DISSERTATION

GLOBAL ANALYSIS REVEALS DIFFERENTIAL REGULATION OF MRNA DECAY IN HUMAN INDUCED PLURIPOTENT STEM CELLS

Submitted by

Ashley T. Neff

Graduate Degree Program in Cell and Molecular Biology

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Fall 2013

Doctoral Committee:

Advisor: Jeffrey Wilusz
Co-advisor: Carol J. Wilusz

Douglas H. Thamm
Michael Weil

ABSTRACT

GLOBAL ANALYSIS REVEALS DIFFERENTIAL REGULATION OF MRNA DECAY IN HUMAN INDUCED PLURIPOTENT STEM CELLS

Induced Pluripotent Stem (iPS) cells are able to proliferate indefinitely while maintaining the capacity for unlimited differentiation and these properties are reflected by global changes in gene expression required for reprogramming of differentiated cells. Although the rate of transcription is an important regulator of steady-state mRNA levels, mRNA decay also plays a significant role in modulating the expression of cell-specific genes. The contribution of regulated mRNA decay towards establishing and maintaining pluripotency is largely unknown. To address this, we sought to determine global mRNA decay rates in iPS cells and the genetically-matched fibroblasts (HFFs) they were derived from. Using a microarray based approach, we determined half-lives for 5,481 mRNAs in both cell lines and identified three classes of mRNAs whose decay is differentially regulated in iPS cells compared to HFFs.

We found that replication-dependent histone mRNAs are more abundant and more stable in iPS cells, resulting in increased histone protein abundances. This up-regulation of histone expression may facilitate the unique chromatin dynamics of pluripotent cells. A large set of C2H2 ZNF mRNAs are also stabilized in iPS cells compared to HFFs, possibly through reduced expression of miRNAs that target their coding regions. As many of these mRNAs encode transcriptional repressors, stabilization of these transcripts may support the overall increased expression of C2H2 ZNF transcription factors in early embryogenesis. Finally, we found that mRNAs containing C-rich elements in their 3'UTR are destabilized in iPS cells compared to HFFs and many of these mRNAs encode factors important for development. Interestingly, we also identified the Poly(C)-Binding Protein (PCBP) family as differentially regulated in iPS cells and investigated their possible involvement in regulation of the mRNAs in our dataset identified as destabilized in iPS cells and having C-rich 3'UTR elements.

Thus, we identified several interesting classes of mRNAs whose decay is differentially regulated in iPS cells compared to HFFs and our results highlight the importance of post-transcriptional control in stem cell gene expression. Coordinated control of mRNA decay is evident in pluripotency and characterization of the mechanisms involved would further contribute to our limited understanding of pluripotent gene expression and possibly identify additional targets for reprogramming.

ACKNOWLEDGEMENTS

First and foremost, I am thankful to Drs. Carol and Jeffrey Wilusz for their mentorship and support throughout my time training in their lab. They helped me develop technical skills and critical thinking necessary for posing and answering investigative questions, creating a tool set I will continue to use and build upon. They have encouraged me through several of the obstacles encountered as a graduate student growing into a researcher and finding a place within the scientific community, and I am grateful for their words of advice.

I also thank our collaborators, Dr. Ju Youn Lee and Dr. Bin Tian for their analysis of our microarray data. They provided a level of bioinformatic expertise that was crucial for carrying out this project. I am also grateful for the additional perspective and guidance provided by graduate committee members Dr. Douglas Thamm and Dr. Michael Weil. I appreciate that their input was always positive and constructive. In addition, I am thankful for previous graduates of the Wilusz lab, Dr. Kevin Sokoloski and Dr. Jerome Lee for answering upwards of a million questions during our time in the lab together and continuing to act as sounding boards when I need advice.

Finally, I am grateful for the unwavering support from my family and husband, Michael. They have endured and celebrated with me throughout this process and their backing has been instrumental. I know they are always proud of me, regardless of whether or not they understand the research, and I am so thankful for their encouragement.

TABLE OF CONTENTS

Chapter 1: Introduction

Pluripotent stem cells are a unique cell type that can proliferate indefinitely and undergo unrestricted differentiation (Thomson et al., 1998; Till and McCulloch, 1961). As stem cells progress towards differentiation, global changes in gene expression occur wherein distinct gene sets are up- or down-regulated to support a lineage-specific phenotype. DNA expression arrays have allowed genome-wide profiling of changes in gene expression upon differentiation into a number of cell types including adipocytes (Gerhold et al., 2002; Jessen and Stevens, 2002), myoblasts (Moran et al., 2002), and neurons (Loring et al., 2001). These analyses have identified several transcription factors whose expression can be manipulated to perform directed differentiation (Irion et al., 2008). In 2007, human adult somatic fibroblasts were reprogrammed back to an undifferentiated state using retrovirus-mediated transfection of pluripotency transcription factors OCT3/4, KLF4, SOX2, and MYC, generating Induced Pluripotent Stem (iPS) cells (Takahashi et al., 2007). Expression of these four transcription factors is sufficient to coordinately regulate the global changes necessary to achieve a pluripotent state. However, more recent studies indicate that post-transcriptional mechanisms, including mRNA decay, are also important for establishing stem cell gene expression (Porciuncula et al., 2013; Qin et al., 2012; Richards et al., 2004; Shibayama et al., 2009; Wang et al., 2013). As a prelude to discussion of the contribution of mRNA stability to gene expression, general features of the typical mRNA life cycle are presented below, giving particular emphasis to mRNA decay and the factors and enzymes important for regulated transcript degradation.

1.1 The typical mRNA life cycle

The life of an mRNA begins with transcription in the nucleus and continues through several processing steps before reaching its mature, translatable form in the cytoplasm. While the lifecycle

dynamics of individual mRNAs can have distinguishing characteristics, the general processes from synthesis to degradation are largely universal for eukaryotic mRNAs.

1.1.1 Transcription

In eukaryotic cells, genetic information encoded in DNA is transcribed into mRNA when RNA Polymerase II (Pol II) binds the promoter region of a gene to form a preinitiation complex including several general and specific transcription factors (Murakami et al., 2013; Safer et al., 1985). Binding of the preinitiation complex activates a conformational change in DNA structure necessary for transcription to proceed (Wilson and Spillman, 1982). Chromatin remodeling factors and helicases unwind the DNA duplex allowing Pol II access to the first two nucleotides of the template strand followed by formation of a highly efficient elongation complex (Kettenberger et al., 2004). Elongation of nascent RNAs continues in short bursts (Rajala et al., 2010) through the 3' Untranslated Region (UTR) until termination is signaled by a mechanism coupled with cleavage and polyadenylation (McCracken et al., 1997; Whitelaw and Proudfoot, 1986).

1.1.2 Capping

The 5' end of nascent RNAs undergoes co-transcriptional enzymatic events to add an inverted 7-methylguanosine (m^7G) residue that protects the transcript from 5'→3' exonuclease XRN1 (Hsu and Stevens, 1993) and interacts with additional proteins to influence translation and decay processes (Topisirovic et al., 2011). The cap also interacts with nuclear Cap-Binding Proteins 80/20 (CBP80/20) to influence pre-mRNA splicing (Izaurralde et al., 1994), export to the cytoplasm (Izaurralde et al., 1995), and translation (Gallie, 1991).

1.1.3 Splicing

Pre-mRNAs contain regions of sequence that will constitute the mature message that are frequently, with rare exceptions, interspersed with non-coding introns that must be spliced out. Splicing takes place co-transcriptionally in the nucleus and is performed by a large complex known as the splicesome (Nilsen, 2003). The splicesome recognizes conserved splice sites flanking regions to be removed from the message and two nucleophilic attacks are then carried out to remove the intron and join the two exons (Inoue et al., 1986; Query et al., 1994). At that point, a large multi-protein complex known as the Exon-Junction Complex (EJC) is deposited upstream of the splice site and acts as a mark of successful processing recognized by mRNA surveillance machinery in the nucleus and cytoplasm (Le Hir et al., 2001; Palacios et al., 2004). Exons may be alternatively spliced, increasing the number of proteins that can be encoded by a single gene. These splicing patterns can include or exclude entire exons or portions of individual exons. The differences in function performed by protein isoforms compared to their standard counterparts can range from subtle to dramatic and their expression is determined by an array of factors that influence splice site choice (Imamura et al., 2010; Martemyanov et al., 2008; Mott and Pak, 2012). Splicing patterns of several genes are altered in embryonic stem (ES) cells compared to cardiac precursor cells (Salomonis et al., 2009), although mechanisms leading to splice site selection are unknown.

1.1.4 Cleavage and polyadenylation

Processing of the 3' end of nascent mRNAs coincides with transcription termination and is the final step of maturation prior to export from the nucleus. Following transcription of the coding region, RNA Pol II continues to transcribe the 3'UTR encoding cleavage and polyadenylation signals (Montell et al., 1983; Zarkower et al., 1986). A conserved AAUAAA sequence and Downstream Element (DSE) bind Cleavage and Polyadenylation Specificity Factor (CPSF; Murthy and Manley, 1992), Cleavage Stimulation

Factor (CstF), and Cleavage Factors I and II (CFI and CFII; Takagaki et al., 1989). Poly(A) Polymerase (PAP) is recruited to the site of cleavage and adds 150-200 adenosine nucleotides (Bardwell et al., 1990; Sheets and Wickens, 1989). Several mRNAs in stem cells exhibit differential poly(A) tail status compared to differentiated cells, being either polyadenylated or non-adenylated, although the mechanisms governing mRNA poly(A) tail status are unknown (Yang et al., 2011).

1.1.5 Export and translation

mRNAs must be exported from the nucleus once they have reached maturation. Exit occurs through a bidirectional channel known as the Nuclear Pore Complex (NPC) that allows the passage of various macromolecules (Mattaj and Englmeier, 1998). The mRNA is heavily decorated with proteins by the time it has completed processing and many of these proteins act as marks of quality assurance that classify transcripts appropriate for export and translation (Schmid and Jensen, 2010). Spliced transcripts are exported from the nucleus more efficiently than unspliced mRNAs, suggesting that mRNA export and splicing are somehow linked (Valencia et al., 2008). In the cytoplasm, the cap of mRNAs is bound by Eukaryotic Translation Initiation Factor 4E (eIF4E). An interaction between eIF4E and Cytoplasmic Poly(A)-Binding Protein (PABPC) on the poly(A) tail bridged by eIF4G leads to transcript circularization and stabilization that stimulates translation (Figure 1.1[1]; Wells et al., 1998). Ribosomes assemble on the mRNA and synthesize nascent peptides until a stop codon signifying translation termination is reached (Bienz et al., 1981).

[1] Figures presented in Chapter 1 were adapted from those that appeared in: Neff *et al.* mRNA stability. In: Meyers, R.A., ed. Encyclopedia of Molecular Cell Biology, Wiley Blackwell, Chapter 4, 2013.

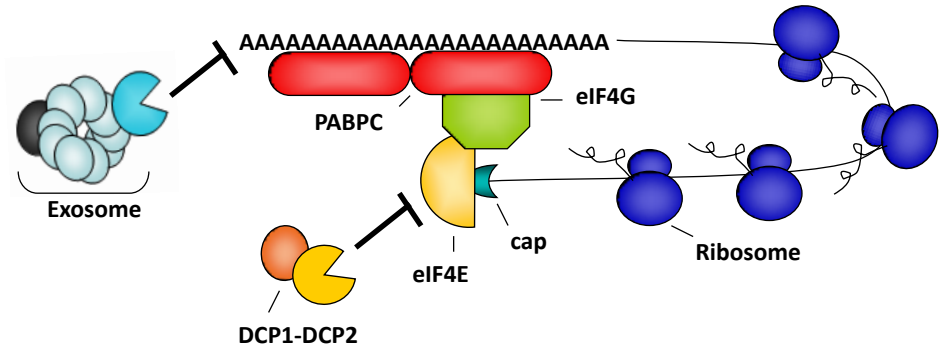


Figure 1.1 Closed-loop conformation of mRNA. Circularization of mRNAs are formed by an interaction between PABPC on the 3' poly(A) tail and eIF4E on the 5' cap bridged by eIF4G enhances translation and protects mRNA termini from exonucleases.

1.1.6 mRNA decay

Most eukaryotic mRNA decay occurs through the deadenylation-dependent pathway resulting in either 5'→3' or 3'→5' degradation (Figure 1.2). This process is initiated by removal of the poly(A) tail which serves to disrupt the closed-loop formation and render mRNA vulnerable to exonucleolytic activities of XRN1 in the 5'→3' direction or in the 3'→5' direction by the exosome or by DIS3L2 (Malecki et al., 2013). Several decay factors and enzymes are involved in the degradation of cellular transcripts and their roles are further discussed here.

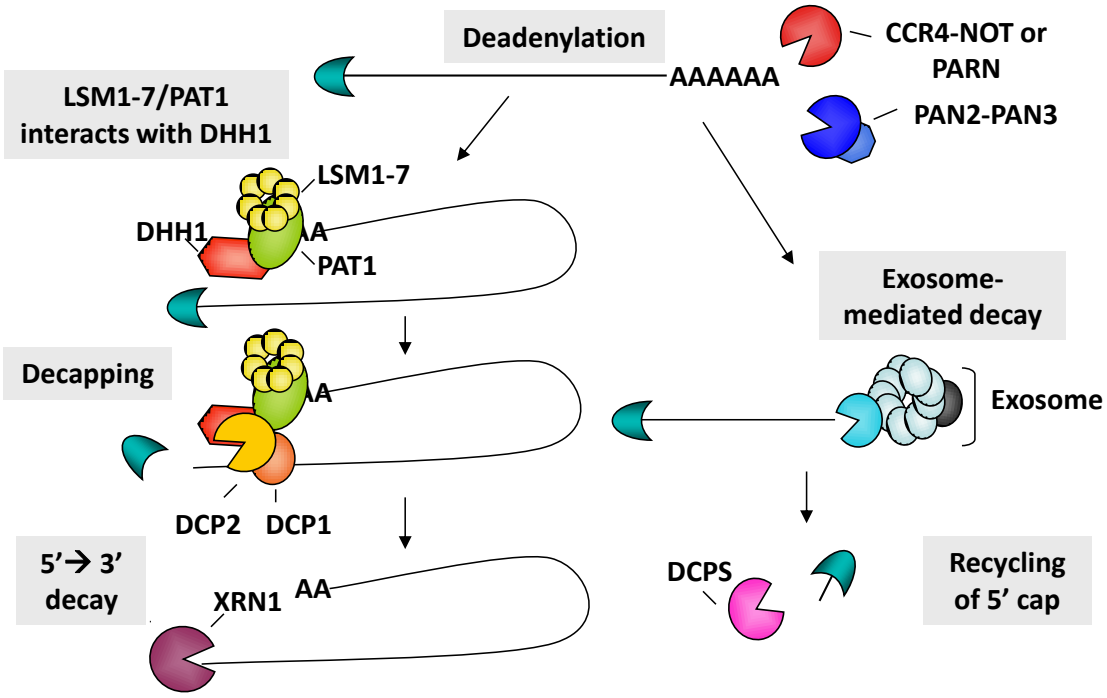


Figure 1.2 Deadenylation-dependent pathway of mRNA decay in eukaryotic cells. Decay is initiated upon removal of the poly(A) tail by a deadenylase. In 5'→3' decay, LSM1-7/PAT1 recruits the decapping complex DCP1/2, leaving the mRNA susceptible to decay by XRN1. In exosome-mediated decay, the exosome degrades deadenylated mRNAs in the 3'→5' direction and the cap is recycled by DCPS.

1.1.6.1 Deadenylation

Rapid shortening of the poly(A) tail is performed in a controlled manner by deadenylases that hydrolyze adenosine nucleotides in a Mg^{2+}-dependent fashion releasing 5'-AMP as the product (Beese and Steitz, 1991). The mammalian deadenylases identified to date include Poly(A)-specific Ribonuclease (PARN; Copeland and Wormington, 2001), PAN2 (Boeck et al., 1996), and CCR4-NOT subunits CNOT7 and CNOT8 (Bianchin et al., 2005; Daugeron et al., 2001) which are members of the RNase D deadenylase superfamily. RNase D-type proteins are in the DEDD superfamily that share three sequence motifs containing four invariant Asp and Glu amino acids surrounded by a region of conserved residues. The activity of these nucleases is catalyzed by a mechanism involving two metal ions (Steitz and Steitz, 1993). CCR4-NOT subunits CNOT6, CNOT6L (Mittal et al., 2011), 2' phosphodiesterase (2'PDE; Rorbach

et al., 2011), Angel (Dupressoir et al., 2001), and Nocturnin (Baggs and Green, 2003) deadenylases make up the Exonuclease-Endonuclease-Phosphatase (EEP) superfamily (Zuo and Deutscher, 2001). EEP-type deadenylases are related through conserved catalytic Asp and His residues within their exonuclease domains. These proteins are thought to have at least some functional redundancy as individual knockdown studies show little effect on cell viability (Mittal et al., 2011; Wahle and Winkler, 2013).

When mRNA is in a circular configuration, deadenylase PARN is blocked from gaining access to the poly(A) tail and 5' cap due to inhibition by eIF4E in the translation initiation complex (Ramirez et al., 2002). Upon removal of this complex, an interaction with the 5' cap stimulates PARN activity leading to removal of the poly(A) tail. PARN is the only deadenylase able to interact with both mRNA termini to influence tail shortening (Wu et al., 2005). PARN has an important role in maintaining particular poly(A) tail lengths during early development where mRNAs are kept translationally silent until polyadenylation and translation are induced (Copeland and Wormington, 2001). This status is maintained through an interaction between PARN and Cytoplasmic Polyadenylation Element Binding protein CPEB that becomes disrupted upon maturation (Kim and Richter, 2006). Although PARN appears to have a major role in higher eukaryotes, homologs do not exist in either *Saccharomyces cerevisiae* (Collart, 2003; Daugeron et al., 2001; Tucker et al., 2001) or *Drosophila melanogaster* (Grönke et al., 2009; Temme et al., 2010).

PAN2 functions as a dimer with PAN3 to trim the lengths of nascent mRNAs from lengths of approximately 200 nucleotides to about 30 nucleotides (Uchida et al., 2004). Like CCR4-NOT function, PAN2 activity is stimulated by the presence of PABPC (Khanna and Kiledjian, 2004). This shortened tail then acts as a substrate for the CCR4-NOT complex (Wahle and Winkler, 2013).

The CCR4-NOT complex is a very large multi-subunit protein whose comprising factors allow it to function in several diverse cell processes (Collart and Panasenko, 2012). Scaffold proteins CNOT1-5, CCR4-Associated Factors CAF130, and CAF40 function with two deadenylase components CNOT6 or

CNOT6L, and CNOT7, or CNOT8. CCR4-NOT can be recruited by a variety of factors to destabilize mRNAs. The RNA stability factor Tristetraprolin (TTP) also recruits the CCR4-NOT complex to target transcripts (Fabian et al., 2013; Sandler et al., 2011) while miRNA-associated protein GW182 directly interacts with the complex to mediate deadenylation (Fabian et al., 2011).

Significant reduction in length of the poly(A) tail during deadenylation leaves the transcript with a short oligo(A) tail that facilitates further degradation of the mRNA. Once the closed-loop formation has been fully disrupted, the transcript is susceptible to decay from both the 3'→5' and 5'→3' directions. Deadenylation-independent decapping is discussed in Section 1.1.6.4.

1.1.6.2 Exosome-mediated decay

Transcript degradation in the 3'→5' direction is mediated by a large, multi-subunit decay complex known as the exosome. In humans, the ~2 MDa cytoplasmic exosome is comprised of three S1-family proteins (EXOSC1-3) and six RNase PH-domain proteins (EXOSC4-9) that form a catalytically inactive core and two RNase D-like enzymes, DIS3 and EXOSC10 (Brouwer et al., 2001; Chlebowski et al., 2011; Mitchell et al., 1997). DIS3 gives the complex its catalytic properties. Exonucleolytic activity is conferred through an RNase II region (Dziembowski et al., 2007) while the PIN domain performs endonucleolytic reactions (Lebreton et al., 2008). The nuclear exosome includes an additional subunit, M-Phase Phosphoprotein 6 (MPP6; Milligan et al., 2008; Schilders et al., 2005). In the cytoplasm, the exosome is recruited to the 3' end of mRNAs destined for degradation through an interaction with SKI7, a component of the SKI complex which also contains the RNA helicase SKI2 (Araki et al., 2001; Houseley et al., 2006).

Exonuclease activity of the exosome requires a 3' OH and region of unstructured RNA which is often provided by the remnants of the poly(A) tail following deadenylation (Schmid and Jensen, 2008). The inactive core forms a ring-shaped structure through which mRNA is threaded while being degraded

(Tsanova and van Hoof, 2010). The crystal structure of the yeast nuclear exosome reveals that the pore unwinds RNA into a single stranded conformation and the DIS3 homologue catalyzes decay of nucleotides as the strand leaves the pore complex (Makino et al., 2013). When only a few nucleotides remain on the mRNA substrate, Scavenger Decapping enzyme DCPS cleaves and recycles the 5' cap (Liu et al., 2002). The yeast DCPS homologue Dcs1p is also able to influence decay from the 5' direction by acting as a co-factor for 5'→3' exonuclease XRN1 (van Dijk et al., 2003).

Interestingly, a paralogue of the enzymatic subunit DIS3 known as DIS3L2 does not associate with the exosome complex and functions in the cytoplasm (Malecki et al., 2013). Studies of DIS3L2 in mouse embryonic stem cells reveals that this 3'→5' exonuclease has a preference for 3' uridylated substrates.

1.1.6.3 Decapping and 5'→3' exoribonucleolytic decay

Following deadenylation, mRNAs may be decapped and degraded in the 5'→3' direction by the highly processive enzyme XRN1 (Coller and Parker, 2004; Muhlrad et al., 1994). The 3' end of deadenylated transcripts is recognized and bound by the ring-shaped LSM1-7/PAT1 complex (Tharun, 2009; Tharun et al., 2000). In conjunction with DEAD box helicase DHH1, the complex then recruits the Decapping Protein DCP1-DCP2 complex for removal of the 5' m^7G cap structure (Coller et al., 2001; Fischer and Weis, 2002). DCP2 is a member of the Nudix superfamily of hydrolases and is the catalytic member of the complex with co-activator DCP1 (Dunckley and Parker, 1999). DCP2 is the best studied decapping enzyme but other Nudix hydrolases, including NUDT16, also appear to have a role in regulating decapping of a distinct subset of mRNAs (Song et al., 2013). Once a transcript has been decapped, the 5' monophosphorylated mRNA becomes a substrate of 5'→3' exonucleolytic decay by XRN1 (Muhlrad et al., 1994).

As mentioned previously, the closed loop formation of mRNA protects the 5' cap from decapping therefore it is not surprising that cap-binding protein eIF4E and poly(A) tailing-binding protein PABPC negatively regulate decapping activity (Ramirez et al., 2002; Vilela et al., 2000). Conversely, there are also factors that stimulate decapping known as Enhancers of Decapping proteins EDC3 and EDC4/HEDLS/Ge-1 (Simon et al., 2006). These proteins, as well as LSM14-homologue SCD6, interact with the DCP1-DCP2 complex to encourage an active conformation (Harigaya et al., 2010; Ling et al., 2011). Interestingly, the decapping complex is also required for miRNA-mediated gene silencing by RNA-Induced Silencing Complex (RISC) component GW182, demonstrating interplay between these two post-transcriptional processes (Rehwinkel et al., 2005).

1.1.6.4 Deadenylation-independent decapping

Although the predominant mechanism of mRNA decay is deadenylation-dependent, there are also other minor pathways that are deadenylation-independent. An important class of mRNAs regulated in this manner are those encoding replication-dependent histones. Instead of a poly(A) tail, these mRNAs end in a conserved 3' stem loop. Regulation of histone mRNAs is discussed in detail in Chapter 4. Briefly, addition of an oligo(U) tract to the 3' end of histone messages recruits the LSM1-7/PAT1 complex, which in turn recruits DCP1-DCP2 followed by decapping and bi-directional decay in the 5'→3' and 3'→5' directions (Herrero and Moreno, 2011; Mullen and Marzluff, 2008).

Deadenylation-independent decapping also serves to rid the cell of aberrant mRNAs through nuclear surveillance pathways including Nonsense-Mediated Decay (NMD, Huang and Wilkinson, 2012), Non-Stop Decay (NSD; Klauer and van Hoof, 2012), and No-Go Decay (NGD; Harigaya and Parker, 2010). These mechanisms have been reviewed at length and will not be included in discussion here. However, much of the quality control degradation performed on these abnormal transcripts occurs through endonucleolytic cleavage.

1.1.6.5 Endonucleolytic cleavage

Endonucleolytic cleavage differs from deadenylation-dependent decay in that it instantly renders a transcript unusable rather than slowly inactivating it over time. Once the mRNA has been cleaved creating two fragments, the unprotected ends are subject to 5'→3' decay by XRN1 or 3'→5' decay by the exosome (Figure 1.3). Many endonucleases also exhibit sequence specificity rather than targeting universal mRNA features and appear to only regulate a subset of transcripts.

The SMG6 protein functions in the NMD pathway and cleaves mRNAs near the site of premature termination codons (Eberle et al., 2009) while RNase L and ZC3H12A function in pathways important for immune response. RNase L combats viral infection by rapidly cleaving both cellular and viral RNAs, often leading to apoptosis (Li et al., 1998). Like RNase L, ZC3H12A is induced during an infection and targets mRNAs encoding cytokines such as IL-6 (Schoenberg, 2011). Polysome-associated endonuclease PMR1 regulates a subset of mRNAs required for cell mobility (Bremer et al., 2003). Perhaps the best studied endonuclease of late is the Argonaute 2 (AGO2) protein. As discussed in Section 1.2.2 below, AGO2 is a component of RISC required for cleavage of targets following their base-pairing with the miRNA component (Liu et al., 2004; Meister et al., 2004).

