluted in 1Xassay buffer (provided with the kit) to make a final

concentration of 0.5µg/µl in 200ul volume. 50µl of fresh cell lysate (0.5µg/µl), NADP+, isocitrate dehydrogenase and diluted Substrate Solution (all provided with kit) were added to each well of a 96-well plate and incubated at 37°C for 15 minutes in the dark. The absorbance at 340nm was measured once every minute with 5 seconds mix time for 15 reads at 37°C using a Biorad xMark plate reader. The aconitase activity was determined from the slope of a graph by plotting the absorbance value over time. Since total cell lysates were used for this assay, aconitase activities of cells were then normalized to citrate synthase activities, which were determined using a citrate synthase assay kit (SIGMA, CS0720) according to the manufacturer's instructions.

4.3 Results

4.3.1 *FAST-1* knockdown validation

To achieve a stable *FAST-1* knockdown, FRDA and normal fibroblast cell lines were transduced with lentiviral (LV) vectors expressing shRNAs targeting *FAST-1* (pLKO.1-puro-CMV-tGFP). Several precautions were taken in order to control off-target effects and specificity, as follows: (i) pLKO.1-puro Non-Target shRNA was used as the negative control; (ii) The normal fibroblast cell line was also transduced with LV *FAST-1* to assess the specificity of any *FAST-1* knockdown effect on FRDA cells.

After 10 days of puromycin selection, total RNA was extracted from the transduced cells and *FAST-1* cDNA was synthesized using cloned AMV first-strand cDNA synthesis kit (Invitrogen) as previously described (section 2.4.5). The efficacy of shRNA knockdown was assessed by qRT-PCR. Analysis of *FAST-1* expression levels in the *FAST-1* knockdown control and FRDA fibroblast cell lines revealed that *FAST-1* levels decreased to 63% ($P<0.01$) and 43% ($P<0.01$) respectively (Figure 4.2).

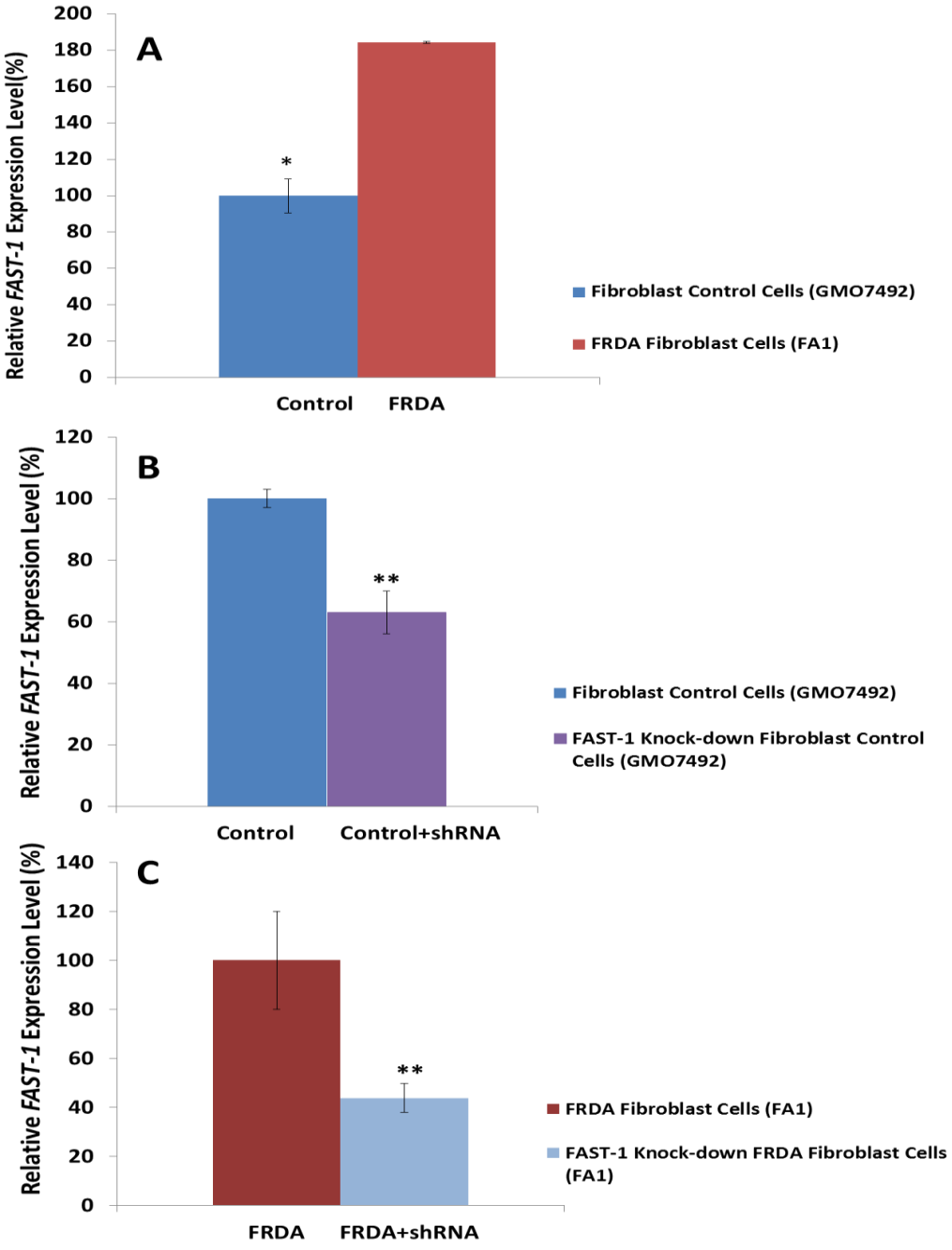

Figure 4.2: QRT-PCR analysis of *FAST-1* expression levels in FRDA fibroblasts and *FAST-1* knockdown fibroblast cells. Real-time RT-PCR showed approximately twice as much *FAST-1* levels in FRDA cell lines versus non-FRDA controls. The efficacy of shRNA knockdown was assessed by qRT-PCR, demonstrating a significant reduction in the levels of *FAST-1* in **B)** control fibroblast and **C)** *FAST-1* knockdown FRDA fibroblast cells. *P* values were generated by t-test. . *p<0.05, **p<0.01, error bars=SEM.

4.3.2 Measurement of *FXN* expression levels following transduction of FRDA fibroblast cells with LV *FAST-1*

After confirming efficient *FAST-1* knockdown in FRDA fibroblast cells, *FXN* mRNA expression was measured in *FAST-1* knockdown cells. Results from qRT-PCR analysis revealed that the levels of *FXN* expression in *FAST-1* knockdown FRDA cells had increased by 1.5-fold ($P<0.05$) when compared to FRDA fibroblast cells. In contrast, *FXN* expression in normal fibroblast cells treated with LV *FAST-1* showed a non-statistically significant 0.8-fold decrease in the *FXN* expression. *FXN* expression remained unchanged in FRDA fibroblast cells treated with scrambled control shRNA (Figure 4.3).

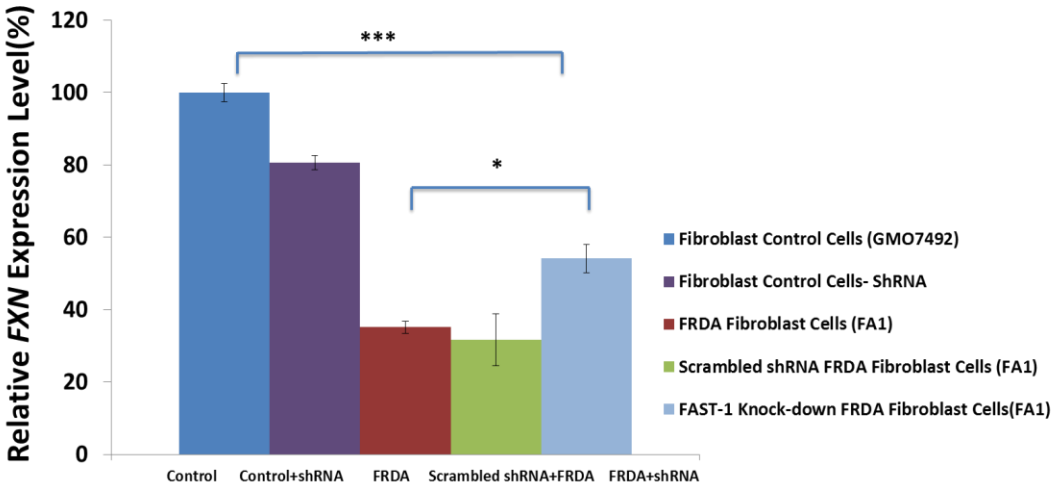

Figure 4.3: QRT-PCR analysis of *FXN* expression levels in control and *FAST-1* knockdown cell lines. Knocking down *FAST-1* in FRDA fibroblast cells increased *FXN* expression, but not to the level of control cells, however it did not significantly affect *FXN* expression in control cells. *FXN* expression in FRDA cells treated with scrambled shRNA remained unchanged. *P* values were generated by t-test. *p<0.05, ***p<0.001, error bars=SEM.

4.3.3 Restoring aconitase activity following knockdown of *FAST-1*

As a biomarker of frataxin function within cells, aconitase activity was measured in *FAST-1* knockdown fibroblast cells 1 month after treatment with LV *FAST-1* using the Aconitase Assay Kit (Cayman) as previously described in section 4.2.3. Normal FRDA fibroblast cells (GMO7492) and FRDA cells treated with scrambled shRNA were used as the controls. All data was calibrated to the mean of the aconitase activity in normal fibroblast cells, which was set to 100%. The aconitase activity in FRDA cells was determined to be 49% (P<0.05) compared to normal control fibroblast cell levels. However, after knocking down *FAST-1* in FRDA fibroblast cells, aconitase activity significantly increased to a value of 195% (P<0.01). No significant difference was detected in FRDA fibroblast cells treated with scrambled shRNA (Figure 4.4).

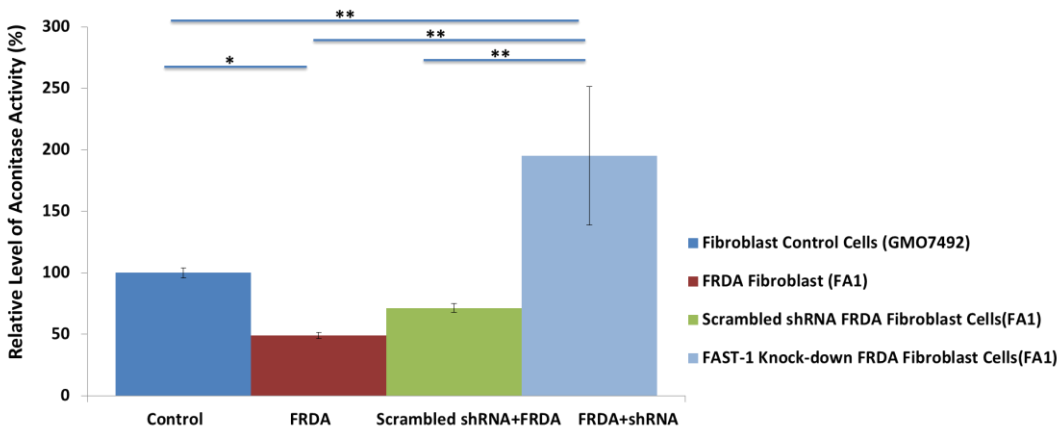

Figure 4.4: Aconitase activity levels in control and *FAST-1* knockdown cell lines. Aconitase activity in normal fibroblast cells was considered as 100%. Aconitase activity was decreased in FRDA fibroblast cells, however, after knocking down *FAST-1*, aconitase activity significantly increased. No significant difference was detected in FRDA fibroblast cells treated with scrambled shRNA. *P* values were generated by t-test. *p<0.05, **p<0.01, error bars=SEM.

4.4 Discussion

Our understanding of antisense transcript biological functions and roles in gene expression regulation is not yet complete. Different high-throughput studies that have investigated sense and antisense transcription correlations have come to differing conclusions. Some studies suggest inversely correlated expression patterns between sense and antisense transcripts; while other studies report only weak inverse correlations (Alcid and Tsukiyama, 2014; Murray *et al.*, 2015; Huber *et al.*, 2016). Furthermore, a recent study inhibited antisense transcription of 162 genes in *S. cerevisiae* and investigated its impact on gene expression. They observed antisense-dependent gene regulation in one quarter of the studied genes. They also found that antisense regulation was correlated with antisense transcription overlapping the sense TSS (Huber *et al.*, 2016). Recently four highly conserved chromatin remodelling complexes were identified in *S. cerevisiae*, which directly repress antisense RNAs. Repression of these antisense RNAs are crucial for normal levels of mRNA transcription and their derepression leads to decreased sense mRNA levels (Alcid and Tsukiyama, 2014). With regard to FRDA, *FAST-1* transcription has been shown to overlap with the *FXN* TSS and have an inverse correlation with the *FXN* gene expression levels (De Biase *et al.*, 2009; Sandi and Pook, 2015). As described in chapter 3, overexpression of *FAST-1* in non-FRDA cells was shown to result in decreased *FXN* expression, suggesting that *FAST-1* might be involved in the modulation of the *FXN* epigenetic status, inducing heterochromatin formation and *FXN* gene silencing. The next step in answering our research question ‘does *FAST-1* have a role in the molecular mechanisms of FRDA disease?’ was to knock down *FAST-1* in FRDA fibroblast cells. To target *FAST-1* transcripts, we used shRNA lentiviral particles to transfect FRDA fibroblasts.

Knocking down *FAST-1* in FRDA fibroblast cells resulted in a 57% reduction in the level of *FAST-1* transcript expression and it was accompanied by 1.5-fold increase in the *FXN* transcript expression level. Not surprisingly, *FXN* expression did not show complete rescue to the same level as control fibroblast cells, because the expanded GAA repeat was still present in the cells. Indeed, targeting antisense RNAs has been shown to upregulate the expression of sense transcripts. Antisense non-coding RNA in the INK4 locus (ANRIL) is an antisense transcript whose increased expression level is associated with $p14^{ARF}$, $p15^{INK4b}$ and $p16^{INK4a}$ silencing via direct interaction with both polycomb repressive comlex-1 and -2. Knocking down ANRIL leads to a significant increase in $p14^{ARF}$, $p15^{INK4b}$ and $p16^{INK4a}$ transcripts (Yap *et al.*, 2010; Li and Chen, 2013; Wan *et al.*, 2013). Knocking down of antisense transcripts to *p21*, *BDNF*, *GDNF*, *EphB2* (Modarresi *et al.*, 2012b) and Oct-4 (Hawkins and Morris, 2010) upregulates the corresponding mRNAs. Suppression of the *p21* antisense transcript was followed by increased *p21* sense transcript and loss of the suppressive H3K27me3 mark at the *p21* promoter (Morris *et al.*, 2008). Several studies suggest antisense RNAs can recruit PCR2 complex and direct gene silencing (Zhao *et al.*, 2010). Thus, it can be speculated that inhibiting *FAST-1* transcript may reduce the heterochromatin repression and allow access to transcription factors to increase the *FXN* transcript expression.

