



The localisation of non-coding RNA in the human cell

A thesis submitted in partial fulfilment of the requirements for
the degree of Doctor of Philosophy

By

Catherine Gayaneh Heath

BSc in Molecular Biology

Department of Molecular Biology and Biotechnology

University of Sheffield

Sheffield, United Kingdom

December 2018

Acknowledgements

I would principally like to thank my supervisor, Professor Stuart Wilson, for guiding me through the project and offering support and advice throughout my time as a PhD student. I would further like to thank Dr Nicolas Viphakone for the endless training he provided, practically and intellectually! I am also exceptionally grateful to Jacob Parker and Dr Ian Sudbery for the bioinformatics analysis presented in this thesis. My colleagues in the Wilson lab, Niko, Simon, Janet, Agata, Marcus, Llywelyn, Dani, Vicky, Helen, James and Ivo, have been instrumental in my project by providing lots of valuable advice and maintaining a positive and entertaining atmosphere in the lab. Last but not least, huge thanks to Matt Parker for helping me to understand bioinformatics and for all the support he has given me throughout these four years.

Abstract

RNA comes in a variety of classes that perform diverse functions throughout the cell. RNA is trafficked to compartments within the nucleus or from the nucleus to the cytoplasm by pathways that are generally specific to the class of RNA. However, recent studies have revealed overlaps between the RNA cargoes that are transported by these pathways. iCLIP data shows proteins that are essential to the mRNA export pathway are not bound exclusively to mRNA but to a variety of non-coding RNAs. Despite the binding of mRNA export proteins to tRNA, we show that the mRNA export pathway does not play a direct role in the export of tRNA to the cytoplasm in human cells. We also learn that the canonical human tRNA export pathways requires further investigation since depletion of the main proteins involved has little effect on tRNA localisation. The depletion of mRNA export proteins bound to nuclear-restricted RNAs, *scaRNA2* and *XIST*, did not affect the intra-nuclear localisation of the RNAs but did have an impact on the levels of *XIST* RNA. The RNA-binding protein hnRNP U is associated with many of the same non-coding RNAs as the mRNA export proteins. hnRNP U is involved in the localisation of specific RNA transcripts to the chromatin and could be involved in preventing their export to the cytoplasm. The newly developed Auxin-Inducible Degron (AID) system was employed to deplete the cell of hnRNP U and assess RNA localisation on a global scale. We demonstrate that hnRNP U is a general chromatin-retention factor for many non-coding RNA classes, some of which are degraded in other cellular compartments after their release from the chromatin.

Table of Contents

Acknowledgements	1
Abstract	2
Abbreviations	6
Chapter 1 - Introduction	9
1.1 The nuclear pore.....	10
1.2 Ran-dependent RNA export	12
1.2.1 The Ran cycle.....	12
1.2.2 tRNA processing and export	16
1.2.3 The Crm1/Exportin-1 export pathway	21
1.3 Messenger RNA export	24
1.4 Additional RNA classes	30
1.4.1 Small Cajal body-specific RNA	30
1.4.2 Long non-coding RNA	30
1.5 hnRNP U	34
1.5.1 hnRNP U functions	36
1.5.2 hnRNP U and Disease.....	39
1.5.3 hnRNP U-like proteins.....	42
1.6 RNA Surveillance.....	43
1.6.1 The exosome complex	43
1.6.2 Nonsense-mediated decay (NMD)	45
1.7 Project Aims	47
Chapter 2 - Materials and Methods	48
2.1 Miscellaneous Buffers	48
2.2 Tissue Culture	49

2.3 Protein Biology	53
2.4 RNA Biology	56
2.5 Cell Biology	62
2.6 DNA and Molecular Biology	64
2.7 Statistical Testing	67
Chapter 3 - The role of TREX and Nxf1 in the cellular localisation of non-coding RNA	68
3.1 TREX and Nxf1 associate with tRNAs	70
3.2 The depletion of TREX does not perturb the cellular localisation of tRNA, determined by RNA fractionation	73
3.3 The depletion of TREX and Nxf1 does not perturb the cellular localisation of tRNA as determined by FISH	80
3.4 The protein levels of Exportin-T and 5 do not change considerably following the depletion of TREX and Nxf1	82
3.5 TREX and Nxf1 are not directly involved in the intra-nuclear localisation of <i>scaRNA2</i>	85
3.6 TREX and Nxf1 are not directly involved in the Intra-nuclear localisation of <i>XIST</i>	87
3.7 TREX but not NXF1 is required to maintain XIST RNA levels	90
3.8 Chapter 3 Summary	92
Chapter 4 – Using the Auxin-Inducible Degron system to deplete hnRNP U from human cells.....	94
4.1 Attempts to deplete hnRNP U in human cells using RNAi	96
4.2 Perturbing the function of hnRNP U using a dominant-negative mutant.....	99
4.3 Generating an Auxin-Inducible Degron (AID) tagged hnRNP U cell line	103
4.3.1 Adding the mini-AID tag to hnRNP U	106
4.3.2 Tagging hnRNP U with the full-length AID tag	111
4.3.3 Validation of the hnRNP U-flAID cell line	116

4.3.4 Comparing the hnRNP U-flAID cell line to existing hnRNP U RNAi systems	120
4.4 Chapter 4 Summary	126
Chapter 5 - hnRNP U is a general chromatin-retention factor for non-coding RNA	127
5.1 hnRNP U is required to maintain the levels of nuclear non-coding RNAs	128
5.2 hnRNP U depletion can result in chromatin-associated ncRNA release and degradation	134
5.3 A global approach demonstrates the general role of hnRNP U in the retention of non-coding RNA on the chromatin	145
5.4 Chapter 5 Summary	167
Chapter 6 - Discussion	168
6.1 Export pathways in tRNA export	168
6.2 Intra-nuclear localisation of non-coding RNA	170
6.3 hnRNP U and non-coding RNA localisation	171
6.4 Conclusion	177
References	178

Abbreviations

AID	Auxin-inducible degron
ATP	Adenosine triphosphate
AUX/IAA	Auxin/Indole-3-acetic acid
CBC	Cap-binding complex
CTD	C-terminal domain
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
EJC	Exon-junction complex
FG Nups	Phenylalanine-glycine Nups
FISH	Fluorescence in-situ hybridisation
flAID	Full-length AID
GAP	GTPase-activating protein
GEF	GTP exchange factor
GTP	Guanosine triphosphate
hnRNP	heterogeneous nuclear ribonucleoprotein
HRP	Horseradish peroxidase
i/eCLIP	individual-nucleotide resolution/enhanced cross-linking and immunoprecipitation
IP	Immunoprecipitation
LFC	Log ₂ fold change
lincRNA	long intergenic non-coding RNA

lncRNA	long non-coding RNA
m ⁷ G	7-methylguanosine
miRNA	microRNA
mRNA	messenger RNA
ncRNA	non-coding RNA
NES	Nuclear export signal
NLS	Nuclear localisation signal
NMD	Nonsense-mediated decay
NPC	Nuclear pore complex
Nups	Nucleoporins
ORF	Open reading frame
osTIR1	TRANSPORT INHIBITOR RESPONSE 1 from <i>Oryza Sativa</i>
PCR	Polymerase chain reaction
PHAX	Phosphorylated adaptor for RNA export
Poly(A)	Poly(adenine)
pre-mRNA	precursor mRNA
qRT-PCR	Reverse transcription and quantitative PCR
Ran	Ras-related nuclear protein
RBP	RNA-binding protein
RIP	RNA-immunoprecipitation
RNA	Ribonucleic acid
RNAi	RNA-interference

RNA pol II/III	RNA polymerase II/III
rpm	revolutions per minute
rRNA	ribosomal RNA
scaRNA	small Cajal body-specific RNA
SEM	Standard error of the mean
shRNA	small/short hairpin RNA
siRNA	small interfering RNA
snoRNA	small nucleolar RNA
snRNA	small nuclear RNA
snRNPs	small nuclear ribonucleoproteins
TREX	Transcription and export complex
tRNA	transfer RNA
(v/v)	volume/volume
(w/v)	weight/volume
XpoT	Exportin-T
Xpo5	Exportin-5

Chapter 1 - Introduction

In eukaryotes, ribonucleic acid (RNA) is produced by transcription involving RNA polymerases in the nucleus of the cell. Many transcripts are then processed to become mature and functional RNAs. There are many classes (or biotypes) of RNA in the human cell and some require export from the nucleus to the cytoplasm for further processing or to carry out their function. This requires a variety of export factors and receptors. The majority of RNAs are exported by exportin proteins, members of the karyopherin protein family, and the RanGTP pathway (Güttler & Görlich 2011). However, additional proteins need to associate for efficient export to occur, such as PHAX for small nuclear RNA (snRNA) export (Ohno et al. 2000). Messenger RNA (mRNA) is exported by the export receptor Nxf1 in a Ran-independent manner. This process requires the Transcription and Export complex (TREX) to deliver the mRNA to Nxf1 in a “handover” event (Heath et al. 2016). Some RNA classes are retained in the nucleus, such as small Cajal body-specific RNAs (scaRNAs) and some long non-coding RNAs (lncRNAs). However, these RNAs still require localisation to their site of function in the nucleus. Some export pathways have been linked to nuclear localisation of RNAs as well as the abundant heterogeneous nuclear ribonucleoprotein U (hnRNP U). RNAs have such diverse and important roles in the cell, so it follows that they are comprehensively surveyed for defects and their levels regulated throughout their lifetime. The known pathways of RNA export, localisation and surveillance will be discussed in this chapter which will lead into the research undertaken for this thesis.

1.1 The nuclear pore

Active and passive export/import across the nuclear membrane occurs through the Nuclear Pore Complex (NPC), thought to be between 60 and 120 kDa depending on the species (Cronshaw et al. 2002)(Figure 1.1). NPCs traverse the nuclear membrane and are made up of 29 core nucleoporins (Nups) and around 18 NPC-associated proteins. Two thirds of these proteins are conserved between yeast and rats, highlighting the essentiality of the complex (Cronshaw et al. 2002). There are different classes of Nups in the NPC, some that provide a scaffold function and some that are essential for export (Beck & Hurt 2017). Nups containing repeats of phenylalanine and glycine (FG Nups) are present in the middle of the pore and can bind to many transport proteins, such as karyopherins and Nxf1, due to their flexible nature. This facilitates the translocation of transport proteins and their cargo through the NPC (Fribourg et al. 2001; Bayliss et al. 2002; Beck & Hurt 2017). On the cytoplasmic side of the NPC, flexible protein domains called cytoplasmic fibrils bring cytoplasmic proteins into close proximity with the nuclear pore (Beck & Hurt 2017). On the nucleoplasmic side, there is a nuclear “basket” that contains a variety of Nups and is influential in maintaining the NPC structure (Goldberg & Allen 1992; Meszaros et al. 2015). Despite the substantial conservation of NPCs, they are thought to be dynamic structures in controlling import and export as associated proteins and fibrils can move throughout the NPC (Tran & Wente 2006). Efficient export and import of protein and RNA is essential for cellular function, and the different methods of export are detailed below.

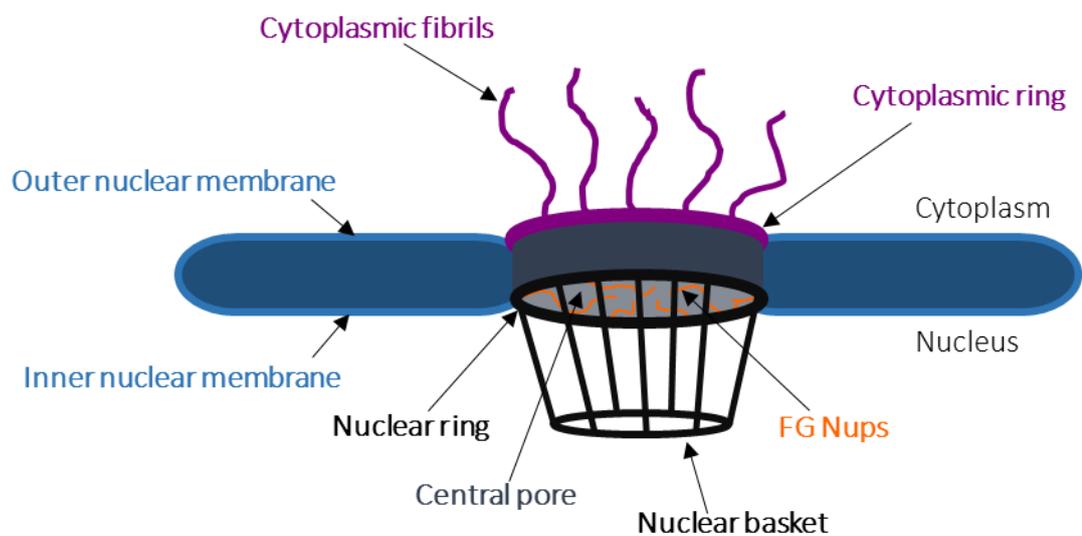


Figure 1.1 The structure of the Nuclear Pore Complex (NPC)

A diagram summarising the main features of the NPC. The NPC is composed of at least 29 core nucleoporin proteins (nups) and additional NPC-associated proteins. These proteins form the central pore, the nuclear basket and cytoplasmic fibrils which stem from the nuclear and cytoplasmic rings respectively. Phenylalanine and glycine repeat nups (FG Nups) are present in the middle of the central pore to facilitate transport of export/import proteins and their cargo through the NPC.

1.2 Ran-dependent RNA export

1.2.1 *The Ran cycle*

Most nuclear export and import systems in the eukaryotic cell use the Ras-related nuclear protein (Ran) and its inherent GTPase activity to drive export or import (Moore & Blobel 1993). A set of over 20 proteins that translocate the membrane and export or import cargo are called exportins and importins respectively and are collectively termed karyopherins (Görlich et al. 1997). All karyopherins are members of the Importin family, which are sequentially related to Importin β , and interact with Ran to facilitate translocation of their cargo (Figure 1.2.1a) (Moore & Blobel 1993; Görlich et al. 1997; Güttler & Görlich 2011). Importins/exportins generally have a specific type of cargo or a subset of cargos that they import or export. However, it seems likely that their cargoes overlap as they use the same RanGTP gradient (Izaurralde et al. 1997; Güttler & Görlich 2011). There is an asymmetric distribution of RanGTP between the nucleus and the cytoplasm with over 1000 times more in the nucleus, to maintain the import-export cycle shown in Figure 1.2.1a (Izaurralde et al. 1997).

In the nucleus, exportins associate with RanGTP and their cargo which contains a Nuclear Export Signal (NES). A NES can be a variety of sequences or peptides such as leucine-rich repeats in proteins recognised by Crm1/Exportin-1 (Güttler & Görlich 2011). The exportin-cargo-RanGTP complex translocates to the cytoplasm through the NPC and dissociates upon the hydrolysis of Ran's GTP to GDP initiated by the GTPase activating protein (GAP) (Kutay et al. 1997). RanGAP is associated with cytoplasmic fibrils from the NPC to efficiently release the cytoplasmic cargo (Izaurralde et al. 1997). The exportin can then translocate back into the nucleus to pick up new cargo. RanGDP also translocates back where it can be turned into RanGTP, catalysed by the GTP exchange factor (GEF/RCC1) to export further cargo alongside exportins.

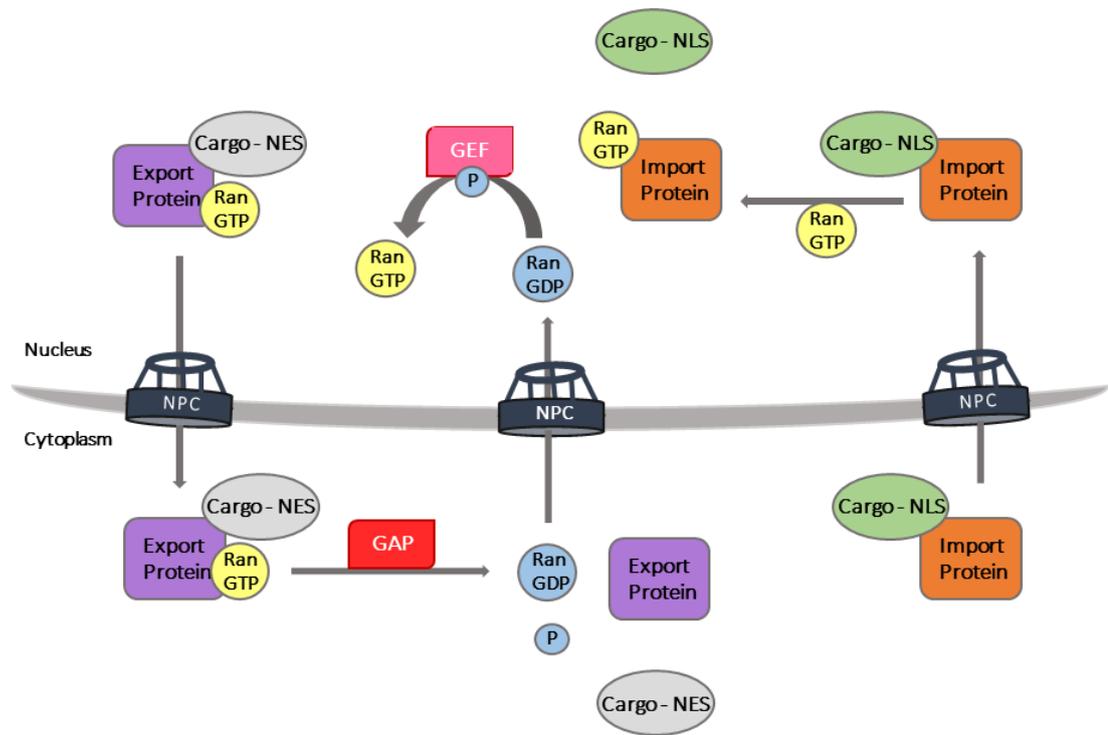


Figure 1.2.1a The RanGTP cycle is necessary for the export and import of various cargoes and the recycling of components for export and import to continue

A diagram of the RanGTP cycle showing how cargo can be exported and imported. Cargo containing a nuclear export signal (NES) is bound to an export protein associated with RanGTP and travels to the cytoplasm through the nuclear pore complex (NPC). The GTP associated with Ran is hydrolysed to GDP by the GTPase activating protein (GAP) which causes release of the cargo. RanGDP returns to the nucleus where guanidine exchange factor (GEF) converts the RanGDP to RanGTP. Cargo with a nuclear localisation signal (NLS) is imported by its associated import protein which releases its cargo once RanGTP binds on the nucleoplasmic side. The import protein associated with RanGTP returns to the cytoplasm where GAP hydrolyses RanGTP to RanGDP which then dissociates from the import protein.

Importins associate with their cargo containing a Nuclear Localisation Signal (NLS) in the absence of RanGTP in the cytoplasm before translocating through the NPC (Melchior 2001; Güttler & Görlich 2011). The association of RanGTP on the nucleoplasmic side causes the complex to dissociate, releasing the cargo (Floer et al. 1997). The importin-RanGTP complex can then return to the cytoplasm and be hydrolysed to RanGDP by GAP to be recycled and import further cargoes (Güttler & Görlich 2011). The export of tRNA, snRNA, miRNA and rRNA all require RanGTP but with different karyopherins and export adaptors associated (Figure 1.2.1b).

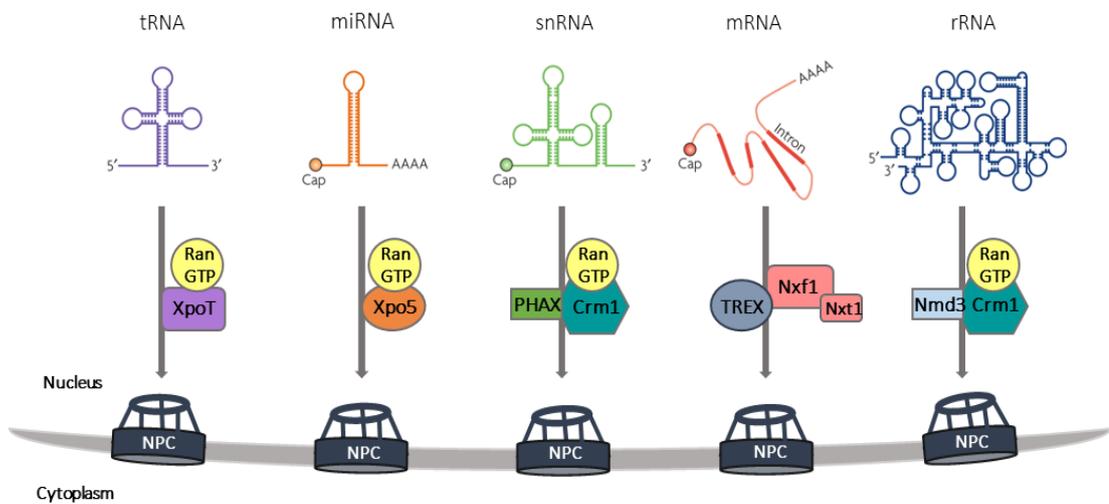


Figure 1.2.1b The nuclear export of RNA classes requires specific export proteins

This figure is adapted from Figure 1 in Kohler and Hurt 2007.

A simplified diagram of the export pathways of the main cellular RNAs. The proteins required to export each type of RNA are shown on the grey arrow. XpoT/5 stands for Exportin-T and Exportin-5 respectively. NPC represents the Nuclear Pore Complex.

1.2.2 tRNA processing and export

Transfer RNA (tRNA) is one of the most abundant classes of RNA in the cell and there are predicted to be over 600 tRNA genes in the human genome (Chan & Lowe 2009). tRNAs are essential for translation to decode mRNA and form amino acid chains which will eventually form a protein through secondary, tertiary and quaternary structures. The crucial nature of tRNA function goes some way in explaining why there are over 600 tRNA genes but only 20 encoded amino acids. There are also 22 tRNAs encoded in the mitochondrial genome for use in mitochondrial-based protein synthesis and mutations in these genes are associated with disease (Suzuki et al. 2011).

tRNA transcripts undergo processing in the nucleus before being exported to the cytoplasm to carry out their function (Figure 1.2.2) (Chatterjee, Nostramo, et al. 2017). Processing does not necessarily have a defined order, but all steps are intimately linked. Following transcription by RNA polymerase III (RNA pol III), the 5' and 3' trailing ends of the pre-tRNA are cleaved by RNAses and a CCA trinucleotide is added to the 3' end (Hopper & Phizicky 2003; Phizicky & Hopper 2010). tRNA nucleosides are modified in a variety of ways, such as the addition of methyl groups, and there is usually between 11 and 13 modifications per tRNA (Saikia et al. 2010). The functions of these modifications are starting to be revealed and they have far-reaching effects on translation and homeostasis in the cell (Schimmel 2018). A small number of human tRNAs contain introns, approximately 6% (Lowe & Eddy 1997), which are spliced out by the SEN complex and the tRNA is re-ligated by the hRTCB ligase complex (De Robertis et al. 1981; Lund & Dahlberg 1998; Paushkin et al. 2004; Popow et al. 2011; Popow et al. 2014). Splicing of tRNA introns is coupled with end-processing of the tRNA, all of which are must occur correctly before export to the cytoplasm (Lund & Dahlberg 1998; Paushkin et al. 2004). Conversely in yeast, pre-tRNA splicing occurs on the surface of the mitochondria, so unspliced tRNA is exported (Yoshihisa et al. 2007; Phizicky & Hopper 2010; Huang & Hopper 2015).

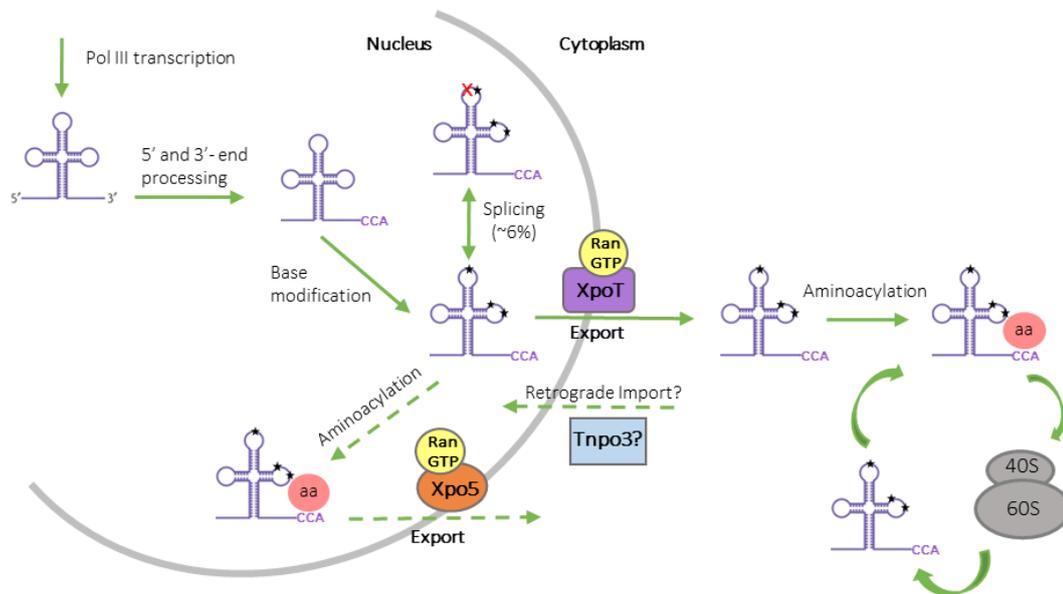


Figure 1.2.2 A diagram of the human tRNA processing and export pathway

This figure is adapted from the yeast pathway shown in Phizicky and Hopper, 2010.

tRNAs are transcribed by RNA polymerase III (RNA Pol III) and processed in the nucleus. The removal of tRNA introns in the anticodon loop, splicing, is represented by the red cross. The proteins involved in export are shown on the green arrows. The dashed arrows indicate minor or unresolved steps. aa signifies the addition of an amino acid to the tRNA before use in translation at the 80S ribosome.

In vertebrates, tRNAs are widely considered to be exported from the nucleus to the cytoplasm by Exportin-T (XpoT), a conserved member of the Importin- β family of proteins (Kutay et al. 1998; Gert Jan Arts et al. 1998). Yeast and plant orthologues exist as Los1 and PAUSED (Hopper et al. 2011). As a karyopherin, XpoT is dependent on RanGTP binding to export tRNA through the nuclear pore (Gert Jan Arts et al. 1998). The tRNA is released in the cytoplasm following GTP hydrolysis as shown in Figure 1.2.1a (Kutay et al. 1998). Correct end-processing and modification of tRNAs must occur before export is licensed by XpoT as demonstrated by the XpoT-tRNA structure in yeast (Lipowsky et al. 1999; Cook et al. 2009).

Some tRNA species can be exported via another importin- β family member, Exportin-5 (Xpo5), although this is believed to be a minor export pathway in comparison to XpoT (Calado et al. 2002; Bohnsack et al. 2002). Xpo5 was initially discovered as an exporter of double-stranded RNA binding proteins and is the main export pathway for tRNAs in *Drosophila* as there is no known XpoT orthologue (Lippai et al. 2000; Brownawell & Macara 2002; Shibata et al. 2006). Xpo5 is more commonly known as the precursor micro RNA (pre-miRNA) exporter which are further processed in the cytoplasm and loaded into the RNA-induced silencing complex (RISC) where it can “silence” RNA expression (Lund et al. 2004; Bohnsack et al. 2004; Okamura et al. 2015). pre-miRNA precursors are 5' m⁷G capped and polyadenylated before export, much like mRNAs (Cai et al. 2004). Additionally, Xpo5 has recently been implicated in exporting the 60S ribosomal subunit to the cytoplasm (Wild et al. 2010). tRNAs with 5' m⁷G caps have recently been found in yeast and HeLa cells so perhaps Xpo5 works to export 5' capped tRNAs specifically as XpoT does not export capped tRNA (Ohira & Suzuki 2016; Cook et al. 2009). Interestingly, Xpo5 seems to preferentially bind processed tRNAs that have already been aminoacylated (Mingot et al. 2013; Huang & Hopper 2015).

Amino acids need to be attached to tRNAs, aminoacylation, before they can function in translation. This is generally thought to occur in the cytoplasm as the majority of aminoacyl-tRNA synthetases are present there (Kaminska et al. 2009). However, some tRNAs may be aminoacylated in the nucleus as some aminoacyl-tRNA synthetases and associated complexes are present in the nucleus (Lund & Dahlberg

1998; Galani et al. 2001; Kaminska et al. 2009). Aminoacyl-tRNAs seem unlikely to be bound by XpoT as its binding to tRNA is dependent on the recognition of the 3' CCA which may be inhibited by the presence of an amino acid (Cook et al. 2009). Therefore, Xpo5 could provide the alternative function of exporting tRNAs that have been aminoacylated in the nucleus.

tRNAs are also thought to be imported back into the nucleus through “retrograde transport” by members of the importin- β family, Mtr10 in yeast and its mammalian orthologue Transportin-3 (Tnp3) (Zhou et al. 2011; Chatterjee, et al. 2017). Evidence for retrograde tRNA transport in mammals has come from HIV studies where Tnp3 was found to bind tRNA and translocate to the nucleus (Zaitseva et al. 2006; Zhou et al. 2011). Additionally, retrograde transport of tRNAs was shown to increase during stress conditions, when cells from rats were starved (Shaheen et al. 2007). However, this observation is hotly contested (Chafe et al. 2011). Perhaps a retrograde pathway would serve to regulate translation in times of stress or provide another quality control step. Xpo5 may then re-export aminoacylated tRNAs when required. A direct interaction between Mtr10 and tRNA has not yet been demonstrated and studies in vertebrate cells needs to go a lot further to define a potential tRNA retrograde transport pathway (Huang & Hopper 2015).

Despite XpoT being considered the main pathway for tRNA export in vertebrates, the *XpoT* gene is not essential in human cell lines, supporting the role of Xpo5 in tRNA export and indicating that other pathways are likely to exist (Blomen et al. 2015; T. Wang et al. 2015; Hart et al. 2015). In yeast, the Crm1/Exportin-1 pathway was identified as a potential tRNA export pathway through a genome-wide mutant screen. *crm1* mutants accumulated end-processed but intron-containing tRNAs, similarly to *los1* mutants, indicative of a block in export as splicing on the mitochondrial face has not occurred (Wu et al. 2015). Mutants also accumulated unspliced tRNAs at the nuclear periphery as seen by FISH. The Mex67-Mtr2 mRNA export pathway was also identified as a potential tRNA export pathway as mutants gave the same tRNA accumulation phenotype (Wu et al. 2015). This observation is now further supported by an additional study, demonstrating nuclear accumulation of a subset of intron-containing tRNAs following *Mex67* knockout (Chatterjee,

Majumder, et al. 2017). Additionally, interaction between Mex67 and tRNA was shown and overexpression of Mex67-Mtr2 in *los1* mutant cells rescued the accumulation phenotype (Chatterjee, Majumder, et al. 2017). The mRNA export pathway may therefore export a specific subset of tRNAs, which could be similar to the role that Crm1 has in tRNA export.

1.2.3 *The Crm1/Exportin-1 export pathway*

Another karyopherin, Crm1 or Exportin-1, is responsible for the export of a variety of cargoes. Crm1 binds to its cargo through adaptor proteins specific for the cargo and associates with RanGTP before translocation to the cytoplasm (Figure 1.2.1b).

Following the discovery of the Nuclear Export Signal (NES), a short peptide sequence containing leucine residues, Crm1 was identified as the protein necessary for the efficient export of NES-containing proteins through recognition of the NES sequence (Wen et al. 1995; Fischer et al. 1995; Fornerod et al. 1997; Fukuda et al. 1997; Stade et al. 1997). Crm1 also exports Snurportin-1 so it can import further small nuclear ribonucleoproteins (snRNPs) back into the nucleus to carry out their function in RNA splicing (Huber et al. 1998; Paraskeva et al. 1999). The export of some RNA species is also carried out by Crm1, such as ribosomal RNA (rRNA) with its ribosomal subunit and small nuclear RNAs (snRNA).

Small nuclear RNAs (snRNA) are approximately 150 nucleotides long with a 5' m⁷G cap. snRNAs are transcribed by RNA polymerase II and III and function in intron splicing as part of the spliceosome (Henry et al. 1998; Will & Lührmann 2001; Wang et al. 2017). The most common RNA pol II transcribed snRNAs, U1, U2, U4, and U5 are exported to the cytoplasm where they are assembled into snRNP protein complexes. snRNPs are reimported into the nucleus by Snurportin-1 and further modified and assembled in the Cajal bodies (Huber et al. 1998; Will & Lührmann 2001; Tycowski et al. 2004; Jády et al. 2003). There are many less common snRNAs, such as U4atac and U12, which can also function in snRNPs. Subfamilies of snRNAs perform other functions in the cell such as U3 small nucleolar RNA (snoRNA) in rRNA processing (Kass et al. 1990; Kiss 2002; Boulon et al. 2004).

The 5' cap plays an essential role in snRNA export following association of the Cap-Binding Complex (CBC) (Izaurralde et al. 1995). The presence of the CBC allows Crm1 to associate with snRNAs, mediated through the adaptor protein Phosphorylated Adaptor for RNA export (PHAX) (Ohno et al. 2000; Segref et al. 2001; Fischer et al. 1995; Izaurralde et al. 1995; Fornerod et al. 1997; Kitao et al. 2008). The phosphorylation of PHAX and the association of RanGTP is also required for complex formation (Ohno et al. 2000; Segref et al. 2001; Kitao et al. 2008). Following export,

Crm1 is recycled after RanGTP hydrolysis and PHAX is imported following dephosphorylation (Ohno et al. 2000; Kitao et al. 2008).

Crm1 is also the main exporter of ribosomal subunits, 60S and 40S, containing rRNA and associated proteins (Ho et al. 2000; Thomas & Kutay 2003). 60S export requires the NES-containing Nmd3 export adaptor, as well as other export adaptors such as the Mex67-Mtr2 heterodimer in yeast (Ho et al. 2000; Thomas & Kutay 2003; Yao et al. 2007). Mex67-Mtr2 has also been implicated in 40S export alongside other adaptors (Faza et al. 2012). rRNA biogenesis is generally conserved. RNA polymerase I transcribes a pre-rRNA precursor that is processed into mature 18S, 5.8S and 28S rRNA and RNA pol III transcribes the smallest 5S rRNA. rRNA associates with ribosomal proteins to form the 40S and 60S subunits in the nucleus before export and further maturation in the cytoplasm (Zemp & Kutay 2007). Crm1 has also been associated with the export of a subset of mRNAs that contain AU-rich elements (ARE) alongside adaptor proteins (Brennan et al. 2000).

The PHAX and Crm1 proteins are also needed for intra-nuclear localisation of RNA. PHAX is required for the first localisation step of U3 snoRNA to Cajal bodies and then Crm1 becomes important to further localise U3 to the nucleolus (Boulon et al. 2004). This seems to be the case for all types of snoRNAs, which are not exported during their biogenesis (Pradet-Balade et al. 2011). Human telomerase RNA (*TERC/hTR*) is a key component of the telomerase enzyme that acts as a template for telomere extension (Greider & Blackburn 1989). *TERC* resembles the scaRNA subclass, discussed in Chapter 1.4, and localises to Cajal bodies which is essential for its function in telomere elongation (Zhu et al. 2004; Jady et al. 2004; Cristofari et al. 2007; Venteicher et al. 2009). *TERC* also associates with PHAX but is not exported to the cytoplasm in human cells, suggesting that PHAX may play a role in localising *TERC* to Cajal bodies within the nucleus similarly to snoRNAs (Kiss 2002; Boulon et al. 2004; Pradet-Balade et al. 2011). The Crm1-PHAX export pathway is involved in the intra-nuclear localisation of different types of RNA, indicating that export pathways may not be restricted to transporting RNAs between the nucleus and the cytoplasm but within the nucleus as well.

Therefore, the seemingly distinct export pathways may overlap in the export of some RNA species. Mex67 and Crm1 pathways overlap with regard to ribosomal subunits and mRNAs and both pathways have been implicated in tRNA export (Wu et al. 2015; Chatterjee et al. 2017). Perhaps export pathways can export all cargoes, but show a preference for certain ones to maintain cellular homeostasis.

1.3 Messenger RNA export

Messenger RNA (mRNA) is a well-studied class of RNA, due to their functional importance but despite their low abundance in the cell (FANTOM Consortium 2006). mRNAs are transcribed by RNA polymerase II and are processed by three main processing steps, the first of which is the addition of an m⁷G cap which protects the 5' end of the transcript (Topisirovic et al. 2011). The transcript can undergo splicing to remove introns which generally occurs co-transcriptionally (Girard et al. 2012). Finally, the 3' end of the transcript is cleaved away from the associated polymerase and a poly(A) chain is added to protect the 3' end of the transcript (Proudfoot 2011). These processing steps are deeply linked with transcription so the integrity of the transcript is reliant on all processes functioning correctly (McCracken et al. 1997). If a transcript has not been properly transcribed or processed, it is susceptible to degradation by nuclear and cytoplasmic machinery (Chapter 1.6) (Trcek et al. 2013; Kilchert et al. 2016). mRNAs provide the code by which proteins are made in translation so correct processing of mRNA is essential for the production of functional proteins.

mRNAs are exported to the cytoplasm by the export receptor Nxf1-Nxt1, a heterodimer, aided by the conserved Transcription and Export complex (TREX) (Heath et al. 2016). Unusually, this is a Ran-independent mechanism and relies on the association of a variety of proteins and handover events to transport the mRNA to the cytoplasm. The TREX complex is made up of the stoichiometric Tho complex and additional export factors, adaptors and co-adaptors that dynamically associate throughout the export process (Figure 1.3.1a-b) (Heath et al. 2016). The Alyref protein is the most well studied export adaptor and RNAi of Alyref has led to the discovery of additional export adaptors and demonstrated the redundancy between them. The depletion of Alyref does not lead to a substantial block in mRNA export, however, depletion of Uif or Chtop proteins alongside Alyref results in a major block in the export of mRNA (Hautbergue et al. 2009; Chang et al. 2013). Export adaptors can also be cell-type specific such as Luzp4 in the testes, which complements the export block seen when Alyref and Uif are depleted together (Viphakone et al. 2015).

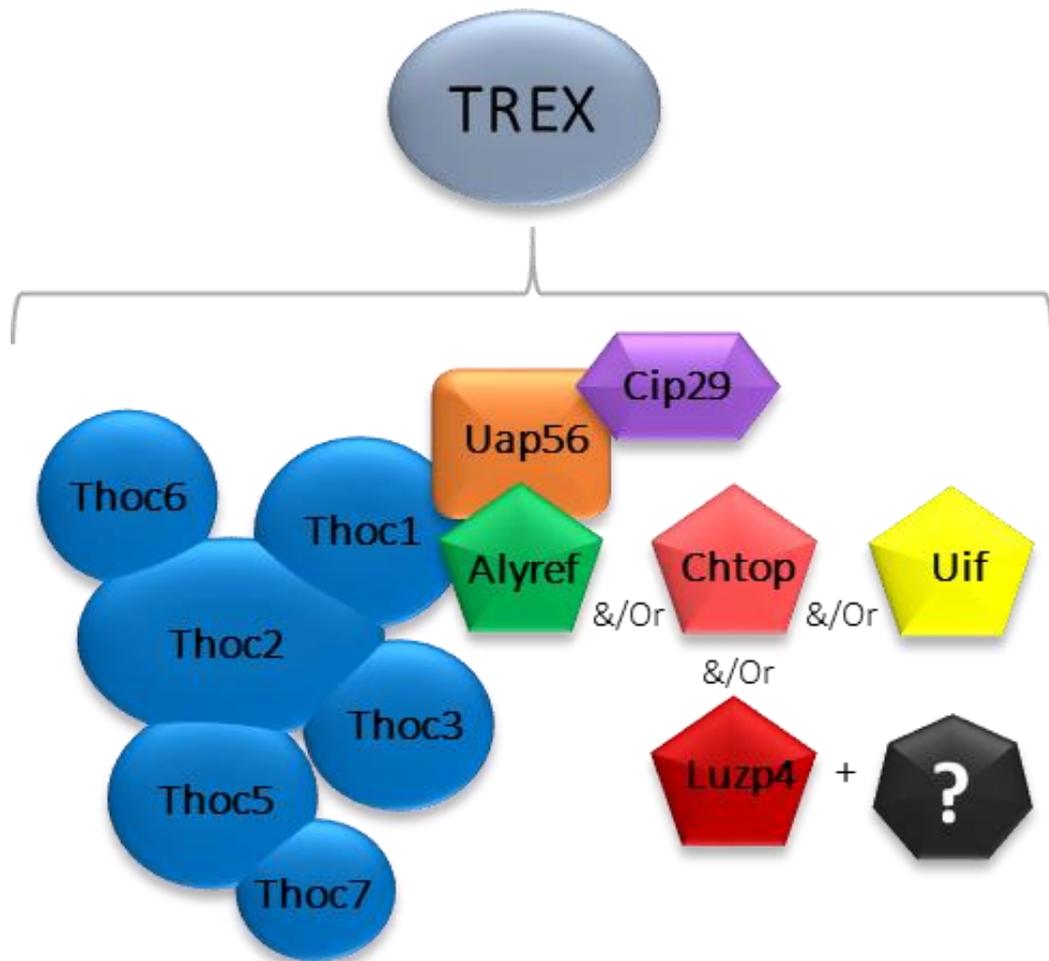


Figure 1.3.1a The structure of the TREX complex

A schematic of the TREX complex made up of the Tho complex (blue) and export adaptors and co-adaptors. TREX is likely to be a dynamic complex with different adaptors associating at different times and in different states. The question mark denotes export factors that are yet to be discovered, but are likely to exist.

TREX subunits		
Protein name	Function in export	Other functions
Uap56/ DDX39B	<i>mRNP assembly factor</i> – ATPase activity “loads” proteins onto mRNA and into the mRNP	- Splicing factor - Associated with EJC
Alyref	<i>Export adaptor</i> – necessary for mRNA handover to Nxf1	- Associated with EJC - Transcriptional co-activator
Uif	<i>Export adaptor</i> – functionally redundant with Alyref in mRNA export	- Associated with EJC
Luzp4	<i>Export adaptor</i> – complements Alyref function in cancer cells	- Associated with EJC
Chtop	<i>Export co-adaptor</i>	- Associated with EJC - PRMT1 binding partner - Transcriptional co-activator
Cip29/ Sarnp	<i>Stimulates Uap56 ATPase/helicase activity</i>	- Cell-cycle protein
Poldip3	<i>TREX-associated</i>	- Cell growth regulator
Zc3h11a	<i>Probable export co-adaptor</i>	

Figure 1.3.1b The function of TREX complex components

A table of known and probable TREX components and their functions.

The DEAD-box helicase Uap56 (or DDX39B) is crucial for spliceosome assembly and mRNA export (Shen et al. 2007; Shen et al. 2008). Following recruitment of Uap56 to the mRNA during splicing, its RNA-dependent ATPase activity allows the recruitment of other TREX components, Alyref and Chtop to the mRNA (Luo et al. 2001; Taniguchi & Ohno 2008). The binding of Uap56 and Alyref to the mRNA occurs in a mutually exclusive manner so the mRNA is handed over to Alyref in the mRNP (Figure 1.3.2) (Chang et al. 2013).

TREX is part of the complex of various proteins surrounding the mRNA, the mRNP, which contains SR proteins that assist in mRNA export and is essential for splicing (Zahler et al. 1992; Huang & Steitz 2001; Huang et al. 2003). As its name suggests, TREX associates with the mRNA during transcription and processing which is thought to provide various quality control steps to prevent export of improper mRNA (Sträßer et al. 2002). TREX was initially thought to only associate with the 5' m⁷G cap of mRNA but it is now considered to associate along the entirety of the mRNA through its associations with splicing factors, the EJC and the 3' end processing complex, mainly through Alyref associations (Cheng et al. 2006; Gromadzka et al. 2016; Chi et al. 2013; Johnson et al. 2009). However, the full extent of TREX binding to mRNA is not completely known but current iCLIP studies (individual-nucleotide resolution cross-linking and immunoprecipitation) will help to more accurately annotate the lifecycle and export of mRNA. Interestingly, the balance between nuclear export and degradation of the mRNA may be controlled by the competition between Alyref or Mtr4, a component of exosome-targeting complexes, binding to the CBC (J. Fan et al. 2017)(Chapter 1.6). If there is plentiful Alyref and TREX binding at this stage, the mRNA is licensed for export.

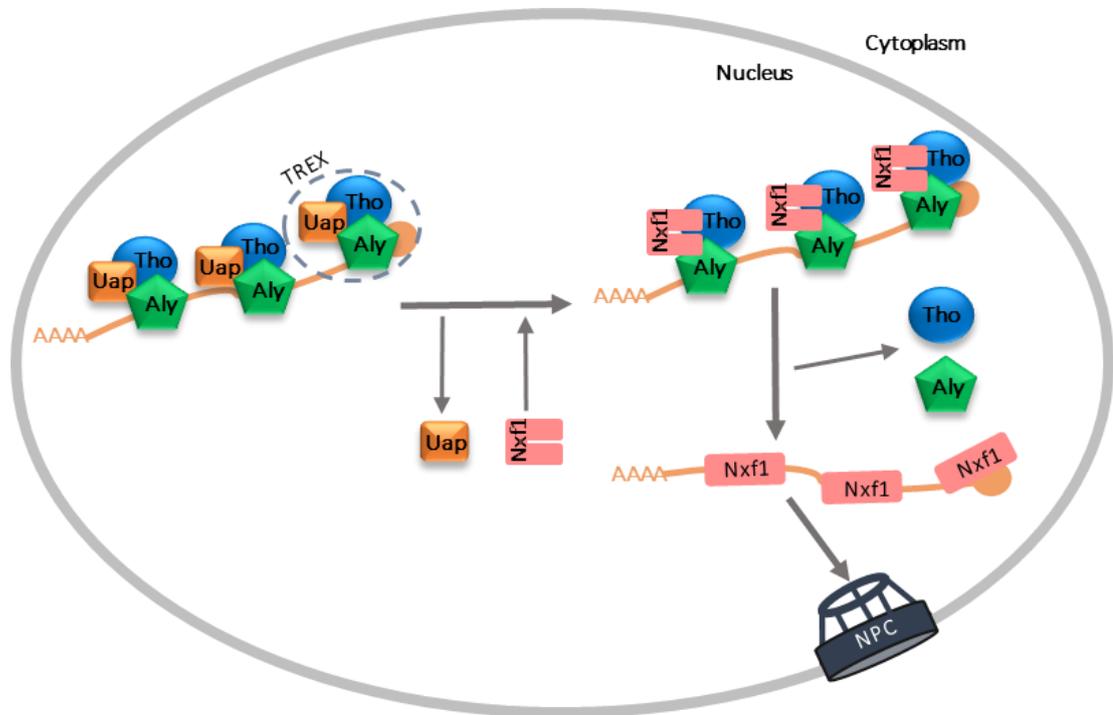


Figure 1.3.2 The dynamic TREX complex is vital for the handover of mRNA to the export receptor Nxf1

The TREX complex is loaded onto the mRNA and into the mRNP by the ATPase Uap56 (Uap). This causes the dissociation of Uap56 from the mRNP and the recruitment of “locked” Nxf1, , and Nxt1 (not shown) via Alyref. Thoc5 and Alyref “unlock” the structure of Nxf1, , allowing Nxf1 to bind mRNA and export it to the cytoplasm through the nuclear pore complex (NPC). TREX components dissociate before complete translocation through the NPC.

Effective mRNA export is dependent on TREX complex components as they handover the RNA to Nxf1-Nxt1 for export (Figure 1.3.2) (Hautbergue et al. 2008; Viphakone et al. 2012). Alyref and Thoc5 bind to Nxf1 and remodel the Nxf1-Nxt1 complex to expose the RNA binding domain (RBD) and RNA recognition motif (RRM) of Nxf1 that is hidden inside the complex (Viphakone et al. 2012). Nxf1 can now bind directly to the RNA. The binding of Alyref and Nxf1 to RNA is mutually exclusive, similarly to Alyref and Uap56 binding. Methylation of the arginines in Alyref reduces its affinity for RNA, allowing Nxf1 to displace Alyref from the mRNP (Hung et al. 2010). This mutually exclusive binding event allows efficient handover of the RNA from TREX to Nxf1-Nxt1 for export (Hautbergue et al. 2008).

Various members of TREX and the mRNP dissociate during translocation and new cytoplasmic factors associate. The RNA helicase Dbp5 (or DDX19B) and Gle1 are associated with the cytoplasmic fibrils of the NPC and cause dissociation of Nxf1 from the mRNP once translocation has occurred (Heath et al. 2016). Nxf1 can then be re-imported and associate with RNA for export once again.

1.4 Additional RNA classes

1.4.1 *Small Cajal body-specific RNA*

Small Cajal body-specific RNAs (scaRNA) function in nuclear structures called Cajal bodies and guide the modification of snRNAs (Jády et al. 2003; Tycowski et al. 2004). Some scaRNAs are capped at the 5' end and are 3' end processed (Tycowski et al. 2004; Enwerem et al. 2015). scaRNAs are targeted to Cajal bodies through a conserved tetranucleotide sequence, termed the CAB box (Richard et al. 2003; Machyna et al. 2014). The constitutive Cajal body protein, Coilin, and additional proteins such as WRAP53 (TCBP1/WDR79) act to localise and maintain scaRNAs at the Cajal body (Tycowski et al. 2009; Enwerem et al. 2015). scaRNAs do not seem to be exported to the cytoplasm at any stage, but require intra-nuclear pathways to localise them to the Cajal bodies.

1.4.2 *Long non-coding RNA*

Long non-coding RNAs (lncRNA) are a relatively new class of RNA and their individual functions in the cell are extremely diverse and tissue specific (Cabili et al. 2011). There are thought to be over 9000 lncRNAs in the human genome (The ENCODE Project Consortium 2012). lncRNAs are over 200 nucleotides long, transcribed by RNA polymerase II and are generally processed similarly to mRNAs although not normally translated (Banfai et al. 2012; Cech & Steitz 2014). However, some lncRNAs are found associated with ribosomes, presumably directed there due to their mRNA-like appearance (Ingolia et al. 2011; Carlevaro-Fita et al. 2016). Many lncRNAs are nuclear restricted but can be localised throughout the cell. lncRNAs can act as scaffolds, regional organisation factors and guides for cellular processes among many other functions (Engreitz et al. 2016).

The scaffold functions of lncRNAs are mediated through their capacity to bind to multiple proteins through their distinct RNA domains (Guttman & Rinn 2012; Engreitz et al. 2016). This phenomenon is notably seen with the lncRNA *XIST*. *XIST* RNA is essential for establishing X-chromosome inactivation in female mammals which is required for dosage compensation between the sexes (Cerase et al. 2015). *XIST* RNA coats the soon-to-be inactive X-chromosome and drives heterochromatin

formation by recruiting histone methylation complexes, such as polycomb repressor complex 2 (PRC2) which introduces H3K27me3 marks on the chromatin (Brown et al. 1992; Clemson et al. 1996; Heard et al. 2001; Peters et al. 2002; Mermoud et al. 2002; Plath 2003; Silva et al. 2003; Kohlmaier et al. 2004). This results in almost complete transcriptional silencing and generation of the Barr body (Barr & Carr 1962; Daly et al. 1977). The silencing activity of *XIST* RNA is mediated by its interactions with different proteins through its distinct, modular domains which function independently of one and other (Wutz et al. 2002). This includes binding to PRC2, SPEN, hnRNP K, hnRNP U as well as a range of DNA binding proteins which are brought together by the function of *XIST* as an RNA scaffold (Wutz et al. 2002; Hasegawa et al. 2010; Rinn & Chang 2012; Monfort et al. 2015; McHugh et al. 2015; Chu et al. 2015; Engreitz et al. 2016).

The lncRNA *HOTAIR* also demonstrates a scaffold function (Engreitz et al. 2016). *HOTAIR* regulates gene expression from the *HOXD* locus to ensure proper cellular differentiation and patterning during embryonic development (Rinn et al. 2007). *HOTAIR* associates with two complementary histone modification complexes, PRC2 and the LSD1/CoREST/REST complex, through its 5' and 3' ends respectively, leading to transcriptional silencing of the *HOXD* locus (Rinn et al. 2007; Tsai et al. 2010). Interestingly, increased *HOTAIR* expression can lead to greater invasiveness and metastasis of breast cancer so can act as a potential biomarker for cancer development (Gupta et al. 2010; Rinn & Chang 2012). *XIST* and *HOTAIR*, alongside many other lncRNAs, effectively demonstrate the emerging functions of lncRNAs as RNA scaffolds that coordinate cellular processes.

The lncRNA *FIRRE* is expressed from the X-chromosome and acts as a regional organisation factor in the chromatin (Hacisuleyman et al. 2014; Engreitz et al. 2016). *FIRRE* contains multiple copies of an RNA-Repeating Domain (RRD) which allows it to interact with other organisational factors throughout the chromatin such as the protein hnRNP U (Hacisuleyman et al. 2014; Hacisuleyman et al. 2016). *FIRRE* can therefore interact with proteins associated with loci on different chromosomes and bring genes into spatial proximity with each other. This activity is seen with genes involved in adipogenesis being brought together by *FIRRE*, presumably allowing the

coordinated regulation of this cellular process (Sun et al. 2013; Hacısuleyman et al. 2014; Hacısuleyman et al. 2016; Engreitz et al. 2016). The depletion of *FIRRE* from mammalian cells results in the dissociation of these loci from each other and the inhibition of adipogenesis (Sun et al. 2013; Hacısuleyman et al. 2014). Other lncRNAs are likely to provide similar roles in regional organisation as many contain repetitive sequences which could act to bring spatially separated loci together to coordinate gene expression.

The highly expressed lncRNAs *NEAT1* and *MALAT1/NEAT2* demonstrate additional roles of lncRNAs in the cell. Both *NEAT1* and *MALAT1* are associated with actively transcribed genes throughout the genome but are enriched at the transcriptional start site (TSS) and the transcriptional termination site (TTS) respectively (West et al. 2014; Engreitz et al. 2014; Engreitz et al. 2016). *MALAT1* also localises to speckles, which are nuclear bodies that contain numerous splicing and RNA processing factors (Hutchinson et al. 2007; Spector & Lamond 2011). As *MALAT1* is associated with both transcriptionally active genes and nuclear speckles, *MALAT1* may provide a link between the transcriptional machinery and the nuclear speckles for efficient RNA processing to occur (Engreitz et al. 2016). In this regard, *MALAT1* influences alternative splicing of pre-mRNA through the regulation of SR proteins in the nuclear speckles (Tripathi et al. 2010)

NEAT1 on the other hand is essential for the formation of paraspeckles. Paraspeckles are dynamic nuclear bodies that contain proteins involved in transcription and RNA processing and act to retain heavily modified mRNAs in the nucleus (Clemson et al. 2009; Sunwoo et al. 2009; Fox & Lamond 2010; Mao et al. 2011). *NEAT1* initiates the formation of paraspeckles and recruits proteins such as paraspeckle component 1 (PSP1) and the splicing factor PSF to form the nuclear body (Clemson et al. 2009; Mao et al. 2011). The depletion of *NEAT1* from HeLa cells results in the dispersal of paraspeckle proteins and the loss of paraspeckle foci (Clemson et al. 2009). *NEAT1* therefore acts as an indispensable architectural RNA to establish the formation of paraspeckles by acting as a scaffold to recruit the relevant proteins (Clemson et al. 2009; Sunwoo et al. 2009).