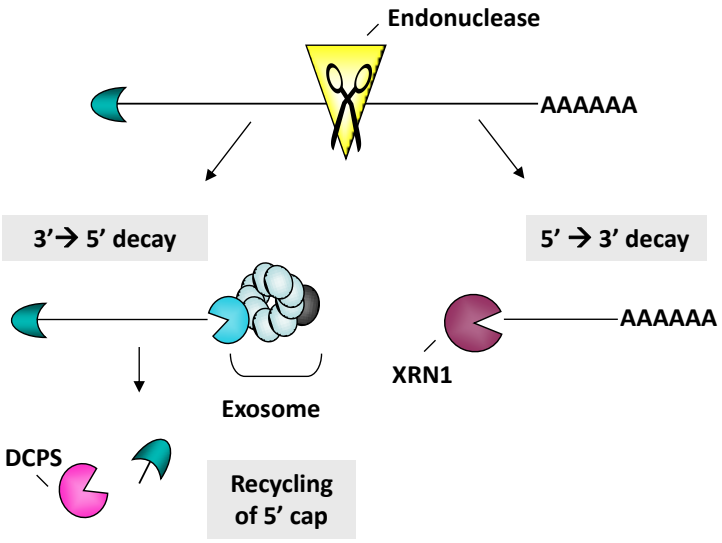


Figure 1.3 Endonucleolytic cleavage pathway. Degradation by endonucleolytic cleavage results in two RNA fragments subject to exosome-mediated decay 3'→ 5' decay and 5' → 3' decay by XRN1.

1.2 Regulated mRNA stability contributes to changes in gene expression

While degradation of mRNAs is likely very similar between cell types, differential regulation of mRNA stability factors such as RNA-binding proteins (RBPs) and miRNAs help define stem cell- or differentiation-specific expression patterns. The general mRNA features, a 5' m^7G cap and 3' poly(A) tail, are able to influence mRNA stability (as described in Section 1.1) but further control is demonstrated through *cis*-acting elements encoded within the transcript. *Cis*-acting elements can reside in the 5'UTR and coding region of an mRNA but they are more frequently found in the 3'UTR (Lee and Gorospe, 2011). Alternative polyadenylation can influence the presence of sequence elements in the 3'UTR (Lu and Bushel, 2013; Touriol et al., 1999), as demonstrated in mouse embryos where mRNA 3'UTRs are lengthened as development progresses (Ji et al., 2009). Usage of downstream poly(A) sites allows the 3'UTR to include additional elements that regulate mRNA stability in a developmentally-dependent fashion (Ji et al., 2009). RBPs and miRNAs interact with 3'UTR elements to either stimulate

or inhibit mRNA decay through a variety of mechanisms (Schoenberg and Maquat, 2012). Figure 1.4 below illustrates examples by which *trans*-acting factors recruit or inhibit mRNA decay machinery. The role of *cis*-acting elements and their associated *trans*-acting factors are discussed here, highlighting those important for stem cell gene expression.

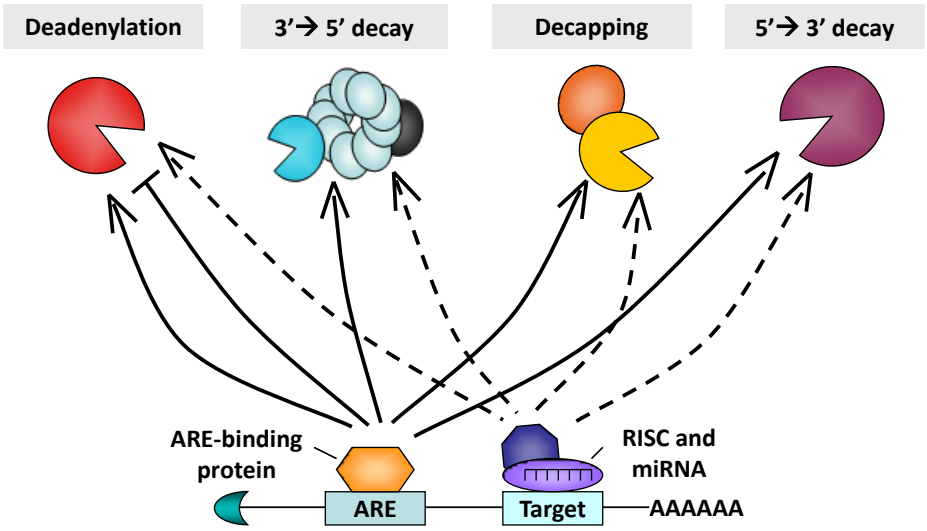


Figure 1.4 *Trans-acting* factors can either recruit or inhibit mRNA decay machinery. Binding of proteins to AU-rich elements (AREs) can result in stimulation or repression of deadenylation and also elicit decapping and exonucleolytic decay. miRNAs typically recruit rather than inhibit decay machinery.

1.2.1 Cis-*acting elements and their associated RNA-binding proteins*

Although mRNA stability elements and factors are widely studied, their influences specific to stem cell gene expression are largely uncharacterized. Regardless, several sequence motifs important for regulated mRNA decay have been identified throughout the transcriptome and their elucidation has strengthened our overall understanding of post-transcriptional control in gene expression. The best-characterized *cis*-acting element is the AU-rich element (ARE). AREs are 15-150 base sequences that contain one or more AUUUA motifs surrounded by U-rich context (Chen et al., 2001; Stoecklin et al.,

2003). These are generally found in the 3'UTR of mRNAs that require transient expression such as cytokines, transcription factors, and cell cycle modulators, representing up to 9% of all transcripts (Anderson, 2008; Gingerich et al., 2004). Most proteins that bind AREs (ARE-BPs) recruit deadenylases such as CCR4-NOT and PARN to the message to elicit destabilization. Well-characterized destabilizing ARE-BPs include the TTP family of CCCH-type ZNF proteins ZNF36, ZPF36L1, and ZFP36L2 (Lai et al., 2006), KHSRP (Gherzi et al., 2004), and AUF1 (hnRNP D; Li et al., 2009). Conversely, transcript stabilization is conferred wherein binding of ARE-BPs such as HuR (ELAVL1; Dean et al., 2001; Peng et al., 1998), TIA1, and TIAR (Gueydan et al., 1999) protect the mRNA from deadenylase activity.

GU-rich elements (GREs) are motifs that consist of UG-repeats or U-rich regions interspersed with guanine residues (e.g. UGUUUGU; Rattenbacher et al., 2010). Global functional analysis estimates that at least 5% of human genes contain GREs (Halees et al., 2011), with many being important for regulating transcription, nucleic acid metabolism, and developmental processes (Vlasova and Bohjanen, 2008). Many of these mRNAs bearing GU-rich elements are regulated through binding of CELF1, also known as CUGBP1 (Vlasova et al., 2008; Vlasova-St Louis et al., 2013). Destabilization of transcripts occurs in one mechanism through a direct interaction between CELF1 and PARN deadenylase (Moraes et al., 2006). Expression patterns of CELF1 in various tissues illustrate that this protein is developmentally regulated (Blech-Hermoni et al., 2013; Matsui et al., 2012). Interestingly, stabilizing RBP HuR has similar binding preferences as CELF1 and may compete for binding in some cases (López de Silanes et al., 2004), demonstrating the complexity of regulated mRNA decay.

Pyrimidine-rich or C-rich elements are among the least understood of *cis*-acting elements identified to date but their role in mRNA stability upon binding of Poly(C)-Binding Proteins (PCBPs) has been extensively described for human *α-globin* mRNA (Kong et al., 2006) and collagen I (*COL1A1)* mRNA (Czyzyk-Krzeska and Bendixen, 1999). Interestingly, PCBP1 is essential for maintaining a transcriptionally-silent state in mouse oocytes (Xia et al., 2012) and PCBP4 is a tumor suppressor (Zhu

and Chen, 2000) whose down-regulation likely supports pluripotency. The involvement of C-rich elements and PCBPs in regulated mRNA decay is discussed in more detail in Chapter 6.

Although stem cell-specific *cis*-acting elements and *trans*-acting factors have yet to be fully characterized, a recent study found that more than 200 RBPs are over-expressed in mouse ES cells compared to differentiated cells, although whether these proteins modulate mRNA stability has yet to be fully determined (Kwon et al., 2013). Regardless, many of the validated proteins were not previously annotated as RBPs, suggesting that there are more RNA-protein interactions that regulate stem cell gene expression than previously appreciated. LIN28 is a well-known, developmentally-regulated RBP (Richards et al., 2004) that binds pluripotency transcription factor *OCT4* mRNA in human ES cells to maintain stem cell gene expression (Qiu et al., 2010). LIN28 is also a negative regulator of miRNA let-7 biogenesis to repress differentiation (Viswanathan et al., 2008) and is likely one of several RBPs with essential roles in stem cell maintenance.

1.2.2 MicroRNAs

MicroRNAs are small non-coding RNAs ~22 nucleotides in length that interact with mRNAs to regulate the expression for as many as 30% of human genes (Neilson et al., 2007). Overall expression of miRNAs is lower in early development but levels of unique subsets are also regulated in cell-specific patterns (Lakshmipathy et al., 2007; Smith-Vikos and Slack, 2012; Zhao and Srivastava, 2007). Like RBPs, this differential expression allows miRNAs to facilitate maintaining transcriptome profiles that support specific cell functions, such as pluripotency.

MicroRNAs are derived from sequences encoded in introns or portions of primary mRNAs and biogenesis requires several proteins that process miRNAs into their mature form. Primary miRNA (pri-miRNA) transcripts are bound by DGCR8, the RNA-binding competent of the Microprocessor (Gregory et al., 2004; Han et al., 2004). Interestingly, DGCR8 expression is required for silencing of self-renewal

during differentiation of ES cells (Wang et al., 2007). The catalytic component, Drosha, then cleaves the pri-miRNA to a reduced length. The fragment is exported to the cytoplasm where it is further processed by DICER to yield the final ~22 nucleotide length (Lee et al., 2003). Mature miRNAs are incorporated into RISC which contains core proteins AGO2 (Hammond et al., 2001) and GW182 (Liu et al., 2005). The sequence of the miRNA guides RISC to the complementary binding site on target mRNAs. Perfect complementarity of sequences activates the AGO2 component of RISC leading to endonucleolytic cleavage ad subsequent bi-directional degradation by XRN1 and the exosome (Figure 1.5; Meister et al., 2004). However, this mechanism is more prevalent in plant genomes and infrequent in humans (Bracken et al., 2011; Karginov et al., 2010; Shin et al., 2010). Imperfect base-pairing results in activation of the GW182 component of RISC which interacts with deadenylases to stimulate shortening of the poly(A) tail and induce mRNA silencing (Braun et al., 2011). GW182 serves as a docking platform to guide deadenylase complexes PAN2-PAN3 and CCR4-NOT to their miRNA targets (Braun et al., 2011; Fabian et al., 2011).

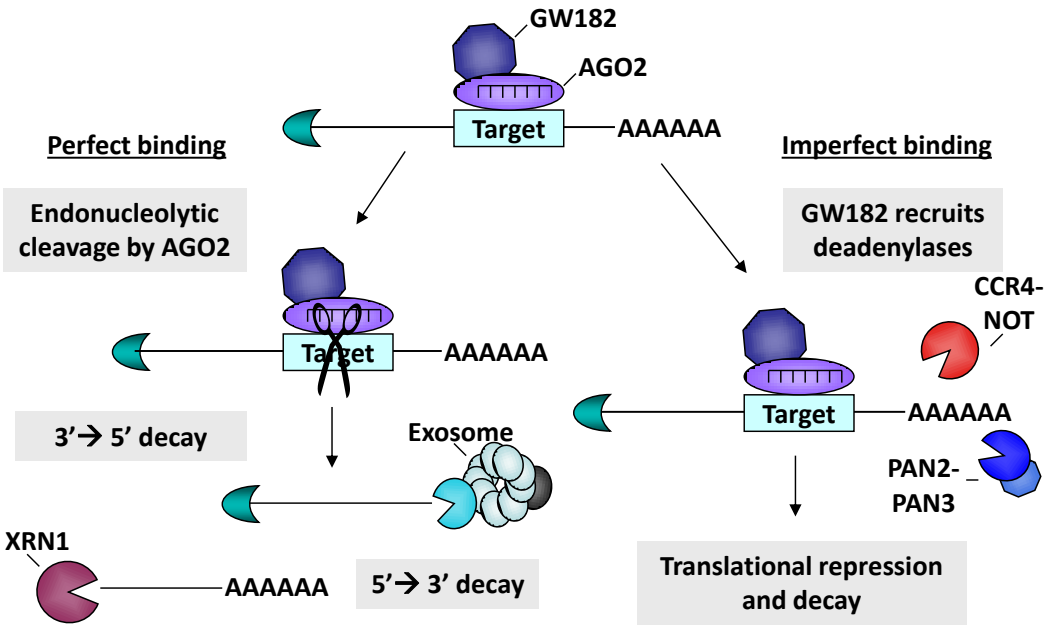


Figure 1.5 MicroRNA-mediated regulation of mRNA stability. A complex (RISC) containing AGO2 and GW182 is guided to target mRNAs by miRNAs. The complementarity of binding determines whether the transcript undergoes endonucleolytic cleavage and degradation or translational repression. In humans, translational repression occurs more frequently than cleavage.

Expression of GW182 is important during early embryo development in *Drosophila* (Schneider et al., 2006) while proper AGO2 expression is also required for various stages of development (Alisch et al., 2007; Deshpande et al., 2005; Lykke-Andersen et al., 2008), illustrating the importance of the miRNA pathway for regulating gene expression. Further, several miRNAs have been implicated in the regulation of pluripotent gene expression including the miR-290-295 and miR-302/367 families that are over-expressed in stem cells (Kaspi et al., 2013; Lüningschrör et al., 2012) and the let-7 and miR-200 families that are down-regulated (Peter, 2009; Porciuncula et al., 2013). Many miRNAs have also been used to increase reprogramming efficiency upon induction of pluripotency (Bao et al., 2013; Leonardo et al., 2012; Wang et al., 2013). These examples demonstrate the ability of miRNAs to control global changes in gene expression and identifying additional pluripotency-associated miRNA families and their mRNA targets would be insightful to stem cell biology.

1.3 Rationale and hypothesis

As discussed in this chapter, the contribution of mRNA decay to pluripotency is largely unknown, despite the notable importance of post-transcriptional control in regulated gene expression. By investigating mRNA decay in human iPS cells compared to the fully differentiated cells they were derived from, we can add to our limited understanding of RNA stability mechanisms that regulate stem cell gene expression. Specifically, we hypothesized that determination of global mRNA half-lives in iPS cells and HFFs (Human Foreskin Fibroblasts) would reveal differentially regulated mRNAs that may be coordinately controlled and we hoped to identify novel regulatory mechanisms that are specific to either stem cells or differentiated cells. We expected that these results would further our understanding of pluripotency and possibly provide additional means to enhance reprogramming.

Chapter 2: Materials and methods

2.1 Cell culture

2.1.1 Cell line maintenance

2.1.1.1 Human foreskin fibroblasts

Human Foreskin Fibroblasts (HFFs) were purchased in a matched set with induced Pluripotent Stem (iPS) cells from System Biosciences and cultured according to company recommendations. Fibroblast culture medium included Dulbecco's Modified Eagle Medium (DMEM, Cellgro) supplemented with 10% Fetal Bovine Serum (FBS, Atlas), 2 mM L-glutamine (Hyclone), 0.1 mM non-essential amino acids (Cellgro), 50 U/mL penicillin, and 50 µg/mL streptomycin (Hyclone). Cells were incubated at 37°C with 5% CO_2 and passaged when 90% confluent at a split ratio of 1:3 using 0.25% trypsin-EDTA (Hyclone) in Phosphate Buffered Saline (PBS, Hyclone). Cryovial cell stocks were made in 50% fibroblast culture medium, 40% FBS, and 10% DMSO (Sigma). Figure 2.1 below is an image of HFF cells at ~80% confluence under 100x magnification on a Diaphot 200 microscope (Nikon) with CoolSNAP camera and RS Image software (Roper Scientific).

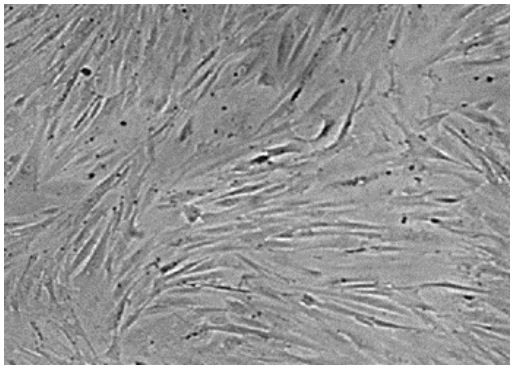


Figure 2.1 Cultured human foreskin fibroblasts. Human Foreskin Fibroblasts (HFFs) at ~80% confluence viewed at 100x magnification.

2.1.1.2 Mouse embryonic fibroblasts

Mouse Embryonic Fibroblasts (MEFs) were purchased from American Type Culture Collection (ATCC) and maintained using the fibroblast protocol described in Section 2.1.1.1. To prepare cells for use as a non-dividing feeder layer, MEFs were grown to confluence in a T75 flask and treated with 50 µL of 1 mg/mL mitomycin C (Sigma) in 6 mL of fibroblast medium for 3 hours at 37°C to effectively halt cell cycle progression. Following treatment, cells were seeded at a density of $1.5x10^5$ cells per well onto 6-well plates coated with 0.1% gelatin (Andwin Scientific). Alternatively, arrested cells were frozen in 50% fibroblast medium, 40% FBS, and 10% DMSO in aliquots of $9x10^5$ cells, enough for all the wells of a single 6-well plate. Feeder-layer MEF cells were allowed to adhere overnight and were ready for iPS cell plating the following day. Figure 2.2 shows a MEF feeder layer on 0.1% gelatin at 100x magnification.

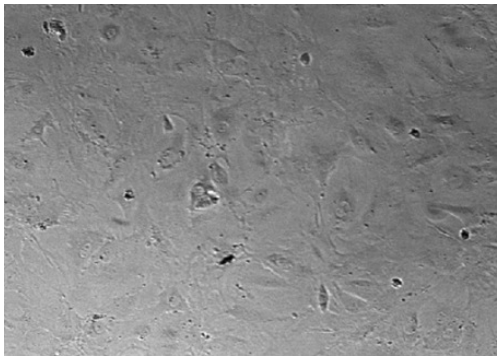


Figure 2.2 Cultured mouse embryonic fibroblasts. Mouse Embryonic Fibroblast (MEF) feeder layer on 0.1% gelatin viewed at 100x magnification.

2.1.1.3 Human induced pluripotent stem cells cultured on MEF feeder layer

Stocks of iPS cells reprogrammed from HFFs were purchased from System Biosciences (SC101A-1, Lot #090725-08). The reprogramming was achieved using lentiviral expression of pluripotency genes OCT4, SOX2, c-MYC, and KLF4, the same cocktail used in Takahashi et al., 2007. The iPS cell line was certified as pluripotent by the company prior to purchase through staining of stem cell markers NANOG,

OCT4, SSEA3, and TRA-1-60. To establish the stock cryovial of $2x10^5$ iPS cells, cells were seeded at a density of approximately $5x10^4$ cells per well of a MEF feeder layer 6-well plate in medium containing DMEM/F12 (Cellgro) supplemented with 20% Knockout Serum Replacement (Gibco), 2 mM L-glutamine, 0.1 mM non-essential amino acids, 0.1 mM 2-mercaptoethaol, 10 ng/mL recombinant human Basic Fibroblast Growth Factor (bFGF, Gibco), 50 U/mL penicillin, and 50 µg/mL streptomycin (complete DMEM/F12 medium). Knockout Serum Replacement is a defined, serum-free formulation that replaces FBS in stem cell culture medium to maintain pluripotency. Cells were incubated at 37°C with 5% CO_2 and medium was changed daily. Cells were split when the edges of colonies became close in proximity. Passaging of iPS cells was performed at a split ratio of 1:10 – 1:20 using StemPro Accutase (Invitrogen) and DMEM/F12 supplemented with 15 mM HEPES (pH 7.0), 2 mM L-glutamine, 50 U/mL penicillin, and 50 µg/mL streptomycin (DMEM/F12). Cryovial cell stocks were made in 50% complete DMEM/F12 medium, 40% FBS, and 10% DMSO. Figure 2.3 below shows iPS cell colonies cultured on a MEF feeder layer at 100x magnification.

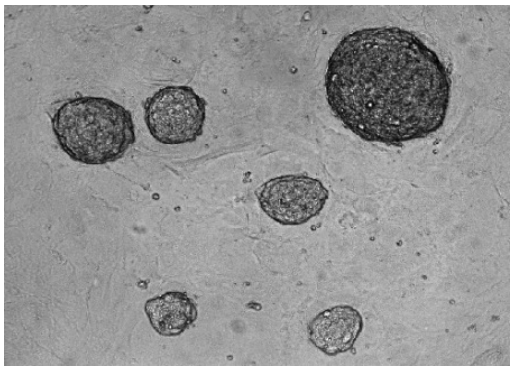


Figure 2.3 Cultured induced pluripotent stem cells on feeder layer. Induced Pluripotent Stem (iPS) cells on a feeder layer viewed at 100x magnification.

2.1.1.4 Human induced pluripotent stem cells cultured on Matrigel (feeder-free)

To eliminate contamination with the feeder layer MEFs, iPS cells were adapted to feeder-free conditions following establishment and expansion. For the initial adaptation, cells were seeded at a high density onto plates coated with 0.3 mg/mL Matrigel Basement Membrane Matrix Growth Factor Reduced (Matrigel, BD Biosciences) diluted in DMEM/F12. Cell loss following plating typically ranged from 50 – 75% and a recovery time of approximately 5 – 7 days was expected before noticeable formation of distinct colonies. Cells were cultured in mTeSR1 medium (Stemcell Technologies) supplemented with 50 U/mL penicillin and 50 µg/mL streptomycin. Passaging was performed using Dispase Solution (Stemcell Technologies) at a concentration of 1 mg/ml in 1x Hank's Balanced Salt Solution (Gibco). Cryovial cell stocks were made in 50% mTeSR1 medium, 40% FBS, and 10% DMSO. As depicted in Figure 2.4, iPS cell colony morphology is slightly different when cultured on Matrigel as compared to those cultured on a feeder layer. Cells are more columnar and colonies tend to be flatter. This difference in colony appearance is expected according to the mTeSR1 technical manual.

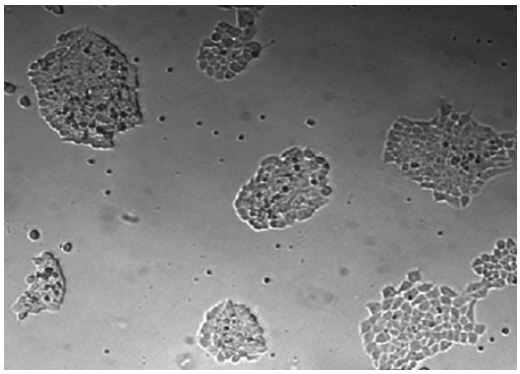


Figure 2.4 Cultured induced pluripotent stem cells on Matrigel. Induced Pluripotent Stem (iPS) cells on Matrigel viewed at 100x magnification.

To maintain cells in a pluripotent state, medium was changed daily and cells were passaged before colony borders began to merge. As differentiation often begins in cells skirting the colonies, it

was important to sufficiently treat cells with Dispase which slowly detaches colonies from Matrigel by first loosening the edges. When cells on the edges were lifted with the majority of the colony intact, thorough rinsing of wells with DMEM/F12 ensured removal of these border cells with the most likelihood of encountering differentiation signals. After washing, adherent colonies were pipetted more aggressively several times in DMEM/F12 to detach and reduce colony size, lessening potential for embryoid body formation.

After iPS cells were adapted to feeder-free conditions on Matrigel, total RNA was collected from culture dishes to determine whether MEF cells had been removed. To do this, cDNA was made and analyzed by PCR (as described in Section 2.2) for the presence of a mouse-specific product (*mRhoC*). While humans also express *RhoC* mRNA, these primers (generously provided by Dr. Jerome Lee) were designed to specifically detect the mouse homologue and do not amplify non-specific products. Using C2C12 mouse myoblasts as a positive control, we saw that although *mRhoC* was readily detectable in iPS cell samples collected from colonies cultured on a feeder layer (Figure 2.5A), PCR products were not amplified using any of four cultures adapted to Matrigel (Figure 2.5B). In contrast, human *GAPDH* was detected in cDNA from iPS cells under both culture conditions. GAPDH was chosen as the reference gene for all PCR experiments as it is a housekeeping gene with a relatively long half-life whose expression is fairly equal between each of the cell types. All experiments described throughout this thesis were performed using iPS cells cultured on Matrigel.

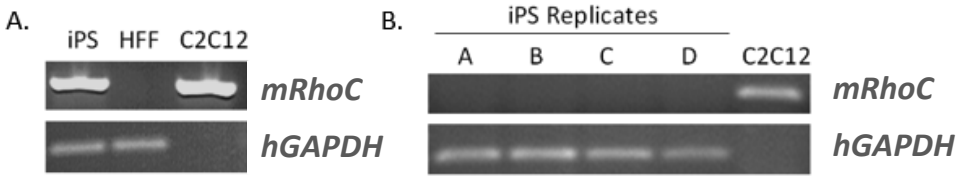


Figure 2.5 Feeder layer contamination is removed following adaption to feeder-free conditions. RT-PCR analysis of cDNA prepared from iPS cells cultured on A) a MEF feeder layer or B) Matrigel for detection of mouse (*mRhoC*) contaminants on a 1% agarose gel visualized by ethidium bromide staining and UV exposure.