Aconitase enzyme activity is considered to be a good indicator of frataxin function within cells, such that low levels of mitochondrial frataxin are associated with reduced activity of the Fe-S containing enzyme aconitase (Bulteau *et al.*, 2004). A recent study demonstrated that *FXN* gene delivery to FRDA cells resulted in long-term overexpression of *FXN* mRNA and frataxin protein. The increased frataxin expression in FRDA cells restored aconitase enzyme activity to near normal levels

(Khonsari *et al.*, 2016). Because our results of the *FAST-1* knockdown study demonstrated an increase in the *FXN* expression level, we decided to investigate the aconitase activity in the *FAST-1* knockdown FRDA cells. We showed that knocking down *FAST-1* in FRDA cells resulted in a significant increase (from 49% to 195%) in aconitase activity. In the future, investigating the histone modification profiles and CTCF occupancy at 5′UTR of the *FXN* locus in *FAST-1* knockdown cells may further help us to elucidate the role of *FAST-1* in FRDA.

Chapter 5: Quantification of *FAST-1* expression in the YG8LR FRDA mouse model

5.1 Introduction

To understand the exact molecular mechanisms of FRDA disease and to test potential therapies, developing mouse models is considered essential. Pook and colleagues have previously generated several human *FXN* YAC transgenic mouse models: Y47R, containing normal-sized GAA repeats (9), and YG8R and YG22R containing expanded GAA repeats (90+120 and 120, respectively). A novel line of mice, designed YG8sR and containing a single *FXN* transgenic locus with 120 GAA repeats has subsequently been derived from YG8R mice. *FAST-1* expression was previously investigated in fibroblasts and neural stem cells (NSCs) derived from these YAC transgenic mouse models (Y47R, YG8sR, YG8R and YG22R). QRT-PCR analysis of fibroblast cells detected increased levels of *FAST-1* in YG8R and YG22R mouse models compared to the Y47R-derived fibroblast cells. Investigating the *FAST-1* expression pattern in NSCs showed a tendency to decrease in the YG8R and YG8sR NSCs cells, indicating that the *FAST-1* expression may be cell type dependent. The analysis of *FAST-1* expression in tissue samples derived from YG8R and YG8sR mice demonstrated that *FAST-1* increased only in the brain and cerebellum tissues of the YG8R mouse model. Although YG8sR mice were derived from a founder YG8R mouse, they showed a different *FAST-1* gene expression pattern from YG8R mice. *FAST-1* expression was significantly decreased in selected YG8sR cells and tissues samples (Sandi and Pook, 2015). In this chapter, *FAST-1* expression was investigated in the brain, cerebellum and heart tissues of a new YG8sR-derived mouse model, designated YG8LR, containing approximately 410 GAA repeats. Histone modifications at the *FXN* transgene were also studied using antibodies against the characteristic heterochromatin marker trimethylated histone 3 lysine 9 (H3K9me3) and euchromatin marker acetylated histone 3 lysine 9 (H3K9ac).

5.2 Material and methods

5.2.1 Quantification of *FAST-1* and *FXN* mRNA levels in FRDA YAC transgenic mouse tissues

Total RNA was extracted from 30mg frozen mouse tissue. Initially, tissue samples were homogenised in 1ml of Trizol by a plastic tissue grinder. The rest of the procedure was performed as described in section 3.2.5. As previously described in section 2.4.5 cDNA was synthesised using cloned AMV kit and qRT-PCR measurements were performed using *Hprt* as the endogenous control and two set of primers that recognize human *FXN* and *FAST-1* cDNA.

5.2.2 Chromatin immunoprecipitation (ChIP) analysis

30mg of frozen cerebellum tissue was homogenised in cold PBS, in a final volume of 1ml and ChIP was performed as previously described (section 2.4.14). For sonication, Bioruptor® Pico sonication System (Diagenode) was used to achieve the desired length of DNA fragments (100bp-500bp). Chromatin samples from mouse cerebellum tissue were sheared for 20 cycles of 30sec on / 30sec off, using 0.1 ml Bioruptor® Microtubes. Real-time PCR was performed in triplicate using a QuantStudio 7 Flex Real-Time PCR instrument (Applied Biosystems). Each qPCR reaction was carried out in triplicates in a 10µl reaction volume by using 2µl of the eluted immunoprecipitated DNA, 5µl of Fast SYBR Green Master Mix (Applied Biosystems) and 12.5pmol of each of the respective forward and reverse primers (Table 2.6). Relative quantification values were normalised to input and minus antibody samples, and finally determined in relation to a control region.

5.3 Results

5.3.1 Quantification of *FXN* and *FAST-1* levels in 3-month old FRDA YAC transgenic mouse tissues

FXN and *FAST-1* expression levels were determined in brain, cerebellum and heart tissues from 3-month old YG8LR and Y47R mice. Y47R expression values were set as 100% and used as the basis for comparative expression results.

Analysis of the 3-month old YG8LR mice (both male and female) compared to Y47R controls revealed that *FXN* mRNA levels reduced to 11% (P<0.001) in the brain, 7% (P<0.001) in the cerebellum, and 11% (P<0.001) in the heart tissues. In contrast, *FAST-1* expression levels increased to 264% (P<0.001) in the brain, 320% (P<0.001) in the cerebellum, and 188% (P<0.001) in the heart tissues, (Figure 5.1).

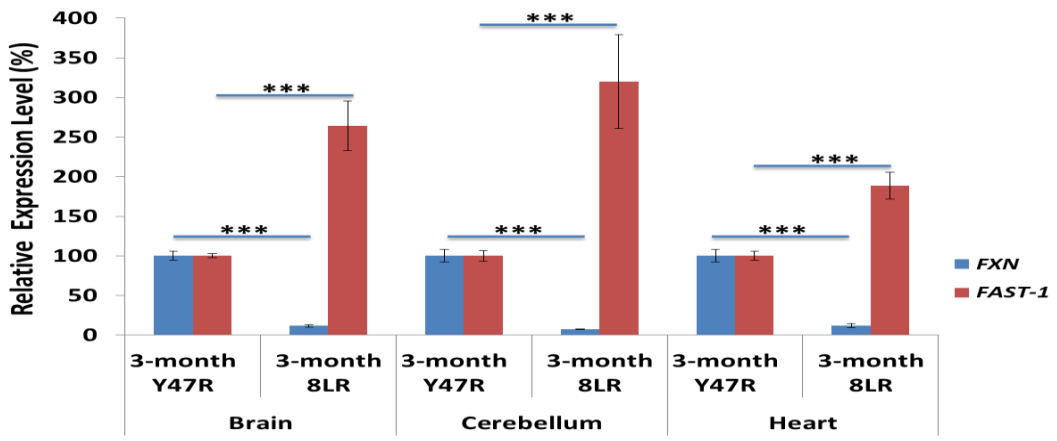

Figure 5.1: Relative *FXN* and *FAST-1* expression levels in brain, cerebellum and heart tissues of the 3-month old YG8LR males and females together. *FXN* and *FAST-1* expression levels are shown relative to the corresponding Y47R mouse tissue value taken as 100% and statistical differences between YG8LR and Y47R control mice are indicated by p values. (n=6), *P* values were generated by t-test. ***p<0.001, error bars=SEM.

To evaluate any potential gender-specific differences in *FXN* and *FAST-1* mRNA expression levels, male and female groups were analysed separately. Analysis of the *FXN* and *FAST-1* expression levels in the 3-month old male YG8LR mice revealed that *FXN* mRNA levels decreased to 15% ($P<0.01$) in the brain, 5 % ($P<0.05$) in the cerebellum, and 6% ($P<0.01$) in the heart tissues, while *FAST-1* expression increased to 320% ($P<0.01$) in the brain, 402 % ($P<0.01$) in the cerebellum, and 240% ($P<0.001$) in the heart tissues (Figure 5.2).

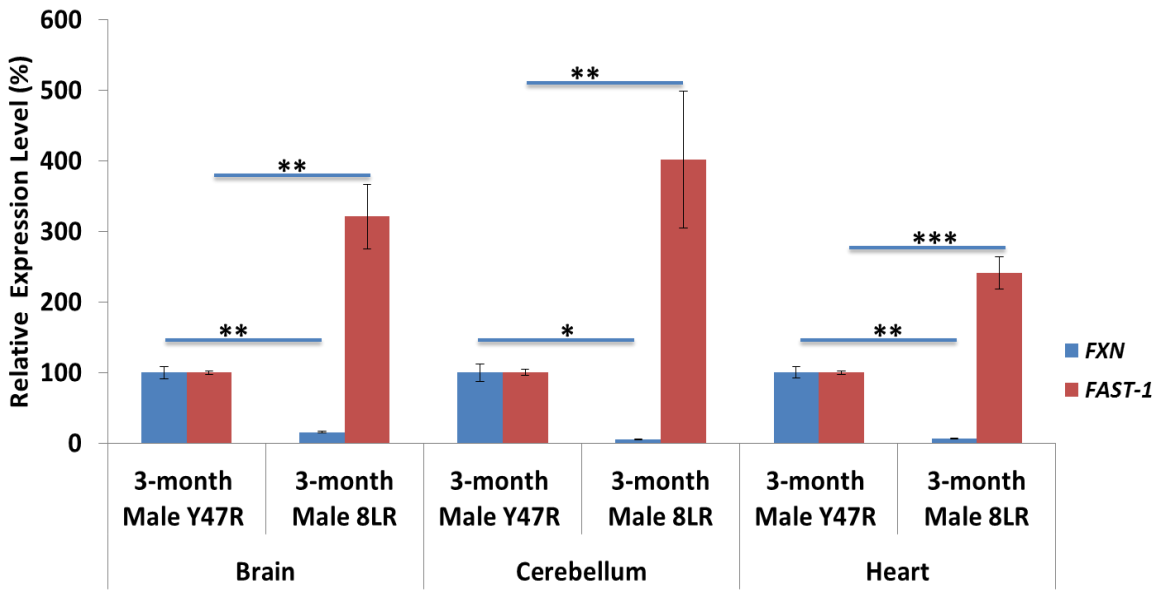

Figure 5.2: Relative *FXN* and *FAST-1* expression levels in brain, cerebellum and heart tissues of the 3-month old YG8LR males. *FXN* and *FAST-1* expression levels are shown relative to the corresponding Y47R mouse tissue value taken as 100% and statistical differences between YG8LR and Y47R control mice are indicated by p values. *P* values were generated by t-test. (n=3) *p<0.05, **p<0.01, ***p<0.001, error bars=SEM.

The same pattern was detected in the 3-month old female YG8LR mice. Analysis of the female mice revealed decreased levels of *FXN* mRNA to 7% (*P*<0.001) in the brain, 8 % (*P*<0.001) in the cerebellum, and 16% (*P*<0.001) in the heart tissues, and increased levels of *FAST-1* expression to 207% (*P*<0.05) in the brain, 238% (*P*<0.001) in the cerebellum, and 136% (*P*<0.05) in the heart tissues (Figure 5.3).

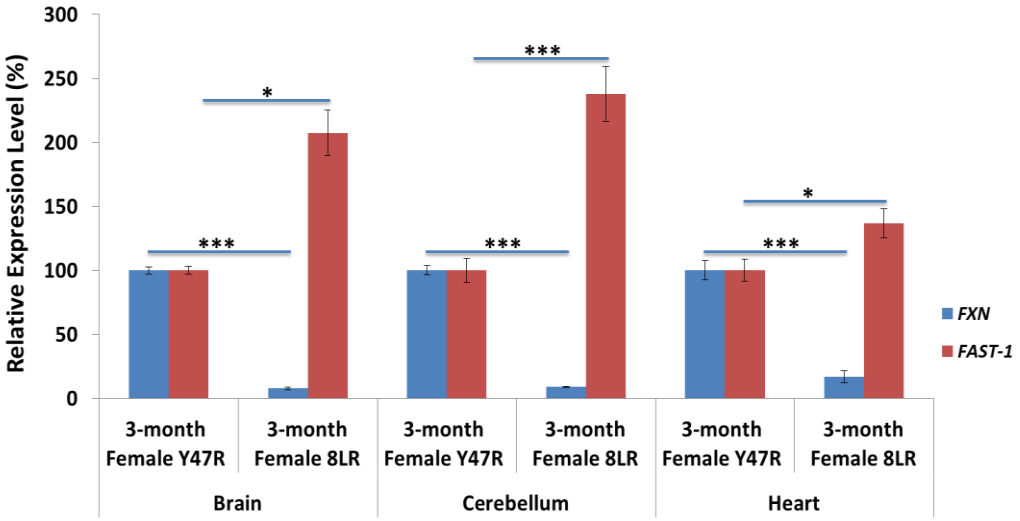

Figure 5.3: Relative *FXN* and *FAST-1* expression levels in brain, cerebellum and heart tissues of the 3-month old YG8LR females. *FXN* and *FAST-1* expression levels are shown relative to the corresponding Y47R mouse tissue value taken as 100% and statistical differences between YG8LR and Y47R control mice are indicated by p values. *P* values were generated by t-test. (n=3) *p<0.05, ***p<0.001, error bars=SEM.

5.3.2 Quantification of *FXN* and *FAST-1* levels in 9-month FRDA YAC transgenic mouse tissues

Analysis of 9-month old YG8LR mice males and females together revealed that the *FXN* mRNA expression levels were significantly decreased in the brain, cerebellum and heart tissues to approximately 5% (P<0.05), 7% (P<0.01) and 8% (P<0.001) respectively, and *FAST-1* expression increased to 172% (P<0.001), 271% (P<0.001)

and 177% (P<0.01) in the brain, cerebellum and heart tissues, respectively, compared to Y47R controls (Fig 5.4).