Through these scaffold and organisational functions, lncRNAs are heavily involved in maintaining cellular homeostasis, making them crucial to the cell. There are likely to be many more functional lncRNAs to be discovered and annotated, which will give us more understanding of the diverse roles these RNAs play in the cell.

1.5 hnRNP U

RNAs require localisation to a particular site in the cell to carry out their function. Alongside annotated export pathways, individual proteins or specific protein complexes may be involved in directing and localising these RNAs. Heterogeneous nuclear ribonucleoprotein U (hnRNP U) is necessary for the localisation of specific non-coding RNAs to the chromatin, which could reveal a global mechanism for non-coding RNA localisation if investigated further. hnRNP U, and theoretically other similar proteins, may act to localise and retain RNAs at their functional cellular location on a universal scale.

hnRNP U is a member of the heterogeneous nuclear ribonucleoprotein (hnRNP) family that consists of at least 16 main RNA-binding proteins (RBPs). hnRNPs have a variety of functions in the cell, from transcription and RNA biogenesis to translation (Geuens et al. 2016). The hnRNP U protein itself has a range of functions, facilitated by its collection of domains that allow it to bind DNA, RNA and proteins such as RNA polymerase II (Figure 1.5) (Dreyfuss et al. 1984; Kiledjian & Dreyfuss 1992; Fackelmayer et al. 1994; Kim & Nikodem 1999; Kukalev et al. 2005). hnRNP U also contains an ATP or GTP hydrolysis domain. Unpublished data from the lab suggests that it preferentially binds GDP and shows no or little hydrolysis activity. However, it has been published that hnRNP U has ATPase activity induced by the presence of RNA (Nozawa et al. 2017). This ATPase activity seems to be necessary for hnRNP U oligomerisation which was observed during the protein's initial discovery (Romig et al. 1992; Nozawa et al. 2017). Whether this ATPase activity has a function outside of oligomerisation of hnRNP U itself remains to be seen.

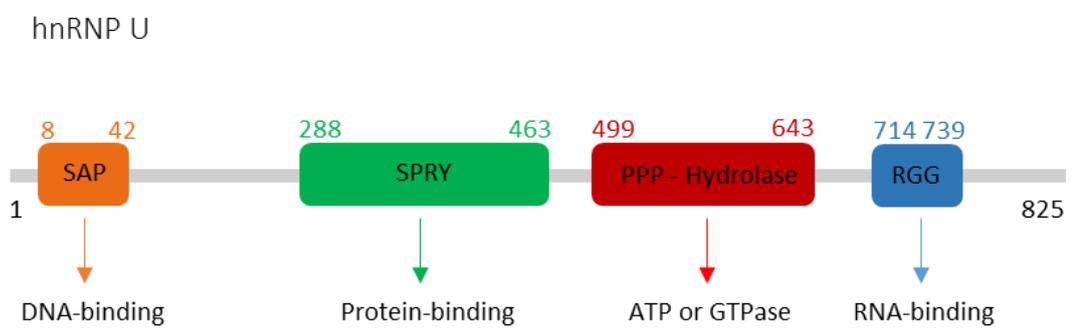


Figure 1.5 The identified protein domains of hnRNP U

This figure is adapted from Hall et al, 2014. A schematic of the domain structure of hnRNP U, including the amino acid numbers and a brief description of the annotated domains.

1.5.1 hnRNP U functions

Localisation of RNA

Recent studies of hnRNP U have demonstrated a key role of hnRNP U in the lifecycle of lncRNAs. An eCLIP study found hnRNP U preferentially bound to lncRNAs and the expression of lncRNAs is particularly affected by the depletion of hnRNP U in the mouse heart (Xiao et al. 2012; Ye et al. 2015). lncRNA *FIRRE* is retained in the nucleus by the interaction between its multiple RNA-Repeating Domains (RRD) and hnRNP U (Hacisuleyman et al. 2014; Hacisuleyman et al. 2016). Additionally, other transcripts containing an RRD are also retained in the nucleus by hnRNP U, suggesting that this may be a global mechanism in RNA retention coordinated by hnRNP U (Hacisuleyman et al. 2016).

The localisation of the lncRNA *XIST* has been heavily linked to hnRNP U. hnRNP U is necessary to bind and localise *XIST* RNA to the soon-to-be inactive X-chromosome so silencing can occur (Helbig & Fackelmayer 2003; Pullirsch et al. 2010; Hasegawa et al. 2010; Smeets et al. 2014; McHugh et al. 2015). hnRNP U depletion results in *XIST* RNA becoming diffuse throughout the nucleus rather than being localised to the X-chromosome, as determined by FISH, which prevents the chromosome being inactivated (Hasegawa et al. 2010). hnRNP U eCLIP data shows hnRNP U bound preferentially to the exons of *XIST*, further demonstrating its role with the mature and functional form of *XIST* (The ENCODE Project Consortium 2012). Other proteins that interact with *XIST* for efficient X-chromosome silencing do not seem to be involved in *XIST* localisation specifically, making hnRNP U unique in this function. For example, SPEN is needed to recruit PRC2 that is necessary to maintain X-chromosome inactivation (Monfort et al. 2015; McHugh et al. 2015).

hnRNP U may also be necessary to localise the lncRNA *PANDA* to cell senescence genes to induce transcriptional silencing during proliferation (Puvvula et al. 2014). The silencing of senescence-promoting genes is coordinated by the recruitment of Polycomb Repressor Complexes (PRCs), which is likely to be coordinated by hnRNP U and *PANDA*, much like X-chromosome silencing through *XIST* (Hasegawa et al. 2010; Puvvula et al. 2014; McHugh et al. 2015; Moindrot et al. 2015).

Repetitive RNA that is associated with chromatin is also localised by hnRNP U. C_oT-1 RNA is made up of repetitive sequences such as LINE1 (L1) and is normally localised to foci of euchromatin (Hall et al. 2014). However, following overexpression of a dominant negative mutant of hnRNP U, the C_ot-1 RNA signal becomes dispersed, much like *XIST* mislocalisation after hnRNP U depletion (Hall et al. 2014; Hasegawa et al. 2010).

The role of hnRNP U in the localisation of lncRNAs *FIRRE*, *XIST*, *PANDA* and repetitive C_ot-1 RNA points toward a general mechanism of non-coding RNA localisation to chromatin directed by hnRNP U. This mechanism may extend to other RNAs as well as overexpression of hnRNP U has been shown to retain HIV1 mRNA in the nucleus (Valente & Goff 2006).

Nuclear and chromatin structure

hnRNP U is also known as Scaffold-Attachment Factor A (SAF-A). SAF-A was discovered as a key component of the nuclear scaffold and matrix structures as it binds to scaffold-attachment region (SAR) DNA in chromatin (Romig et al. 1992; Fackelmayer et al. 1994; Göhring & Fackelmayer 1997). These scaffolds maintain the architecture of the nucleus and so hnRNP U is cleaved during apoptosis to break the link between DNA and proteins to effectively collapse the cell (Göhring et al. 1997; Kipp et al. 2000). More recently, hnRNP U has been associated with coordinating chromatin structure with the help of chromatin-associated RNAs in specific regions and globally throughout the genome (Nozawa et al. 2017; H. Fan et al. 2017). Both studies show hnRNP U as a promoter of chromatin decompaction and a regulator of various 3D-chromatin structures.

Gene expression

hnRNP U has been implicated in transcription, and more specifically in transcription elongation. hnRNP U is found at the promoters of actively transcribed genes and is associated with the C-terminal domain (CTD) of RNA polymerase II, signifying a role in transcription (Carty & Greenleaf 2002; Kim & Nikodem 1999; Kukalev et al. 2005; Obrdlik et al. 2008; H. Fan et al. 2017). hnRNP U is part of various complexes that

regulate transcription elongation. In complex with TFIIF, hnRNP U was shown to inhibit elongation by preventing phosphorylation of the RNA pol II CTD (Kim & Nikodem 1999). However, in complex with Actin and the histone-acetyltransferase PCAF, hnRNP U is needed for active transcription elongation through association with the phosphorylated form of the RNA pol II CTD (Kukalev et al. 2005; Obrdlik et al. 2008). Additionally, hnRNP U is involved in inhibiting the transcriptional activity of the Wilm's Tumour 1 (WT1) transcription factor (Spraggon et al. 2007). Generally, it seems that hnRNP U may be required to regulate transcription through its association with various protein complexes and transcription factors.

Splicing

hnRNP U is also involved in the splicing process. hnRNP U has been shown to preferentially bind the introns of pre-mRNAs (Xiao et al. 2012), and plays a role in the regulation of exon inclusion and exclusion. Out of the alternative splicing events detected following hnRNP U depletion, hnRNP U is necessary for exon inclusion rather than exclusion in the majority of cases (Huelga et al. 2012; Ye et al. 2015). However, exon exclusion is also activated by hnRNP U, implicating hnRNP U in general exon definition (Huelga et al. 2012; Xiao et al. 2012). The role of hnRNP U in alternative splicing has clinical relevance due to its involvement in the splicing of the *Survival of Motor Neurone* gene (*SMN2*), splicing in the development of the mouse heart and in the T-cell signalling pathway (Xiao et al. 2012; Ye et al. 2015; Meininger et al. 2016). This demonstrates the far-reaching role of hnRNP U in gene expression in vertebrates. hnRNP U also has an indirect role in splicing through the maturation of U2 snRNP (Xiao et al. 2012). The depletion of hnRNP U results in more mature 17S U2 snRNPs rather than 12S and 15S intermediate snRNPs and an increased number of Cajal bodies, the final site of snRNP maturation (Krämer et al. 1999; Nesic 2004; Xiao et al. 2012). This may have effects on splicing in general and may cause the alternative splicing events discussed above.

1.5.2 *hnRNP U and Disease*

hnRNP U has been linked to a variety of diseases due to its numerous roles in gene expression and cellular homeostasis. The protein is important during development as studies in mice have shown that the loss of hnRNP U leads to defects in embryonic development (Roshon & Ruley 2005; Ye et al. 2015). The loss of hnRNP U in the developing mouse heart results in death after 14 days, partly to do with the role of hnRNP U in alternative splicing (Ye et al. 2015). In humans, heterozygous mutations in the HNRNPU gene and deletions in the chromosomal region of HNRNPU, 1q44, can leave patients with a variety of nervous-system related symptoms. These most commonly include intellectual disability and epilepsy but have also been reported to cause organ abnormalities, facial dysmorphism and developmental delay (Caliebe et al. 2010; Thierry et al. 2012; Depienne et al. 2017; Yates et al. 2017; Raun et al. 2017; Bramswig et al. 2017). Variants occur throughout the HNRNPU gene but many are in the SPRY domain (Bramswig et al. 2017; Yates et al. 2017). The majority of the variants in one study are predicted to cause a loss-of-function of hnRNP U due to the mutation producing a premature stop codon (Yates et al. 2017). The human clinical studies and mouse models show that hnRNP U is vital for successful life and reduction or interference with the protein can have profound effects on the organism.

The clinical symptoms described above are heavily linked to Central Nervous System (CNS) development. Other evidence points to hnRNP U being crucial in the CNS, such as its involvement in Amyotrophic Lateral Sclerosis (ALS). When overexpressed, hnRNP U seems to prevent cell death in the brain caused by the toxic effects of the TDP-43 protein (Suzuki et al. 2015). A circular form of the hnRNP U transcript, ciRNA, is highly enriched in neuronal nuclei, further linking hnRNP U to brain disease and the maintenance of a functioning CNS (Reddy et al. 2016) .

Genomic integrity is essential for preventing disease and is maintained by repairing damaged DNA and preserving telomeres at the ends of chromosomes. hnRNP U is linked to both these processes. hnRNP U is phosphorylated in response to double-stranded DNA breaks and localises to the site of DNA-damage through binding to poly(ADP-ribose) chains (PAR) that are present on the damaged DNA (Berglund &

Clarke 2009; Britton et al. 2009; Britton et al. 2014; Gupte et al. 2017). hnRNP U is found to be associated with both the non-homologous end-joining (NHEJ) complex at double-stranded breaks and the base-excision repair (BER) machinery through its interaction with the key NEIL1 protein (Hegde et al. 2012; Hegde et al. 2016). hnRNP U association can help direct the repair of the double-strand breaks by the NHEJ method rather than the more inaccurate BER process (Hegde et al. 2012; Hegde et al. 2016). Additionally, the phosphorylation of hnRNP U by PLK1 is required to prevent mitotic defects in HeLa cells which can lead to significant DNA damage (Douglas et al. 2015).

hnRNP U is also directly associated with the human telomerase complex (hTERT) that is made up of the telomerase enzyme and an associated RNA template, *TERC/hTR*. hTERT is essential to maintain the telomeres of chromosomes during replication and prevent disease (Greider & Blackburn 1987; Greider & Blackburn 1989; Fu & Collins 2007). hnRNP U is associated with both the RNA transcript and the enzyme, but its role in the activity of the complex is not completely known (Fu & Collins 2007; Xiao et al. 2012). Overexpression of hnRNP U results in shortened telomeres implying that it is involved in regulating the complex in some way (Fu & Collins 2007). The hTERT complex and the *TERC* RNA are localised throughout the cell at various cell-cycle stages. However, *TERC* is found in the Cajal bodies the majority of the time and this localisation is essential for telomerase function (Zhu et al. 2004; Tomlinson et al. 2006; Cristofari et al. 2007). hnRNP U has been linked to Cajal body function through its role in U2 snRNP maturation and has a direct influence on the number of Cajal bodies in the cell (Xiao et al. 2012)(Chapter 1.5.1). hnRNP U may have an even bigger role in Cajal body function through coordination of the localisation of various RNAs such as snRNAs and *TERC* through the Cajal body, perhaps similarly to PHAX as discussed in Chapter 1.2.2.

The main functions of hnRNP U are summarised in Figure 1.5.2.

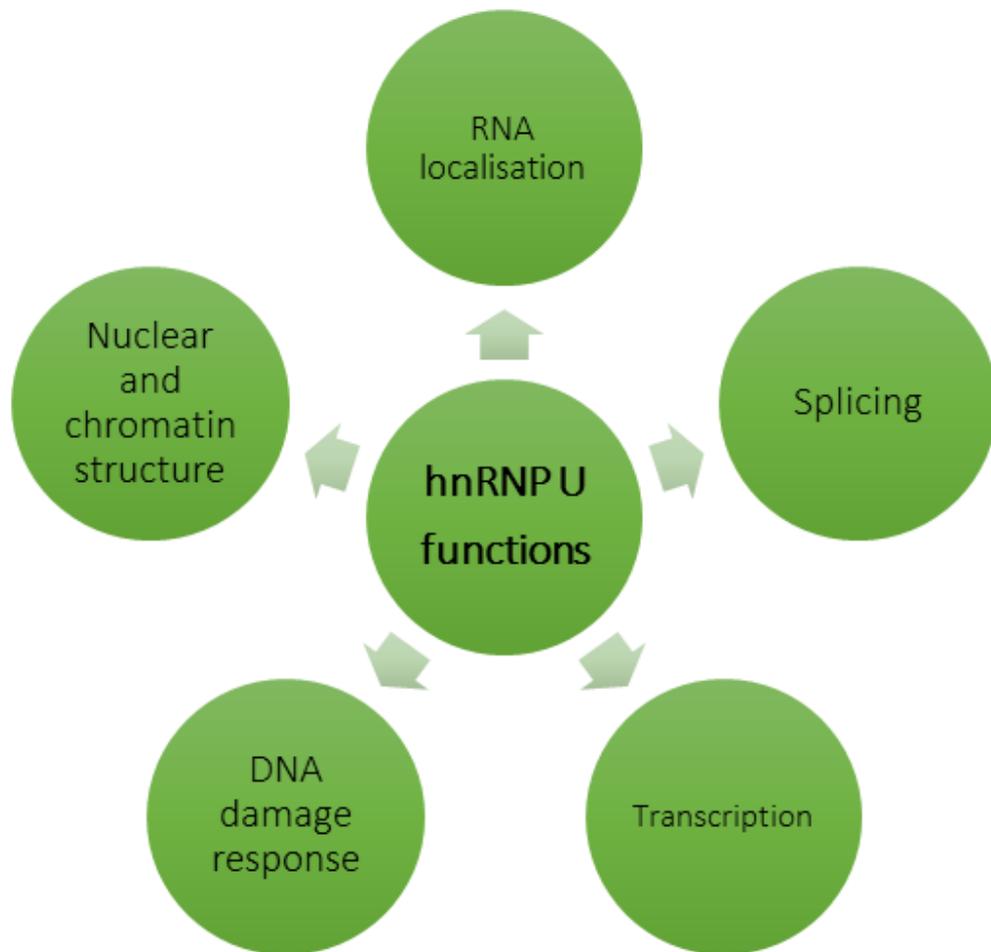


Figure 1.5.2 The main functions of hnRNP U

A summary of the main functions of hnRNP U in the cell.

1.5.3 hnRNP U-like proteins

As well as being part of the large hnRNP family of proteins, hnRNP U has its own subfamily of hnRNP U-like proteins containing hnRNP U-like 1 and hnRNP U-like 2. These proteins share sequence and protein structure similarity and are all fairly equivalent in size. hnRNP U-like 1, also known as E1B-AP5, shares the central region with hnRNP U as well as the RGG domain and putative ATP/GTPase domain (Gabler et al. 1998). hnRNP U and hnRNP U-like 1/2 proteins are conserved in mice and also show domain similarity between the proteins (Sakaguchi et al. 2016). hnRNP U-like 1 has been shown to rescue the hnRNP U depletion phenotype of *XIST* RNA mislocalisation, suggesting some potential redundancy between the proteins (Sakaguchi et al. 2016). However, there are still a range of differences between the proteins which may allow them to carry out different functions in the cell (Gabler et al. 1998). hnRNP U-like 1 has been more thoroughly linked to the DNA-damage response and directly in mRNA export (Barral et al. 2005; Blackford et al. 2008; Polo et al. 2012; Hong et al. 2013; Gabler et al. 1998; Bachi et al. 2000). hnRNP U-like 2 has not been as extensively studied so its role in the cell is not yet clear. However, hnRNP U-like 2 may share similar functions to hnRNP U in regard to lncRNAs as it has been shown to influence the expression of lncRNA *CRNDE* in cancer cells (Jiang et al. 2017). The abundance of the proteins in human cells is markedly different, with hnRNP U being significantly more abundant than hnRNP U-like 1 and even more abundant than hnRNP U-like 2 (Boisvert et al. 2012). Overall, the role that these proteins play in the cell may be synergistic or redundant as well as having individual roles. Their roles in a variety of cellular processes makes them an important family to study and could have far-reaching implications in health and disease.

1.6 RNA Surveillance

Despite the vast array of proteins involved in transcribing, processing and exporting RNAs, many transcripts are still defective or require regulation. These RNAs need to be degraded to prevent any improper action or saturation of RNA that may be detrimental to the cell. This involves conserved RNA surveillance mechanisms which include the role of the exosome complex and the nonsense-mediated decay (NMD) pathway. Additionally, some RNAs are processed by the exosome as well as being degradation substrates.

1.6.1 The exosome complex

The exosome complex was discovered in yeast in the late 1990s and was shown to be conserved in humans (Mitchell et al. 1997). The complex is made up of a catalytically-inert 9 subunit core, EXO-9, with the catalytic activity provided by associated endo and exonucleases. The core complex has a central channel, made up of 6 proteins, with a three subunit “cap” that sits on top of the channel (Liu et al. 2006; Kilchert et al. 2016). The cap, made of Rrp40, Rrp4 and Csl4, is thought to recruit substrates and coordinate their access to the core channel and the nucleases of the exosome complex (Liu et al. 2006; Bonneau et al. 2009; Wasmuth & Lima 2012).

The nucleases associate with EXO-9 depending on the substrate and the cellular location (Mitchell et al. 1997; Tomecki et al. 2010; Lykke-Andersen et al. 2011; Schneider & Tollervey 2014). The 3' to 5' exoribonuclease Rrp6 associates with the top of the core channel complex in the nucleus in yeast but in all compartments in humans, with the strongest enrichment in the nucleoli (Mitchell et al. 1997; Tomecki et al. 2010). The 3' to 5' endo and exoribonuclease Dis3/Rrp44 associates with the bottom of the channel in both the nucleus and the cytoplasm in yeast (Mitchell et al. 1997). However, two Dis3 orthologues exist in humans, hDis3 and hDis3L/Rrp44 which form part of the exosome in the nucleoplasm and cytoplasm respectively (Tomecki et al. 2010).

The exosome functions to degrade aberrant RNA transcripts and process specific non-coding RNAs. Irregular mRNAs can be degraded by the exosome in the

cytoplasm and also in the nucleus when they are export-defective (Anderson & Parker 1998; J. Fan et al. 2017). Other unusual RNA transcripts such as promoter upstream transcripts (PROMPTs), cryptic unstable and stable unannotated transcripts (CUTs and SUTs respectively) are degraded as well as unmodified tRNAs (Schneider et al. 2007; Lykke-Andersen et al. 2011; Lubas et al. 2011). LncRNAs have recently been shown to be exosome substrates, suggesting that any RNA transcript could be an exosome substrate (Lubas et al. 2015; Schlackow et al. 2017; J. Fan et al. 2017). Conversely, the exosome is also required for the processing of non-coding RNAs such as 5.8S rRNA and 3' end processing of snoRNAs and snRNAs (Mitchell et al. 1997; Allmang et al. 1999; Lubas et al. 2011; Schuller et al. 2018).

To recruit substrates and achieve substrate specificity, exosome adaptor complexes such as SKI, TRAMP and NEXT complexes and activators are needed. The SKI complex, harbouring RNA helicase activity, acts in the cytoplasm to recruit mRNAs to the exosome in yeast (van Hoof et al. 2000; Araki et al. 2001). The TRAMP complex is made up of Mtr4p, a DEAD-box RNA helicase, the zinc-knuckle protein Air2p and the non-canonical poly(A) polymerase Trf4p (Schmidt & Butler 2013; Falk et al. 2014). The putative human TRAMP complex is made of a human Mtr4, a zinc-knuckle protein ZCCHC7 and a human Trf4-2 protein (Lubas et al. 2011). The complex unwinds its substrate and adds an unstructured poly(A) tail which then targets it for degradation (Shcherbik et al. 2010; Schmidt & Butler 2013). The TRAMP complex functions in the nucleolus in humans and so it targets nucleolar rRNAs to the exosome (Schmidt & Butler 2013).

The NEXT complex also contains the Mtr4 protein, alongside a zinc-knuckle protein ZCCHC8 and the RNA-binding protein RBM7 (Lubas et al. 2011). The NEXT complex targets PROMPTS and 3' extended snRNAs for degradation in the nucleoplasm (Lubas et al. 2011; Lubas et al. 2015). The cellular localisation of these two complexes provides substrate specificity to the exosome, despite sharing the Mtr4 protein. NEXT also links transcription termination with exosome degradation through its association with the cap-binding complex (CBC) (Andersen et al. 2013). Recent work suggests that RBM7, as part of NEXT, is generically bound to RNA pol II transcripts in the event that the transcript needs to be degraded (Lubas et al. 2015).

The exosome and its specificity as an RNA surveillance mechanism is yet to be fully understood, but it is clear that the exosome plays a major role in cellular homeostasis.

1.6.2 Nonsense-mediated decay (NMD)

Nonsense-mediated decay is a mechanism of RNA degradation that is generally employed for irregular mRNA transcripts. These transcripts may have premature stop codons (PTCs) or extended 3' untranslated regions (UTRs) (Peltz et al. 1993; Muhrad & Parker 1999). This process is heavily linked to translation and is therefore mainly cytoplasmic, although some debate remains about the presence of a nuclear NMD pathway. Following detection of an NMD substrate during translation, UPF proteins 1, 2 and 3 assemble on the transcript and recruit de-capping and degradation enzymes, DCP1-DCP2 and Xrn1 exonuclease respectively (Peltz et al. 1993; Muhrad & Parker 1999; Behm-Ansmant & Izaurralde 2006). This causes the removal of the 5' m⁷G cap on the mRNA and 5'-3' degradation by Xrn1. The exosome can also degrade the transcript from the 3' end following deadenylation (Anderson & Parker 1998). As well as the afore mentioned proteins, the Exon-Junction Complex (EJC) is required for effective NMD in mammalian cells, and Upf3 is associated with this complex (Behm-Ansmant & Izaurralde 2006).

Non-coding transcripts have also been found to be substrates of NMD, despite the low likelihood of being translated. Many non-coding RNA biotypes are detected in polysomes in yeast and are partially translated (Smith et al. 2014). LncRNAs such as Xrn1-sensitive unstable transcripts (XUTs) are found in polysomes and are degraded by NMD when they have extended 3' ends, suggesting that they are somewhat translated and therein marked for NMD (Van Dijk et al. 2011; Smith et al. 2014; Wery et al. 2016). This process also involves the RNA helicase function of Mtr4, noted mainly as an important adaptor for the exosome complex (Lubas et al. 2011; Wery et al. 2016). More defined lncRNAs such as *XIST* and *MALAT1* are also found in polysomes, despite their nuclear localisation, further indicating that non-coding RNAs are present in the cytoplasm and may be translated. These lncRNAs can then be subject to degradation by NMD as cycloheximide (CHX) treatment to inhibit translation leads to stabilisation of the RNAs (Carlevaro-Fita et al. 2016). Although

NMD is heavily defined as an mRNA surveillance pathway, non-coding RNAs present in the cytoplasm are also turned over by this mechanism, making it a more broad method of RNA regulation.

The exosome and NMD RNA degradation pathways show alternative methods of degradation, different cellular localisations and have distinct proteins involved. However, it is unlikely that particular RNA biotypes can only be degraded by one method and the coordinated effort of the exosome, NMD and other RNA surveillance pathways keeps any detrimental RNA in check.

1.7 Project Aims

Despite RNA classes utilising specific export pathways, there is a variety of evidence to suggest that they overlap and perhaps compensate for each other. The role of export pathways and other proteins in intra-nuclear localisation of RNA is also becoming more evident.

Unpublished iCLIP data from the Wilson lab shows mRNA export proteins bound to a variety of non-coding RNAs, such as tRNA and lncRNA, in the human cell. We aim to establish what role the mRNA export pathway plays in the lifecycle of these ncRNAs by investigating a potential function in the export of tRNA and the intra-nuclear localisation of other non-coding RNAs. We will analyse the localisation of tRNA using cellular fractionation and FISH following depletion of canonical tRNA export proteins alongside the depletion of TREX components and Nxf1. If the mRNA export pathway is involved in global tRNA export, we predict that we will see accumulation of tRNA within the nucleus of cells defective in mRNA export. The intra-nuclear localisation of *scaRNA2* and the lncRNA *XIST* will also be investigated using FISH in TREX and Nxf1 depleted cells.

The protein hnRNP U is bound to many classes of RNA and is involved in the nuclear localisation of specific non-coding transcripts. The absence of hnRNP U can result in lncRNA mislocalisation. mRNA export proteins are bound to many of the same lncRNAs as hnRNP U and the TREX protein Alyref is known to interact with hnRNP U. Perhaps these RNAs mislocalise as the mRNA export pathway can now export them in the absence of hnRNP U. This effect could be universal for lncRNAs and other RNAs associated with the chromatin. To test this hypothesis, we aim to establish a reliable and reproducible method of depleting hnRNP U in human cells to assess RNA localisation changes. We will build on published studies by getting a global view of hnRNP U function in RNA localisation by carrying out an RNA-sequencing experiment. The work here will go some way in dissecting the interactions between RNA export and localisation pathways and help to uncover global mechanisms in RNA localisation.

Chapter 2 - Materials and Methods

2.1 Miscellaneous Buffers

10x PBS pH7.4: 1.4 M NaCl, 27 mM KCl, 0.1 M Na₂HPO₄, 18 mM KH₂HPO₄

IP lysis buffer: 50 mM HEPES-NaOH pH7.5, 0.1 M NaCl, 1 mM EDTA pH8, 0.5% Triton-X-100, 10% Glycerol

5x TBE: 4.4 M Tris, 4.4 M Boric Acid, 0.1 M EDTA pH8

2.2 Tissue Culture

Cell lines

HEK-293T – Human Embryonic Kidney cell line that contains the large T antigen from the SV40 virus

Flp-In T-Rex 293 – Cell line used to make Tet-inducible stable cell lines. Contains one Flp Recombination Target (FRT) site to allow integration of a hairpin of choice.

HeLa – Human cervical epithelial carcinoma cell line

HCT116 - Human colorectal carcinoma cell line

HCT116-TIR1 – Human colorectal carcinoma cell line that contains the TRANSPORT INHIBITOR RESPONSE 1 protein (TIR1) from *Oryza Sativa* (os) or “Asian rice”. (Eaton et al. 2018)

Growth media

293T, HeLa and HCT116 cells were cultured in Dulbecco's Modified Eagle Medium (DMEM, *life technologies*) supplemented with 10% FCS (v/v) (*life technologies*) and 1% Penicillin-Streptomycin (v/v) (Pen/Strep, *Invitrogen*) and grown at 37 °C in 5% CO₂. The Flp-In T-Rex 293 cells were cultured in DMEM with 10% Tet-free FCS (v/v) (*life technologies*), 1% Pen/Strep (v/v), 100 µg/ml Zeocin (*Invitrogen*) and 15 µg/ml Blastidicin (*Invitrogen*). The generated inducible Flp-In cell lines were selected and cultured in DMEM with 10% Tet-free FCS (v/v), 1% Pen/Strep (v/v) with additional 100 µg/ml Hygromycin (*Invitrogen*) and 15 µg/ml Blastidicin to maintain selection. When generating a cell line, Initial transfection and pre-selection medium contains Tet-free FCS and Pen/Strep but without selective antibiotics. Stable cell lines are induced with Tetracycline at a final concentration of 1 µg/ml. The approximate percentage knockdown at the protein level was calculated using volume analysis and shown in the table below.

RNAi cell line	Length of tet-induced knockdown (hrs)	~Knockdown at the protein level (%)
Control	Equivalent to test cell line	N/A
Alyref	72	53
Alyref/Chtop	48	53 and 65
Nxf1	72	94

HCT116-TIR1 cells were maintained in DMEM with 10% FCS (v/v), 1% Pen/Strep (v/v) and 10 µg/ml Blasticidin. The generated hnRNP U Auxin-Inducible Degron (AID) cell line was selected and cultured in the maintenance media supplemented with 150 µg/ml Hygromycin, 800 µg/ml G418/Neomycin (*ThermoFisher*), 10 µg/ml Blasticidin and 1 µg/ml Puromycin (*ThermoFisher*). Indole-3-acetic acid/Auxin (*SIGMA-ALDRICH*), was added to the AID cell line at a final concentration of 500 µM. Cell lines were maintained by passaging twice a week using Trypsin/EDTA (*Labtech*) and diluting in fresh media. Cell lines were replaced with low passage number cells once they had reached passage 30.

Small interfering RNA (siRNA) Transfections

Cells were plated the day before to give ~50% confluency at the time of transfection. Lipofectamine RNAiMax reagent (*ThermoFisher*) was used to transfect between 5 and 25 nM siRNA in accordance with the manufacturer's instructions. Cells were then grown for 48 hours before the cells were harvested or the transfection was repeated again. In some cases, depending on the siRNA and experiment, the cells were trypsinised and split down to between 1/3 and a 1/2 of the cell number just before re-transfecting. Cells were then used at 72-96 hour time points. The approximate percentage knockdown at the protein level was calculated using volume analysis and shown in the table below.

Target transcript	siRNA sequence/source	Length of knockdown (hrs)	~ Knockdown at the protein level (%)
Control - dsRED	5'- CACCGUGAAGCUGAAGGUG	Equivalent to test knockdown	N/A
Exportin 5	5'- CAAGCAAACUGUCGAGUAG	72	93
Exportin T	Ambion Silencer Select – s22234 (ThermoFisher)	72	75
Alyref	5'- GGAACUCUUUGCUGAAUUU-3'	72	76
Chtop	5'- GACAACCAAUUGGAUGCAUUAU	72	85
Nxf1	5'- UGAGCAUGAUUCAGAGCAA	72	84
hnRNP U	SMARTPool: ON-TARGETplus – 3192 (Dharmacon)	96	53
Rrp40	5'- UGGCAUAGUGACAGCUAAA	72	Only tested at RNA level – Average of 75% across all cellular fractions

The knockdown time when using RNAi cell lines or siRNA was decided upon as a balance between the greatest level of depletion at the protein level and the health of the cells so they can be effectively processed. For example, fractionation of Alyref/Chtop RNAi cells after 48 hours generally resulted in nucleoplasmic RNA leakage to the cytoplasm. The earliest possible time point is generally used as the longer the time taken for depletion, the less likely you are to see direct effects of depleting the protein and the more likely you are to see indirect effects on cellular proliferation and metabolism. The time taken for successful depletion is also dependent on the half-life of the protein and how well the mRNA is degraded following siRNA targeting.

DNA Transfections

Cells were plated the day before to give a ~70% confluency at the time of transfection. Turbofect (*Thermo Scientific*) was used to transfect DNA in accordance with the manufacturer's instructions.

Colony-formation assay

A 6cm dish of each condition was seeded and if required, treated with auxin for 24 hours at 500 μ M final concentration. The dishes were then trypsinised and a 6-well plate seeded for each condition with 200 cells per well. The dishes were left to grow for 13 days at 37 °C, with the auxin being refreshed twice a week and the selection antibiotics refreshed once a week. Cells were washed once in 1x PBS and then 1 ml of Crystal Violet solution (0.5% Crystal Violet powder (w/v) in 100% methanol) was left on each well for 5 minutes. The solution was washed off twice with 1 ml of water and then the dishes were left to dry for 15 minutes before images were obtained using the BIO-RAD Chemidoc system and colonies counted. All colonies counted contained more than 100 cells.

Generating inducible expression cell lines using Flp-In T-Rex 293 cells

The Flp-In System (*life technologies*) was used to make all stable hairpin or construct expressing cell lines using Flp-In T-Rex 293 cells (*Invitrogen*). These cells contain an FRT site for recombination of your desired construct and a Tet-repressor to allow control of its expression. Tetracycline can be added to the cell growth media to induce expression of the construct.

2.3 Protein Biology

Buffers

1x SDS-PAGE Resolving gel buffer: 375 mM Tris pH8.8, 0.0375% (w/v) SDS

1x SDS-PAGE Stacking gel buffer: 125 mM Tris pH6.8, 0.0375% (w/v) SDS

1x SDS-PAGE Loading buffer: 50 mM Tris pH6.8, 100 mM Dithiothreitol (DTT), 0.1% (w/v) bromophenol blue, 10% (v/v) glycerol, 2% (w/v) SDS

1x SDS-PAGE Running buffer: 25 mM Tris, 250 mM Glycine, 0.1% (w/v) SDS

10x TBS pH7.6: 0.2 M Tris, 1.37 M NaCl

TBST: 1x TBS, 0.2% Tween-20 (v/v)

Milk solution: 5% milk powder (w/v), 0.2% Tween-20 (v/v), 1x TBS

ECL1: 100 mM Tris-HCL pH8.5, 2.5 mM Luminol, 400 μ M p-coumaric acid

ECL2: 100 mM Tris pH8.5, 5.3 mM Hydrogen Peroxide

Sodium Dodecyl Sulphate -Polyacrylamide Gel Electrophoresis of proteins (SDS-PAGE)

SDS-PAGE gels of 8-12% acrylamide/bisacrylamide resolving gel were prepared depending on the size of the protein of interest with a 5% acrylamide/biscrylamide stacking gel on top. All gels were prepared using a 30% acrylamide/bisacrylamide stock (*BIO-RAD*) with their corresponding buffers. 0.2% (v/v) TEMED and 0.1% (w/v) ammonium persulfate were added to the gel solution to facilitate polymerisation in *BIO-RAD* Mini-Protean II gel casts. Cells were lysed in IP lysis buffer supplemented with protease inhibitors (*SIGMAFAST* Protease inhibitor cocktail tablets), 1x RiboSafe RNase Inhibitor (*Bioline*) and 1 mM DTT and centrifuged at 13200 rpm for 5 minutes at 4 °C. The protein concentration of samples was determined using the Bradford protein assay (*SIGMA-ALDRICH*). Samples were made up with 1x SDS-loading buffer supplemented with 10% (v/v) β -mercaptoethanol and boiled for 5 minutes before loading alongside a protein ladder (PageRuler Plus Prestained Protein Ladder,

Fermentas). The gels were run in SDS-PAGE running buffer at approximately 25mA per gel for an hour.

Western blot

SDS-PAGE gels were transferred to nitrocellulose membranes using a BIO-RAD fast-transfer machine at 25 volts for 15 minutes. Membranes were blocked in milk solution for 1 hour before primary antibodies diluted in milk solution were added. Membranes were washed in TBST, probed with secondary antibodies for 30 minutes and washed extensively again in TBST. Proteins with antibodies bound were detected by Enhanced ChemiLuminescence (ECL) using a mixture of equal volumes of ECL1 and ECL2 washed over the membrane. The membrane was immediately exposed using a BIO-RAD Chemidoc system. The secondary antibodies conjugated to HRP were inactivated using 0.02% Azide (From 1 M 6.5% stock) in TBST for 10 minutes to detect antibodies raised in different organisms.

Antibody	Source	Species	Type
Tubulin	SIGMA-ALDRICH T5168 - Clone B-5-1-2	Mouse	Monoclonal
Chtop	Clone KT64	Rat	Monoclonal
Alyref	SIGMA-ALDRICH A9979 - Clone 11G5	Mouse	Monoclonal
Exportin-5	Abcam (57491)	Mouse	Monoclonal
Exportin-T	Abcam (168832)	Rabbit	Polyclonal
SSRP1	Biolegend - Clone 10D1	Mouse	Monoclonal
Myc	Roche – Clone 9E10	Mouse	Monoclonal
FLAG	SIGMA-ALDRICH F3165 – Clone M2	Mouse	Monoclonal
Nxf1	Abcam (50609) – Clone 53H8	Mouse	Monoclonal
hnRNP U	Abcam (10297) – Clone 3G6	Mouse	Monoclonal
hnRNP U-L1	In house	Rabbit	Polyclonal
Histone H3	Abcam (1791)	Rabbit	Polyclonal
hnRNP A1	Merck-Millipore 04-1469 – Clone 9H10	Mouse	Monoclonal
TDP-43	Proteintech 10782-2-AP	Rabbit	Polyclonal
eiF4A3	Abcam (180573) – Clone EPR14301(B)	Rabbit	Monoclonal
PSF	SIGMA-ALDRICH P2860 – Clone B92	Mouse	Monoclonal
Coilin	Bethyl Laboratories A303-760A	Rabbit	Polyclonal
α-Mouse	Promega W4021, HRP Conjugate	Goat	Polyclonal
α-Rabbit	Promega W4011, HRP Conjugate	Goat	Polyclonal
α-Rat	ThermoFisher 62-9520, HRP Conjugate	Goat	Polyclonal

Anti-FLAG co-Immunoprecipitation (co-IP)

100 μ l of Protein-G Dynabeads (*Invitrogen*) per condition were incubated with 4 μ g FLAG antibody (*M2 clone, SIGMA-ALDRICH*) in IP lysis buffer with 1% (w/v) Bovine Serum Albumin (BSA) on a wheel for 2 hours. Meanwhile, 3 x 15cm dishes of cells per condition were lysed in IP lysis buffer, supplemented with protease inhibitors and 1 mM DTT, and the cell debris removed by centrifuging at 17000 x g for 10 minutes at 4 °C. The protein concentration of the resulting supernatant was calculated using the Bradford protein assay. The beads were washed three times in IP lysis buffer and 5 mg of protein extract from each condition at the same concentration was added. The beads were rotated for 2 hours at 4 °C. The beads were washed three times in IP lysis buffer with 1 mM DTT and the bound protein complexes eluted in 60 μ l of IP lysis buffer with 100 μ g/ml 3x FLAG peptide (*SIGMA-ALDRICH*), rotating at 4 °C for 1 hour. The resulting elution was loaded onto two 10% polyacrylamide gels, as described above, alongside input samples that were 0.8% of the amount of protein added to the beads. The gels were transferred to a nitrocellulose membrane and Western blotting carried out as above.

2.4 RNA Biology

Buffers

Sucrose lysis buffer: 0.5 M Sucrose, 3 mM CaCl₂, 2 mM MgCl₂, 10% glycerol (v/v), 10 mM Tris-HCl pH8.0

NRB buffer: 20 mM HEPES pH7.5, 75 mM NaCl, 1 mM DTT, 50% Glycerol (v/v), Protease Inhibitors

NUN buffer: 20 mM HEPES pH7.5, 300 mM NaCl, 1 mM DTT, 10 mM MgCl₂, 1 M Urea, 1% NP-40 (v/v)

Buffer A: 10 mM HEPES pH7.5, 10 mM KCl, 10% glycerol (v/v), 4 mM MgCl₂, 1 mM DTT, Protease Inhibitors

RIPA buffer: 50 mM HEPES pH7.5, 15 mM NaCl, 1 mM DTT, 10% glycerol (v/v), 1% NP-40 (v/v), 0.1% SDS (w/v), 0.5% Sodium deoxycholate (w/v), Protease inhibitors

20x SSPE pH7.4: 3 M NaCl, 0.18 M NaH₂PO₄, 20 mM EDTA pH8,

5x Denhardt's: 0.2% Ficoll (w/v), 0.2% Polyvinylpyrrolidone (w/v), 0.2% BSA (w/v)

8% Urea-acrylamide gel: 8 M Urea, 8% Acrylamide/Bis solution 19:1 (40%), 0.5x TBE, 0.07% APS (w/v), 0.07% (v/v) TEMED

2x RNA sample buffer: 95% formamide (v/v), 20 mM EDTA pH8, 0.05% bromophenol blue (w/v), 0.05% xylene cyanol (w/v)

Oligohybridisation buffer: 6x SSPE, 5x Denhardt's, 0.2% SDS (w/v)

Stripping buffer: 0.1x SSPE, 0.1% SDS (w/v)

RIP Lysis buffer: 50 mM HEPES pH7.5, 150 mM NaCl, 10% glycerol (v/v), 1% NP-40 (v/v), 0.1% SDS (w/v), 0.5% sodium deoxycholate (w/v)

High salt RIP lysis buffer: 50 mM HEPES pH7.5, 500 mM NaCl, 10% glycerol (v/v), 1% NP-40 (v/v), 0.1% SDS (w/v), 0.5% sodium deoxycholate (w/v)

Proteinase K (PK) buffer: 100 mM Tris-HCl pH7.5, 50 mM NaCl, 10 mM EDTA pH8

RNA extraction

RNA was isolated using TRIzol reagent (*Invitrogen*) as per the manufacturer's instructions. The RNA was resuspended in water and DNase treated with 4 units of TurboDNase (*ThermoFisher*) with TurboDNase buffer and Ribosafe RNase inhibitors (*Bioline*) for 1 hour at 37 °C. The RNA was re-extracted using RNA acidic phenol (pH5.5) and re-precipitated in 86 mM sodium acetate pH5.8, 100% ethanol and glycogen. The pellet was washed in 75% ethanol and re-suspend in RNase-free water as required.

Chromatin, Nucleoplasmic and Cytoplasmic RNA fractionation

Cells were washed in 1x PBS, trypsinised and transferred to a 15 ml round-bottomed tube in pre-warmed media. Cells were pelleted at 500 x g for 3 minutes and washed twice in 1x PBS. The pellet was gently re-suspended in 8-10x (v/v) sucrose lysis buffer supplemented with 0.5% Triton-X (v/v), 1 mM DTT, Protease and RNase Inhibitors. The sample was transferred to a 2 ml round-bottomed tube and spun at 500 x g for 5 minutes at 4 °C. The supernatant was retained and spun at full speed for 1 minute. A sample of the resulting supernatant was taken as the cytoplasmic protein sample and 750 µl of TRIzol LS was added to the rest. The remaining nuclear-pellet was washed twice in 1 ml sucrose lysis buffer, once in 1x PBS supplemented with 0.5 M sucrose and once more in sucrose lysis buffer. The nuclear pellet was resuspended in 300 µl of NRB buffer and an equal volume of NUN buffer was added. The sample was left on ice for 5 minutes with inversion every minute. The sample was spun at 1600 x g for 5 minutes at 4 °C and 50 µl of supernatant was taken as a nucleoplasmic protein sample and 250 µl added to TRIzol LS as the nucleoplasmic RNA sample. The remaining chromatin pellet was resuspended in Buffer A, transferred to a new tube and spun at 1600 x g as before. For protein extraction, the chromatin pellet was resuspended in 100 µl of RIPA buffer supplemented with 500 units of Benzonase (*Insight Biotechnology*) and left at room temperature for 45 minutes. The chromatin sample was then spun at 16100 x g for 10 minutes at 4 °C and 50 µl of the supernatant retained. For RNA extraction, the chromatin pellet was made up to 250

µl with Buffer A and 750 µl of TRIzol LS added. All RNA samples were purified using TRIzol as normal.

Urea-polyacrylamide gel electrophoresis of RNA and Northern blot

Samples were boiled with 1x RNA sample buffer and run on an 8% Urea-polyacrylamide gel in 0.5x TBE for approximately two and a half hours. Gels were transferred to a Hybond N+ membrane using a MINI Trans-blot system (BIO-RAD) for 2 hours at 40 volts which was then UV crosslinked twice at 1200 µjoules cm⁻². Membranes were blocked in oligohybridisation buffer for 1 hour at 37 °C. 250 nM of DNA probe was labelled with [γ -³²P]ATP (*PerkinElmer*) using polynucleotide kinase for 1 hour at 37 °C. The labelled probe was added to the membrane in oligohybridisation buffer and hybridised overnight at 37 °C. The membrane was washed in 6x SSPE, dried and exposed to a phosphor film. The film was exposed using a Typhoon FLA 7000 laser scanner and analysed using Quantity One software. Boiling stripping buffer was used to remove the probe so the blot could be hybridised again.

UV RNA-Immunoprecipitation (RIP)

100 µl of Protein G Dynabeads (*Invitrogen*) per condition were washed in RIP lysis buffer and rotated overnight at 4 °C with 2 µg of FLAG antibody per condition in RIP lysis buffer with 1% BSA (v/v). 4 x 15 cm dishes of cells per condition were washed with ice-cold 1x PBS and UV irradiated with 150 mjoules cm⁻². The cells were then sonicated 6 times (10 seconds on and 30 seconds off) at 10 000 amplitudes in RIP lysis buffer supplemented with 1 mM DTT, protease inhibitors, RNase inhibitors and TurboDNase. The lysate was cleared by spinning at 17000 x g for 10 minutes at 4 °C. The beads were washed in RIP lysis buffer and the cleared cell lysis added to the beads, following the removal of a sample to act as an input. The samples were rotated overnight at 4 °C. The supernatant was removed and beads washed twice with RIP lysis buffer, twice with high-salt RIP lysis buffer and twice with PK buffer. The beads were resuspended in 200 µl PK buffer with 5% Proteinase K (v/v) (*Roche*) and incubated at 37 °C for 1 hour at 1100 rpm. The eluates were transferred to fresh tubes and the RNA extracted using TRIzol LS as above.

Reverse transcription and quantitative Polymerase Chain Reaction (qRT-PCR)

~2.5 µg of RNA was used to generate cDNA by reverse transcription using the BioScript kit (*Bioline*). RNA samples were mixed with 200 ng of poly(dN)⁶ random primer and 0.5 mM dNTPS and incubated at 70 °C for 5 minutes. 100 units of BioScript enzyme and 1x reaction buffer were added and reverse transcription carried out using the following conditions:

25 °C for 10 minutes, 42 °C for 1 hour and 85 °C for 5 minutes.

The resulting cDNA was diluted by 5-10x with RNase-free water and used in quantitative PCR with SensiMix (*Bioline*). 1-4 µl of cDNA was mixed with 500 nM of selected primers (see table below) and 1x SensiMix and cycled on a Rotor-gene 6000 (*Qiagen*) using the following conditions:

95 °C for 10 minutes and then 45 cycles of 95 °C for 10 seconds, 59 °C for 15 seconds and 72 °C for 25 seconds.

The resulting data was analysed using the corresponding Rotor-gene software using the comparative quantitation method. Transcripts from “housekeeping” genes were used to normalise the levels of tested transcripts to the amount of cDNA present in the sample. The levels of the normalising transcripts chosen remain constant regardless of the treatment to the cell, so the levels are proportional to the amount of cDNA in the sample. The housekeeping transcripts were chosen depending on the nature of the experiment. For qRT-PCR following protein depletion using Alyref/Chtop or Nxf1 RNAi cell lines, normalisation was carried out using U1 snRNA transcripts as published (Chang et al. 2013; Viphakone et al. 2015). For qRT-PCR of total RNA extracted from hnRNP U depleted cells by using siRNA or the auxin-inducible degron cell line, normalisation was carried out using U6 snRNA transcripts. U1 could not be used in this case as the levels are reduced following hnRNP U depletion where as U6 levels remain constant (Figure 5.1.3). For RT-qPCR of fractionated RNA samples, normalisation was carried out using the 18S rRNA transcripts, as used commonly (Topisirovic et al. 2009; Werner et al. 2017) .

qRT-PCR Primers:

Transcript	Forward primer	Reverse primer	Associated Paper
18S	GTGGAGCGATTTGTCTGGTT	CGGACATCTAAGGGCATCAC	
U6	TGCTCGCTTCGGCAGCACAT	AATATGGAACGCTTCACGAA	
U1	ACCTGGCAGGGGAGATACCA	GGGGAAAGCGCGAACGCAGT	
U2	GATTTTTGGAGCAGGGAGATGG	TACTGCAATACCAGGTCGATGC	
U4	TATCCGAGGCGCGATTATTG	CCAGTGCCGACTATATTGCAAG	
U4atac	AACCATCCTTTTCTTGGGGT	GCACCAAATAAAGCAAAGC	
XIST	AAAGCTCACTACCACTGGGC	CTTGACGTGTGGTGGTTGTT	
FIRRE	CTGGCAGCAGAGACTAAGGT	GTGTTTGCAAGCCAGGTACA	
TERC	CCTAACTGAGAAGGGCGTAGG	GCTCTAGAATGAACGGTGGAAG	
NEAT1	GCTTGGAACCTTGCTTCAAG	GGTGGGTAGGTGAGAGGTCA	
NEAT1_2	GGCCAGAGCTTTGTTGCTTC	GGTGCGGGCACTTACTFACT	
TUG1	GTGCAAACCTGAGGATGCTCC	GGTACCAGGTCTGTAGGCTG	
MALAT1	CTAGGACTGAGGAGCAAGCG	CTCGCTCCTTCTGGAATCC	
DANCR	TTGTCACAACCTCGGAGGTGG	GTGACCTTGCTATAGCGCCT	
PANDA	TGCACACATTTAACCCGAAG	CCCCAAAGCTACATCTATGACA	(Puvvula et al. 2014)
hnRNP U	CTGGGAATCGTGGCGGATAT	TTCCTCTGTTGGGCATTCC	
EGR1	GAAGAACTTGACATGGCTGTT TC	CCTCCCTCTACTGGAGTGGA A	
ACTB mRNA	GAAGGAGATCACTGCCCTGG	CGGACTCGTCATACTCCTGC	
MC1R	GGACCGCTACATCTCCATCT	GCATAGCCAGGAAGAAGACC	
HIST1H3H	TGCTGATCAGAAAGCTGCCT	CTGGATGTCCTTGGGCATGA	
HIST1H2BC	ATCACCTCCAGGGAGATCCA	GAGCCTTTGGGGTTAGGTGT	
FEN1	GGCTCAAACCACTTCTCAGG	TGGTGGCTCTCTCAAGGT	
HHLA3	CAGACCCCAAGAGAGCATTTC	GGGCAGGAACAAATCACAAT	(Xiao et al. 2012)
MAP3K7	CTCCATCCCAATGGCTTATC	TTTTTGCATTGCTGGTAGTAAG	(Xiao et al. 2012)
FIP1L1	GAGGATACGAATGGGACTTGA	TGGAAGCCCAGTCTTGAACA	(Xiao et al. 2012)
RAC1	GGTAGATGGAAAACCGGTGA	CTTTGCACGGACATTTTCAA	(Xiao et al. 2012)
ST7	ATGCAGAAAGCCTGGAGAGA	CCTCAGGGCCTGCTTAAAT	(Xiao et al. 2012)
ATXN2	CAACTCAGTACGGGGCTCAT	GACTGGGTGCAGGATGACTT	(Xiao et al. 2012)
ACTB pre-mRNA	TCAAGGTGGGTGTCTTTCCT	CCTGCTTGCTGATCCACATC	
NFkB1A pre-mRNA	TCCTCAACTCCAGAACAACC	CAGAAAGGATCTGGGGTGAC	(Yoshimoto et al. 2017)
DUSP1 pre-mRNA	GAGCTGTGCAGGAAACAGTC	CTGTGGCAGGGACACCTACT	(Yoshimoto et al. 2017)
ID2 pre-mRNA	CCTGTAGCACTGCTGTTGGA	GGAGGAAGCCACAGTTTGAA	(Yoshimoto et al. 2017)

18S was used due to concerns about the effects of variation in the levels of U6 snRNA between the different fractions where as 18S levels remain consistent between fractions. As these transcripts remain relatively unchanged in all respective conditions and fractions, the normalisation procedure should not mask any data as the given values are similar to the the raw qPCR values before normalisation.

RNA-sequencing (RNA-seq)

Sequencing libraries were prepared and sequenced by Novogene, Hong Kong. This involved ribosomal RNA depletion before sequencing. Long non-coding RNA paired-end 150bp sequencing was carried out using Illumina technology. The resulting data was analysed by Mr Jacob Parker and Dr Ian Sudbery. The Ensemble gene ids and the corresponding RNA biotype was taken from the Ensemble human genome version 85.

2.5 Cell Biology

Buffers

Fixing solution: 4% formaldehyde (v/v), 1x PBS

Permeabilisation solution: 0.5% Triton-X 100 (v/v), 1x PBS

BSA solution: 1% BSA (w/v), 1x PBS

20x SSC: 3 M NaCl, 0.34 M Sodium citrate, 0.23 M NaH₂PO₄

Hybridisation buffer: 10% Dextran sulfate (w/v), 10% formamide (v/v), 2x SSC, 0.2% Bovine serum albumin (w/v), 125 µg/ml *E.coli* total RNA, 500 µg/ml sheared salmon sperm DNA

Cell fixation

Cells grown on coverslips were washed in 1x PBS and fixed in fixing solution for 30 minutes. The cells were washed and incubated in permeabilisation solution on ice for 10 minutes. The cells were washed twice in 1x PBS and then stored at 4 °C.