2.1.1.5 HeLa cells

HeLa cells were used in some experiments for comparison to the human primary fibroblast and stem cell lines. These cells were incubated at 37°C with 5% CO_2 and cultured in DMEM supplemented with 10% FBS, 2mM L-glutamine, 100 U/mL penicillin, and 100 µg/mL streptomycin. Cells were passaged at 70 – 95% confluence at a split ratio of 1:5 – 1:20 using 0.25% trypsin-EDTA and PBS. Cryovial cell stocks were made using 50% culture medium, 40% FBS, and 10% DMSO.

2.1.2 Cell cycle analysis

To measure the percentage of cells in S phase in an asynchronous population, HFFs were cultured on chamber slides and iPS cells were cultured on chamber slides coated with 0.3 mg/mL Matrigel. Cells were treated with 10 µM Bromodeoxyuridine (BrdU, BD Biosciences) in culture medium for 30 minutes at 37°C then fixed with 70% ethanol in 50 mM glycine (pH 2.0) for 30 minutes at -20°C. To expose DNA for staining, cells were treated with 1.5 M HCl for 30 minutes at room temperature. After rinsing, slides were blocked in 3% Bovine Serum Albumin (BSA)/0.5% Tween-20 in PBS for 1 hour at room temperature. Anti-BrdU (BU-1, Thermo Scientific) diluted 1:500 in 1% BSA/PBS was applied to the cells and left overnight at 4°C. Following rinsing with PBS, secondary antibody Alexa Fluor 594 goat anti-mouse IgG (Invitrogen) was applied at a concentration of 1:2,500 in 1% BSA/PBS for 2 hours at room temperature. Slides were then mounted using ProLong Gold antifade reagent with DAPI (Life Technologies Corporation) and visualized using an Olympus IX71 microscope. Positive staining with anti-BrdU identified cells in S phase and the percentage was determined based on the total number of cells counted. Three replicates of slides were analyzed for each cell line and calculations were based on counting ~100 cells per slide.

2.1.3 Growth rate measurement

The growth rates of HFF and iPS cells were determined by measuring the number of cells in each population over a period of five days. To do this, cells were seeded into five dishes at an equal seed density. Each day, cells were detached and suspended into a single-cell suspension using 0.25% trypsin-EDTA and PBS. For iPS cells, 0.05% trypsin and pipetting was necessary to diffuse the colonies. The total number of live cells collected for each of three replicates per day was determined following 0.4% Trypan Blue Solution (Cellgro) staining using a hemocytometer.

2.2 RNA preparation and assays

2.2.1 RNA isolation and quantification

Total RNA was collected from cells using TRIzol (Invitrogen) according to the manufacturer's protocol. Briefly, RNA was extracted from TRIzol following addition of chloroform and transfer of the upper aqueous phase. Nucleic acids were precipitated in isopropanol using glycogen (Fermentas) as a carrier and then treated with DNase1 (Fermentas) for 20 minutes at 37°C. Following phenol:chloroform extraction, RNAs were precipitated by the addition of 0.33 volumes 10 M ammonium acetate and 2.5 volumes 100% ethanol using glycogen as a carrier. After drying, RNA pellets were resuspended in RNase-free distilled water (dH_20) and concentration was measured on a NanoDrop 2000c spectrophotometer (Thermo Fisher Scientific). For routine RNA preparations, samples were then diluted to 1 µg/µl in RNase-free dH_20.

2.2.2 Measuring RNA per cell

To estimate the amount of total RNA per cell, HFF and iPS cells were collected and counted using a hemocytometer. Total RNA was isolated from a known number of cells using TRIzol and quantification on a NanoDrop 2000c spectrophotometer allowed calculation of µg RNA/cell.

2.2.3 Reverse transcription

Complementary DNA (cDNA) was made using random hexamers (Integrated DNA Technologies Inc) and ImProm-II Reverse Transcriptase (Promega) following the manufacturer's protocol. Briefly, 1 µg of total RNA was annealed to 500 ng random hexamer in a 5 µL reaction at 70°C for 5 minutes followed by rapid cooling to 4°C. Reverse transcription was then carried out by adding 15 µL of a prepared mixture containing 4 µL 5x Reaction Buffer (250 mM Tris-HCl (pH 8.3), 375 mM KCl, and 50 mM DTT), 1 µL 10 mM dNTPs, 2.4 µL 25 mM $MgCl_2$, 5.6 uL RNase-free dH_20, 1 µL RiboLock RNase Inhibitor (Fermentas), and 1 µL ImProm-II Reverse Transcriptase and incubated in a thermal cycler at 25°C for 10 minutes, 50°C for 60 minutes and 70°C for 15 minutes.

2.2.4 Semi-quantitative PCR

Semi-quantitative PCR was carried out using the GoTaq Flexi Polymerase kit by Promega. In a 25 µL reaction, 2 µL of cDNA was mixed with 5 µL 5x Green GoTaq Reaction Buffer, 0.5 µL forward primer, 0.5 µL reverse primer, 0.5 µL 10 mM dNTPs, 2 µL 25 mM $MgCl_2$, 14 µL dH_20 and 0.5 µL Taq polymerase. Amplification was performed in a Gene Amp PCR System 2700 thermal cycler (Applied Biosystems) and PCR products were visualized on 1-2% agarose gels containing EtBr with GeneRuler 1kb DNA Ladder or GeneRuler DNA Ladder Mix (Thermo Fisher Scientific). Quantification of band intensities was performed using ImageQuant software (GE Healthcare). The primers used for semi-quantitative PCR are listed in Table 2.1 below.

Table 2.1 List of primers used for RT-PCR experiments.

Primer name	Gene ID	Gene	Forward sequence	Reverse sequence	Annealing	Elongation	Cycles
mRhoC F/R	11853	Rhoc	AAAGCTTCCTCAACCCTCCCA	CTACCCAAAGCAGAAACCCCCA	50°C	1 min	32
GAPDH F/R	2597	GAPDH	AAGGTGAAGGTCGGAGTCAA	AATGAAGGGGTCATTGATGG	60°C	1 min	32
PCBP4 Exon 9 F/R	57060	PCBP4	CTCCTGCAAATGGTGGAAAT	TCGGATGGTAGGGGATAGTG	60°C	30 sec	30
PCBP4 Exon 15 F/R	57060	PCBP4	TGTCACCATCACTGGCTCTC	CAGCCTTCTTGCTCCCATTA	60°C	30 sec	30

2.2.5 Quantitative PCR

To perform quantitative PCR (qPCR), a 3x master mix was prepared for each reaction using 15 µL of iQ SYBR Green Supermix (Bio-Rad), 1.2 µL of each primer (2.5mM), 1.2 µL of dH_2O and 2.4 µL of cDNA. Reactions were divided into three 10 µL aliquots for three replicates. Thermal cycling was performed using the CFX96 Real-Time PCR Detection System and analyzed using CFX Manager software (Bio-Rad). All primer pairs were standardized using 5- or 10-fold cDNA serial dilutions to allow determination of PCR efficiency. All primers were required to meet the following standards: PCR efficiency was between 80-120%, the correlation coefficient of the standard curve was >0.98, and produced a single product of the correct size as visualized by agarose gel electrophoresis. For relative abundance data, target gene expression was normalized to reference gene *GAPDH* mRNA. To calculate this output, the CFX Manager software applied the Pfaffl Method (Pfaffl, 2001) to determine expression relative to the reference gene, taking into account individual primer amplification efficiencies.

Primers were designed and chosen using various web-based programs such as Primer3 Input (Version 0.4.0; http://frodo.wi.mit.edu/), Primer_BLAST (http://www.ncbi.nlm.nih.gov/tools/primer-blast/), and qPrimerDepot (http://primerdepot.nci.nih.gov/). Primer sets were designed using parameters for GC content of 40-60%, annealing temperature between 58°C and 60°C, and length of ~20 nucleotides. Sequences used as input for primer design corresponded to the open reading frame of each transcript. As DNAse treatment was routinely performed on isolated total RNA samples, primers were not specifically designed to span introns. The primers used for qPCR are listed in Table 2.2 below. For histone mRNA detections, primer sets that recognize core histones H2A, H2B, H3, and H4 and linker histone H1 were generously provided by Dr. Paul Laybourn and detect multiple transcripts within each histone family (Bogenberger and Laybourn, 2008), therefore a specific gene ID was not provided for these oligos.

Table 2.2 List of primers used for qRT-PCR experiments.

Primer name	Gene ID	Gene name	Forward sequence	Reverse sequence	Efficiency (%)
DGCR8_3 F/R	54487	DGCR8	CAAGCAGGAGACATCGGACAAG	CACAATGGACATCTTGGGCTTC	108.1
DUSP7_2 F/R	1849	DUSP7	CTGAGGCCATCAGCTTCATT	TAGGCGTCGTTGAGTGACAG	115.4
FOS_1 F/R	2353	FOS	GTGGGAATGAAGTTGGCACT	CTACCACTCACCCGCAGACT	95.8
GAPDH F/R	2597	GAPDH	AAGGTGAAGGTCGGAGTCAA	AATGAAGGGGTCATTGATGG	112.1
H1 F/R	N/A	H1	CCGGTGTCCGAGCTCATTACTAAA	GCTTTCTTGAGAGCGGCCAAAGAT	99.5
H2A F/R	N/A	H2A	AGCTCAACAAGCTTCTGGGCAA	TTGTGGTGGCTCTCGGTCTTCTT	110.7
H2B F/R	N/A	H2B	TGCGCCCAAGAAGGGTTCTAAA	ACGAAGGAGTTCATGCCCA	117.5
H3 F/R	N/A	H3	TGCTCATCCGCAAACTGCCATT	AGTGACACGCTTGGCGTGAATA	105.5
H4F/R	N/A	H4	ACCGTAAAGTACTGCGCGACAA	TTCTCCAGGAACACCTTCAGCA	98.4
LATS2_3 F/R	26524	LATS2	AAGAGCTACTCGCCATACGCCTTT	AGCTTTGGCCATTTCTTGCTCCAG	94.2
PCBP1_1 F/R	5093	PCBP1	GTAACGAGCCCAACTCCCCCGA	CCTCCGAGATGTTGATCCGCGC	93.8
PCBP2_1 F/R	5094	PCBP2	GCTGCACCAGTTGGCAATGCAA	AGCCTTTCACCTCTGGAGAGCTGG	107.9
PCBP3_3 F/R	54039	PCBP3	TATAGCCTGCTTATGGGGGA	AGGCCACCTTCTGAGACCTT	95.9
PCBP4_1 F/R	57060	PCBP4	GGCGAGACTGTAAAGCGAAT	TGGAGACTGCATGGAAGACA	111.7
TOB2_1 F/R	10766	TOB2	CACCCTGGAGGGAGAAGC	AGAGAATCAGCACAGGGCAC	108.5
TUT1_1 F/R	64852	TUT1	TACTTCCAGACATTTGGCCC	CTGTGCTTGGGCTGTGATAA	98.6
WEE1_4 F/R	7465	WEE1	ATTTCTCTGCGTGGGCAGAAG	CAAAAGGAGATCCTTCAACTCTGC	108.7
ZNF134_1 F/R	7693	ZNF134	GAGTCATGCCCTCTCAGGC	CTACGAACTGTGATGGCGG	108.2
ZNF43_1 F/R	7594	ZNF43	TGGCCAAAAGTCTTGGGTAA	TGATCACCTGTCTGGAGCAA	114.2
ZNF627_2 F/R	199692	ZNF627	TTGAGGATGTGGCTGTGAAC	CAGAAGCCAGGTTCCTGAAG	97.8

2.2.6 Measuring mRNA half-life

To determine mRNA half-lives, cells were treated with 5 µg/mL actinomycin D (Invitrogen) starting 20 minutes before the collection of the 0 minute time point. Cells were collected in TRIzol at 0, 15, 30, 60, 120 and 240 minutes after the 20 minute pre-incubation. Total RNA was isolated in TRIzol and cDNA was made as described in Sections 2.2.1 and 2.2.3. Quantitative PCR was used to measure the abundance of specific mRNAs at each time point. The mRNA abundances were then fit to an exponential curve to provide a decay rate constant (k) and a half-life was derived based on the equation $T_{1/2} = \ln(2)/k$.

2.2.7 Preparation and hybridization of RNAs for microarrays

For the microarrays, transcription shut-off was performed in four replicates for each cell line. To confirm that transcription had been effectively inhibited, decay rates of *FOS* and *TUT1* mRNAs were measured using *GAPDH* mRNA as a reference gene. After appropriate half-lives were calculated, the replicates were compared to one another to choose the three that gave the lowest standard deviation. These were then sent to Erin Petrilli at the CSU Genomics and Proteomics Core who used a Bioanalyzer (Agilent) to verify RNA quality and then processed the samples for hybridization to 18 Affymetrix Human Gene 1.0 ST microarrays (one per time point per replicate). Datasets were sent to collaborators at Rutgers University – New Jersey Medical School where Dr. Ju Youn Lee performed bioinformatic analysis under the mentorship of Dr. Bin Tian.

2.2.8 Inhibition of DNA synthesis using hydroxyurea

To measure half-lives of histone mRNAs following the end of S phase, cells were treated with 5 μg/mL of hydroxyurea (HU, Sigma) in complete medium and incubated at 37°C with 5% CO_2. Thirty minutes following the addition of HU, the 0 minute time point was collected in TRIzol. Additional samples were collected at 15, 30, 60, 120 and 240 minutes. RNA isolation and qRT-PCR analysis of half-life was performed as previously described in Sections 2.2.1, 2.2.3, 2.2.5, and 2.2.6.

2.2.9 Measuring microRNA abundance

Measurement of microRNA (miRNA) abundance by qRT-PCR was performed as previously described (Git et al., 2010). Approximately 1 μg of total RNA was polyadenylated in a 25 μL reaction using the Poly(A) Tailing Kit (Ambion) according to the manufacturer's protocol. After phenol:chloroform extraction, cDNA was made using 500 ng of Poly(T) adapter (Integrated DNA Technologies Inc) and ImProm-II Reverse Transcriptase. The resulting cDNA as treated with RNase H

(Fermentas) at 37°C for 1 hour and 12.5 ng was added to each reaction for qPCR analysis. MicroRNA abundances were measured using a universal reverse primer and miRNA-specific forward primers and normalized to human *5S* rRNA. MicroRNA targets were chosen based on those identified in Schnall-Levin et al. (2011) and primer sequences were provided in the Supplemental Materials of Git et al. (2010). Oligonucleotide sequences used for quantification of miRNA abundance are listed in Table 2.3 below.

Table 2.3 List of oligonucleotides used for quantification of miRNAs by qRT-PCR.

Oligonucleotide	**Sequence**	**Efficiency (%)**
Poly(T) adapter	GCGAGCACAGAATTAATACGACTCACTATAGGTTTTTTTTTTTT(A/G/C)(A/G/C/T)	
Universal Reverse	GCGAGCACAGAATTAATACGACTC	
Human 5S rRNA	ACCGGGTGCTGTAGGCT	98.5
hsa-miR-23a	ATCACATTGCCAGGGATTTCC	100.2
hsa-miR-23b	ATCACATTGCCAGGGATTACCAC	105.5
hsa-miR-181a	AACATTCAACGCTGTCGGTGAGT	113.6
hsa-miR-181b	AACATTCATTGCTGTCGGTGGGTT	105.1
hsa-miR-181d	AACATTCATTGTTGTCGGTGGGT	109.4
hsa-miR-199a-5p	CCCAGTGTTCAGACTACCTGTT(C)	93.8

2.3 Protein preparation and assays

2.3.1 Preparation of whole cell lysates

Whole cell lysates were prepared by washing cells twice in PBS and lysing in Radio Immunoprecipitation (RIPA) buffer (50 mM Tris-Cl (pH 7.4), 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, 1% Na-deoxycholate, 0.1% SDS, 1 mM PMSF and 1x Protease Inhibitor Cocktail Tablets (Roche)). Extracts were sonicated three times for 3 seconds each with resting on ice for 30 seconds and insoluble material was removed by centrifugation at maximum speed, or 16,100 x g, for 5 minutes at 4°C.

Protein concentrations were measured using the Bio-Rad Protein Assay (Bio-Rad) according to the manufacturer's protocol. Briefly, 1 mL of reagent diluted 1:5 in dH_2O was added to 20 µL of lysate

diluted 1:10 in dH_2O. After 10 minutes of incubation at room temperature, the absorbance was measured at 595 nm wavelength on an Ultrospec 2000 spectrophotometer (Pharmacia Biotech). A standard curve was generated using known concentrations of BSA and then was applied to the sample absorbance values to determine sample concentration.

2.3.2 Measuring protein per cell

To estimate the amount of protein per cell, cells were counted prior to protein extraction and protein abundance was determined by two methods. First, a known number of cells were lysed in RIPA buffer and protein concentration was measured using the Bio-Rad Protein Assay (described in Section 2.3.1). Second, equal numbers of cells were pelleted and resuspended in 1x SDS protein dye (diluted from 6x SDS protein dye (0.375 M Tris-Cl (pH 6.8), 12% SDS, 60% glycerol, 0.6 M DTT, and 0.06% bromophenol blue)) and separated on a 15% SDS-PAGE gel. The gel was Coomassie stained as described in Section 2.3.4 and compared to a Coomassie stained gel resolving equal µg amounts of whole cell lysates.

2.3.3 Western blot analysis

To determine protein abundance, western blot analysis was performed using whole cell lysates from HFF and iPS cells. Protein samples of 25 µg each were prepared for SDS-PAGE by adding 6x SDS protein dye and boiling at 95°C for 2 minutes. Samples for detection of core histones H2A, H2B, H3, and H4 were separated on 15% SDS-PAGE gels and transferred to 0.2 µm PVDF Immobilon-P^{SQ} Membrane (Millipore) in 1x Transfer Buffer (25 mM Tris and 192 mM glycine, pH 8.3) containing 35% methanol. All other proteins were resolved on 8% or 10% SDS-PAGE gels in 1x SDS-PAGE Buffer (25 mM Tris, 192 mM glycine and 0.1% SDS) and transferred to 0.45 µm PVDF Immobilon Transfer Membranes (Millipore) in 1x Transfer Buffer containing 20% methanol. Membranes for the detection of CELF1 (also known as

CUGBP1) were blocked and incubated in 5% non-fat dry milk in 1x PBS and 0.05% Tween-20 (PBST), while all others were blocked and incubated in 5% non-fat dry milk in 1x Tris Buffered Saline (TBS) and 0.1% Tween-20 (TBST). Antibody incubations took place for 1.5 hours at room temperature followed by three 10 minute washes with 1xTBST. SuperSignal West Pico Chemiluminescent Substrate (Thermo Fisher Scientific) was used for detection in concert with a ChemiDoc XRS+ System (Bio-Rad). Quantification was performed using ImageLab3.0 software (Bio-Rad). Antibodies used for western blot analysis are listed in Table 2.4 below.

Table 2.4 List of antibodies used for western blot analysis.

Protein target	Product name	Company	Type	Dilution
Alpha-tubulin	Anti-a-Tubulin	Sigma	mouse monoclonal	1:15,000 - 1:20,000
AUF1	Anti-AUF1	Millipore	rabbit polyclonal	1:5,000
CUGBP1/CELF1	CUG-BP1 (3B1)	Santa Cruz Biotechnology Inc.	mouse monoclonal	1:5,000
GAPDH	Anti-GAPDH	Millipore	mouse monoclonal	1:20,000
H1	Histone H1 (AE-4)	Santa Cruz Biotechnology Inc.	mouse monoclonal	1:150
H2A	Anti-Histone H2A (acidic patch)	Millipore	rabbit polyclonal	1:250
H2A	Histone H2A (C-19)	Santa Cruz Biotechnology Inc.	goat polyclonal	1:1,000
H2B	Anti-Histone H2B Antibody	Millipore	rabbit polyclonal	1:1,000
H3	Anti-Histone H3 Antibody	Millipore	rabbit polyclonal	1:1,000
H4	Histone H4 Antibody	Bethyl Laboratories, Inc.	rabbit polyclonal	1:10,000
HuR	HuR (3A2)	Hybridoma	mouse monoclonal	1:50
KHSRP	KHSRP Antibody	Novus Biologicals	rabbit polyclonal	1:5,000
PARN	Anti-PARN serum	BIOO Scientific/in house	rabbit polyclonal	1:5,000
PCBP1	Anti-PCBP1	MBL International	rabbit polyclonal	1:1,000
PCBP1	Anti-PCBP1	Abnova	mouse polyclonal	1:1,000
PCBP2	hnRNP E2 (23-G)	Santa Cruz Biotechnology Inc.	mouse monoclonal	1:2:50
PCBP2	Anti-PCBP2	Abnova	rabbit polyclonal	1:5,000
PCBP3	Anti-PCBP3	Sigma	rabbit polyclonal	1:1,000
PCBP3	Anti-PCBP3	MBL International	rabbit polyclonal	1:15,000
PCBP4 (Isoform C)	PCBP4 (B-25)	Santa Cruz Biotechnology Inc.	rabbit polyclonal	1:150
PCBP4 (Isoform A)	Anti-PCBP4 (196-210)	Sigma	rabbit polyclonal	1:1,000
PTBP1	PTBP1	Abcam	goat polyclonal	1:500
PUM2	Pumilio 2 Antibody	Bethyl Laboratories, Inc.	rabbit polyclonal	1:5,000
SLBP	SLBP Antibody (2C4-1C8)	Novus Biologicals	mouse monoclonal	1:500
XRN1	XRN1 (C-1)	Santa Cruz Biotechnology Inc.	mouse monoclonal	1:5,000
ZFP36L2	Anti-ZFP36L2	Genway	rabbit polyclonal	1:500
	Donkey anti-goat IgG-HRP	Santa Cruz Biotechnology Inc.	Secondary	1:5,000 - 1:10,000
	Goat anti-mouse IgG-HRP	Santa Cruz Biotechnology Inc.	Secondary	1:10,000 - 1:20,000
	Goat anti-rabbit IgG (H+L)-HRP	Bio-Rad	Secondary	1:2,000 - 1:20,000

2.3.4 Coomassie staining of SDS-PAGE gels

To visualize histone proteins in a total lysate sample, 25 µg of protein was resolved on a 15% SDS-PAGE gel which was then fixed in 50% methanol/10% acetic acid for 10 minutes. The gel was stained with 0.5% Coomassie Blue R-250 for 2 hours then destained in 40% methanol/10% acetic acid and washed in MilliQ water. The gel was imaged with a ChemiDoc XRS+ System and then allowed to dry between two sheets of cellophane after which a scanned image was taken using PhotoSuite (MGI). Core histones (H2A, H2B, H3, and H4) were the predominant bands with molecular weights between 10 – 15 kDA. Quantification was performed using ImageQuant software.

2.3.5 Two dimensional protein electrophoresis

2.3.5.1 Lysate preparation

Cells were lysed in RIPA buffer containing Protease Inhibitor Cocktail and 100x Halt Phosphatase Inhibitor Cocktail (Pierce) at a final concentration of 1x. Lysate samples were sonicated and protein concentration was measured as previously described in Section 2.3.1. Aliquots were then made of 300 µg lysate each and stored at -80°C until use. Pellets were resuspended in 125 µL Rehydration Buffer (8 M urea, 2 M thiourea, 2% CHAPS, 0.3% DTT, 0.2% Triton-X 100, and ~0.002% bromophenol blue) containing 2.5% carrier ampholytes, pH 3-10 (Fluka).

2.3.5.2 Isoelectric focusing

Prior to loading for two dimensional (2D) protein electrophoresis, protein samples were centrifuged at maximum speed for 5 minutes. Samples were loaded on ReadyStrip Immobilized pH Gradient (IPG) Strips, pH 3-10 (Bio-Rad) using the Bio-Rad Protean IEF Cell. The first dimension was run following 12 hours of active rehydration at 50 V and focusing was performed initially at 250 V for 15 minutes then for a total of 30 kV-hours.

2.3.5.3 Second dimension analysis

After focusing, strips were washed in Equilibration Buffer (6 M urea, 30% glycerol, 2% SDS, and 24 mM Tris-Cl (pH 6.8)) containing 2% DTT followed by Equilibration Buffer containing 2.5% iodoacetamide. Strips were then loaded onto precast Mini-Protean TGX Gels (Bio-Rad) with Precision Plus Protein Standard Plugs, unstained (Bio-Rad). To seal the strip and plug into the gel, an agarose overlay (0.5% agarose and ~0.003% bromophenol blue in 1x SDS-PAGE Running Buffer) was applied to the top of each well. After the overlay had solidified, the gel was run in 1x SDS-PAGE Buffer at 150 V for approximately 1.5 hours. Proteins were transferred to 0.45 µm PVDF Immobilon Transfer Membranes using a semi-dry transfer apparatus and 1x Transfer Buffer containing 20% methanol at 18 V for 20-25 minutes.

To mark the ladder, membranes were stained in 0.1% Ponceau S in 1% acetic acid for 5 minutes followed by two 5 minutes washes in 5% acetic acid and two 5 minute washes in dH_2O. A pencil was used to mark the ladder and then membranes were blocked in 5% milk in TBST. Antibody detection was performed as previously described in Section 2.3.3. The antibodies used for 2D protein electrophoresis are listed below in Table 2.5.

Table 2.5 List of antibodies used for two dimensional protein electrophoresis.