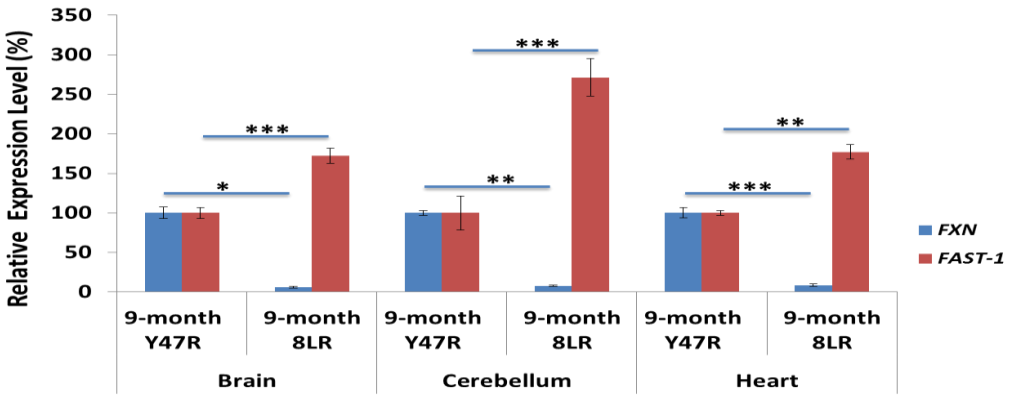

Figure 5.4: Relative *FXN* and *FAST-1* expression levels in brain, cerebellum and heart tissues of the 9-month old YG8LR males and females together. *FXN* and *FAST-1* expression levels are shown relative to the corresponding Y47R mouse tissue value taken as 100% and statistical differences between YG8LR and Y47R control mice are indicated by p values. *P* values were generated by t-test. (n=6) *p<0.05, **p<0.01, ***p<0.001, error bars=SEM.

Significant reduction in the *FXN* mRNA levels were also observed in the brain (5%, P<0.05), cerebellum (8%, P<0.001) and heart (3%, P<0.01) of the 9-month old YG8LR males. In addition, *FAST-1* expression level increased to 172% (*P*<0.01) in the brain, 346% (*P*<0.01) in the cerebellum, and 236% (*P*<0.001) in the heart tissues compared to Y47R controls (Figure 5.5).

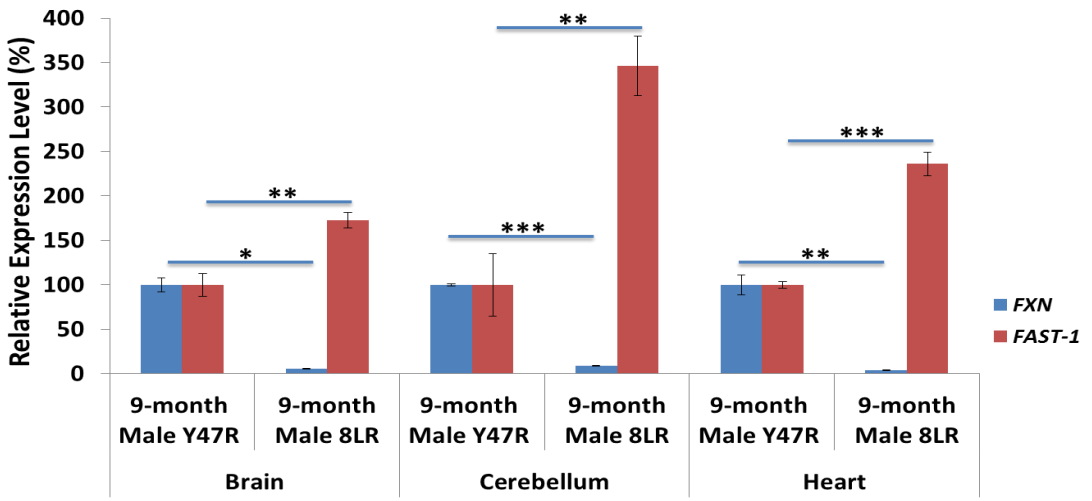

Figure 5.5: Relative *FXN* and *FAST-1* expression levels in brain, cerebellum and heart tissues of the 9-month old YG8LR males. *FXN* and *FAST-1* expression levels are shown relative to the corresponding Y47R mouse tissue value taken as 100% and statistical differences between YG8LR and Y47R control mice are indicated by p values. *P* values were generated by t-test. (n=3) *p<0.05, **p<0.01, ***p<0.001, error bars=SEM.

The same trend was also observed in the tissue of the 9-month old YG8LR females. The levels of *FXN* mRNA expression in the 9-month old YG8LR females were decreased to approximately 6% (P<0.001) in the brain, 6% (P<0.001) in the cerebellum, and 12% (P<0.001) in the heart tissues. In addition, The *FAST-1* expression levels were increased to 170% (P<0.05) in the brain, 196% (P<0.01) in the cerebellum, and 117% (P<0.05) in the heart tissues of the 9-month old YG8LR females compared with the Y47R control (Figure 5.6).

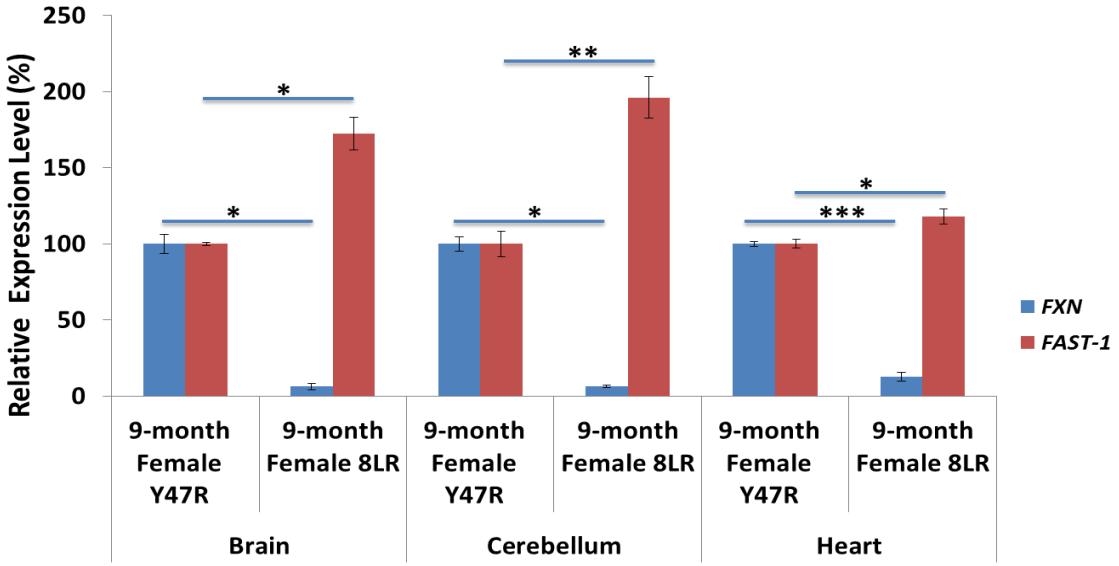

Figure 5.6: Relative *FXN* and *FAST-1* expression levels in brain, cerebellum and heart tissues of the 9-month old YG8LR females. *FXN* and *FAST-1* expression levels are shown relative to the corresponding Y47R mouse tissue value taken as 100% and statistical differences between YG8LR and Y47R control mice are indicated by p values. *P* values were generated by t-test. (n=3) *p<0.05, **p<0.01, ***p<0.001, error bars=SEM.

5.3.3 Investigation of histone modifications in YG8LR mice

Since posttranslational histone modifications play an important role in FRDA, further investigations were carried out to assess levels of acetylation and trimethylation of histone H3 at lysine 9 in 9-month old cerebellum tissue from YG8LR mice. Histone modifications in cerebellum tissues from three male and three female YG8LR mice were investigated by ChIP analysis of the *FXN* promoter region as previously described in section 5.2.2. The immunoprecipitated DNA was amplified by qPCR and results were normalised to input and -Ab values were taken into account (see Table 2.6 for primers used). All of the *FXN* promoter region values from the Y47R control

mouse tissue were taken as 100%. The results showed a significant decrease in H3K9 acetylation and increase in H3K9 trimethylation in both male and female YG8LR mice (Figure 5.7).

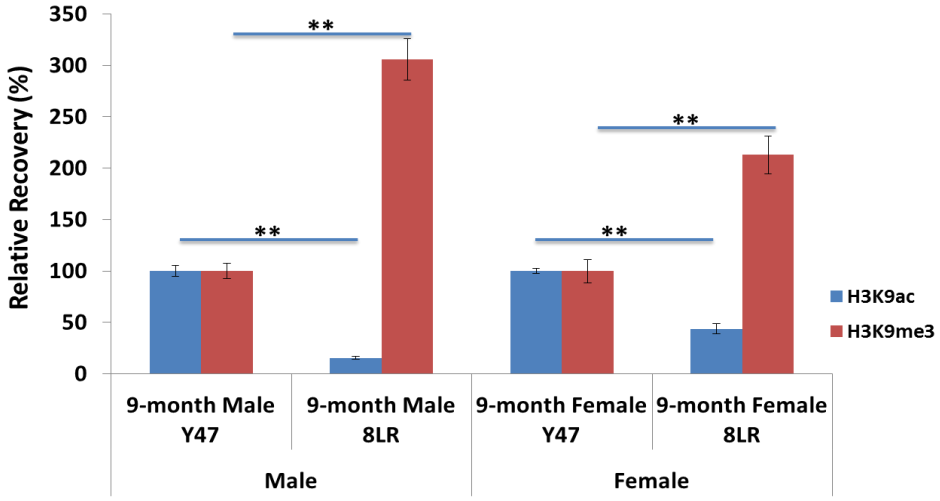

Figure 5.7: ChIP with anti-H3K9me3 and anti-H3K9ac in cerebellum tissue of 9-month old YG8LR mice. ChIP data showed enrichment of H3K9me3 and a significant decrease of H3K9 acetylation at the 5′UTR of the *FXN* gene in both male and females. One ChIP sample was analysed in triplicate from one mouse per group. *P* values were generated by t-test. **p<0.01, error bars=SEM.

5.4 Discussion

Recently, a new GAA repeat expanded mouse model, YG8LR, was derived from the natural breading of YG8sR YAC transgenic mice. The investigation of the molecular and behavioural phenotypes of this new mouse model revealed that features of FRDA patients are replicated in these mice. So far, most *FAST-1* studies have been performed in cultured cells. Therefore, we were keen to investigate the *FAST-1* expression pattern in the most clinically relevant FRDA tissues, brain, cerebellum and heart tissues, but obtaining good quality autopsy tissues from FRDA patients is very difficult. To characterise our new human FRDA YAC transgenic mouse model, YG8LR, and to investigate the expression pattern of the human *FAST-1* transcript across brain, cerebellum and heart tissues, we used a previously established robust qRT-PCR based method to quantify *FAST-1* in YG8LR mouse tissues compared with Y47 control mouse tissues. To evaluate potential differences across time, we examined groups of mice that were 3-months and 9-months old. Similar decreases in the *FXN* expression and increases in the *FAST-1* expression were identified at 3 months and 9 months of age, suggesting that the molecular mechanisms are fixed by 3 months of age and do not vary thereafter. To evaluate any potential gender-specific differences in *FXN* and *FAST-1* mRNA expression levels, 3 male and 3 female mice were analysed separately. Both male and female YG8LR mice express significantly less *FXN* mRNA in all three tissues, brain, cerebellum and heart, compared with control mice Y47R at both 3-months and 9-months of age. In contrast, *FAST-1* levels were found to be significantly increased in YG8LR mice, specifically in the cerebellum tissue, in both males and females, at both 3-months and 9-months of age. These findings are in a good agreement with our previous results from *FAST-1* studies in YG8R mouse models, but different to YG8sR *FAST-1*

studies. Previous studies in our lab identified significantly decreased levels of *FAST-1* expression in brain and cerebellum tissues of the YG8sR mice. On the other hand, YG8R mice tissues exhibited significantly increased levels of *FAST-1* expression relative to control mice. Although YG8sR and YG8LR mice are both derived from an YG8R founder, YG8sR gene expression pattern was found to be different from YG8R and YG8LR mice, which may be due to the GAA repeat expansion differences in these groups of mice (Sandi and Pook, 2015).

A plausible link between heterochromatin and antisense transcription in FRDA led us to investigate the acetylation and methylation of the conserved lysine residue of histone H3 in YG8LR cerebellum tissue in both male and female mice. Our promoter ChIP analysis in both male and female mouse cerebellum tissues showed increased trimethylation of H3K9, a hallmark of silenced chromatin, and H3K9 deacetylation in the *FXN* promoter region. These results are consistent with the findings of previous studies which also reported similar histone modifications in both FRDA human and YG8R mouse brain tissue (Al-Mahdawi *et al.*, 2008). Our similar findings in cell models and mouse tissues indicate that *FAST-1* has a definite role to play in FRDA disease mechanisms (De Biase *et al.*, 2009). It also indicates that YG8LR mice are good models to represent FRDA disease.

Chapter 6: General discussion

FRDA is an autosomal recessive neurodegenerative disorder. It is the most common hereditary form of ataxia with an estimated prevalence of 1 in 50,000 in Caucasians. Most individuals with FRDA have a GAA repeat expansion within the first intron of the *FXN* gene in both alleles, which results in a severe deficiency of the encoded protein, frataxin. After almost 20 years since the identification of the genetic mutation underlying FRDA, the exact mechanisms by which GAA repeat expansions lead to the *FXN* gene silencing still cannot be explained. However, several potential mechanisms have been proposed based on experimental evidence. It has been suggested that abnormal triplex conformation, such as sticky DNA adopted by GAA repeat expansion and/or RNA.DNA hybrids, might impede the transcription of the *FXN* gene (Sakamoto *et al.*, 2001a; Grabczyk *et al.*, 2007). In addition, several previous studies of FRDA mouse models, human tissues and cells have highlighted the possibility of epigenetic changes being also either directly or indirectly involved in *FXN* gene silencing (Al-Mahdawi *et al.*, 2008). Significant levels of histone deacetylation increased H3K9 methylation, significant enrichment of the repressive marks at the 5′UTR of the *FXN* gene, elevated levels of the non-coding RNA (*FAST-1*) and hypermethylation of CpG regions upstream of the GAA repeat are the epigenetic changes that have been identified in FRDA (Herman *et al.*, 2006; Greene *et al.*, 2007; Kim *et al.*, 2011). Enrichment of these repressive chromatin marks in the FRDA alleles provides evidence of the heterochromatin formation in the promoter region involving the critical +1 nucleosome and/or the first intron of the *FXN* gene. Bidichandani and colleagues showed that in FRDA patient-derived fibroblasts, the *FXN Antisense Transcript* (*FAST-1*) was expressed at higher levels. Previous *in vitro* studies in our lab also detected high levels of *FAST-1* in the YG8R and YG22R-derived fibroblast cells (Sandi and Pook, 2015). In addition, it has been reported that