Immunostaining

The cells were saturated with BSA solution for 1 hour before incubation with the primary antibody in BSA solution at room temperature. The cells were washed three times in 1x PBS and incubated with the secondary fluorescent antibody for 30 minutes in BSA solution. Three 1x PBS washes were applied and then mounted onto slides using mounting medium containing DAPI (*VECTASHIELD*).

Fluorescent in situ hybridisation (FISH)

The cells were hybridised at 37 °C overnight with 10-50 pmol Cy3-labelled probe in hybridisation buffer. The cells were washed three times in 2x SSC and 10% formamide (v/v). The coverslips were mounted as above. An *XIST* RNA Stellaris FISH probe (*LGC Biosearch*) was used for *XIST* staining and Cy3-labelled probes to tRNA and scaRNA2 were manually designed and ordered from Eurofins.

FISH probes:

Transcript	FISH probe sequence/source	Associated Paper
tRNA-Lys	5' Cy3- CGAACCCACGACCCTGAGATTAAGAGTCTCATGC	(Chafe et al, 2011)
scaRNA2	5'Cy3- CCGGCCTCGTCTATCTGATCAATTCATCACTTCT	
XIST	Stellaris Human XIST Quaser 570 dye (<i>LGC Biosearch</i>)	

2.6 DNA and Molecular Biology

Buffers

6x DNA loading buffer: 0.25% Xylene cyanol (w/v), 0.25% Bromophenol blue (w/v), 30% glycerol (v/v)

Escherichia coli (E.coli) growth media and selection conditions

Luria broth (LB): 10 g/l Tryptone, 10 g/l NaCl, 5 g/l Yeast extract

LB Agar plates: LB as above with 2% agar (w/v)

Antibiotic selection conditions:

Antibiotic	Final concentration
Ampicillin	100 µg/ml
Kanamycin	50 µg/ml

Polymerase Chain Reaction (PCR)

PCR reactions were set up with a variety of DNA polymerases depending on the purpose of the PCR. In general, around 50 ng of DNA template, 0.5 mM dNTPs, 0.2-1 µM forward and reverse primers, 1x reaction buffer and 2-5 units of polymerase were used depending on the manufacturer's instructions. Reactions were cycled between 30 and 35 times, with the denaturation temperature around 98 °C, the annealing temperature >59 °C and an extension temperature of 72 °C. The extension time depended on the enzyme and the template size but annealing and denaturation lasted between 15 and 30 seconds.

Colony PCR

To screen for successful *E.coli* transformants, colonies were picked using a pipette tip, dotted onto a fresh agar plate and then shaken into a PCR tube. PCR can then be carried out as above but with an initial denaturation step of 98 °C for 10 minutes.

Successful transformants can then be grown in LB overnight and the plasmids purified as described below.

Agarose gel electrophoresis of DNA

0.5-2% agarose (w/v) depending on the size of DNA fragments was dissolved in 0.5x TBE before adding ethidium bromide (*BIO-RAD*) to a final concentration of 10 µg/ml. Gels were left to set in moulds before loading DNA samples made up with 1x loading buffer. Gels were run in 0.5x TBE at around 100 volts until the desired DNA bands had separated enough. DNA bands were visualised using a *BIO-RAD* Chemidoc transilluminator by exposing to UV light.

Extraction of DNA fragments from an agarose gel

The QIAquick gel extraction kit (*Qiagen*) was used to purify excised DNA as per manufacturer's instructions.

Phenol: Chloroform DNA purification

100 µl of phenol: chloroform pH6.7 was added to 100 µl of DNA sample, vortexed for 1 minute and spun at 17000 x g for 3 minutes. The aqueous phase was then transferred to a fresh tube containing glycogen and 86 mM sodium acetate pH5.8 and 300 µl of 100% ethanol was added. The DNA sample was then mixed and incubated at -20 °C for a minimum of 30 minutes. Following incubation, the sample was centrifuged for 30 minutes at 17000 x g, the resulting pellet washed with ice cold 75% ethanol and spun for 10 minutes at 17000 x g at 4 °C. The resulting pellet was dried and resuspended in the appropriate amount of water.

Transformation of chemically competent E.coli

The DH5α strain of *E.coli* (*Invitrogen*) were made competent using the Rubidium chloride method and stored in the -80 °C freezer (Green & Rogers 2013). Competent cells were thawed on ice, the appropriate amount of DNA added and left on ice for 30 minutes. The cells were then heat-shocked for 30 seconds at 42 °C and returned to ice for 2 minutes. 900 µl of LB medium was added to the cells and were left to

recover for 1 hour at 37 °C whilst shaking at 200 rpm. 100 µl of the cells were plated onto a pre-warmed agar plate containing the appropriate antibiotic and incubated overnight at 37 °C.

Extraction of plasmids from E.coli

Transformed *E.coli* were grown in 3-50 ml LB containing the appropriate antibiotic overnight at 37 °C with shaking at 200 rpm. QIAGEN mini and midi-prep kits were used in accordance with the manufacturer's instructions to purify plasmids from *E.coli* grown in 3-5 ml or 50 ml of LB respectively.

Molecular cloning using the Gibson assembly method

The Gibson Assembly Master Mix Kit (*New England Biolabs*) was used to assemble DNA fragments when cloning CRISPR-Cas9 guide RNA and repair template vectors. The Cas9 plasmid used was pX330 (*addgene*)(Ran et al. 2013) and pUC19 was used to clone the repair templates. The master mix was used to assemble DNA fragments as per the manufacturer's instructions and were then transformed into *E.coli* using the standard transformation procedure above.

DNA sequencing

DNA was sequenced using the Source Bioscience Sanger sequencing service using their stock primers or primers sent with the DNA samples.

2.7 Statistical Testing

A two-tailed unpaired T-test was carried out for all statistical analyses. A one-sample T-test was carried out when comparing the condition relative to the control when the control is set as 1. The statistical significance level, or P-value, is indicated on the graphs using the following asterisk marks:

P-value	Number of asterisks
$P = > 0.05$	Not indicated on graph
$P = < 0.05$	*
$P = < 0.01$	**
$P = < 0.001$	***

Chapter 3 - The role of TREX and Nxf1 in the cellular localisation of non-coding RNA

Non-coding RNAs (ncRNA) make up the majority of RNA in the cell, providing diverse cellular roles that are heavily dependent on their localisation. Although the most abundant classes of ncRNA, transfer RNA (tRNA) and ribosomal RNA (rRNA), are exported to the cytoplasm for their function, many classes of ncRNA are retained in the nucleus and in specific nuclear bodies. Small Nucleolar RNAs (snoRNA) and Small Cajal body-specific RNAs (scaRNA) localise where their name suggests and many long non-coding RNAs (lncRNA) are retained in the nucleus. However, the factors needed and the mechanisms that are necessary for the localisation and consequently the function of these various RNAs has not been fully defined.

iCLIP (individual-nucleotide resolution Cross-Linking and Immuno-Precipitation) is a technique used to map where proteins are bound to RNA at an individual nucleotide resolution (König et al. 2011; Huppertz et al. 2014). The procedure involves irradiating cells with UV light, causing crosslinking of proteins to the RNA they are bound to through an irreversible covalent bond. A protein of interest can then be immunoprecipitated from cell extracts, bringing with it the associated RNA. Following processing steps, complementary DNA (cDNA) from the RNA associated is sequenced and iCLIP traces are created that detail where the protein binds throughout the transcriptome.

Cell lines expressing FLAG-tagged Alyref, Chtop and Nxf1 proteins were used by Dr Nicolas Viphakone to generate iCLIP data to investigate the association on mRNA export proteins with RNA. Interestingly, TREX components and the export receptor Nxf1 are directly associated with many types of non-coding RNA as well as with mRNA as expected. Therefore, the mRNA export pathway and its individual components might serve alternative functions in ncRNA localisation.

tRNAs were among the classes of ncRNA associated with TREX and Nxf1 in the iCLIP data set, perhaps signifying a role in tRNA export as well as mRNA export. In vertebrates, tRNAs are widely considered to be exported from the nucleus to the

cytoplasm by Exportin-T, a member of the β -importin family of proteins (Kutay et al. 1998; G J Arts et al. 1998; Gert Jan Arts et al. 1998). Some tRNA species can also be exported via a member of the same protein family, Exportin-5, although this is believed to be a minor export pathway (Calado et al. 2002; Bohnsack et al. 2002). However, organisms lacking the functional orthologues of these proteins are still viable (Hunter et al. 2003; Takano et al. 2005). This suggests that other, alternative methods of export must exist. The mRNA export pathway, alongside other RNA export pathways, has been recently implicated in tRNA export by an RNAi screen in yeast (Wu et al. 2015). A knockout of Mex67, the yeast orthologue of the mRNA export receptor Nxf1, results in nuclear accumulation of tRNA, an additional indication that other RNA export pathways may be involved in tRNA export (Wu et al. 2015). Further investigation in budding yeast confirmed the use of Mex67-Mtr2 as a tRNA export receptor (Chatterjee, Majumder, et al. 2017).

As well as tRNAs, ncRNA classes such as scaRNAs and lncRNAs were also found to be associated with TREX and Nxf1. The majority of these RNAs are retained in the nucleus so the purpose of the association is more likely to be for a process other than nuclear export. As well as the principal function of the Crm1-PHAX pathway in snRNA export to the cytoplasm, the pathway has been implicated in intra-nuclear localisation of U3 snRNA to Cajal bodies and snoRNAs to the nucleolus (Boulon et al. 2004; Pradet-Balade et al. 2011). This alternative function suggests that the Crm1-PHAX pathway may not only be defined as a nuclear export pathway, but may have wider roles in localisation. The same may well apply to the mRNA export pathway, especially when it comes to nuclear-retained ncRNAs.

3.1 TREX and Nxf1 associate with tRNAs

The initial step in determining the possible role of the mRNA export pathway in tRNA export was to confirm a direct interaction between mRNA export proteins and tRNA that was previously seen in the lab's recent iCLIP data (Figure 3.1.1). An example of a spliced and unspliced tRNA is shown and they are among the tRNAs with the highest density of iCLIP tags. FLAG-Alyref and FLAG-Nxf1 show peaks of binding to particular nucleotides of the two example tRNAs shown. These peaks are above the background level seen in the FLAG-Flp-In iCLIP set, whereas FLAG-Chtop does not have any significant peaks. The difference in binding of Alyref and Chtop to tRNAs demonstrates the dynamic nature of the TREX complex or could indicate a TREX-independent function of Alyref.

To confirm the direct interaction of Nxf1 and Alyref with tRNA, a UV RNA-immunoprecipitation (RIP) experiment was carried out using FLAG-tagged Nxf1, Alyref and GFP expressing stable cell lines. UV irradiation causes crosslinking between RNA and proteins that are associated in the cell so you can assess specific interactions. FLAG-GFP was used as a control condition as it should not bind specifically to human proteins or RNA. The RNA bound by FLAG-Nxf1, FLAG-Alyref and FLAG-GFP was analysed using a Northern blot, probed with two specific tRNA probes (Figure 3.1.2). Both tRNAs are convincingly present in the RNA associated with FLAG-Alyref and FLAG-Nxf1 in the +UV condition, demonstrating a specific and direct interaction between the tRNA and Alyref and Nxf1. There is only a very low level of tRNA associated with FLAG-GFP following UV irradiation and in the FLAG-Alyref and FLAG-Nxf1 -UV conditions. The weak signal seen in these conditions is likely to be the "background" level of non-specific tRNA association with the proteins. Thus, the iCLIP and RIP experiments together demonstrate a direct interaction between the mRNA export proteins Alyref and Nxf1 and tRNA.

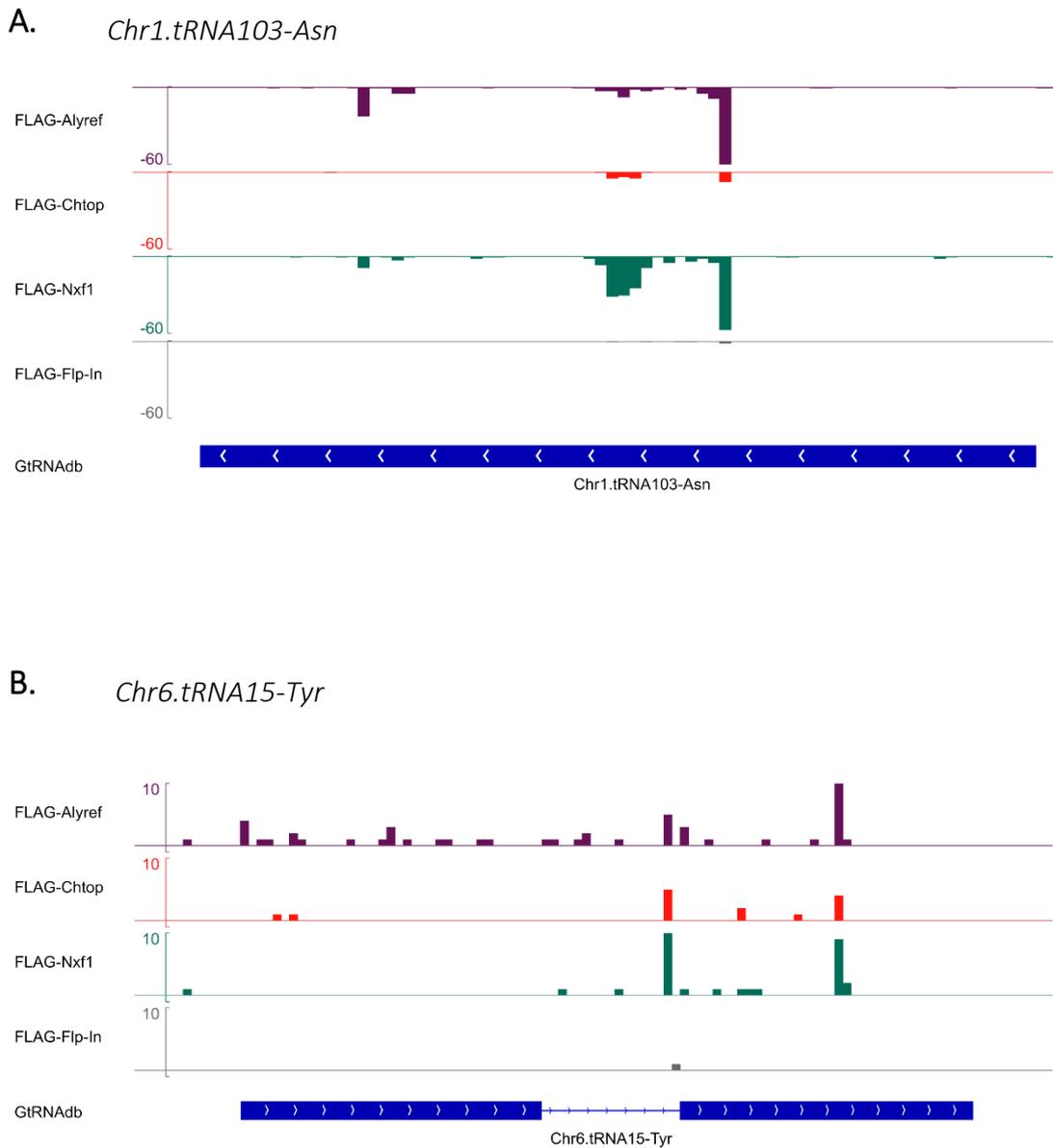


Figure 3.1.1 FLAG-Alyref and FLAG-Nxf1 are bound to tRNAs

iCLIP binding profile of FLAG-Alyref, FLAG-Chtop and FLAG-Nxf1 on (A) tRNA103-Asn and (B) tRNA15-Tyr. FLAG-Flp-In is the background control condition. The peaks represent the average peaks of three data sets. The scale is shown on the left hand side and the negative symbol denotes transcription of the gene from the negative strand. The corresponding gene from the GtRNA database (Chan & Lowe 2009) is shown as a blue bar with the thick areas of the bar representing exons.

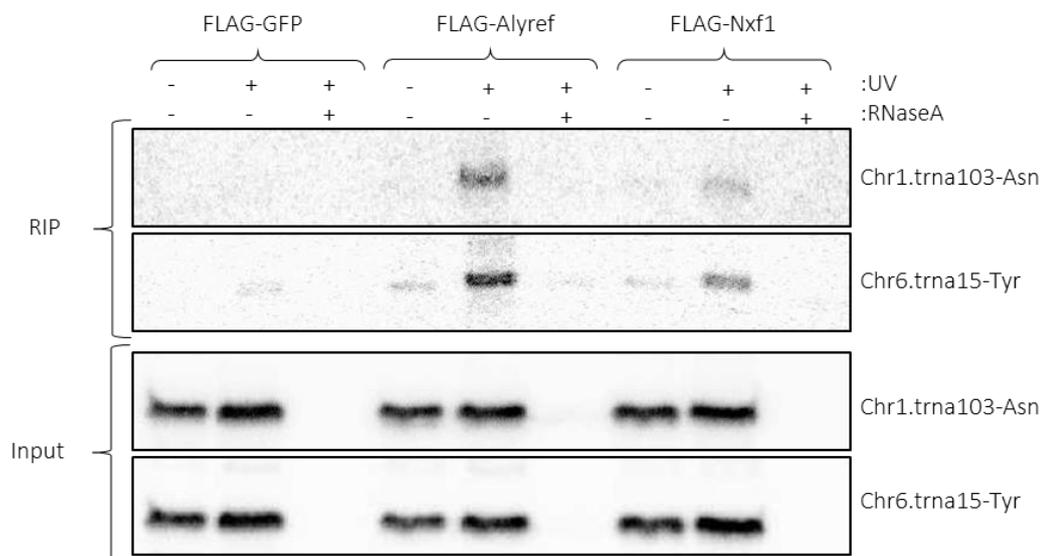


Figure 3.1.2 UV RIP confirms that FLAG-Alyref and FLAG-Nxf1 are bound to tRNAs

A Northern blot following RNA immunoprecipitation (RIP) of FLAG-GFP, FLAG-Alyref and FLAG-Nxf1 proteins. Cell samples that have been UV irradiated or RNase A treated are indicated by the + symbol. All RNA from the RIP was run on the gel and 5 μ g run as the input sample. The Northern was probed with 32 P – labelled probes designed for specific tRNAs. This blot is representative of one RIP experiment.

3.2 The depletion of TREX does not perturb the cellular localisation of tRNA, determined by RNA fractionation

Due to the known function of TREX and Nxf1 in the export of mRNAs, their function in the lifecycle of tRNA would most likely be in tRNA export from the nucleus to cytoplasm. Depleting 293T cells of the canonical tRNA export proteins should lead to a nuclear accumulation and/or cytoplasmic depletion of tRNA. Therefore, if the mRNA export pathway is involved in tRNA export, depletion of the TREX complex alongside the canonical pathway should have a cumulative effect on the cellular mislocalisation of tRNA. TREX subunit and Nxf1 protein levels were depleted using RNA-interference (RNAi) stable-cell lines. The cell lines can be induced to express an RNAi hairpin that is processed to a siRNA that will then target the mRNA of the protein of interest. The exportin proteins were depleted using transfected siRNA (Figure 3.2.1). All proteins were depleted by Nuclear and cytoplasmic RNA was extracted following RNAi or siRNA treatment and analysed by Northern blot (Figure 3.2.2). The Northern blot was probed with a tRNA probe and a 5.8S rRNA probe as a loading control. Northern blot analyses were performed on three biological replicate samples. Representative datasets are shown in figures 3.2.2 to 3.2.4

An initial inspection of tRNA abundance in the tested conditions did not reveal any differences, suggesting that depletion of any of the four proteins has little effect on tRNA localisation. To confirm this observation, comparative quantitation was necessary to determine the relative amount of tRNA in each fraction of each condition. The relative abundance of tRNA in the conditions can be represented as a ratio of the abundance in the cytoplasm (C) compared to the nucleus (N). All conditions are normalised to the ratio in the control conditions, Control siRNA or Control RNAi. A ratio of less than 1 in the non-control conditions indicates a nuclear enrichment of tRNA compared to the control where as a ratio of more than 1 indicates a cytoplasmic enrichment. The C:N ratio decreases mildly when Exportin-5 is depleted and decreases more following Exportin-T or Exportin-T and 5 depletion together (Figure 3.2.3).

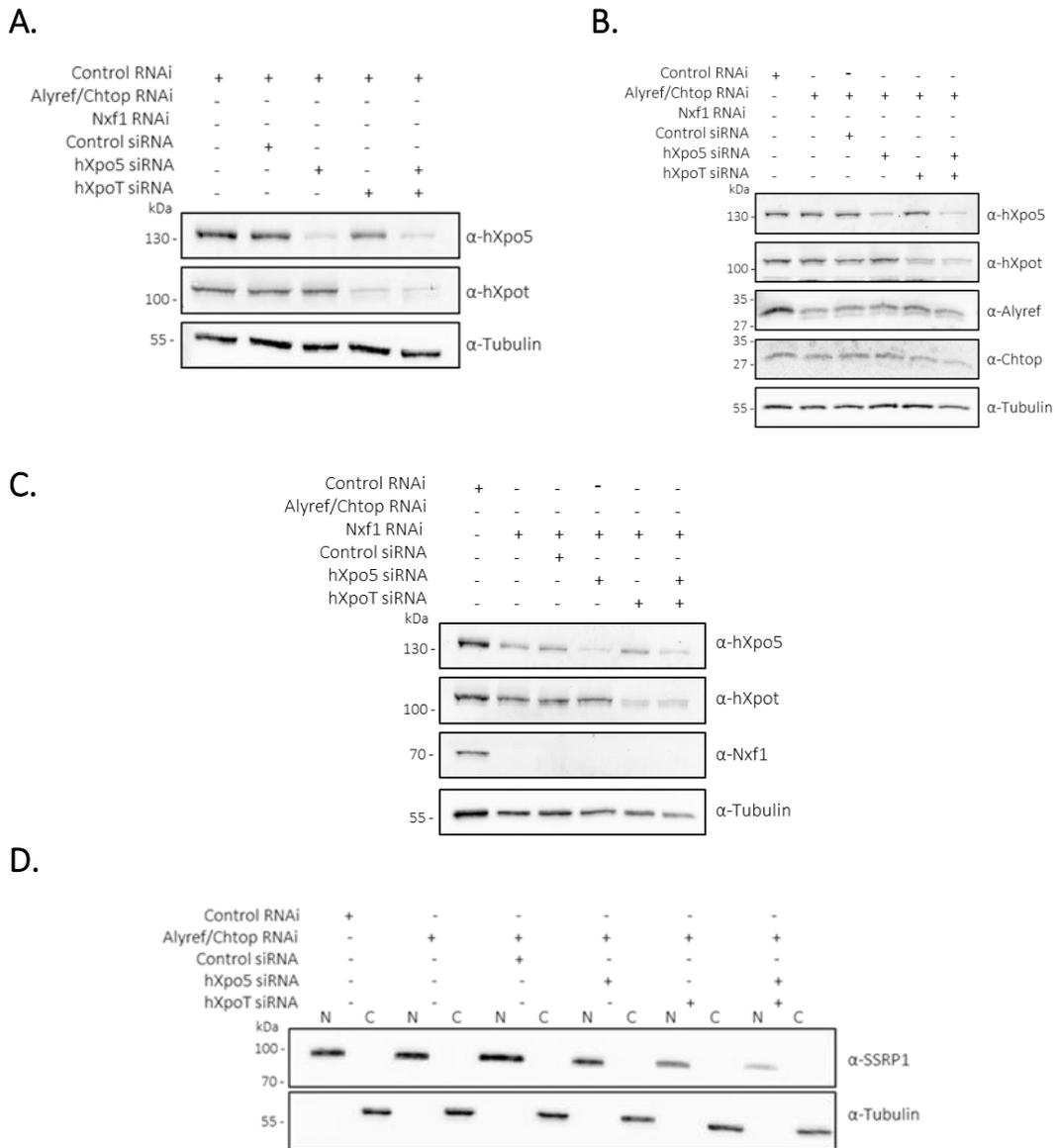


Figure 3.2.1 Validation of RNAi conditions and cellular fractionation by Western blot

Western blots of the TREX, Nxf1 and exportin protein knockdowns in the (A) Control RNAi (B) Alyref/Chtop RNAi and (C) Nxf1 RNAi cell backgrounds. Tubulin is shown as a loading control on blots A-C. RNAi treatments were carried out for 72 hours for Control and Nxf1 and 48 hours for Alyref/Chtop. All siRNA treatments were carried out for 72 hours, with 10 nM transfections at 0 and 48 hours. (D) The cleanliness of nuclear and cytoplasmic fractionation was assessed by Western blotting, using protein extracts from the same cells used for the RNA extraction. The cytoskeletal protein Tubulin and the nucleosome organisation factor SSRP1 are used as the cytoplasmic and nuclear markers respectively.

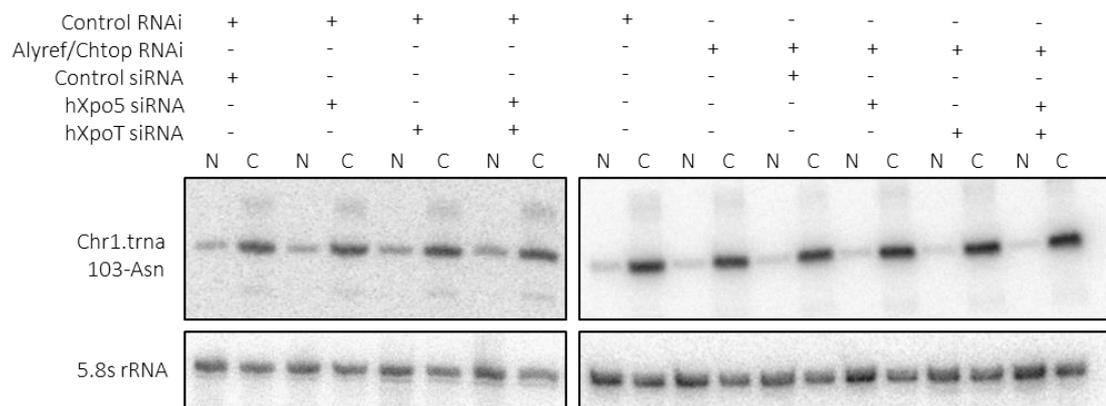


Figure 3.2.2 Depletion of the canonical tRNA export pathway and the mRNA export pathway does not result in a visible effect on tRNA localisation by Northern blot

Northern blots of nuclear (N) and cytoplasmic (C) RNA from RNAi/siRNA conditions described in Figure 3.2.1. Blots were hybridised with a ^{32}P -labelled tRNA probe and a 5.8S rRNA probe as a loading control. 5 μg of RNA was loaded in each lane. The blots were repeated three times with three sets of biological repeats and a representative image is shown in this figure. Quantification of the representative blots are shown in Figures 3.2.3 and 3.2.4.

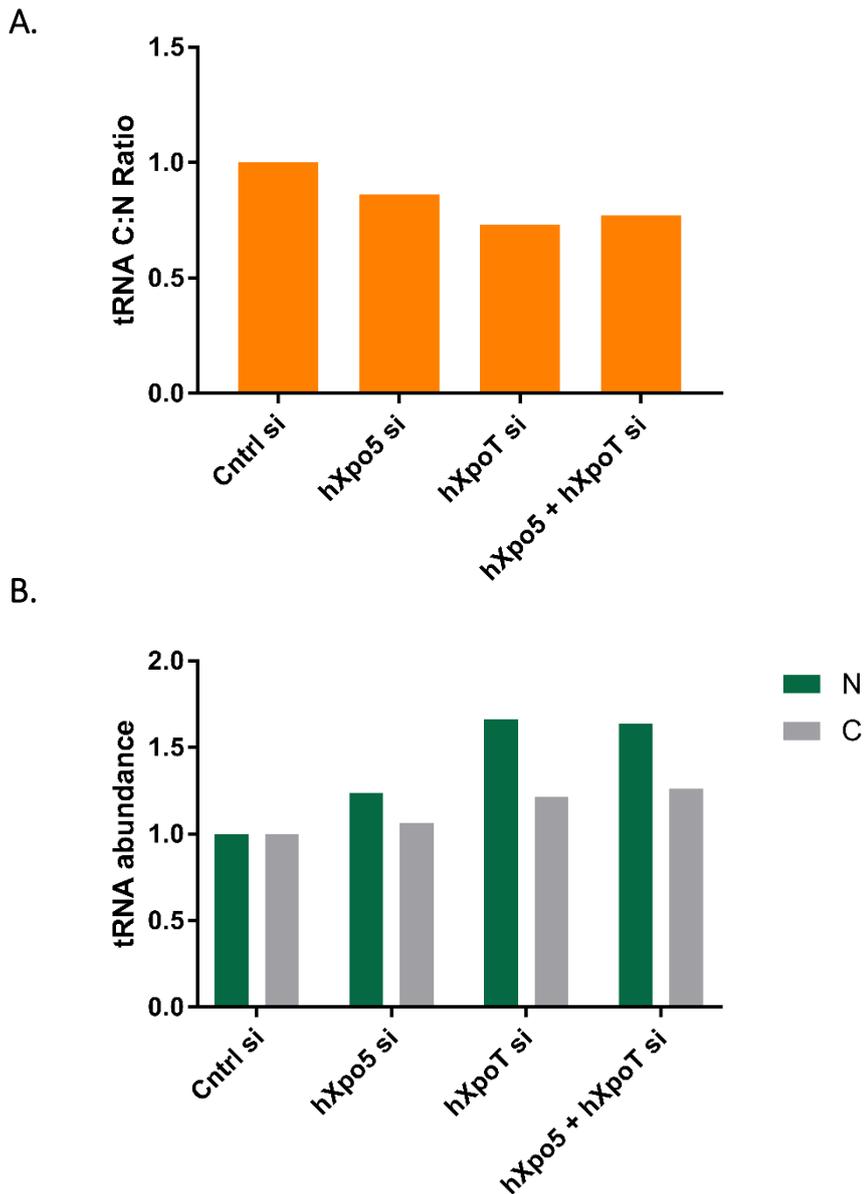


Figure 3.2.3 Quantification of the abundance of tRNA following exportin protein depletion reveals minor changes in abundance

Quantification of the abundance of tRNA from the representative Northern blots in Figure 3.2.2, calculated by using pixel density or “volume analysis”. The data presented here is therefore from one biological repeat. The abundance is represented as (A) a cytoplasmic:nuclear ratio (C:N) per condition and (B) of each fraction independently. The C:N ratio and the abundance in each independent fraction is shown relative to the Control siRNA ratio or abundance. The amount of tRNA in each fraction was normalised to the corresponding 5.8S rRNA control. All siRNA treatments were applied to Control RNAi cells.

This suggests a nuclear accumulation and/or cytoplasmic reduction in tRNA. This result endorses the current view of Exportin-T as the major and Exportin-5 as the minor player in tRNA export as there is a larger effect on tRNA localisation following Exportin-T depletion.

However, this is a relatively minor decrease in ratio considering that these two proteins are widely understood to be the main export proteins for tRNA in vertebrates. If these are the major tRNA export proteins in humans, a much starker result would be expected. On the other hand, maybe the cell can function relatively normally with the residual amounts of Exportin-T and 5 protein left after siRNA treatment (Figure 3.2.1).

To understand why the ratios may be decreasing, the relative abundance of tRNA in each fraction independently is shown (Figure 3.2.3.B). Following Exportin-5 depletion, there is a minor increase in the relative amount of tRNA in the nuclear fraction. Following Exportin-T and Exportin-T and 5 depletion together there is a greater increase in tRNA in the nucleus. However, in all cases there is also an increase in the amount of tRNA in the cytoplasm. The increase in the nucleus is larger than in the cytoplasm, giving the decrease in C:N ratios, suggesting that some tRNA is accumulating in the nucleus. However, as there is an increase in tRNA in both fractions, perhaps the change in abundance is due to factors other than export effects, such as a general transcriptional upregulation of tRNA genes.

This experiment was repeated with three biological replicates and the same relationship between the conditions was seen. However, the variation in quantified values from the volume analysis was too great to present on the graph so a representative Northern blot is shown and quantified.

Following depletion of Exportin-T, Exportin-5 and TREX combined (Figure 3.2.4), a decrease in C:N ratio is not seen and the relative abundances suggest a decrease in the amount of nuclear tRNA and an increase in the amount of cytoplasmic tRNA. The C: N ratios actually show an increased ratio when TREX is depleted alongside the two exportin proteins, suggesting a minor increase in export of tRNA to the cytoplasm.

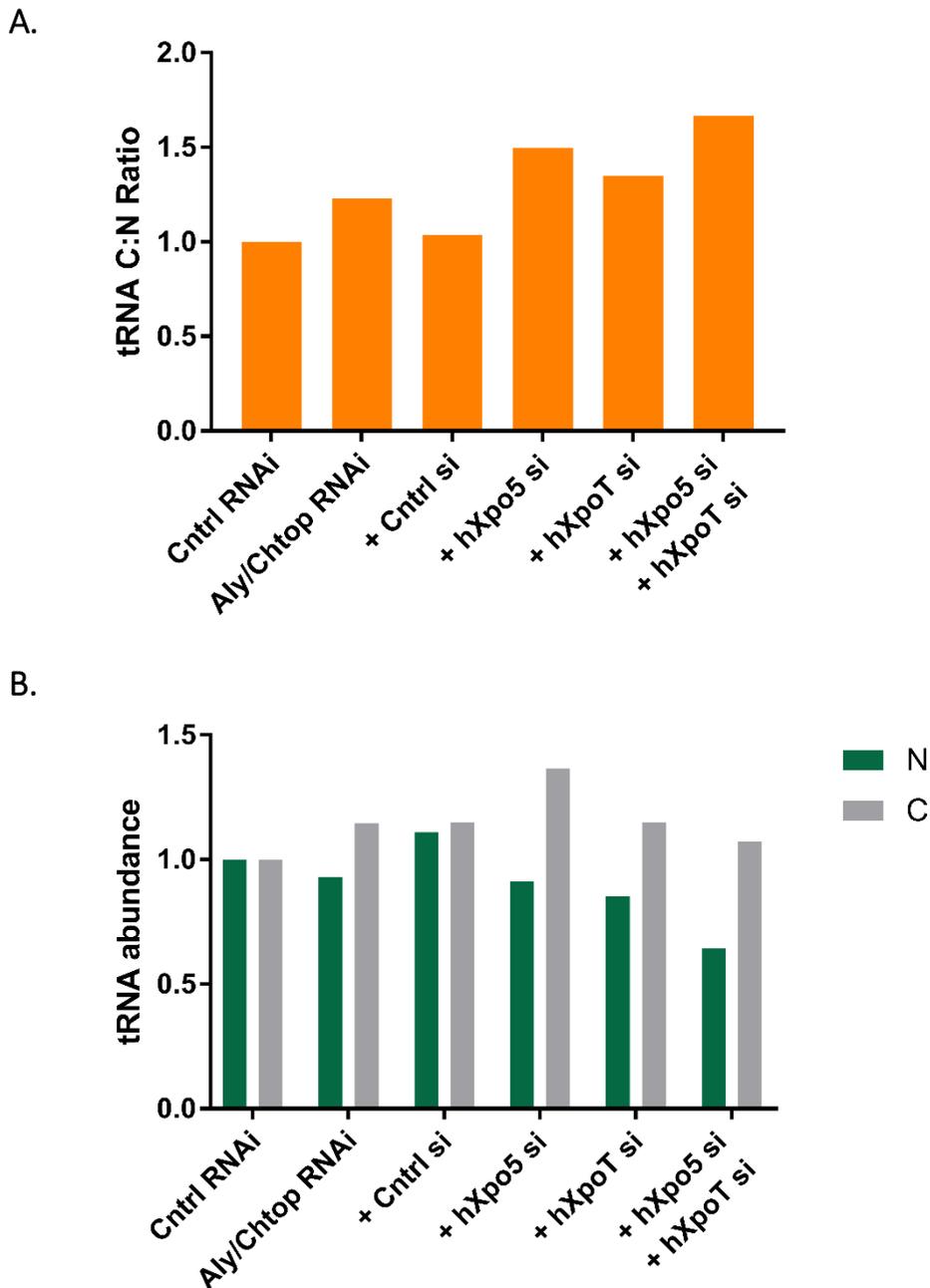


Figure 3.2.4 Quantification of the abundance of tRNA does not reveal a considerable effect on tRNA levels following exportin and TREX protein depletion together

Quantification of tRNA abundance and conversion to tRNA C:N ratio was carried out as in Figure 3.2.3. The data presented here is therefore from one biological repeat. All siRNA treatments shown were applied to Alyref/Chtop RNAi cells. Control siRNA treatment was applied to Control and Alyref/Chtop RNAi cells. **(A)** C:N ratios and **(B)** independent abundances are shown relative to the Control RNAi condition.

However, the Northern blot with the Alyref/Chtop RNAi background conditions was repeated with three biological replicates and the minor effects seen here were not reproducible. Therefore, a representative Northern blot is shown and quantified. Even if these effects were reproducible, they show very minor changes to tRNA abundance in the nucleus and cytoplasm. The majority of the effects discussed did not reach above a 1.5 fold change over control, even when the supposed canonical tRNA export pathway was depleted (Figure 3.2.2).

However, a positive control for nuclear accumulation of transcripts is needed to establish whether the type of assay is preventing the visualisation of any effects. The RNAi conditions and cell lines used were the same as published data which show a block in mRNA export and accumulation of mRNA in the nuclear fraction (Chang et al. 2013). As mRNA accumulates in the nucleus under these conditions, a probe to a specific housekeeping mRNA or to general poly(A)⁺ tails could have been used to confirm that you can see general nuclear accumulation on a Northern blot following cellular fractionation. However, a much larger Northern blot would have been needed to detect mRNAs. Following Alyref/Chtop RNAi, protein depletion was confirmed by Western blot but the published mRNA nuclear accumulation phenotype was not assayed by poly(A)⁺ FISH (Chang et al. 2013). This control would have also helped to confirm the results seen. To rule out the possibility that it is the type of assay that is preventing us from seeing any considerable effects on tRNA localisation, an alternative technique was employed in Chapter 3.3.

3.3 The depletion of TREX and Nxf1 does not perturb the cellular localisation of tRNA as determined by FISH

In parallel to the Northern blot, Fluorescence In Situ Hybridisation (FISH) was carried out to look directly at tRNA localisation in cells following protein depletion (Figure 3.3). The same conditions were used as in Chapter 3.2 but with the Nxf1 RNAi cell line in addition. Nxf1, the mRNA export receptor, acts at the late stage of mRNA export, allowing us to compare the effect of depleting early stage export proteins, Alyref and Chtop, and late stage Nxf1.

Cy3-labelled tRNA-Lys is predominately seen in the cytoplasm, as shown when overlaid with DAPI staining of DNA, similarly to the localisation of tRNA seen by the Northern blot (Figure 3.2.2). Exportin-T and/or Exportin-5 reduction doesn't have a clearly visible effect on tRNA localisation in Control RNAi cells. In the Alyref/Chtop and Nxf1 RNAi conditions with Exportin-T siRNA, the nuclei appear redder compared to the control siRNA condition, perhaps suggesting slight nuclear accumulation.

The images taken with the light microscope were not of high enough quality to quantify the Cy3-tRNA signal distribution using standard quantification methods such as ImageJ. Taking images using a confocal microscope would have improved the resolution and the contrast between areas of the cell with Cy3-tRNA staining and those without. Images taken in this fashion may have then been quantifiable by intensity scanning. However, the potentially subtle effects of steady-state tRNA mislocalisation may still not be apparent. Similarly to Chapter 3.2, positive controls are needed to confirm that the FISH assay can demonstrate nuclear accumulation and that the RNAi and siRNA conditions are resulting in expected phenotypes as well as showing protein depletion (Figure 3.2.1). A general poly(A)⁺ FISH probe should have been used to confirm nuclear accumulation of mRNA following Alyref/Chtop and Nxf1 RNAi as published (Viphakone et al. 2012; Chang et al. 2013).

Overall, the FISH assay seems to show a similar result to the Northern blot; depletion of any of the proteins, canonical tRNA export exportins and mRNA export proteins, does not greatly effect steady-state tRNA localisation.

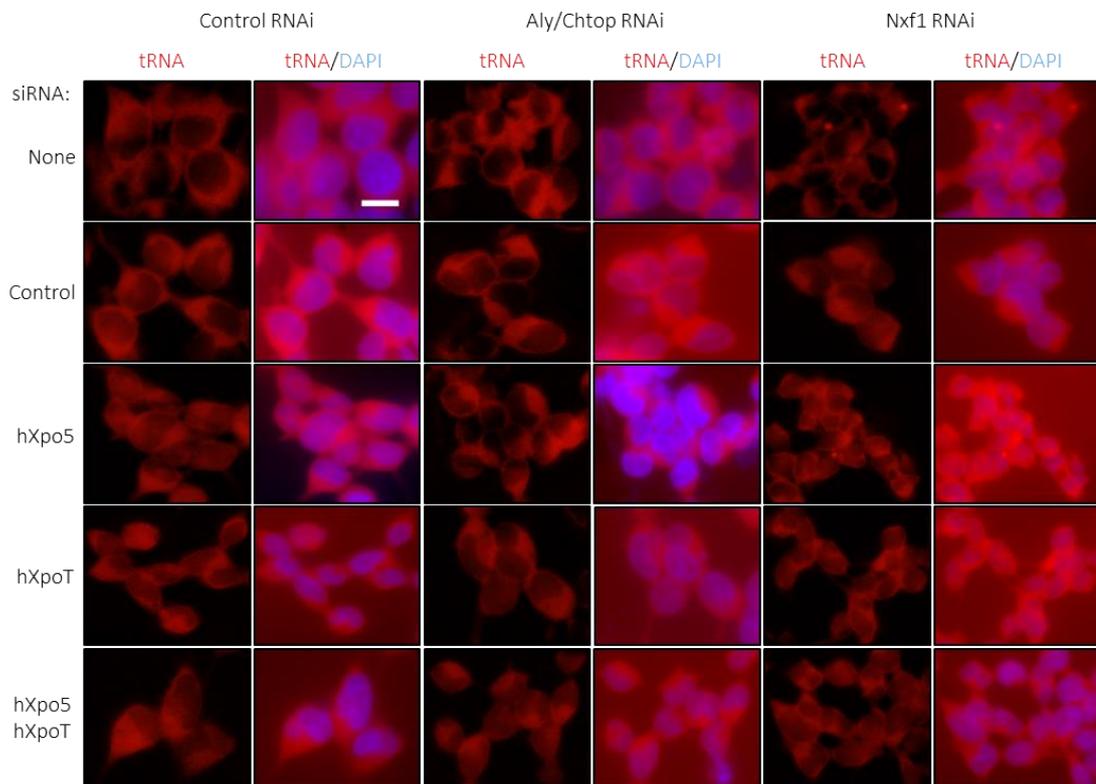


Figure 3.3. tRNA FISH following depletion of exportin proteins individually or alongside TREX and Nxf1 does not demonstrate tRNA mislocalisation

RNAi and siRNA treatments were carried out as described in Figure 3.2.1 before cells were fixed and hybridised with a Cy3-labelled tRNA probe. Cells were also stained with DAPI. Whole mount images were taken using a Leica light microscope at 63x magnification using DAPI and Cy3 filter cubes. An overlay of DAPI and Cy3-tRNA images is also shown. The white bar represents 10 μ m.

3.4 The protein levels of Exportin-T and 5 do not change considerably following the depletion of TREX and Nxf1

A link between mRNA export proteins and Exportin-T and 5 may be found by considering any change in the level of exportin proteins following depletion of TREX and Nxf1 (Figure 3.4). If there is substantial upregulation in protein levels, this might indicate a shared function as one protein compensates for the loss of the other. The levels of exportin proteins are constant in the Alyref/Chtop RNAi condition, suggesting that there isn't a direct link between the two pathways. However, defects in mRNA export can mask any upregulation as there may be less exportin mRNA exported and therefore less protein expressed. Therefore, Alyref RNAi on its own is included as the block on mRNA export is not as strong. Alyref and Nxf1 RNAi result in a slight increase in the amount of Exportin-T, indicating a potential compensatory mechanism. However, in comparison to the upregulation of Chtop following Alyref and Nxf1 depletion as they function in the same pathway, I suspect it is not a significant result.

Nxf1 RNAi results in less Exportin-5 protein which is likely to be due to the defect in mRNA export, which Exportin-5 mRNA may be more sensitive to than Exportin-T. On the other hand, the increase in Exportin-T levels following Nxf1 RNAi may actually be greater if it wasn't for the reduction in mRNA export efficiency. This assay may not be particularly helpful in determining any links between the pathways due to the function of the proteins being depleted. Ideally, the Western could be carried out in reverse by depleting the exportin proteins and looking at TREX and Nxf1 levels. Although, this could have the same issues if there was an effect on translation if less tRNA was reaching the cytoplasm. However, as my data in Chapter 3.2 and 3.3 presents, there seems to be little effect on tRNA localisation following depletion of Exportin-T and 5 so the reverse Western may have been more useful in forming a link between the pathways.

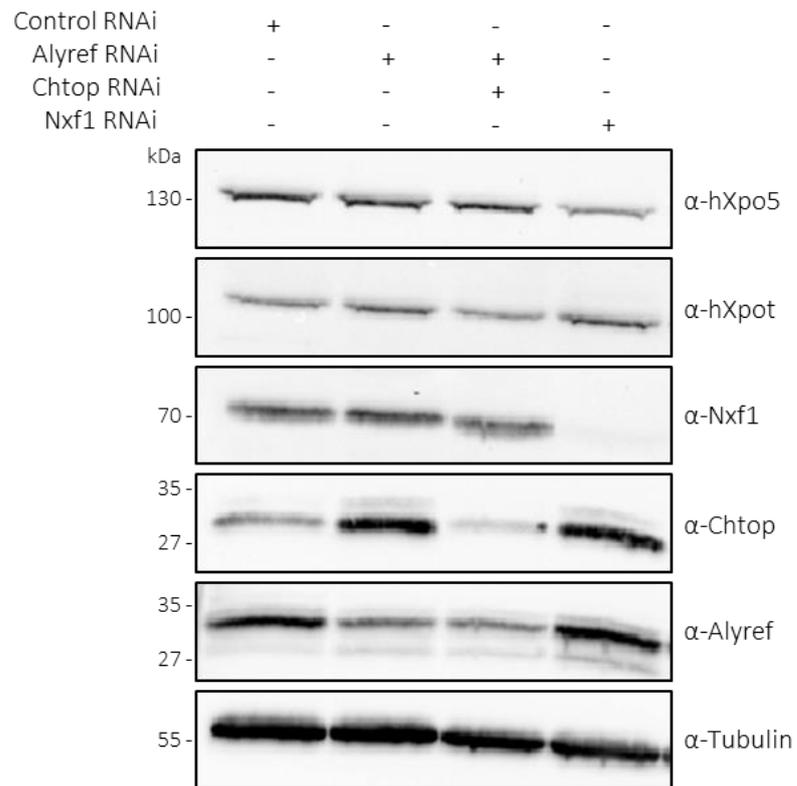


Figure 3.4. Upregulation of exportin proteins is not seen when mRNA export proteins are depleted

Western blot following Alyref, Alyref/Chtop and Nxf1 depletion using RNAi cell lines. Protein depletion was carried out for 72, 48 and 72 hours respectively. The control RNAi cell line was induced for 72 hours also. Tubulin is shown as a loading control.

If an effect on tRNA localisation was revealed by cellular fractionation and FISH, Figure 3.4 would have been necessary to confirm that any changes seen in the Alyref/Chtop RNAi condition were not due to the effects of disrupting the expression of the Exportin-T and Exportin-5 transcripts following an mRNA export block. As there doesn't seem to be a substantial effect, the results seen can be taken at face value.

3.5 TREX and Nxf1 are not directly involved in the intra-nuclear localisation of *scaRNA2*

From the iCLIP data generated in the lab, scaRNAs are bound by FLAG-Alyref, FLAG-Chtop and FLAG-Nxf1. An example scaRNA, *scaRNA2*, is shown in Figure 3.5.A. As scaRNAs function in the nucleus and are localised to Cajal bodies (Richard et al. 2003), the connection between the mRNA export pathway and scaRNAs could be in intra-nuclear transport, perhaps similarly to the Crm1-PHAX pathway in U3 snoRNA localisation (Boulon et al. 2004). scaRNA localisation can be investigated by assessing the co-localisation of scaRNA with Coilin protein, a fundamental Cajal body component (Raška et al. 1991). *scaRNA2* was chosen as an example scaRNA for this experiment as it had the highest density of CLIP tags.

The extent of *scaRNA2* and Coilin co-localisation was assessed following depletion of the mRNA export proteins using siRNAs in HeLa cells (Figure 3.5.B). In the control conditions, the intensity peaks of *scaRNA2* and Coilin signals strongly overlap as expected. Following Nxf1 and TREX knockdown, the peaks still overlap, indicating that *scaRNA2* localisation to Cajal bodies has not been disrupted. Interestingly, the peaks of *scaRNA2* seem less defined in the Nxf1 siRNA condition, suggesting that Nxf1 may be needed for efficient or complete localisation to the Cajal bodies or to maintain nuclear levels of *scaRNA2*. The total levels of *scaRNA2* could be assessed in TREX and Nxf1 knockdown conditions, although this would not help to conclude whether there was less *scaRNA2* out of the total amount reaching the Cajal bodies or if there was less *scaRNA2* present in the cell. Nonetheless, *scaRNA2* is still generally localised to the Cajal bodies following Nxf1 depletion so the widespread localisation mechanisms do not require Nxf1 or TREX. To confirm this result however, poly(A)⁺ FISH should also have been carried out to confirm the phenotypic effect of mRNA accumulation in the nucleus following Alyref/Chtop and Nxf1 depletion (Viphakone et al. 2012; Chang et al. 2013). Despite confirming adequate protein depletion by Western blot, we cannot be certain that Alyref/Chtop and Nxf1 depletion is affecting the cells as expected.

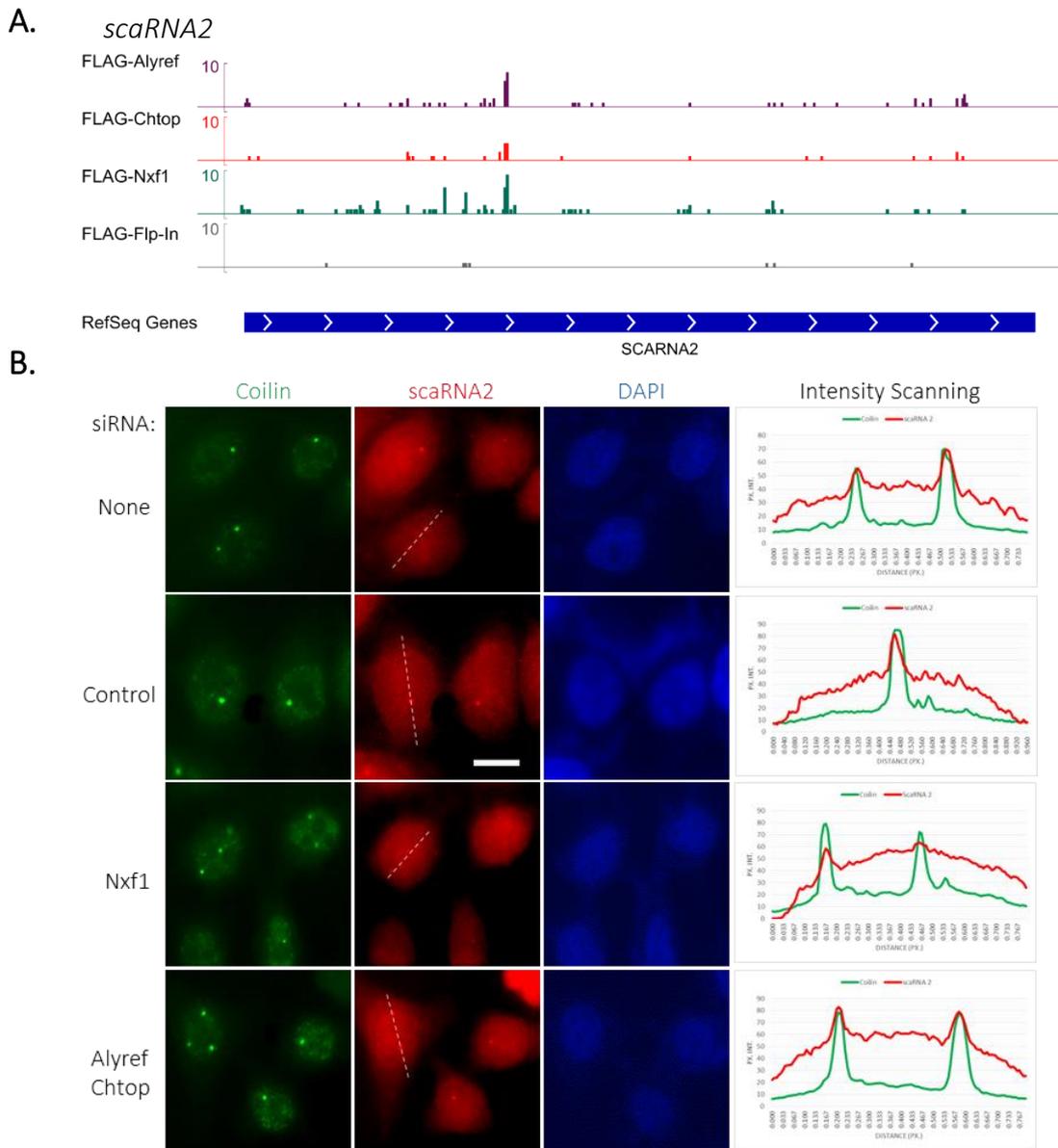


Figure 2.5 TREX and Nxf1 depletion does not affect the intra-nuclear localisation of *scaRNA2* to Cajal bodies

(A) iCLIP binding profile of FLAG-Alyref, FLAG-Chtop and FLAG-Nxf1 on *scaRNA2*. FLAG-Flp-In is the background control condition. The scale is shown on the left hand side. (B) Whole mount images of HeLa cells transfected with 10 nM of corresponding siRNA(s) for 72 hours. A Cy3-*scaRNA2* probe was used to visualise *scaRNA2* and a Coilin antibody was used to stain the Cajal bodies. The secondary antibody was conjugated with GFP. Images were taken with a Leica light microscope at 63x magnification using DAPI, Cy3 and GFP filter cubes. Intensity scanning was used to quantify the amount of each signal, *scaRNA2* or Coilin, along a strip across the cell, represented by the dashed white line. The white bar represents 10 μ m.

3.6 TREX and Nxf1 are not directly involved in the Intra-nuclear localisation of *XIST*

From the iCLIP data, TREX and Nxf1 are bound to lncRNAs such as *XIST* and *FIRRE* (Figure 3.6.1). FLAG-Alyref and FLAG-Chtop proteins are bound to a greater extent than FLAG-Nxf1, but are all above the background level seen in the control FLAG-Flp-In cell line. The lower density of FLAG-Nxf1 iCLIP tags on these RNAs when compared to FLAG-Alyref and FLAG-Chtop possibly reflects the nuclear enrichment of *XIST* and *FIRRE* as they are unlikely to be exported to the cytoplasm in normal cellular conditions. Interestingly, the binding of all three proteins to both lncRNAs seems to be enriched on the exonic regions of the RNA transcript. As these lncRNAs are mainly nuclear restricted, the role of the mRNA export pathway in lncRNA biogenesis and function may be in intra-nuclear transport as discussed in Chapter 3.5.