Protein target	**Product name**	**Company**	**Type**	**Dilution**
PCBP1	Anti-PCBP1	MBL International	rabbit polyclonal	1:1,000
PCBP2	hnRNP E2 (23-G)	Santa Cruz Biotechnology Inc.	mouse monoclonal	1:250
PCBP3	Anti-PCBP3	Sigma	rabbit polyclonal	1:1,000
PCBP4 (Isoform C)	PCBP4 (B-25)	Santa Cruz Biotechnology Inc.	rabbit polyclonal	1:150
PCBP4 (Isoform A)	Anti-PCBP4 (196-210)	Sigma	rabbit polyclonal	1:1,000
alpha-tubulin	Anti-α-Tubulin	Sigma	mouse monoclonal	1:15,000
GAPDH	Anti-GAPDH	Millipore	mouse monoclonal	1:20,000
	Goat anti-mouse IgG-HRP	Santa Cruz Biotechnology Inc.	Secondary	1:15,000
	Goat anti-rabbit IgG (H+L)-HRP	Bio-Rad	Secondary	1:2,000 - 1:15,000

2.4 Bioinformatic analysis of microarray data

As bioinformatic analyses described here were performed in large part by Dr. Ju Youn Lee and Dr. Carol Wilusz, explanation of methodology has been adapted from descriptions that appeared in: Neff *et al*. Global analysis reveals multiple pathways for unique regulation of mRNA decay in induced pluripotent stem cells. Genome Research. Vol. 22, No. 8, pg. 1457-1467, August 2012.

2.4.1 Half-life analysis by microarray

Microarray data were first processed by the Affymetrix Power Tools (APT) program, using the GC-bin method for background correction. All probe set values were then normalized to the fifth-percentile value of the same array. Transcripts whose probe sets with Detection Above Background (DABG) P-value <0.05 for at least two out of three replicates at the 0 min time point were considered "expressed" and used for subsequent analysis.

Global mRNA half-lives were determined from microarray data by Dr. Ju Youn Lee who applied a nonlinear least squares model (NLS; Wang et al., 2002). Similar to that described in Section 2.2.6, the abundance values of each transcript at time points following actinomycin D treatment were fitted to a first-order exponential decay curve over time to determine the decay rate constant (k). Half-life ($T_{1/2}$) was defined as ln(2)/k. All three replicates were used to generate a single decay curve for each transcript and half-lives were deemed reliable for those that fit the exponential decay curve with P-value <0.05.

2.4.2 Gene Ontology analysis

Gene Ontology (GO) analysis was performed by Dr. Ju Youn Lee and Dr. Carol Wilusz. Gene lists were uploaded to DAVID (Huang et al., 2009) along with a background list consisting of all the genes for which half-lives were generated in both cell types. Terms that were significantly over-represented in the list of interest were selected based on P-value. For analysis of domains enriched in ZNF mRNAs, DAVID

was used to retrieve lists of genes associated with terms defined by Simple Modular Architecture Research Tool (SMART; Letunic et al., 2009). The accession numbers for these terms are C2H2-ZNF #SM00355 or KRAB #SM00349. Half-lives for genes in each list were then used to generate a box-and-whiskers plot and P-values were determined using the Kolmogorov-Smirnov Test.

2.4.3 Cis-*element analysis*

Analysis of *cis*-elements over-represented in mRNAs was performed by Dr. Ju Youn Lee who examined 5' untranslated regions (UTRs), ORFs, and 3' UTRs for hexamers that were significantly enriched in the most stable (top 10% in half-life) and the least stable (bottom 10% in half-life) transcripts. Sequences and their annotations were retrieved from the RefSeq database and the significance of a hexamer was calculated by the Fisher's exact test comparing its frequencies in the most stable and least stable mRNA sets. A significance score was assigned to each hexamer equaling $-\log_{10}$(P-value)*s where s equals -1 if the hexamer is more common in the least stable transcripts and s equals 1 if otherwise.

Chapter 3: Global analysis reveals multiple pathways for unique regulation of mRNA decay in human induced pluripotent stem cells

3.1 Introduction

Large-scale approaches to studying mRNA decay have been used for more than a decade (Bernstein et al., 2002; Holstege et al., 1998; Wang et al., 2002) and have allowed researchers to both further characterize molecular mechanisms important for gene expression as well as identify cross-talk events that were previously unsuspected (Ghosh, 2000; Liang et al., 2004). This chapter will highlight the use of global mRNA decay analyses for understanding regulated gene expression as well as discuss our own results following genome-wide determination of mRNA half-lives in induced pluripotent stem cells and the differentiated cells they were generated from.

3.1.1 Global approaches for characterization of sequences and factors involved in mRNA decay

The use of global approaches to study cell networks has been encouraged by the advancement and accessibility of high-throughput technologies. At the start of this project, there had only been a limited number of studies investigating mRNA decay on a global scale in mammalian cells (Dolken et al., 2008; Friedel et al., 2009; Lam et al., 2001; Raghavan et al., 2002; Sharova et al., 2009; Yang et al., 2003) and little effort had been made to compare rates of decay between different cell types. One study focused on mouse Embryonic Stem (ES) cells prior to and following differentiation (Sharova et al., 2009) and provided interesting insights into stem cell-specific mRNA decay rates. However, there are several significant differences between mouse and human stem cells (Dowell, 2011; Zheng-Bradley et al., 2010), supporting that a similar study in human cells would give additional information. The dramatic changes in gene expression that occur during reprogramming must involve widespread coordinated changes in

mRNA decay rates and a global approach would allow us to identify sets of transcripts that showed differential decay in pluripotent and fully differentiated cell types.

Global analysis of mRNA degradation has provided insight into gene expression regulation by allowing the grouping of mRNAs based on their decay rates. Genes that share similar regulation of their mRNA half-lives during a cellular response often encode proteins involved in particular pathways or functions (Raghavan et al., 2002; Yang et al., 2003). Gene Ontology (GO) analysis may be used to identify cellular processes regulated at the level of mRNA decay. Studies investigating mRNA decay dynamics during development and differentiation indicate that mRNA decay is spatiotemporally regulated within an organism (Sharova et al., 2009; Thomsen et al., 2010). Additionally, changes in mRNA half-life in response to various growth or stress conditions highlight the importance of mRNA decay to regulating gene expression (Miller and Olivas, 2011; Munchel et al., 2011; Rabani et al., 2011). One of our goals in determining global mRNA decay rates in pluripotent and differentiated cells was to identify cellular processes that are differentially regulated at the level of mRNA decay through the use of GO analysis. It seemed likely that some of these processes may facilitate the maintenance or achievement of pluripotency.

Transcripts with similar decay rates may also be enriched for sequence motifs serving as *cis*-acting stability elements that interact with RNA-binding proteins (RBPs) and/or miRNAs. Identification of common *cis*-elements could give clues to mechanism(s) of targeted transcript stabilization or destabilization. Previous global analyses identified high association of AU-rich elements with mRNA instability (Schwanhäusser et al., 2011, 2013; Sharova et al., 2009; Thomsen et al., 2010) although presence of these elements alone is not predictive of stability, suggesting combinatorial effects of other RBPs and miRNAs (Yang et al., 2003). GU-rich motifs are also important *cis*-acting elements associated with mRNA destabilization (Lee et al., 2010; Vlasova et al., 2008). Characterization of this element identified CELF1 (also known as CUGBP1) as a *trans*-acting factor responsible for binding mRNAs at these

sequences to elicit rapid decay (Vlasova et al., 2008). By studying the sequences of mRNAs exhibiting differential decay in iPS and differentiated cells, we hoped to identify novel *cis*-acting elements through which transcript stability is regulated to support stem cell gene expression.

Bioinformatic analysis of genome-wide decay datasets has recently led to observations of coordinated regulation between transcription and decay networks. In yeast, impairment of RNA polymerase II activity affects the degradation of a select set of transcripts (Shalem et al., 2011) and genes with similar promoters have similar decay patterns (Dori-Bachash et al., 2012). Coupling of transcription and decay is also evident in mammalian genomes (Dori-Bachash et al., 2012). Interestingly, increased mRNA degradation is often correlated with increased mRNA abundance, suggesting feedback to transcription processes. This cross-talk between pathways allows for a quick but transient response to the environment (Dori-Bachash et al., 2012). As stem cells must be able to respond efficiently to differentiation cues, it seems likely that they would also demonstrate network cross-talk.

Our objective was to determine global mRNA half-lives in HFFs and iPS cells to (i) fully characterize mRNA decay rates in these two genetically identical cell types for the first time, (ii) identify differentially regulated mRNAs, and (iii) characterize pathways of gene expression regulated at the level of mRNA decay. Together, these results would characterize the contribution of regulated mRNA decay to stem cell gene expression.

3.1.2 Methods for performing global analysis of mRNA decay rates

Initial experiments to measure mRNA decay were first performed by impairing RNA polymerase activity and this method continues to be used today (Ross, 1995). Drugs such as actinomycin D (Bleyman and Woese, 1969) and 5,6-Dichlorobenzimidazole 1-b-D-ribofuranoside (DRB; Tamm et al., 1976) are used to inhibit transcription, eliminating the addition of nascent transcripts to the total

population such that any decrease in mRNA abundance can be attributed to mRNA degradation. The first global analysis of mRNA decay rates in human cells was performed almost serendipitously when lymphoma cells were used to investigate the effects of potential anti-cancer drug flavopiridol. Microarray analysis performed on cells following treatment revealed that the drug inhibited transcription and thus allowed for the generation of mRNA half-lives for 2,794 genes (Lam et al., 2001). In recent years, methods utilizing metabolic labeling of nascent mRNAs have been developed since prolonged transcription inhibition can have cytotoxic effects which may influence gene expression in unpredictable ways (Sawicki and Godman, 1971). Furthermore, recent studies indicate that transcription and degradation events are coupled (Dori-Bachash et al., 2012; Shalem et al., 2011), therefore inhibiting transcription almost certainly has indirect effects on mRNA decay rates. In metabolic labelling, chemically modified bases such as 4-thiouridine (4sU) or 5-bromouridine (5BrU) are incorporated into newly transcribed mRNAs allowing for the selective separation of labeled transcripts from previously existing unlabeled pools (Dolken et al., 2008; Tani et al., 2012). Although half-lives generated from metabolic labeling are less likely to be influenced by generalized transcriptome impairment, the isolation of RNA prior to analysis is more challenging and the algorithms used for analysis are distinct from those used in microarray-based half-life assays (Friedel and Dölken, 2009).

Global decay rates are calculated by determining the change in abundance for individual transcripts over time. High-density oligonucleotide microarrays have traditionally been used to generate mRNA half-lives by plotting changes in probe intensities over time (Raghavan and Bohjanen, 2004). More recently, high-throughput sequencing of RNA samples has been used wherein the number of sequencing reads are used to determine transcript abundance (Imamachi et al., 2013). Although this method provides more data than microarrays, both are adequate for determining global mRNA decay rates.

3.1.3 Rationale for experimental design

To determine global mRNA decay rates in HFF and iPS cells, we decided to use actinomycin D for several reasons: 1) the method was already established and in practice in the lab; 2) we felt that we could minimize effects of inhibition of transcription by treating cells for only a relatively short period of time. Functional annotation of unstable transcripts with half-lives less than 2 hours in mouse ES cells identified by Sharova et al. revealed that short mRNA half-lives correspond to genes involved in cell cycle regulation as well as pluripotency and early development (Sharova et al., 2009). Therefore, we felt that a short time period would be long enough to reliably determine a large proportion of shorter mRNA half-lives but also allow for extrapolation to estimate decay rates for longer lived transcripts. In addition, 3) our results would be more directly comparable to those generated previously in our lab (Lee et al., 2010) and to those from other labs, including the mouse ES cell decay study (Sharova et al., 2009). Finally, 4) the bioinformaticist working on the project with us had already developed algorithms for this type of analysis (Lee et al., 2010, 2012), thus making data analysis much simpler. We decided to use microarrays to assess mRNA abundances because other methods that use deep-sequencing of RNA samples at various time points generate a vast amount more information than is required to determine half-lives and at the time when this experiment was performed, deep-sequencing was significantly more expensive than microarrays.

We chose to use induced pluripotent stem cells because there are no ethical implications that restrict their use in research and characterized cell lines along with genetically matched precursor fibroblast cell lines were commercially available (System Biosciences). As all the genetic information in these two cell types is identical, any differences in mRNA decay rates could be attributed to either the differentiated or pluripotent phenotype. Although these iPS cells are not exactly the same as ES cells (Bilic and Belmonte, 2012), they provide a very convenient system for studying stem cells and were

therefore used in this study investigating global mRNA decay in pluripotency. The combined results and discussion of global mRNA decay analysis in HFFs and iPS cells are presented below[2].

3.2 Generation and quality control of RNA samples for microarray hybridization

Cells were treated with actinomycin D to inhibit transcription and total RNA was collected at 0, 15, 30, 60, 120 and 240 minute time points starting 20 minutes following treatment. This 20 minute delay was chosen to allow the inhibitor enough time to effectively halt transcription and is routinely used in determination of half-lives in mammalian cells. By plotting the change in mRNA abundance over time and fitting to an exponential decay curve, a decay rate constant (k) was provided to calculate mRNA half-life ($T_{1/2}$) where $T_{1/2} = \ln(2)/k$.

To minimize variation, triplicate sample sets were selected from experiments performed simultaneously in quadruplicate (replicates A, B, C, and D). For each cell line, 24 culture dishes (four replicates for six time points) were treated with actinomycin D and processed for RNA isolation from TRIzol at the same time. To choose the three replicates with the lowest standard deviation, half-lives of two test mRNAs were determined by qRT-PCR. *FOS* mRNA was selected to represent short-lived mRNAs as previous studies indicate this transcript has a half-life ranging from 20 – 60 minutes in human monocytes (Sariban et al., 1988). *FOS* mRNA abundance was assessed at 0, 15, 30 and 60 minutes. To assess the quality of samples taken at later time points (0, 60, 120 and 240 minutes), the moderately unstable *TUT1* mRNA was chosen. In mouse C2C12 myoblasts, *TUT1* mRNA half-life is ~93 ± 10 minutes (Lee et al., 2010). After measuring transcript abundances at each time point and calculating half-lives, triplicate combinations of replicates (ABC, ABD, ACD, and BCD) were compared in each cell line to select the replicates with the lowest standard deviation (Appendix A1). Based on these data, we found that

[2] The majority of the results presented in Chapter 3 appeared in: Neff *et al*. Global analysis reveals multiple pathways for unique regulation of mRNA decay in induced pluripotent stem cells. Genome Research. Vol. 22, No. 8, pg. 1457-1467, August 2012.

HFF replicates A, B, and C and iPS replicates A, C, and D had the lowest standard deviation among time points and were therefore chosen to represent each cell line. It is important to note that although experiments described here were performed using HFFs cultured on plastic, analysis of HFFs cultured on Matrigel showed no measurable effect on mRNA abundances (Appendix A2).

Figure 3.1 below shows the half-lives of *FOS* and *TUT1* mRNAs in HFF (Figure 3.1A) and iPS cells (Figure 3.1B) as determined using the selected replicates. These data indicate that the half-lives measured for test mRNAs in each cell line were within the expected range. Consistent with published data (Sariban et al., 1988), the half-life of *FOS* mRNA in HFFs was ~34 ± 2 minutes and ~42 ± 13 minutes in iPS cells. In both cell lines, *TUT1* mRNAs were slightly more stable than in C2C12 myoblast cells (Lee et al., 2010) with a half-life of ~116 ± 13 minutes in HFFs and ~161 ± 15 minutes in iPS cells. These differences are not surprising given that this comparison is between different cell types and species. Nevertheless, the results support that *TUT1* mRNA is more stable than *FOS* mRNA, as predicted, in both HFF and iPS cells. Both sets of samples were used to proceed with global analysis.

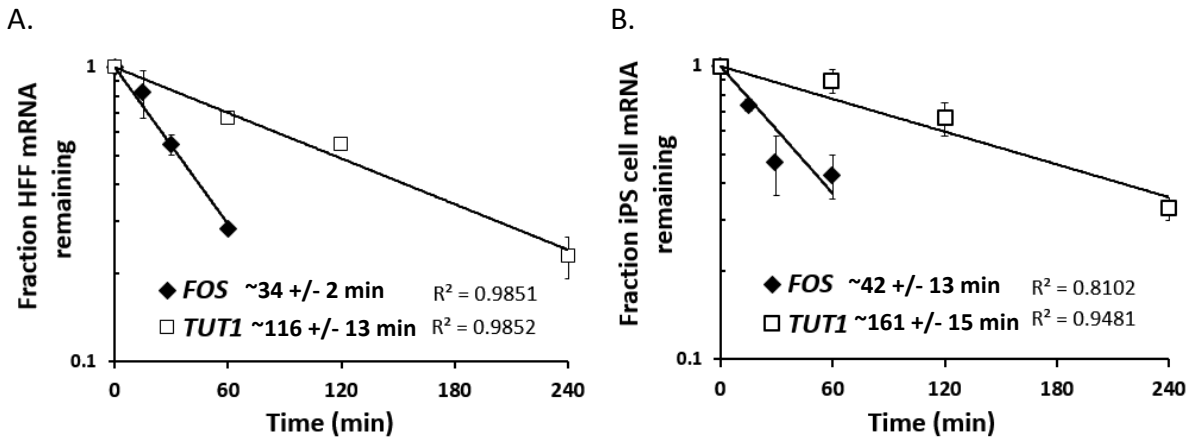


Figure 3.1 Determination of half-lives for test mRNAs *FOS* and *TUT1* allows selection of microarray sample replicates. Half-lives were calculated for *FOS* and *TUT1* mRNAs in (A) HFF and (B) iPS cells following actinomycin D treatment. mRNA levels were measured by qRT-PCR at each time point and normalized to *GAPDH* mRNA. The standard deviations were derived from three independent replicates. Half-lives for each mRNA and the R^2 values are denoted to the right of the cell-line keys.

Total RNA samples were used to generate cDNA probes which were hybridized to Affymetrix Human Gene Chip 1.0 ST microarrays (Erin Petrilli, Colorado State University Genomics and Proteomics Core Facility). Probe intensities were measured to indicate transcript abundance and data files were sent to collaborators Dr. Bin Tian and Dr. Ju Youn Lee at the Rutgers University – New Jersey Medical School for analysis.

3.3 Estimation of mRNA half-lives in HFF and iPS cells from gene expression data

Microarray data were first processed by Dr. Ju Youn Lee with the Affymetrix Power Tools (APT) program using the GC-bin method for background correction. All probe set values were normalized to the fifth-percentile value of the same array and transcripts whose probe sets had a Detection Above Background (DABG) P-value <0.05 for at least two out of three replicates at the 0 minute time point were considered "expressed" and used for subsequent analysis. As shown in Figure 3.2, 20,434 mRNAs in iPS cells and 19,791 mRNAs in HFFs were expressed at the 0 minute time point with 19,190 mRNAs expressed in both cell lines.

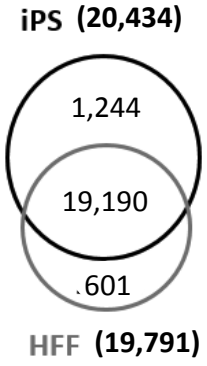


Figure 3.2 Venn diagram showing the number of mRNAs expressed in HFF and iPS cells at the 0 minute time point as determined by microarray.

Following normalization and processing of the data (described in Chapter 2), the 0 minute time point from each replicate was then compared to determine correlation of expression between sample sets. As seen in Figure 3.3, there was high correlation within replicates of each cell line (0.99) and less correlation between the two cell lines (0.88). These results were expected and provided sample quality assurance.

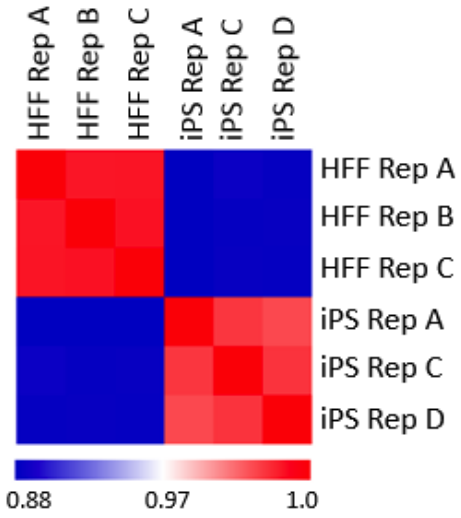


Figure 3.3 Correlation of microarray array data between and within replicates at the 0 minute time point. Heat map showing the correlation between microarrays for the 0 minute time point of each HFF and iPS cell replicate using Robust Multi-array Average (RMA) normalization. Raw intensity values were background corrected, $\log_2$ transformed, and quantile normalized.

Global mRNA half-lives were determined by applying the Nonlinear Least Squares (NLS) model (Wang et al., 2002). Similar to calculations performed using qRT-PCR data, microarray expression values for each transcript at time points following actinomycin D treatment were fit to a first-order exponential decay curve to determine the mRNA decay rate constant (k) ($T_{1/2} = \ln(2)/k$). All three replicates were used to generate a single decay curve for each transcript and half-lives were deemed reliable for those that fit the exponential decay curve with a P-value <0.05. Figure 3.4 below shows decay curves generated for *HIST1H14B* mRNA to calculate a half-life of ~81 minutes (P-value = 2.59×10^{-7}) in HFFs and ~277 minutes (P-value = 6.42×10^{-5}) in iPS cells within a 95% Confidence Interval (CI).

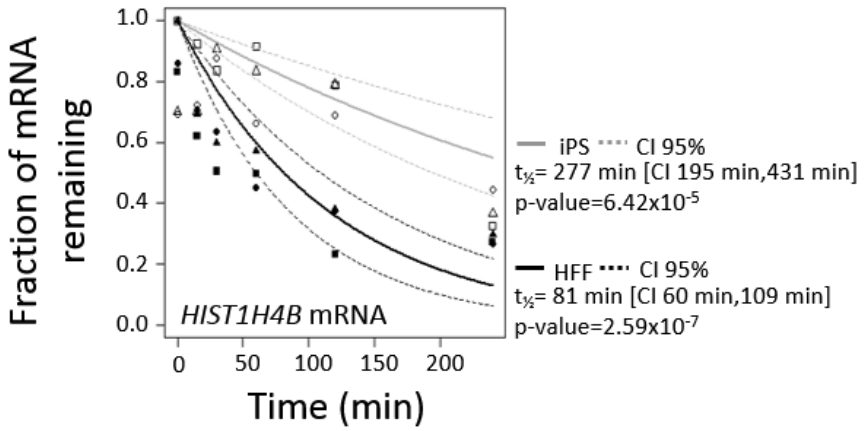


Figure 3.4 Exponential decay curves for *HIST1H4B* mRNA in HFF and iPS cells as determined by microarray analysis. Each symbol denotes *HIST1H4B* mRNA abundance in a single replicate with closed symbols representing abundance in HFFs and open symbols representing mRNA abundance in iPS cells. Three replicates were performed in each cell type at each time point. All three replicates were used to generate a single half-life in both cell types. The P-value for each cell type indicates how well the data fit to the exponential decay curve (solid line).

When this method was applied to all detectable transcripts, reliable half-lives were determined for 8,283 HFF mRNAs and 10,445 iPS mRNAs (Figure 3.5). The microarray data were deposited in the NCBI Gene Expression Omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/geo/) under accession number GSE33417.

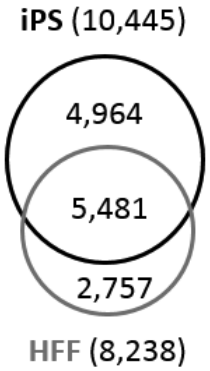


Figure 3.5 Venn diagram showing the number of mRNA half-lives in HFF and iPS cells determined by microarray. Half-lives for 5,481 mRNAs were determined in both cell lines.

Half-lives of 5,481 mRNAs were determined in both cell types and this subset was used for all subsequent analyses. When the half-life distribution of these common transcripts was graphed (Figure 3.6), we found that generally speaking, mRNAs decay slightly faster in iPS cells than in HFFs as indicated by median half-lives of 8.6 hours and 9.2 hours, respectively. Although this difference is not large, it is highly significant with P-value = 1.59×10^{-10}. The median half-life in mouse embryonic fibroblasts measured by Sharova et al. was approximately 7 hours and differentiation caused this half-life median to change. Spontaneous differentiation induced by depletion of leukemia inhibitory factor from culture medium (Williams et al., 1988) resulted in a median half-life of 5.5 hours while inducing neuron formation through addition of retinoic acid (Bain et al., 1995; Slager et al., 1993) lengthened the median to 8.6 hours (Sharova et al., 2009). Our data are consistent with the median decay rates of 10 hours described for human hepatocellular carcinoma cells (HepG2) and primary fibroblasts (Bud8; Yang et al., 2003) and 9 hours in mouse fibroblasts (Schwanhäusser et al., 2011, 2013). When comparing the median half-life to that established in yeast and bacteria, investigators reasoned that the median mRNA half-life was proportional to length of cell cycle for each cell type (Yang et al., 2003). Hence, it seems reasonable that the faster cell cycle of stem cells contributes to their slightly shorter median mRNA half-life.

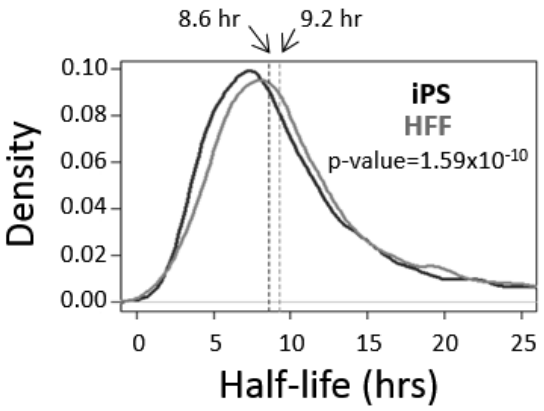


Figure 3.6 The median mRNA half-life in iPS cells is slightly shorter than the median mRNA half-life in HFFs. Graph showing the distribution of 5,481 mRNA half-lives in HFFs (gray) and iPS cells (black). The median half-life in each cell line is denoted by a dotted line.