higher levels of *FAST-1* were associated with the severe CTCF depletion and coincidentally heterochromatin formation in the 5′UTR of the *FXN* gene. It is important to note that FRDA cells do not have a generalized defect in CTCF binding (De Biase *et al.*, 2009). Bidichandani and colleagues also reported that knocking down CTCF by siRNA in wild type fibroblasts resulted in high levels of *FAST-1* and reduced *FXN* transcription. These results inspired us to investigate any possible role that *FAST-1* might have in FRDA. Our results from overexpressing *FAST-1* in three different non-FRDA cell lines demonstrated that *FAST-1* can act *in trans* in a similar manner to the *cis*-acting *FAST-1* overexpression that has previously been identified in FRDA fibroblasts. This mechanism of action, where an antisense RNA acts both in *cis* and *trans* and induces epigenetic silencing of its partner sense gene, has been reported for the *p15* naturally occurring antisense transcript (Yu *et al.*, 2008). We also observed that *FAST-1* overexpressing HeLa cells had a reduced occupancy of CTCF, a decreased level of H3K9ac and increased levels of H3K9me3 at the 5′UTR of the *FXN* gene. This is interesting, because the start of *FAST-1* transcription overlaps with the CTCF binding site in the 5′UTR of the *FXN* gene. The effects of *FAST-1* overexpression on CTCF, H3K9ac and H3K9me3 at the *FXN* 5′UTR locus indicated a potential causative role for *FAST-1* in FRDA, as opposed to the increased *FAST-1* expression that is observed in FRDA merely being the passive result of loss of CTCF binding at the *FXN* 5′UTR. Thus, it is plausible that increased *FAST-1* RNA molecules may displace CTCF from the *FXN* 5′UTR, resulting in heterochromatin formation and *FXN* gene silencing. The eviction of CTCF from its binding site by the increased transcription of an antisense non-coding RNA in chicken lysozyme gene has been reported by Bonifer and colleagues. They also demonstrated that CTCF eviction was associated with nucleosome repositioning

over the CTCF binding site, which prevented the re-association of CTCF with its binding site and resulted in a permanent eviction of CTCF/cohesin complex (Lefevre *et al.*, 2008; Phillips and Corces, 2009). Their findings showed that non-coding RNA transcription is indispensable for chromatin remodelling and physical displacement of CTCF (Ong and Corces, 2008). CTCF is a highly-conserved zinc finger multifunctional chromatin protein, which is a key regulator of chromatin architecture and fundamental genomic processes, including insulation, X chromosome inactivation, gene activation, gene repression and chromatin looping (Bonev and Cavalli, 2016). It has been proposed that CTCF is involved in the boundary formation and the partitioning of distinct epigenetic domains. Loss of CTCF-dependent chromatin boundaries is a common event in epigenetic silencing of tumor suppressor genes. Results from TNR disease studies propose an important role for CTCF to establish the local chromatin structure at the repeats and restrict the repressive chromatin structure to the repeat region (Cho *et al.*, 2005; Libby *et al.*, 2008; De Biase *et al.*, 2009). As mentioned earlier, we demonstrated that *FAST-1* overexpressing HeLa cells have a reduced occupancy of CTCF associated with H3K9m3 enrichment and decreased H3K9ac at the 5′UTR of the *FXN* gene. These results concur with other studies that show that the loss of CTCF is associated with the significant enrichment of H3K9me3 around CTCF binding sites and consequently causes chromatin compaction (Pavlaki *et al.*, 2017). Based on these findings, it is plausible to speculate that *FAST-1* overexpression may displace CTCF from the 5′UTR of the *FXN* gene, which would result in secondary changes in chromatin structure that are incompatible with CTCF binding to DNA (Figure 6.1).

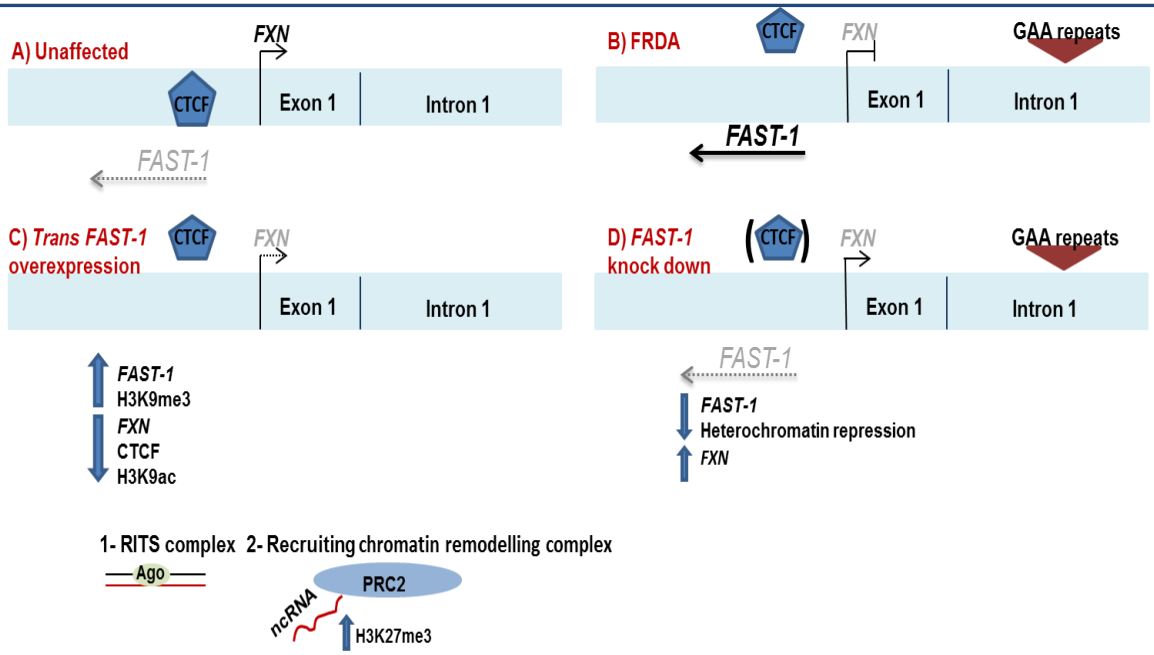

Figure 6.1: Role of *FAST-1* in FRDA. A) In unaffected individuals, the FXN expression is normal, *FAST-1* expression level is low and CTCF is bound to its binding site at 5'UTR of the *FXN* gene. B) In FRDA patients, *FAST-1* expression is elevated and CTCF is depleted from the 5'UTR. C) Following *FAST-1* overexpression, the *FXN* gene and protein expression were decreased. *FAST-1* overexpressing HeLa cells showed reduced occupancy of CTCF, a decreased level of H3K9ac and an increased level of H3K9me3- a hallmark of silenced chromatin- at the 5'UTR of the *FXN* gene. It is plausible to speculate that *FAST-1* overexpression through yet unknown mechanisms (RITS, PRC2) triggers heterochromatin formation and subsequently displaces CTCF from 5'UTR of the *FXN* gene. D) Knocking down *FAST-1* in FRDA fibroblast cells resulted in a 1.5-fold increase in the *FXN* gene expression. Knocking down *FAST-1* may possibly reduce the heterochromatin repression leading to increase in frataxin expression.

It has been demonstrated that embedding short GAA repeat expansions in a completely different sequence context induces heterochromatin formation and silencing of a linked reporter gene. It is well established that it is the intrinsic property of the expanded GAA repeat that causes heterochromatin formation and exert its gene silencing effect (Saveliev *et al.*, 2003). It is not clear yet how GAA repeat expansion results in CTCF dislodgement. The expanded GAA repeat drives chromatin changes in intron 1 and this repeat-proximal heterochromatin spreads from the expanded GAA in intron 1 to the upstream regions of the *FXN* gene

encompassing the *FXN* promoter and CTCF binding site (Herman *et al.*, 2006; Chutake *et al.*, 2014a). One possibility is that the heterochromatin emanating from the expanded GAA repeat displaces CTCF from its binding site. Considering the reduced occupancy of CTCF, H3K9m3 enrichment and decreased H3K9ac levels - hallmarks of heterochromatin formation - at the 5′UTR of the *FXN* gene in *FAST-1* overexpressing HeLa cells, it can be speculated that *FAST-1* RNA, via subsequent heterochromatin, may indirectly displace CTCF. Indeed, several studies have reported that association of H3K9me3 heterochromatic mark with loss of CTCF binding is a common event in epigenetic silencing of cancer-related genes such as *p16* and *p53* (Witcher and Emerson, 2009; Soto-Reyes and Recillas-Targa, 2010). When the chromosomal boundaries are destabilised through dissociation of CTCF and long-range epigenetic organisations are lost, repressive chromatin can spread passively into the promoter region of the *FXN* gene and inhibit its transcription (Witcher and Emerson, 2009; De Biase *et al.*, 2009). In addition, given the enrichment of CTCF at the border of lamina-associated domains (LADs), loss of CTCF can affect the position of genomic loci relative to the nuclear lamina (NL), a location with a generally repressive environment (Guelen *et al.*, 2008; Yáñez-Cuna and van Steensel, 2017). With this in mind, it has been reported that the majority of expanded *FXN* alleles are positioned at the NL. Thus, CTCF loss and subsequent heterochromatin formation may contribute to *FXN* relocation to the NL and its repression (Silva *et al.*, 2015). Recent studies suggest that CTCF has a contradictory function to its classical enhancer-blocker function. It has been proposed that CTCF can tether distant chromatin sites together, create a loop and physically bridge enhancer-promoter interaction (Oti *et al.*, 2016; Hu and Tee, 2017). In addition, it has been demonstrated that gene looping plays an important role in restricting divergent

transcription of non-coding RNA. A protein factor, Ssu72, is found to be essential for gene looping and its mutation results in a significant increase in divergent non-coding RNA production (Tan-Wong *et al.*, 2012). The Ssu72 protein has been shown to bind to cohesin, a protein which strongly co-localises with CTCF (Kim *et al.*, 2010). Based on these findings, it can be speculated that CTCF may facilitate enhancer-promoter interactions at the 5′UTR of the *FXN* gene to drive the *FXN* expression. *FAST-1* overexpression and resultant local depletion of CTCF could abolish this CTCF-mediated enhancer-promoter interaction and consequently reduce *FXN* gene expression. In addition, CTCF can directly interact with the largest subunit of RNA pol II (LS Pol II) (Chernukhin *et al.*, 2007). It can also enhance pausing of RNA pol II and facilitate remodelling of the chromatin and interact with positive elongation factor B to facilitate efficient transcription elongation (Laitem *et al.*, 2015). It has been reported that, in FRDA cell lines, the RNA pol II level is decreased throughout the *FXN* gene but it remains unchanged in the promoter region (Kim *et al.*, 2011). Transcription in the antisense direction is more likely to terminate due to RNA pol II stalling (Bornelöv *et al.*, 2015). Therefore, loss of CTCF in FRDA may contribute to disrupted transcription and repressive chromatin structure. If we consider non-coding transcript to be transcribed in an enhancer-independent manner, then loss of CTCF may promote *FAST-1* transcription in FRDA. Non-coding RNAs have also been found to interact with transcription factors and regulate local gene expression. The growth arrest-specific 5 (*Gas5*) gene encodes a non-coding RNA that can interact with DNA binding domain of the glucocorticoid receptor (GR) and suppress glucocorticoid-activated genes. Measuring occupancy changes of transcription factor GR following overexpression of *Gas5* demonstrated that GR was depleted at its target promoters leading to decrease in GR-induced mRNA expression (Kino *et al.*,

2010; Hudson and Ortlund, 2014). Two transcription factors, SRF and TFAP2, have been found to regulate expression of the *FXN* gene. They bind directly to a region that is located close to the start of *FAST-1* antisense transcript (Li *et al.*, 2010). Therefore, it is possible that *FAST-1* acts as a decoy, reducing the availability of these transcription factors required for the *FXN* gene transcription.

To investigate the molecular mechanisms that may be involved in reducing *FXN* mRNA and protein level in *FAST-1* overexpressing cells, we performed MethylScreen analysis to determine the general DNA methylation status (5mC and 5hmC) at the CpG sites U6 and U11 of the *FXN* upstream GAA intron 1 region. Our result revealed that *FAST-1* overexpressing cells exhibited an FRDA-like increased DNA methylation profile at these CpG sites. Previous studies in our lab showed that the majority of the methylated DNA in the *FXN* upstream GAA repeat region comprises 5hmC rather than 5mC, and 5hmC is known to promote DNA demethylation (Guo *et al.*, 2011). Thus, 5hmC levels can be considered as an indicator of DNA demethylation progression. In contrast, no alterations in the DNA methylation status have previously been detected at the CTCF binding site in the *FXN* 5′UTR region of FRDA cells or tissues (Greene *et al.*, 2007; Al-Mahdawi *et al.*, 2008). Furthermore, it has been reported that a natural occurring antisense transcript could induce DNA demethylation, and its overexpression dramatically abolishes DNA methylation (Imamura *et al.*, 2004). Therefore, one can postulate that increased levels of *FAST-1* might explain the reduced DNA methylation status at the CTCF binding site in the *FXN* 5′UTR, which is concurrent with *FXN* gene silencing. Antisense transcription is a very common phenomenon in mammalian transcriptomes. Although natural antisense transcripts can show mRNA-like characteristics such as having a poly A tail and a 5′ cap structure, the function of

these transcripts is not well understood (Villegas and Zaphiropoulos, 2015). Experimental evidence suggests that non-coding RNA can influence gene expression by recruiting chromatin remodeling complexes to specific alleles and mediating heterochromatin formation and gene silencing. It has been shown that the PRC2 complex associates with almost 3000 NATs in mouse embryonic stem cells (Zhao *et al.*, 2010; Halley *et al.*, 2013). Indeed, antisense transcripts can act as scaffolds or directly recruit PRC2 to chromatin in *cis* and *trans* manners to induce trimethylation of H3K27 and consequently establish a silent chromatin state. According to this mechanism, even a relatively low abundance of antisense transcript can mediate transcriptional repression (Holoch and Moazed, 2015). In addition, recent studies suggest that there is a cooperative mechanism between H3K27 and H3K9 methylation marks, where H3K9me3 can crosstalk with the Polycomb H3K27me3 modification to cooperate in gene silencing. Indeed, H3K27me3-bound PRC2 stabilizes H3K9me3-anchored HP1 and reinforces heterochromatin formation and ensuing gene repression. Significant enrichment of H3K9 and H3K27 methylation and HP1 has been reported at the *FXN* locus of FRDA-derived fibroblast cells. Here in *FAST-1* overexpressing HeLa cells, we also identified significant enrichment of H3K9me3 at the 5′UTR of the *FXN* gene. In addition, it has been reported that 5hmc and H3K27me3 are tightly correlated across a variety of somatic tissues and high levels of both marks have been identified in FRDA (Haffner *et al.*, 2013). Moreover, knowing that genomic repeats are targets of Polycomb complexes and repeat sequences might provide a binding platform for polycomb group proteins, it is conceivable to speculate that the expanded GAA repeat and high levels of *FAST-1* either directly or indirectly recruit PRC2, and

through a cooperative mechanism dictate heterochromatin formation and *FXN* gene silencing in FRDA (Leeb *et al.*, 2010; Al-Mahdawi *et al.*, 2013).