XIST RNA is localised to the silenced X chromosome (Brown et al. 1992), so the role of the mRNA export pathway in the intra-nuclear localisation of *XIST* can be investigated in a similar way to *scaRNA2* in Chapter 3.5. TREX and Nxf1 proteins were depleted using stable RNAi cell lines and FISH was carried out using a Cy3-labelled probe that hybridises to *XIST* RNA (Figure 3.6.2.A). Heterogeneous nuclear ribonucleoprotein U (hnRNP U) has been comprehensively shown to be necessary for *XIST* localisation to the X chromosome (Hasegawa et al. 2010), and is therefore shown as a positive control to demonstrate *XIST* mislocalisation following its depletion (Figure 3.6.2.B). hnRNP U depletion using 96 hours of siRNA treatment results in a diffuse *XIST* signal as there are fewer distinct foci compared to the control siRNA condition where the bulk of the Cy3 signal is seen in one focal point in each cell, the X chromosome. Following TREX and Nxf1 depletion, the *XIST* signal remains largely localised at one point, suggesting that TREX and Nxf1 are not involved in the intra-nuclear localisation of *XIST* to the X chromosome.

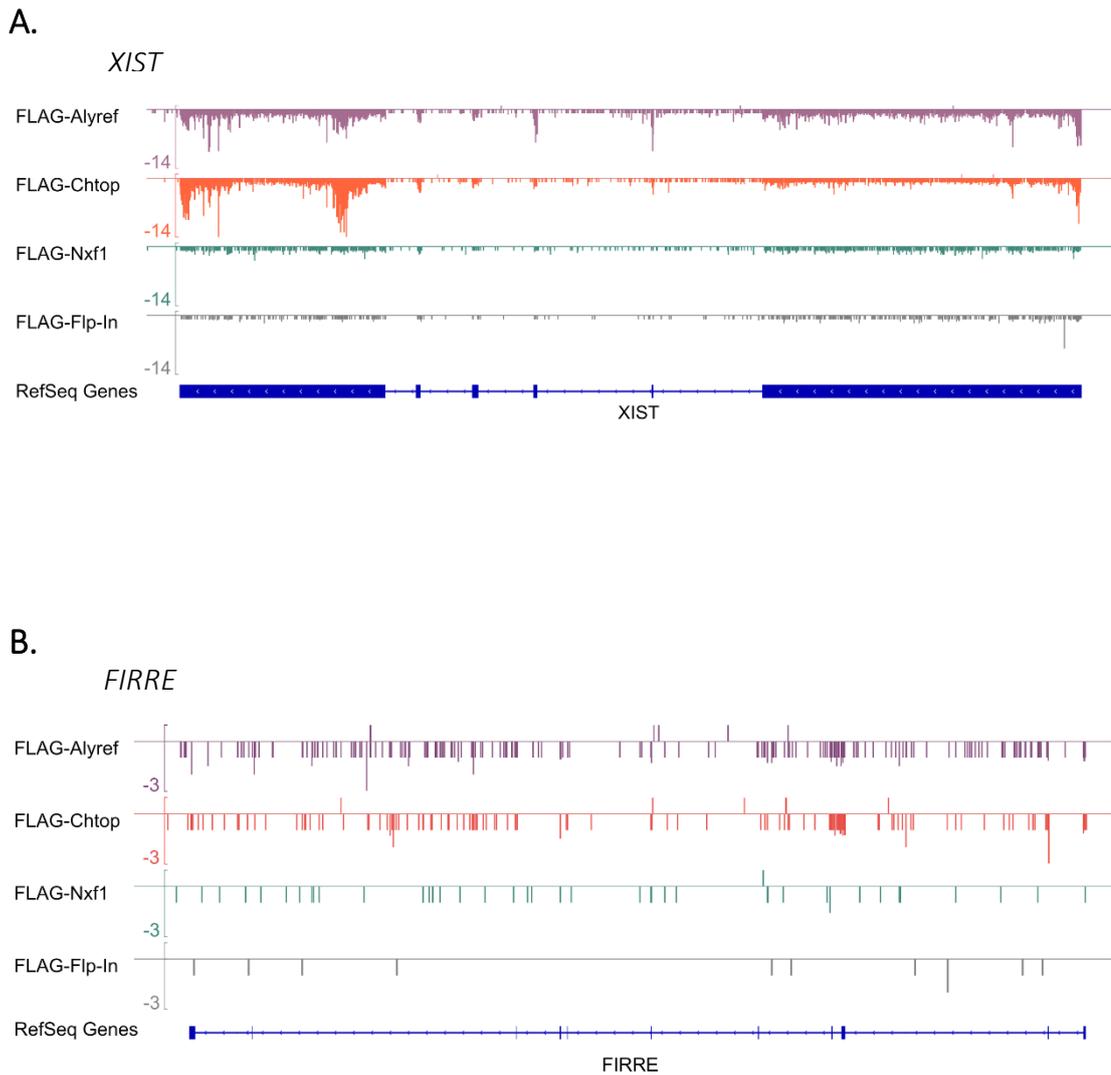


Figure 3.6.1. FLAG-Alyref, FLAG-Chtop and FLAG-Nxf1 are bound to lncRNAs *XIST* and *FIRRE*

iCLIP binding profiles of FLAG-Alyref, FLAG-Chtop and FLAG-Nxf1 on **(A)** *XIST* and **(B)** *FIRRE* lncRNAs. FLAG-Flp-In is the background control condition. The RefSeq gene is displayed as a blue bar with the thicker areas of the bar representing exons. The scale is shown on the left hand side and the negative symbol denotes transcription of the gene from the negative strand.

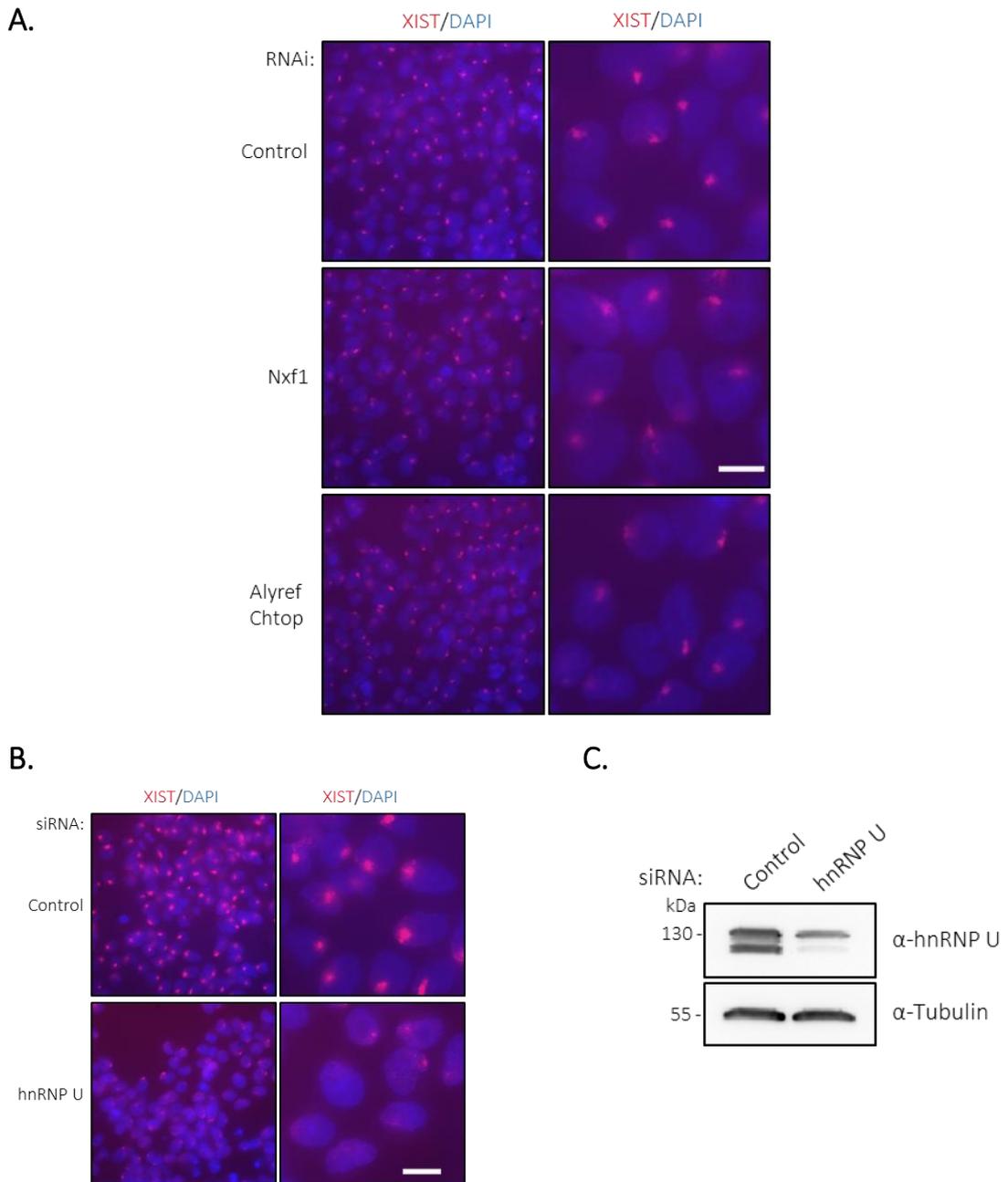


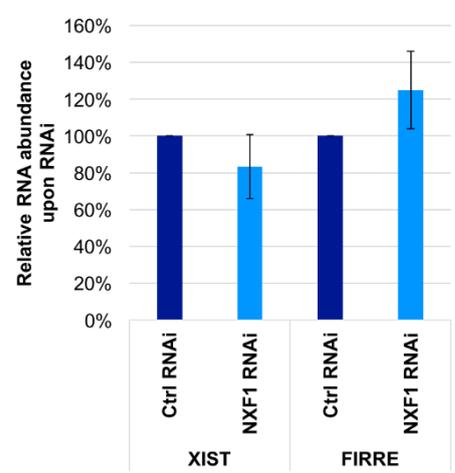
Figure 3.6.2. TREX and Nxf1 depletion does not lead to mislocalisation of *XIST* RNA

Treated 293T cells were hybridised with Cy3-labelled *XIST* RNA probes and stained with DAPI. Whole mount images were taken using a Leica light microscope at 40x and 63x magnification with DAPI and Cy3 filter cubes. The white bar represents 10 μ m. **(A)** *XIST* FISH in Control, Nxf1 and Alyref/Chtop RNAi cell lines. RNAi treatments were carried out for 72, 72 and 48 hours respectively before the cells were fixed and stained. **(B)** *XIST* FISH in 293T cells transfected with 25 nM Control or hnRNP U siRNA for 96 hours. **(C)** Western blot following hnRNP U siRNA treatment. Tubulin is shown as a loading control.

3.7 TREX but not NXF1 is required to maintain XIST RNA levels

In Chapter 3.6, the localisation of *XIST* was shown to not be effected by TREX and Nxf1 depletion. However, in Figure 3.6.2, there looks to be a weaker Cy3-*XIST* RNA signal in the Alyref/Chtop RNAi condition so the abundance of *XIST* and another lncRNA, *FIRRE*, was assessed in both RNAi conditions (Figure 3.7). The abundance of *XIST* or *FIRRE* is not affected by the depletion of Nxf1, however, the depletion of TREX results in a significant decrease in the abundance of these lncRNAs. TREX may therefore be required to maintain the levels of these nuclear lncRNAs in the cell, independently of the mRNA export pathway with Nxf1.

A.



B.

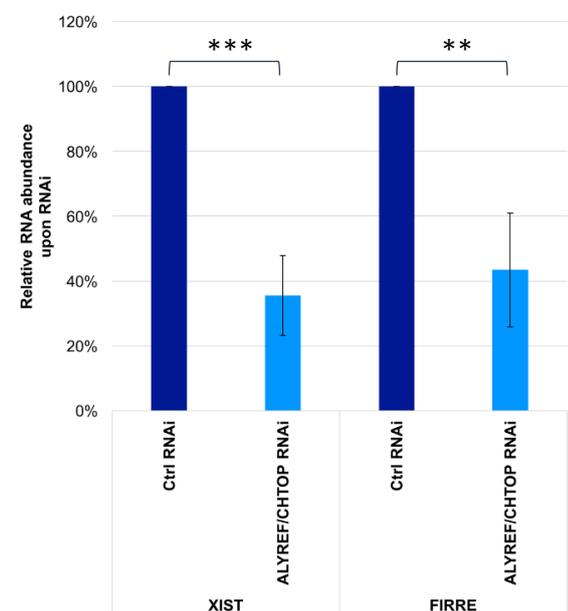


Figure 3.7. TREX but not Nxf1 is required to maintain the levels of *XIST* and *FIRRE*

Figure 3.7 was kindly provided by Dr Nikolas Viphakone who carried out the experiment.

The abundance of lncRNAs *XIST* and *FIRRE* was assessed by qRT-PCR in (A) Nxf1 RNAi and (B) Alyref/Chtop RNAi conditions. RNAi treatment was carried out for 72 hours for the Nxf1 RNAi condition and 48 hours for the Alyref/Chtop RNAi condition. The mean abundance of three replicates is shown and the error bars represent the SEM. All abundances are normalised to the abundance of U1 in the sample and are shown relative to the mean abundance in the Control RNAi condition. A t-test was carried out as described in Chapter 2 to determine the statistical significance.

3.8 Chapter 3 Summary

iCLIP data from the lab shows that key proteins in mRNA export are bound to a variety of non-coding RNA species as well as mRNAs. The role of mRNA export proteins in the life of these RNAs was assessed by looking at a potential role in their export and localisation. The export of tRNAs from the nucleus to the cytoplasm was not perturbed by the loss of mRNA export proteins, suggesting an alternative role of TREX and Nxf1 in their lifecycle. mRNA export proteins may be required for correct processing of tRNA, most of which occurs in the nucleus, or for their stability and prevention of degradation.

However, very little effect was seen on tRNA localisation following depletion of the supposed canonical vertebrate tRNA export proteins either. We would expect a relative increase in the amount of tRNA in the nucleus and a reduction in the cytoplasm if tRNA export was blocked. Exportins T and 5 have not been definitively established as the main exporters of tRNA in humans as most protein work has been carried out through injection of the human proteins into *Xenopus* oocytes (G J Arts et al. 1998; Kutay et al. 1998; Calado et al. 2002). Perhaps, the proteins may not function equivalently in humans or additional proteins may exist to export tRNA. This would be an interesting area of research to pursue, possibly through a genome-wide screen. Conversely, perhaps the cell is capable of conducting tRNA export efficiently with the residual amount of Exportin-T and 5 left after depletion using siRNA (Figure 3.2.1). If this is this case, an inducible knockout cell line would be needed for this experiment as a conventional knockout line of these proteins alongside mRNA export proteins may not survive.

Additionally, perhaps steady-state tRNA is too stable to see any change in localisation using the Northern blotting and FISH methods employed here. A reporter tRNA that can be followed kinetically may be required to accurately assess the effect of depleting mRNA and tRNA export pathways on tRNA localisation. Furthermore, appropriate positive controls should have been used to confirm the results seen in this chapter.

The intra-nuclear localisation of *scaRNA2* and *XIST* RNA was also investigated, after strong iCLIP binding profiles of TREX and Nxf1 on these RNAs were seen. The localisation of either of these RNAs to their site of function did not seem to be effected by the loss of the mRNA export proteins, much like tRNAs. However, the abundance of *XIST* and *FIRRE* was reduced upon TREX knockdown but not upon Nxf1 depletion. This indicates a role of Alyref and Chtop in maintaining *XIST* RNA levels in the cell, independent of the mRNA export pathway. Perhaps Alyref and Chtop are needed for correct processing of lncRNAs and Nxf1 is bound to *XIST* RNA in preparation for export but is prevented of doing so due to other cellular components retaining *XIST* in the nucleus. TREX and or Nxf1 could be involved in processing *scaRNAs* rather than localising them, possibly similarly to lncRNAs and tRNAs.

Chapter 4 – Using the Auxin-Inducible Degron system to deplete hnRNP U from human cells

iCLIP data from the Wilson lab shows TREX and Nxf1 bound to many non-coding RNAs, despite their established role in mRNA nuclear export. The potential function of TREX and Nxf1 in the localisation of tRNA, *XIST* RNA and *scaRNA2* was investigated in the previous chapter but the mRNA export proteins do not seem to be required for the general localisation of these non-coding RNAs. The majority of tRNAs are localised to the cytoplasm whereas *XIST* and *scaRNA2* are nuclear retained, highlighting a profound difference between these types of RNA despite being bound by mRNA export proteins. Many proteins bind to non-coding RNAs, such as lncRNAs, but the function is not necessarily known (Li et al. 2015). Therefore, there may be other factors involved that link mRNA export proteins to various RNA classes and instigate different functions for those interactions. To establish a role of the mRNA export proteins in the lifecycle of the non-coding RNAs they are bound to, we investigated the hnRNP U protein as it is known to interact with mRNAs and non-coding RNAs alongside RNA-binding proteins such as the mRNA export protein Alyref (Fackelmayer et al. 1994; Göhring et al. 1997; Xiao et al. 2012; Li et al. 2015). hnRNP U interacts with many nuclear-retained lncRNAs, but does not seem to bind tRNAs which are exported, so could provide a link between mRNA export proteins binding to nuclear-retained lncRNAs (Xiao et al. 2012).

Therefore, an alternative theory of nuclear retention may be put forward. Perhaps TREX and Nxf1 are bound to non-coding RNAs to process them or export them to the cytoplasm but there are factors, such as hnRNP U, that prevent export occurring. As shown in Figure 3.6.2 and published data, hnRNP U is necessary to localise *XIST* RNA to the X chromosome, maintain lncRNA *FIRRE* in the nucleus and keep non-coding repeated RNA sequences, *Cot-1* RNA, associated with the chromatin. (Hasegawa et al. 2010; Hacısuleyman et al. 2014; Hall et al. 2014). The diffuse *XIST* RNA signal seen when hnRNP U is depleted may include some *XIST* RNA reaching the cytoplasm much like *FIRRE*. Perhaps these RNAs reach the cytoplasm as the mRNA export pathway can export them in the absence of hnRNP U. This effect could be universal for

lncRNAs and other RNAs associated with the chromatin. To test this hypothesis, hnRNP U can be depleted from human cells and ncRNA localisation assessed. If RNAs are found to “leak” to the cytoplasm, the role of the mRNA export pathway in this process could then be investigated. Primarily, a robust mechanism to deplete hnRNP U had to be established which is discussed in this chapter.

4.1 Attempts to deplete hnRNP U in human cells using RNAi

The first attempt at depleting hnRNP U was by using a SMARTpool of siRNA (Dharmacon ON-TARGETplus 3192). This consists of a pool of four hnRNP U-targeting siRNAs to maximise depletion of the mRNA. However, only a ~53% knockdown at the protein level was achieved, as determined by volume analysis, despite the 80% knockdown at the RNA level (Figure 4.1). This level of knockdown was only possible after 96 hours of siRNA treatment, which may lead to indirect effects of hnRNP U depletion on the cell such as affecting cellular proliferation and metabolism. Ideally, assays to monitor metabolism and proliferation could be done alongside siRNA treatment to indicate whether any effects seen are to do with the loss of hnRNP U or general cellular changes. Despite this consideration, the published effect of *XIST* RNA mislocalisation throughout the nucleoplasm was still observed following siRNA treatment, so this level of protein depletion may be enough to assess hnRNP U function (Figure 4.1.C) (Hasegawa et al. 2010). There also seems to be a general decrease in *XIST* RNA abundance, indicating that this level of hnRNP U depletion is having a profound effect on *XIST* RNA.

However, there were still reservations about using a partial knockdown so we attempted to make an inducible stable cell line that expressed a small-hairpin RNA (shRNA) to target hnRNP U mRNA using the Flp-In system. Despite the desired effects seen with transient expression of the hairpin, no colonies survived the selection process on two occasions. This may be because the Flp-In system can be “leaky”; the hairpin is expressed at low levels despite the lack of tetracycline induction. Therefore, the cells may not be able to survive with even a small loss in hnRNP U over a long period of time.

Further investigation shows that hnRNP U is likely to be essential for humans. hnRNP U is necessary for cell growth in five human cell lines and is deemed essential in multiple haploid cancer cell lines (Hart et al. 2015; Blomen et al. 2015; T. Wang et al. 2015). Additionally, hnRNP U is essential for embryonic development and survival in mice (Roshon & Ruley 2005; Ye et al. 2015). This could explain the lack of colonies whilst making a stable RNAi cell line due to the “leaky” hairpin expression.

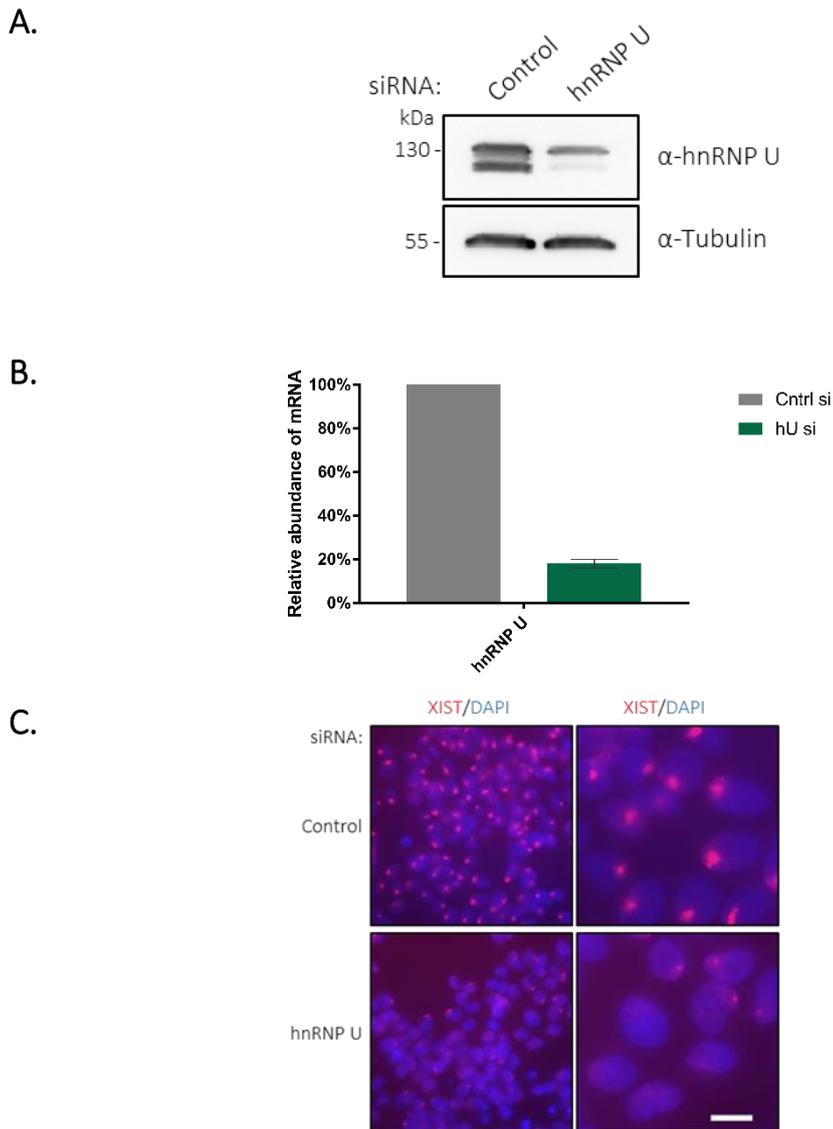


Figure 4.1 hnRNP U siRNA treatment resulted in a partial knockdown at the protein level but still demonstrated a hnRNP U-depletion phenotype in 293T cells

293T cells were transfected with 25 nM hnRNP U siRNA at 0 and 48 hours and were harvested at 96 hours. **(A)** Western blot after siRNA treatment. Tubulin is shown as a loading control. **(B)** qRT-PCR data quantifying hnRNP U mRNA levels following siRNA treatment. Three replicates were carried out and the average abundance with SEM plotted. The abundance of hnRNP U mRNA is shown relative to control siRNA treatment and all abundances are normalised to the abundance of U6 in the sample. **(C)** FISH using a Cy3-labelled *XIST* RNA probe following siRNA treatment as above. Whole mount Images were taken at 40x and 63x magnification using a Leica light microscope with DAPI and Cy3 filters. The white scale bar represents 10 μ m.

Furthermore, mouse hnRNP U has been estimated to have a half-life of 149 hours which is much longer than the majority of the proteins we have depleted using the Flp-In system. For example, in the same study, Alyref and Nxf1 were shown to have half-lives of 54 and 52 hours respectively (Schwanhüusser et al. 2011).

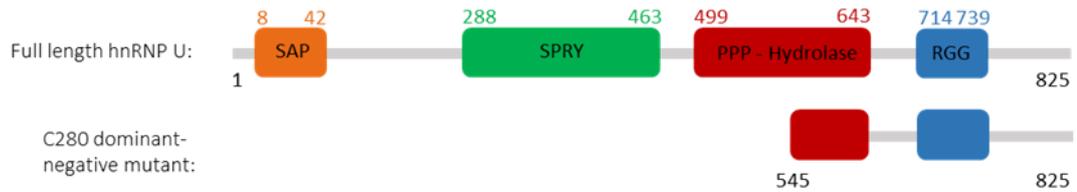
In addition, hnRNP U seems to be hugely abundant, ranked the 111th most abundant protein in HeLa cells (Schwanhüusser et al. 2011; Boisvert et al. 2012). Other labs have also had difficulty depleting the protein and hnRNP U proved the hardest to deplete out of the hnRNPs studied (Valente & Goff 2006; Huelga et al. 2012). The combination of the abundance, half-life and essentiality of hnRNP U makes it difficult to deplete in human-cells using RNAi. Therefore, an alternative approach was needed to prevent hnRNP U functioning in the cell so the resulting effects can be investigated.

4.2 Perturbing the function of hnRNP U using a dominant-negative mutant

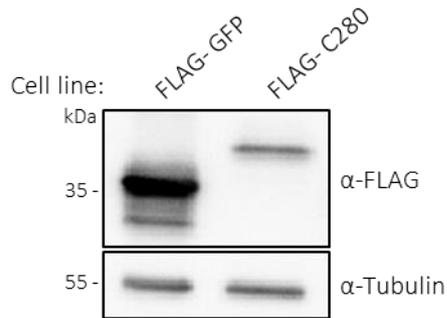
Following RNAi attempts, an alternative technique was employed to try and disturb hnRNP U function. Overexpression of a truncated form of hnRNP U, just the C-terminal 280 residues (Figure 4.2.1.A), has been shown to disrupt the ability of C_oT-1 RNA to associate with chromatin (Hall et al. 2014). This is a comparable effect to that seen with *XIST* RNA (Hasegawa et al. 2010). The truncated hnRNP U, C280, acts as a dominant-negative mutant presumably by binding to RNA through its RGG box but being unable to perform any other functions such as DNA and protein binding through its SAP and SPRY domains respectively. The overexpression of C280 would therefore titrate away the functional full-length protein. As C280 expression releases RNA from the chromatin, we decided to use this approach to disrupt hnRNP U function and investigate ncRNA localisation.

A stable cell line was generated using the Flp-In system that expresses FLAG-hnRNP U-C280 upon induction with tetracycline (Figure 4.2.1.B). A FLAG-GFP expressing cell line is presented as a comparison and shows that FLAG-C280 is only expressed at a low level when compared to FLAG-GFP. Consequently, FLAG-C280 was transiently expressed which gives a higher expression level in comparison to the cell line (Figure 4.2.1.C). FISH of *XIST* RNA was then carried out to confirm that overexpression of the C280 form of hnRNP U causes release of *XIST* from the chromatin (Figure 4.2.2). This would then liken it to the effect seen with hnRNP U RNAi and could be a more complete way of studying hnRNP U function. YFP-Tubulin (YFP-Tub) was transfected to control for any effects of the DNA transfection procedure and to indicate the cells that have been successfully transfected. As Tubulin is a cytoplasmic protein, the YFP-Tub signal is mainly present in the cytoplasm. A much lower mass of YFP-Tub was transfected alongside FLAG-C280 so cells that have taken up YFP-Tub are likely to have also taken up FLAG-C280. As YFP-Tub can be visualised in the green channel, you can be relatively confident about which cells have received FLAG-C280. After transfecting 293T cells with FLAG-C280, the majority of *XIST* RNA was still localised to focal points in the nucleus as seen in the YFP-Tub transfected control.

A.



B.



C.

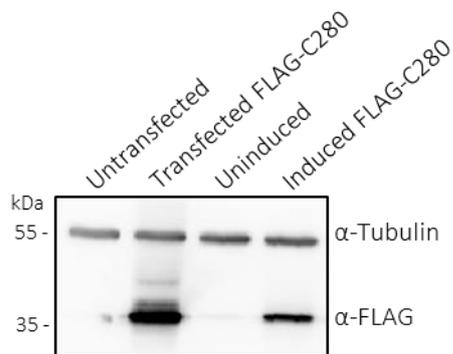


Figure 4.2.1 The C280 dominant-negative mutant of hnRNP U can be expressed in a cell line or transiently through DNA transfection

(A) A schematic diagram of the hnRNP U C280 construct described in Hall et al, 2014, compared to the full-length protein. The amino acid numbers are shown with the annotated domains. (B) Western blot showing FLAG-C280 expression after 48 hours in an inducible 293T stable cell line compared to FLAG-GFP in the same system. Tubulin is shown as a loading control. (C) Western blot of protein from cells transfected with FLAG-C280 for 48 hours compared to the inducible stable cell line in B. Tubulin is shown as a loading control.

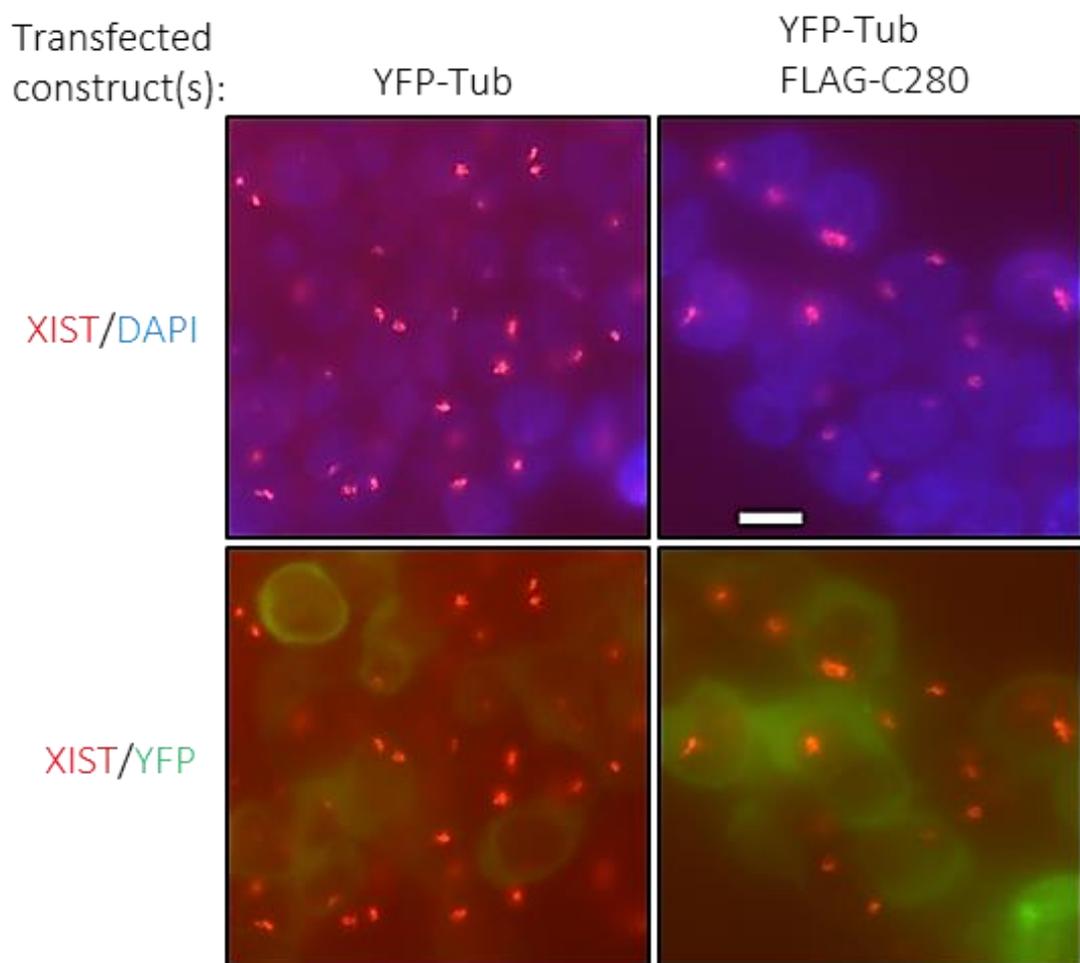


Figure 4.2.2 Overexpressing the C280 dominant-negative mutant of hnRNPU does not result in *XIST* RNA mislocalisation

293T cells were transfected with 700ng YFP-Tub or 600 ng FLAG-C280 and 100 ng YFP-Tub per well of a 24-well plate. The cells were grown for 48 hours before fixing. FISH was carried out using a Cy3-labelled *XIST* RNA probe and cells were DAPI stained. Whole mount images were taken using a Leica light microscope at 63x magnification using the DAPI, Cy3 and GFP filter cubes. The images from different filter channels were overlaid. The white scale bar represents 10 μ m.

XIST RNA has not been mislocalised following C280 overexpression so using the dominant-negative method is not analogous to depleting the hnRNP U protein. It is unlikely that any stronger effects would be seen with this system over using siRNA. A third approach was then employed to study the function of hnRNP U in human cells.

4.3 Generating an Auxin-Inducible Degron (AID) tagged

hnRNP U cell line

A natural mechanism to degrade proteins in plants has been recently developed into a tool that can be used in other organisms, the auxin-inducible degron (AID) system (Nishimura et al. 2009) (Figure 4.3). Once established, exposure to the auxin hormone IAA causes the rapid degradation of a protein of interest, creating an inducible system that can be utilised, in theory, for any protein.

In plants, the proteins that are degraded in this way are plant AUX/IAA transcriptional repressors (IAAs). All IAAs in plants contain a conserved sequence, degron II, which is recognised by SCF-TIR1-auxin that then recruits ubiquitination enzymes to mark the protein for degradation (Ramos et al. 2001; Dharmasiri et al. 2003; Kepinski & Leyser 2004). Therefore, a degron II/AID domain can be expressed together with the protein of interest to cause its degradation upon addition of auxin (Nishimura et al. 2009) (Figure 4.3). There are human orthologues of the proteins in the SCF and E2 complexes as they are highly conserved, however, there is no TIR1 (TRANSPORT INHIBITOR RESPONSE 1) orthologue to form the SCF-TIR1 complex. Consequently, TIR1 needs to be expressed in the cells which has been shown to work in yeast, chickens and human cells (Nishimura et al. 2009). However, the mechanism to stably introduce the degron II domain to the DNA sequence relies on homologous recombination which is not particularly effective in human cells. Therefore, a CRISPR-Cas9 approach was developed to add a degron II/AID DNA sequence to the end of the gene of interest so a fusion protein is expressed (Cong et al. 2013; Mali et al. 2013; Natsume et al. 2016). This method still relies on homologous recombination, but a nick in the DNA is already made by Cas9 where the recombination with the AID sequence needs to take place. The Cas9 enzyme is directed to the DNA sequence to be cut by guide RNAs (gRNA) that have a complementary sequence to the target DNA (Ran et al. 2013).

For use in the human HCT116 cell line, selected as the majority of its genome is diploid, the degron II domain from IAA17 was used as it has a short half-life of approximately 10 minutes (Dreher et al. 2006). Furthermore, TIR1 from *Oryza sativa*

(rice) had to be used instead of *Arabidopsis thaliana* TIR1 due to the higher temperature at which human cells grow (Nishimura et al. 2009). The osTIR1 sequence used is also codon optimised for efficient expression of the protein in human cells.

Using this system, we set out to make a hnRNP U auxin-inducible degron (AID) cell line that upon addition of auxin/IAA, the majority of the hnRNP U protein in the cell will be rapidly degraded by the proteasome (Figure 4.3).

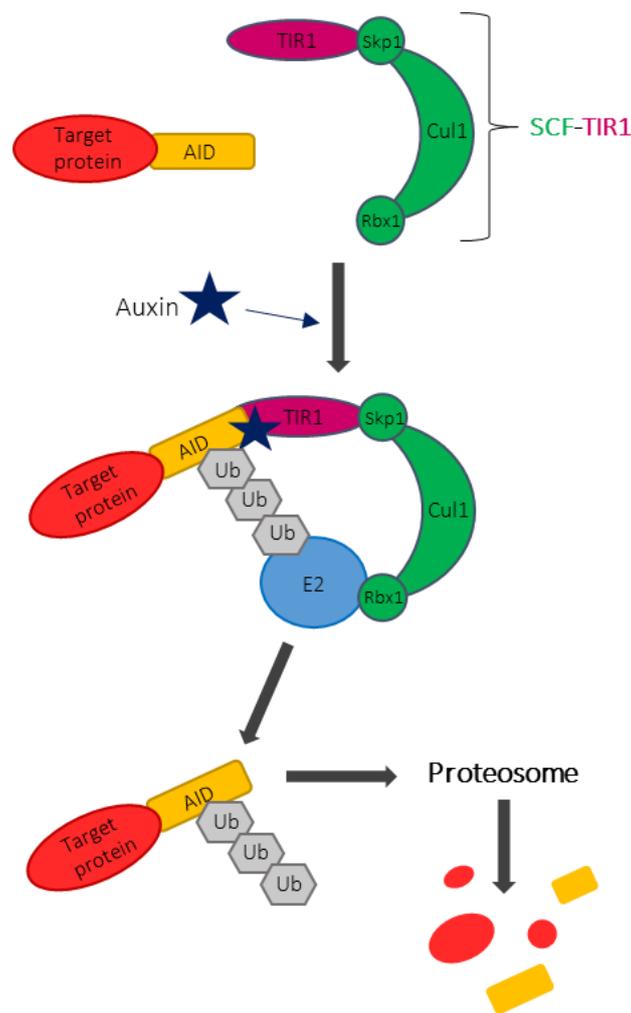


Figure 4.3. A schematic diagram of the auxin-inducible degron (AID) system

This figure is adapted from Nishimura et al. 2009 and Natsume et al. 2016.

The SCF ubiquitin-protein ligase complex containing the TRANSPORT INHIBITOR RESPONSE 1 (TIR1) protein binds to auxin plant hormones such as Indole-3-acetic acid (IAA), represented as Auxin in the figure. Once TIR1 is bound to auxin, the SCF-TIR1 complex recruits the E2 complex that in turn polyubiquitinates the AID part of the protein, represented as Ub. These proteins are then rapidly degraded by the proteasome.

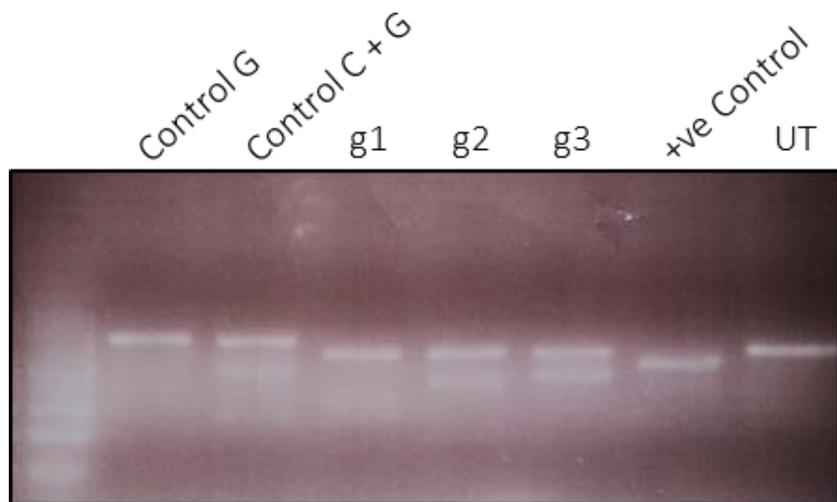
4.3.1 Adding the mini-AID tag to *hnRNP U*

Guide RNAs (gRNA) to the end of the open reading frame (ORF) of the HNRNPU gene were designed using the Benchling tool (Benchling 2018) and cloned into a plasmid containing the Cas9 enzyme sequence (Ran et al. 2013). The efficiency of the guide RNAs to direct Cas9 to the desired location and cause a DNA nick was determined using the SURVEYOR nuclease assay as described in Ran et al. 2013 (IDT, Surveyor mutation detection kit). If a nick had occurred, an Indel (insertion or deletion of bases) is likely to form which will be recognised and digested by SURVEYOR nuclease S following an annealing reaction. Therefore, two different sized fragments would be seen on the agarose gel if the Cas9 had caused a nick in the DNA as some of the DNA would be digested (Figure 4.3.1.1.A). All three guides caused DNA to be nicked and indels to form but we continued with guide 3 (g3) as it appeared to have the greatest amount of DNA in the lower band suggesting more DNA had been nicked.

Additionally, guide 3 targets Cas9 to the most C-terminal DNA sequence of HNRNPU out of all the guides which is more suited to the addition of a C-terminal AID tag.

Following a successful SURVEYOR assay, a plasmid containing the DNA sequence of the end of the HNRNPU ORF that leads sequentially into the AID tag had to be generated (Figure 4.3.1.1.B). This is the homologous repair template (HR) for the cell once the Cas9 has nicked the DNA. The AID tag will then succeed the HNRNPU sequence in frame and be expressed as a fusion protein. At this stage, a mini-AID degron tag was used, which consists of 3 copies of a small fragment from the IAA17 degron sequence (Nishimura et al. 2009; Nishimura & Kanemaki 2014; Natsume et al. 2016). The mini-AID sequence is preceded by three FLAG tags and followed by a self-cleaving P2A cleavage site (Y. Wang et al. 2015; Liu et al. 2017). After the P2A site, a protein that confers resistance to an antibiotic is encoded which allows you to select for successful tagging of the HNRNPU gene by the addition of the antibiotic to the cell culture medium. The P2A site allows the hnRNP U-mini-AID fusion protein and the antibiotic resistance-conferring protein to function separately after they have been expressed as a single polypeptide. HR templates with two selectable markers had to be generated as both copies of the HNRNPU gene would need to be tagged.

A.



B.

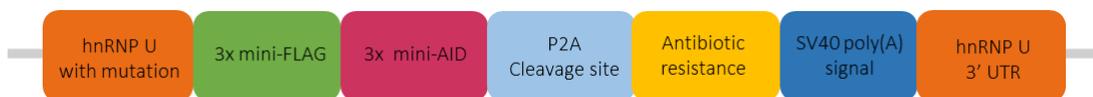


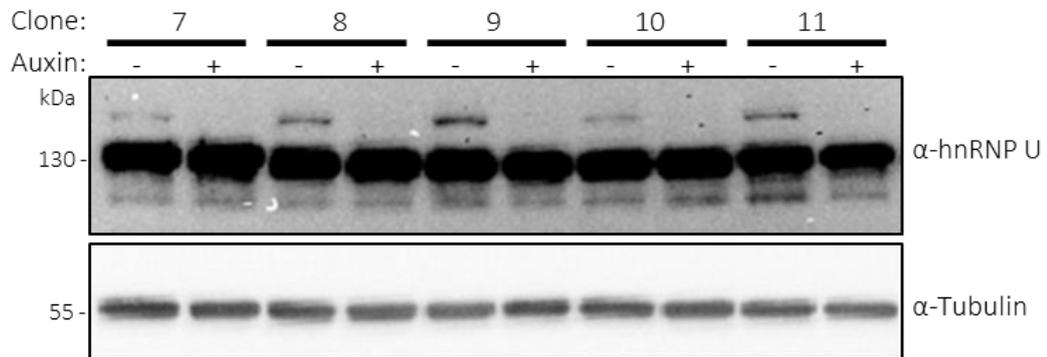
Figure 4.3.1.1 Cloned gRNAs 1-3 can be used to target Cas9 to the end of the HNRNPU gene

(A) DNA gel electrophoresis of products from the SURVEYOR assay, Ran et al. 2013, using the IDT Surveyor mutation detection kit. g1,2 and 3 represent the three guide RNAs tested, +ve control is a known functioning guide, UT represents untransfected cells and controls G and control C and G are provided in the IDT kit. (B) A schematic of the mini-AID HR template that will be used to repair the DNA following successful nicking of the HNRNPU gene.

The sequence targeted by the guide RNAs and the Cas9 enzyme had to be mutated in the repair template so it was not also nicked by Cas9. The Benchling programme (Benchling 2018) was used to decide on the most efficient mutations to generate in HNRNPU so that the guide RNA would no longer recognise the sequence but the amino acid encoded would remain the same.

HCT116-TIR1 cells were then transfected with the Cas9/guide RNA plasmid and two HR templates conferring different antibiotic resistances, Hygromycin and Neomycin/G418. Successful tagging of the protein was selected for by adding Hygromycin and G418 to the media. 13 clones were selected, expanded and tested for auxin responsiveness (Figure 4.3.1.2.A). There is a weak band at the predicted size of tagged hnRNP U that is much reduced when auxin is added to the cells in all clones. However, there is a much stronger band at the endogenous size of hnRNP U, 130 kDa, suggesting that there is still some untagged hnRNP U despite the selection using both antibiotics. This suggests that there may be a third copy of the hnRNP U gene in HCT116-TIR1 cells. Additionally, a few clones were also probed with an anti-FLAG antibody to check that the higher band on the blot was in fact AID-tagged hnRNP U (Figure 4.3.1.2.B). The larger band is seen in all clones when the blot is probed with anti-FLAG and the band is depleted when auxin is added, demonstrating the presence of AID-tagged hnRNP U in these cells. However, there is also a more dominant band present on the FLAG blot at the size of the endogenous hnRNP U and a band seen in between the two strongest bands. This suggests that the mini-AID tag may be falling off the protein as there are three sizes of bands detected by the anti-FLAG antibody as the three FLAG tags fall off. The unstable tag is probably contributing to the large band seen on the anti-hnRNP U blot at the endogenous size, but it is unlikely that it completely accounts for the much bigger signal seen at the size of wild-type hnRNP U. Both eventualities are likely to be occurring; the mini-AID tag is unstable and falling off the hnRNP U protein and there are at least three copies of the HNRNPU gene in HCT116-TIR1 cells.

A.



B.

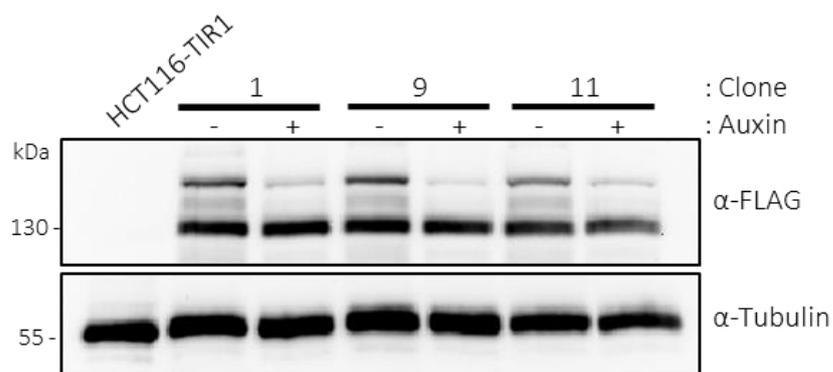


Figure 4.3.1.2 Tagging HNRNPU with the mini-AID tag reveals there are likely to be three copies of HNRNPU in HCT116-TIR1 cells and the mini-AID tag is unstable

(A) Western blot of protein samples from selected clones, treated with auxin, following transfection of the Cas9/gRNA3 plasmid and two mini-AID HR templates into HCT116-TIR1 cells. Tubulin is shown as a loading control. (B) Western blot of protein samples from A, probed with anti-FLAG antibody. FLAG is present at the size of untagged hnRNP U, 130 kDa, as well as the larger tagged protein. Tubulin is shown as a loading control.

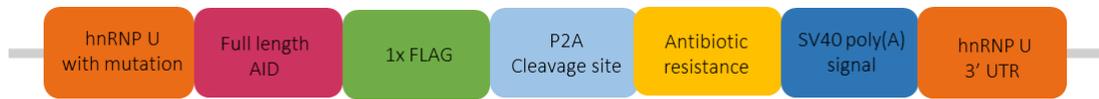
Consequently, the next steps to make the cell line were to try a different AID tag and to generate a third, and potentially a fourth, HR template with other antibiotic resistance genes. Additionally, the cell lines would be made in the HCT116 cells without the presence of osTIR1. The constant expression of osTIR1 could potentially interfere with hnRNP U and target it for degradation before the addition of auxin. Therefore, there may be a selective pressure against tagging all copies of HNRNPU if the protein is degraded even without the presence of auxin. osTIR1 could then be re-integrated following the tagging of all HNRNPU genes and a lower level of osTIR1 expression may be preferable for the cell line to function. Alternatively, there is a doxycycline-inducible osTIR1 that could be used (Natsume et al. 2016).

4.3.2 Tagging hnRNP U with the full-length AID tag

Due to the instability of the mini-AID tag, an alternative tag was devised and named full-length AID (flAID) (Figure 4.3.2.1.A). This tag contains one copy of the complete IAA17 degron sequence but it contains fewer glycine linkers and a C-terminal FLAG tag. The flAID tag was cloned into the mini-AID vectors with alternative resistance genes. This resulted in four HR templates containing the full-length AID tag and resistance genes for Hygromycin, Neomycin/G418, Puromycin and Zeocin. Only two homology repair vectors and the Cas9/gRNA plasmid were transfected into HCT116 at one time as the addition of more than two selection antibiotics at once killed all the cells before colonies could form.

Despite the high possibility that there are three copies of the HNRNPU gene in HCT116 cells, transfecting only two HR vectors produced clones that expressed only tagged hnRNP U (Figure 4.3.2.1.B). Clones 1, 2 and 4 have all copies of HNRNPU tagged whereas clones 3, 5 and 6 do not. The clones with complete tagging of HNRNPU probably have two copies of the gene tagged with a full-length AID tag encoding the same antibiotic resistance gene. At this stage, we did not foresee a problem and continued to transfect the osTIR1-containing plasmid into Clone 1 and selected for stable integration using the Blasticidin resistance marker. All resulting clones expressed osTIR1, demonstrated by the presence of the myc tag, and all responded to the addition of auxin as hnRNP U disappeared (Figure 4.3.2.2.A). However, the cells had reverted to having untagged hnRNP U, the band at 130 kDa, over the course of expanding Clone 1 to transfect with osTIR1 and the selection process for osTIR1 expression. A protein sample of Clone 1 cells cultured for a few weeks has a more prevalent band of untagged hnRNP U than newly defrosted Clone 1 cells as they have lost the flAID tag over time (Figure 4.3.2.2.B). This is likely to have occurred due to a selective pressure to lose the flAID tag from the third copy of the HNRNPU gene. As two copies of the gene are tagged with flAID that expresses the same resistance marker, there is no selection pressure to keep both copies tagged to survive. This also suggests that there is an advantage to not having the tag on the HNRNPU gene, perhaps due to it destabilising the protein that is likely to be essential.

A.



B.

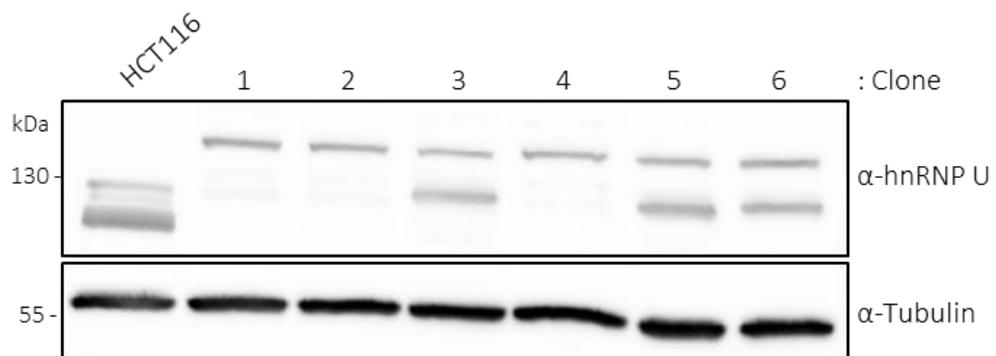
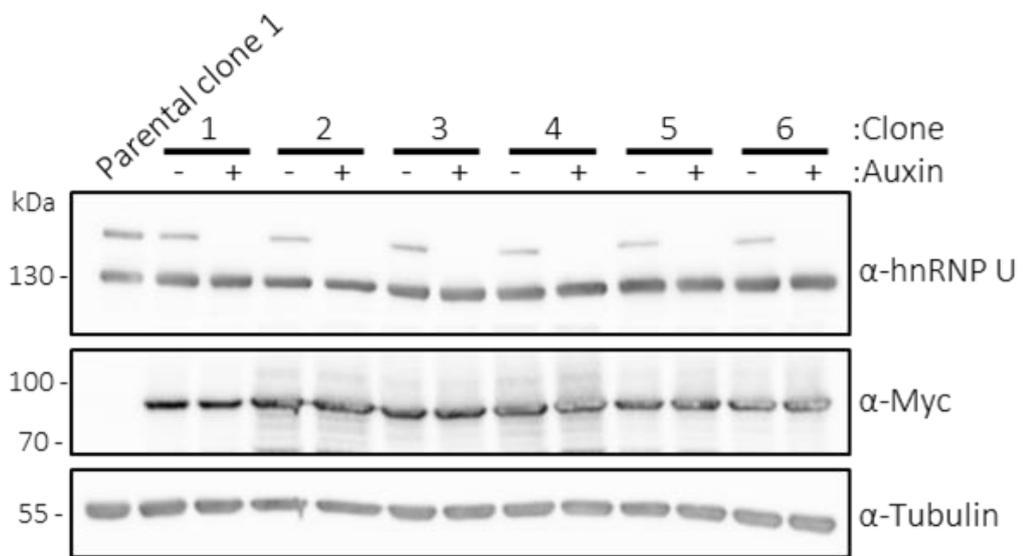


Figure 4.3.2.1 Using the full-length AID tag allows all copies of HNRNPU to be tagged in some clones in the absence of osTIR1

(A) A schematic of the full-length-AID (fAID) tag used as the HR template following Cas9 nicking. (B) Western blot of protein samples from selected clones following transfection of the gRNA3/Cas9 plasmid and two fAID HR templates into HCT116 cells. The HR templates contained Hygromycin and Neomycin/G418 resistance genes. Tubulin is shown as a loading control.

A.



B.