One possible explanation for this difference in overall decay rates could be due to differential expression of mRNA decay enzymes. Although it was not feasible to investigate expression of all mRNA decay enzymes, we measured the abundances of mRNA decay factors for which antibodies were readily available, namely XRN1 and PARN, to determine whether they are differentially regulated in the two cell types. XRN1 is a cytoplasmic exoribonuclease that functions to decay a large proportion of cellular transcripts in the 5'→3' direction (Bashkirov et al., 1997; Muhlrad et al., 1994). XRN1 has been shown very recently to be essential for coupling mRNA decay rates to transcription in yeast, highlighting its wide-ranging impact on gene expression (Haimovich et al., 2013). PARN is a deadenylase that often acts at the first step of mRNA decay (Körner and Wahle, 1997) and has vital roles for oogenesis and embryogenesis (Copeland and Wormington, 2001; Kang and Han, 2011; Radford et al., 2008). Western blot analysis of these decay factors in HFF and iPS cells is shown in Figure 3.7. The abundance of each protein was normalized to GAPDH levels and iPS abundance was quantified relative to HFF expression. Although this experiment was performed twice, one replicate was developed on film and therefore not quantified. Results from both replicates indicated that expression of XRN1 is relatively similar between

cell lines and PARN is ~2-fold more abundant in iPS cells. The increased abundance of PARN in iPS cells could contribute to the slightly faster turnover of mRNAs in these cells but other factors, such as the abbreviated pluripotent cell cycle, could also influence overall mRNA decay rates. These results are supported by mouse ES decay data wherein more than 70 mRNAs encoding decay factors showed decreased abundance upon differentiation, although XRN1 and PARN were not represented (Sharova et al., 2009). In our abundance data at the 0 minute time point, expression of *XRN1* and *PARN* mRNA was slightly increased in HFFs by 1.09- and 1.17-fold change, respectively, compared to iPS cells although half-lives were not generated in either cell line.

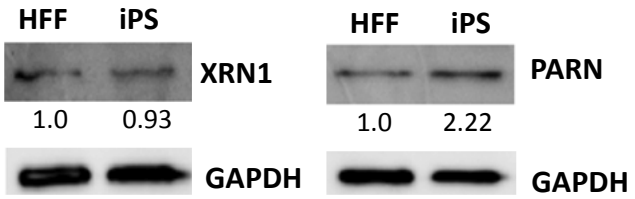


Figure 3.7 Differential expression of decay factors may contribute to differences in median mRNA half-life. Western blot analysis of decay factors XRN1 and PARN in HFF and iPS cells using 25µg of whole cell lysate on a 10% SDS-PAGE gel transferred to polyvinylidene difluoride (PVDF) membranes. Quantification using ImageQuant software is based on a single replicate where levels of XRN1 or PARN are normalized to levels of GAPDH and relative to expression in HFF cells.

We next sought to identify stable and unstable transcripts in each cell line by evaluating HFF and iPS mRNA half-life distributions for 5,481 transcripts (Figure 3.8). Transcripts in the 10th percentile were considered unstable while those in the 90th percentile for each cell line were designated as stable. As these stable and unstable mRNA sets exhibited differential regulation compared to the other 4,385 transcripts (80%) in the dataset, Gene Ontology analysis was performed to determine functional terms associated with these distinct mRNA subsets.

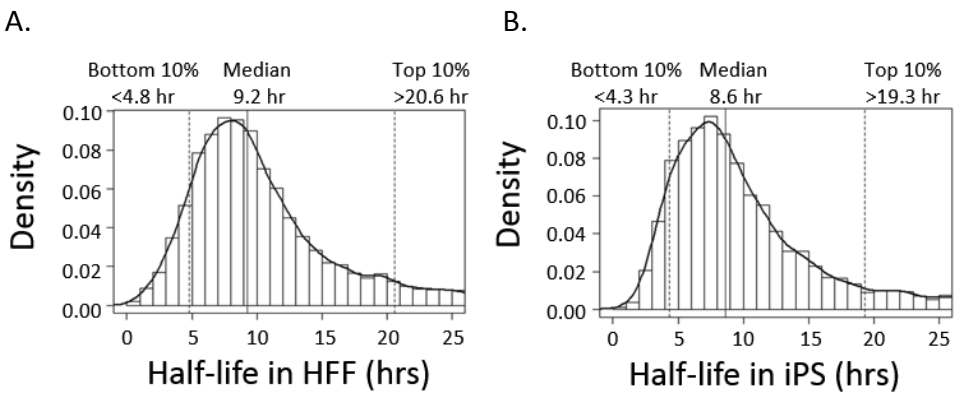


Figure 3.8 Graphing the mRNA half-life distribution in HFF and iPS cells identified 10th, median, and 90th percentiles for 5,481 mRNAs. Half-life distribution of 5,481 mRNAs in A) HFFs and B) iPS cells determines unstable (bottom 10%) and stable (top 10%) mRNAs in each dataset.

3.4 Gene Ontology terms associated with instability or stability

We next assigned functional classifications to the stable and unstable mRNAs in each cell line. Common pathways associated with stable and unstable transcripts could indicate a mechanism of coordinate regulation. To perform Gene Ontology analysis, gene lists were uploaded to DAVID (Huang et al., 2009a) along with a background list consisting of all the genes for which half-lives were generated in both cell types. Table 3.1 shows the top 20 terms significantly over-represented in unstable mRNAs and terms enriched in stable mRNAs using a similar P-value cutoff for each cell line. The results are discussed below.

Table 3.1 Gene Ontology terms associated with unstable or stable mRNAs in HFF and iPS cells.

Unstable in HFF

GO ID	P-value	Term
GO:0006357	7.45E-06	regulation of transcription from RNA polymerase II
GO:0006366	1.59E-05	transcription from RNA polymerase II promoter
GO:0034621	2.41E-05	cellular macromolecular complex subunit organization
GO:0034622	4.10E-05	cellular macromolecular complex assembly
GO:0009892	6.94E-05	negative regulation of metabolic process
GO:0031324	6.94E-05	negative regulation of cellular metabolic process
GO:0010605	1.96E-04	negative regulation of macromolecule metabolic
GO:0045934	1.96E-04	negative regulation of nucleobase, nucleoside,
GO:0051276	1.96E-04	chromosome organization
GO:0009890	3.27E-04	negative regulation of biosynthetic process
GO:0010558	3.27E-04	negative regulation of macromolecule biosynthetic
GO:0016481	3.27E-04	negative regulation of transcription
GO:0006325	3.33E-04	establishment or maintenance of chromatin
GO:0010629	5.09E-04	negative regulation of gene expression
GO:0045892	5.63E-04	negative regulation of transcription, DNA-
GO:0051253	5.63E-04	negative regulation of RNA metabolic process
GO:0006334	5.88E-04	nucleosome assembly
GO:0031497	5.88E-04	chromatin assembly
GO:0006511	8.25E-04	ubiquitin-dependent protein catabolic process
GO:0019941	8.25E-04	modification-dependent protein catabolic process

Stable in HFF

GO ID	P-value	Term
GO:0007600	9.69E-07	sensory perception
GO:0050890	4.06E-06	cognition
GO:0007606	5.15E-06	sensory perception of chemical stimulus
GO:0007608	5.21E-05	sensory perception of smell
GO:0006811	5.67E-05	ion transport
GO:0007268	5.30E-04	synaptic transmission
GO:0006812	7.56E-04	cation transport

Unstable in iPS

GO ID	P-value	Term
GO:0006366	1.51E-13	transcription from RNA polymerase II promoter
GO:0006357	7.40E-12	regulation of transcription from RNA polymerase II
GO:0010628	6.68E-08	positive regulation of gene expression
GO:0045941	6.68E-08	positive regulation of transcription
GO:0010558	9.90E-08	negative regulation of macromolecule biosynthetic
GO:0016481	9.90E-08	negative regulation of transcription
GO:0045934	9.90E-08	negative regulation of nucleobase, nucleoside,
GO:0010629	1.31E-07	negative regulation of gene expression
GO:0010605	3.00E-07	negative regulation of macromolecule metabolic
GO:0045893	3.05E-07	positive regulation of transcription, DNA-dependent
GO:0051254	3.05E-07	positive regulation of RNA metabolic process
GO:0007049	3.16E-07	cell cycle
GO:0045935	3.63E-07	positive regulation of nucleobase, nucleoside,
GO:0009890	4.60E-07	negative regulation of biosynthetic process
GO:0009892	5.78E-07	negative regulation of metabolic process
GO:0031324	5.78E-07	negative regulation of cellular metabolic process
GO:0010557	8.58E-07	positive regulation of macromolecule biosynthetic
GO:0051276	9.63E-07	chromosome organization
GO:0045892	1.20E-06	negative regulation of transcription, DNA-dependent
GO:0051253	1.20E-06	negative regulation of RNA metabolic process

Stable in iPS

GO ID	P-value	Term
GO:0007606	3.42E-06	sensory perception of chemical stimulus
GO:0007608	6.92E-06	sensory perception of smell

3.4.1 Gene Ontology terms associated with instability or stability in HFFs

Table 3.1 shows that unstable HFF mRNAs were enriched for functions associated with "transcription" and "chromatin assembly" amongst other more general terms. Numerous studies have shown that mRNAs encoding transcriptional regulators have a high propensity to be unstable (Schwanhäusser et al., 2011, 2013; Yang et al., 2003). Regulation of transcription also involves chromatin modulation so it is not surprising to have found terms related to both functions represented here.

In stable HFF mRNAs, we found an enrichment of terms associated with neurophysiological processing such as "sensory perception" and "cognition". "Ion transport" was also over-represented. Previous studies have indicated that transcripts encoding housekeeping genes tend to be stable as their expression does not need to change in response to signals (Schwanhäusser et al., 2011, 2013). While ion transport could be considered a housekeeping function, the reason for enrichment of neurophysiological processing terms is not clear.

3.4.2 Gene Ontology terms associated with instability or stability in iPS cells

As in HFFs, we found that relatively short-lived iPS cell mRNAs encoded proteins enriched with functions related to "transcription". To a lesser extent, we also saw that functions in "cell cycle" and "chromosome organization" are represented. Pluripotent stem cells are characterized by their abbreviated cell cycle and unique chromatin structure so it was interesting to find these functions in our unstable mRNA dataset. Unstable mRNAs in mouse ES cells were enriched for genes with functions in "transcription factors", "cell cycle", "apoptosis" and, "development" (Sharova et al., 2009), consistent with our analysis. As shown in stable HFF mRNAs, we also found that stable iPS cell mRNAs are enriched for GO terms related to "sensory perception" but the significance of this is unclear.

The smaller number of enriched terms in the stable mRNA set could possibly be explained by the wide range of functions performed by proteins encoded by stable mRNAs as compared to the relatively smaller set of functions performed by proteins encoded by unstable mRNAs (Halees et al., 2011; Khabar, 2005). Additionally, half-life P-values were much lower for unstable mRNAs in general than for stable mRNAs that had half-lives of >19 hours extrapolated from a 4-hour time course.

3.5 Sequence elements associated with instability or stability

Having identified stable and unstable mRNAs in each cell type, we next wanted to determine whether specific sequence elements that might modulate mRNA decay rates were over-represented in either set. This information could give clues as to the identity of *trans*-acting factors that might be specifically regulating mRNA decay in a specific cell type. Analysis of hexamer-sized *cis*-elements over-represented in 5' UTR, ORF, and 3' UTR of mRNAs was performed on the dataset containing stable or unstable HFF and iPS cell transcripts. Sequences and their annotations were retrieved from the RefSeq database and the significance of a hexamer was calculated by the Fisher's exact test comparing its frequency in the most stable and least stable mRNA sets. A significance score was assigned to each hexamer equaling $-\log_{10}(\text{P-value})*s$, where s equals -1 if the hexamer is more common in the least stable transcripts, and s equals 1 if otherwise. A heat map representation of hexamer sequences found within each region of unstable and stable mRNAs in HFF and iPS cells is shown in Figure 3.9 below.

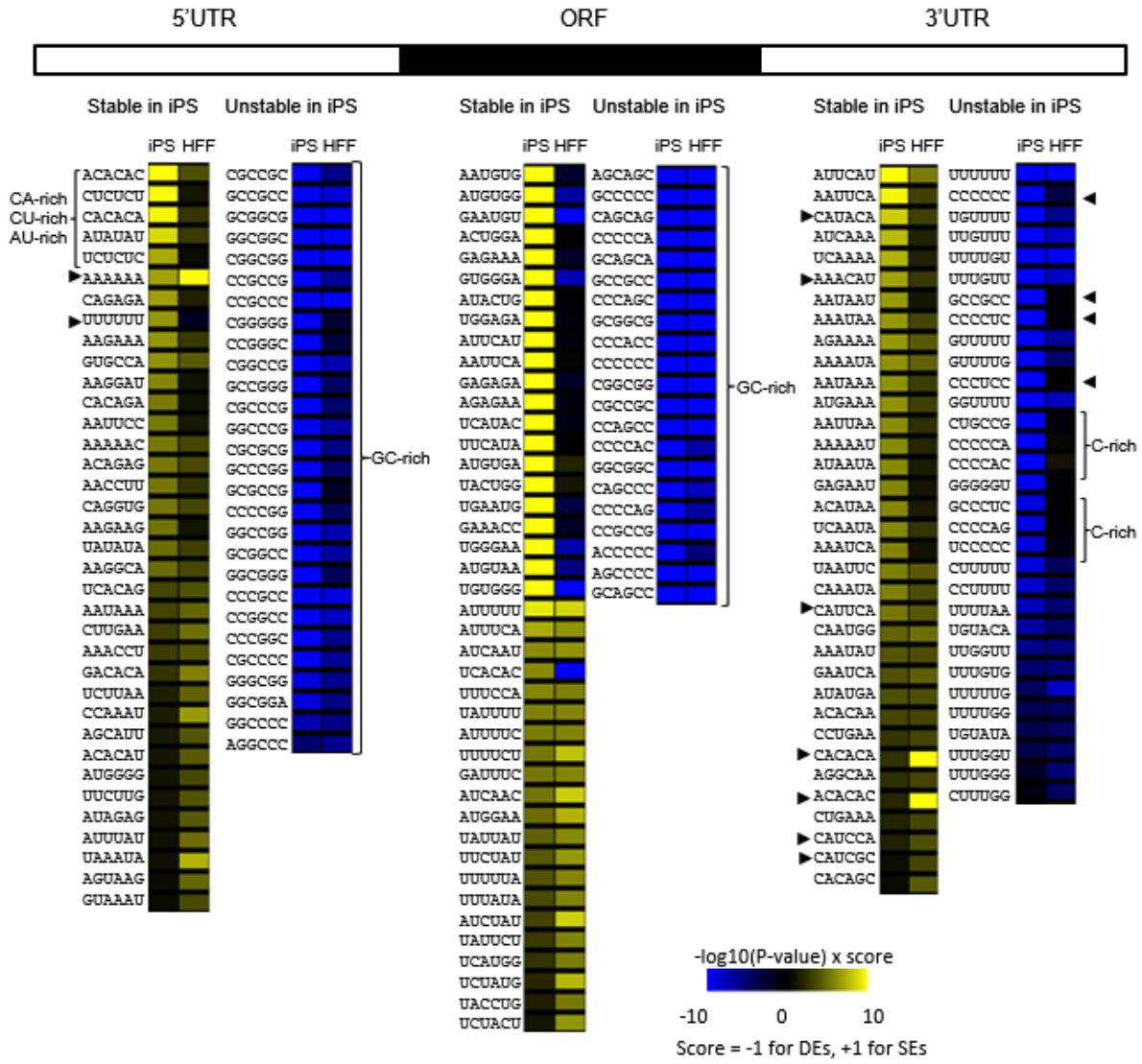


Figure 3.9 Hexamers enriched in the 5'UTR, ORF, and 3'UTR of stable and unstable iPS cell transcripts compared to enrichment in HFF transcripts. Hexamers over-represented in stable mRNAs (Stabilizing Elements; SEs) are depicted in yellow while hexamers over-represented in unstable mRNAs (Destabilizing Elements; DE) are in blue. Specific elements of interest are indicated by brackets and arrowheads.

3.5.1 5'UTR elements are over-represented in unstable and stable mRNAs

Cis-acting elements in the 5'UTR of mRNAs such as Internal Ribosome Entry Sites (IRES), Upstream Open Reading Frames (uORFs), Terminal Oligopyrimidine (TOP) sequences, and secondary structures can influence mRNA activities through a variety of mechanisms. The role of 5'UTR *cis*-elements is predominantly in regulation of translation, but they can also influence other mRNA metabolism events including decay (de la Cruz et al., 2002; Dibbens et al., 1999; Vilela et al., 1999). We were encouraged to find the enrichment of hexamer sequences in the 5'UTR of unstable and stable HFF and iPS cell mRNAs that could potentially be important for maintaining differentiation or pluripotency. These 5'UTR elements are discussed further here.

3.5.1.1 GC-rich 5'UTR elements are over-represented in unstable mRNAs in both cell types

Analysis of 5'UTR hexamer sequences indicated that GC-rich elements in this region were associated with instability in both cell lines, although more so in iPS cells. These data are consistent with previous findings that CpG dinucleotides are over-represented in unstable mRNAs in mouse ES cells (Sharova et al., 2009). In general, the 5'UTR of mRNAs is more GC-rich than the 3'UTR (Pesole et al., 1997), but it was interesting to find that these elements were correlated with transcript instability. Higher GC-content in this region can lead to more complex RNA structure that inhibits ribosomal scanning resulting in transcript degradation (Linz et al., 1997). Therefore, it is possible that these GC-rich hexamer sequences in the 5'UTR of unstable mRNAs are able to elicit instability through this mechanism.

3.5.1.2 Di- and mononucleotide repeat 5'UTR elements are over-represented in stable iPS cell mRNAs

Sequences in the 5'UTR that correlated with increased stability in both cell types included the hexamer AAAAAA, although the correlation was stronger in HFFs. Pre-AUG poly(A) tracts less than 12

nucleotides in length bind translation initiation factors whereas elements longer than 12 nucleotides likely bind poly(A)-binding protein to repress translation (Gilbert et al., 2007; Xia et al., 2011). Lsm1-7 binds poly(A) tracts 10 to 21 nucleotides in length in the 5'UTR of orthopoxvirus mRNAs, preventing 3'→5' decay and decapping (Bergman et al., 2007). These examples demonstrate that even simple homopolymer repeats can have various functions. A UUUUUU element was also identified in stable iPS cell mRNAs, but little correlation with stability was seen in HFFs. Additionally, we found that increased mRNA stability is associated with 5'UTR CA-, CU-, and AU- repeat elements in iPS cells, although this association is barely detected in HFFs. Interestingly, these appear to be actual repeat elements rather than general nucleotide enrichments since related elements such as CACCAC are not over-represented. It is possible that these repeat elements recruit iPS cell-specific RNA-binding proteins, such as HNRNP L which binds CA-repeat elements to regulate decay and translation processes (Lee et al., 2009, 2004).

3.5.2 GC-rich ORF elements are strongly associated with instability in both cell types

Unstable mRNAs in both cell lines showed over-representation of GC-rich elements within the coding regions. The presence of this class of elements could lead to the formation of secondary structures that impede ribosome progression. Surveillance mechanisms recognize stalled ribosomes and initiate decay of the transcript to release the translation machinery so that it may continue synthesizing proteins (Doma and Parker, 2006). However, it is important to note that studies in differentiated cells (HeLa and 293T) investigating the stability of GC-rich and GC-poor genes under the same promoter and UTR sequences found that mRNA decay rates are not correlated with GC content (Kudla et al., 2006).

3.5.3 3'UTR elements are over-represented in stable and unstable mRNAs

We expected to find sequence elements that correlate with stability in the 3'UTR since this is typically where most *cis*-acting stability elements have been located to date (Schoenberg and Maquat, 2012). The stability of an mRNA is in large part specified by the recruitment of *trans*-acting factors such as miRNAs and RNA-binding proteins. Identification of over-represented hexamer sequences could give clues as to which miRNAs and RBPs might be involved in modulating mRNA decay in HFF and iPS cells. Indeed, we found enrichment of sequences in both stable and unstable mRNAs in both cell lines and those are discussed in more detail below.

3.5.3.1 U-rich and GU-rich 3'UTR elements are over-represented in unstable mRNAs in both cell lines

As shown in Figure 3.9 above, U-rich and GU-rich elements are over-represented in unstable HFF mRNAs. In iPS cells, U-rich elements were correlated with instability to a higher degree compared to HFFs. Association of U-rich elements with instability is not surprising as destabilization induced by AU- and GU-rich sequences is well studied (Schoenberg and Maquat, 2012). Transcripts bearing these elements usually encode transcription factors, growth factors, and cytokines (Halees et al., 2011; Khabar, 2005). Such elements are reported to bind to CELF proteins (Vlasova et al., 2008; Vlasova-St Louis et al., 2013) and ELAV-like proteins including HuR (Brennan and Steitz, 2001; Dean et al., 2001); both well-known regulators of mRNA decay. U-rich *cis*-elements are also able to confer regulation in a developmentally-dependent manner, allowing temporal control (Liu et al., 2009). However, we were surprised to note that typical AU-rich elements (e.g. UAUUUAU and variants thereof) were not over-represented in either iPS or HFF unstable mRNAs. At this time we do not understand why AU-rich elements were not detected but it is possibly due to the under-representation of very unstable mRNAs in our dataset. Half-lives <2 hours were only determined for 69 mRNAs in HFFs and 26 mRNAs in iPS

cells. Low abundance due to instability may have prevented a higher proportion of these transcripts from meeting our criteria for good half-life determination.

3.5.3.2 C-rich 3'UTR elements are exclusively over-represented in unstable iPS cell mRNAs

In addition to U-rich elements, unstable iPS cell mRNAs also show enrichment of C-rich hexamers. As C-rich elements are typically associated with transcript stabilization (Holcik and Liebhaber, 1997; Wang et al., 1995), their identification here in unstable mRNAs was surprising. Several proteins are known to bind these RNA elements including HNRNP K (Siomi et al., 1993), PTBP1 (Kosinski et al., 2003), NOVA proteins (Buckanovich and Darnell, 1997; Yang et al., 1998), and the family of Poly(C)-Binding Proteins (PCBPs; Kiledjian et al., 1995; Makeyev and Liebhaber, 2000). Further investigation of C-rich elements and their potential role in regulating decay in stem cells is presented in Chapter 6.

3.5.3.3 CAU-containing 3'UTR elements are over-represented in stable iPS cell mRNAs

Additionally, we found that CAU-containing hexamer sequences (CAUACA, AAACAU, CAUUCA, CAUCCA, and CAUCGC) are enriched in the 3'UTR of stable mRNAs in iPS cells. A similar element in the 3' region of *GLUT1* mRNA stabilizes the transcript during glucose deprivation and hypoxia (Boado and Pardridge, 2002). Such elements are also found in the 3'UTR of *VEGF* mRNA (Claffey et al., 1998) and may be bound by RNA-binding proteins HNRNP L (Shih and Claffey, 1999), HuR (Levy et al., 1998; Tang et al., 2002), or PTB (Coles et al., 2004) to confer transcript stabilization.

3.5.3.4 CA-rich 3'UTR elements are over-represented in stable HFF mRNAs

In stable HFF mRNA 3'UTRs we found an enrichment of CA-repeat elements (CACACA and ACACAC) that are not over-represented in stable iPS cell mRNAs. Although these elements may be bound by proteins such as HNRNP L to destabilize mRNAs (Lee et al., 2009, 2004) through miRNA-

independent endonucleolytic cleavage (Bracken et al., 2011), binding of HNRNP L to CA-rich elements can also stabilize transcripts by inhibiting miRNA activity (Jafarifar et al., 2011). As these *cis*-elements are not well characterized, it is difficult to speculate on their contribution to pluripotency or differentiation in HFF and iPS cells.

3.6 Identification of mRNAs that show significant differences in decay rate between HFF and iPS cells

To identify differentially regulated transcripts, the 5,481 mRNAs whose half-lives were determined in both cell lines were compared to calculate half-life fold-change differences. The half-life difference was graphed to determine the most significantly stabilized (top 10%, 90th percentile) and most significantly destabilized (bottom 10%, 10th percentile) mRNAs in iPS cells compared to HFFs (Figure 3.10). Both of these data sets (each containing 548 mRNAs) were then analyzed in more detail using bioinformatic approaches to discern possible mechanisms of transcript stabilization and/or destabilization.

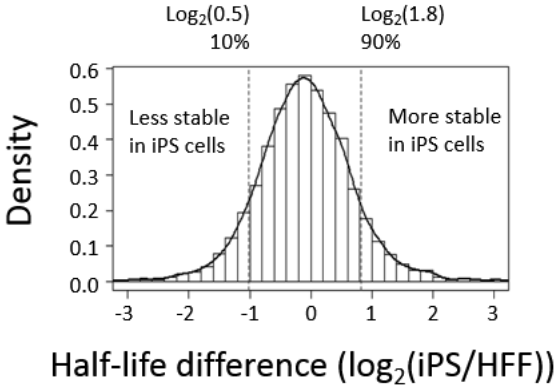


Figure 3.10 Plotting the distribution of fold-change in half-life difference between HFF and iPS cells reveals differentially regulated mRNAs. For 5,481 mRNAs, the fold-change in half-life difference was graphed to identify mRNAs destabilized (bottom 10%) and stabilized (top 10%) in iPS cells compared to HFFs.