To further investigate the possible role of *FAST-1* in *FXN* gene expression, *FAST-1* was knocked down using lentiviral shRNA particles. Our results showed that knocking down *FAST-1* in FRDA fibroblast cells increases *FXN* gene expression, but not to the level of normal cells (Figure 6.1). Therefore, it can be concluded that since *FAST-1* is associated with epigenetic repression of the *FXN* gene, inhibition of *FAST-1* may be an approach to increase the *FXN* transcripts and stimulate subsequent protein expression. Considering the nature of NATs and the fact that many currently available drugs would not affect the activity of non-coding RNA molecules, developing new methods to disrupt the function of NATs seems necessary. Knocking down NATs using siRNA, using oligonucleotides containing locked nucleic acid modifications (AntagoNAT) or by blocking their interaction with their associated histone modifying enzymes, are the potential methods for targeting NATs (Halley *et al.*, 2013). Using an AntagoNAT strategy to knock down the BDNF (brain-derived neurotrophic factor) NAT has been shown to have promising results in both cell cultures and *in vivo*. Knocking down this antisense transcript, considerably increased sense mRNA and protein expression and modified chromatin marks (Modarresi *et al.*, 2012a). One of the main advantages of targeting NATs is their locus-specific gene upregulation effects. However, there are formidable biological barriers that need to be addressed. These include increasing the resistance to nuclease degradation, reducing the base pair length of the oligonucleotide to increase cellular uptake and decreasing the toxicity effects whilst still maintaining high binding affinity for the target nucleic acid (Wahlestedt, 2013).

In conclusion, a full-length *FAST-1* transcript with a total length of 523bp in size was previously identified in our lab. Mapping the 3′ and 5′ ends of *FAST-1* transcript onto the genome precisely localised them to nucleotides -359 and 164 of the *FXN* gene, respectively. Of particular interest is the fact that the 5′ end at position 164 coincides with a known CTCF binding site. A poly (A) signal was also identified in the *FAST-1* sequence at nucleotide positions of -341 to -346 (Sandi and Pook, 2015). In this research project, I aimed to investigate the potential role of *FXN* antisense transcript-1 (*FAST-1*) in FRDA. Therefore, I first overexpressed *FAST-1* in HEK293, HeLa and fibroblast cell lines and I observed a consistent decrease in *FXN* expression in each cell type compared to control cells. I also identified a positive correlation between *FAST-1* copy number and *FAST-1* expression and a negative correlation between *FAST-1* expression and *FXN* expression. Additionally, I found that *FAST-1* overexpression is associated with CTCF depletion and heterochromatin formation at the 5′UTR of the *FXN* gene. Furthermore, I performed short hairpin RNA (shRNA) knock down of *FAST-1* in FRDA fibroblasts, detecting significantly increased *FXN* expression, but not to the same levels as the control fibroblasts. Studying clinically relevant FRDA tissues from GAA repeat expansion-containing *FXN* transgenic mice revealed that elevated levels of *FAST-1* were associated with increased histone marks for silenced chromatin and decreased *FXN* expression. Overall, our results suggest a role for *FAST-1* in the epigenetic regulation of *FXN* gene expression (Figure 6.1). Further studies are required to fully understand the details of this mechanism and to develop potential new approaches for FRDA clinical therapies.

References

ADINOLFI, S., IANNUZZI, C., PRISCHI, F., PASTORE, C., IAMETTI, S., MARTIN, S. R., BONOMI, F., PASTORE, A. 2009. Bacterial frataxin CyaY is the gatekeeper of iron-sulfur cluster formation catalyzed by IscS. *Nature Structural & Molecular Biology,* 16**,** 390-396.

AL-MAHDAWI, S., PINTO, R. M., ISMAIL, O., VARSHNEY, D., LYMPERI, S., SANDI, C., TRABZUNI, D., POOK, M. 2008. The Friedreich ataxia GAA repeat expansion mutation induces comparable epigenetic changes in human and transgenic mouse brain and heart tissues. *Human Molecular Genetics,* 17**,** 735-746.

AL-MAHDAWI, S., PINTO, R. M., RUDDLE, P., CARROLL, C., WEBSTER, Z., POOK, M. 2004. GAA repeat instability in Friedreich ataxia YAC transgenic mice. *Genomics,* 84**,** 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., POOK, M. A. 2006. GAA repeat expansion mutation mouse models of Friedreich ataxia exhibit oxidative stress leading to progressive neuronal and cardiac pathology. *Genomics,* 88**,** 580-590.

AL-MAHDAWI, S., SANDI, C., MOURO PINTO, R., POOK, M. A. 2013. Friedreich ataxia patient tissues exhibit increased 5-hydroxymethylcytosine modification and decreased CTCF binding at the FXN locus. *PLoS One,* 8**,** 74956.

ALCID, E. A., TSUKIYAMA, T. 2014. ATP-dependent chromatin remodeling shapes the long noncoding RNA landscape. *Genes & Development,* 28**,** 2348-2360.

ALLFREY, V., FAULKNER, R., MIRSKY, A. 1964. Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. *Proceedings of the National Academy of Sciences,* 51**,** 786-794.

ARMSTRONG, J. S., KHDOUR, O., HECHT, S. M. 2010. Does oxidative stress contribute to the pathology of Friedreich's ataxia? A radical question. *The FASEB Journal,* 24**,** 2152-2163.

BABCOCK, M., DE SILVA, D., OAKS, R., DAVIS-KAPLAN, S., JIRALERSPONG, S., MONTERMINI, L., PANDOLFO, M., KAPLAN, J. 1997. Regulation of Mitochondrial Iron Accumulation by Yfh1p, a Putative Homolog of Frataxin. *Science,* 276**,** 1709-1712.

BARTOLOMEI, M. S., ZEMEL, S., TILGHMAN, S. M. 1991. Parental imprinting of the mouse H19 gene. *Nature,* 351**,** 153.

BAYOT, A., REICHMAN, S., LEBON, S., CSABA, Z., AUBRY, L., STERKERS, G., HUSSON, I., RAK, M., RUSTIN, P. 2013. Cis-silencing of PIP5K1B evidenced in Friedreich's ataxia patient cells results in cytoskeleton anomalies. *Human Molecular Genetics,* 22**,** 2894-2904.

BELL, A. C., FELSENFELD, G. 2000. Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. *Nature,* 405**,** 482-485.

BELL, A. C., WEST, A. G., FELSENFELD, G. 1999. The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell,* 98**,** 387-396.

BENCZE, K. Z., KONDAPALLI, K. C., COOK, J. D., MCMAHON, S., MILLÁN-PACHECO, C., PASTOR, N., STEMMLER, T. L. 2006. The structure and function of frataxin. *Critical Reviews in Biochemistry and Molecular Biology,* 41**,** 269-291.

BENINI, M., FORTUNI, S., CONDÒ, I., ALFEDI, G., MALISAN, F., TOSCHI, N., SERIO, D., MASSARO, D. S., ARCURI, G., TESTI, R. 2017. E3 Ligase RNF126 Directly Ubiquitinates Frataxin, Promoting Its Degradation: Identification of a Potential Therapeutic Target for Friedreich Ataxia. *Cell Reports,* 18**,** 2007-2017.

BIDICHANDANI, S., CASTRO, A., CHUTAKE, Y. 2011. RNA-mediated transcriptional gene silencing in Friedreich ataxia. *Current Opinion in Biotechnology,* 22**,** 16.

BIDICHANDANI, S. I., ASHIZAWA, T., PATEL, P. I. 1998. The GAA Triplet-Repeat Expansion in Friedreich Ataxia Interferes with Transcription and May Be Associated with an Unusual DNA Structure. *The American Journal of Human Genetics,* 62**,** 111-121.

BONEV, B., CAVALLI, G. 2016. Organization and function of the 3D genome. *Nature Reviews Genetics,* 17**,** 661-678.

BORNELÖV, S., KOMOROWSKI, J., WADELIUS, C. 2015. Different distribution of histone modifications in genes with unidirectional and bidirectional transcription and a role of CTCF and cohesin in directing transcription. *BMC Genomics,* 16**,** 300.

BOROS, J., ARNOULT, N., STROOBANT, V., COLLET, J.-F., DECOTTIGNIES, A. 2014. Polycomb repressive complex 2 and H3K27me3 cooperate with H3K9 methylation to maintain heterochromatin protein 1α at chromatin. *Molecular and Cellular Biology,* 34**,** 3662-3674.

BRADLEY, J. L., HOMAYOUN, S., HART, P. E., SCHAPIRA, A. H. V., COOPER, J. M. 2004. Role of Oxidative Damage in Friedreich's Ataxia. *Neurochemical Research,* 29**,** 561-567.

BULTEAU, A.-L., O'NEILL, H. A., KENNEDY, M. C., IKEDA-SAITO, M., ISAYA, G., SZWEDA, L. I. 2004. Frataxin acts as an iron chaperone protein to modulate mitochondrial aconitase activity. *Science,* 305**,** 242-245.

BÜRK, K. 2017. Friedreich Ataxia: current status and future prospects. *Cerebellum & Ataxias,* 4**,** 4.

BURKE, L. J., ZHANG, R., BARTKUHN, M., TIWARI, V. K., TAVOOSIDANA, G., KURUKUTI, S., WETH, C., LEERS, J., GALJART, N., OHLSSON, R. 2005. CTCF binding and higher order chromatin structure of the H19 locus are maintained in mitotic chromatin. *The EMBO Journal,* 24**,** 3291-3300.

BURNETT, R., MELANDER, C., PUCKETT, J. W., SON, L. S., WELLS, R. D., DERVAN, P. B., GOTTESFELD, J. M. 2006. DNA sequence-specific polyamides alleviate transcription inhibition associated with long GAA· TTC repeats in Friedreich's ataxia. *Proceedings of the National Academy of Sciences,* 103**,** 11497-11502.

CAMPUZANO, V., MONTERMINI, L., LUTZ, Y., COVA, L., HINDELANG, C., JIRALERSPONG, S., TROTTIER, Y., KISH, S. J., FAUCHEUX, B., TROUILLAS, P. 1997. Frataxin is reduced in Friedreich ataxia patients and is associated with mitochondrial membranes. *Human Molecular Genetics,* 6**,** 1771-1780.

CAMPUZANO, V., MONTERMINI, L., MOLTÒ, M. D., PIANESE, L., COSSÉE, M., CAVALCANTI, F., MONROS, E., RODIUS, F., DUCLOS, F., MONTICELLI, A. 1996. Friedreich's ataxia: autosomal recessive disease caused by an intronic GAA triplet repeat expansion. *Science,* 271**,** 1423-1427.

CASTALDO, I., PINELLI, M., MONTICELLI, A., ACQUAVIVA, F., GIACCHETTI, M., FILLA, A., SACCHETTI, S., KELLER, S., AVVEDIMENTO, V. E., CHIARIOTTI, L. 2008. DNA methylation in intron 1 of the frataxin gene is related to GAA repeat length and age of onset in Friedreich ataxia patients. *Journal of Medical Genetics,* 45**,** 808-812.

CAVADINI, P., ADAMEC, J., TARONI, F., GAKH, O., ISAYA, G. 2000. Two-step Processing of Human Frataxin by Mitochondrial Processing Peptidase PRECURSOR AND INTERMEDIATE FORMS ARE CLEAVED AT DIFFERENT RATES. *Journal of Biological Chemistry,* 275**,** 41469-41475.

CHAMBERLAIN, S., SHAW, J., ROWLAND, A., WALLIS, J., SOUTH, S., NAKAMURA, Y., VON GABAIN, A., FARRALL, M., WILLIAMSON, R. 1988. Mapping of mutation causing Friedreich's ataxia to human chromosome 9. *Nature,* 334**,** 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. 2013. Heterochromatinization induced by GAA-repeat hyperexpansion in Friedreich's ataxia can be reduced upon HDAC inhibition by vitamin B3. *Human Molecular Genetics,* 22**,** 2662-2675.

CHANDLER, S. P., KANSAGRA, P., HIRST, M. C. 2003. Fragile X (CGG) n repeats induce a transcriptional repression in cis upon a linked promoter: evidence for a chromatin mediated effect. *BMC Molecular Biology,* 4**,** 3.

CHEN, J., SUN, M., KENT, W. J., HUANG, X., XIE, H., WANG, W., ZHOU, G., SHI, R. Z., ROWLEY, J. D. 2004. Over 20% of human transcripts might form sense–antisense pairs. *Nucleic Acids Research,* 32**,** 4812-4820.

CHEN, X., MARIAPPAN, S., CATASTI, P., RATLIFF, R., MOYZIS, R. K., LAAYOUN, A., SMITH, S. S., BRADBURY, E. M., GUPTA, G. 1995. Hairpins are formed by the single DNA strands of the fragile X triplet repeats: structure and biological implications. *Proceedings of the National Academy of Sciences,* 92**,** 5199-5203.

CHERNUKHIN, I., SHAMSUDDIN, S., KANG, S. Y., BERGSTRÖM, R., KWON, Y.-W., YU, W., WHITEHEAD, J., MUKHOPADHYAY, R., DOCQUIER, F., FARRAR, D. 2007. CTCF interacts with and recruits the largest subunit of RNA polymerase II to CTCF target sites genome-wide. *Molecular and Cellular Biology,* 27**,** 1631-1648.

CHO, D. H., THIENES, C. P., MAHONEY, S. E., ANALAU, E., FILIPPOVA, G. N., TAPSCOTT, S. J. 2005. Antisense transcription and heterochromatin at the DM1 CTG repeats are constrained by CTCF. *Molecular Cell,* 20**,** 483-489.

CHU, Y., YUE, X., YOUNGER, S. T., JANOWSKI, B. A., COREY, D. R. 2010. Involvement of argonaute proteins in gene silencing and activation by RNAs complementary to a non-coding transcript at the progesterone receptor promoter. *Nucleic Acids Research,* 38**,** 7736-7748.

CHUNG, D. W., RUDNICKI, D. D., YU, L., MARGOLIS, R. L. 2011. A natural antisense transcript at the Huntington's disease repeat locus regulates HTT expression. *Human Molecular Genetics,* 20**,** 3467–3477.

CHUTAKE, Y. K., COSTELLO, W. N., LAM, C., BIDICHANDANI, S. I. 2014a. Altered nucleosome positioning at the transcription start site and deficient transcriptional initiation in Friedreich ataxia. *Journal of Biological Chemistry,* 289**,** 15194-15202.

CHUTAKE, Y. K., LAM, C., COSTELLO, W. N., ANDERSON, M., BIDICHANDANI, S. I. 2014c. Epigenetic promoter silencing in Friedreich ataxia is dependent on repeat length. *Annals of Neurology,* 76**,** 522-528.