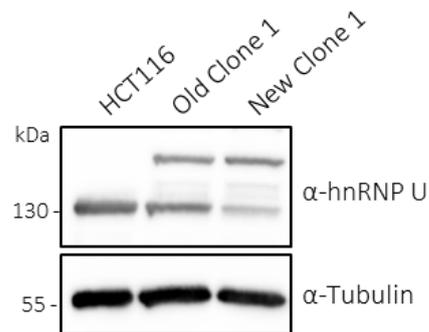


Figure 4.3.2.2. The flAID-tagged hnRNP U protein can be reverted to untagged hnRNP U over time

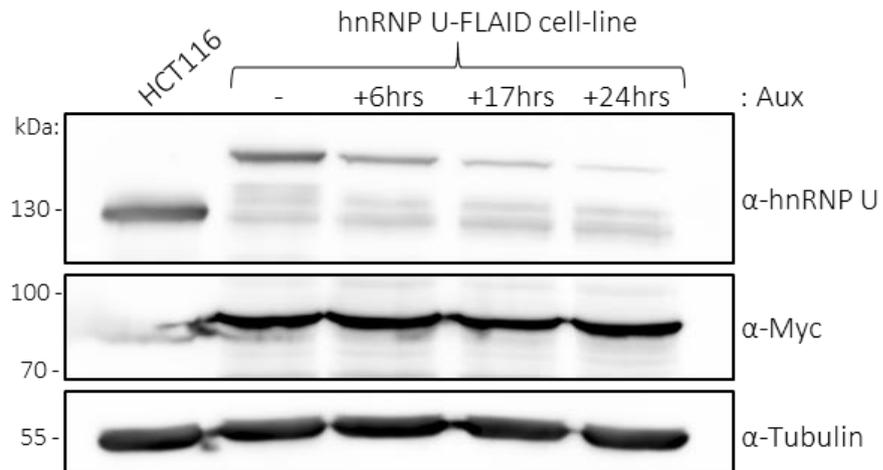
(A) Western blot of protein extracts from selected clones, treated with auxin, following the transfection of osTIR1 into Clone 1 from Figure 4.3.2.1. The presence of myc demonstrates the expression of osTIR1 and Tubulin is a loading control. (B) A Western blot of protein samples from Clone 1 in Figure 4.3.2.1 that have been cultured for a few weeks, old clone 1, and from newly defrosted cells, new clone 1. Tubulin is shown as a loading control.

To get to the final and functional cell line, a full-length AID tag with Puromycin resistance was transfected into the polyclonal population containing flAID tags with Hygromycin and Neomycin/G418 resistance and then the osTIR1 transfected subsequently (Figure 4.3.2.3.A).

A time course of auxin addition shows that flAID-tagged hnRNP U protein takes 24 hours of exposure to auxin to be almost completely degraded, much longer than suggested for most proteins (Natsume et al. 2016). As previously discussed in Chapter 4.1, hnRNP U is exceptionally abundant in the cell which is likely to be why it takes longer to be degraded by the proteasome. A hnRNP U-like 1-flAID cell line was generated by Llywelyn Griffith and a time course of auxin addition is also shown (Figure 4.3.2.3.B). The hnRNP U-L1-flAID cell line contains the doxycycline-inducible osTIR1 so doxycycline was added to the cells 24 hours before the addition of auxin. It takes approximately 2 hours for the hnRNP U-L1-flAID protein to be degraded, in contrast to 24 hours for hnRNP U-flAID. hnRNP U-L1 is much less abundant in mammalian cells than hnRNP U (Schwanhüusser et al. 2011; Boisvert et al. 2012), which may contribute to the differences seen between the two cell lines. However, other factors may influence the degradation time such as the ability of the SCF complex to recognise the AID tag and the proteasome to degrade the marked tag when the degron is expressed as a fusion protein with hnRNP U and hnRNP U-L1.

Overall, we have shown that the auxin-degron tagging method can be used to induce the depletion of the majority of the hnRNP U protein in HCT116 cells in a relatively short space of time in comparison to RNAi systems. The reduction in protein level is far greater than that of any RNAi method used, making this one of the most practical and effective methods of studying hnRNP U function in human cells. However, the functionality of the tagged hnRNP U protein and the state of the cell line must be validated to make sure that any effects seen are due to the depletion of the protein from the cells and not the direct effect of tagging the protein.

A.



B.

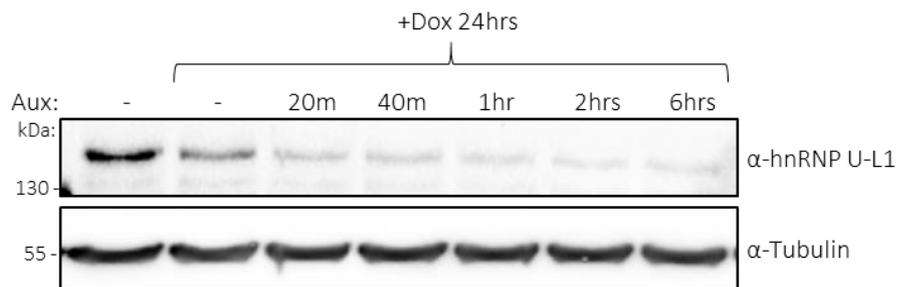


Figure 4.3.2.3 Completed hnRNP U and hnRNP U-like 1 flAID cell lines

The Western blot in **(B)** was kindly provided by Llywelyn Griffith.

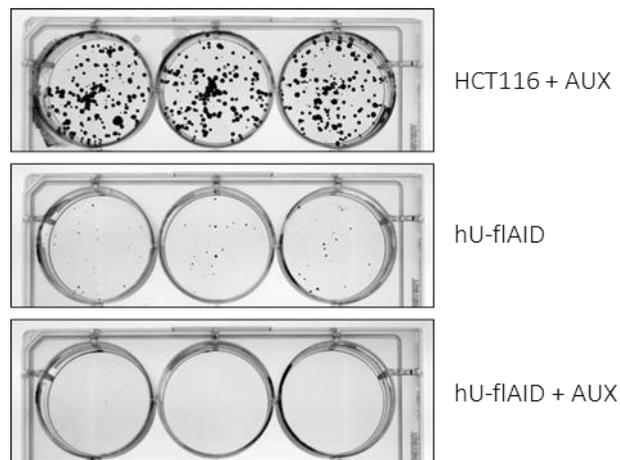
(A) Western blot showing a time course of auxin treatment on the completed hnRNP U cell line and the resulting levels of hnRNP U-flAID. The expression of Myc tag confirms the presence of osTIR1. Tubulin is shown as a loading control. **(B)** Western blot showing a time course of auxin treatment on the completed hnRNP U-L1 cell line. 24 hours of doxycycline treatment was applied beforehand to induce the expression of osTIR1. Tubulin is shown as a loading control.

4.3.3 Validation of the hnRNP U-flAID cell line

Following the generation of the functional hnRNP U degron cell line, there were a few potential issues that needed to be addressed before continuing with the planned experiments. This was to make sure that any results gained were as a result of the loss of the hnRNP U protein and not for any other reason such as growth rate or the flAID-tag disrupting the function of the protein.

A colony formation assay was carried out to compare the survival ability of the hnRNP U-flAID cell line and regular HCT116 cells (Figure 4.3.3.1). The hnRNP U-flAID cell line treated with auxin was also included to give an indication of whether hnRNP U is essential for survival in HCT116 cells. There were no colonies of hnRNP U-flAID in any of the wells after 13 days of auxin treatment, suggesting that HCT116 cells cannot survive without hnRNP U so hnRNP U is likely to be essential in human cells. HCT116 cells treated with auxin formed on average 132.5 colonies which is a reasonable number as 200 cells were plated per well. Their survival and growth seems typical. In comparison, on average only 24.3 colonies of hnRNP U-flAID were countable, suggesting that their survival ability is impaired. This could be due to the reduction in the amount of hnRNP U-flAID protein compared to the amount of hnRNP U in HCT116 cells, as seen on the Western blot (Figure 4.3.2.3.A). The reduction in protein level is also seen in the hnRNP U-L1-flAID cell line (data not shown), but to a much greater extent. hnRNP U siRNA treatment, which gave a ~53% protein depletion (Figure 4.1.A), still demonstrated the published phenotype of releasing *XIST* RNA from the chromatin (Figure 4.1.C). Therefore, even a small decrease in hnRNP U protein levels could have a strong effect on the cells. Depletion of hnRNP U over a long time is likely to result in indirect effects of depleting the protein and not just the direct result of having a minimal amount of hnRNP U in the cell. The hnRNP U-flAID cell line may have acquired a different rate of proliferation and metabolism as it has less hnRNP U than the wild type HCT116 cells. This could be assayed by a more direct quantitative approach such as an MTT assay. Therefore, perhaps the timing of cellular proliferation has been altered in the cell line, leading to reduced numbers of colonies over 13 days.

A.



B.

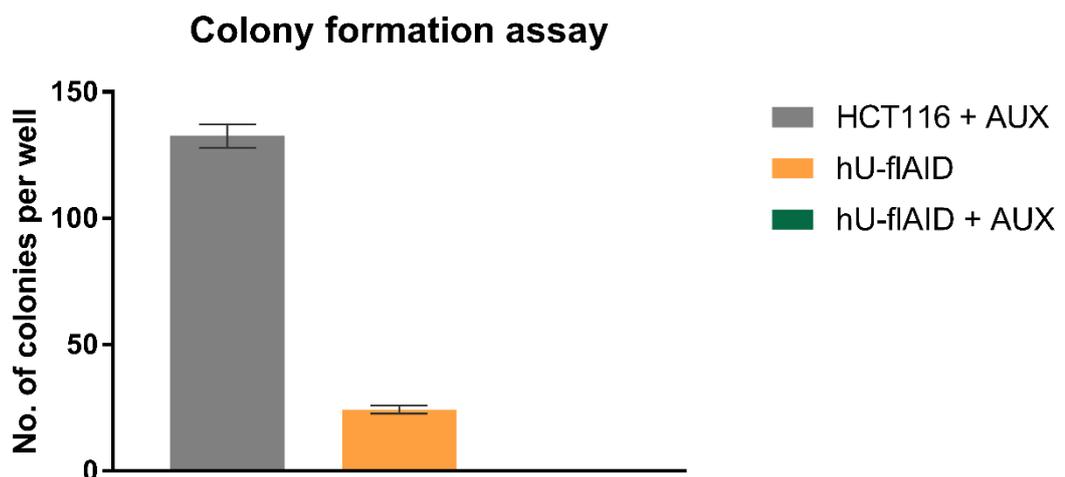


Figure 4.3.3.1 A colony formation assay reveals the decreased survival ability of the hnRNP U-flAID cell line and that hnRNP U is essential to HCT116 cells

Cells from all three conditions were plated at a density of 200 cells per well in a 6-well plate and incubated for 13 days before staining with crystal violet and counting. (A) Three of the six replicate wells are shown per condition. (B) Quantification of the number of colonies from 6 wells in each condition. The average number of colonies per well and the SEM are plotted. Auxin treatment of the hnRNP U-flAID cell line resulted in no colonies being formed.

Following the discovery of the impaired growth of the hU-flAID cell line, the function of hnRNP U when tagged with flAID was assessed to make sure its functionality is not impaired. Unfortunately, HCT116 cells do not express *XIST* RNA so we could not look for *XIST* RNA mislocalisation. Alternatively, a co-immunoprecipitation (co-IP) experiment was carried out to check that hnRNP U-flAID can still form known functional interactions with other proteins.

A FLAG co-IP was carried out as the antibody to endogenous hnRNP U was ineffective. A cell line expressing FLAG-hnRNP U upon tetracycline induction was used as a comparison to hnRNP U-flAID which has a FLAG tag following the AID sequence. A FLAG-GFP expressing cell line was used as a control for co-IP specificity. Both FLAG-hnRNP U and FLAG-GFP have three N-terminal FLAG tags. The interaction between FLAG-hnRNP U or hnRNP U-flAID with proteins known to interact or unlikely to interact was tested (Figure 4.3.3.2). Other hnRNPs, Alyref and TDP-43 were co-immunoprecipitated at similar levels by both FLAG-hnRNP U and hnRNP U-flAID proteins. Functional interactions are therefore likely to be maintained with the hnRNP U-flAID protein, such as binding to TDP-43 (Suzuki et al. 2015). Nxf1 and eIF4A3 are not immunoprecipitated with either FLAG-hnRNP U or hnRNP U-flAID, further demonstrating their similarity in terms of protein:protein interactions. As the interactions between hnRNP U or hnRNP U-flAID and other proteins are unchanged, hnRNP U-flAID is likely to function in a comparable manner to untagged hnRNP U so the cell line can be used to look at the effects of depleting hnRNP U.

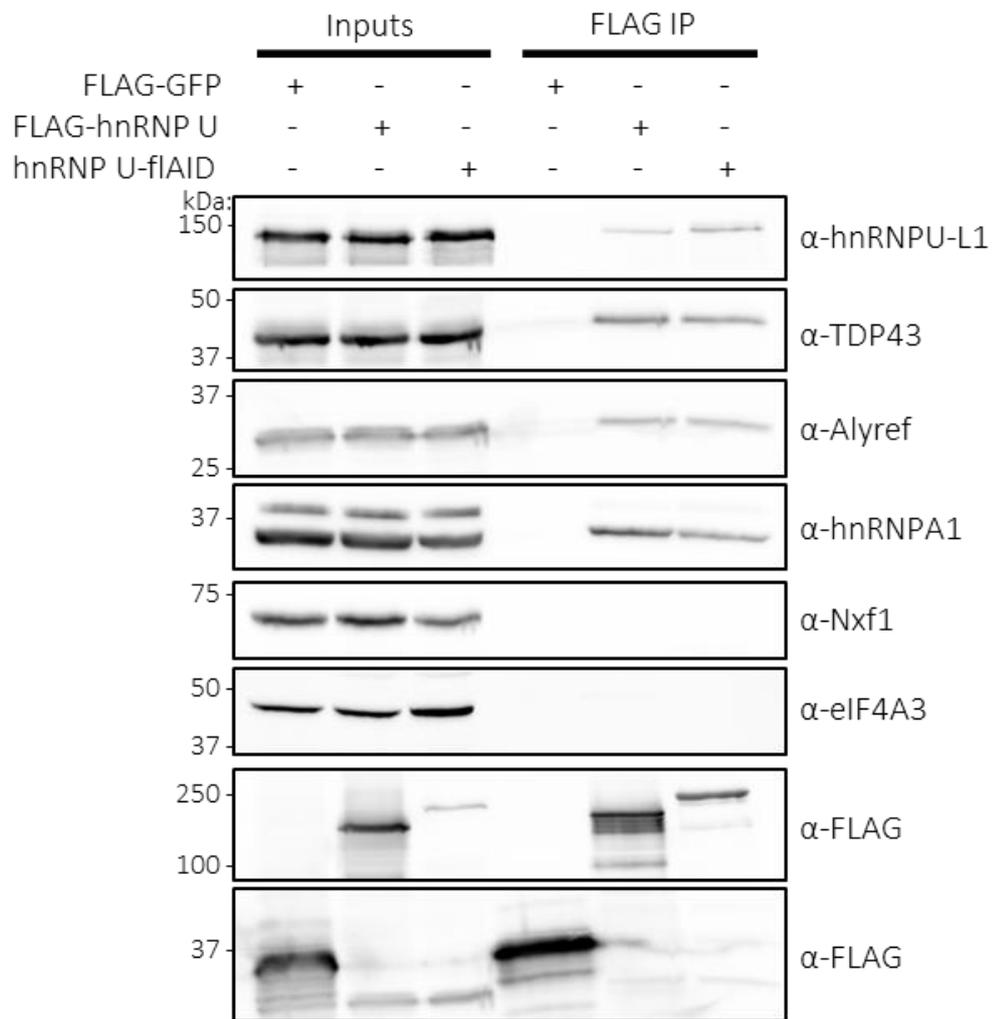


Figure 4.3.3.2 A co-IP of FLAG-hnRNP U and hnRNP U-flAID cell lines reveals no substantial difference in the proteins associated

Protein lysates from cells expressing FLAG-GFP, FLAG-hnRNP U and hnRNP U-flAID were used to carry out a co-IP using beads bound with anti-FLAG antibody. The eluates were run alongside input samples (~0.1%), collected before the lysates were added to the beads. The resulting Western blots were probed with antibodies to proteins known to interact with hnRNP U and some thought to be unlikely to interact. FLAG-GFP was used as a control for co-IP specificity. The expression of all three proteins is shown by the α -FLAG blots.

4.3.4 Comparing the hnRNP U-flAID cell line to existing hnRNP U RNAi systems

To understand how the hnRNP U-flAID cell line functions, we decided to see if it shows the same phenotypes as cells depleted of hnRNP U using RNAi. Published data and our own data from hnRNP U siRNA treatments was used.

From comparing the relative abundances of RNAs following hnRNP U depletion using the flAID cell line and siRNA treatment, there does not seem to be much agreement between the two data sets (Figure 4.3.4.1). The levels of most of the non-coding RNAs tested are significantly different between the two depletion methods and some even show opposing effects such as RNA polymerase II-transcribed snRNAs. snRNAs seem to be upregulated following hnRNP U siRNA treatment but downregulated following hnRNP U-flAID depletion. Messenger RNAs seem to be more consistent with a lack of change in both systems, with EGR1 as the exception, perhaps because hnRNP U does not function as greatly in their lifecycle. However, the siRNA treatment was carried out in 293T cells whereas the hnRNP U-flAID cell line has a HCT116 background, which may account for some variation. Although, the stark difference seen with the snRNAs suggests to me that it is the method of protein depletion that is causing this disparity. Perhaps the length of time that the cell is depleted of hnRNP U plays a role as the siRNA treatment takes 96 hours whereas treating the hnRNP U-flAID cell line with auxin only takes 24 hours.

Data from the ENCODE project (The ENCODE Project Consortium 2012) contains RNA-seq experiments using small hairpin RNAs (shRNA), equivalent to siRNAs, in K562 and HepG2 human cancer cell lines. From this data, analysed by Dr Ian Sudbery, a list of RNAs that are significantly increased or decreased in abundance following hnRNP U shRNA treatment was generated. RNA transcripts that are classed as “UP” were significantly more abundant following hnRNP U shRNA treatment and the effect on those transcripts following hnRNP U-flAID depletion is shown (Figure 4.3.4.2). The selection of transcripts tested had high fold changes over control shRNA in the ENCODE data and were biotypes that were interesting to look at.

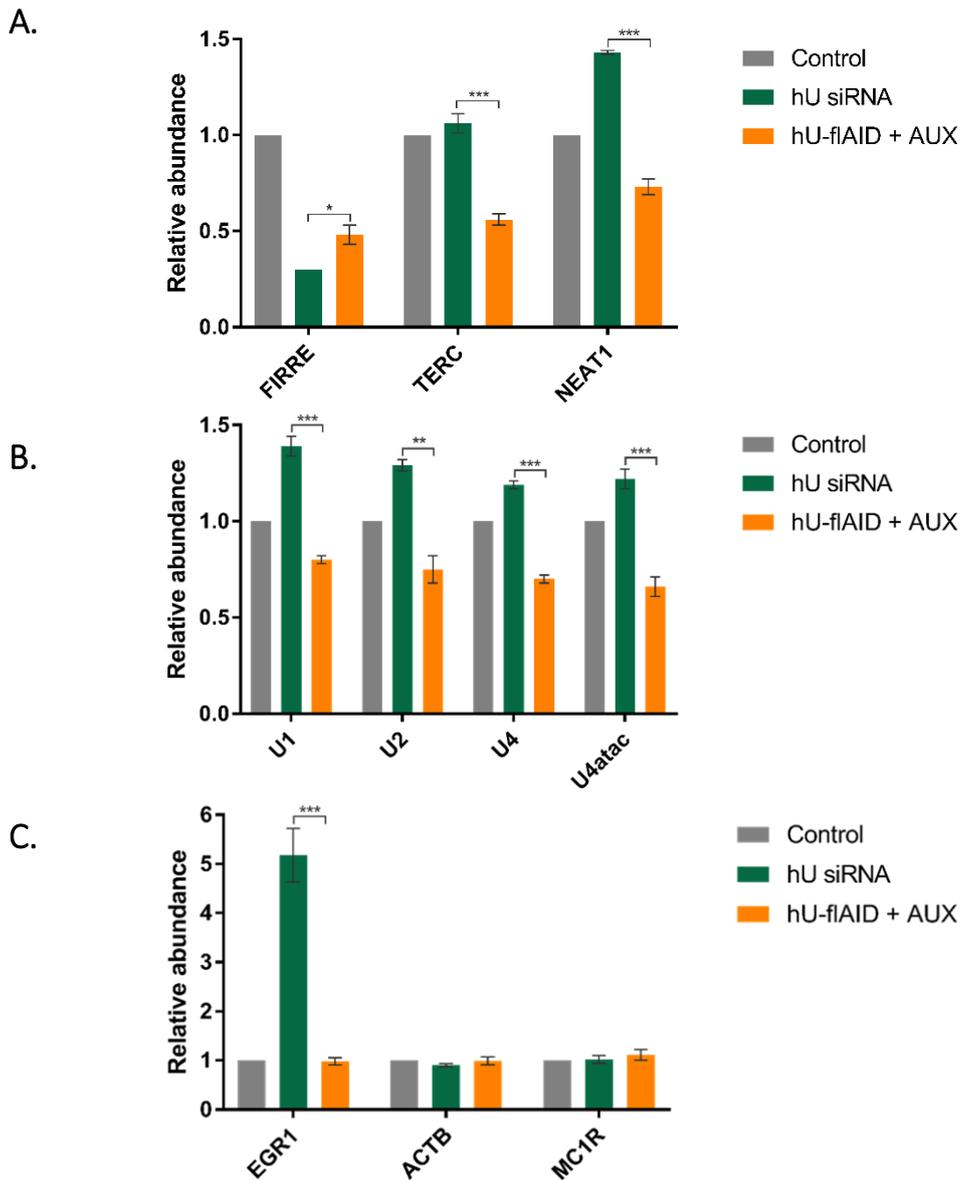


Figure 4.3.4.1 Treating 293T cells with hnRNP U siRNA is not comparable to depletion of hnRNP U using the AID system in HCT116 cells

qRT-PCR data of RNA from 293T cells treated with hnRNP U siRNA for 96 hours and hnRNP U-flAID cells treated with auxin for 24 hours. Control siRNA and HCT116 + auxin RNA samples were used as the respective controls. (A)(B) and (C) show the mean abundances of some lncRNAs, snRNAs and mRNAs respectively. The mean was taken from three or four replicates for siRNA treatment and the flAID cell line respectively. All abundances are shown relative to the mean control abundance. Error bars represent the SEM and all abundances are normalised to the abundance of U6 in the sample.

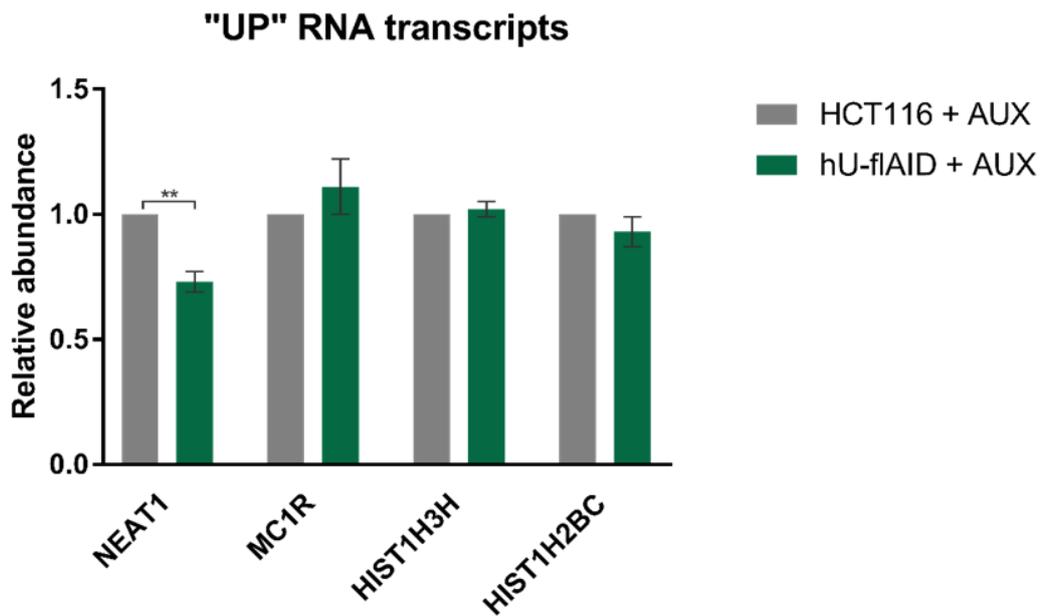


Figure 4.3.4.2 hnRNP U depletion using the AID system is not equivalent to hnRNP U RNAi data from ENCODE

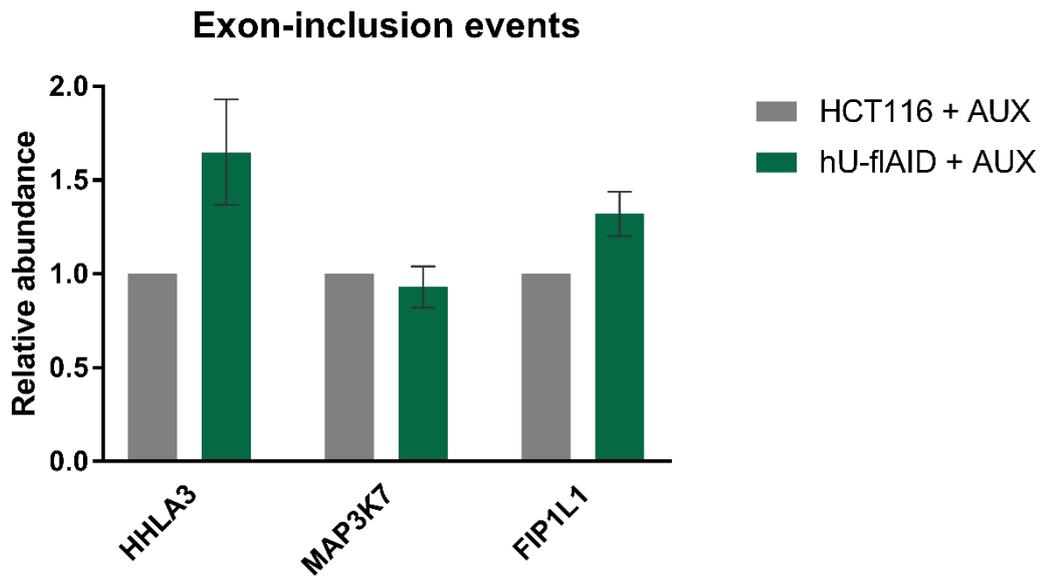
qRT-PCR data of RNA transcripts that are annotated as being significantly upregulated from the ENCODE RNA-seq data following hnRNP U RNAi in two human cell lines. These transcripts are annotated as “UP”. hU-flAID cells were treated with auxin for 24 hours before RNA extraction. The mean abundance of four replicate RNA samples is shown with abundances shown relative to the control (HCT116 + AUX) abundance. Error bars represent the SEM and all abundances are normalised to the abundance of U6 in the sample.

However, none of the transcripts tested were significantly greater in abundance following hnRNP U-flAID depletion and the opposite effect is seen with *NEAT1* RNA. The increase in abundance of *NEAT1* RNA that is seen in the RNA-seq data mirrors the effect seen with the hnRNP U siRNA data in 293T cells (Figure 4.3.4.1), further suggesting that it is the method of hnRNP U depletion that results in differing abundances of RNA transcripts.

We also attempted to replicate some splicing changes seen with hnRNP U siRNA in HeLa cells (Xiao et al. 2012). Specific exon-skipping and exon-inclusion events were seen following hnRNP U RNAi, which were analysed in the hnRNP U-flAID cells following auxin treatment (Figure 4.3.4.3). qRT-PCR primers specific to the skipped or included exons skipped were used so you would expect a decrease in abundance of skipped exons and an increase in abundance of included exons following hnRNP U-flAID depletion. Neither the published exon-skipping nor inclusion events could be significantly replicated in the hnRNP U-flAID cell line. Once again, this study was carried out using a different cell line although it is more likely that the method of hnRNP U depletion is resulting in different effects.

The final attempt at validating the effects seen in the hnRNP U-flAID cell line with published RNAi data came from the same paper as the splicing changes. Xiao et al showed that hnRNP U siRNA treatment resulted in an increase in the number of Cajal bodies per cell, which we attempted to replicate in the hnRNP U-flAID cell line (Xiao et al. 2012). The α -Coilin antibody was used to stain Cajal bodies and the number per cell was counted in control, hnRNP U-flAID and hnRNP U-L1-flAID cell lines (Figure 4.3.4.4). The number of Cajal bodies per cell increases following hnRNP U-flAID depletion compared to the number per cell in HCT116 cells, with the majority having 2 or 3 compared to 1 or 2 respectively. This is in accordance with the published data (Xiao et al. 2012), suggesting that the method of depletion of hnRNP U or the cell line used does not alter the effect seen on nuclear bodies. In contrast, the number of Cajal bodies per cell when hnRNP U-L1-flAID is depleted is dramatically reduced, shifting to the majority having 0 or 1 Cajal bodies per cell. This helps to demonstrate the difference between the two proteins and shows that the addition of the flAID tag to these proteins doesn't produce the same effect.

A.



B.

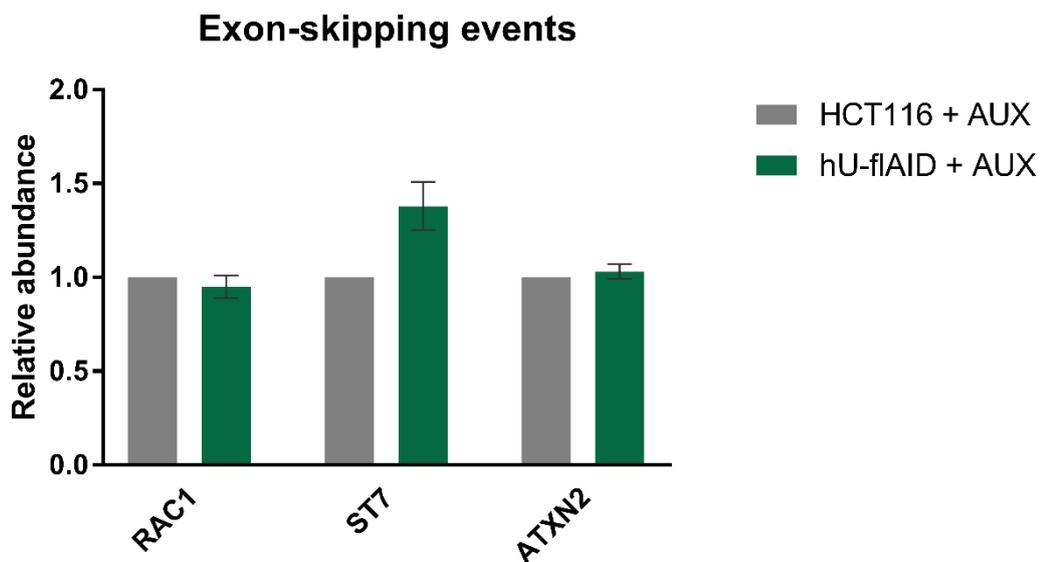


Figure 4.3.4.3 Published splicing changes following hnRNP U RNAi are not reproducible in hnRNP U-flAID cells

qRT-PCR data of RNA transcripts taken from Xiao et al, 2012 that were shown to demonstrate (A) exon inclusion or (B) exon skipping following hnRNP U RNAi. hU-flAID cells were treated with auxin for 24 hours before RNA extraction. The mean transcript abundances of four RNA samples is shown, relative to the control (HCT116 + AUX) abundance. Error bars represent the SEM and all abundances are normalised to the abundance of U6 in the sample.

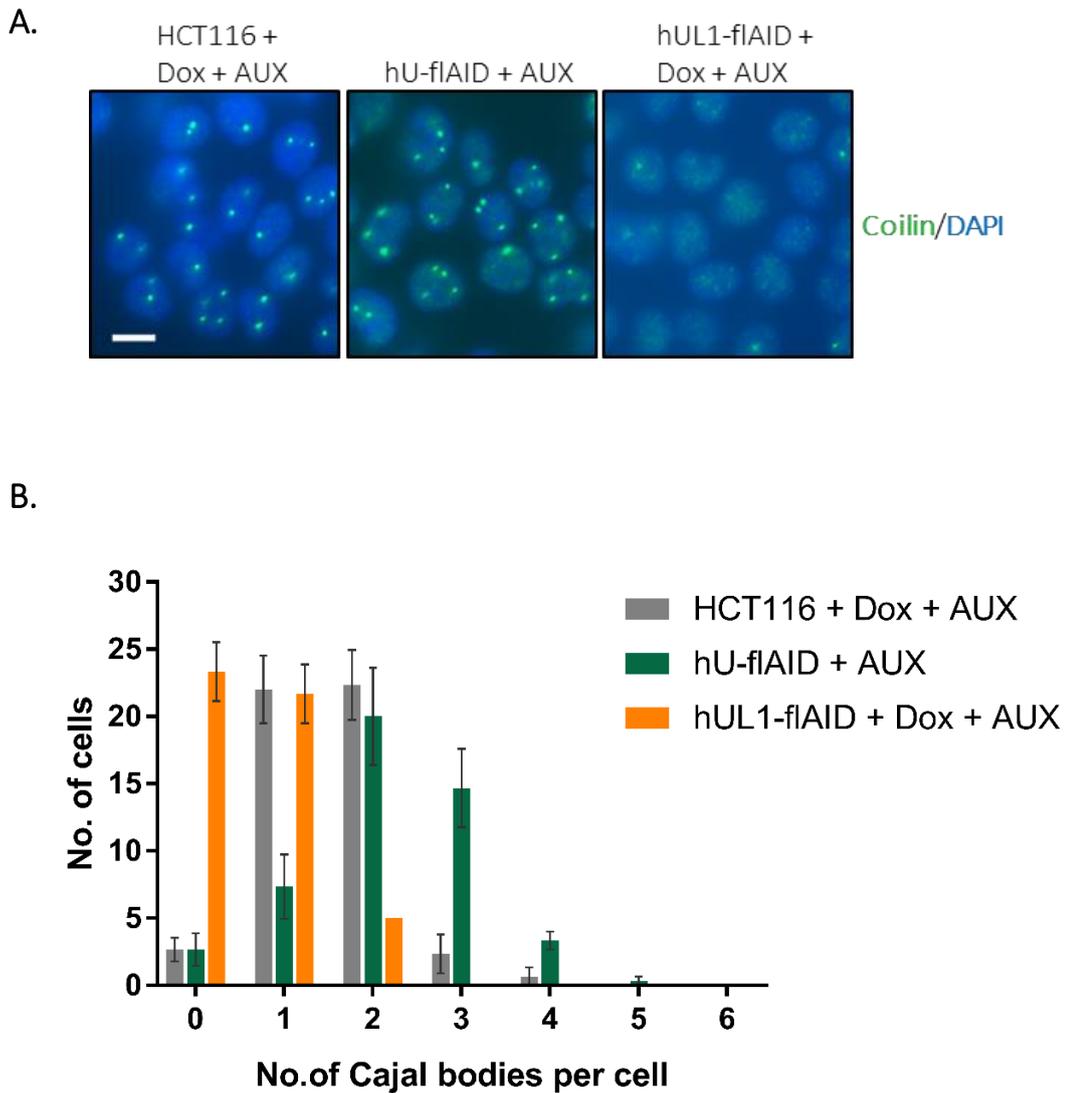


Figure 4.3.4.4 The number of Cajal bodies per cell increases following hnRNP U-flAID depletion, in accordance with hnRNP U RNAi

This experiment was carried out by Llywelyn Griffith and the cell images and data were collected by Professor Stuart Wilson.

(A) hnRNP U-flAID and hnRNP U-L1-flAID cells treated with auxin (AUX) for 24 hours and auxin and doxycycline (Dox) for 2 hours respectively were immunostained using anti-Coilin antibody to detect Cajal bodies alongside control cells (HCT116 + Dox + AUX). The white scale bar represents 10 μ M. The number of Cajal bodies per cell in 50 cells in triplicate were counted. (B) The mean number of cells containing a number of Cajal bodies and the SEM are plotted on the graph.

4.4 Chapter 4 Summary

In this chapter, the making and validating of a human cell line that can be depleted of hnRNP U in 24 hours is detailed. hnRNP U is one of the most abundant proteins in the human cell and is difficult to deplete by conventional methods, especially in a short amount of time. Additionally, a knockout cell line would have probably not survived as it is likely to be essential. This led us to the auxin-inducible degron system that allows degradation of hnRNP U in HCT116 cells when required. The degron-tagging of the hnRNP U protein reduced the cell survivability and the abundance of hnRNP U was reduced even without the addition of auxin. This could be due to the constitutive expression of osTIR1 in the cell or the presence of the tag making the protein slightly less stable. In the hnRNP U-like1-flAID cell line, the expression of osTIR1 by adding doxycycline results in less hnRNP U-L1-flAID even before auxin is added (Figure 4.3.4.3). This suggests that osTIR1 presence has an effect, which could explain the reduction in hnRNP U in the hnRNP U-flAID cell line. However, using the doxycycline-inducible osTIR1 in hnRNP U-flAID cells did not eliminate the destabilising effect (data not shown), suggesting it is the presence of the tag that causes the destabilisation. Nonetheless, the flAID tag did not impair known protein:protein interactions or confer new associations that wild type hnRNP U does not have. Therefore, this is an effective system for studying hnRNP U function in human cells.

It also seems that this method of hnRNP U protein depletion has different effects to that seen using RNAi, in the context of our experiments in 293T cells and published data. This could be down to the variety of human cell lines used to study hnRNP U or more likely, it is to do with the contrasting methods. Although effects on the levels of RNA transcripts in the cells gave differing results between RNAi and degron tagging, the effect on Cajal bodies remained the same. Perhaps therefore, steady state levels of RNAs change a significant amount over the course of siRNA treatment but effects on nuclear bodies remain constant. Although the process has not been straightforward, we have shown that the AID system can be used to rapidly deplete human cells of hnRNP U and will be extremely useful to study its function in the cell.

Chapter 5 - hnRNP U is a general chromatin-retention factor for non-coding RNA

Following the successful creation of the hnRNP U-flAID cell line, the function of hnRNP U in the cell can be investigated. There are likely to be various undiscovered functions of hnRNP U in the cell due to its range of known protein domains, allowing it to bind DNA, proteins and RNA. A CLIP-seq study found hnRNP U associated with many types of RNA, such as snRNAs, lncRNAs and pre-mRNA introns (Xiao et al. 2012). These types of RNA are largely nuclear and often chromatin-associated. hnRNP U can bind to chromatin-associated RNAs and regulate chromatin structure (Nozawa et al. 2017; H. Fan et al. 2017). The nuclear localisation of lncRNAs *XIST* and *FIRRE*, and most likely *PANDA*, is controlled by hnRNP U (Hasegawa et al. 2010; Hacısuleyman et al. 2014; Puvvula et al. 2014), which indicates a potential role of hnRNP U in general ncRNA nuclear or chromatin localisation and retention.

In this chapter, the levels and cellular localisation of a variety of RNAs following hnRNP U depletion is assessed. Following initial qRT-PCR-based experiments, RNA-seq was employed to get a comprehensive view of the effects of hnRNP U depletion on the cellular localisation of RNA.

5.1 hnRNP U is required to maintain the levels of nuclear non-coding RNAs

Initially, the abundance of non-coding RNAs were measured following hnRNP U depletion (Figure 5.1.1). The abundance of lncRNAs such as *FIRRE* and *TERC* are heavily reduced following hnRNP U depletion whereas lncRNAs such as *DANCR* and *PANDA* are not. As hnRNP U seems to perform nuclear functions, perhaps the differences in lncRNA sensitivity to hnRNP U loss could be to do with their cellular localisation. The RNA from control and hnRNP U depleted cells were fractionated into chromatin-associated, nucleoplasmic and cytoplasmic RNA samples and the cellular distribution of these lncRNAs was quantified (Figure 5.1.2). The lncRNAs that show substantial chromatin enrichment (Figure 5.1.2.A) correspond to the transcripts with reduced levels following hnRNP U loss. Conversely, the lncRNAs that show less nuclear restriction (Figure 5.1.2.B) are unaffected by hnRNP U loss. This demonstrates the specificity of hnRNP U action on nuclear and chromatin-associated lncRNAs.

Polymerase II transcribed snRNAs also show reduced steady state levels following the loss of hnRNP U, in contrast to U6 snRNA which is polymerase III transcribed (Figure 5.1.3). Perhaps hnRNP U is playing a role in the transcription or co-transcriptional processing of RNA pol II snRNAs. hnRNP U has been shown to be necessary for U2 snRNP maturation (Xiao et al. 2012), further linking hnRNP U to snRNA biogenesis.

Interestingly, the steady state levels of a variety of spliced mRNA and pre-mRNA transcripts are not significantly affected by hnRNP U depletion (Figure 5.1.4). As pre-mRNAs are chromatin-associated but are not affected by hnRNP U depletion, hnRNP U seems to play a role in the maintenance of non-coding RNA specifically.

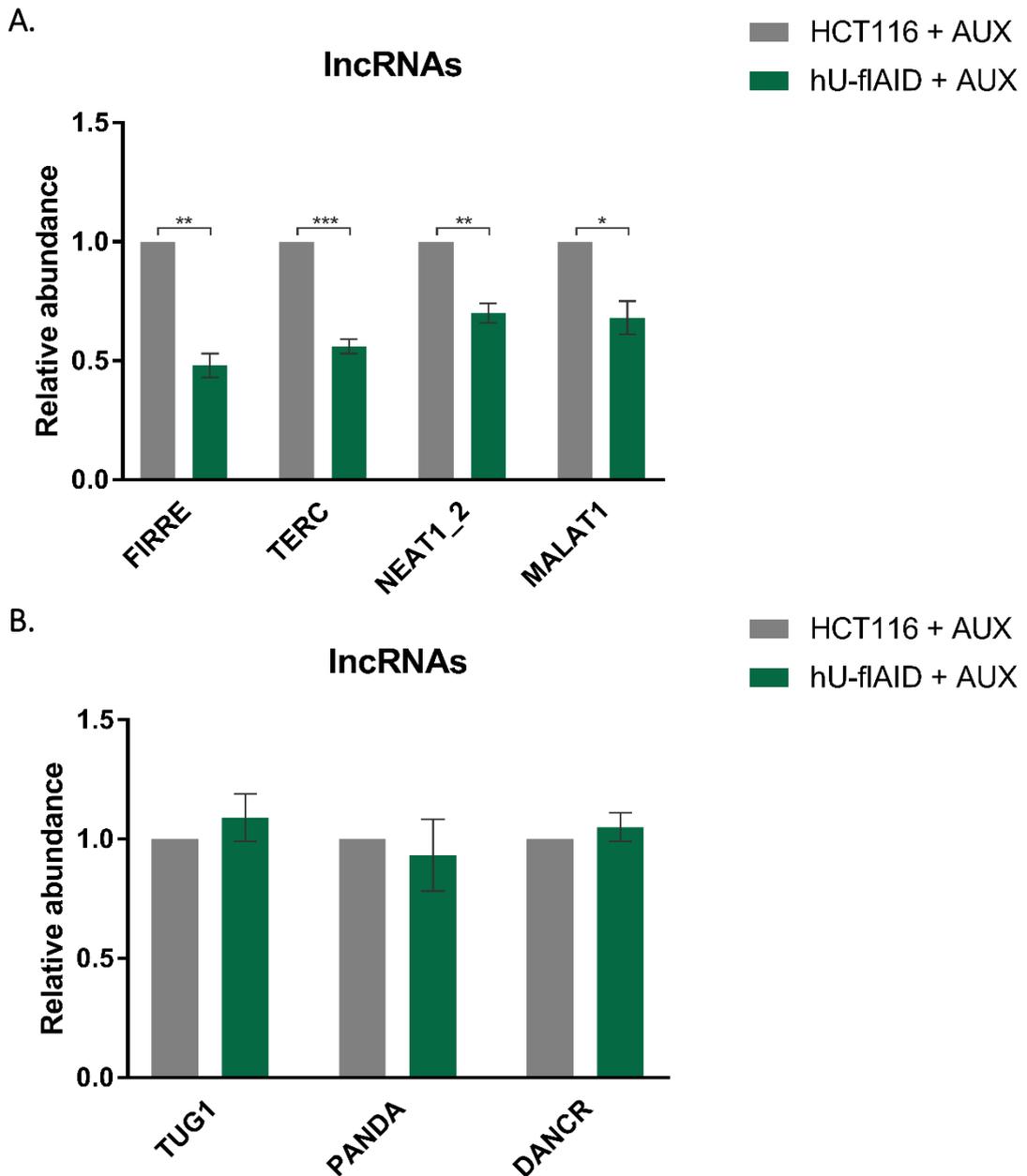


Figure 5.1.1 Nuclear-restricted lncRNA levels are reduced following the loss of hnRNP U

qRT-PCR data from four samples of control (HCT116 + AUX) and hnRNP U depleted cells after the addition of auxin for 24 hours (hU-flAID + AUX). The average abundance of each transcript is shown relative to the control. Error bars represent the SEM and all abundances are normalised to the abundance of U6 in the sample. A t-test was carried out as described in Chapter 2 to determine the statistical significance. The average abundance of lncRNAs (A) localised to the chromatin and (B) localised throughout the cell are shown.

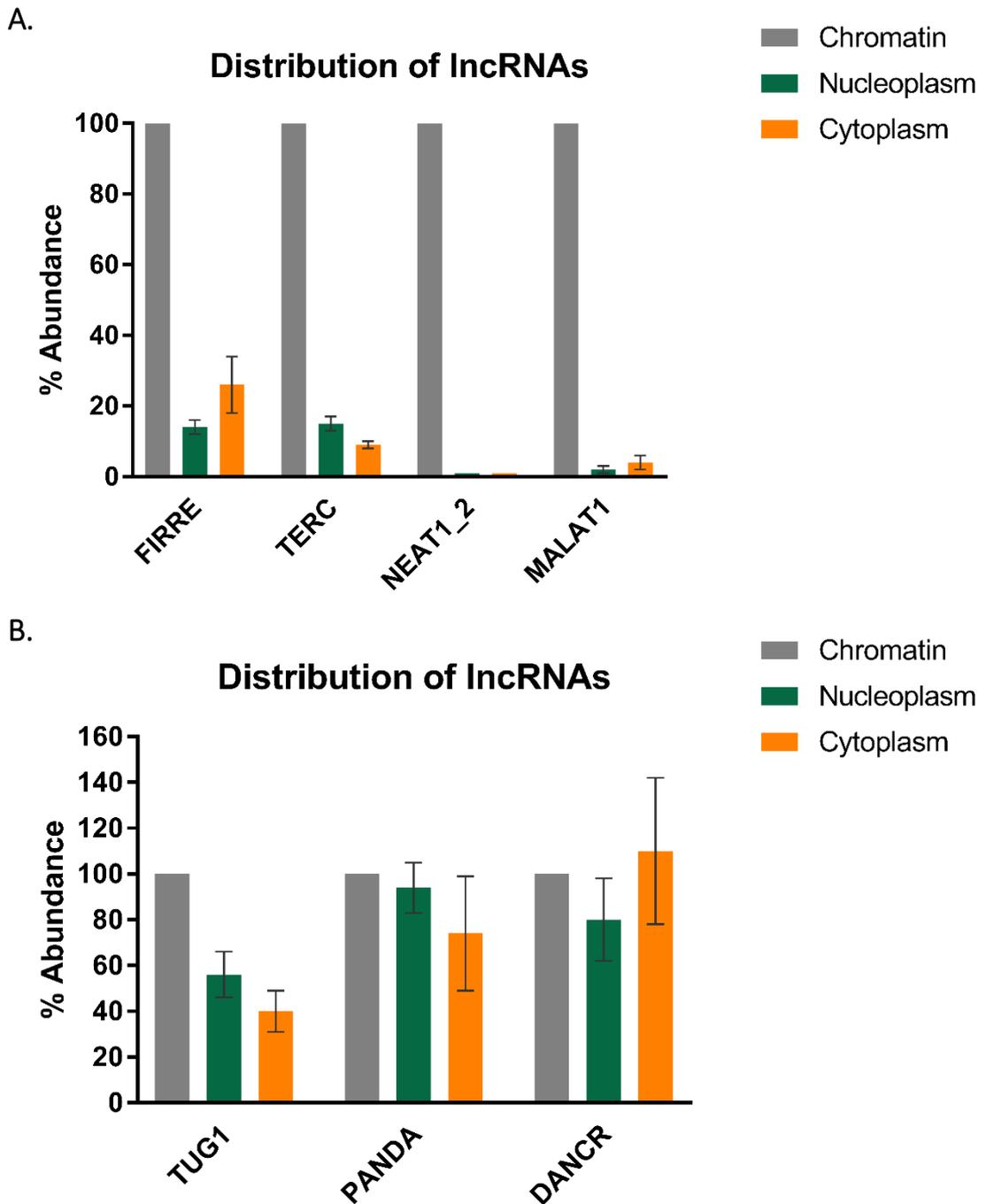


Figure 5.1.2 lncRNAs show different distributions between cellular compartments

qRT-PCR was carried out using three sets of fractionated RNA samples from HCT116 cells. The mean percentage abundance of the transcript is shown relative to the abundance in the chromatin fraction, 100%. All values are normalised to the abundance of 18S rRNA in the sample. Error bars represent the SEM.

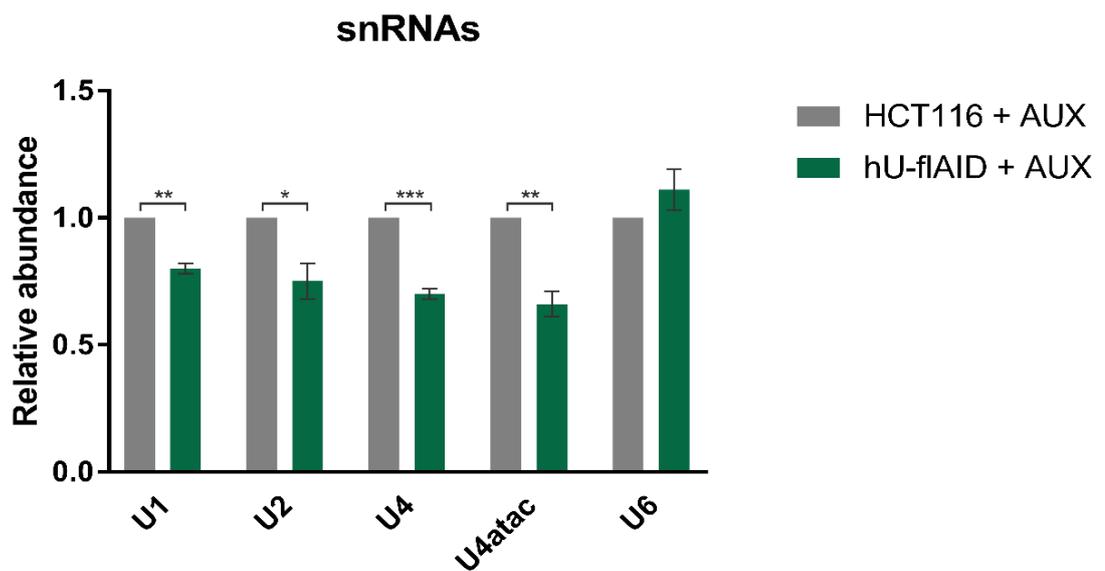


Figure 5.1.3 RNA pol II transcribed snRNA levels are reduced following the loss of hnRNP U

qRT-PCR data was collected and plotted as described in Figure 5.1.1. HCT116 and hU-flAID cells were treated with auxin for 24 hours. Abundances are normalised to the abundance of U6 in the sample, except for U6 itself where 18S rRNA was used. The abundance snRNAs are shown relative to the control.

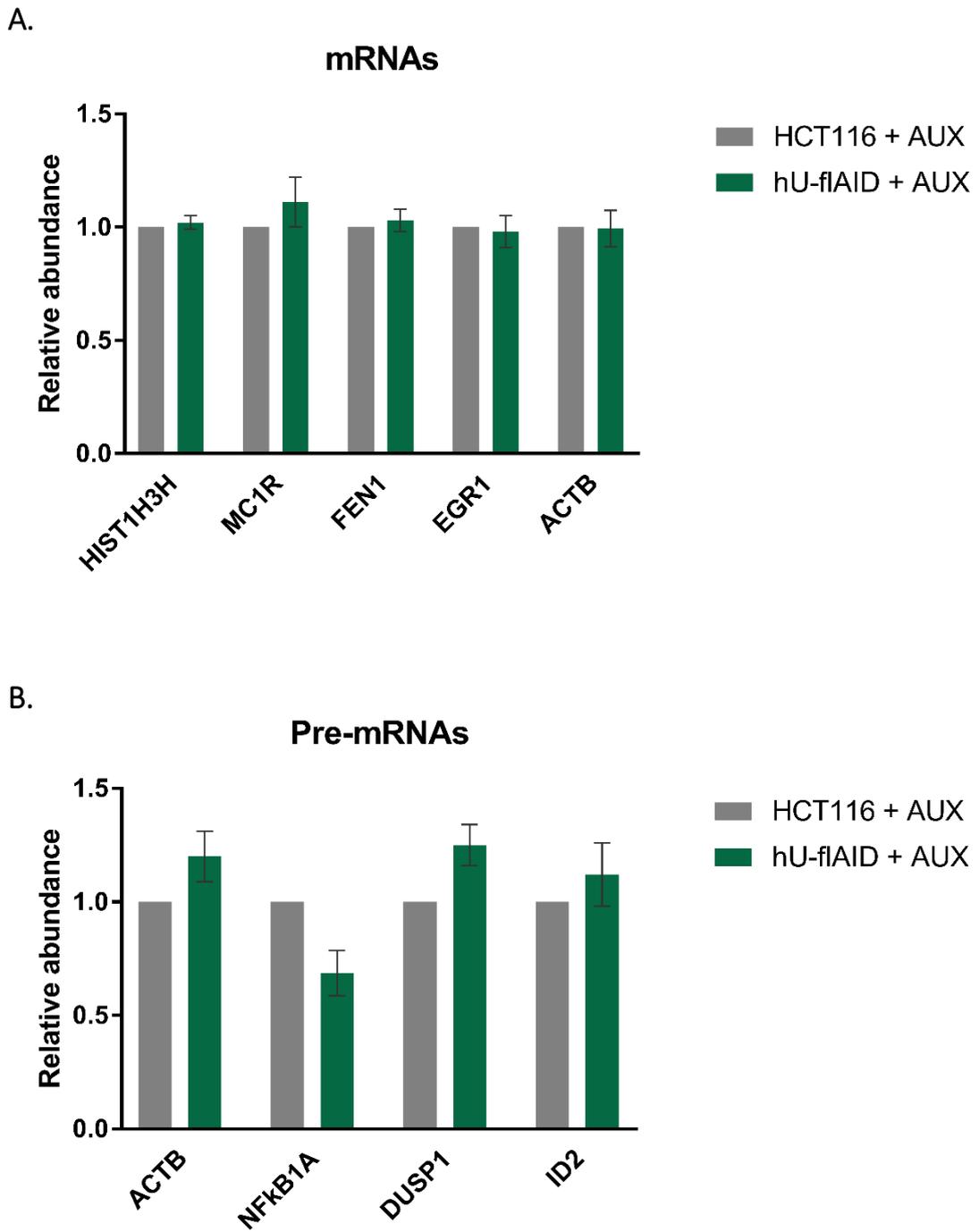


Figure 5.1.4 mRNA and pre-mRNA levels are not significantly different following the loss of hnRNP U

qRT-PCR data was collected, normalised and plotted as described in Figure 5.1.1. HCT116 and hU-flAID cells were treated with auxin for 24 hours. The abundance of (A) spliced mRNAs and (B) un-spliced pre-mRNAs are shown relative to the control.

hnRNP U could be involved in co-transcriptional processing of the affected nuclear ncRNAs as it is bound to the CTD of RNA pol II (Carty & Greenleaf 2002) and is heavily bound to chromatin (Göhring & Fackelmayer 1997). The inefficient processing of these RNA transcripts following hnRNP U loss may then cause them to be degraded. hnRNP U could also be involved directly in transcription as the protein has been linked to transcription elongation (Kim & Nikodem 1999; Obrdlik et al. 2008). Additionally, due to the reduced levels of snRNAs in hnRNP U depleted cells and the published roles in splicing, hnRNP U loss could be causing aberrant splicing and therefore activating nuclear surveillance mechanisms to degrade the transcript. However, as the levels of pre-mRNAs do not seem to be affected, global splicing changes do not seem to be occurring. As hnRNP U has been commonly described as a localisation factor in ncRNA lifecycles, we pursued cellular mislocalisation as the cause of the decrease in abundance of various ncRNAs. Perhaps mislocalisation triggers their degradation by nuclear RNA surveillance mechanisms.