3.6.1 Functional analysis reveals histone and C2H2 ZNF mRNAs are stabilized in iPS cells

The functional annotation tool DAVID (Huang et al., 2009a, 2009b) was used to identify enrichment of functionally-related genes within the set of mRNAs differentially regulated in iPS cells compared to HFFs. Table 3.2 summarizes these results. Although the dataset of destabilized mRNAs did not reveal significant term enrichment (data not shown), mRNAs that were destabilized in iPS cells were more likely to encode proteins associated with the plasma membrane (P-value =$1.2x10^{-4}$). Analysis of the mRNAs that were stabilized in iPS cells relative to HFFs identified several enrichments for GO terms including "nucleosome", "regulation of transcription", and "DNA binding" (P-value <$1x10^{-7}$). Transcripts that were associated with increased stability in mouse ES cells compared to differentiated cells were also enriched for functions including "transcription regulation" and "chromatin binding" (Sharova et al., 2009). Many of the transcripts contained in the group linked with the "nucleosome" GO term encode histones while many zinc finger (ZNF) protein mRNAs were associated with terms "regulation of transcription" and "DNA binding." To identify functional domains over-represented in these mRNAs, the SMART protein domain annotation tool (Letunic et al., 2012; Schultz et al., 1998) was used. Consistent with the identified GO terms, we found enrichment of histones (P-value =$3.1x10^{-4}$) and significant over-representation of KRAB (Krüppel-Associated Box) and C2H2 (Krüppel-like domain) zinc finger domains found in C2H2 ZNF mRNAs (KRAB P-value = $1.2x10^{-23}$; C2H2 zinc finger P-value = $7.4x10^{-18}$).

Table 3.2 List of Gene Ontology and SMART terms associated with mRNAs stabilized in iPS cells. Functional analysis revealed histone and KRAB C2H2 ZNF mRNAs are stabilized in iPS cells.

ID	Term	p-value
GO:Cellular Component		
GO:0000786	nucleosome	$1.1x10^{-8}$
GO:0032993	protein DNA Complex	$7.9x10^{-7}$
GO:0000785	chromatin	$4.2x10^{-4}$
GO:Biological Process		
GO:0006355	regulation of transcription, DNA-dependent	$3.1x10^{-8}$
GO:0051252	regulation of RNA metabolic process	$3.9x10^{-8}$
GO:0006350	transcription	$5.2x10^{-8}$
GO:0045449	regulation of transcription	$1.6x10^{-6}$
GO:0006334	nucleosome assembly	$1.2x10^{-5}$
GO:0031497	chromatin assembly	$2.6x10^{-5}$
GO:0065004	protein-DNA complex assembly	$3.7x10^{-5}$
GO:0034728	nucleosome organization	$3.7x10^{-5}$
GO:0006323	DNA packaging	$6.6x10^{-4}$
GO:0006333	chromatin assembly or disassembly	$7.3x10^{-4}$
GO:Molecular Function		
GO:0003677	DNA binding	$1.5x10^{-8}$
GO:0008270	zinc ion binding	$5.6x10^{-8}$
GO:0046914	transition metal ion binding	$4.8x10^{-7}$
GO:0046872	metal ion binding	$7.7x10^{-5}$
GO:0043169	cation binding	$8.5x10^{-5}$
SMART		
SM00349	Kruppel associated box (KRAB)	$1.2x10^{-23}$
SM00355	C2H2 zinc finger	$7.4x10^{-18}$
SM00414	histone 2A	$3.1x10^{-4}$

Given the dramatic stabilization of mRNAs encoding histones and C2H2 ZNF proteins, regulation of histone mRNA stability was further investigated and is presented in Chapter 4 while the mechanism of C2H2 ZNF mRNA stabilization is explored in Chapter 5.

3.6.2 Sequence analysis reveals mRNAs with U-rich or C-rich 3'UTR elements are destabilized in iPS cells

As determined through *cis*-element analysis of unstable iPS mRNAs, we also found significant enrichment of U-rich and C-rich elements in the 3'UTR of transcripts significantly destabilized in iPS cells compared to HFFs. Using the top 50 hexamer sequences identified in these mRNAs, we found that the most abundantly represented sequences are UUUUUU (P-value = $4.74x10^{-27}$), CCCCCC (P-value =

1.40x10^{-18}), and a CUG-containing motif (Figure 3.11). We investigated the regulation of U-rich and C-rich element containing mRNAs in more detail (Section 3.6.3 and Chapter 6, respectively). Although we elected not to study the CUG-rich motif any further, these elements in the 3'UTR of mRNAs may be bound by Muscleblind-like splicing regulator 1 (MBNL1) which destabilizes transcripts (Masuda et al., 2012; Warf and Berglund, 2007). Interestingly, MBNL proteins are negative regulators of pluripotency that repress stem cell-specific alternative splicing of transcription factor *FOXP1* mRNA. Knockdown of MBNL proteins promotes pluripotent splicing patterns and enhances reprogramming of fibroblasts (Han et al., 2013).

Hexamers clustered for sequence logos:

UUUUU	4.74E-27	UGUUU	7.59E-17
UUGUU	6.26E-16	UUUUG	5.76E-15
UUUGU	3.99E-14	GUUUU	5.38E-13
GUUUU	3.17E-12	GGUUU	2.36E-11
CUUUUU	1.30E-09	CCUUUU	4.39E-09
UUUUUA	6.05E-09	UUUUAA	2.02E-08
UGGUU	1.24E-07		

CCCCCC	1.40E-18	CCCCUC	4.80E-13
CCCUCC	1.17E-11	CCCCCA	4.92E-11
CCCCAC	5.00E-11	GCCCUC	8.21E-10
CCCCAG	9.68E-10	UCCCCC	1.05E-09
CUCCCC	5.04E-09	CACCCC	1.18E-08
CCCCCU	1.76E-08	GCCCCC	7.10E-08
CCACCC	4.15E-07		

GCCGCC	1.58E-13	CUGCCG	2.52E-11
UGCCGG	2.06E-09	GGCCUG	3.59E-09
CCUGCC	3.83E-09	GGGCCU	4.56E-09
UGGGCC	8.82E-09	GGCCGC	1.20E-08
GCCUGC	3.69E-08	CUGCCC	4.93E-08
GGUGCC	7.30E-08	GCCGGC	8.22E-08
CCGCCC	8.59E-08	CCGCCG	2.05E-07
GCUGCC	2.98E-07	GCGCCU	3.36E-07
GGCGCC	4.20E-07	UGCUGC	4.33E-07

Figure 3.11 U-rich, C-rich, and CUG-containing motifs are over-represented in the 3'UTR of mRNAs destabilized in iPS cells relative to HFFs. Sequence logos were generated from the top 50 hexamer sequences over-represented in the dataset of 548 destabilized iPS mRNAs.

3.6.3 qRT-PCR confirms destabilization of mRNAs containing U-rich sequence elements

The half-lives of two mRNAs containing U-rich sequence elements in their 3'UTR, *WEE1* and *LATS2*, as measured by qRT-PCR analysis in actinomycin d-treated cells were consistent with the hexamer data (Figure 3.11) indicating that these elements are more highly associated with instability in iPS cells than HFFs (Figure 3.12). This increased stability was also seen in mouse ES data where *WEE1* mRNA half-life was increased from ~108 minutes to ~126 minutes and *LATS2* mRNA half-life was increased from ~180 minutes to ~252 minutes upon differentiation with retinoic acid (Sharova et al., 2009). Our microarray data indicated that HFF and iPS cells exhibit half-lives of 246 minutes and 199 minutes for *WEE1* mRNA while *LATS2* mRNA had a half-life of 503 minutes in iPS cells. Although *LATS2* was identified as a stabilized mRNA in an earlier analysis of the microarray data, it did not meet the criteria for good half-life determination in the final analysis described here in HFFs. *WEE1* encodes a checkpoint kinase that negatively regulates mitotic entry through phosphorylation of cyclin B/Cdk2 (McGowan and Russell, 1995) and *LATS2* is a tumor suppressor required for embryonic development (McPherson et al., 2004). Interestingly, knockdown of *LATS2* by RNAi was recently shown to increase reprogramming efficiency (Qin et al., 2012).

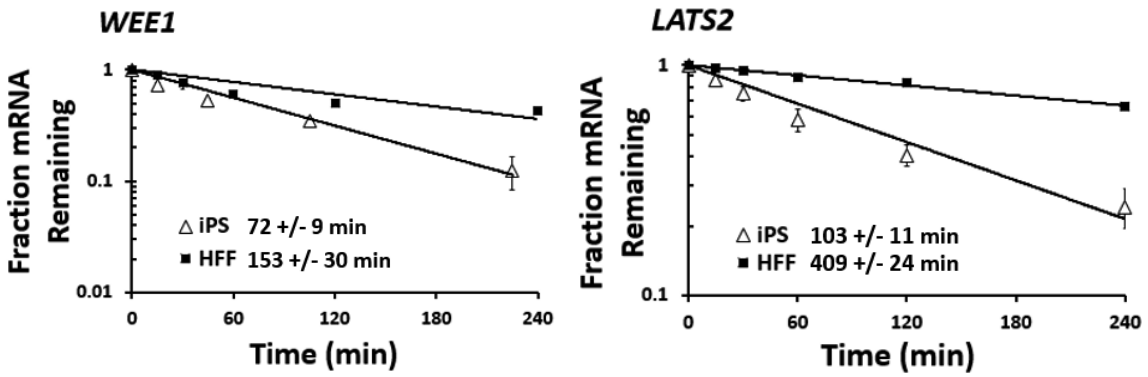


Figure 3.12 *WEE1* and *LATS2* mRNAs containing U-rich 3'UTR elements are destabilized in iPS cells compared to HFFs. Half-lives were assessed in HFF and iPS cells for *WEE1* and *LATS2* mRNAs containing U-rich 3'UTR elements following actinomycin D treatment. mRNA levels were measured by qRT-PCR at each time point and normalized to *GAPDH* mRNA. The standard deviations were derived from three independent replicates. Half-lives for each mRNA are denoted to the right of the cell-line keys.

3.7 Several RNA-binding proteins are more abundant in iPS cells

The expression of PTBP1 and several RNA-Binding Proteins (RBPs) known to interact with U-containing sequence elements was measured using whole cell lysates and western blot analysis. The abundance values, normalized to α-TUBULIN and relative to HFF expression, and standard deviation derived from three replicates are summarized in Figure 3.13 which also includes the preferred binding sequence of each protein. Specifically, we assayed expression of CELF1 (CUGBP1; Vlasova et al., 2008), ELAVL1 (HuR; Meisner and Filipowicz, 2010), PUM2 (Miller and Olivas, 2011), KHSRP (Gherzi et al., 2010), PTBP1 (Kosinski et al., 2003), HNRNP D (AUF1; Gratacós and Brewer, 2010), and ZFP36L2 (Hudson et al., 2004). We found that all seven of these proteins, which (except for ELAVL1) are known to confer instability, are elevated in iPS cells relative to HFFs to varied degrees. Increased abundance of KHSRP in iPS cells was also verified in an independent set of matched cell lines (A. Jalkanen personal communication). Interestingly, repression of PTBP1 in fibroblasts is sufficient to reprogram the cells into neurons (Xue et al., 2013) supporting that altered abundance of RBPs can have profound effects on gene expression. The increased abundance of destabilizing RBPs that recognize U-rich elements correlates well with the fact that U-rich elements are more associated with instability in iPS cells (Figure 3.6).

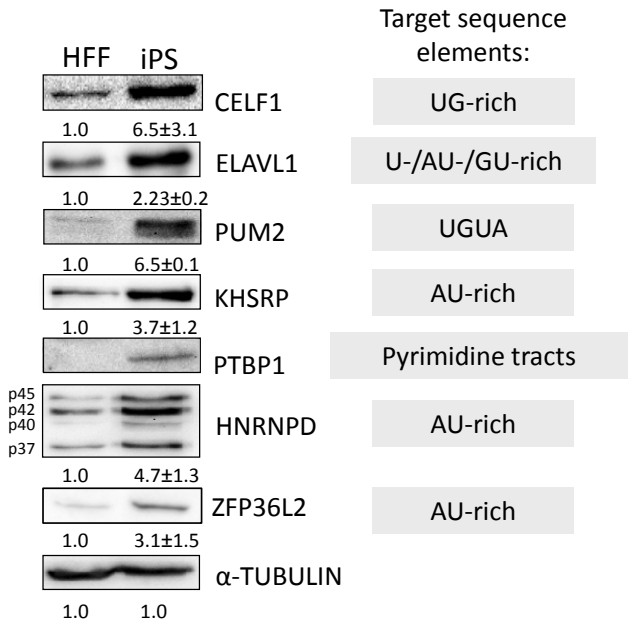


Figure 3.13 Several RNA-binding proteins are more abundant in iPS cells compared to HFFs. Western blot analysis of PTBP1 and various RNA-binding proteins known to interact with U-containing sequences was performed in HFF and iPS cells using 25µg of whole cell lysate on a 10% SDS-PAGE gel transferred to PVDF membrane. Quantification using ImageQuant software was based on a three replicates where levels of each protein are normalized to levels of α-TUBULIN and relative to expression in HFF cells. The preferred target sequence elements of the respective RNA-binding proteins are listed at the right.

3.8 Transcription rates and decay rates are inversely correlated

In addition to providing mRNA decay rates, our microarray analysis allowed mRNA abundances to be assessed from the 0 minute time points. We wondered whether changes in mRNA stability might have a predictable effect on mRNA abundances. In order to detect a correlation, we plotted the fold-change in half-life against the fold-change in abundance (Figure 3.14). Somewhat surprisingly, we found a negative correlation meaning that when mRNAs are destabilized, their abundance actually tends to increase rather than decrease. The only explanation for this is that transcription can be increased to compensate for the increased rates of decay. These results are consistent with recent studies demonstrating that regulation of transcription and degradation processes are linked (Dori-Bachash et al., 2012; Haimovich et al., 2013; Shalem et al., 2011).

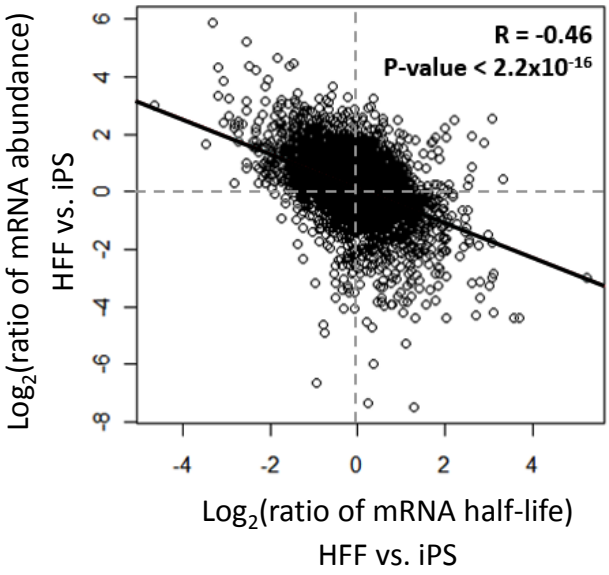


Figure 3.14 mRNA decay rates negatively correlate with mRNA abundance. Scatter plot of differences in mRNA abundance (y-axis) and differences in mRNA half-life (x-axis) for 5,481 genes quantified as log_2(ratio of HFF/iPS) indicate that mRNA half-life correlates negatively with mRNA abundance.

3.9 Concluding remarks

Establishment of this large dataset of mRNA half-lives has emphasized the utility of global mRNA decay analyses and opened up several new avenues for investigation. By determining genome-wide mRNA decay rates in HFF and iPS cells, we identified three groups of differentially regulated transcripts. Using Gene Ontology and functional domain analyses, we revealed that replication-dependent histone mRNAs and C2H2 ZNF mRNAs are significantly stabilized in iPS cells compared to HFFs. Given the clear ability of histones and transcription factors to influence gene expression, their stabilization was investigated further and is discussed in more detail in Chapters 4 and 5, respectively. Sequence analysis of hexamer motifs enriched in unstable mRNAs also revealed possible mechanisms of coordinated regulation. U-rich, C-rich, and CUG-containing elements were over-represented in the 3'UTR of mRNAs destabilized in iPS cells. We focused on C-rich elements primarily because their association with

instability has not been reported in any cell types other than iPS cells. Possible mechanisms regulating destabilized mRNAs bearing this element are discussed further in Chapter 6.

This dataset has provided a foundation for expanding this project into characterizing the mechanisms by which coordinated regulation is achieved and has also already been used by others to study mRNA decay rates (Dori-Bachash et al., 2012). Although we have focused our efforts on just three groups of differentially regulated mRNAs, there are certainly other groups that may be worthy of further study. It was intriguing to find that CA-repeat elements were enriched in the 3'UTR of stable HFF mRNAs but not stable iPS cell mRNAs, possibly demonstrating a mechanism of coordinated regulation of stability. Moreover, there is also likely a great deal to be learned from examining regulated decay of certain individual transcripts whose turnover differs between HFF and iPS cells. For example, HES1 encodes a transcription factor which represses stem cell differentiation and its mRNA is destabilized significantly in iPS cells (169 minutes versus 261 minutes in HFFs) but shows a ~3-fold increase in abundance. The *HES1* mRNA has a long A-rich region in the 5'UTR as well as C-rich, AU-rich, and GU-rich elements in its 3'UTR, any or all of which might contribute to its post-transcriptional regulation.

Chapter 4: Histone mRNAs are stabilized in iPS cells

4.1. Introduction

Several studies have revealed dramatic differences in chromatin status and histone expression between stem cells and differentiated cells that are clearly important for the achievement and maintenance of pluripotency (Boyer, 2009; Delgado-Olguin and Recillas-Targa, 2011; Meshorer and Misteli, 2006). Two very obvious differences are (i) that stem cells express a unique subset of histones (Yang et al., 2011; Zhang et al., 2012) and (ii) that the epigenetic profile of stem cells is very different from that of differentiated cells (Larson and Yuan, 2012; Song et al., 2012). In the context of these differences, we were intrigued to find that the stability of many histone mRNAs is enhanced in iPS cells as compared to HFFs (discussed in Chapter 3). As an introduction to the further analysis of these results, the function and regulation of histone gene expression is discussed below, highlighting the unique aspects of these processes in pluripotency.

4.1.1 Histones are essential for genome packaging and gene expression regulation

One role of histone proteins is to efficiently compartmentalize genomic DNA. It is estimated that each human diploid cell contains two meters of DNA which must be packaged as chromatin into a nucleus that is less than 10 microns across. Histones exert organizational control over this large amount of genetic information through the formation of nucleosomes that allow the DNA in chromosomes to be packed into deliberate, rather than random, positions. Histone proteins enable genome compaction through their interaction with DNA and chromatin remodeling factors. Each nucleosome contains 147 bases of DNA wrapped in two turns around an octamer containing two molecules each of core histones H2A, H2B, H3, and H4. DNA is effectively locked into place by linker histone H1 (Luger et al., 1997).

In addition to reducing the physical space requirement of the genome, chromatin structure is also established to support cell- and stage-specific gene expression patterns by spatiotemporally controlling access of the transcription machinery to certain gene loci. The composition of the nucleosomes at specific positions can have profound effects on transcription and RNA processing (van Bakel et al., 2013; Koerber et al., 2009). In this respect, it is important to note that each of the core histones as well as the linker histone are encoded by multiple genes resulting in a staggering number of possible arrangements (Lichtler et al., 1982; Marzluff et al., 2002). Moreover, each histone protein can experience multiple post-translational modifications ranging from acetylation to ubiquitination to further increase the number of permutations (Grewal and Moazed, 2003; Lennartsson and Ekwall, 2009; Rossetto et al., 2012).

4.1.2 Histone expression

Expression of replication-dependent histone mRNAs is tightly coupled to DNA synthesis and is therefore cell cycle-regulated (Marzluff and Duronio, 2002). As stem cells have abbreviated cell cycles (Becker et al., 2006), the temporal regulation of histone expression is also altered compared to differentiated cell types. Although histone genes are highly conserved within histone families, stem cells express a unique subset of histones that likely support pluripotency (Yang et al., 2011). Regulation of histone mRNA biology, aspects of which have been characterized from transcription through decay, may provide some clues to stem cell-specific patterns of histone gene expression.

4.1.2.1 Histone gene transcription and maturation of pre-mRNAs

The replication-dependent histone proteins are encoded by a large family of highly conserved genes that are present in chromosomal clusters (Tripputi et al., 1986). The genes encoding the four core histones exhibit the most homology between one another with few differences at the amino acid level

while the linker histone H1 show more variation between the five subtypes (Marino-Ramirez et al., 2006; Thatcher and Gorovsky, 1994). Although not all mechanisms of histone transcription induction have been characterized, it is evident that the expression of replication-dependent histones is tightly linked to the cell cycle wherein histone mRNAs are transcribed only during S phase (Marzluff and Duronio, 2002). As such, histone genes exhibit cell cycle-dependent promoter activity. Histones are the most significantly up-regulated set of genes in the G1/S transition (Medina et al., 2012) with transcription rates increasing 3 – 10-fold during this period (Harris et al., 1991) and histone mRNA stability increasing as much as 5-fold (Heintz et al., 1983).

Transcription of histone mRNAs at the G1/S boundary is triggered by phosphorylation of $p220^{NPAT}$ by cyclin E/Cdk2 (Ma et al., 2000). Nuclear protein $p220^{NPAT}$ is a co-activator of histone transcription factor HINFP which associates with HINFP upstream of transcription start sites to activate transcription of histone genes (Medina et al., 2008). The activity of cyclin E/Cdk2 effectively links histone regulation to cell cycle progression (Miele et al., 2005). The $p220^{NPAT}$/cyclin E/Cdk2 interaction constitutes a forward-feed loop that sustains HINFP expression for synthesis of histone pre-mRNAs throughout S phase (Xie et al., 2009). FLASH is another component of histone transcription machinery that associates with promoter sequences and is also required for cell cycle progression (Barcaroli et al., 2006). The role of FLASH is essential for proper histone transcription during embryogenesis (De Cola et al., 2012).

Histone genes do not contain introns so their post-transcriptional processing is limited to 3' end formation (Dominski and Marzluff, 1999; Martin et al., 1997). Unlike typical mRNAs, mammalian replication-dependent histone transcripts are not polyadenylated and instead end in a highly conserved stem loop structure that controls their post-transcriptional fate. Stem Loop Binding Protein (SLBP) interacts with this 3' stem loop element to regulate histone mRNA processing, translation, and decay (Figure 4.1). Binding of SLBP to histone pre-mRNA initiates 3' end processing by causing a

conformational change to expose the Histone Downstream Element (HDE; Jaeger et al., 2006). The U7 snRNP containing U7 snRNA, five Sm proteins, and two LSm proteins (LSM10 and LSM11) is then able to bind the HDE (Bond et al., 1991; Pillai et al., 2001; Schaufele et al., 1986) through base pairing of the U7 snRNA (Georgiev and Birnstiel, 1985; Mowry and Steitz, 1987). U7 snRNP also recruits Zinc Finger Protein ZFP100 that stabilizes the U7 snRNP/SLBP/stem loop complex through an interaction between the complex and SLBP (Dominski et al., 2002). LSM11 of the U7 snRNP interacts with FLASH to recruits CPSF73, a factor that also participates in the standard polyadenylation event on mRNAs, for endonucleolytic cleavage five nucleotides after the conserved stem loop sequence (Dominski et al., 2003, 2005; Yang et al., 2013). 3' exonuclease 3'hExo/ERI1 is also recruited to precisely trim the cleaved end (Yang et al., 2006, 2009). Association of 3'hExo/ERI1 causes further conformational changes to the 3'UTR which structurally limits the amount of trimming the exonuclease can perform (Tan et al., 2013). The result is a mature histone mRNA ending with a highly conserved structure consisting of a 24-base hairpin loop that contains a UUUC tetraloop (Zanier et al., 2002).

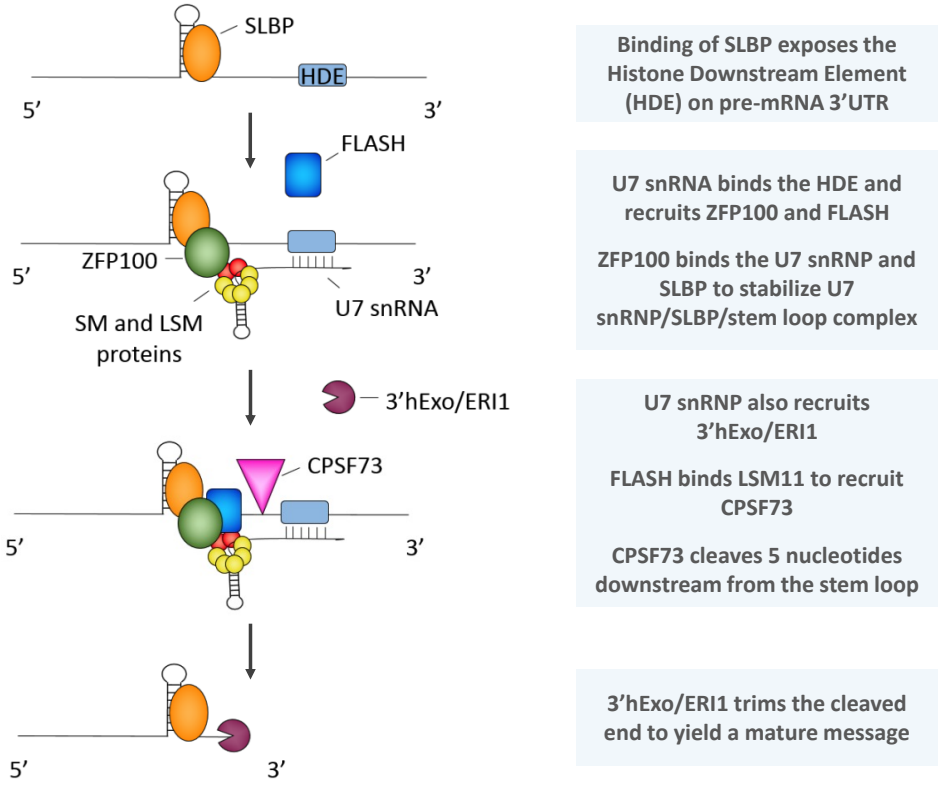


Figure 4.1 Histone mRNA 3′ end processing. SLBP binds the stem loop of replication-dependent histone mRNAs to expose the histone downstream element (HDE) which is then bound by U7 snRNP and stabilized by ZFP100. CPSF73 cleaves five nucleotides after the stem loop and 3′hExo/ERI1 trims the cleaved end to yield a mature histone transcript.