CHUTAKE, Y. K., LAM, C. C., COSTELLO, W. N., ANDERSON, M. P., BIDICHANDANI, S. I. 2016. Reversal of epigenetic promoter silencing in Friedreich ataxia by a class I histone deacetylase inhibitor. *Nucleic Acids Research,* 44**,** 5095–5104.

CLARK, R. M., DE BIASE, I., MALYKHINA, A. P., AL-MAHDAWI, S., POOK, M., BIDICHANDANI, S. I. 2007. The GAA triplet-repeat is unstable in the context of the human FXN locus and displays age-dependent expansions in cerebellum and DRG in a transgenic mouse model. *Human Genetics,* 120**,** 633-640.

COSSÉE, M., DÜRR, A., SCHMITT, M., DAHL, N., TROUILLAS, P., ALLINSON, P., KOSTRZEWA, M., NIVELON-CHEVALLIER, A., GUSTAVSON, K.-H., KOHLSCHÜTTER, A., MÜLLER, U., MANDEL, J.-L., BRICE, A., KOENIG, M., CAVALCANTI, F., TAMMARO, A., DE MICHELE, G., FILLA, A., COCOZZA, S., LABUDA, M., MONTERMINI, L., POIRIER, J., PANDOLFO, M., INSTITUTIONEN FÖR GENETIK OCH, P., MEDICINSKA OCH FARMACEUTISKA, V., MEDICINSKA, F., UPPSALA, U. 1999. Friedreich's ataxia: Point mutations and clinical presentation of compound heterozygotes. *Annals of Neurology,* 45**,** 200-206.

COSSÉE, M., PUCCIO, H., GANSMULLER, A., KOUTNIKOVA, H., DIERICH, A., LEMEUR, M., FISCHBECK, K., DOLLÉ, P., KŒNIG, M. 2000. Inactivation of the Friedreich ataxia mouse gene leads to early embryonic lethality without iron accumulation. *Human Molecular Genetics,* 9**,** 1219-1226.

COSSEE, M., SCHMITT, M., CAMPUZANO, V., REUTENAUER, L., MOUTOU, C., MANDEL, J.-L., KOENIG, M. 1997. Evolution of the Friedreich's Ataxia Trinucleotide Repeat Expansion: Founder Effect and Premutations. *Proceedings of the National Academy of Sciences of the United States of America,* 94**,** 7452-7457.

CUDDAPAH, S., JOTHI, R., SCHONES, D. E., ROH, T.-Y., CUI, K., ZHAO, K. 2009. Global analysis of the insulator binding protein CTCF in chromatin barrier regions reveals demarcation of active and repressive domains. *Genome Research,* 19**,** 24-32.

DE BIASE, I., CHUTAKE, Y. K., RINDLER, P. M., BIDICHANDANI, S. I. 2009. Epigenetic silencing in Friedreich ataxia is associated with depletion of CTCF (CCCTC-binding factor) and antisense transcription. *PLoS One,* 4**,** 7914.

DE BIASE, I., RASMUSSEN, A., MONTICELLI, A., AL-MAHDAWI, S., POOK, M., COCOZZA, S., BIDICHANDANI, S. I. 2007. Somatic instability of the expanded GAA triplet-repeat sequence in Friedreich ataxia progresses throughout life. *Genomics,* 90**,** 1-5.

DE MICHELE, G., CAVALCANTI, F., CRISCUOLO, C., PIANESE, L., MONTICELLI, A., FILLA, A., COCOZZA, S. 1998. Parental gender, age at birth and expansion length influence GAA repeat intergenerational instability in the X25 gene: pedigree studies and analysis of sperm from patients with Friedreich's ataxia. *Human Molecular Genetics,* 7**,** 1901-1906.

DE MICHELE, G., FILLA, A., CAVALCANTI, F., TAMMARO, A., MONTICELLI, A., PIANESE, L., DI SALLE, F., PERRETTI, A., SANTORO, L., CARUSO, G., COCOZZA, S. 2000. Atypical Friedreich ataxia phenotype associated with a novel missense mutation in the X25 gene. *Neurology,* 54**,** 496-499.

DELATYCKI, M. B., WILLIAMSON, R., FORREST, S. M. 2000. Friedreich ataxia: An overview. *Journal of Medical Genetics,* 37**,** 1-8.

DHE-PAGANON, S., SHIGETA, R., CHI, Y.-I., RISTOW, M., SHOELSON, S. E. 2000. Crystal structure of human frataxin. *Journal of Biological Chemistry,* 275**,** 30753-30756.

DION, V., WILSON, J. H. 2009. Instability and chromatin structure of expanded trinucleotide repeats. *Trends in Genetics,* 25**,** 288-297.

DUPONT, C., ARMANT, D. R., BRENNER, C. A. Epigenetics: definition, mechanisms and clinical perspective. Seminars in Reproductive Medicine, 2009. © Thieme Medical Publishers, 27**,**351-357.

DÜRR, A., COSSEE, M., AGID, Y., CAMPUZANO, V., MIGNARD, C., PENET, C., MANDEL, J.-L., BRICE, A., KOENIG, M. 1996. Clinical and genetic abnormalities in patients with Friedreich's ataxia. *New England Journal of Medicine,* 335**,** 1169-1175.

ELGIN, S. C., GREWAL, S. I. 2003. Heterochromatin: silence is golden. *Current Biology,* 13**,** 895-898.

ELGIN, S. C., REUTER, G. 2013. Position-effect variegation, heterochromatin formation, and gene silencing in Drosophila. *Cold Spring Harbor Perspectives in Biology,* 5**,** a017780.

ESSEBIER, A., WOLF, P. V., CAO, M. D., CARROLL, B. J., BALASUBRAMANIAN, S., BODÉN, M. 2016. Statistical enrichment of epigenetic states around triplet repeats that can undergo expansions. *Frontiers in Neuroscience,* 10**,** 92.

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. 2012. FXN methylation predicts expression and clinical outcome in Friedreich ataxia. *Annals of Neurology,* 71**,** 487-497.

FAGHIHI, M. A., WAHLESTEDT, C. 2009. Regulatory roles of natural antisense transcripts. *Nature Reviews Molecular Cell Biology,* 10**,** 637-643.

FERRI, E., PETOSA, C., MCKENNA, C. E. 2016. Bromodomains: Structure, function and pharmacology of inhibition. *Biochemical Pharmacology,* 106**,** 1-18.

FILIPPOVA, G. N., FAGERLIE, S., KLENOVA, E. M., MYERS, C., DEHNER, Y., GOODWIN, G., NEIMAN, P. E., COLLINS, S. J., LOBANENKOV, V. V. 1996. An exceptionally conserved transcriptional repressor, CTCF, employs different combinations of zinc fingers to bind diverged promoter sequences of avian and mammalian c-myc oncogenes. *Molecular and Cellular Biology,* 16**,** 2802-2813.

FILIPPOVA, G. N., THIENES, C. P., PENN, B. H., CHO, D. H., HU, Y. J., MOORE, J. M., KLESERT, T. R., LOBANENKOV, V. V., TAPSCOTT, S. J. 2001. CTCF-binding sites

flank CTG/CAG repeats and form a methylation-sensitive insulator at the DM1 locus. *Nature Genetics,* 28**,** 335-343.
GELLERA, C., CASTELLOTTI, B., MARIOTTI, C., MINERI, R., SEVESO, V., DIDONATO, S., TARONI, F. 2007. Frataxin gene point mutations in Italian Friedreich ataxia patients. *Neurogenetics,* 8**,** 289-299.
GERHARDT, J., BHALLA, A. D., BUTLER, J. S., PUCKETT, J. W., DERVAN, P. B., ROSENWAKS, Z., NAPIERALA, M. 2016. Stalled DNA Replication Forks at the Endogenous GAA Repeats Drive Repeat Expansion in Friedreich's Ataxia Cells. *Cell Reports,* 16**,** 1218-1227.
GIBSON, T. J., KOONIN, E. V., MUSCO, G., PASTORE, A., BORK, P. 1996. Friedreich's ataxia protein: phylogenetic evidence for mitochondrial dysfunction. *Trends in Neurosciences,* 19**,** 465-468.
GIJSELINCK, I., VAN MOSSEVELDE, S., VAN DER ZEE, J., SIEBEN, A., ENGELBORGHS, S., DE BLEECKER, J., IVANOIU, A., DERYCK, O., EDBAUER, D., ZHANG, M. 2016. The C9orf72 repeat size correlates with onset age of disease, DNA methylation and transcriptional downregulation of the promoter. *Molecular Psychiatry,* 21**,** 1112-1124.
GRABCZYK, E., MANCUSO, M., 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 Research,* 35**,** 5351-5359.
GRABCZYK, E., USDIN, K. 2000a. Alleviating transcript insufficiency caused by Friedreich's ataxia triplet repeats. *Nucleic acids research,* 28**,** 4930-4937.
GRABCZYK, E., USDIN, K. 2000c. 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 Research,* 28**,** 2815-2822.
GREENE, E., ENTEZAM, A., KUMARI, D., USDIN, K. 2005. Ancient repeated DNA elements and the regulation of the human frataxin promoter. *Genomics,* 85**,** 221-230.
GREENE, E., MAHISHI, L., ENTEZAM, A., KUMARI, D., USDIN, K. 2007. Repeat-induced epigenetic changes in intron 1 of the frataxin gene and its consequences in Friedreich ataxia. *Nucleic Acids Research,* 35**,** 3383-3390.
GROH, M., LUFINO, M., WADE-MARTINS, R., GROMAK, N. 2014. R-loops associated with triplet repeat expansions promote gene silencing in Friedreich ataxia and fragile X syndrome. *PLoS Genet,* 10**,** 1004318.
GUDDE, A. E., VAN HEERINGEN, S. J., DE OUDE, A. I., VAN KESSEL, I. D., ESTABROOK, J., WANG, E. T., WIERINGA, B., WANSINK, D. G. 2017. Antisense transcription of the myotonic dystrophy locus yields low-abundant RNAs with and without (CAG) n repeat. *RNA Biology,* 14**,** 1374-1388.
GUELEN, L., PAGIE, L., BRASSET, E., MEULEMAN, W., FAZA, M. B., TALHOUT, W., EUSSEN, B. H., DE KLEIN, A., WESSELS, L., DE LAAT, W. 2008. Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature,* 453**,** 948-951.
GUO, J. U., SU, Y., ZHONG, C., MING, G.-L., SONG, H. 2011. Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. *Cell,* 145**,** 423-434.
HAFFNER, M., PELLAKURU, L., GHOSH, S., LOTAN, T., NELSON, W. G., DE MARZO, A., YEGNASUBRAMANIAN, S. 2013. Tight correlation of 5-hydroxymethylcytosine and Polycomb marks in health and disease. *Cell Cycle,* 12**,** 1835-1841.
HALLECK, M. S., GURLEY, L. R. 1981. Histone acetylation and heterochromatin content of cultured Peromyscus cells. *Experimental Cell Research,* 132**,** 201-213.
HALLEY, P., KHORKOVA, O., WAHLESTEDT, C. 2013. Natural antisense transcripts as therapeutic targets. *Drug Discovery Today. Therapeutic Strategies,* 10**,** 119.
HANAUER, A., CHERY, M., FUJITA, R., DRIESEL, A., GILGENKRANTZ, S., MANDEL, J. 1990. The Friedreich ataxia gene is assigned to chromosome 9q13-q21 by mapping of tightly linked markers and shows linkage disequilibrium with D9S15. *American Journal of Human Genetics,* 46**,** 133.

HAWKINS, P. G., MORRIS, K. V. 2010. Transcriptional regulation of Oct4 by a long non-coding RNA antisense to Oct4-pseudogene 5. *Transcription,* 1**,** 165-175.

HE, Y., VOGELSTEIN, B., VELCULESCU, V. E., PAPADOPOULOS, N., KINZLER, K. W. 2008. The Antisense Transcriptomes of Human Cells. *Science,* 322**,** 1855-1857.

HEIDENFELDER, B. L., MAKHOV, A. M., TOPAL, M. D. 2003. Hairpin formation in Friedreich's ataxia triplet repeat expansion. *Journal of Biological Chemistry,* 278**,** 2425-2431.

HERMAN, D., JENSSEN, K., BURNETT, R., SORAGNI, E., PERLMAN, S. L., GOTTESFELD, J. M. 2006. Histone deacetylase inhibitors reverse gene silencing in Friedreich's ataxia. *Nature Chemical Biology,* 2**,** 551-558.

HEROLD, M., BARTKUHN, M., RENKAWITZ, R. 2012. CTCF: insights into insulator function during development. *Development,* 139**,** 1045-1057.

HOLLOWAY, T. P., ROWLEY, S. M., DELATYCKI, M. B., SARSERO, J. P. 2011. Detection of interruptions in the GAA trinucleotide repeat expansion in the FXN gene of Friedreich ataxia. *Biotechniques,* 50**,** 182-186.

HOLOCH, D., MOAZED, D. 2015. RNA-mediated epigenetic regulation of gene expression. *Nature Reviews Genetics,* 16**,** 71-84.

HU, Z., TEE, W.-W. 2017. Enhancers and chromatin structures: regulatory hubs in gene expression and diseases. *Bioscience Reports,* 37**,** 13.

HUBER, F., BUNINA, D., GUPTA, I., KHMELINSKII, A., MEURER, M., THEER, P., STEINMETZ, L. M., KNOP, M. 2016. Protein abundance control by non-coding antisense transcription. *Cell Reports,* 15**,** 2625-2636.

HUDSON, W. H., ORTLUND, E. A. 2014. The structure, function and evolution of proteins that bind DNA and RNA. *Nature Reviews Molecular Cell Biology,* 15**,** 749-760.

HUERTAS, P., AGUILERA, A. 2003. Cotranscriptionally formed DNA: RNA hybrids mediate transcription elongation impairment and transcription-associated recombination. *Molecular Cell,* 12**,** 711-721.

IMAMURA, T., YAMAMOTO, S., OHGANE, J., HATTORI, N., TANAKA, S., SHIOTA, K. 2004. Non-coding RNA directed DNA demethylation of Sphk1 CpG island. *Biochemical and Biophysical Research Communications,* 322**,** 593-600.

KANDURI, C. Kcnq1ot1: a chromatin regulatory RNA. Seminars in Cell & Developmental Biology, 2011. Elsevier, 22**,**343-350.

KANHERKAR, R. R., BHATIA-DEY, N., CSOKA, A. B. 2014. Epigenetics across the human lifespan. *Frontiers in Cell and Developmental Biology,* 2**,** 49.