5.2 hnRNP U depletion can result in chromatin-associated ncRNA release and degradation

Before assessing the role of hnRNP U in RNA localisation, the localisation of the hnRNP U protein itself was determined (Figure 5.2.1). HCT116 cells were fractionated into chromatin-associated, nucleoplasmic and cytoplasmic protein fractions. The majority of the hnRNP U protein is present in the two nuclear fractions, underlining its predominant role in the nucleus. The proteins Tubulin and Histone 3 are only present in the cytoplasmic and chromatin fractions respectively as expected, demonstrating a successful protein fractionation. Alyref and SSRP1 are also nuclear, further confirming a clean preparation.

To investigate non-coding localisation following hnRNP U depletion, the RNA from fractionated cells was collected and specific transcripts investigated using qRT-PCR. The quality of the RNA fractionations were first assessed using transcripts known to localise to a particular fraction (Figures 5.2.2-3). Chromatin-associated RNAs such as pre-mRNA and *NEAT1_2* lncRNA were heavily represented in the chromatin fractions whereas mRNAs and the lncRNA *DANCR* were localised throughout the cell. Enrichment of these transcripts in the particular fractions is expected biologically, representing successful fractionations.

The abundance of lncRNAs, such as *FIRRE*, *NEAT1_2* and *TERC* are generally reduced in all fractions following hnRNP U depletion (Figures 5.2.4-5), but most prominently in the chromatin fraction where they are mainly localised (Figure 5.1.2). The levels of lncRNA *TUG1* however are mainly reduced in the nucleoplasmic fraction. By qRT-PCR, *TUG1* does not localise as comprehensively to the chromatin compared to lncRNAs that are mostly reduced in the chromatin (Figure 5.1.2), perhaps explaining why the levels of *TUG1* in the nucleoplasm are more effected.

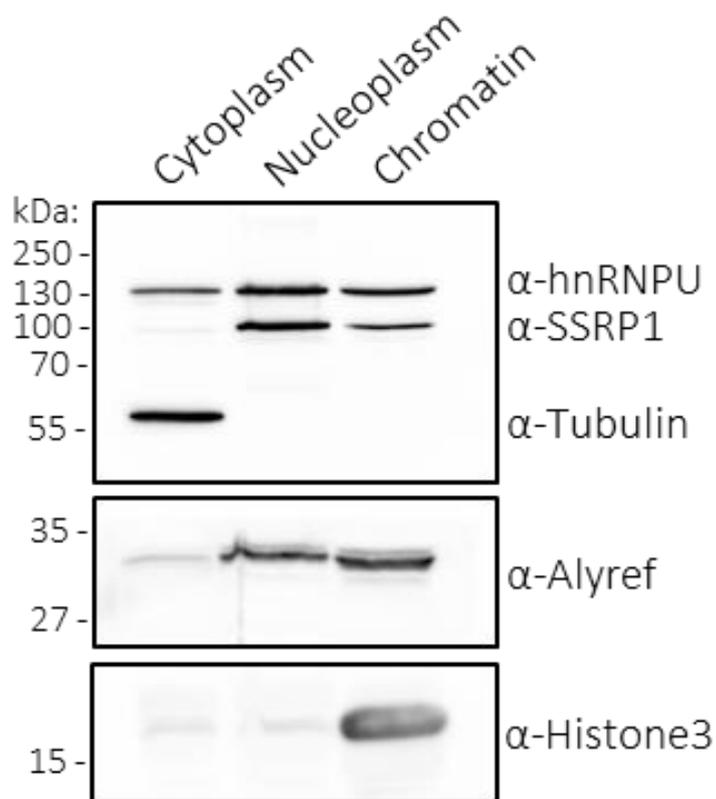
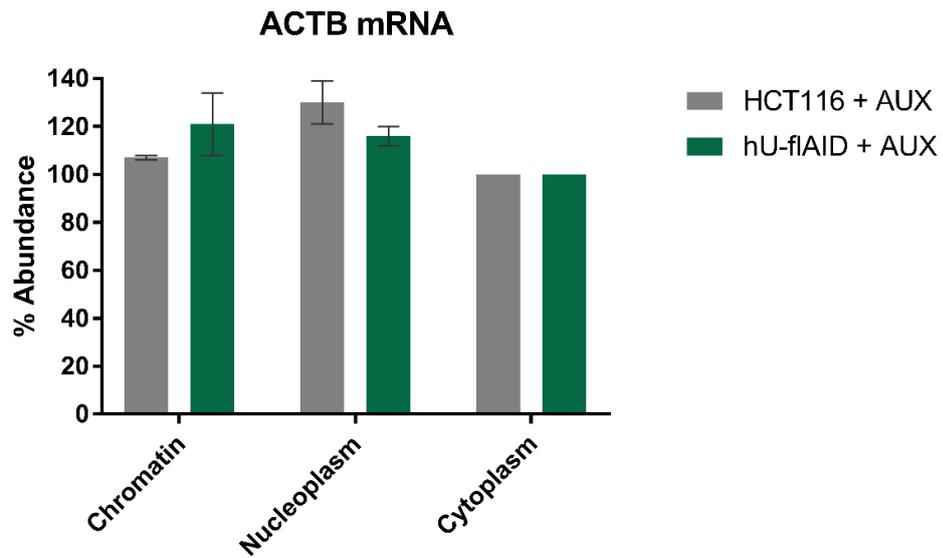


Figure 5.2.1 hnRNP U is mainly present in nuclear fractions

Western blot of chromatin, nucleoplasmic and cytoplasmic protein fractions from HCT116 cells. The majority of hnRNP U protein is in the two nuclear fractions. The blot was probed for control proteins (Tubulin, Histone 3, Alyref and SSRP1) as well as hnRNP U to check for successful fractionation.

A.



B.

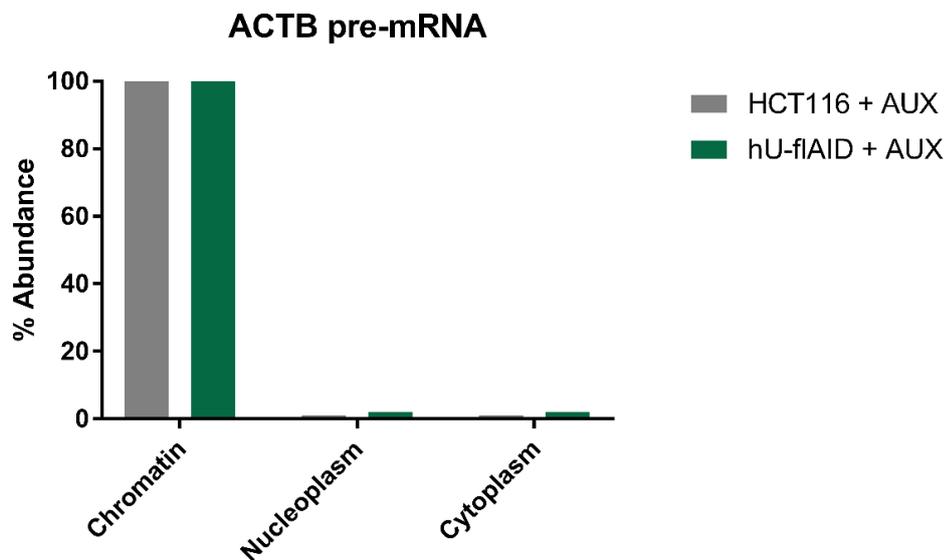
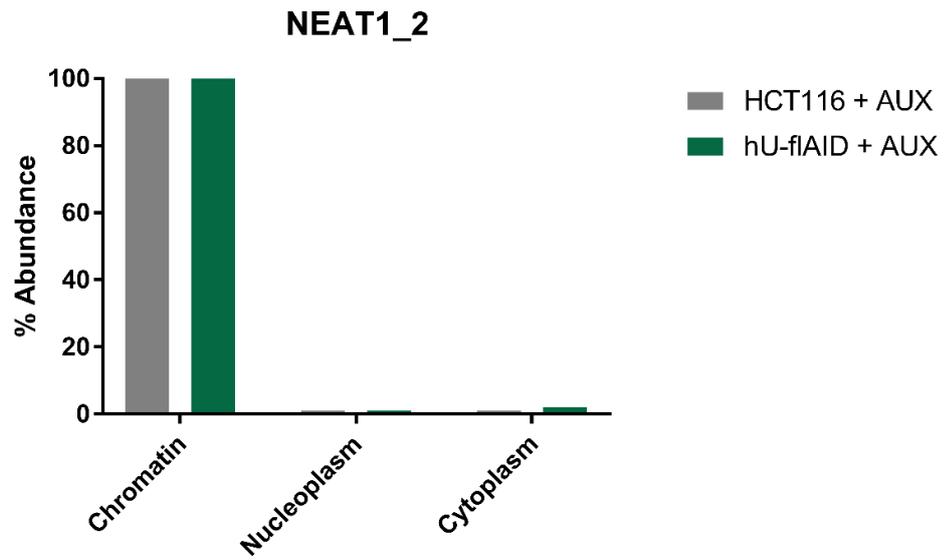


Figure 5.2.2 The relative abundance of ActinB spliced and un-spliced mRNA in cellular fractions demonstrates effective fractionation

qRT-PCR was carried out on three sets of fractionated RNA samples per condition. HCT116 and hU-flAID cells were treated with auxin for 24 hours. The mean percentage abundance of the transcript is shown relative to the abundance in the fraction predicted to contain the most of the transcript, 100%. All values are normalised to the abundance of 18S rRNA in the sample. Error bars represent the SEM. The abundance of (A) spliced ActinB (ACTB) mRNA and (B) unspliced ActinB pre-mRNA in each fraction is shown.

A.



B.

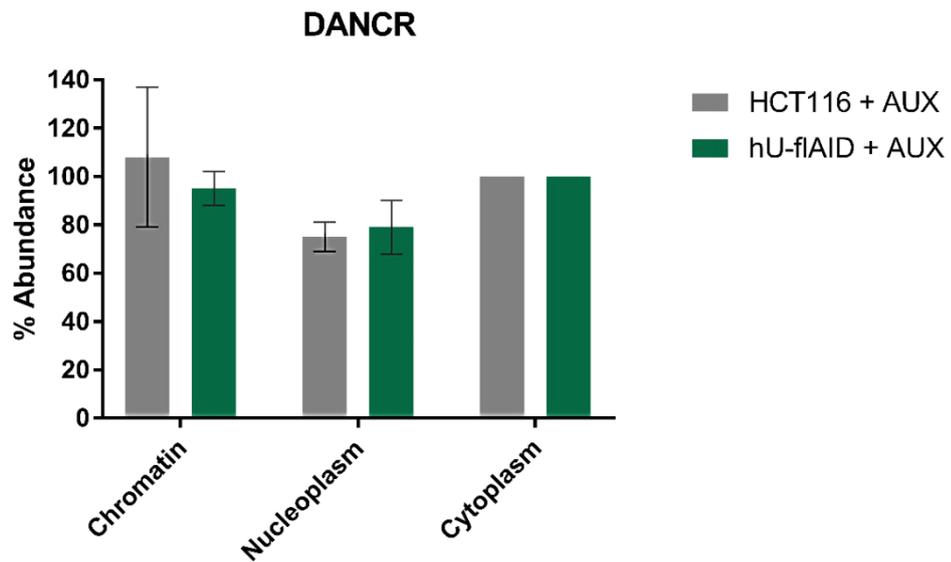


Figure 5.2.3 The relative abundance of lncRNAs in cellular fractions demonstrates effective fractionation

This data was ascertained and represented as in Figure 5.2.2. HCT116 and hU-flAID cells were treated with auxin for 24 hours. (A) Nuclear lncRNA *NEAT1_2* and (B) cytoplasmic lncRNA *DANCR* are shown.

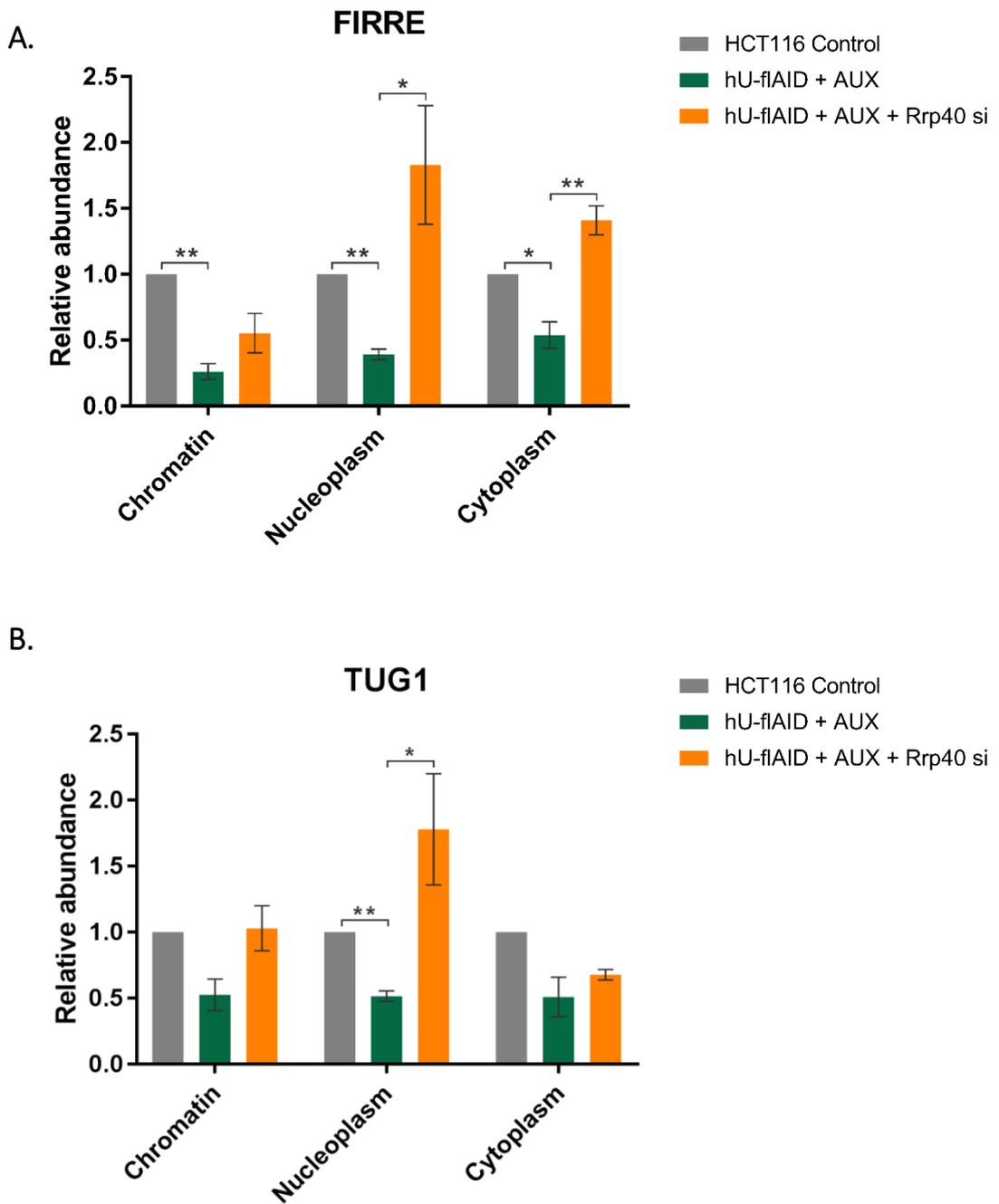


Figure 5.2.4 The levels of lncRNAs *FIRRE* and *TUG1* are reduced following the loss of hnRNP U but the levels are restored following disruption of the exosome

qRT-PCR was carried out on three sets of fractionated RNA samples per condition. HCT116 and hU-flAID cells were treated with auxin for 24 hours and 25 nM siRNA treatment carried out for 72 hours. The mean abundance is shown relative to the amount in the HCT116 + AUX and HCT116 + AUX + Cntrl si controls for hU-flAID + AUX and hU-flAID + AUX + Rrp40 samples respectively. All values are normalised to the abundance of 18S rRNA in the sample. Error bars represent the SEM. The abundance of lncRNAs (A) *FIRRE* and (B) *TUG1* in each fraction is shown.

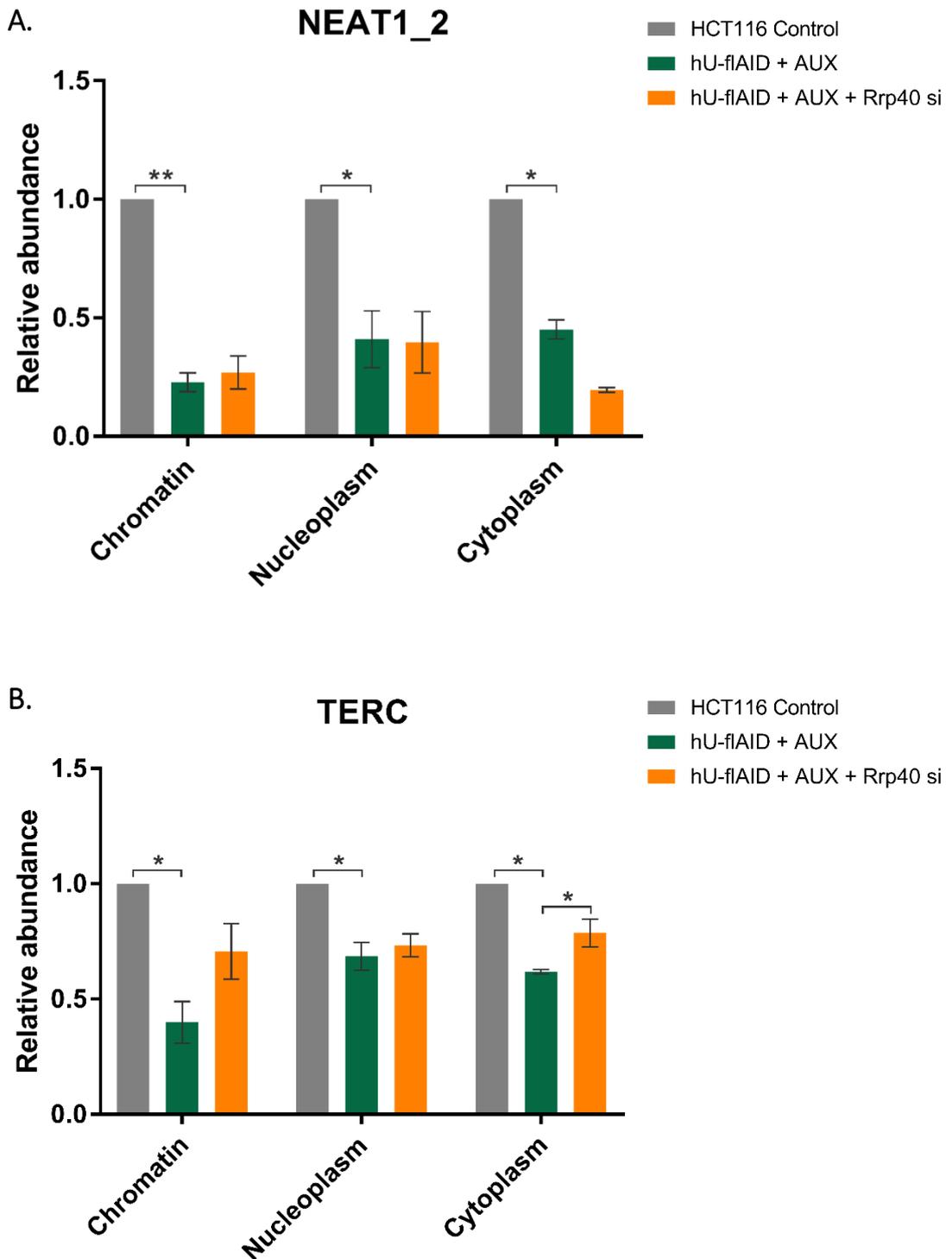


Figure 5.2.5 The levels of lncRNAs *NEAT1_2* and *TERC* are reduced following the loss of hnRNP U and the levels are not always restored following disruption of the exosome

Data was collected and represented as in Figure 5.2.4. HCT116 and hU-flAID cells were treated with auxin for 24 hours and siRNA treatment carried out for 72 hours. The abundance of lncRNAs (A) *NEAT1_2* and (B) *TERC* in each fraction is shown.

Following hnRNP U depletion, NFkB1A and ID2 pre-mRNAs are reduced when looking at fractionated RNA samples, with the majority being lost in the chromatin fraction much like the nuclear lncRNAs tested (Figure 5.2.6). ActinB pre-mRNA on the other hand is not reduced following hnRNP U loss (Figure 5.2.7). Perhaps the stability or higher expression level of “housekeeping” mRNAs such as ActinB prevents its loss during the course of depleting hnRNP U in the cell line. U1 snRNA levels are also decreased in the chromatin fraction (Figure 5.2.8), indicating that a wide range of non-coding RNA classes show reduced levels in the chromatin fraction following hnRNP U depletion.

The data so far presented suggests that the depletion of hnRNP U mainly effects non-coding RNA in the chromatin fraction. This could indicate that transcription of these genes is directly affected. However, you would then expect a similar reduction in RNA levels in all three fractions after hnRNP U loss following turnover of already transcribed RNAs. Although, the short timescale over which the cells are depleted of hnRNP U may not allow complete turnover of the transcripts already present in the cell. The reduction in the abundance of chromatin-associated lncRNAs, pre-mRNAs and snRNA U1 in the chromatin fraction and a less significant change in the other fractions could suggest leakage from the chromatin to the nucleoplasm and cytoplasm. This has already been demonstrated with lncRNAs *XIST* and *FIRRE* (Hacisuleyman et al. 2014; Hasegawa et al. 2010). Additionally, the abundance of pre-mRNAs ActinB and NFkB1A seems to increase in the nucleoplasmic and cytoplasmic fractions (Figures 5.2.6-7), although not shown to be statistically significant. Furthermore, an increase in abundance of the RNAs may not be seen in the nucleoplasm and cytoplasm due to degradation.

To investigate the degradation hypothesis, the exosome complex was disabled along with the depletion of hnRNP U to see if RNA levels were restored now that they can no longer be effectively degraded (orange bars in Figures 5.2.4-8). This was done using siRNA treatment against Rrp40 mRNA, a core component of the exosome complex (Kilchert et al. 2016). lncRNAs, alongside mRNAs, have been shown to be human exosome substrates (H. Fan et al. 2017) and Rrp40 depletion increases the levels of lncRNAs in the nucleoplasm of human cells (Schlackow et al. 2017).

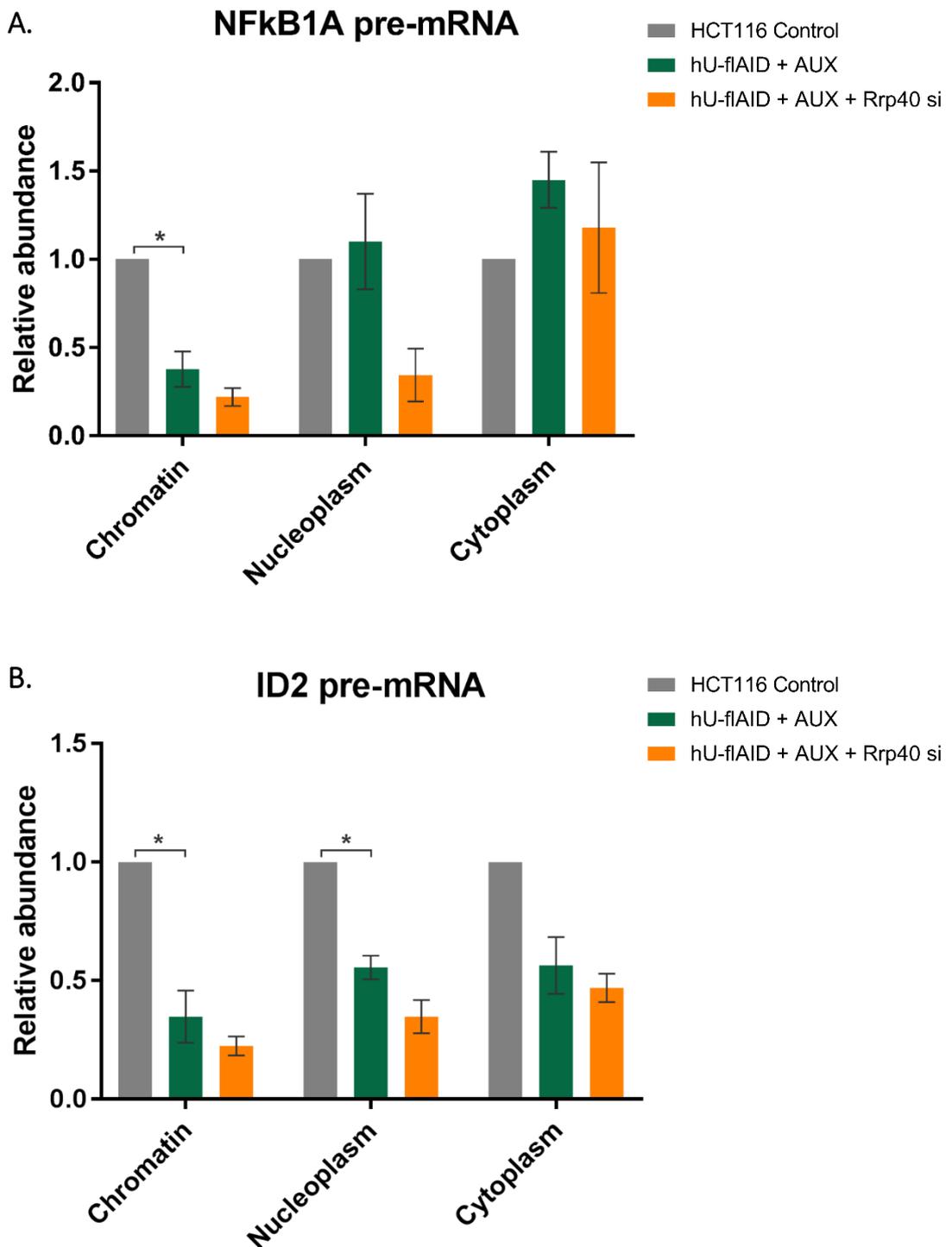


Figure 5.2.6 NFkB1A and ID2 pre-mRNA levels are reduced following the loss of hnRNP U and the levels are not restored following disruption of the exosome

Data was collected and represented as in Figure 5.2.4. HCT116 and hU-flAID cells were treated with auxin for 24 hours and siRNA treatment carried out for 72 hours. The abundance of pre-mRNAs (A) NFkB1A and (B) ID2 in each fraction is shown.

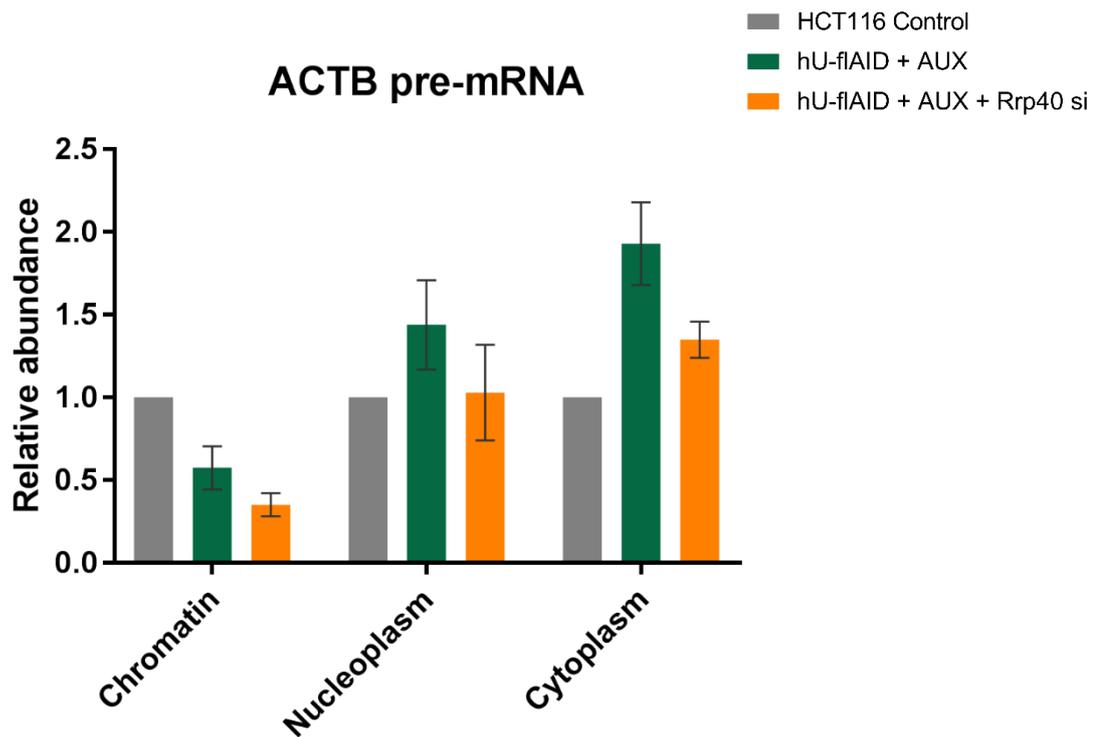


Figure 5.2.7 The levels of ActinB pre-mRNA are not significantly different following the loss of hnRNP U and its abundance is not affected by disruption of the exosome

Data was collected and represented as in Figure 5.2.4. HCT116 and hU-flAID cells were treated with auxin for 24 hours and 25 nM siRNA treatment carried out for 72 hours. The abundance of ActinB pre-mRNA (ACTB) in each fraction is shown.

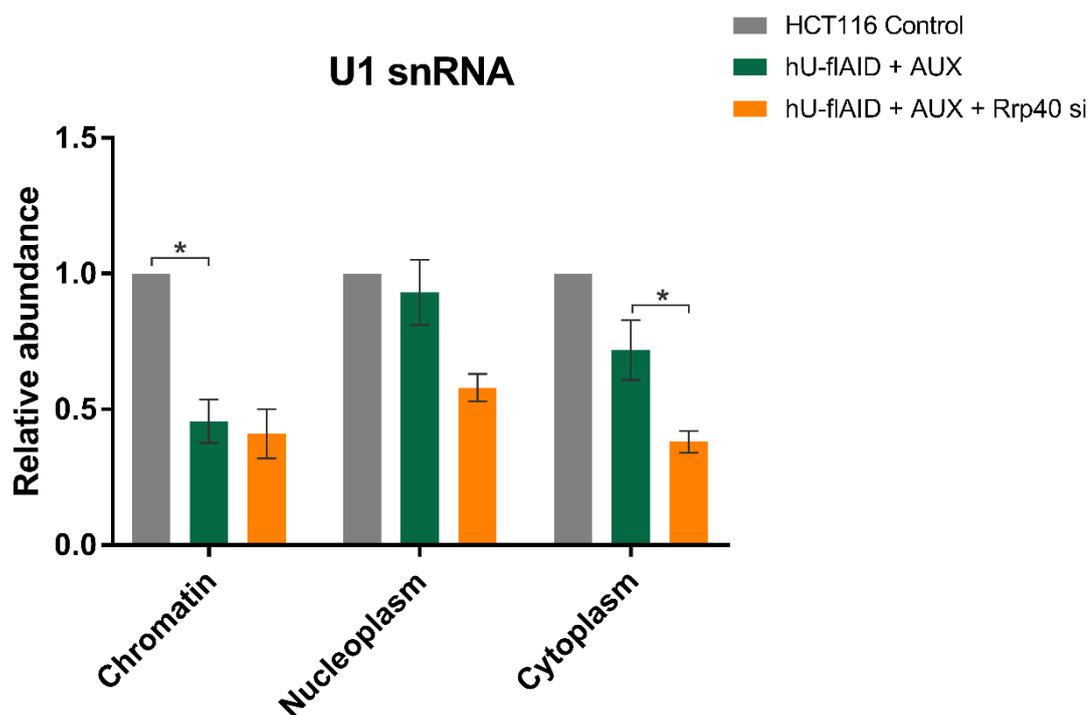


Figure 5.2.8 U1 snRNA levels are reduced in the chromatin fraction by the loss of hnRNP U and its abundance is not affected by disruption of the exosome

Data was collected and represented as in Figure 5.2.4. HCT116 and hU-flAID cells were treated with auxin for 24 hours and 25 nM siRNA treatment carried out for 72 hours. The abundance of U1 snRNA in each fraction is shown.

The abundance of lncRNAs *FIRRE* and *TUG1* showed a significant increase in the nucleoplasm once the exosome had been disrupted (Figure 5.2.4), endorsing the theory that they are released from the chromatin following hnRNP U depletion and then degraded in the nucleoplasm (Schlackow et al. 2017). *FIRRE* was also stabilised in the cytoplasmic fraction, suggesting that *FIRRE* may leak out to the cytoplasm as previously seen by FISH (Hacisuleyman et al. 2014). Interestingly, lncRNA *NEAT1_2* levels were not restored at all following exosome disruption (Figure 5.2.5). This could be due to the selectivity of exosome substrates and it may be degraded by an alternative method. None of the pre-mRNAs tested or U1 snRNA were affected by the disruption of the exosome, suggesting that they may not be substrates of the nuclear exosome or that they are more heavily protected from degradation compared to lncRNAs. Alternatively, transcription of *NEAT1_2*, U1 snRNA and the pre-mRNAs tested may be altered by hnRNP U loss.

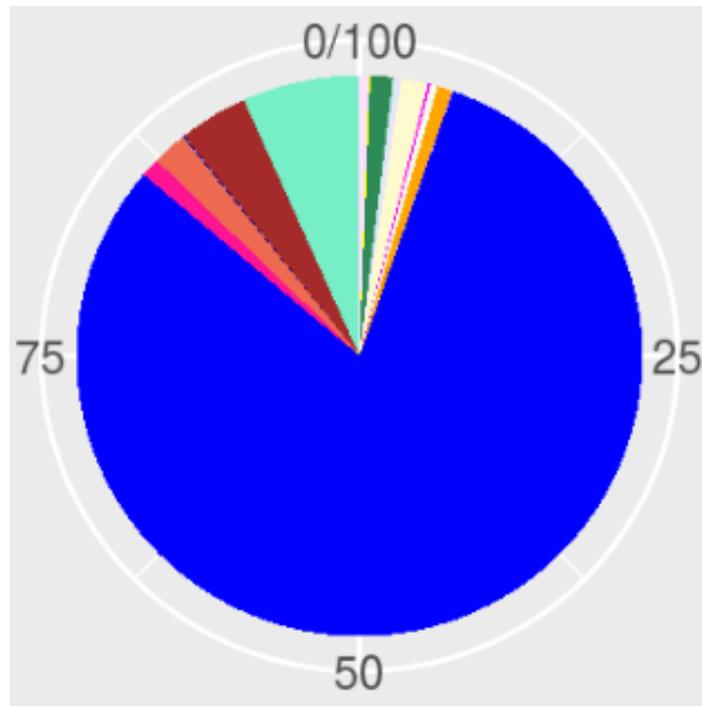
Despite some common conclusions made about the effect of hnRNP U depletion on various classes of RNAs, there is still a lot of variation within the classes despite only a few transcripts being tested. Additionally, only a subset of classes of non-coding and chromatin-associated RNAs have been investigated. Therefore, to get a global view on what is happening to the localisation of RNAs following hnRNP U depletion, RNA-seq was carried out.

5.3 A global approach demonstrates the general role of hnRNP U in the retention of non-coding RNA on the chromatin

To get a comprehensive view of what happens to RNA following hnRNP U depletion, RNA sequencing was carried out with the same fractionated samples used in Chapter 5.2. The RNA from the chromatin, nucleoplasm and cytoplasm in control and hnRNP U depleted cells were sequenced but not the Rrp40 siRNA conditions. Three replicates of each condition were sequenced.

The RNA-seq data was analysed by Jacob Parker and Dr Ian Sudbery who consequently produced the figures in this chapter. The RNA-seq data is generally discussed in terms of biotypes (or classes) of RNA to more comprehensively assess how RNA is affected by hnRNP U depletion. The general distribution of RNA biotypes out of all the expressed genes in the control samples was assessed and each biotype is shown as a percentage of all expressed genes (Figure 5.3.1). Protein coding genes make up the vast majority of the annotated expressed genes in this sequencing experiment and non-coding RNAs make up less than 25% of all sequenced genes. Antisense and long and long intergenic RNAs (lnc/lincRNA) make up the majority of the non-coding RNAs sequenced.

To look at differential expression of genes, the threshold for classing a gene as up or down regulated is a \log_2 fold change (LFC) of equal to or greater than 1 or -1 respectively. This is equivalent to at least doubling or halving in abundance. The same threshold was used in all comparisons. Initially, comparisons were made between the same fractions in the control condition compared to the hnRNP U knockdown.



PercentageBiotype

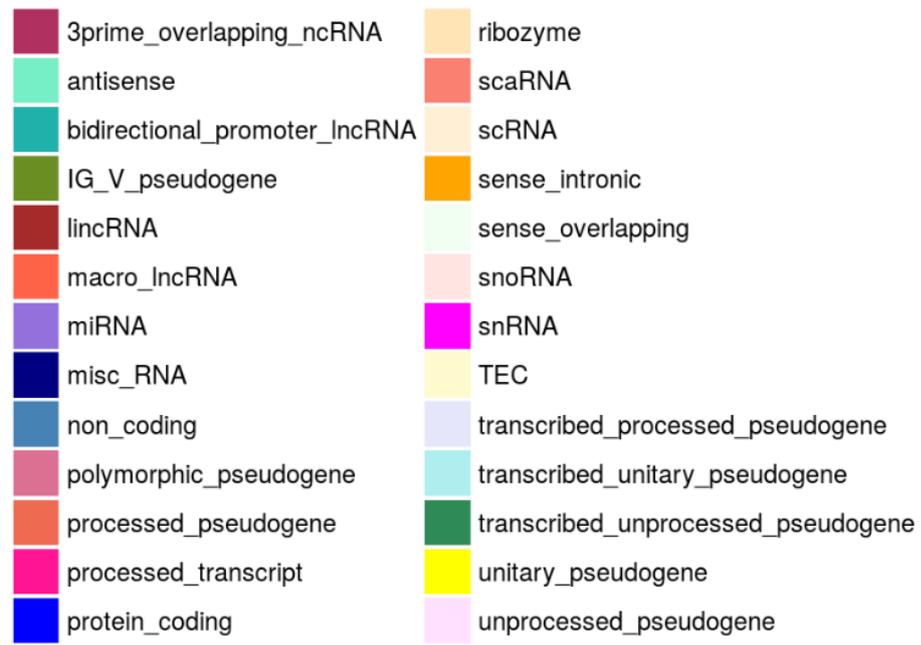


Figure 5.3.1 The distribution of RNA biotypes out of all annotated expressed genes in control samples

The percentage of RNA biotypes out of all the annotated expressed genes in the control (HCT116 + AUX) samples.

Chromatin

In the chromatin fraction, 2212 genes are differentially expressed in the hnRNP U knockdown condition. The vast majority of these genes, 78%, are downregulated rather than upregulated (Figure 5.3.2a). The percentage of the up and downregulated genes that are classed as a particular biotype is also shown (Figure 5.3.2b). The upregulated genes have a similar RNA biotype distribution as the general distribution (Figure 5.3.1). Although protein coding genes still make up the majority of the downregulated genes, non-coding RNAs such as lincRNAs, snRNAs and antisense RNAs made up a much larger proportion of the downregulated genes than in the upregulated genes and the general distribution (Figure 5.3.2b-c). In fact, there is at least double the proportion of most non-coding biotypes in the downregulated set and there are four times as many lincRNA genes in the downregulated set than in the upregulated set (Figure 5.3.2c). In general, these non-coding biotypes must be vulnerable to the loss of hnRNP U in the chromatin fraction.

On the other hand, differentially expressed protein coding genes are generally upregulated (Figure 5.3.2c). As the depletion of hnRNP U results in splicing misregulation (Xiao et al. 2012; Ye et al. 2015; Meininger et al. 2016)(Chapter 1.5.1), perhaps the upregulation of protein coding genes in the chromatin fraction is due to failed or improper splicing of pre-mRNA transcripts. This may lead to the transcripts not being released from the chromatin for further processing and export and therefore become enriched in the chromatin fraction. The data from the chromatin fraction highlights a profound difference between the effects of hnRNP U depletion on coding and non-coding RNA biotypes.

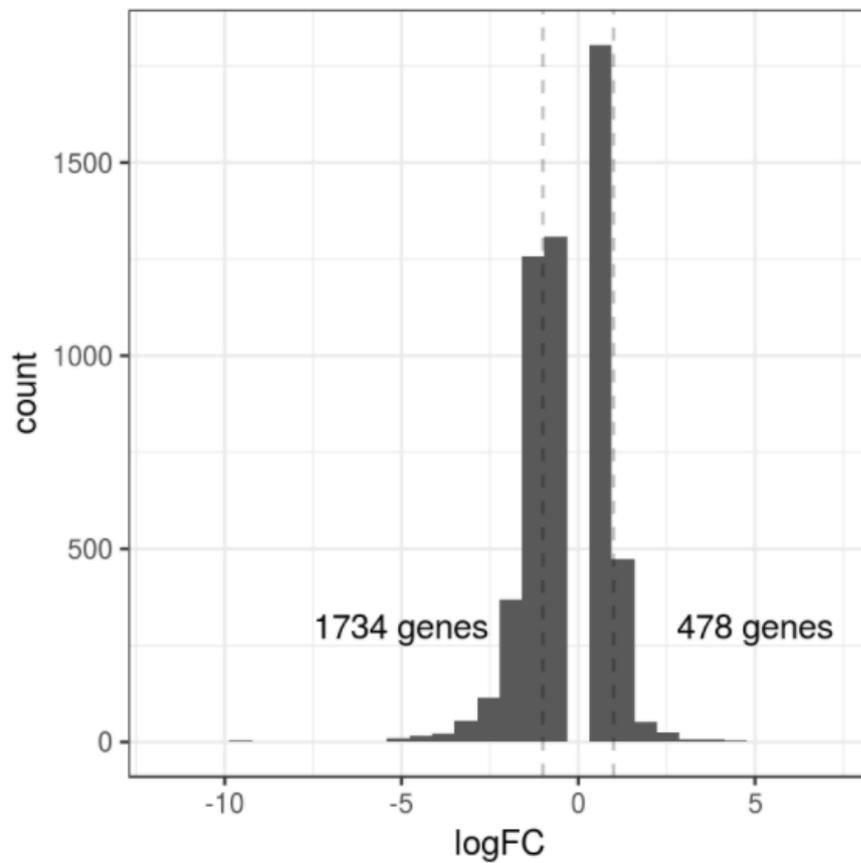
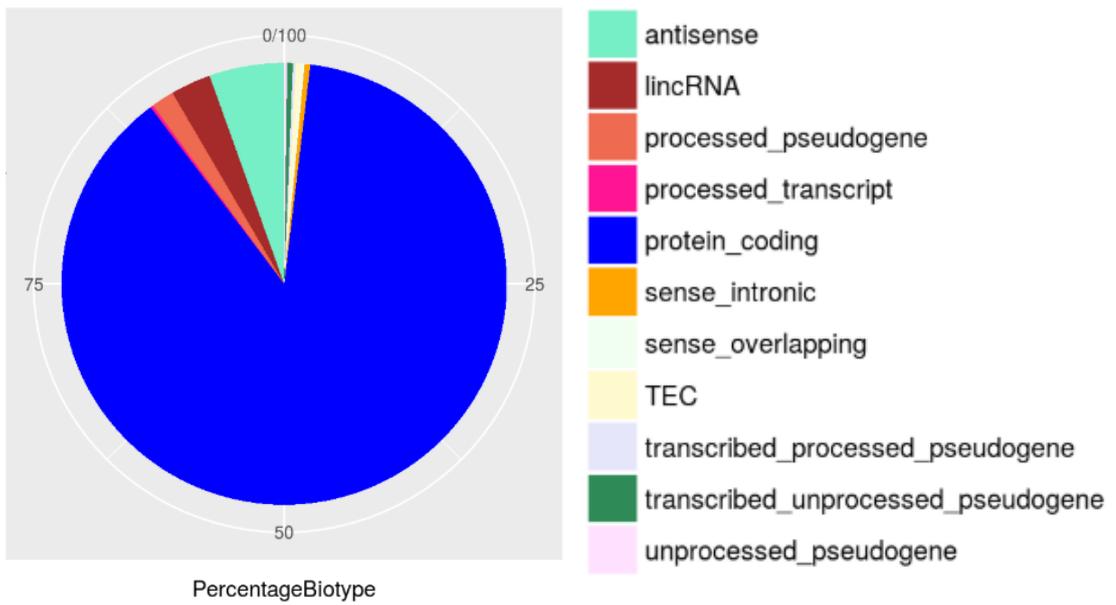


Figure 5.3.2a hnRNP U depletion results in over 2000 differentially expressed genes in the chromatin fraction

Following RNA-seq, the expression of genes in the hnRNP U knockdown chromatin samples were compared to the chromatin control samples. This graph represents the number of differentially expressed genes (count) by their log fold changes (logFC). LogFC is \log_2 fold change. The dotted line indicates the significance threshold of a logFC of -1 or +1 which is halving or doubling in expression level. The values shown, 1738 and 478, are the number of downregulated and upregulated genes that are above the respective significance thresholds.

A. Upregulated genes



B. Downregulated genes

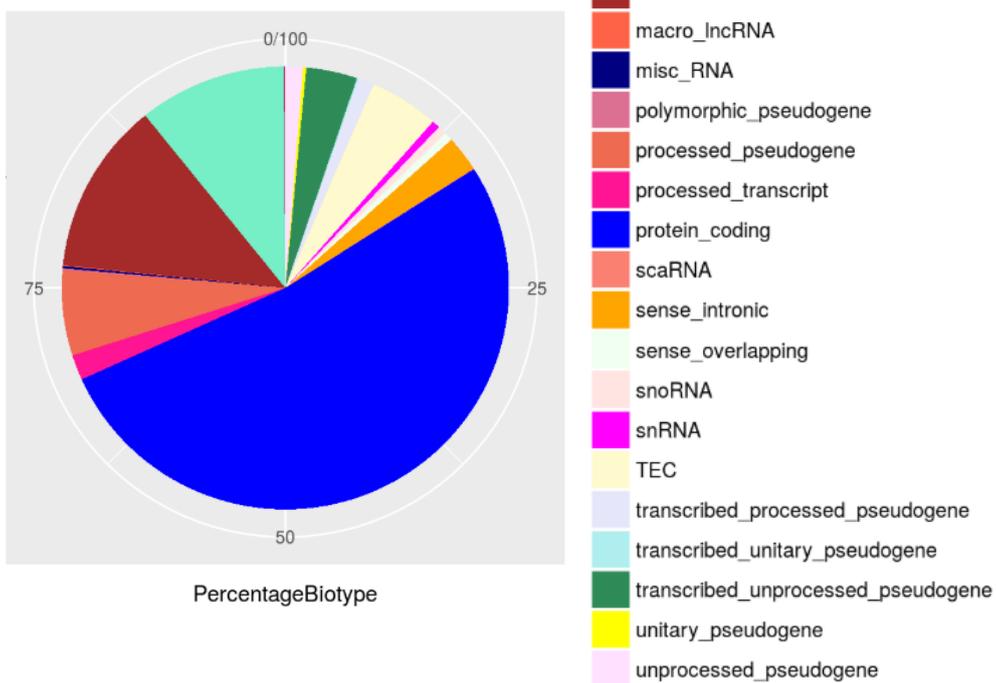


Figure 5.3.2b Non-coding RNAs make up a larger proportion of the downregulated genes than the upregulated genes in the chromatin fraction

The percentage of the significantly (A) upregulated and (B) downregulated genes in 5.3.2a shown as RNA biotypes.

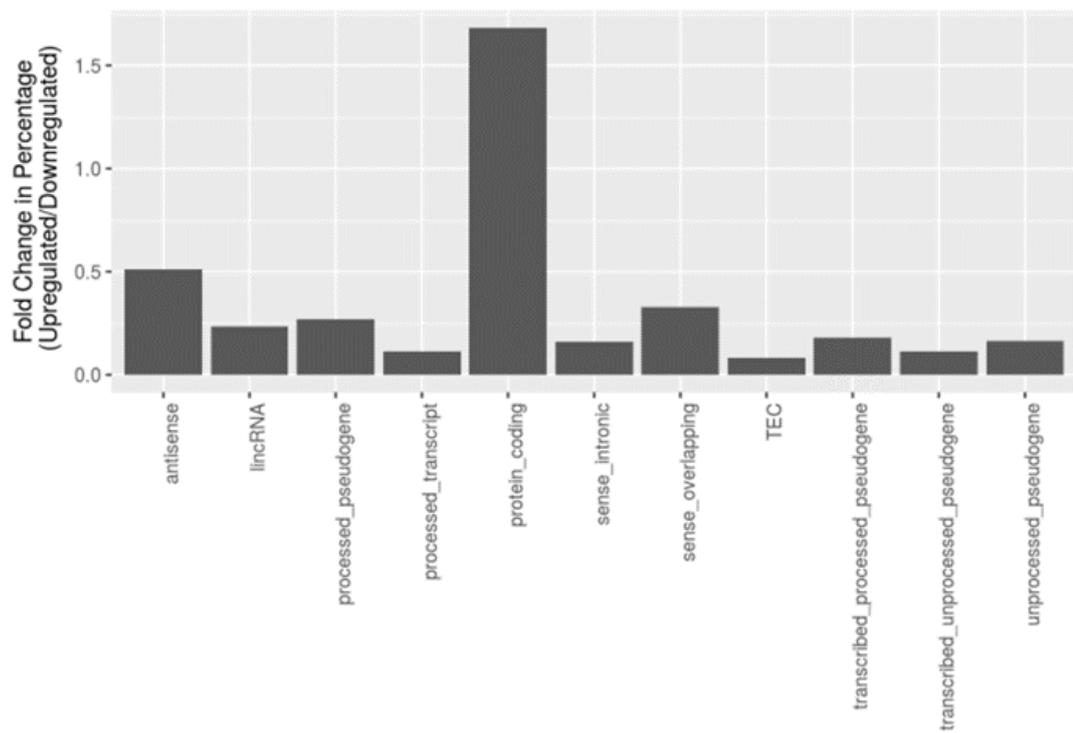


Figure 5.3.2c Non-coding RNAs make up a larger proportion of the downregulated genes than the upregulated genes in the chromatin fraction

The percentages from 5.3.2b represented as the percentage of the biotype in the upregulated set of genes compared to the downregulated set of genes. A fold change of greater than 1 indicates more of the biotype is upregulated where as a fold change of less than one indicates more of the biotype is downregulated.

The downregulation of non-coding RNAs is probably due to degradation in the absence of hnRNP U, as previously shown in Schlackow et al. 2017 and with lncRNAs studied in Chapter 5.2. A transcriptional downregulation of non-coding RNAs cannot be discounted however, although such a strong effect over many non-coding RNA biotypes specifically suggests a post or co-transcriptional effect. Additionally, if the loss of hnRNP U had an effect on RNA polymerase II transcription, we might have expected a similar distribution of up and down regulated genes for protein coding and non-coding genes.

For the chromatin fraction, the RNA-seq data is in alignment with the qRT-PCR data shown earlier in this chapter. There is widespread downregulation of nuclear non-coding RNAs after depletion of hnRNP U with the bulk of the loss in the chromatin fraction.

Nucleoplasm

In the nucleoplasmic fraction of the hnRNP U knockdown samples, 1069 genes are differentially expressed with 60.7% of them being downregulated (Figure 5.3.3a). There are fewer downregulated genes in the nucleoplasmic fraction compared to the chromatin and the proportion of upregulated to downregulated genes is more even. This shows that there is less of an impact on the downregulation of genes in the nucleoplasm compared to the chromatin, potentially as many non-coding RNAs localise to the chromatin.

In the nucleoplasm, the protein coding genes make up the largest proportion of up and down regulated genes as seen previously in the chromatin fraction and the general distribution. However, in contrary to the chromatin fraction, the non-coding RNA biotypes now make up a greater proportion in the upregulated genes than the downregulated genes (Figure 5.3.3b-c). The majority of the non-coding RNA biotypes are enriched in the upregulated genes compared to the downregulated, such as lincRNAs which have an almost two fold enrichment (Figure 5.3.3c).

A greater proportion of non-coding RNAs, such as lincRNAs and antisense RNAs, are downregulated in the chromatin fraction and a greater proportion are upregulated in the nucleoplasm. This supports the hypothesis of a general movement or mislocalisation of non-coding RNA from the chromatin to the nucleoplasmic fraction.