Both transcription and processing of histone pre-mRNAs are performed within sub-nuclear foci close to histone gene clusters known as Histone Locus Bodies (HLBs). These bodies arise at active sites of histone transcription and are present through interphase but disappear at mitosis (Nizami et al., 2010). Several histone processing factors localize to these foci including FLASH, $p220^{NPAT}$, LSM10/ LSM11, and U7 snRNP (Machyna et al., 2013). During *Drosophila melanogaster* development, factors are recruited to HLBs in a hierarchical fashion (White et al., 2011). FLASH and LSM11 interact directly which guides FLASH to the processing bodies (Burch et al., 2011); localization of FLASH is important for the establishment of these foci (Barcaroli et al., 2006). LSM10/LSM11 are also important for localization of the U7 snRNP to histone locus bodies (Godfrey et al., 2009). Together, these components work to convert newly transcribed histone mRNAs into their mature forms.

4.1.2.2 Histone mRNA translation

A mature histone mRNA is exported from the nucleus while SLBP is still bound (Sullivan et al., 2009) so that it may facilitate translation (Whitfield et al., 2004). Synthesis of histone proteins is initiated when SLBP is phosphorylated at up to 23 sites by multiple serine/threonine kinases, allowing recruitment of SLBP-Interacting Protein 1 (SLIP1; Bansal et al., 2013). Binding of SLIP1 to SLBP facilitates bridging of the 5' end and 3' stem loop of histone mRNA (Cakmakci et al., 2008) similar to circularization of polyadenylated transcripts that occurs though cap and poly(A) tail binding factors (Wells et al., 1998). A physical interaction between SLBP, eIF4G, and eIF3 further stimulates translation of histone mRNAs (Ling et al., 2002). Synthesis of histone H4 proteins begins when eIF4E binds the first of two structural elements in the open reading frame to allow positioning of the ribosome at the start codon (Martin et al., 2011). Translation of replication-dependent histones ceases at the end of S phase, largely due to destabilization of transcripts resulting in their decay (Pandey and Marzluff, 1987).

4.1.2.3 Histone mRNA decay

Replication-dependent histone mRNAs are degraded upon completion of DNA synthesis. Just as transcripts are stabilized at the start of S phase to accumulate histone mRNAs (Heintz et al., 1983), stability is again altered at the end of S phase leading to rapid degradation (Pandey and Marzluff, 1987). Because replication-dependent histone transcripts end in a 3' stem loop instead of a poly(A) tail, these mRNAs escape the widely used deadenylation-dependent degradation pathway (discussed in Chapter 1) and are instead degraded by a deadenylation-independent mechanism that requires active translation of the message (Kaygun and Marzluff, 2005a). As described in Section 4.1.2.2, mRNAs in this state are circularized through an interaction between SLBP with eIF4G and eIF3 (Ling et al., 2002) bridged by SLIP1 (Cakmakci et al., 2008). Once DNA synthesis is complete, checkpoint kinase ATR phosphorylates UPF1, a key regulator of Nonsense-Mediated Decay (NMD), which promotes the interaction between UPF1 and

the histone stem loop/SLBP complex (Kaygun and Marzluff, 2005b). Binding of UPF1 results in the recruitment of TUTase ZCCHC11 and oligouridylation of the 3'end (Schmidt et al., 2011), providing an oligo(U) tract that permits binding of the LSM1-7 complex. Similar to the decay mechanism for deadenylated mRNAs, binding of LSM1-7 recruits DCP1/DCP2 to initiate decapping and decay by XRN1 in the 5'→3' directions (Coller et al., 2001; Mullen and Marzluff, 2008). Although decapping mediated by oligouridylation is the predominant mechanism of histone mRNA decay (Su et al., 2013), simultaneous decay in the 3'→5' direction can also occur through recruitment of 3'hExo/Eri-1 by the LSM1-7 complex leading to exonucleolytic degradation of the stem loop structure (Hoefig et al., 2013) and subsequent 3'→5' degradation by the exosome (Mullen and Marzluff, 2008).

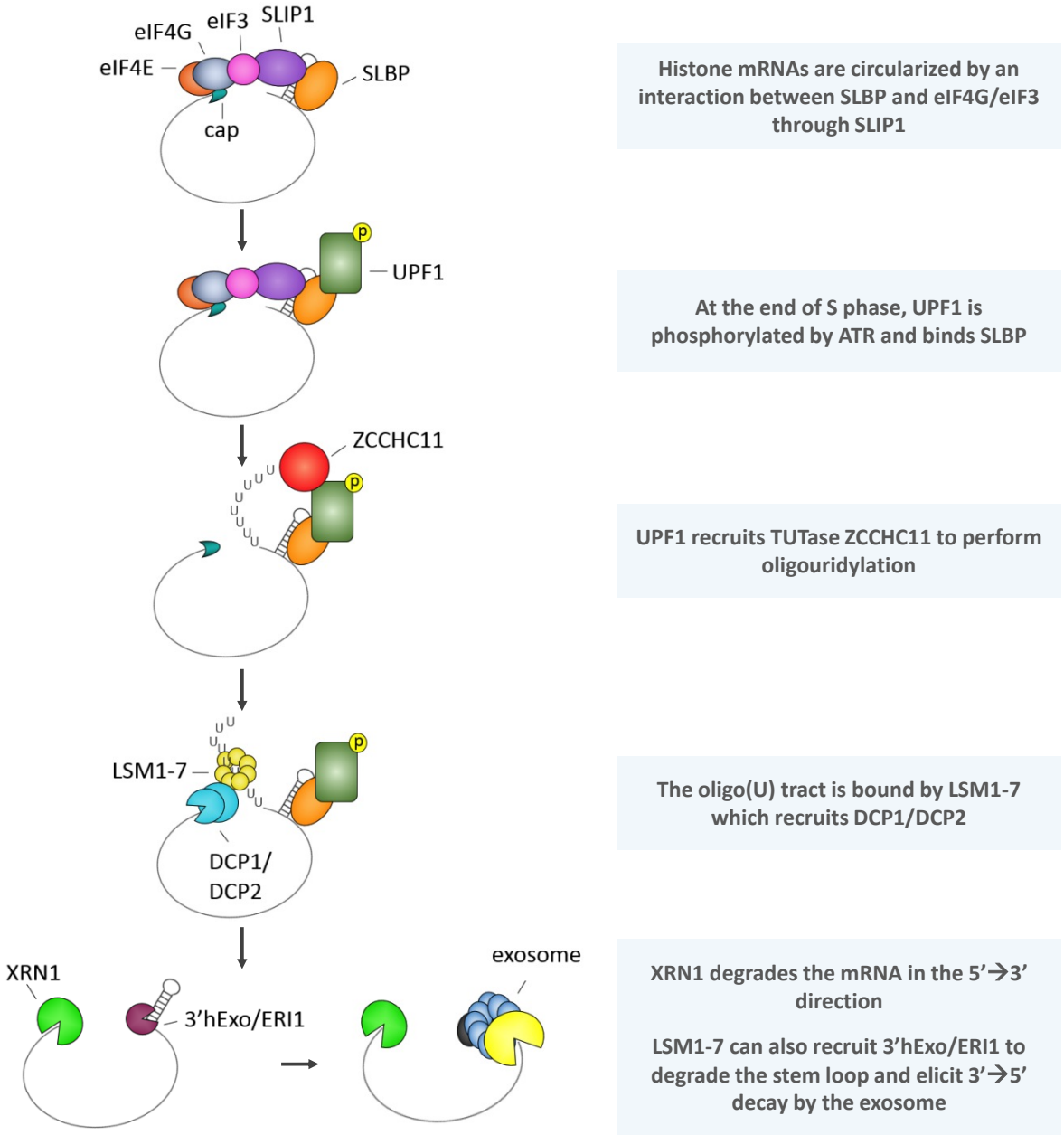


Figure 4.2 Histone mRNA decay. At the end of S phase, checkpoint kinase ATR phosphorylates UPF1 to promote its interaction with the stem loop/SLBP mRNP complex of actively translating histone mRNAs. UPF1 recruits ZCCHC11 to oligouridylate the 3' end providing a substrate for the LSM1-7 complex. In the predominant mechanism of histone mRNA degradation, recruitment of DCP1/DCP2 by LSM1-7 leads to 5'→3' decay by XRN1. In addition, recruitment of 3'hExo/ERI1 by LSM1-7 degrades the stem loop structure and allows simultaneous decay in the 3'→5' direction by the exosome.

4.1.2.4 SLBP expression correlates with histone mRNA expression

Given its importance at every step of histone gene expression, it comes as no surprise that SLBP expression is also dependent on the cell cycle with levels increasing 10- to 20-fold shortly before S phase and decreasing following DNA synthesis (Whitfield et al., 2000). Although histone and SLBP expression parallel one another, studies have shown that expression of SLBP outside of S phase activates histone pre-mRNA processing, indicating that SLBP is a master regulator of histone expression (Wang et al., 1996). SLBP remains associated with histone mRNAs to facilitate processing and translation. At the S/G2 transition when histones are targeted for decay, PIN1, a prolyl isomerase important for regulation of cell cycle and differentiation (Lu et al., 2007), induces SLBP polyubiquitination to cause dissociation of SLBP-histone mRNA complexes (Krishnan et al., 2012). SLBP is then phosphorylated by cyclin A/Cdk1 (Koseoglu et al., 2008) on two threonine residues leading to degradation of the protein (Zheng et al., 2003). Studies on the role of SLBP in oogenesis also show that expression of this protein is developmentally regulated (Allard et al., 2005).

4.1.2.5 Histone profiles are established through regulated expression of histone subtypes

Although all replication-dependent histone mRNAs share the highly conserved 3' stem loop that modulates their metabolism at every step (Marzluff et al., 2008), their expression is also controlled at other levels. The complement of histone proteins is unique to different cell types (Lin et al., 2004; Yang et al., 2011; Zhang et al., 2012) suggesting that in addition to the coordinated control that facilitates their production during S phase, there are further layers of control specific to each histone subtype. Most histone subtype studies have investigated expression of linker H1 subtypes (Happel and Doenecke, 2009). Levels of HIST1H1T are specifically regulated in male germ cells to maintain a proper gene expression profile (Lin et al., 2004) and liver cells demonstrate distinct changes in H1 subtype expression in response to stress (Jeong et al., 2003). The control of expression of individual core histone subtypes is

virtually unstudied, primarily due to the challenges of distinguishing such closely related genes. However, subtype-specific regulation of transcription, processing, mRNA decay, translation, or protein decay could all influence the final complement of histone proteins in the cell.

4.1.2.6 Histone variants

Although replication-dependent histones are coordinately regulated through the various pathways outlined above, there are additional, replication-independent histones that have alternate mechanisms for expression. Variant histones, or replication-independent histones, are different from typical histones in that their mRNAs end with a polyadenylated tail rather than the conserved 3' stem loop structure. Without this regulatory element, variant histone mRNAs are not subject to the same cell cycle control. These histones are also transcribed from loci separate from those of the standard histones, further contributing to their atypical regulation of transcription. Histone variants have specific roles in chromatin metabolism such as DNA repair and active transcription and their expression often fluctuates based on cell cycle, differentiation and development (Law and Cheung, 2012). Just as stem cells express a unique complement of core and linker histone genes compared to differentiated cells (Yang et al., 2011), they also express distinct histone variants (Binda et al., 2013; Cao et al., 2013) that establish inherent differences in stem cell chromatin compared to the genome of differentiated cells.

4.1.3 Pluripotent cells have a unique histone profile

Although histones are obviously important for chromatin compaction in every cell type, their roles in modulating genes expression vary widely and, as alluded to above, the relative abundance of different subtypes and of the replication-independent variants also differs. As such, stem cells have several unique features with regards to their histone profile and chromatin conformation. First, the complement of histones in stem cells is distinctive. Of the five canonical H1 histone subtypes,

HIST1H1A, HIST1H1D, and HIST1H1B have elevated expression in stem cells (Saeki et al., 2005; Terme et al., 2011; Yang et al., 2011). In addition, expression of the variant histone H1F0 is low during pluripotency (Terme et al., 2011) but increases when differentiation is induced (van Hemert et al., 1992). The stem cell histone H2A profile is also different. Histone MacroH2A variants have a role in terminal differentiation and act as a barrier to reprogramming of differentiated cells into induced pluripotent stem cells (Gaspar-Maia et al., 2013). Their role in the formation of chromatin signatures specific to stages of development or disease has been the focus of several literature reviews to date (Law and Cheung, 2012; Millar, 2013; Szenker et al., 2011; Vardabasso et al., 2013). Second, stem cells have a relatively open chromatin configuration supported in part by their abbreviated cell cycle. The shortened G1 and predominant S phase allow stem cells to maintain chromatin plasticity (Becker et al., 2006) where the truncated G1 allows insufficient time for the complete compaction of heterochromatin (Hindley and Philpott, 2013). The chromatin structure of pluripotent stem cells is less tightly bound and more euchromatic in formation while differentiated cell types exhibit more compact chromatin exposing fewer regions of chromosome to transcription factors (Gaspar-Maia et al., 2011). Third, stem cells exhibit an exceptional epigenetic profile that supports their key features of infinite self-renewal with the ability for unlimited differentiation, by allowing the cell to be in "poised" for differentiation (Fisher and Fisher, 2011; Spivakov and Fisher, 2007). Pluripotent cells exhibit transcriptional hyperactivity with increased expression of chromatin remodeling factors and transcriptional machinery (Efroni et al., 2008) but also employ chromatin regulators that suppress lineage-specific gene sets (Lessard and Crabtree, 2010). This hyper-dynamic quality of pluripotent chromatin is a hallmark of stem cells wherein genomic plasticity must be maintained should differentiation be induced (Meshorer et al., 2006). In addition to expressing unique core, linker, and variant histones, many stem cell-specific histone modifications define pluripotent chromatin architecture. For example, the importance of histone acetylation is highlighted by the fact that Histone DeACetylase (HDAC) inhibitors can enhance reprogramming of cells

into a pluripotent state (Plath and Lowry, 2011) presumably by blocking the removal of acetyl groups that are associated with condensed chromatin structure. DNA methylation on cytosine residues is a critical epigenetic regulator of stem cell differentiation and embryogenesis (Lister et al., 2009). Characterizations of highly dynamic and stage-specific chromatin features have been used to identify states of differentiation and development (Guibert and Weber, 2013; Larson and Yuan, 2012). In the poised state, any combination of histone modifications has the potential to cause chromatin re-organization during stem cell differentiation (Binder et al., 2013).

Global analysis of mRNA decay rates in HFF and iPS cells (discussed in Chapter 3) revealed that many replication-dependent histone mRNAs exhibited significant stabilization in a pluripotent state. As the chromatin status of stem cells is unique from that of differentiated cells and plays a significant role in defining and maintaining pluripotent gene expression, we hypothesized that differential regulation of histone mRNA stability might be an important factor in defining the histone profiles in iPS cells. We therefore further investigated histone mRNA stabilization and the effects on overall histone expression. Our results are discussed below[3].

4.2 Histone mRNA half-lives are increased in iPS cells

We found that a significant proportion of replication-dependent histone mRNAs are stabilized in iPS cells. For the 50 transcripts encoding histone proteins which half-lives were determined for (data for three transcripts out of 53 total were duplicates of the same gene with the same assessed half-life), an arbitrary cut-off of a 2-fold difference was used to find that 23 histone mRNAs were stabilized >2-fold, 26 exhibited a less than 2-fold change, and a single histone mRNA, *HIST3H2BB*, showed a 2.7-fold destabilization in iPS cells. These data are summarized in Table 4.1 below where the most stabilized

[3] Many of the results presented in Chapter 4 appeared in: Neff *et al*. Global analysis reveals multiple pathways for unique regulation of mRNA decay in induced pluripotent stem cells. Genome Research. Vol. 22, No. 8, pg. 1457-1467, August 2012.

transcripts are highlighted in light shading while the destabilized transcript *HIST3H2BB* is in dark shading. All four core histones and the linker histone were represented in the stabilized set of mRNAs.

Table 4.1 Summary of histone mRNA half-lives as determined by microarray. Twenty-three mRNAs were stabilized >2-fold (light shading) and 1 transcript was destabilized >2-fold (dark shading) in iPS cells compared to HFFs.

Gene name	iPS Half-life (min)	iPS P-value	HFF Half-life (min)	HFF P-value	Fold Change
HIST1H3I	273	1.05E-02	27	1.05E-05	10.06
HIST1H1E	536	3.32E-03	64	3.77E-06	8.38
HIST2H2AC	354	9.33E-03	47	4.76E-07	7.49
HIST2H2AB	363	6.63E-03	59	2.05E-06	6.11
HIST1H1A	1122	2.50E-02	193	4.66E-05	5.82
HIST1H4K	533	2.73E-02	93	5.01E-05	5.74
HIST2H3A	451	6.71E-04	97	4.15E-05	4.65
HIST1H3J	435	6.47E-03	102	2.42E-05	4.26
HIST1H2BE	997	1.05E-03	242	2.43E-05	4.13
HIST1H2AB	282	3.29E-03	70	2.64E-06	4.05
HIST2H3D	448	7.09E-04	112	5.09E-05	4.00
HIST1H3G	1808	1.41E-02	478	3.12E-05	3.79
HIST1H4J	502	1.82E-02	135	2.00E-05	3.72
HIST1H3D	366	6.09E-03	102	4.18E-05	3.60
HIST1H4B	277	6.42E-05	81	2.59E-07	3.42
HIST2H3D	488	5.12E-03	147	1.82E-04	3.32
HIST1H1D	274	4.70E-03	85	1.84E-05	3.24
HIST2H2AA3	373	1.30E-03	117	9.70E-05	3.20
HIST1H3F	443	4.11E-03	150	1.07E-05	2.95
HIST1H4E	343	1.18E-02	132	4.67E-06	2.60
HIST1H3B	193	6.14E-03	87	1.10E-04	2.23
HIST1H4C	323	3.89E-02	157	4.23E-05	2.06
HIST1H2AK	413	1.36E-02	202	6.36E-05	2.05
HIST1H4A	621	2.75E-02	319	6.12E-05	1.95
HIST1H1C	372	3.93E-02	192	4.60E-04	1.94
HIST1H2BK	345	1.67E-02	186	1.10E-03	1.85
HIST2H2BE	399	1.52E-02	223	5.82E-03	1.79
HIST1H2AI	342	1.12E-03	191	1.06E-04	1.79
HIST1H2AL	596	3.59E-03	344	6.76E-07	1.73
HIST1H3E	865	2.11E-02	519	1.05E-03	1.67
HIST1H3A	336	1.89E-02	208	3.78E-05	1.62
HIST1H2BG	735	4.61E-02	496	1.73E-03	1.48
HIST1H4D	273	3.09E-02	192	4.30E-06	1.42
HIST2H2BE	510	1.96E-02	362	6.13E-03	1.41
HIST4H4	714	4.65E-04	555	1.05E-03	1.29
HIST1H3H	300	7.10E-04	253	5.40E-04	1.18
HIST1H2BD	329	3.20E-03	292	5.06E-04	1.13
HIST1H4I	686	4.08E-03	644	8.90E-03	1.06
HIST2H2BF	248	1.16E-04	236	2.48E-05	1.05
HIST1H2AH	337	1.18E-02	402	2.69E-05	0.84
HIST1H2AC	245	9.74E-03	292	2.43E-03	0.84
HIST1H2BJ	288	3.90E-04	348	5.04E-04	0.83
HIST1H2AE	465	2.67E-04	637	1.16E-02	0.73
HIST1H4H	238	1.86E-03	332	1.15E-02	0.72
HIST2H2BF	244	5.19E-05	351	7.87E-04	0.69
HIST1H2AG	376	9.80E-04	589	3.39E-03	0.64
HIST2H4A	290	9.93E-04	487	1.42E-02	0.59
HIST1H2BN	527	3.69E-03	909	2.75E-02	0.58
HIST1H2BL	501	1.44E-02	922	9.00E-03	0.54
HIST3H2BB	289	5.28E-04	784	1.16E-03	0.37

When compared to all mRNAs that half-lives were determined for, we saw that the histone mRNAs as a group exhibit more stabilization in iPS cells versus HFFs ($P=2.2x10^{-11}$; Figure 4.3A). Figure 4.3B shows the half-life for each histone family when the half-lives of all 50 individual transcripts are averaged. As depicted by the error bars indicating standard deviation between related transcripts, histone mRNAs within the same family exhibited a wide range in half-life. This was most evident for histone H3 in iPS cells where estimated half-lives ranged from ~193 minutes (*HIST1H3B*) to ~1800 minutes (*HIST1H3G*) and histone H2B in HFFs where *HIST1H2BK* mRNA had a half-life of ~186 minutes while *HIST1H2BL* mRNA was more stable with a half-life of ~922 minutes. Despite the half-life variation seen within each histone set, the averages shown here indicated that histone mRNAs are generally more stable in iPS cells than in HFFs.

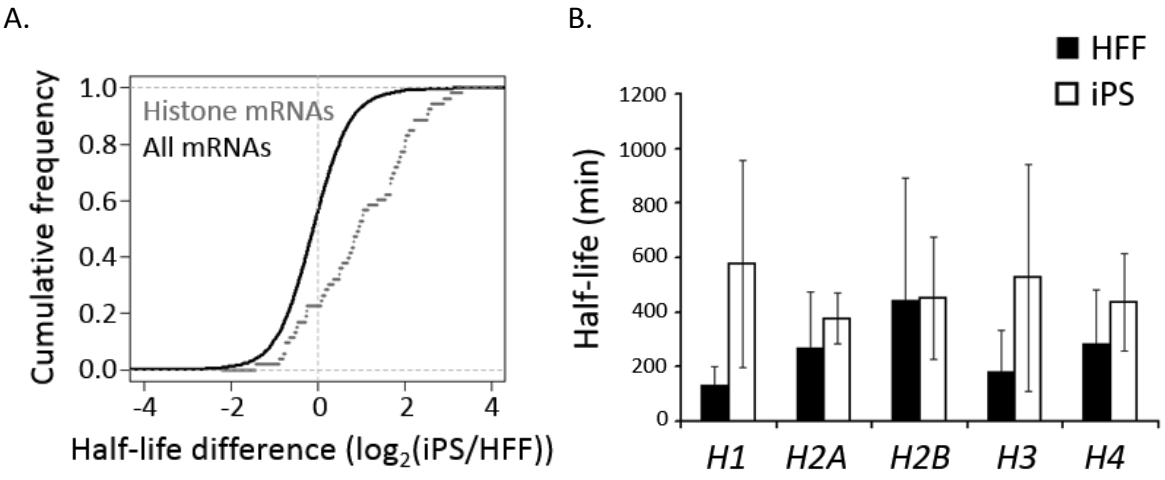


Figure 4.3 Histone half-lives are increased in iPS cells compared to HFFs as determined by microarray. A) Comparison of the change in half-life for histone mRNAs in HFF and iPS cells to changes for all 5,481 half-lives and B) Bar graph presentation of averaged mRNA half-lives for each histone family as determined by microarray.

To validate this observation, histone mRNA half-lives were measured by qRT-PCR in HFF and iPS cells using the same RNA samples that were used to generate the microarray data. Primer sets to recognize core histones H2A, H2B, H3, and H4 and linker histone H1 were generously provided by Dr.

Paul Laybourn and detect multiple transcripts within each histone family (Bogenberger and Laybourn, 2008). However, when the detection of each transcript was predicted based on homology to the primer sequences, we noted that several transcripts escape detection (shown in Table 4.2 below). Specifically, forward and reverse primer sequences ~24 bases in length were aligned to mRNA sequences for all individuals within a histone family and analyzed for the number of base mismatches. Transcripts containing more than four mismatches in a row or several mismatches close in proximity in one or both primer oligos were unlikely to be effectively detected by the primer set.

Table 4.2 Histone transcripts detected by histone family primers.

Histone Fami	Detected	Not Detected
H1	HIST1H1e	HIST1H1a
		HIST1H1b
		HIST1H1c
		HIST1H1d
		HIST1H1t
H2A	HIST1H2Ac	HIST1H2Aa
	HIST1H2Ag	HIST1H2Ab
	HIST1H2Ah	HIST1H2Ae
	HIST1H2Ai	HIST2H2Ab
	HIST1H2Ak	
	HIST1H2Al	
	HIST1H2Am	
	HIST2H2Aa3	
	HIST2H2Ac	
	HIST3H2A	
H2B	HIST1H2Bb	HIST1H2Bk
	HIST1H2Bc	
	HIST1H2Bd	
	HIST1H2Be	
	HIST1H2Bf	
	HIST1H2Bg	
	HIST1H2Bh	
	HIST1H2Bj	
	HIST1H2Bl	
	HIST1H2Bm	
	HIST1H2Bn	
	HIST1H2Bo	
	HIST2H2Ba	
	HIST2H2Bf	
	HIST3H2Bb	
H3	HIST1H3c	HIST1H3a
	HIST1H3d	HIST1H3b
	HIST1H3e	HIST1H3g
	HIST1H3f	HIST1H3h
	HIST1H3i	HIST2H3a
	HIST1H3j	HIST2H3d
		HIST3H3
H4	HIST1H4a	HIST1H4c
	HIST1H4b	HIST1H4h
	HIST1H4d	HIST4H4a
	HIST1H4e	
	HIST1H4f	
	HIST1H4i	
	HIST1H4j	
	HIST1H4k	
	HIST1H4l	
	HIST4H4	

Summarized from the microarray data, Table 4.3 below shows the mean $\log_2$ (intensity) values at the 0 minute time point averaged from three replicates for each cell line. Although our qRT-PCR primers only detected the *HIST1H1E* transcript, this mRNA was the most abundant of the H1 family members in both cell lines.