KARMODIYA, K., KREBS, A. R., OULAD-ABDELGHANI, M., KIMURA, H., TORA, L. 2012. H3K9 and H3K14 acetylation co-occur at many gene regulatory elements, while H3K14ac marks a subset of inactive inducible promoters in mouse embryonic stem cells. *BMC Genomics,* 13**,** 424.

KAYTOR, M. D., BURRIGHT, E. N., DUVICK, L. A., ZOGHBI, H. Y., ORR, H. T. 1997. Increased trinucleotide repeat instability with advanced maternal age. *Human Molecular Genetics,* 6**,** 2135-2139.

KHALIL, A. M., GUTTMAN, M., HUARTE, M., GARBER, M., RAJ, A., MORALES, D. R., THOMAS, K., PRESSER, A., BERNSTEIN, B. E., VAN OUDENAARDEN, A. 2009. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proceedings of the National Academy of Sciences,* 106**,** 11667-11672.

KHONSARI, H., SCHNEIDER, M., AL-MAHDAWI, S., CHIANEA, Y. G., THEMIS, M., PARRIS, C., POOK, M. A. 2016. Lentivirus-meditated frataxin gene delivery reverses genome instability in Friedreich ataxia patient and mouse model fibroblasts. *Gene Therapy,* 23**,** 846-856.

KIM, E., NAPIERALA, M., DENT, S. Y. 2011. Hyperexpansion of GAA repeats affects post-initiation steps of FXN transcription in Friedreich's ataxia. *Nucleic Acids Research,* 39**,** 8366-8377.

KIM, H. S., BAEK, K. H., HA, G. H., LEE, J. C., KIM, Y. N., LEE, J., PARK, H. Y., LEE, N. R., LEE, H., CHO, Y. 2010. The hsSsu72 phosphatase is a cohesin-binding protein that

regulates the resolution of sister chromatid arm cohesion. *The EMBO Journal,* 29**,** 3544-3557.

KIM, S., YU, N.-K., KAANG, B.-K. 2015. CTCF as a multifunctional protein in genome regulation and gene expression. *Experimental & Molecular Medicine,* 47**,** 166.

KINO, T., HURT, D. E., ICHIJO, T., NADER, N., CHROUSOS, G. P. 2010. Noncoding RNA Gas5 is a growth arrest and starvation-associated repressor of the glucocorticoid receptor. *Science Signaling,* 3**,** ra8.

KLENOVA, E., NICOLAS, R., PATERSON, H., CARNE, A., HEATH, C., GOODWIN, G., NEIMAN, P., LOBANENKOV, V. 1993. CTCF, a conserved nuclear factor required for optimal transcriptional activity of the chicken c-myc gene, is an 11-Zn-finger protein differentially expressed in multiple forms. *Molecular and Cellular Biology,* 13**,** 7612-7624.

KOHWI, Y., KOHWI-SHIGEMATSU, T. 1991. Altered gene expression correlates with DNA structure. *Genes & Development,* 5**,** 2547-2554.

KOUTNIKOVA, H., CAMPUZANO, V., FOURY, F., DOLLÉ, P., CAZZALINI, O., KOENIG, M. 1997. Studies of human, mouse and yeast homologues indicate a mitochondrial function for frataxin. *Nature Genetics,* 16**,** 345-351.

KRANS, A., KEARSE, M. G., TODD, P. K. 2016. Repeat-associated non-AUG translation from antisense CCG repeats in fragile X tremor/ataxia syndrome. *Annals of Neurology,* 80**,** 871-881.

KUMARI, D., BIACSI, R. E., USDIN, K. 2011. Repeat expansion affects both transcription initiation and elongation in friedreich ataxia cells. *Journal of Biological Chemistry,* 286**,** 4209-4215.

LACHNER, M., O'CARROLL, D., REA, S., MECHTLER, K., JENUWEIN, T. 2001. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature,* 410**,** 116-120.

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., 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. *Human Molecular Genetics,* 16**,** 3174-3187.

LAITEM, C., ZABOROWSKA, J., TELLIER, M., YAMAGUCHI, Y., CAO, Q., EGLOFF, S., HANDA, H., MURPHY, S. 2015. CTCF regulates NELF, DSIF and P-TEFb recruitment during transcription. *Transcription,* 6**,** 79-90.

LEEB, M., PASINI, D., NOVATCHKOVA, M., JARITZ, M., HELIN, K., WUTZ, A. 2010. Polycomb complexes act redundantly to repress genomic repeats and genes. *Genes & Development,* 24**,** 265-276.

LEFEVRE, P., WITHAM, J., LACROIX, C. E., COCKERILL, P. N., BONIFER, C. 2008. The LPS-induced transcriptional upregulation of the chicken lysozyme locus involves CTCF eviction and noncoding RNA transcription. *Molecular Cell,* 32**,** 129-139.

LI, C. H., CHEN, Y. 2013. Targeting long non-coding RNAs in cancers: progress and prospects. *The International Journal of Biochemistry & Cell Biology,* 45**,** 1895-1910.

LI, K., SINGH, A., CROOKS, D. R., DAI, X., CONG, Z., PAN, L., HA, D., ROUAULT, T. A. 2010. Expression of human frataxin is regulated by transcription factors SRF and TFAP2. *PLoS One,* 5**,** 12286.

LI, L., MATSUI, M., COREY, D. R. 2016. Activating frataxin expression by repeat-targeted nucleic acids. *Nature Communications,* 7**,** 10606.

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. 2008. CTCF cis-regulates trinucleotide repeat instability in an epigenetic manner: a novel basis for mutational hot spot determination. *PLoS Genet,* 4**,** 1000257.

LIBRI, V., YANDIM, C., ATHANASOPOULOS, S., LOYSE, N., NATISVILI, T., LAW, P. P., CHAN, P. K., MOHAMMAD, T., MAURI, M., TAM, K. T. 2014. Epigenetic and neurological effects and safety of high-dose nicotinamide in patients with Friedreich's ataxia: an exploratory, open-label, dose-escalation study. *The Lancet,* 384**,** 504-513.

LOBANENKOV, V., NICOLAS, R., ADLER, V., PATERSON, H., KLENOVA, E., POLOTSKAJA, A., GOODWIN, G. 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**,** 1743-1753.

LÓPEZ CASTEL, A., NAKAMORI, M., TOMÉ, S., CHITAYAT, D., GOURDON, G., THORNTON, C. A., PEARSON, C. E. 2010. Expanded CTG repeat demarcates a boundary for abnormal CpG methylation in myotonic dystrophy patient tissues. *Human Molecular Genetics,* 20**,** 1-15.

LORINCZ, M. C., DICKERSON, D. R., SCHMITT, M., GROUDINE, M. 2004. Intragenic DNA methylation alters chromatin structure and elongation efficiency in mammalian cells. *Nature Structural & Molecular Biology,* 11**,** 1068-1075.

LUFINO, M. M., SILVA, A. M., NÉMETH, A. H., ALEGRE-ABARRATEGUI, J., RUSSELL, A. J., WADE-MARTINS, R. 2013. A GAA repeat expansion reporter model of Friedreich's ataxia recapitulates the genomic context and allows rapid screening of therapeutic compounds. *Human Molecular Genetics,* 22**,** 5173-5187.

MARCHESE, F. P., HUARTE, M. 2014. Long non-coding RNAs and chromatin modifiers: their place in the epigenetic code. *Epigenetics,* 9**,** 21-26.

MCMACKIN, M. Z., HENDERSON, C. K., CORTOPASSI, G. A. 2017. Neurobehavioral deficits in the KIKO mouse model of Friedreich's ataxia. *Behavioural Brain Research,* 316**,** 183-188.

MIRANDA, C. J., SANTOS, M. M., OHSHIMA, K., SMITH, J., LI, L., BUNTING, M., COSSÉE, M., KOENIG, M., SEQUEIROS, J., KAPLAN, J. 2002. Frataxin knockin mouse. *FEBS letters,* 512**,** 291-297.

MIRKIN, S. M. 2006. DNA structures, repeat expansions and human hereditary disorders. *Current Opinion in Structural Biology,* 16**,** 351-358.

MODARRESI, F., FAGHIHI, M. A., LOPEZ-TOLEDANO, M. A., FATEMI, R. P., MAGISTRI, M., BROTHERS, S. P., VAN DER BRUG, M. P., WAHLESTEDT, C. 2012a. Inhibition of natural antisense transcripts in vivo results in gene-specific transcriptional upregulation. *Nature Biotechnology,* 30**,** 453-459.

MODARRESI, F., FAGHIHI, M. A., LOPEZ-TOLEDANO, M. A., FATEMI, R. P., MAGISTRI, M., BROTHERS, S. P., VAN DER BRUG, M. P., WAHLESTEDT, C. 2012b. Natural Antisense Inhibition Results in Transcriptional De-Repression and Gene Upregulation. *Nature Biotechnology,* 30**,** 453.

MONTERMINI, L., MONTICELLI, A., TURANO, M., FILLA, A., DE MICHELE, G., COCOZZA, S., ANDERMANN, E., LABUDA, M., RICHTER, A., PANDOLFO, M., CAVALCANTI, F., PIANESE, L., IODICE, L., FARINA, G. 1997. The Friedreich ataxia GAA triplet repeat: Premutation and normal alleles. *Human Molecular Genetics,* 6**,** 1261-1266.

MONTICELLI, A., GIACCHETTI, M., DE BIASE, I., PIANESE, L., TURANO, M., PANDOLFO, M., COCOZZA, S. 2004. New clues on the origin of the Friedreich ataxia expanded alleles from the analysis of new polymorphisms closely linked to the mutation. *Human Genetics,* 114**,** 458-463.

MORRIS, K. V. 2012. *Non-coding RNAs and epigenetic regulation of gene expression: Drivers of natural selection*, Horizon Scientific Press**.**

MORRIS, K. V., SANTOSO, S., TURNER, A.-M., PASTORI, C., HAWKINS, P. G. 2008. Bidirectional transcription directs both transcriptional gene activation and suppression in human cells. *PLoS Genet,* 4**,** 1000258.

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. 2006. Bidirectional expression of CUG and CAG expansion transcripts and intranuclear polyglutamine inclusions in spinocerebellar ataxia type 8. *Nature Genetics,* 38**,** 758-769.

MURRAY, S. C., HAENNI, S., HOWE, F. S., FISCHL, H., CHOCIAN, K., NAIR, A., MELLOR, J. 2015. Sense and antisense transcription are associated with distinct chromatin architectures across genes. *Nucleic Acids Research,* 43**,** 7823-7837.

MUSCO, G., STIER, G., KOLMERER, B., ADINOLFI, S., MARTIN, S., FRENKIEL, T., GIBSON, T., PASTORE, A. 2000. Towards a structural understanding of Friedreich's ataxia: the solution structure of frataxin. *Structure,* 8**,** 695-707.

NAUMANN, A., HOCHSTEIN, N., WEBER, S., FANNING, E., DOERFLER, W. 2009. A distinct DNA-methylation boundary in the 5′-upstream sequence of the FMR1 promoter binds nuclear proteins and is lost in fragile X syndrome. *The American Journal of Human Genetics,* 85**,** 606-616.

OHHATA, T., HOKI, Y., SASAKI, H., SADO, T. 2008. Crucial role of antisense transcription across the Xist promoter in Tsix-mediated Xist chromatin modification. *Development,* 135**,** 227-235.

ONG, C.-T., CORCES, V. G. 2008. Modulation of CTCF insulator function by transcription of a noncoding RNA. *Developmental Cell,* 15**,** 489-490.

OTI, M., FALCK, J., HUYNEN, M. A., ZHOU, H. 2016. CTCF-mediated chromatin loops enclose inducible gene regulatory domains. *BMC Genomics,* 17**,** 252.

PANDOLFO, M. 1998. Molecular genetics and pathogenesis of Friedreich ataxia. *Neuromuscular Disorders,* 8**,** 409-415.

PANDOLFO, M. 2002. Iron metabolism and mitochondrial abnormalities in Friedreich ataxia. *Blood Cells, Molecules, and Diseases,* 29**,** 536-547.

PANDOLFO, M. 2008. Friedreich ataxia. *Archives of Neurology,* 65**,** 1296-1303.

PANDOLFO, M. 2009. Friedreich ataxia: The clinical picture. *Journal of Neurology,* 256**,** 3-8.

PANDOLFO, M., HAUSMANN, L. 2013. Deferiprone for the treatment of Friedreich's ataxia. *Journal of Neurochemistry,* 126**,** 142-146.

PASTORE, A., PUCCIO, H. 2013. Frataxin: a protein in search for a function. *Journal of Neurochemistry,* 126**,** 43-52.

PATEL, P. I., ISAYA, G. 2001. Friedreich Ataxia: From GAA Triplet–Repeat Expansion to Frataxin Deficiency. *The American Journal of Human Genetics,* 69**,** 15-24.

PAVLAKI, I., DOCQUIER, F., CHERNUKHIN, I., TEIF, V. B., KLENOVA, E. 2017.Poly (ADP-ribosyl) ation dependent changes in CTCF-chromatin binding and gene expression in breast cells. *bioRxiv* doi: 10.1101/175448.

PEARSON, C. E., EDAMURA, K. N., CLEARY, J. D. 2005. Repeat instability: mechanisms of dynamic mutations. *Nature Reviews Genetics,* 6**,** 729-742.

PELECHANO, V., STEINMETZ, L. M. 2013. Gene regulation by antisense transcription. *Nature Reviews Genetics,* 14**,** 880-893.

PERDOMINI, M., BELBELLAA, B., MONASSIER, L., REUTENAUER, L., MESSADDEQ, N., CARTIER, N., CRYSTAL, R. G., AUBOURG, P., PUCCIO, H. 2014. Prevention and reversal of severe mitochondrial cardiomyopathy by gene therapy in a mouse model of Friedreich's ataxia. *Nature Medicine,* 20**,** 542-547.

PHILLIPS, J. E., CORCES, V. G. 2009. CTCF: master weaver of the genome. *Cell,* 137**,** 1194-1211.

PIANESE, L., CAVALCANTI, F., DE MICHELE, G., FILLA, A., CAMPANELLA, G., CALABRESE, O., CASTALDO, I., MONTICELLI, A., COCOZZA, S. 1997. The effect of parental gender on the GAA dynamic mutation in the FRDA gene [2]. *American Journal of Human Genetics,* 60**,** 460-463.

PIANESE, L., TURANO, M., LO CASALE, M. S., DE BIASE, I., GIACCHETTI, M., MONTICELLI, A., CRISCUOLO, C., FILLA, A., COCOZZA, S. 2004. Real time PCR quantification of frataxin mRNA in the peripheral blood leucocytes of Friedreich ataxia patients and carriers. *Journal of Neurology, Neurosurgery and Psychiatry,* 75**,** 1061-1063.