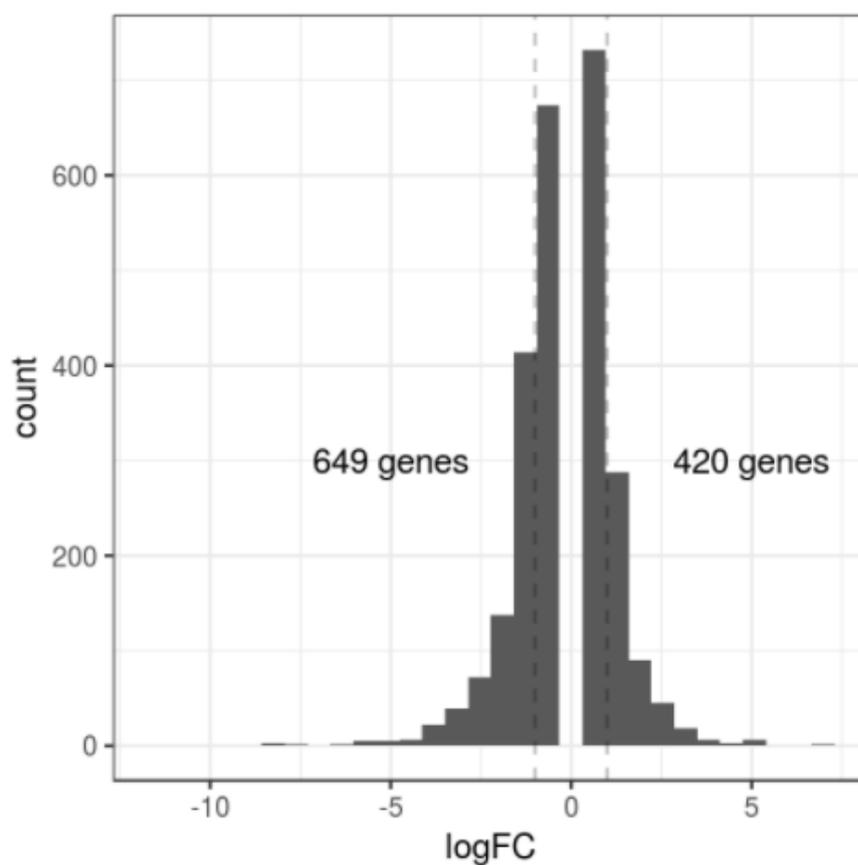
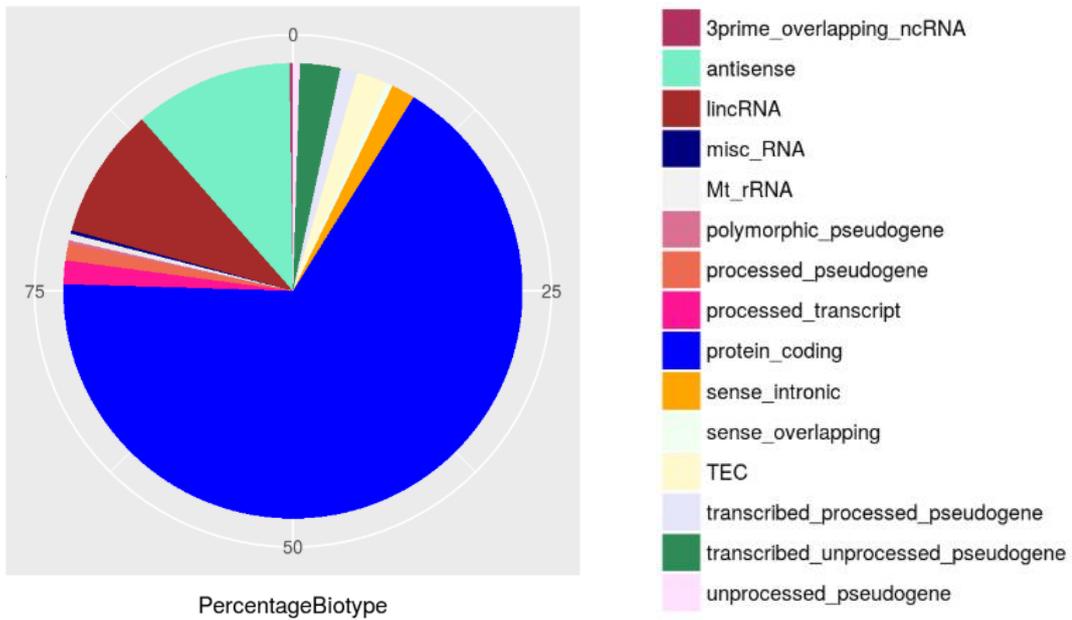


Figure 5.3.3a hnRNP U depletion results in over 1000 differentially expressed genes in the nucleoplasmic fraction

Following RNA-seq, the expression of genes in the hnRNP U knockdown nucleoplasmic samples were compared to the control nucleoplasmic samples. The figure is annotated as in Figure 5.3.2a and the same thresholds used.

A. Upregulated genes



B. Downregulated genes

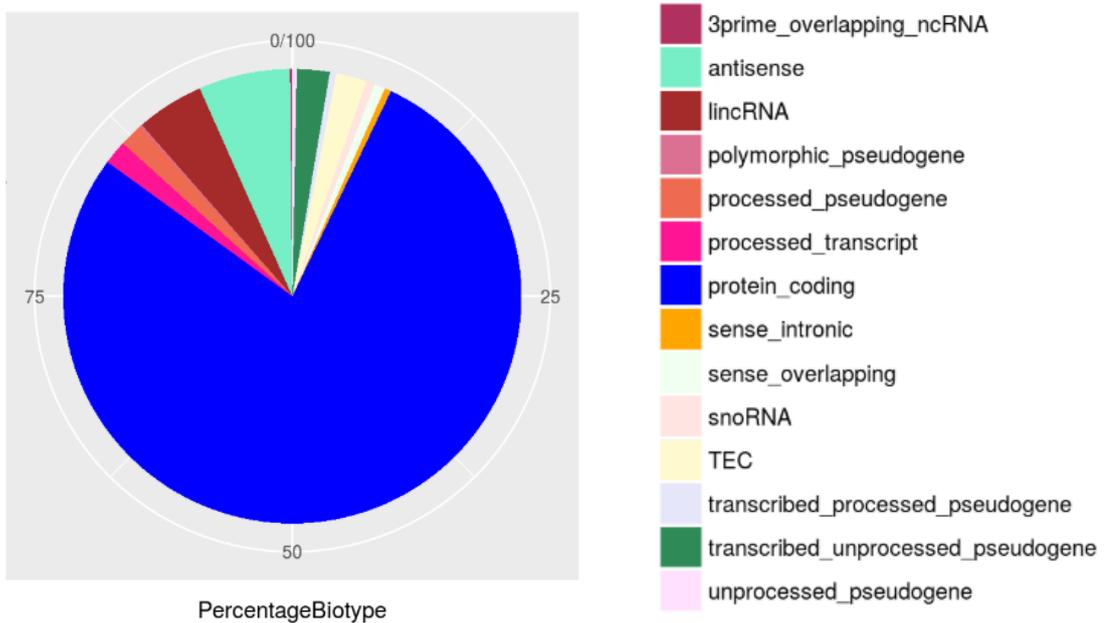


Figure 5.3.3b Many non-coding RNA biotypes make up a larger proportion of the upregulated genes than the downregulated genes in the nucleoplasmic fraction, in contrast to the chromatin fraction

The percentage of the significantly (A) upregulated and (B) downregulated genes in 5.3.3a shown as RNA biotypes.

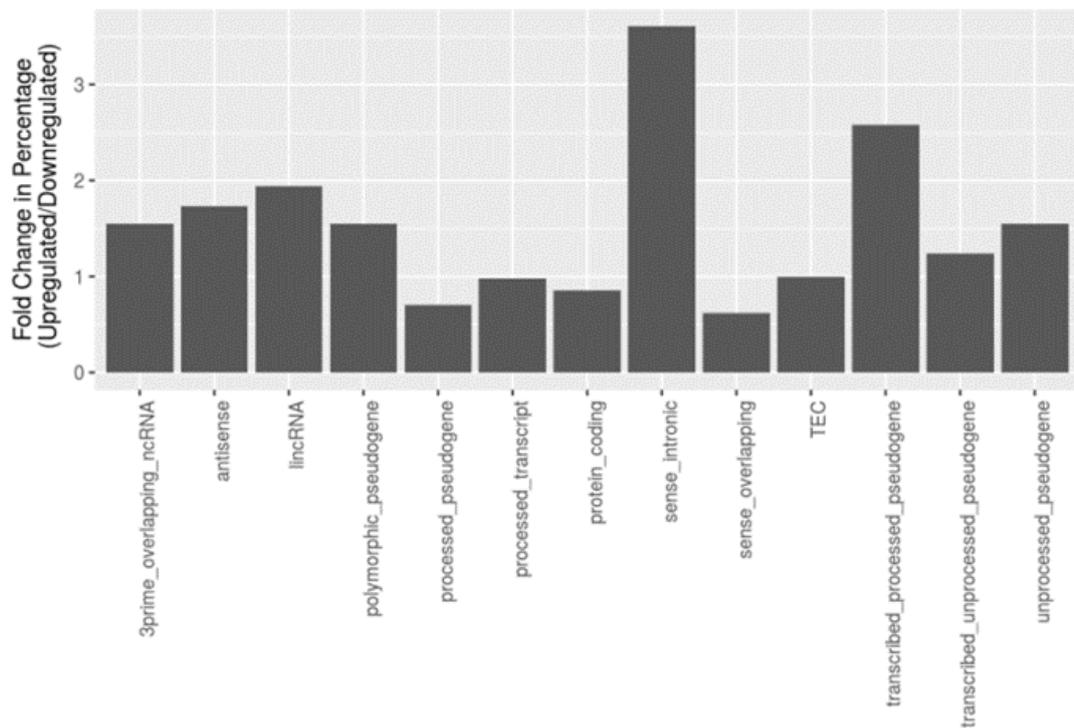


Figure 5.3.3c Many non-coding RNA biotypes make up a larger proportion of the upregulated genes than the downregulated genes in the nucleoplasmic fraction, in contrast to the chromatin fraction

The percentages from 5.3.3b represented as the percentage of the biotype in the upregulated set of genes compared to the downregulated set of genes. A fold change of greater than 1 indicates more of the biotype is upregulated where as a fold change of less than one indicates more of the biotype is downregulated.

Cytoplasm

1503 genes are differentially expressed in the cytoplasmic fraction following hnRNP U depletion, with 60.4% of them being downregulated (Figure 5.3.4a). This is an almost identical proportion of downregulated genes as in the nucleoplasm (Figure 5.3.3a). The percentage of each biotype is also similar to the nucleoplasm (Figure 5.3.4b), but there is a smaller enrichment of many non-coding RNA biotypes in the upregulated over the downregulated genes (Figure 5.3.4c). This suggests that there is less of an upregulation of non-coding RNAs in the cytoplasm compared to the nucleoplasm. Perhaps this is due to less non-coding RNA being present in the cytoplasm as it is degraded before much export can occur, even if the RNA is mislocalised.

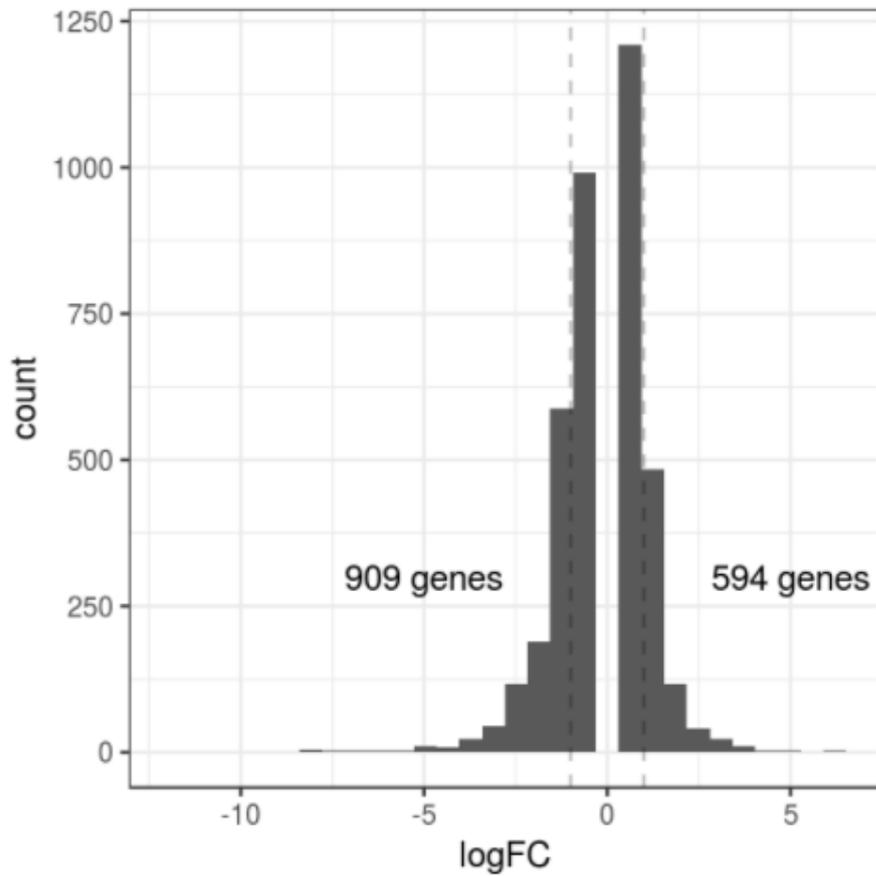


Figure 5.3.4a hnRNP U depletion results in around 1500 differentially expressed genes in the cytoplasm

Following RNA-seq, the expression of genes in the hnRNP U knockdown cytoplasmic samples were compared to the control cytoplasmic samples. The figure is annotated as in Figure 5.3.2a and the same thresholds used.

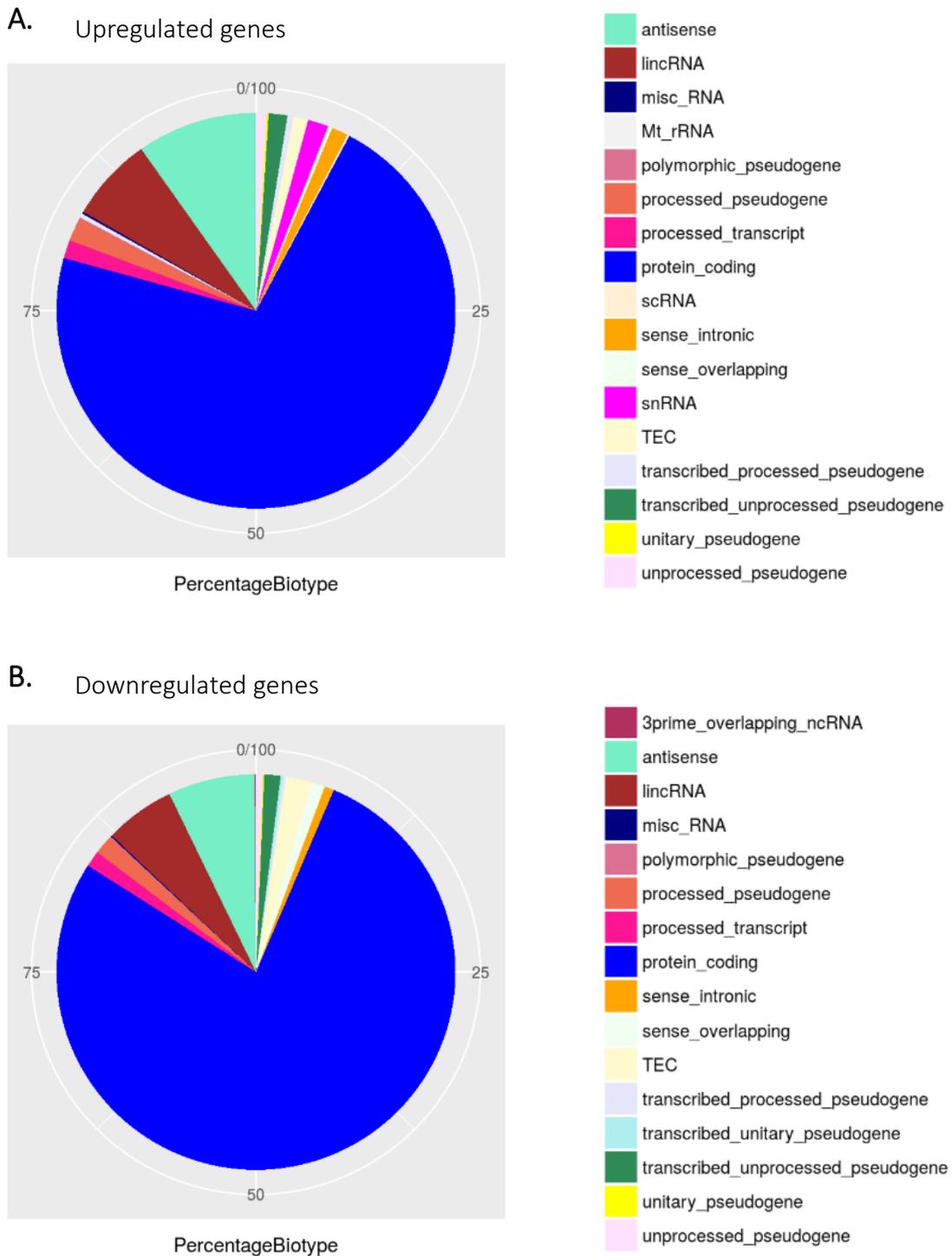


Figure 5.3.4b Many non-coding RNA biotypes make up a slightly larger proportion of the upregulated genes than the downregulated genes in the cytoplasmic fraction, similarly to the nucleoplasmic fraction

The percentage of the significantly (A) upregulated and (B) downregulated genes in 5.3.4a shown as RNA biotypes.

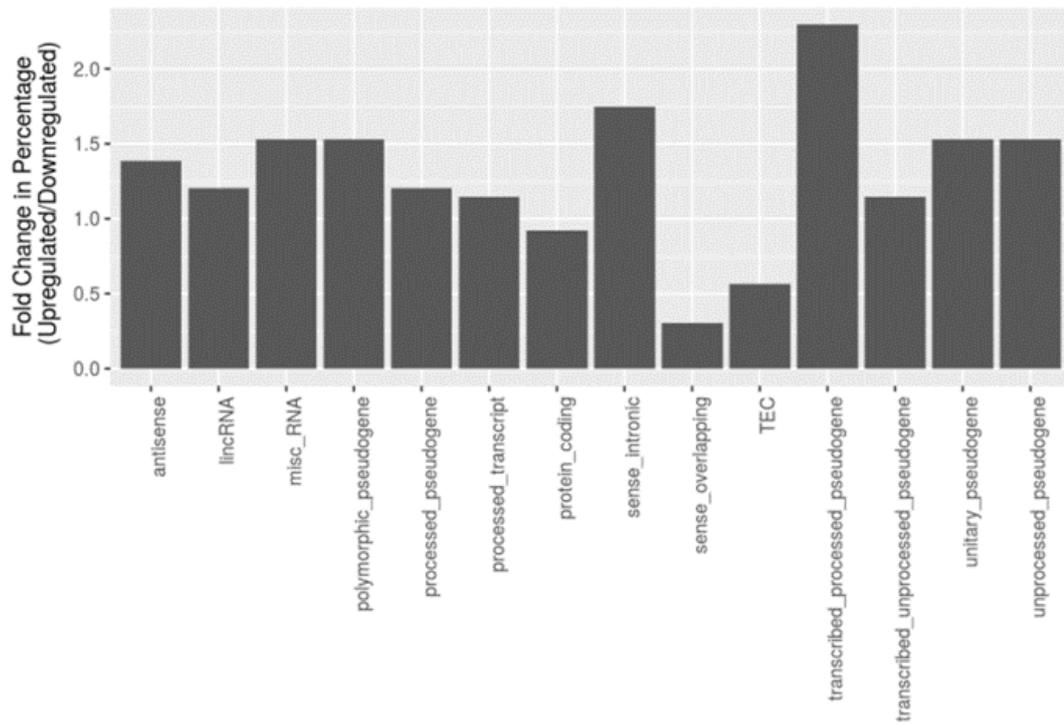


Figure 5.3.4c Many non-coding RNA biotypes make up a slightly larger proportion of the upregulated genes than the downregulated genes in the cytoplasmic fraction, similarly to the nucleoplasmic fraction

The percentages from 5.3.4b represented as the percentage of the biotype in the upregulated set of genes compared to the downregulated set of genes. A fold change of greater than 1 indicates more of the biotype is upregulated where as a fold change of less than one indicates more of the biotype is downregulated.

Chromatin vs Nucleoplasm and Cytoplasm

The genes that were differentially expressed in the hnRNP U knockdown compared to the control between two fractions was also assessed, inter-fraction comparisons, to give an indication of the relative amounts of an RNA between two fractions.

The differential expression analysis of RNAs in hnRNP U knockdown cells compared to control between the chromatin and nucleoplasmic fractions revealed 215 downregulated genes and 78 upregulated genes (Figure 5.3.5a). This means there are 215 genes that have a lower expression level in the chromatin than in the nucleoplasm following hnRNP U depletion when compared to the control expression levels. The 78 upregulated genes imply that 78 genes have a higher expression level in the nucleoplasm than in the chromatin following hnRNP U depletion when compared to the control expression levels. The biotype percentages in the down and upregulated genes are very similar to that of the intra-fraction chromatin comparison (data not shown). On the surface, this suggests that non-coding RNA is being lost from the chromatin to the nucleoplasm, supporting the mislocalisation theory.

Gene ontology analysis of the downregulated genes in this comparison show that snRNA and lincRNA biotypes are significantly affected; the biotype is downregulated more than would be expected by chance. For example, 9 out of the 29 snRNAs sequenced and 24 out of the 782 lincRNAs sequenced are significantly downregulated in the chromatin compared to the nucleoplasm following hnRNP U loss. However, with inter-fraction comparisons, differential downregulation between two compartments only represents a ratio of change between the two fractions. A downregulation between the chromatin and the nucleoplasm could be due to the gene being downregulated in the chromatin only, being upregulated in the nucleoplasm only, being downregulated in both fractions but going down more in the chromatin or a mix of all situations. Therefore, it does not actually tell us if the RNA is mislocalising from the chromatin to the nucleoplasm after hnRNP U depletion.

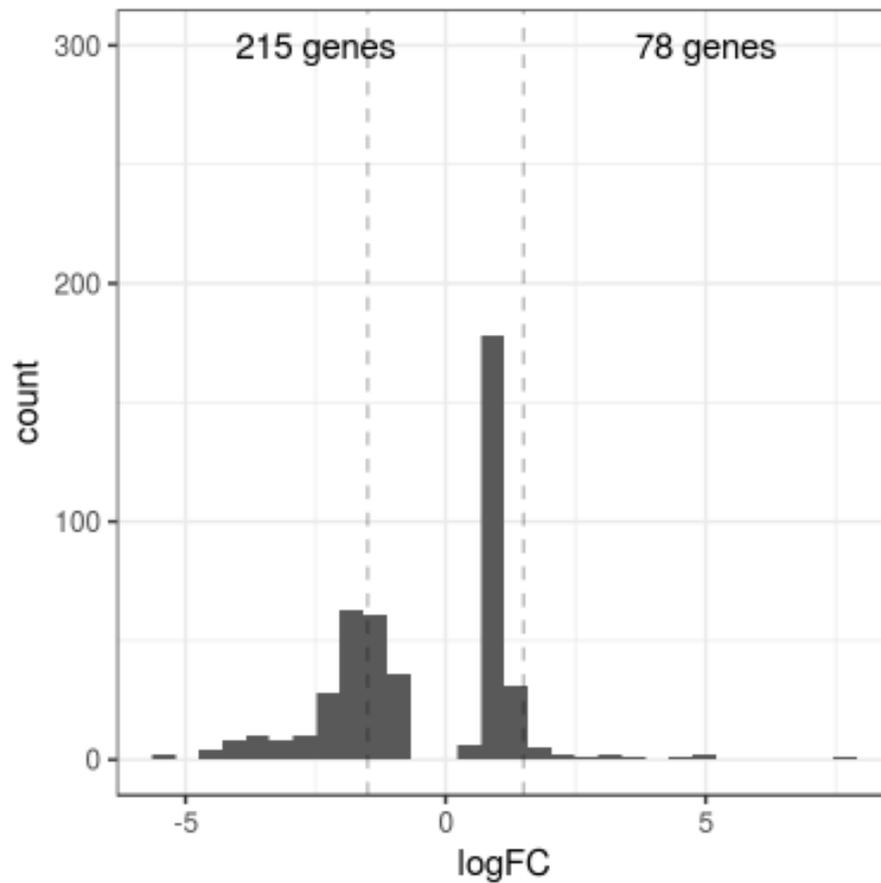


Figure 5.3.5a hnRNP U depletion results in 293 differentially expressed genes between the chromatin and nucleoplasmic fractions

A graph showing the expression of genes in the hnRNP U knockdown samples compared to the control in the chromatin fraction compared to the nucleoplasmic fraction. The figure is annotated as in Figure 5.3.2a and the same thresholds used.

To establish what is happening to this set of downregulated genes between the chromatin and nucleoplasmic fractions, the LFCs from the intra-fraction comparisons between control and hnRNP U knockdown can be used to establish their relative levels in each fraction and generate a heat map. Only non-coding biotype genes were used in this analysis. Using non-coding genes that are differentially downregulated between the chromatin and nucleoplasm, the genes that are downregulated in the chromatin (blue) are generally maintained (white) or upregulated (red) in the nucleoplasm and cytoplasm (Figure 5.3.5b). This denotes a change in localisation from the chromatin to the nucleoplasmic and cytoplasmic fractions when hnRNP U is absent from the cell.

Over 1000 genes are differentially expressed in the comparison between the chromatin and cytoplasmic fractions and similar results to the chromatin vs nucleoplasm comparison are seen on the heat map (Figure 5.3.6a-b). The gene ontology analysis shows that snRNA and lincRNA biotypes are significantly downregulated in the chromatin compared to the cytoplasm. Similarly to the chromatin and nucleoplasm comparison, the downregulated genes in the chromatin are generally upregulated or maintained in the nucleoplasm and cytoplasm (Figure 5.3.6b). There are only 4 differentially expressed ncRNA genes between the nucleoplasm and cytoplasm (Figure 5.3.7), indicating that the expression changes occur between the chromatin and either of the two other fractions.

Therefore, hnRNP U has a general role in localising non-coding RNA to the chromatin as its depletion results in a general movement of ncRNA from the chromatin to the nucleoplasm and the cytoplasm.

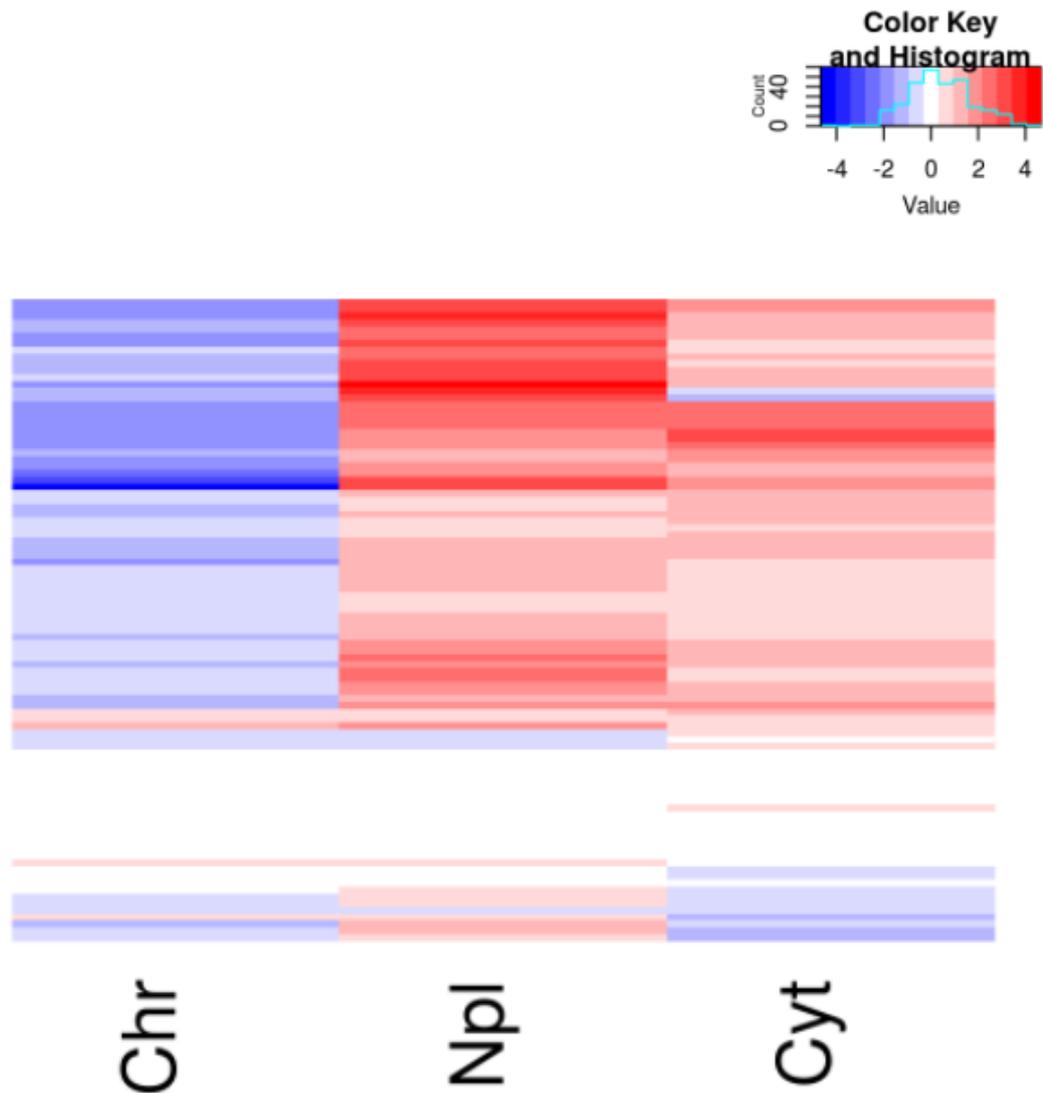


Figure 5.3.5b Downregulated genes between the chromatin and the nucleoplasmic fractions after hnRNP U depletion are generally upregulated in the nucleoplasm and cytoplasm

The LFCs of the downregulated genes from 5.3.5a with non-coding biotypes were used to generate the heat map shown. The LFCs are the values from the hnRNP U knockdown compared to control in each fraction independently; the intra-fraction comparisons. A colour key is shown, where value represents the LFC.

Downregulation (LFC of -1 or less) is demonstrated by the blue spectrum and upregulation (LFC of +1 or more) is demonstrated by the red spectrum. Chr, Npl and Cyt denotes chromatin, nucleoplasmic and cytoplasmic fractions respectively.

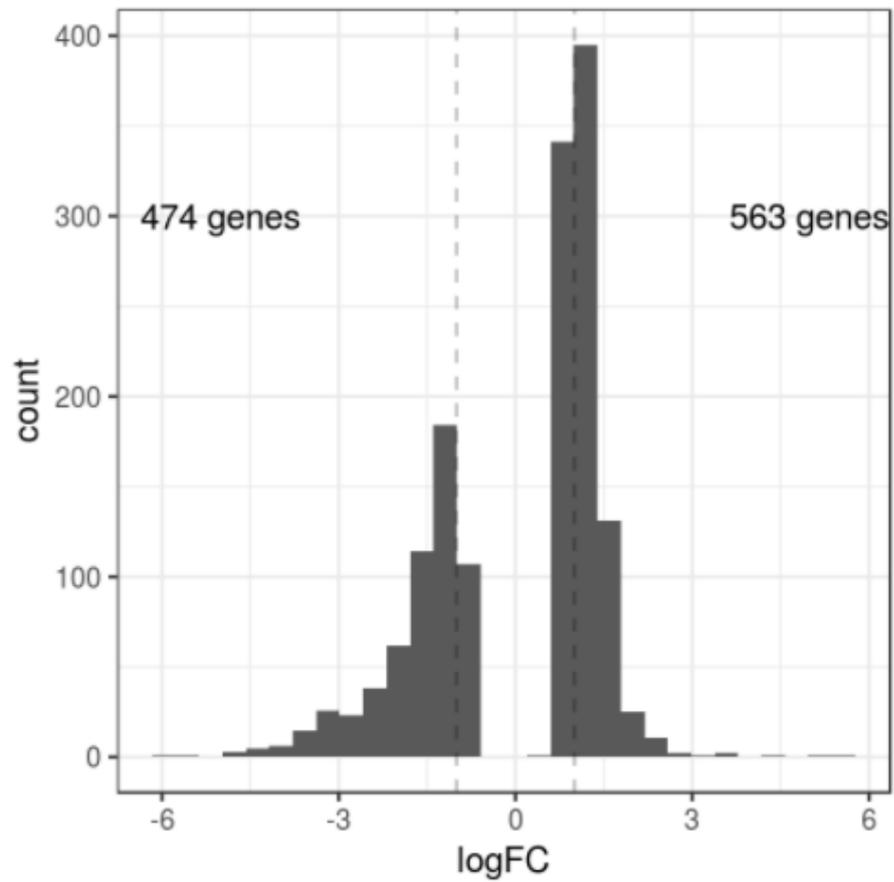


Figure 5.3.6a hnRNP U depletion results in 1037 differentially expressed genes between the chromatin and cytoplasmic fractions

A graph showing the expression of genes in the hnRNP U knockdown samples compared to the control in the chromatin fraction compared to the cytoplasmic fraction. The figure is annotated as in Figure 5.3.2a and the same thresholds used.

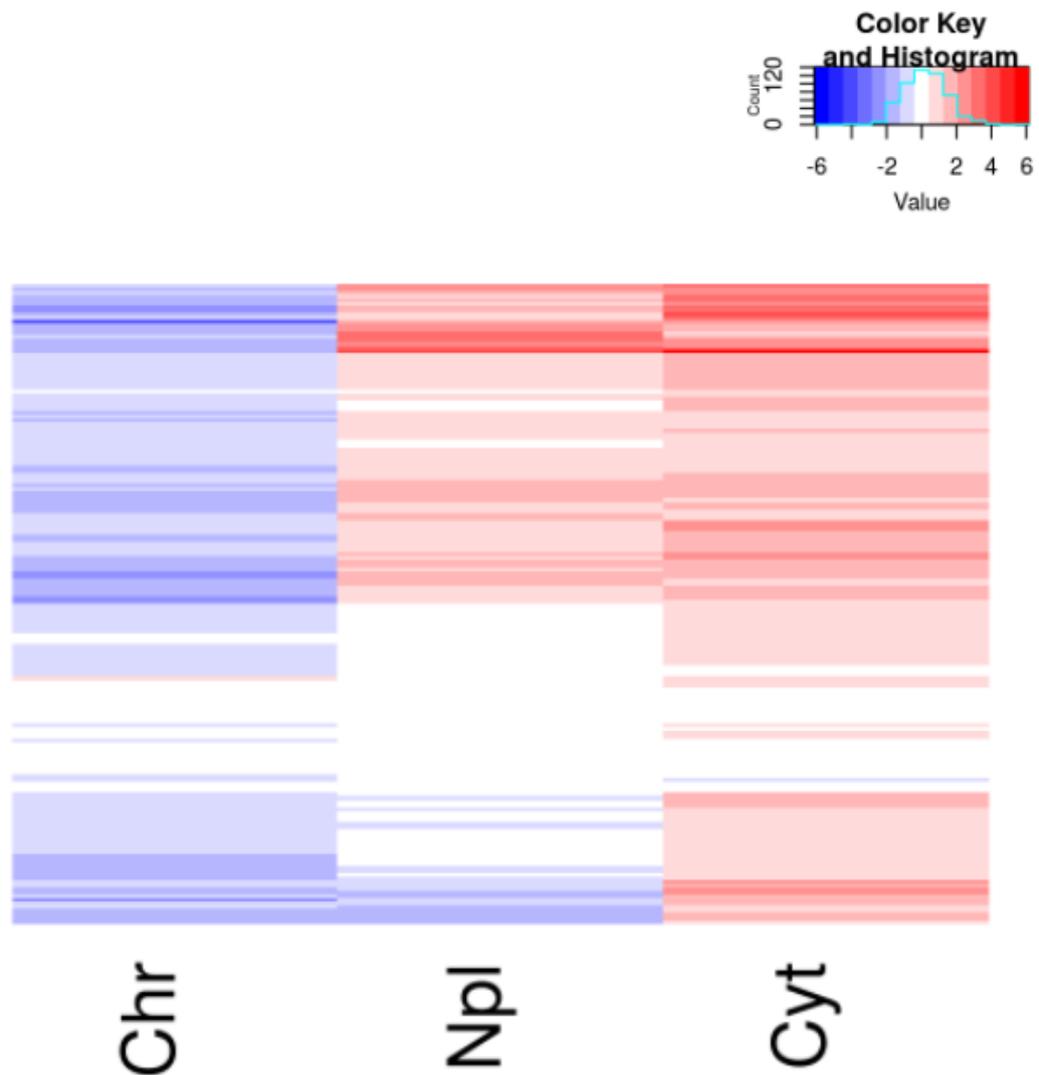


Figure 5.3.6b Downregulated genes between the chromatin and the cytoplasmic fractions after hnRNP U depletion are generally maintained or upregulated in the nucleoplasm and cytoplasm

The LFCs of the downregulated genes from 5.3.6a with non-coding biotypes were used to generate the heat map shown. The heat map values were generated as described in Figure 5.3.5b. The heat map annotations are also as in Figure 5.3.5b.

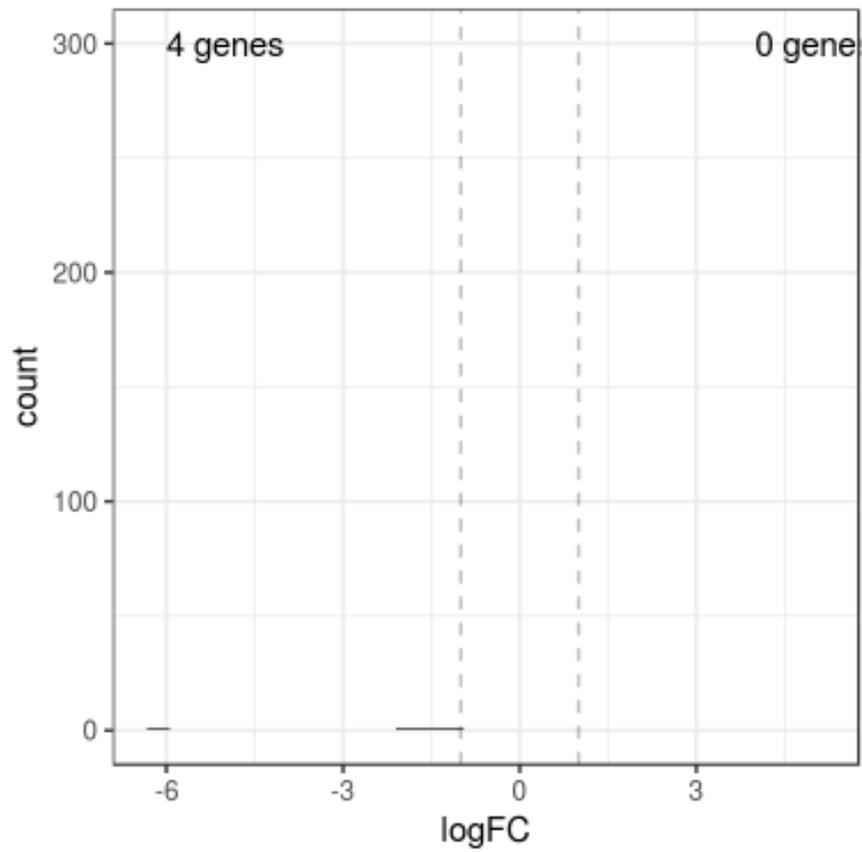


Figure 5.3.7 hnRNP U depletion results in only 4 differentially expressed genes between the nucleoplasmic and cytoplasmic fractions

A graph showing the expression of genes in the hnRNP U knockdown samples compared to the control in the nucleoplasmic fraction compared to the cytoplasmic fraction. The figure is annotated as in Figure 5.3.2a and the same thresholds used.

5.4 Chapter 5 Summary

The function of hnRNP U in RNA localisation was investigated by using an auxin-inducible hnRNP U knockdown cell line to compare the abundance of RNAs in the total cell and in three cellular fractions, chromatin, nucleoplasm and cytoplasm. From total RNA studies, hnRNP U is necessary to maintain the levels of some nuclear RNA classes such as lncRNAs and RNA pol II transcribed snRNAs. mRNAs and pre-mRNAs however did not seem to be generally affected. Following fractionation, the localisation of individual RNAs, such as lncRNAs *FIRRE* and *NEAT1_2*, was assessed after depleting hnRNP U. The RNAs tested were most heavily destabilised in the chromatin fraction, where many of them localise normally. Additionally, after loss from the chromatin, some lncRNAs are degraded in the nucleoplasm as the depletion of the exosome restored their levels. This suggests the RNA is mislocalised from the chromatin to the nucleoplasm where it is degraded.

RNA-seq was carried out to get a comprehensive view of the effect of hnRNP U depletion on RNA localisation, which largely confirmed the results seen with the few RNAs tested by qRT-PCR. Non-coding RNAs are largely downregulated rather than upregulated following hnRNP U loss in the chromatin, in contrast to protein-coding genes. Additionally, it seems that non-coding RNAs that are downregulated in the chromatin are maintained or upregulated in the nucleoplasm and cytoplasm on a genome-wide scale. This demonstrates that hnRNP U retains ncRNAs on the chromatin and its loss causes them to mislocalise to the nucleoplasm and the cytoplasm.

Chapter 6 - Discussion

The previous chapters have revealed more detail about the localisation mechanisms of a variety of non-coding RNAs.

6.1 Export pathways in tRNA export

tRNA is generally understood to be exported to the cytoplasm by XpoT and Xpo5 (Kutay et al. 1998; G J Arts et al. 1998; Calado et al. 2002; Bohnsack et al. 2002). However, recent work has exposed alternative tRNA export pathways such as the Crm1 and Mex67 export pathways in yeast (Wu et al. 2015; Chatterjee, Majumder, et al. 2017). Alyref and Nxf1 proteins are bound to a proportion of tRNAs in human cells, as determined by iCLIP, so the human mRNA export pathway was investigated to determine if it also functions in tRNA export. In human cells that have been depleted of XpoT, Xpo5 and mRNA export proteins individually and in combination, no significant block in tRNA export was witnessed by Northern blot following cellular fractionation or by FISH (Chapters 3.2 and 3.3). This was surprising as the depletion of both canonical export proteins, XpoT and Xpo5, had a minimal effect on tRNA export. Studies on the human pathway have generally been carried out using *Xenopus* oocytes with injection of the human proteins, so direct evidence of XpoT and Xpo5 activity in human tRNA export is not as clear cut as in yeast.

Perhaps there are other major proteins involved that have yet to be discovered. A genome-wide screen for proteins involved in human tRNA export seems necessary to confirm Xpot and Xpo5 as major exporters and more likely, uncover any missing factors. Additionally, a human inducible knockout cell line for XpoT and Xpo5 may be required to see any block in tRNA export as perhaps the knockdown level of the proteins is not enough to witness an effect (Figure 3.2.1). This could be done using the auxin-inducible degron method described in Chapter 4. Additionally, the type of assay may need to be altered. Kinetically tracking a reporter tRNA rather than trying to assess a change in localisation of steady-state tRNA may be much more effective. tRNA may be too stable and abundant to witness any steady-state changes so following a reporter tRNA may be required to establish the role of the mRNA export pathway in the tRNA lifecycle.

The human mRNA export proteins may be required for correct processing of tRNA in the nucleus rather than in export, similarly to most mRNA processing steps requiring Alyref (Heath et al. 2016). A method akin to that used in Wu et al, 2015 could be employed to look at the levels of intron-containing and precursor tRNAs by Northern blot using specific probes following the knockdown of TREX components and Nxf1. Depletion of the mRNA export pathway may result in accumulation of intron-containing tRNA species, suggesting a role in processing. However, this could be difficult with human RNA samples as only ~6% of tRNAs contain introns so they may be hard to detect (Phizicky & Hopper 2010). However, as tRNA is transcribed by RNA pol III rather than pol II like mRNAs, it is less likely that mRNA export proteins are involved in their processing.

6.2 Intra-nuclear localisation of non-coding RNA

Export pathways have also been linked to intra-nuclear localisation of RNA, such as the Crm1 pathway in snoRNA localisation to Cajal bodies (Boulon et al. 2004).

Therefore, the involvement of the mRNA export pathway in intra-nuclear RNA localisation was investigated. *scaRNA2* and *XIST* RNA were selected due to the large peaks of Alyref, Chtop and Nxf1 binding from iCLIP data (Figures 3.5 and 3.6.1).

Neither *scaRNA2* localisation to the Cajal body nor *XIST* localisation to the X chromosome was impeded by the depletion of TREX and Nxf1 in human cells (Figures 3.5 and 3.6.2). Perhaps in both these cases, you would need to look for a cumulative effect to assess a potential role of the mRNA export pathway in localisation. The depletion of TREX components and Nxf1 alongside other localisation proteins for *scaRNA2* and *XIST*, such as WRAP53 and hnRNP U respectively, may uncover a role in intra-nuclear export that is not evident by depletion of the mRNA export pathway alone (Tycowski et al. 2009; Hasegawa et al. 2010).

Additionally, the abundance of *XIST* RNA is significantly reduced following the depletion of TREX proteins but is stable after Nxf1 depletion (Figure 3.7). Therefore, TREX proteins may play a role in processing of non-coding RNA rather than in localisation and their loss could expose the RNA to degradation mechanisms. This could be assessed by investigating any splicing or end-processing defects by qRT-PCR or Northern blot. Ideally, the levels of other lncRNAs would be assessed as well as *scaRNA2* and other non-coding RNA biotypes following TREX and Nxf1 depletion. The function of mRNA export proteins in the stability and processing of lncRNAs such as *XIST* seems more likely than in tRNA processing as many are transcribed by RNA polymerase II.

6.3 hnRNP U and non-coding RNA localisation

As well as annotated export pathways, individual cellular proteins act to localise RNAs in the human cell. hnRNP U has a variety of roles in the cell, including the localisation of some nuclear RNAs to the chromatin (Hasegawa et al. 2010; Hacısuleyman et al. 2014; Hall et al. 2014). To assess this phenomenon further, an auxin-inducible degron cell line was generated to deplete hnRNP U from HCT116 cells in a controlled manner (Chapter 4)(Nishimura et al. 2009; Natsume et al. 2016; Eaton et al. 2018). From qRT-PCR of specific transcripts, many non-coding RNAs are downregulated following hnRNP U depletion and they are mostly depleted in the chromatin fraction. A genome-wide RNA sequencing experiment revealed that this is the case for many lncRNAs and other non-coding RNA biotypes and that these transcripts are generally maintained or upregulated in the nucleoplasmic and cytoplasmic fractions. This suggests that a large scale re-localisation of chromatin-associated RNAs to the nucleoplasm and cytoplasm is taking place. qRT-PCR data also revealed that specific lncRNAs such as *FIRRE* are degraded by the exosome in the nucleoplasm following hnRNP U depletion, indicating that some RNAs released from the chromatin are degraded in the other cellular compartments, in alignment with published data (Schlackow et al. 2017).

RNA sequencing of cells depleted of hnRNP U and the exosome combined would be able to clarify whether this is a generic mechanism or whether only specific transcripts are degraded in this manner. However, other degradation pathways may be acting to degrade RNA that has been released from the chromatin, such as NMD in the cytoplasm. Despite non-coding RNAs not being translated by definition, some evidence suggests that they may undergo an initial round of translation and therefore could be susceptible to NMD (Carlevaro-Fita et al. 2016). Treatment of hnRNP U-depleted cells with cycloheximide would determine if individual RNAs mislocalised by the loss of hnRNP U are degraded by NMD. qRT-PCR of specific transcripts or a genome-wide approach would deal with this question. As the qRT-PCR data has highlighted a link between hnRNP U and the exosome, it would be interesting to establish whether there is a physical association between hnRNP U and exosome components through a co-immunoprecipitation.

Alternatively to mislocalisation and degradation of RNA species, hnRNP U- depletion could be affecting the transcription of chromatin-associated RNAs. Native elongating transcript sequencing (NET-seq) will be used to assess the effect of hnRNP U- depletion on RNA pol II transcription. The data can be used to establish whether the transcription of RNAs that have significantly decreased in abundance in this data set and has been effected. As hnRNP U has previously been implicated in transcription regulation, it is likely that some effect on transcription of non-coding RNAs will be seen. However, the lncRNAs and the non-coding RNA biotypes that are most greatly effected in terms of mislocalisation, are unlikely to be more greatly affected by transcriptional defects than other RNA pol II transcripts. Both the RNA-seq and NET-seq experiments will be able to give a complete picture of the role of hnRNP U in non-coding RNA biogenesis and localisation.

Furthermore, as there is a general mislocalisation of chromatin-associated RNA to other cellular compartments, perhaps the mRNA export pathway is exporting the non-coding RNA to the cytoplasm in the absence of hnRNP U. As there are dense iCLIP tags of mRNA export proteins on lncRNAs and other non-coding RNA biotypes, there may be a default system of export via the mRNA export pathway that is normally prevented by the association of hnRNP U with the RNA. A similar phenomenon has been recently established in terms of competition for degradation or export of RNAs. ZFC3H1, as part of the PAXT exosome-targeting complex, seemingly competes with Alyref for association with RNAs to maintain them in the nucleus or export them to the cytoplasm respectively (Silla et al. 2018). Interestingly, even RNAs known to function in the nucleus such as SNGH19 were exported to the cytoplasm in the absence of ZFC3H1, supporting the idea of a default export pathway that can be prevented by retention factors (Silla et al. 2018). Therefore, co-depletion of hnRNP U and TREX proteins or Nxf1 could establish whether non-coding RNA mislocalisation is dependent on the mRNA export pathway. A co-immunoprecipitation experiment demonstrated the interaction between hnRNP U and Alyref but not Nxf1 (Figure 4.3.3.2). This suggests that TREX may associate with hnRNP U that is bound to RNA but Nxf1 is prevented from doing so. The presence of hnRNP U may inhibit the handover of RNA from TREX to Nxf1 for export. Additionally,

the hnRNP U eCLIP data set generated in Xiao et al, 2012 showed an enrichment of hnRNP U on pre-mRNA introns (Xiao et al. 2012). Perhaps the presence of hnRNP U on pre-mRNA introns is preventing them being exported by Nxf1 before complete processing. An interesting experiment would be to deplete the cells of hnRNP U and see if more RNA is associated with Nxf1. Associations between hnRNP U and other TREX components such as the Tho complex and Chtop should be established as well, as Alyref may be associated with hnRNP U independently of TREX.

If hnRNP U creates a general retention mechanism for chromatin-associated or nuclear RNAs, perhaps specific sequences in RNA are recognised by hnRNP U individually or in a protein complex. The RNA-repeating domain (RRD) of lncRNA *FIRRE* is bound by hnRNP U and is necessary to maintain *FIRRE* in the nucleus in both mice and human cells (Hacisuleyman et al. 2014; Hacisuleyman et al. 2016). Additionally, non-nuclear transcripts are retained in the nucleus when an RRD has been inserted and are mislocalised to the cytoplasm following hnRNP U depletion (Hacisuleyman et al. 2016). hnRNP U may localise transcripts to the nucleus through recognition of a sequence similar to the RRD of *FIRRE*. As lncRNAs are enriched in local repeats (LRs) in comparison to mRNAs (Hacisuleyman et al. 2016), this could provide a generic mechanism for nuclear localisation of ncRNA by hnRNP U.

Nuclear-restricted lncRNAs and some mRNAs are enriched in C or C/T-rich motifs as well as repetitive Alu elements (Lubelsky & Ulitsky 2018; Shukla et al. 2018). The C-rich motifs can be bound by hnRNP K to enhance nuclear enrichment of these RNAs and perhaps hnRNP U acts in a similar manner to maintain RNA in the nucleus and on the chromatin (Lubelsky & Ulitsky 2018).

The RNA-seq data set could be analysed further to look at specific transcripts that are mislocalised within the RNA biotype. Looking at specific transcripts that are affected by the loss of hnRNP U could shed light on the functions of lncRNAs that are currently unknown. For example, many lncRNAs may be involved in chromatin architecture or be scaffolds in cellular processes which could be further elucidated with this data. The role that hnRNP U plays in chromatin organisation could also be developed. How hnRNP U interacts with the chromatin, either through RNA or

directly, may become clearer through the combination of published eCLIP data and this data set (Xiao et al. 2012). The general decrease in the levels of RNA polII transcribed snRNAs is certainly an area to focus on with this data. This generic effect may reveal a function of hnRNP U in the maturation or processing or localisation of all polII snRNAs.

The RNA-seq data has so far only been used to look at expression and mislocalisation of RNA biotypes following depletion of hnRNP U. This data set can also be used to further characterise the role of hnRNP U in splicing (Huelga et al. 2012; Xiao et al. 2012). Additionally, any influence of hnRNP U in 3' poly (A) site choice could be ascertained from the RNA-seq data set. From the ENCODE hnRNP U RNAi data analysed by Dr Ian Sudbery (Chapter 4.3.4), a number of antisense RNA transcripts are upregulated in the hnRNP U RNAi conditions. A few of these antisense transcripts were analysed by qRT-PCR, alongside the corresponding sense transcript, using samples from the hnRNP U-flAID cell line treated with auxin (Figure 6.3). All three transcripts tested demonstrated significant upregulation of the antisense transcript following hnRNP U depletion. hnRNP U could consequently be involved with determining promoter directionality (Grzechnik et al. 2014), with hnRNP U promoting sense transcription in a wild type context. The corresponding sense transcript was analysed to determine if it is general upregulation around the promoter or specifically in the antisense direction.

TRIM52 and LBX2 sense transcripts showed no increase and a decrease in RNA levels respectively, suggesting a directionality effect. The third transcript tested, SLC2A1, showed an upregulation of sense and antisense transcripts, suggesting upregulation of the promoter in general. In contrast, from the RNA-seq data the antisense RNA biotype was borderline significant in the list of downregulated genes in the chromatin compared to the nucleoplasm following hnRNP U depletion. This suggests that hnRNP U may play a role in maintaining antisense transcripts. The promoter directionality phenomenon may be interesting to investigate using this RNA-seq data. Interestingly, promoter directionality has been linked to snRNP activity (Almada et al. 2013).

Sense and antisense RNAs

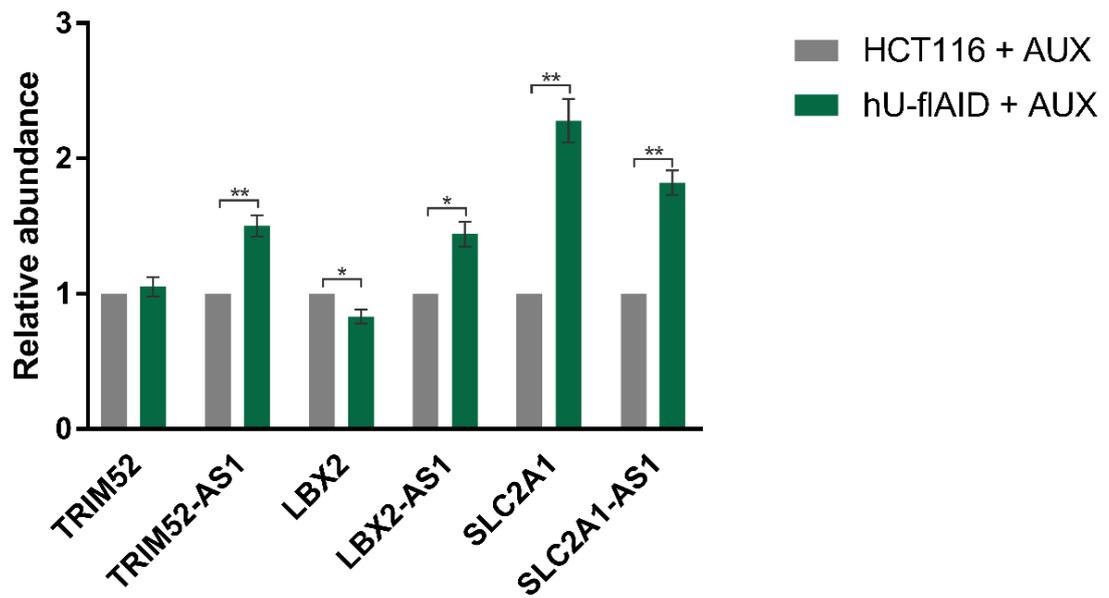


Figure 6.3 hnRNP U may play a role in promoter directionality

qRT-PCR was carried out on four RNA samples per condition. HCT116 and hU-flAID cells were treated with auxin for 24 hours. The mean abundance is shown relative to the mean abundance in the control samples (HCT116 + AUX). All values are normalised to the abundance of U6 snRNA in the individual samples. Error bars represent the SEM. AS1 signifies antisense transcript 1, and the absence of AS1 indicates a sense transcript. A t-test was carried out as described in Chapter 2 to determine the statistical significance.

As discussed in Chapter 1.5.1, hnRNP U influences the biogenesis of snRNPs so a role in promoter directionality could be a knock-on effect of hnRNP U function in snRNP biogenesis.

Despite the effectiveness of depleting hnRNP U using the auxin-inducible degron system (Chapter 4), the typical levels of hnRNP U in the cell line is less than in wild type HCT116 cells. This could have an effect on the cells before the addition of auxin. The process of selecting for cells that have all HNRNPU genes tagged may also select for beneficial adaptations that help to cope with the decreased amount of hnRNP U in the cell line. The hnRNP U-flAID cell line may have altered cell proliferation and metabolism rates due to the indirect effects of the reduced levels of hnRNP U, as seen in the colony formation assay (Chapter 4.3.3). Ideally, the cell line would have nearer wildtype levels of hnRNP U before auxin addition, which may not be possible with this system.

6.4 Conclusion

The complex localisation of the variety of RNAs that exist in a human cell requires an overlap of pathways for efficiency. iCLIP data demonstrates that components of the TREX complex, Alyref and Chtop and the mRNA export receptor Nxf1 are bound to a variety of non-coding RNAs as well as mRNAs. Despite iCLIP binding peaks on tRNAs, the mRNA export pathway does not function significantly in steady-state tRNA nuclear export in human cells although the orthologous yeast Nxf1 protein may play a role (Wu et al. 2015; Chatterjee, Majumder, et al. 2017). Furthermore, depletion of the canonical tRNA export pathways does not significantly affect tRNA localisation as determined by Northern blotting and FISH. These results may highlight differences between the tRNA export pathways in eukaryotes and the human pathway in particular requires further study. However, this experiment may require the use of a reporter tRNA rather than looking at the localisation of tRNAs in steady-state.

TREX proteins and Nxf1 are bound to nuclear-restricted non-coding RNAs, such as *scaRNA2* and *XIST*, but their intra-nuclear localisation is not reliant on the mRNA export pathway. However, the stability of some non-coding RNAs may be dependent on the TREX complex. hnRNP U is required to maintain the levels of many non-coding RNAs in the human cell, potentially due to its function in RNA localisation. hnRNP U is known to localise specific RNAs to the chromatin but data in this thesis establishes a more global requirement of hnRNP U for non-coding RNA localisation to the chromatin. A mechanism of hnRNP U retention and competition with the mRNA export pathway would develop this story further to establish a universal nuclear retention mechanism coordinated by hnRNP U. The topic of non-coding RNA localisation demands more research to fully understand the mechanisms of this process that is vital to cellular homeostasis, which when disrupted can lead to disease.

References

- Allmang, C. et al., 1999. Functions of the exosome in rRNA, snoRNA and snRNA synthesis. *The EMBO Journal*, 18(19), pp.5399–5410.
- Almada, A.E. et al., 2013. Promoter directionality is controlled by U1 snRNP and polyadenylation signals. *Nature*, 499(7458), pp.360–363.
- Andersen, P.R. et al., 2013. The human cap-binding complex is functionally connected to the nuclear RNA exosome. *Nature Structural and Molecular Biology*, 20(12), pp.1367–76.
- Anderson, J.S.J. & Parker, R., 1998. The 3' to 5' degradation of yeast mRNAs is a general mechanism for mRNA turnover that requires the SK12 DEVH box protein and 3' to 5' exonucleases of the exosome complex. *The EMBO Journal*, 17(5), pp.1497–1506.
- Araki, Y. et al., 2001. Ski7p G protein interacts with the exosome and the ski complex for 3'-to-5' mRNA decay in yeast. *The EMBO Journal*, 20(17), pp.4684–4693.
- Arts, G.J. et al., 1998. The role of exportin-t in selective nuclear export of mature tRNAs. *The EMBO Journal*, 17(24), pp.7430–7441.
- Arts, G.J., Fornerod, M. & Mattaj, I.W., 1998. Identification of a nuclear export receptor for tRNA. *Current Biology*, 8(6), pp.305–314.
- Bachi, A. et al., 2000. The C-terminal domain of TAP interacts with the nuclear pore complex and promotes export of specific CTE-bearing RNA substrates. *RNA*, 6(1), pp.136–158.
- Banfai, B. et al., 2012. Long noncoding RNAs are rarely translated in two human cell lines. *Genome Research*, 22(9), pp.1646–1657.
- Barr, M.L. & Carr, D.H., 1962. Correlations between sex chromatin and sex chromosomes. *Acta Cytol*, 6, pp.34–45.
- Barral, P.M. et al., 2005. The interaction of the hnRNP family member E1B-AP5 with p53. *FEBS Letters*, 579(13), pp.2752–2758.

- Bayliss, R. et al., 2002. GLFG and FxFG nucleoporins bind to overlapping sites on importin- β . *Journal of Biological Chemistry*, 277(52), pp.50597–50606.
- Beck, M. & Hurt, E., 2017. The nuclear pore complex: Understanding its function through structural insight. *Nature Reviews Molecular Cell Biology*, 18(2), pp.73–89.
- Behm-Ansmant, I. & Izaurralde, E., 2006. Quality control of gene expression: A stepwise assembly pathway for the surveillance complex that triggers nonsense-mediated mRNA decay. *Genes & Development*, 20(4), pp.391–398.
- Benchling, 2018. Benchling · Better tools, faster research. Available at: <https://benchling.com/>.
- Berglund, F.M. & Clarke, P.R., 2009. hnRNP-U is a specific DNA-dependent protein kinase substrate phosphorylated in response to DNA double-strand breaks. *Biochemical and Biophysical Research Communications*, 381(1), pp.59–64.
- Blackford, A.N. et al., 2008. A Role for E1B-AP5 in ATR Signaling Pathways during Adenovirus Infection. *Journal of Virology*, 82(15), pp.7640–7652.
- Blomen, V.A. et al., 2015. Gene essentiality and synthetic lethality in haploid human cells. *Science*, 350(6264), pp.1092–1096.
- Bohnsack, M.T. et al., 2002. Exp5 exports eEF1A via tRNA from nuclei and synergizes with other transport pathways to confine translation to the cytoplasm. *The EMBO Journal*, 21(22), pp.6205–6215.
- Bohnsack, M.T., Czaplinski, K. & Görlich, D., 2004. Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *RNA*, 10(2), pp.185–191.
- Boisvert, F.-M. et al., 2012. A Quantitative Spatial Proteomics Analysis of Proteome Turnover in Human Cells. *Molecular & Cellular Proteomics*, 11(3), p.M111.011429.
- Bonneau, F. et al., 2009. The Yeast Exosome Functions as a Macromolecular Cage to

- Channel RNA Substrates for Degradation. *Cell*, 139(3), pp.547–559.
- Boulon, S. et al., 2004. PHAX and CRM1 are required sequentially to transport U3 snoRNA to nucleoli. *Molecular Cell*, 16(5), pp.777–787.
- Bramswig, N.C. et al., 2017. Heterozygous HNRNPU variants cause early onset epilepsy and severe intellectual disability. *Human Genetics*, 136(7), pp.821–834.
- Brennan, C.M., Gallouzi, I.E. & Steitz, J.A., 2000. Protein ligands to HuR modulate its interaction with target mRNAs in vivo. *Journal of Cell Biology*, 151(1), pp.1–13.
- Britton, S. et al., 2009. Cell nonhomologous end joining capacity controls SAF-A phosphorylation by DNA-PK in response to DNA double-strand breaks inducers. *Cell Cycle*, 8(22), pp.3717–3722.
- Britton, S. et al., 2014. DNA damage triggers SAF-A and RNA biogenesis factors exclusion from chromatin coupled to R-loops removal. *Nucleic Acids Research*, 42(14), pp.9047–9062.
- Brown, C.J. et al., 1992. The human XIST gene: Analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. *Cell*, 71(3), pp.527–542.
- Brownawell, A.M. & Macara, I.G., 2002. Exportin-5, a novel karyopherin, mediates nuclear export of double-stranded RNA binding proteins. *Journal of Cell Biology*, 156(1), pp.53–64.
- Cabili, M.N. et al., 2011. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes & Development*, 25(18), pp.1915–1927.
- Cai, X., Hagedorn, C.H. & Cullen, B.R., 2004. Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *RNA*, 10(12), pp.1957–1966.
- Calado, A. et al., 2002. Exportin-5-mediated nuclear export of eukaryotic elongation factor 1A and tRNA. *The EMBO Journal*, 21(22), pp.6216–6224.

- Caliebe, A. et al., 2010. Four patients with speech delay, seizures and variable corpus callosum thickness sharing a 0.440 Mb deletion in region 1q44 containing the HNRPU gene. *European Journal of Medical Genetics*, 53(4), pp.179–185.
- Carlevaro-Fita, J. et al., 2016. Cytoplasmic long noncoding RNAs are frequently bound to and degraded at ribosomes in human cells. *RNA*, 22(6), pp.867–82.
- Carty, S.M. & Greenleaf, A.L., 2002. Hyperphosphorylated C-terminal repeat domain-associating proteins in the nuclear proteome link transcription to DNA/chromatin modification and RNA processing. *Molecular & Cellular Proteomics*, 1(8), pp.598–610.
- Cech, T.R. & Steitz, J.A., 2014. The noncoding RNA revolution - Trashing old rules to forge new ones. *Cell*, 157(1), pp.77–94.
- Cerase, A. et al., 2015. Xist localization and function: New insights from multiple levels. *Genome Biology*, 16(1), pp.1–12.
- Chafe, S.C. et al., 2011. Nutrient stress does not cause retrograde transport of cytoplasmic tRNA to the nucleus in evolutionarily diverse organisms. *Molecular Biology of the Cell*, 22(7), pp.1091–1103.
- Chan, P.P. & Lowe, T.M., 2009. GtRNADB: A database of transfer RNA genes detected in genomic sequence. *Nucleic Acids Research*, 37(Database issue), pp.93–97.
- Chang, C.-T. et al., 2013. Chtop is a component of the dynamic TREX mRNA export complex. *The EMBO Journal*, 32(3), pp.473–86.
- Chatterjee, K., Majumder, S., et al., 2017. Sharing the load: Mex67-Mtr2 co-functions with Los1 in primary tRNA nuclear export. *Genes & Development*, 31(21), pp.2186–2198.
- Chatterjee, K., Nostramo, R.T., et al., 2017. tRNA dynamics between the nucleus, cytoplasm and mitochondrial surface: Location, location, location. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, 1861(4), pp.373–386.
- Cheng, H. et al., 2006. Human mRNA export machinery recruited to the 5' end of

- mRNA. *Cell*, 127(7), pp.1389–400.
- Chi, B. et al., 2013. Aly and THO are required for assembly of the human TREX complex and association of TREX components with the spliced mRNA. *Nucleic Acids Research*, 41(2), pp.1294–306.
- Chu, C. et al., 2015. Systematic discovery of Xist RNA binding proteins. *Cell*, 161(2), pp.404–416.
- Clemson, C.M. et al., 2009. An Architectural Role for a Nuclear Non-coding RNA: NEAT1 RNA is Essential for the Structure of Paraspeckles. *Molecular Cell*, 33(6), pp.717–726.
- Clemson, C.M. et al., 1996. XIST RNA paints the inactive X chromosome at interphase: Evidence for a novel RNA involved in nuclear/chromosome structure. *Journal of Cell Biology*, 132(3), pp.259–275.
- Cong, L. et al., 2013. Multiplex Genome Engineering Using CRISPR/VCas Systems. *Science*, 339(6121), pp.819–823.
- Cook, A.G. et al., 2009. Structures of the tRNA export factor in the nuclear and cytosolic states. *Nature*, 461(7260), pp.60–65.
- Cristofari, G. et al., 2007. Human Telomerase RNA Accumulation in Cajal Bodies Facilitates Telomerase Recruitment to Telomeres and Telomere Elongation. *Molecular Cell*, 27(6), pp.882–889.
- Cronshaw, J.M. et al., 2002. Proteomic analysis of the mammalian nuclear pore complex. *Journal of Cell Biology*, 158(5), pp.915–927.
- Daly, R.F. et al., 1977. Structure and Barr body formation of an Xp + chromosome with two inactivation centers. *American Journal of Human Genetics*, 29(1), pp.83–93.
- Depienne, C. et al., 2017. Genetic and phenotypic dissection of 1q43q44 microdeletion syndrome and neurodevelopmental phenotypes associated with mutations in ZBTB18 and HNRNPU. *Human Genetics*, 136(4), pp.463–479.

- Dharmasiri, N. et al., 2003. Auxin Action in a Cell-Free System. *Current Biology*, 13(16), pp.1418–1422.
- Van Dijk, E.L. et al., 2011. XUTs are a class of Xrn1-sensitive antisense regulatory non-coding RNA in yeast. *Nature*, 475(7354), pp.114–119.
- Douglas, P. et al., 2015. Phosphorylation of SAF-A/hnRNP-U serine 59 by polo-like kinase 1 (PLK1) is required for mitosis. *Molecular and Cellular Biology*, 35(15), pp.2699–2713.
- Dreher, K.A. et al., 2006. The Arabidopsis Aux/IAA Protein Family Has Diversified in Degradation and Auxin Responsiveness. *The Plant Cell*, 18(3), pp.699–714.
- Dreyfuss, G., Adam, S. a & Choi, Y.D., 1984. Physical change in cytoplasmic messenger ribonucleoproteins in cells treated with inhibitors of mRNA transcription. *Molecular and Cellular Biology*, 4(3), pp.415–423.
- Eaton, J.D. et al., 2018. Xrn2 accelerates termination by RNA polymerase II, which is underpinned by CPSF73 activity. *Genes & Development*, 32(2), pp.127–139.
- Engreitz, J.M. et al., 2014. RNA-RNA interactions enable specific targeting of noncoding RNAs to nascent pre-mRNAs and chromatin sites. *Cell*, 159(1), pp.188–199.
- Engreitz, J.M., Ollikainen, N. & Guttman, M., 2016. Long non-coding RNAs: spatial amplifiers that control nuclear structure and gene expression. *Nature Reviews - Molecular Cell Biology*, 17(12), pp.756–770.
- Enwerem, I.I. et al., 2015. Cajal body proteins differentially affect the processing of box C/D scaRNPs. *PLOS One*, 10(4), pp.1–21.
- Fackelmayer, F.O. et al., 1994. Nucleic-acid-binding properties of hnRNP-U/SAF-A, a nuclear-matrix protein which binds DNA and RNA in vivo and in vitro. *European Journal of Biochemistry*, 221(2), pp.749–757.
- Falk, S. et al., 2014. The Molecular Architecture of the TRAMP Complex Reveals the Organization and Interplay of Its Two Catalytic Activities. *Molecular Cell*, 55(6),

pp.856–867.

- Fan, H. et al., 2017. The nuclear matrix protein HNRNPU maintains 3D genome architecture globally in mouse hepatocytes. *Genome Research*, 28(2), pp.192–202.
- Fan, J. et al., 2017. Exosome cofactor hMTR4 competes with export adaptor ALYREF to ensure balanced nuclear RNA pools for degradation and export. *The EMBO Journal*, 36(19), pp.2870–2886.
- Faza, M.B. et al., 2012. Role of Mex67-Mtr2 in the Nuclear Export of 40S Pre-Ribosomes. *PLOS Genetics*, 8(8), pp.16–21.
- Fischer, U. et al., 1995. The HIV-1 Rev Activation Domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs. *Cell*, 82(3), pp.475–483.
- Floer, M., Blobel, G. & Rexach, M., 1997. Disassembly of RanGTP karyopherinb complex, an intermediate in nuclear protein import. *Journal of Biological Chemistry*, 272(31), pp.19538–19546.
- Fornerod, M. et al., 1997. CRM1 is an export receptor for leucine-rich nuclear export signals. *Cell*, 90(6), pp.1051–1060.
- Fox, A.H. & Lamond, A.I., 2010. Paraspeckles. *Cold Spring Harbor Perspectives in Biology*, 2(7), pp.a000687–a000687.
- Fribourg, S. et al., 2001. Structural basis for the recognition of a nucleoporin FG repeat by the NTF2-like domain of the TAP/p15 mRNA nuclear export factor. *Molecular Cell*, 8(3), pp.645–656.
- Fu, D. & Collins, K., 2007. Purification of Human Telomerase Complexes Identifies Factors Involved in Telomerase Biogenesis and Telomere Length Regulation. *Molecular Cell*, 28(5), pp.773–785.
- Fukuda, M. et al., 1997. CRM1 is responsible for intracellular transport mediated by the nuclear export signal. *Nature*, 390(6657), pp.308–311.

- Gabler, S. et al., 1998. E1B 55-Kilodalton-Associated Protein: a Cellular Protein with RNA-Binding Activity Implicated in Nucleocytoplasmic Transport of Adenovirus and Cellular mRNAs. *Journal of Virology*, 72(10), pp.7960–7971.
- Galani, K. et al., 2001. The intracellular location of two aminoacyl-tRNA synthetases depends on complex formation with Arc1p. *The EMBO Journal*, 20(23), pp.6889–6898.
- Geuens, T., Bouhy, D. & Timmerman, V., 2016. The hnRNP family: insights into their role in health and disease. *Human Genetics*, 135(8), pp.851–867.
- Girard, C. et al., 2012. Post-transcriptional spliceosomes are retained in nuclear speckles until splicing completion. *Nature Communications*, 3(994), doi:10.1038/ncomms1998.
- Göhring, F. et al., 1997. The novel SAR-binding domain of scaffold attachment factor A (SAF-A) is a target in apoptotic nuclear breakdown. *The EMBO Journal*, 16(24), pp.7361–7371.
- Göhring, F. & Fackelmayer, F.O., 1997. The scaffold/matrix attachment region binding protein hnRNP-U (SAF-A) is directly bound to chromosomal DNA in vivo: a chemical cross-linking study. *Biochemistry*, 36(27), pp.8276–83.
- Goldberg, M.W. & Allen, T.D., 1992. High resolution scanning electron microscopy of the nuclear envelope: Demonstration of a new, regular, fibrous lattice attached to the baskets of the nucleoplasmic face of the nuclear pores. *Journal of Cell Biology*, 119(6), pp.1429–1440.
- Görlich, D. et al., 1997. A novel class of RanGTP binding proteins. *Journal of Cell Biology*, 138(1), pp.65–80.
- Green, R. & Rogers, E.J., 2013. Chemical Transformation of *E. coli*. *Methods in enzymology*, (529), pp.329–336.
- Greider, C.W. & Blackburn, E.H., 1989. A telomeric sequence in the RNA of *Tetrahymena* telomerase required for telomere repeat synthesis. *Nature*, 337(6205), pp.331–337.

- Greider, C.W. & Blackburn, E.H., 1987. The telomere terminal transferase of Tetrahymena is a ribonucleoprotein enzyme with two kinds of primer specificity. *Cell*, 51(6), pp.887–98.
- Gromadzka, A.M. et al., 2016. A short conserved motif in ALYREF directs cap- and EJC-dependent assembly of export complexes on spliced mRNAs. *Nucleic Acids Research*, 44(5), pp.2348–2361.
- Grzechnik, P., Tan-Wong, S.M. & Proudfoot, N.J., 2014. Terminate and make a loop: Regulation of transcriptional directionality. *Trends in Biochemical Sciences*, 39(7), pp.319–327.
- Gupta, R. a et al., 2010. Long noncoding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature*, 464(7291), pp.1071–1076.
- Gupte, R., Liu, Z. & Kraus, W.L., 2017. Parps and adp-ribosylation: Recent advances linking molecular functions to biological outcomes. *Genes & Development*, 31(2), pp.101–126.
- Güttler, T. & Görlich, D., 2011. Ran-dependent nuclear export mediators: A structural perspective. *The EMBO Journal*, 30(17), pp.3457–3474.
- Guttman, M. & Rinn, J.L., 2012. Modular regulatory principles of large non-coding RNAs. *Nature*, 482(7385), pp.339–346.
- Hacisuleyman, E. et al., 2016. Function and evolution of local repeats in the Firre locus. *Nature Communications*, 7(11021), doi:10.1038/ncomms11021.
- Hacisuleyman, E. et al., 2014. Topological organization of multichromosomal regions by the long intergenic noncoding RNA Firre. *Nature Structural and Molecular Biology*, 21(2), pp.198–206.
- Hall, L.L. et al., 2014. Stable COT-1 repeat RNA is abundant and is associated with euchromatic interphase chromosomes. *Cell*, 156(5), pp.907–919.
- Hart, T. et al., 2015. High-Resolution CRISPR Screens Reveal Fitness Genes and Genotype-Specific Cancer Liabilities. *Cell*, 163(6), pp.1515–1526.

- Hasegawa, Y. et al., 2010. The matrix protein hnRNP U is required for chromosomal localization of xist RNA. *Developmental Cell*, 19(3), pp.469–476.
- Hautbergue, G.M. et al., 2008. Mutually exclusive interactions drive handover of mRNA from export adaptors to TAP. *Proceedings of the National Academy of Sciences of the United States of America*, 105(13), pp.5154–9.
- Hautbergue, G.M. et al., 2009. UIF, a New mRNA export adaptor that works together with REF/ALY, requires FACT for recruitment to mRNA. *Current Biology*, 19(22), pp.1918–24.
- Heard, E. et al., 2001. Methylation of histone H3 at Lys-9 is an early mark on the X chromosome during X inactivation. *Cell*, 107(6), pp.727–738.
- Heath, C.G., Viphakone, N. & Wilson, S.A., 2016. The role of TREX in gene expression and disease. *Biochemical Journal*, 473(19), pp.2911–2935.
- Hegde, M.L. et al., 2012. Enhancement of NEIL1 protein-initiated oxidized DNA base excision repair by heterogeneous nuclear ribonucleoprotein U (hnRNP-U) via direct interaction. *Journal of Biological Chemistry*, 287(41), pp.34202–34211.
- Hegde, M.L. et al., 2016. Scaffold attachment factor A (SAF-A) and Ku temporally regulate repair of radiation-induced clustered genome lesions. *Oncotarget*, 7(34), pp.54430–54444.
- Helbig, R. & Fackelmayer, F.O., 2003. Scaffold attachment factor A (SAF-A) is concentrated in inactive X chromosome territories through its RGG domain. *Chromosoma*, 112(4), pp.173–82.
- Henry, R.W. et al., 1998. SNAP19 mediates the assembly of a functional core promoter complex (SNAP(c)) shared by RNA polymerases II and III. *Genes & Development*, 12(17), pp.2664–2672.
- Ho, J.H., Kallstrom, G. & Johnson, A.W., 2000. Nmd3p is a Crm1p-dependent adapter protein for nuclear export of the large ribosomal subunit. *The Journal of Cell Biology*, 151(5), pp.1057–66.

- Hong, Z. et al., 2013. The Role of hnRPU1 Involved in DNA Damage Response Is Related to PARP1. *PLoS One*, 8(4), p.e60208.
- van Hoof, A. et al., 2000. Function of the ski4p (Csl4p) and Ski7p proteins in 3'-to-5' degradation of mRNA. *Molecular and Cellular Biology*, 20(21), pp.8230–8243.
- Hopper, A.K., Pai, D.A. & Engelke, D.R., 2011. Cellular dynamics of tRNAs and their genes. *FEBS Letters*, 584(2), pp.310–317.
- Hopper, A.K. & Phizicky, E.M., 2003. tRNA transfers to the limelight. *Genes & Development*, 17(717), pp.162–180.
- Huang, H.Y. & Hopper, A.K., 2015. In vivo biochemical analyses reveal distinct roles of β -importins and eEF1A in tRNA subcellular traffic. *Genes & Development*, 29(7), pp.772–783.
- Huang, Y. et al., 2003. SR splicing factors serve as adapter proteins for TAP-dependent mRNA export. *Molecular Cell*, 11(3), pp.837–843.
- Huang, Y. & Steitz, J.A., 2001. Splicing factors SRp20 and 9G8 promote the nucleocytoplasmic export of mRNA. *Molecular Cell*, 7(4), pp.899–905.
- Huber, J. et al., 1998. Snurportin1, an m3G-cap-specific nuclear import receptor with a novel domain structure. *The EMBO Journal*, 17(14), pp.4114–4126.
- Huelga, S.C. et al., 2012. Integrative Genome-wide Analysis Reveals Cooperative Regulation of Alternative Splicing by hnRNP Proteins. *Cell Reports*, 1(2), pp.167–178.
- Hung, M.-L. et al., 2010. Arginine methylation of REF/ALY promotes efficient handover of mRNA to TAP/NXF1. *Nucleic Acids Research*, 38(10), pp.3351–61.
- Hunter, C. a et al., 2003. PAUSED encodes the Arabidopsis exportin-t ortholog. *Plant physiology*, 132(4), pp.2135–2143.
- Huppertz, I. et al., 2014. iCLIP: Protein-RNA interactions at nucleotide resolution. *Methods*, 65(3), pp.274–287.

- Hutchinson, J.N. et al., 2007. A screen for nuclear transcripts identifies two linked noncoding RNAs associated with SC35 splicing domains. *BMC Genomics*, 8(39), pp.1–16.
- Ingolia, N.T., Lareau, L.F. & Weissman, J.S., 2011. Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell*, 147(4), pp.789–802.
- Izaurralde, E. et al., 1995. A cap-binding protein complex mediating U snRNA export. *Nature*, 376(6542), pp.709–712.
- Izaurralde, E. et al., 1997. The asymmetric distribution of the constituents of the Ran system is essential for transport into and out of the nucleus. *The EMBO Journal*, 16(21), pp.6535–6547.
- Jády, B.E. et al., 2003. Modification of Sm small nuclear RNAs occurs in the nucleoplasmic Cajal body following import from the cytoplasm. *The EMBO Journal*, 22(8), pp.1878–1888.
- Jády, B.E., Bertrand, E. & Kiss, T., 2004. Human telomerase RNA and box H/ACA scaRNAs share a common Cajal body-specific localization signal. *Journal of Cell Biology*, 164(5), pp.647–652.
- Jiang, H. et al., 2017. Long noncoding RNA CRNDE stabilized by hnRNPUL2 accelerates cell proliferation and migration in colorectal carcinoma via activating Ras/MAPK signaling pathways. *Cell Death and Disease*, 8(6), p.e2862.
- Johnson, S.A., Cubberley, G. & Bentley, D.L., 2009. Cotranscriptional recruitment of the mRNA export factor Yra1 by direct interaction with the 3' end processing factor Pcf11. *Molecular Cell*, 33(2), pp.215–26.
- Kaminska, M. et al., 2009. Dynamic organization of aminoacyl-tRNA synthetase complexes in the cytoplasm of human cells. *Journal of Biological Chemistry*, 284(20), pp.13746–13754.
- Kass, S. et al., 1990. The U3 small nucleolar ribonucleoprotein functions in the first step of preribosomal RNA processing. *Cell*, 60(6), pp.897–908.

- Kepinski, S. & Leyser, O., 2004. Auxin-induced SCFTIR1-Aux/IAA interaction involves stable modification of the SCFTIR1 complex. *Proceedings of the National Academy of Sciences of the United States of America*, 101(33), pp.12381–12386.
- Kilchert, C., Wittmann, S. & Vasiljeva, L., 2016. The regulation and functions of the nuclear RNA exosome complex. *Nature Reviews Molecular Cell Biology*, 17(4), pp.227–239.
- Kiledjian, M. & Dreyfuss, G., 1992. Primary structure and binding activity of the hnRNP U protein: binding RNA through RGG box. *The EMBO Journal*, 11(7), pp.2655–2664.
- Kim, M.K. & Nikodem, V.M., 1999. hnRNP U Inhibits Carboxy-Terminal Domain Phosphorylation by TFIIH and Represses RNA Polymerase II Elongation. *Molecular and Cellular Biology*, 19(10), pp.6833–6844.
- Kipp, M. et al., 2000. Apoptotic cleavage of scaffold attachment factor A (SAF-A) by caspase-3 occurs at a noncanonical cleavage site. *Journal of Biological Chemistry*, 275(7), pp.5031–5036.
- Kiss, T., 2002. Small nucleolar RNAs: An abundant group of noncoding RNAs with diverse cellular functions. *Cell*, 109(2), pp.145–148.
- Kitao, S. et al., 2008. A Compartmentalized Phosphorylation/Dephosphorylation System That Regulates U snRNA Export from the Nucleus. *Molecular and Cellular Biology*, 28(1), pp.487–497.
- Kohlmaier, A. et al., 2004. A chromosomal memory triggered by Xist regulates histone methylation in X inactivation. *PLOS Biology*, 2(7), p.E171.
- König, J. et al., 2011. iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nature Structural and Molecular Biology*, 17(7), pp.909–915.
- Krämer, A. et al., 1999. Architecture of the Mammalian U2 snRNP. *The Journal of Cell Biology*, 145(7), pp.1355–1368.

- Kukalev, A. et al., 2005. Actin and hnRNP U cooperate for productive transcription by RNA polymerase II. *Nature Structural and Molecular Biology*, 12(3), pp.238–244.
- Kutay, U. et al., 1997. Export of importin α from the nucleus is mediated by a specific nuclear transport factor. *Cell*, 90(6), pp.1061–1071.
- Kutay, U. et al., 1998. Identification of a tRNA-specific nuclear export receptor. *Molecular Cell*, 1(3), pp.359–369.
- Li, J.-H. et al., 2015. Discovery of protein-lncRNA interactions by integrating large-scale CLIP-Seq and RNA-Seq datasets. *Frontiers in Bioengineering and Biotechnology*, 2(88), pp.1–11.
- Lipowsky, G. et al., 1999. Coordination of tRNA nuclear export with processing of tRNA. *RNA*, 5(4), pp.539–549.
- Lippai, M. et al., 2000. The Ketel gene encodes a Drosophila homologue of importin- β . *Genetics*, 156(4), pp.1889–1900.
- Liu, Q., Greimann, J.C. & Lima, C.D., 2006. Reconstitution, Activities, and Structure of the Eukaryotic RNA Exosome. *Cell*, 127(6), pp.1223–1237.
- Liu, Z. et al., 2017. Systematic comparison of 2A peptides for cloning multi-genes in a polycistronic vector. *Scientific Reports*, 7(2193), doi:10.1038/s41598-017-02460-2.
- Lowe, T.M. & Eddy, S.R., 1997. TRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Research*, 25(5), pp.955–964.
- Lubas, M. et al., 2011. Interaction Profiling Identifies the Human Nuclear Exosome Targeting Complex. *Molecular Cell*, 43(4), pp.624–637.
- Lubas, M. et al., 2015. The Human Nuclear Exosome Targeting Complex Is Loaded onto Newly Synthesized RNA to Direct Early Ribonucleolysis. *Cell Reports*, 10(2), pp.178–192.
- Lubelsky, Y. & Ulitsky, I., 2018. Sequences enriched in Alu repeats drive nuclear

- localization of long RNAs in human cells. *Nature*, 555(7694), pp.107–111.
- Lund, E. et al., 2004. Nuclear export of microRNA precursors. *Science*, 303(5654), pp.95–98.
- Lund, E. & Dahlberg, J.E., 1998. Proofreading and aminoacylation of tRNAs before export from the nucleus. *Science*, 282(5396), pp.2082–2085.
- Luo, M.L. et al., 2001. Pre-mRNA splicing and mRNA export linked by direct interactions between UAP56 and Aly. *Nature*, 413(6856), pp.644–7.
- Lykke-Andersen, S. et al., 2011. The eukaryotic RNA exosome: Same scaffold but variable catalytic subunits. *RNA Biology*, 8(1), pp.61–6.
- Machyna, M. et al., 2014. The coilin interactome identifies hundreds of small noncoding RNAs that traffic through cajal bodies. *Molecular Cell*, 56(3), pp.389–399.
- Mali, P. et al., 2013. RNA-Guided Human Genome Engineering via Cas9. *Science*, 339(6121), pp.823–826.
- Mao, Y.S. et al., 2011. Direct Visualization of the Co-transcriptional Assembly of a Nuclear Body by Noncoding RNAs. *Nature Cell Biology*, 13(1), pp.95–101.
- McCracken, S. et al., 1997. The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. *Nature*, 385(6614), pp.357–361.
- McHugh, C.A. et al., 2015. The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature*, 521(7551), pp.232–6.
- Meininger, I. et al., 2016. Alternative splicing of MALT1 controls signalling and activation of CD4+ T cells. *Nature Communications*, 7(11292), doi:10.1038/ncomms11292.
- Melchior, F., 2001. Ran GTPase cycle: One mechanism - Two functions. *Current Biology*, 11(7), pp.257–260.
- Mermoud, J.E. et al., 2002. Histone H3 lysine 9 methylation occurs rapidly at the

onset of random X chromosome inactivation. *Current Biology*, 12(3), pp.247–251.

Meszaros, N. et al., 2015. Nuclear Pore Basket Proteins Are Tethered to the Nuclear Envelope and Can Regulate Membrane Nuclear Pore Basket Proteins Are Tethered to the Nuclear Envelope and Can Regulate Membrane Curvature. *Developmental Cell*, 33(3), pp.285–298.

Mingot, J. et al., 2013. EEF1A mediates the nuclear export of snag-containing proteins via the exportin5-aminoacyl-tRNA complex. *Cell Reports*, 5(3), pp.727–737.

Mitchell, P. et al., 1997. The exosome: A conserved eukaryotic RNA processing complex containing multiple 3'→5' exoribonucleases. *Cell*, 91(4), pp.457–466.

Moindrot, B. et al., 2015. A Pooled shRNA Screen Identifies Rbm15, Spen, and Wtap as Factors Required for Xist RNA-Mediated Silencing. *Cell Reports*, 12(4), pp.562–572.

Monfort, A. et al., 2015. Identification of Spen as a crucial factor for Xist function through forward genetic screening in haploid embryonic stem cells. *Cell Reports*, 12(4), pp.554–561.

Moore, M.S. & Blobel, G., 1993. The GTP-binding protein Ran/TC4 is required for protein import into the nucleus. *Nature*, 365(6447), pp.661–663.

Muhlrad, D. & Parker, R., 1999. Aberrant mRNAs with extended 3' UTRs are substrates for rapid degradation by mRNA surveillance. *RNA*, 5(10), pp.1299–1307.

Natsume, T. et al., 2016. Rapid Protein Depletion in Human Cells by Auxin-Inducible Degron Tagging with Short Homology Donors. *Cell Reports*, 15(1), pp.210–218.

Nesic, D., 2004. A role for Cajal bodies in the final steps of U2 snRNP biogenesis. *Journal of Cell Science*, 117(19), pp.4423–4433.

Nishimura, K. et al., 2009. An auxin-based degron system for the rapid depletion of

- proteins in nonplant cells. *Nature Methods*, 6(12), pp.917–922.
- Nishimura, K. & Kanemaki, M.T., 2014. Rapid depletion of budding yeast proteins via the fusion of an auxin-inducible degron (AID). *Current Protocols in Cell Biology*, 64, p.20.9.1-20.9.16.
- Nozawa, R. et al., 2017. SAF-A regulates interphase chromosome structure through oligomerisation with chromatin-associated RNAs. *Cell*, 169(7), p.1214–1227.e18.
- Obrdlik, A. et al., 2008. The Histone Acetyltransferase PCAF Associates with Actin and hnRNP U for RNA Polymerase II Transcription. *Molecular and Cellular Biology*, 28(20), pp.6342–6357.
- Ohira, T. & Suzuki, T., 2016. Precursors of tRNAs are stabilized by methylguanosine cap structures. *Nature Chemical Biology*, 12(8), pp.648–655.
- Ohno, M. et al., 2000. PHAX, a mediator of U snRNA nuclear export whose activity is regulated by phosphorylation. *Cell*, 101(2), pp.187–198.
- Okamura, M., Inose, H. & Masuda, S., 2015. RNA Export through the NPC in Eukaryotes. *Genes*, 6(1), pp.124–149.
- Paraskeva, E. et al., 1999. CRM1-mediated Recycling of Snurportin 1 to the Cytoplasm. *The Journal of Cell Biology*, 145(2), pp.255–264.
- Paushkin, S. V. et al., 2004. Identification of a human endonuclease complex reveals a link between tRNA splicing and pre-mRNA 3' end formation. *Cell*, 117(3), pp.311–321.
- Peltz, S.W., Brown, A.H. & Jacobson, A., 1993. mRNA destabilization triggered by premature translational termination depends on at least three cis-acting sequence elements and one trans-acting factor. *Genes & Development*, 7(9), pp.1737–1754.
- Peters, A.H.F.M. et al., 2002. Histone H3 lysine 9 methylation is an epigenetic imprint of facultative heterochromatin. *Nature Genetics*, 30(1), pp.77–80.

- Phizicky, E.M. & Hopper, A.K., 2010. tRNA biology charges to the front. *Genes & Development*, 24(17), pp.1832–1860.
- Plath, K., 2003. Role of Histone H3 Lysine 27 Methylation in X Inactivation. *Science*, 300(5616), pp.131–135.
- Polo, S.E. et al., 2012. Regulation of DNA-End Resection by hnRNPU-like Proteins Promotes DNA Double-Strand Break Signaling and Repair. *Molecular Cell*, 45(4), pp.505–516.
- Popow, J. et al., 2014. Analysis of orthologous groups reveals archease and DDX1 as tRNA splicing factors. *Nature*, 511(7507), pp.104–107.
- Popow, J. et al., 2011. HSPC117 Is the Essential Subunit of a Human tRNA Splicing Ligase Complex. *Science*, 331(6018), pp.760–764.
- Pradet-Balade, B. et al., 2011. CRM1 controls the composition of nucleoplasmic pre-snoRNA complexes to licence them for nucleolar transport. *The EMBO Journal*, 30(11), pp.2205–2218.
- Proudfoot, N.J., 2011. Ending the message: Poly(A) signals then and now. *Genes & Development*, 25(17), pp.1770–1782.
- Pullirsch, D. et al., 2010. The Trithorax group protein Ash2l and Saf-A are recruited to the inactive X chromosome at the onset of stable X inactivation. *Development*, 137(6), pp.935–943.
- Puvvula, P.K. et al., 2014. Long noncoding RNA PANDA and scaffold-attachment-factor SAFA control senescence entry and exit. *Nature communications*, 5(5323), doi:10.1038/ncomms6323.
- Ramos, J.A. et al., 2001. Rapid degradation of auxin/indoleacetic acid proteins requires conserved amino acids of domain II and is proteasome dependent. *Plant Cell*, 13(10), pp.2349–2360.
- Ran, F.A. et al., 2013. Genome engineering using the CRISPR-Cas9 system. *Nature Protocols*, 8(11), pp.2281–308.

- Raška, I. et al., 1991. Immunological and ultrastructural studies of the nuclear coiled body with autoimmune antibodies. *Experimental Cell Research*, 195(1), pp.27–37.
- Raun, N. et al., 2017. Quantitative phenotypic and network analysis of 1q44 microdeletion for microcephaly. *American Journal of Medical Genetics, Part A*, 173(4), pp.972–977.
- Reddy, A.S. et al., 2016. A Comprehensive Analysis of Cell Type-Specific Nuclear RNA From Neurons and Glia of the Brain. *Biological Psychiatry*, 81(3), pp.252–264.
- Richard, P. et al., 2003. A common sequence motif determines the Cajal body-specific localization of box H / ACA scaRNAs. *The EMBO Journal*, 22(16), pp.4283–93.
- Rinn, J.L. et al., 2007. Functional Demarcation of Active and Silent Chromatin Domains in Human HOX Loci by Noncoding RNAs. *Cell*, 129(7), pp.1311–1323.
- Rinn, J.L. & Chang, H.Y., 2012. Genome regulation by long noncoding RNAs. *Annual Review of Biochemistry*, 81, pp.145–166.
- De Robertis, E.M., Black, P. & Nishikura, K., 1981. Intranuclear location of the tRNA splicing enzymes. *Cell*, 23(1), pp.89–93.
- Romig, H. et al., 1992. Characterization of SAF-A, a novel nuclear DNA binding protein from HeLa cells with high affinity for nuclear matrix/scaffold attachment DNA elements. *The EMBO Journal*, 11(9), pp.3431–3440.
- Roshon, M.J. & Ruley, H.E., 2005. Hypomorphic mutation in hnRNP U results in post-implantation lethality. *Transgenic Research*, 14(2), pp.179–192.
- Saikia, M. et al., 2010. Genome-wide analysis of N¹-methyl-adenosine modification in human tRNAs. *RNA*, 16, pp.1317–1327.
- Sakaguchi, T. et al., 2016. Control of Chromosomal Localization of Xist by hnRNP U Family Molecules. *Developmental Cell*, 39(1), pp.11–12.
- Schimmel, P., 2018. RNA Processing and Modifications: The emerging complexity of

- the tRNA world: Mammalian tRNAs beyond protein synthesis. *Nature Reviews Molecular Cell Biology*, 19(1), pp.45–58.
- Schlackow, M. et al., 2017. Distinctive Patterns of Transcription and RNA Processing for Human lincRNAs. *Molecular Cell*, 65(1), pp.25–38.
- Schmidt, K. & Butler, J.S., 2013. Nuclear RNA surveillance: Role of TRAMP in controlling exosome specificity. *Wiley Interdisciplinary Reviews: RNA*, 4(2), pp.217–231.
- Schneider, C., Anderson, J.T. & Tollervey, D., 2007. The Exosome Subunit Rrp44 Plays a Direct Role in RNA Substrate Recognition. *Molecular Cell*, 27(2), pp.324–331.
- Schneider, C. & Tollervey, D., 2014. Looking into the barrel of the RNA exosome. *Nature Structural and Molecular Biology*, 21(1), pp.17–18.
- Schuller, J.M. et al., 2018. Structure of the nuclear exosome captured on a maturing preribosome. *Science*, 360(6385), pp.219–222.
- Schwanhüsser, B. et al., 2011. Global quantification of mammalian gene expression control. *Nature*, 473(7347), pp.337–342.
- Segref, A., Mattaj, I.W. & Ohno, M., 2001. The evolutionarily conserved region of the U snRNA export mediator PHAX is a novel RNA-binding domain that is essential for U snRNA export. *RNA*, 7(3), pp.351–360.
- Shaheen, H.H. et al., 2007. Retrograde nuclear accumulation of cytoplasmic tRNA in rat hepatoma cells in response to amino acid deprivation. *Proceedings of the National Academy of Sciences of the United States of America*, 104(21), pp.8845–8850.
- Shcherbik, N. et al., 2010. Polyadenylation and degradation of incomplete RNA polymerase I transcripts in mammalian cells. *EMBO Reports*, 11(2), pp.106–111.
- Shen, H. et al., 2008. Distinct activities of the DExD/H-box splicing factor hUAP56 facilitate stepwise assembly of the spliceosome. *Genes & Development*, 22(13), pp.1796–1803.

- Shen, J., Zhang, L. & Zhao, R., 2007. Biochemical characterization of the ATPase and helicase activity of UAP56, an essential pre-mRNA splicing and mRNA export factor. *The Journal of Biological Chemistry*, 282(31), pp.22544–50.
- Shibata, S. et al., 2006. Exportin-5 orthologues are functionally divergent among species. *Nucleic Acids Research*, 34(17), pp.4711–4721.
- Shukla, C.J. et al., 2018. High-throughput identification of RNA nuclear enrichment sequences. *The EMBO Journal*, 37(6), p.e98452.
- Silla, T., Karadoulama, E. & Ma, D., 2018. The RNA Exosome Adaptor ZFC3H1 Functionally Competes with Nuclear Export Activity to Retain Target Transcripts. *Cell Reports*, (23), pp.2199–2210.
- Silva, J. et al., 2003. Establishment of histone H3 methylation on the inactive X chromosome requires transient recruitment of Eed-Enx1 polycomb group complexes. *Developmental Cell*, 4(4), pp.481–495.
- Smeets, D. et al., 2014. Three-dimensional super-resolution microscopy of the inactive X chromosome territory reveals a collapse of its active nuclear compartment harboring distinct Xist RNA foci. *Epigenetics & Chromatin*, 7(8), doi:10.1186/1756-8935-7-8.
- Smith, J. et al., 2014. Translation of small open reading frames within unannotated RNA transcripts in *Saccharomyces cerevisiae*. *Cell Reports*, 7(6), pp.1858–1866.
- Spector, D.L. & Lamond, A.I., 2011. Nuclear Speckles. *Cold Spring Harbor Perspective Biology*, 3(2), pp.1–12.
- Spraggon, L. et al., 2007. hnRNP-U directly interacts with WT1 and modulates WT1 transcriptional activation. *Oncogene*, 26(10), pp.1484–1491.
- Stade, K. et al., 1997. Exportin 1 (Crm1p) is an essential nuclear export factor. *Cell*, 90(6), pp.1041–1050.
- Sträßer, K. et al., 2002. TREX is a conserved complex coupling transcription with messenger RNA export. *Nature*, 417(6886), pp.304–308.

- Sun, L. et al., 2013. Long noncoding RNAs regulate adipogenesis. *PNAS*, 110(9), pp.3387–3392.
- Sunwoo, H. et al., 2009. Men e/b nuclear-retained non-coding RNAs are up-regulated upon muscle differentiation and are essential components of paraspeckles. *Genome Research*, 19(3), pp.347–359.
- Suzuki, H. et al., 2015. Nuclear TDP-43 causes neuronal toxicity by escaping from the inhibitory regulation by hnRNPs. *Human Molecular Genetics*, 24(6), pp.1513–1527.
- Suzuki, T., Nagao, A. & Suzuki, T., 2011. Human Mitochondrial tRNAs: Biogenesis, Function, Structural Aspects, and Diseases. *Annual Review of Genetics*, 45(1), pp.299–329.
- Takano, A., Endo, T. & Yoshihisa, T., 2005. tRNA actively shuttles between the nucleus and cytosol in yeast. *Science*, 309(5731), pp.140–142.
- Taniguchi, I. & Ohno, M., 2008. ATP-dependent recruitment of export factor Aly/REF onto intronless mRNAs by RNA helicase UAP56. *Molecular and Cellular Biology*, 28(2), pp.601–8.
- The ENCODE Project Consortium, 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature*, 489(7414), pp.57–74.
- The FANTOM Consortium, 2006. The Transcriptional Landscape of the Mammalian Genome. *Science*, 309(5740), pp.1559–1563.
- Thierry, G. et al., 2012. Molecular characterization of 1q44 microdeletion in 11 patients reveals three candidate genes for intellectual disability and seizures. *American Journal of Medical Genetics, Part A*, 158(7), pp.1633–1640.
- Thomas, F. & Kutay, U., 2003. Biogenesis and nuclear export of ribosomal subunits in higher eukaryotes depend on the CRM1 export pathway. *Journal of Cell Science*, 116(12), pp.2409–2419.
- Tomecki, R. et al., 2010. The human core exosome interacts with differentially

- localized processive RNases: HDIS3 and hDIS3L. *The EMBO Journal*, 29(14), pp.2342–2357.
- Tomlinson, R., Ziegler, T.D. & Terns, M., 2006. Cell Cycle-regulated Trafficking of Human Telomerase to Telomeres. *Molecular Biology of the Cell*, 17(2), pp.955–965.
- Topisirovic, I. et al., 2011. Cap and cap-binding proteins in the control of gene expression. *Wiley Interdisciplinary Reviews: RNA*, 2(2), pp.277–298.
- Topisirovic, I. et al., 2009. Molecular dissection of the eukaryotic initiation factor 4E (eIF4E) export-competent RNP. *The EMBO Journal*, 28(8), pp.1087–98.
- Tran, E.J. & Wentz, S.R., 2006. Dynamic Nuclear Pore Complexes: Life on the Edge. *Cell*, 125(6), pp.1041–1053.
- Trcek, T. et al., 2013. Temporal and spatial characterization of nonsense-mediated mRNA decay. *Genes & Development*, 27(5), pp.541–551.
- Tripathi, V. et al., 2010. The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Molecular Cell*, 39(6), pp.925–938.
- Tsai, M. et al., 2010. Long Noncoding RNA as Modular Scaffold of Histone Modification Complexes. *Science*, 329(5992), pp.689–693.
- Tycowski, K.T. et al., 2009. A Conserved WD40 Protein Binds the Cajal Body Localization Signal of scaRNP Particles. *Molecular Cell*, 34(1), pp.47–57.
- Tycowski, K.T., Aab, A. & Steitz, J.A., 2004. Guide RNAs with 5' Caps and Novel Box C/D snoRNA-like Domains for Modification of snRNAs in Metazoa. *Current Biology*, 14(22), pp.1985–1995.
- Valente, S.T. & Goff, S.P., 2006. Inhibition of HIV-1 Gene Expression by a Fragment of hnRNP U. *Molecular Cell*, 23(4), pp.597–605.
- Venteicher, A.S. et al., 2009. A human telomerase holoenzyme protein required for Cajal body localization and telomere synthesis. *Science*, 323(5914), pp.644–648.

- Viphakone, N. et al., 2015. Luzp4 defines a new mRNA export pathway in cancer cells. *Nucleic Acids Research*, 43(4), pp.2353–2366.
- Viphakone, N. et al., 2012. TREX exposes the RNA-binding domain of Nxf1 to enable mRNA export. *Nature communications*, 3(1006), doi:10.1038/ncomms2005.
- Wang, J. et al., 2017. Current Research on Non-Coding Ribonucleic Acid (RNA). *Genes*, 8(12), p.E336.
- Wang, T. et al., 2015. Identification and characterization of essential genes in the human genome. *Science*, 350(6264), pp.1096–1101.
- Wang, Y. et al., 2015. 2A self-cleaving peptide-based multi-gene expression system in the silkworm *Bombyx mori*. *Scientific Reports*, 5(16273), doi:10.1038/srep16273.
- Wasmuth, E. V. & Lima, C.D., 2012. Exo- and Endoribonucleolytic Activities of Yeast Cytoplasmic and Nuclear RNA Exosomes Are Dependent on the Noncatalytic Core and Central Channel. *Molecular Cell*, 48(1), pp.133–144.
- Wen, W. et al., 1995. Identification of a signal for rapid export of proteins from the nucleus. *Cell*, 82(3), pp.463–473.
- Werner, M.S. et al., 2017. Chromatin-enriched lncRNAs can act as cell-type specific activators of proximal gene transcription. *Nature Structural and Molecular Biology*, 24(7), pp.596–603.
- Wery, M. et al., 2016. Nonsense-Mediated Decay Restricts lncRNA Levels in Yeast Unless Blocked by Double-Stranded RNA Structure. *Molecular Cell*, 61(3), pp.379–392.
- West, J.A. et al., 2014. The Long Noncoding RNAs NEAT1 and MALAT1 Bind Active Chromatin Sites. *Molecular Cell*, 55(5), pp.791–802.
- Wild, T. et al., 2010. A Protein Inventory of Human Ribosome Biogenesis Reveals an Essential Function of Exportin 5 in 60S Subunit Export. *PLOS Biology*, 8(10), p.e1000522.

- Will, C.L. & Lührmann, R., 2001. Spliceosomal UsnRNP biogenesis, structure and function. *Current Opinion in Cell Biology*, 13(3), pp.290–301.
- Wu, J. et al., 2015. Genome-wide screen uncovers novel pathways for tRNA processing and nuclear-cytoplasmic dynamics. *Genes & Development*, 29(24), pp.2633–2644.
- Wutz, A., Rasmussen, T.P. & Jaenisch, R., 2002. Chromosomal silencing and localization are mediated by different domains of Xist RNA. *Nature genetics*, 30(2), pp.167–74.
- Xiao, R. et al., 2012. Nuclear Matrix Factor hnRNP U/SAF-A Exerts a Global Control of Alternative Splicing by Regulating U2 snRNP Maturation. *Molecular Cell*, 45(5), pp.656–668.
- Yao, W. et al., 2007. Nuclear Export of Ribosomal 60S Subunits by the General mRNA Export Receptor Mex67-Mtr2. *Molecular Cell*, 26(1), pp.51–62.
- Yates, T.M. et al., 2017. De novo mutations in *HNRNPU* result in a neurodevelopmental syndrome. *American Journal of Medical Genetics Part A*, 173(11), pp.3003–3012.
- Ye, J. et al., 2015. hnRNP U protein is required for normal pre-mRNA splicing and postnatal heart development and function. *Proceedings of the National Academy of Sciences*, 112(23), p.201508461.
- Yoshihisa, T. et al., 2007. Cytoplasmic splicing of tRNA in *Saccharomyces cerevisiae*. *Genes to Cells*, 12(3), pp.285–297.
- Yoshimoto, R. et al., 2017. Global analysis of pre-mRNA subcellular localization following splicing inhibition by spliceostatin A. *RNA*, 23(1), pp.47–57.
- Zahler, a. M. et al., 1992. SR proteins: A conserved family of pre-mRNA splicing factors. *Genes & Development*, 6(5), pp.837–847.
- Zaitseva, L., Myers, R. & Fassati, A., 2006. tRNAs promote nuclear import of HIV-1 intracellular reverse transcription complexes. *PLOS Biology*, 4(10), pp.1689–

1706.

Zemp, I. & Kutay, U., 2007. Nuclear export and cytoplasmic maturation of ribosomal subunits. *FEBS Letters*, 581(15), pp.2783–2793.

Zhou, L. et al., 2011. Transportin 3 promotes a nuclear maturation step required for efficient HIV-1 integration. *PLOS Pathogens*, 7(8), p.e1002194.

Zhu, Y. et al., 2004. Telomerase RNA Accumulates in Cajal Bodies in Human Cancer Cells. *Molecular Biology of the Cell*, 15(1), pp.81–90.