POOK, M. A. 2012. DNA methylation and trinucleotide repeat expansion diseases. *In: DNA Methylation-From Genomics to Technology.* TATARINOVA, T., KERTON, O. (eds.)193. InTech.

POOK, M. A., AL-MAHDAWI, S., CARROLL, C. J., COSSÉE, M., PUCCIO, H., LAWRENCE, L., CLARK, P., LOWRIE, M. B., BRADLEY, J. L., COOPER, M. J. 2001. Rescue of the Friedreich's ataxia knockout mouse by human YAC transgenesis. *Neurogenetics,* 3**,** 185-193.

PUCCIO, H., SIMON, D., COSSÉE, M., CRIQUI-FILIPE, P., TIZIANO, F., MELKI, J., HINDELANG, C., MATYAS, R., RUSTIN, P., KOENIG, M. 2001. Mouse models for Friedreich ataxia exhibit cardiomyopathy, sensory nerve defect and Fe-S enzyme deficiency followed by intramitochondrial iron deposits. *Nature Genetics,* 27**,** 181-186.

PUNGA, T., BÜHLER, M. 2010. Long intronic GAA repeats causing Friedreich ataxia impede transcription elongation. *EMBO Molecular Medicine,* 2**,** 120-129.

RADISKY, D. C., BABCOCK, M. C., KAPLAN, J. 1999. The yeast frataxin homologue mediates mitochondrial iron efflux evidence for a mitochondrial iron cycle. *Journal of Biological Chemistry,* 274**,** 4497-4499.

RAI, M., SORAGNI, E., JENSSEN, K., BURNETT, R., HERMAN, D., COPPOLA, G., GESCHWIND, D. H., GOTTESFELD, J. M., PANDOLFO, M. 2008. HDAC inhibitors correct frataxin deficiency in a Friedreich ataxia mouse model. *PloS one,* 3**,** e1958.

RAJESWARI, M. R. 2012. DNA triplex structures in neurodegenerative disorder, Friedreich's ataxia. *Journal of Biosciences,* 37**,** 519-532.

REA, S., EISENHABER, F., O'CARROLL, D., STRAHL, B. D., SUN, Z.-W., SCHMID, M., OPRAVIL, S., MECHTLER, K., PONTING, C. P., ALLIS, C. D. 2000. Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature,* 406**,** 593-599.

ROTIG, A., LONLAY, P. D., CHRETIEN, D., FOURY, F., KOENIG, M., SIDI, D., MUNNICH, A., RUSTIN, P. 1997. Aconitase and mitochondrial iron-sulphur protein deficiency in Friedreich ataxia. *Nature Genetics,* 17**,** 215-217.

RUFINI, A., FORTUNI, S., ARCURI, G., CONDÒ, I., SERIO, D., INCANI, O., MALISAN, F., VENTURA, N., TESTI, R. 2011. Preventing the ubiquitin–proteasome-dependent degradation of frataxin, the protein defective in Friedreich's ataxia. *Human Molecular Genetics,* 20**,** 1253-1261.

SAKAMOTO, N., CHASTAIN, P. D., PARNIEWSKI, P., OHSHIMA, K., PANDOLFO, M., GRIFFITH, J. D., WELLS, R. D. 1999. Sticky DNA: self-association properties of long GAA· TTC repeats in R· R· Y triplex structures from Friedreich's ataxia. *Molecular Cell,* 3**,** 465-475.

SAKAMOTO, N., LARSON, J. E., IYER, R. R., MONTERMINI, L., PANDOLFO, M., WELLS, R. D. 2001a. 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. *Journal of Biological Chemistry,* 276**,** 27178-27187.

SAKAMOTO, N., OHSHIMA, K., MONTERMINI, L., PANDOLFO, M., WELLS, R. D. 2001c. Sticky DNA, a self-associated complex formed at long GAA· TTC repeats in intron 1 of the frataxin gene, inhibits transcription. *Journal of Biological Chemistry,* 276**,** 27171-27177.

SAKSOUK, N., SIMBOECK, E., DÉJARDIN, J. 2015. Constitutive heterochromatin formation and transcription in mammals. *Epigenetics & Chromatin,* 8**,** 3.

SANDI, C., AL-MAHDAWI, S., POOK, M. A. 2013a. Epigenetics in Friedreich's ataxia: Challenges and opportunities for therapy.

SANDI, C., AL-MAHDAWI, S., POOK, M. A. 2013b. Epigenetics in Friedreich's ataxia: challenges and opportunities for therapy. *Genetics Research International,* 2013**,** 852080.

SANDI, C., PINTO, R. M., AL-MAHDAWI, S., EZZATIZADEH, V., BARNES, G., JONES, S., RUSCHE, J. R., GOTTESFELD, J. M., POOK, M. A. 2011. Prolonged treatment with pimelic o-aminobenzamide HDAC inhibitors ameliorates the disease phenotype of a Friedreich ataxia mouse model. *Neurobiology of Disease,* 42**,** 496-505.

SANDI, C., SANDI, M., VIRMOUNI, S. A., AL-MAHDAWI, S., POOK, M. A. 2014. Epigenetic-based therapies for Friedreich ataxia. *Frontiers in Genetics,* 5**,** 165.

SANDI, M., POOK, M. 2015. *Identification and quantification of FXN antisense transcript 1 (FAST-1) in friedreich ataxia.* Dissertation/Thesis, Brunel University London U6 - ctx_ver=Z39.88-2004&ctx_enc=info%3Aofi%2Fenc%3AUTF-8&rfr_id=info%3Asid%2Fsummon.serialssolutions.com&rft_val_fmt=info%3Aofi%2Ff

mt%3Akev%3Amtx%3Adissertation&rft.genre=dissertation&rft.title=Identification+and+quantification+of+FXN+antisense+transcript+1+%28FAST-1%29+in+friedreich+ataxia&rft.DBID=D6B&rft.au=Sandi%2C+Madhavi&rft.au=Pook%2C+M&rft.date=2015&rft.pub=Brunel+University+London&rft.externalDBID=com_2438_23&rft.externalDocID=oai_bura_brunel_ac_uk_2438_12464&paramdict=en-US U7 - Dissertation.

SAVELIEV, A., EVERETT, C., SHARPE, T., WEBSTER, Z., FESTENSTEIN, R. 2003. DNA triplet repeats mediate heterochromatin-protein-1-sensitive variegated gene silencing. *Nature,* 422**,** 909-913.

SCHMUCKER, S., ARGENTINI, M., CARELLE-CALMELS, N., MARTELLI, A., PUCCIO, H. 2008. The in vivo mitochondrial two-step maturation of human frataxin. *Human Molecular Genetics,* 17**,** 3521-3531.

SCHULZ, J., DEHMER, T., SCHÖLS, L., MENDE, H., HARDT, C., VORGERD, M., BÜRK, K., MATSON, W., DICHGANS, J., BEAL, M. 2000. Oxidative stress in patients with Friedreich ataxia. *Neurology,* 55**,** 1719-1721.

SCHWARTZ, J. C., YOUNGER, S. T., JANOWSKI, B. A., COREY, D. R. 2014. MODULATING GENE EXPRESSION WITH agRNA AND GAPMERS TARGETING ANTISENSE TRANSCRIPTS. Google Patents.

SEZNEC, H., SIMON, D., BOUTON, C., REUTENAUER, L., HERTZOG, A., GOLIK, P., PROCACCIO, V., PATEL, M., DRAPIER, J.-C., KOENIG, M. 2005. Friedreich ataxia: the oxidative stress paradox. *Human Molecular Genetics,* 14**,** 463-474.

SEZNEC, H., SIMON, D., MONASSIER, L., CRIQUI-FILIPE, P., GANSMULLER, A., RUSTIN, P., KOENIG, M., PUCCIO, H. 2004. Idebenone delays the onset of cardiac functional alteration without correction of Fe-S enzymes deficit in a mouse model for Friedreich ataxia. *Human Molecular Genetics,* 13**,** 1017-1024.

SHARMA, R., BHATTI, S., GOMEZ, M., CLARK, R. M., MURRAY, C., ASHIZAWA, T., 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. *Human Molecular Genetics,* 11**,** 2175-2187.

SILVA, A. M., BROWN, J. M., BUCKLE, V. J., WADE-MARTINS, R., LUFINO, M. M. 2015. Expanded GAA repeats impair FXN gene expression and reposition the FXN locus to the nuclear lamina in single cells. *Human Molecular Genetics,* 24**,** 3457-3471.

SIMS, R. J., MILLHOUSE, S., CHEN, C.-F., LEWIS, B. A., ERDJUMENT-BROMAGE, H., TEMPST, P., MANLEY, J. L., REINBERG, D. 2007. Recognition of trimethylated histone H3 lysine 4 facilitates the recruitment of transcription postinitiation factors and pre-mRNA splicing. *Molecular Cell,* 28**,** 665-676.

SORAGNI, E., MIAO, W., IUDICELLO, M., JACOBY, D., DE MERCANTI, S., CLERICO, M., LONGO, F., PIGA, A., KU, S., CAMPAU, E. 2014. Epigenetic therapy for Friedreich ataxia. *Annals of Neurology,* 76**,** 489-508.

SOTO-REYES, E., RECILLAS-TARGA, F. 2010. Epigenetic regulation of the human p53 gene promoter by the CTCF transcription factor in transformed cell lines. *Oncogene,* 29**,** 2217-2227.

SWYGERT, S. G., PETERSON, C. L. 2014. Chromatin dynamics: interplay between remodeling enzymes and histone modifications. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms,* 1839**,** 728-736.

TAN-WONG, S. M., ZAUGG, J. B., CAMBLONG, J., XU, Z., ZHANG, D. W., MISCHO, H. E., ANSARI, A. Z., LUSCOMBE, N. M., STEINMETZ, L. M., PROUDFOOT, N. J. 2012. Gene loops enhance transcriptional directionality. *Science,* 338**,** 671-675.

TUFARELLI, C., STANLEY, J. A. S., GARRICK, D., SHARPE, J. A., AYYUB, H., WOOD, W. G., HIGGS, D. R. 2003. Transcription of antisense RNA leading to gene silencing and methylation as a novel cause of human genetic disease. *Nature Genetics,* 34**,** 157-165.

VAN KRUIJSBERGEN, I., HONTELEZ, S., VEENSTRA, G. J. C. 2015. Recruiting polycomb to chromatin. *The International Journal of Biochemistry & Cell Biology,* 67**,** 177-187.

VILLEGAS, V. E., ZAPHIROPOULOS, P. G. 2015. Neighboring gene regulation by antisense long non-coding RNAs. *International Journal of Molecular Sciences,* 16**,** 3251-3266.

VYAS, P. M., TOMAMICHEL, W. J., PRIDE, P. M., BABBEY, C. M., WANG, Q., MERCIER, J., MARTIN, E. M., PAYNE, R. M. 2012. A TAT–Frataxin fusion protein increases lifespan and cardiac function in a conditional Friedreich's ataxia mouse model. *Human Molecular Genetics,* 21**,** 1230-1247.

WAHLESTEDT, C. 2013. Targeting long non-coding RNA to therapeutically upregulate gene expression. *Nature Reviews Drug Discovery,* 12**,** 433-446.

WAN, G., MATHUR, R., HU, X., LIU, Y., ZHANG, X., PENG, G., LU, X. 2013. Long non-coding RNA ANRIL (CDKN2B-AS) is induced by the ATM-E2F1 signaling pathway. *Cellular Signalling,* 25**,** 1086-1095.

WANG, Z., ZANG, C., ROSENFELD, J. A., SCHONES, D. E., BARSKI, A., CUDDAPAH, S., CUI, K., ROH, T.-Y., PENG, W., ZHANG, M. Q. 2008. Combinatorial patterns of histone acetylations and methylations in the human genome. *Nature Genetics,* 40**,** 897-903.

WEST, A. G., GASZNER, M., FELSENFELD, G. 2002. Insulators: many functions, many mechanisms. *Genes & Development,* 16**,** 271-288.

WETH, O., PAPROTKA, C., GÜNTHER, K., SCHULTE, A., BAIERL, M., LEERS, J., GALJART, N., RENKAWITZ, R. 2014. CTCF induces histone variant incorporation, erases the H3K27me3 histone mark and opens chromatin. *Nucleic Acids Research,* 42**,** 11941-11951.

WHITEHEAD, J., PANDEY, G. K., KANDURI, C. 2009. Regulation of the mammalian epigenome by long noncoding RNAs. *Biochimica et Biophysica Acta (BBA)-General Subjects,* 1790**,** 936-947.

WITCHER, M., EMERSON, B. M. 2009. Epigenetic silencing of the p16INK4a tumor suppressor is associated with loss of CTCF binding and a chromatin boundary. *Molecular Cell,* 34**,** 271-284.

YANDIM, C., NATISVILI, T., FESTENSTEIN, R. 2013. Gene regulation and epigenetics in Friedreich's ataxia. *Journal of Neurochemistry,* 126**,** 21-42.

YÁÑEZ-CUNA, J. O., VAN STEENSEL, B. 2017. Genome–nuclear lamina interactions: from cell populations to single cells. *Current Opinion in Genetics & Development,* 43**,** 67-72.

YANG, J., CORCES, V. G. 2011. Chromatin insulators: a role in nuclear organization and gene expression. *Advances in Cancer Research,* 110**,** 43.

YAP, K. L., LI, S., MUÑOZ-CABELLO, A. M., RAGUZ, S., ZENG, L., MUJTABA, S., GIL, J., WALSH, M. J., ZHOU, M.-M. 2010. Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. *Molecular Cell,* 38**,** 662-674.

YU, W., GIUS, D., ONYANGO, P., MULDOON-JACOBS, K., KARP, J., FEINBERG, A. P., CUI, H. 2008. Epigenetic silencing of tumour suppressor gene p15 by its antisense RNA. *Nature,* 451**,** 202-206.

ZHAO, J., OHSUMI, T. K., KUNG, J. T., OGAWA, Y., GRAU, D. J., SARMA, K., SONG, J. J., KINGSTON, R. E., BOROWSKY, M., LEE, J. T. 2010. Genome-wide identification of polycomb-associated RNAs by RIP-seq. *Molecular Cell,* 40**,** 939-953.

ZHAO, J., SUN, B. K., ERWIN, J. A., SONG, J.-J., LEE, J. T. 2008. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science,* 322**,** 750-756.

ZU, T., LIU, Y., BAÑEZ-CORONEL, M., REID, T., PLETNIKOVA, O., LEWIS, J., MILLER, T. M., HARMS, M. B., FALCHOOK, A. E., SUBRAMONY, S. 2013. RAN proteins and RNA foci from antisense transcripts in C9ORF72 ALS and frontotemporal dementia. *Proceedings of the National Academy of Sciences,* 110**,** E4968-E4977.