Table 4.3 Mean log2(intensity) values at the 0 minute time point for all histone H1 mRNAs.

Transcript	HFF mean	iPS mean
HIST1H1A	6.47	4.81
HIST1H1B	8.24	4.57
HIST1H1C	6.84	8.85
HIST1H1D	6.63	7.92
HIST1H1E	9.16	11.72
HIST1H1T	3.23	4.05

In Figure 4.4 we saw that by qRT-PCR, histone mRNA half-lives were indeed increased in iPS cells compared to HFFs. From the data we found that in HFFs, histones H1, H2A, H2B, and H4 have transcript half-lives ranging from 85 – 110 minutes and H3 mRNAs were the least stable with a half-life of ~44 minutes. In iPS cells, all histone mRNAs were significantly stabilized (ANOVA) with half-lives ranging from approximately 150 – 270 minutes. We importantly note that stabilization of H2A transcripts was verified in an independent set of matched HFF and iPS cells also purchased from System Biosciences (SC101A-1, Lot #110415-01), supporting that these differences are not due to the insertion positions of transgenes resulting from lentiviral reprogramming (A. Jalkanen personal communication).

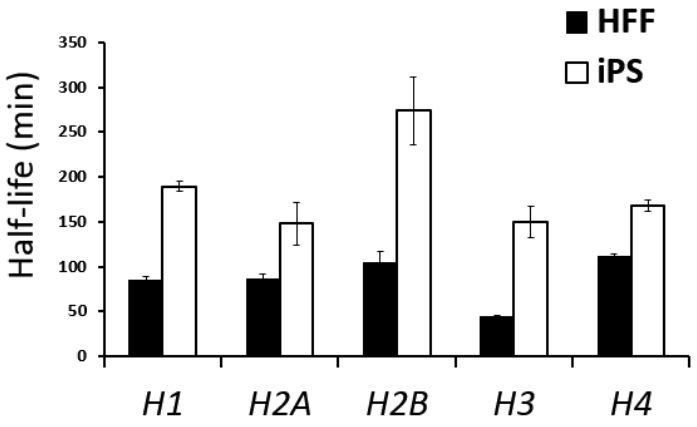


Figure 4.4 Histone mRNA half-lives are increased in iPS cells compared to HFFs as determined by qRT-PCR. Half-lives of individual histone mRNA families were assessed in HFF and iPS cells following actinomycin D treatment. mRNA levels were measured by qRT-PCR and normalized to *GAPDH* mRNA. Standard deviations represents three replicates.

The half-lives estimated here by qRT-PCR were shorter than the averages shown in Figure 4.3B above. Discrepancies between half-lives generated from microarray and qRT-PCR data may be explained by the specificity of the histone primers used which are not all-inclusive. Also, the microarray average weights each histone transcript the same regardless of abundance whereas qRT-PCR will give the most abundant mRNAs most weight, skewing the stability of less stable mRNAs. In addition, the linear range of the qRT-PCR assay is greater than that of the microarray (Etienne et al., 2004; Wang et al., 2006).

4.3 Histone abundance is increased at the mRNA and protein levels

Given that histone mRNAs were quite dramatically stabilized in iPS cells, it was important to determine whether this resulted in an increased abundance of histone mRNA and protein. The 0 minute time point of the microarray data was used to determine the relative abundance of histone transcripts in HFF and iPS cells. As seen in Figure 4.5 below, there was a significant ($P=6.5x10^{-9}$) increase in histone mRNA abundance in iPS cells compared to HFFs that correlated with the half-life increase shown in Figure 4.3 (above).

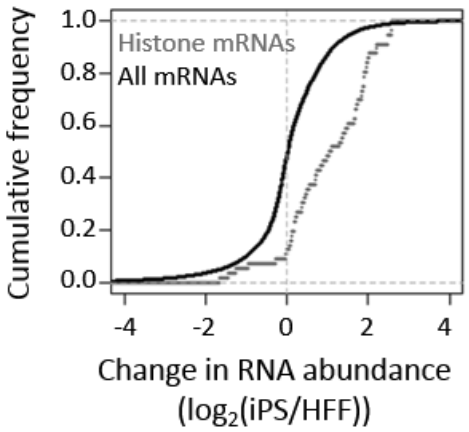


Figure 4.5 Histone mRNAs are more abundant in iPS cells compared to HFFs as determined by microarray. Comparison of the change in abundance for histone mRNAs in HFF and iPS cells to changes for all 19,190 mRNAs expressed in both HFF and iPS cells.

Relative steady-state histone mRNA abundances were also measured in cells that had not been treated with actinomycin D using qRT-PCR. In Figure 4.6, we found that iPS cells have up-regulated expression of histone mRNAs ranging from a 3.7-fold to a >200-fold increase in abundance.

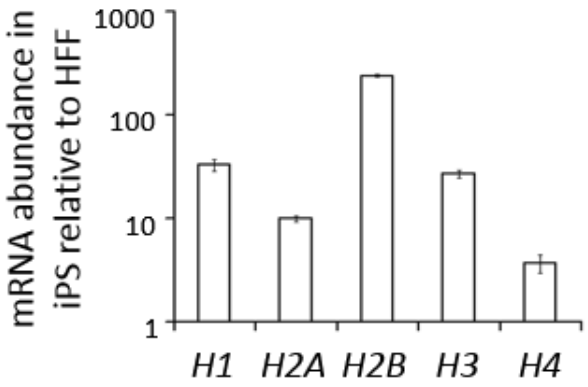


Figure 4.6 Histone mRNAs are more abundant in iPS cells compared to HFFs as determined by qRT-PCR. HFF and iPS cells were not treated with actinomycin D and were assessed for abundance of individual histone mRNA families. mRNA levels were measured by qRT-PCR and normalized to *GAPDH* mRNA. Standard deviations represents three replicates.

We next measured histone protein expression in both cell lines. Histone proteins are so abundant that they may be visualized by Coomassie staining (Irie and Sezaki, 1983). As shown in Figure 4.7, Coomassie staining of HFF and iPS cell extracts containing ~$4.2x10^4$ cells resolved on a 15% SDS-PAGE gel revealed ~10-fold higher levels of core histones in iPS cells (normalized to a ~36 kDa band, asterisk) compared to HFFs. As histone H1 is larger than core histones with a molecular weight of ~32 kDa, it was difficult to distinguish this protein from non-specific proteins of a similar size. Nonetheless, increased abundance of core histone proteins was again consistent with increased transcript stability (Figures 4.3 and 4.4) and abundance (Figures 4.5 and 4.6) observed previously. These results were reproduced in an independent set of matched HFF and iPS cells, confirming increased expression of histone proteins in iPS cells (Appendix A3).

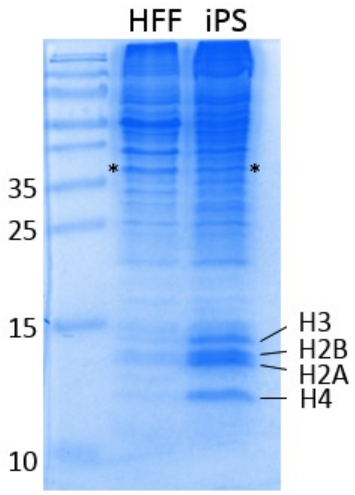


Figure 4.7 Core histone proteins are more abundant in iPS cells compared to HFFs as determined by Coomassie staining. Equal numbers of cells and equal amounts of whole cell lysate were resolved on a 15% SDS-PAGE gel and stained with Coomassie Blue to reveal protein bands corresponding to core histones H2A, H2B, H3, and H4. The asterisk denotes the ~36 kDa band used for quantification normalization.

Western blot analysis was also used to investigate histone protein abundance. We found that protein levels of all four core histones H2A, H2B, H3, and H4 are elevated in iPS cells (Figure 4.8) consistent with Coomassie staining, and that linker histone H1 was also more abundant. However, the fold-change differences for core histones seen here does not reflect the ~10-fold increase determined by Coomassie quantification. As antibodies have several idiosyncrasies that influence their range of detection, we believe that histone abundances determined by staining likely reflect true abundances since the Coomassie stain exhibits more general binding properties that would stain these proteins equally.

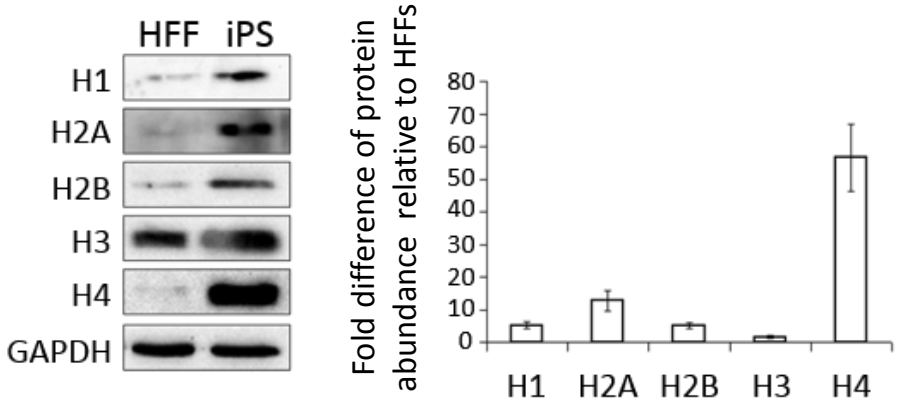


Figure 4.8 Histone proteins are more abundant in iPS cells compared to HFFs as determined by western blot. A) Western blot analysis showing histone protein abundances in HFF and iPS cells. Equal numbers of cells and amounts of whole cell lysate were resolved on a 15% SDS-PAGE gel and transferred to 0.2 µm polyvinylidene difluoride (PVDF) membranes. B) Quantification of iPS histone abundance normalized to levels of GAPDH and relative to HFF as determined by western blot. Standard deviations represent three replicates.

Although we found increased levels of histone protein in iPS cells, the fold-change does not correlate directly with the change in mRNA abundance suggesting there are differences in translation efficiency as well as mRNA stability in iPS cells. Regardless, it is evident that histone mRNAs are stabilized in iPS cells and this finding correlated with increased levels of mRNA and protein.

4.4 Expression of SLBP is increased in iPS cells

As SLBP is a critical regulator of histone mRNA metabolism (discussed in Section 4.1.2), it was thought that SLBP might be expressed at a higher level in iPS cells to facilitate the production of increased levels of histone message. Indeed, by western blot we found that SLBP is undetectable in HFFs and much more abundant in iPS cells (Figure 4.9). In support of this, SLBP mRNA levels were increased 1.89-fold in iPS cells based on the abundance determined from the 0 minute time point of the microarray data. While this analysis does not account for the contribution of other factors involved in histone mRNA metabolism such as LSM1 (mRNA abundance increased 1.54-fold in HFFs at the 0 minute time point), UPF1 (mRNA abundance increased 1.46-fold in iPS cells at the 0 minute time point), and

TUTase ZCCHC11 (mRNA abundance increased 2.40-fold in iPS cells at the 0 minute time point), the elevated levels of SLBP in iPS cells seen here likely contribute to the processing and turnover of the numerous histone mRNAs that must be transcribed, translated, and degraded over each cell cycle.

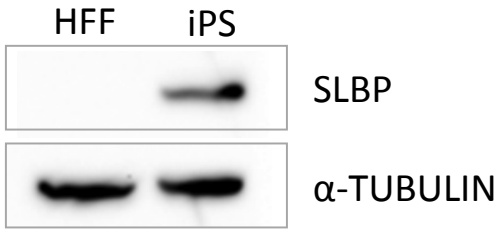


Figure 4.9 SLBP protein expression is increased in iPS cells compared to HFFs. Equal numbers of cells and equal amounts of whole cell lysate were resolved on a 10% SDS-PAGE gel and transferred to PVDF membrane. Expression of SLBP was assessed relative to levels of α-TUBULIN as determined by western blot analysis.

4.5 Inhibition of DNA synthesis leads to faster degradation of histone mRNAs in iPS cells

One of the caveats of the experiments described thus far was that the cells analyzed were an asynchronous population at different phases of the cell cycle. To mimic the coordinated decay of histone mRNAs at the end of S phase, cells were treated with hydroxyurea (HU) to inhibit DNA synthesis. Previous studies suggest that fibroblasts may not be as responsive to HU treatment (CJ Wilusz communication from WF Marzluff); therefore, HeLa cells were included in this analysis for comparison.

Starting 30 minutes following HU treatment, total RNA was collected at time points 0, 15, 30, 60, 120 and 240 minutes. This 30 minute delay was chosen to allow the inhibitor enough time to effectively halt DNA synthesis. We found that in contrast to significant stabilization overall, histone mRNAs were degraded much more quickly in iPS cells than HFFs (Figure 4.10) following inhibition of DNA synthesis. The ineffectiveness of inhibiting DNA synthesis in HFFs may explain the relatively unchanged half-lives of H1 and H2B mRNAs at 68 and 70 minutes, respectively, compared to those determined following

actinomycin D treatment. In HeLa cells, the half-lives of histone transcripts H1, H2A, H2B and H4 were fairly consistent and ranged from 32 – 39 minutes. The half-life for H3 was the shortest in all cells types and was about 23 minutes in HeLa cells. Histone half-lives determined in iPS cells demonstrated more destabilization of transcripts after HU treatment compared to the differentiated cell types. Half-lives of H1, H2B, and H4 mRNAs ranged from 20 – 23 minutes while those for H2A and H3 mRNAs were 14 and 6 minutes, respectively. From this, we found that histone mRNAs were generally more stable in iPS cells than differentiated cells but were degraded more quickly upon inhibition of DNA synthesis. Compared to HeLa cells, the difference was not as dramatic but histone transcripts were still degraded more quickly in iPS cells than in HeLa.

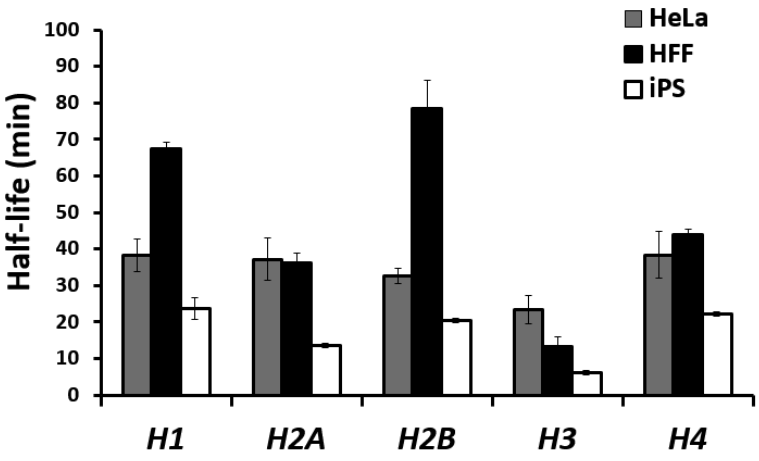


Figure 4.10 Histone mRNAs are degraded faster in iPS cells upon inhibition of DNA synthesis compared to HFF and HeLa cells. Half-lives of histone mRNA families were assessed in HeLa, HFF, and iPS cells following hydroxyurea (HU) treatment. mRNA levels were measured by qRT-PCR and normalized to *GAPDH* mRNA. Standard deviations represents three replicates.

4.6 Differences in cell cycle contribute to differences in histone abundance

Given that stem cells are reported to have an abbreviated cell cycle, and expression of replication-dependent histones is cell cycle regulated, we questioned how much this might contribute to differences in histone gene expression. Specifically, we wondered whether an increased proportion of

cells in S phase might explain the elevated abundance of histone mRNAs and proteins (which only accumulate in S phase in most cell types) in iPS cells. To answer this question, we considered several measures of cell cycle. In the first, we confirmed that our iPS cells have a faster growth rate than HFFs. By measuring cell populations over a period of 5 days, we found that HFF cells have a doubling rate of ~36 hours while iPS cells duplicate every ~22 hours.

We also sought to determine the percentage of each cell population in S phase at any given time. To do this, a BrdU incorporation assay was performed to stain the nuclei of cells that were actively synthesizing DNA within an asynchronous population. BrdU incorporation was assessed by immunofluorescence using anti-BrdU antibodies. By determining the percentage of cells with nuclear BrdU staining (three replicates each), it was found that iPS cells have ~33.2 ± 4.8% cells in S phase while HFFs have ~14.2 ± 1.0% (Figure 4.11). This 2-fold difference certainly contributes to the increased levels of histones observed in iPS cells but cannot account for the 50- to 100-fold higher level of histone mRNA and protein in iPS cells over HFFs.

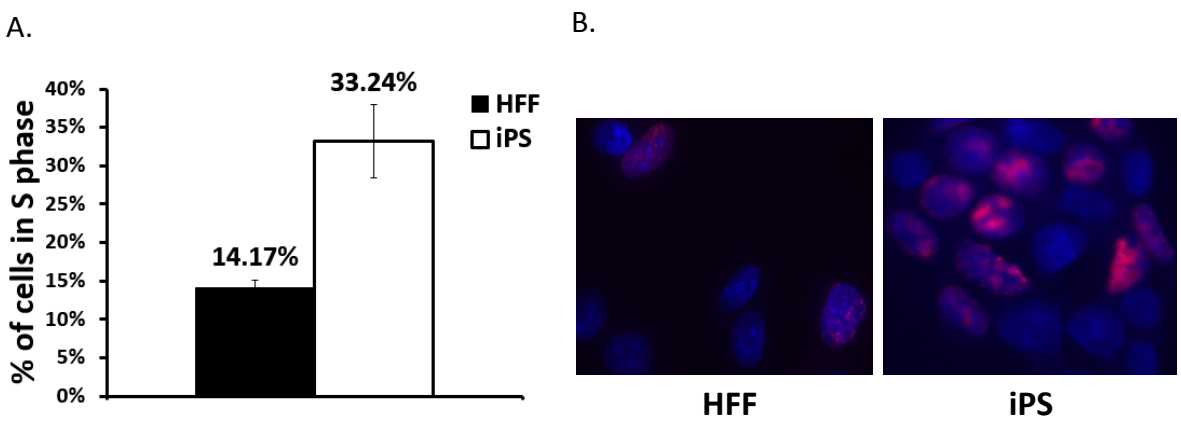


Figure 4.11 A larger proportion of iPS cells in an asynchronous population are in S phase compared to HFFs. The BrdU incorporation assay was used to stain cells in S phase. A) Percentage of cells in S phase within an asynchronous population. Standard deviations represents three replicates consisting of 100 cells each. B) Representative images of BrdU staining in HFF and iPS cells as determined by immunofluorescence. S phase positive cells are stained with anti-BrdU (red) and nuclei are stained with DAPI (blue).

Both of these analyses confirmed that iPS cells have an abbreviated cell cycle (as noted by a faster doubling rate) and more cells in S phase, consistent with previously published reports (Becker et al., 2006; Fluckiger et al., 2006; Savatier et al., 1994). However, it is clear that other features of pluripotent gene expression also influence histone regulation.

4.7 The absolute amount of histone expression in iPS cells is increased

Although each HFF and iPS cell contain the same amount of DNA, there may be significant differences in the overall amount of RNA or protein per cell that could contribute to the perceived increases in abundance of histone mRNA and proteins. Quantification of nucleolar DNA, RNA, proteins, and lipids in iPS cells compared to differentiated fibroblasts indicates that the nucleolar molecular signature of iPS cells is characterized by increases in all of these molecules except lipids (Pliss et al., 2013). We therefore sought to compare the absolute amount of histone mRNA and protein per cell rather than normalizing to equal amounts of total RNA or protein. As shown in Figure 4.12, iPS cells have twice as much RNA per cell as HFFs. Following TRIzol extraction, we found that iPS cells have ~30.6

± 2.1 pg RNA per cell while HFFs have ~14.0 ± 1.8 pg, fairly consistent with the estimation that mammalian cells contain ~20 – 30 pg RNA each (Alberts et al., 1994). Based on these results, the relative histone mRNA abundances determined by qRT-PCR may actually underestimate the fold-change increase in iPS cells as each HFF reaction contained the equivalent of twice as many cells but still had dramatically less histone mRNA.

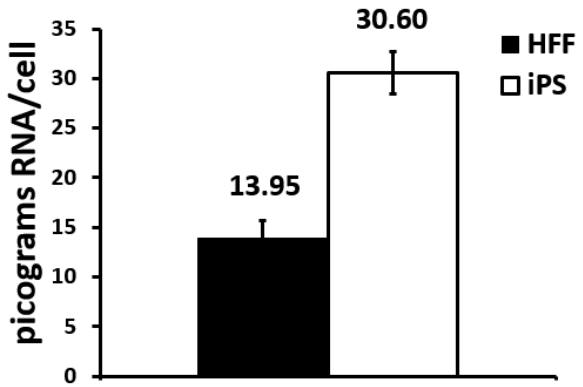


Figure 4.12 iPS cells have twice as much total RNA per cell compared to HFFs. HFF and iPS cells were assessed for the amount of RNA per cell by measuring RNA isolated by TRIzol from a known number of cells (~$3x10^6$ HFFs and ~$4.5x10^5$ iPS cells). Standard deviations represents three replicates.

To determine the amount of protein per cell, two approaches were used. In the first, we measured the amount of protein by preparing whole cell lysate from a known number of cells and measuring concentration using the Bio-Rad Protein Assay. Interestingly, despite obvious differences in cell size and amounts of RNA, Figure 4.13 showed that HFF and iPS cells both have ~600 pg of protein per cell (599.2 ± 15.0 pg in HFF and 601.7 ± 45.2 pg in iPS cells). The typical mammalian cell is estimated to contain ~500 pg of protein (Alberts et al., 1994), not dramatically different from levels measured here.

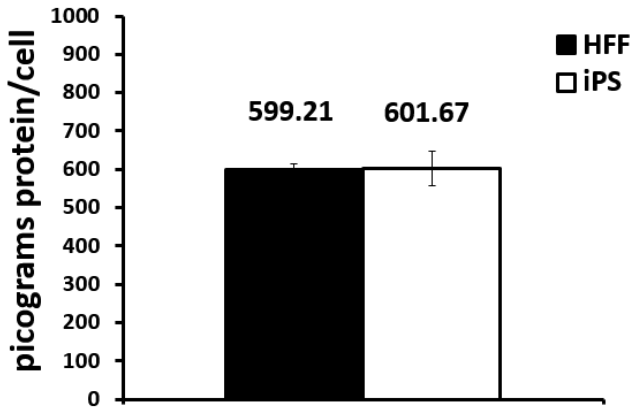


Figure 4.13 HFF and iPS cells have equal amounts of protein per cell compared to HFFs. Cells were assessed for amount of protein per cell using the Bio-Rad Protein Assay to measure lysates made from a known number of cells. Standard deviations represents three replicates.

In an alternative method, equal numbers of cells were resuspended in 1x SDS protein dye and resolved on a 15% SDS-PAGE gel. Following Coomassie staining, the gel was compared to a Coomassie-stained gel loaded with 25 µg of whole cell lysate (Figure 4.7) showing no discernable differences, giving further evidence that HFFs and iPS cells have equal amounts of protein per cell.

Given these analyses indicating that HFF and iPS cells contain the same amount of protein per cell, all comparisons of histone protein abundance can be assumed to represent the same numbers of cells. In conjunction with the Coomassie data, we determined that there is ~10-fold more core histone protein per iPS cell than HFF. Considering the importance of histones in chromatin dynamics and the well-accepted description of pluripotent heterochromatin structure, this finding is very intriguing.

4.8 Discussion

Based on both microarray and qRT-PCR data, we showed that iPS cells express more histone mRNA and protein than HFF cells, supported in part by significant stabilization of replication-dependent transcripts. Assuming that the stabilized mRNAs are translatable and intact, it seems likely that an early step in the decay pathway is inhibited in iPS cells. As TUTase ZCCHC11 (Schmidt et al., 2011) and UPF1

(Kaygun and Marzluff, 2005b) initiate degradation of histone mRNAs, it would be interesting to see how they are regulated in iPS cells compared to HFFs. Analogous to the "poised" nature of pluripotent chromatin, these decay factors must be inactive towards histone transcripts during S phase but rapidly induce destabilization once DNA synthesis is complete, as seen in Figure 4.10. It is possible that these decay proteins exhibit differential post-translational regulation, contributing to differential histone mRNA stability. In our microarray data, the transcripts encoding these proteins showed 2.40-fold and 1.46-fold increases in iPS cells, respectively, but this does not give any information about the protein abundance or activity.

The overall stabilization of histone mRNAs during S phase and dramatic decay immediately upon inhibition of DNA synthesis in iPS cells suggests that these cells exert sharper control over histone regulation than their differentiated counterparts. Consistent with this, in ES cells, histone locus body formation is induced 2 hours prior to S phase whereas induction begins 6 hours before S phase in differentiated cells (Ghule et al., 2008). Further, knockdown of reprogramming factors OCT4 and SOX2 in ES cells leads to down-regulated expression of 20 to 30 histone mRNAs (Greber et al., 2007). Considering the importance of chromatin plasticity to maintaining pluripotent gene expression (see Section 4.1.3), it seems reasonable that iPS cells would dedicate more stringent control to histone mRNAs than HFFs wherein chromatin remodeling is less utilized (Efroni et al., 2008).

We also found that histone mRNAs, even within families, exhibited a wide range in half-lives despite being predominantly regulated through the same conserved 3'UTR element. For instance, H3 transcripts in iPS cells and H2B transcripts in HFFs had half-lives ranging up to 10-fold more stabilized than the least stable transcript. It is possible that other structural or sequence elements within histone transcripts determine stability. Further analysis of the individual mRNAs identified by microarray may reveal more insightful clues to differential histone stability within the same cell line.

From our data, we also find that the fold-change in histone protein expression (~10-fold higher in iPS cells) is much lower than what would be predicted based on mRNA levels. This suggests that iPS cells have more mRNA undergoing less translation than HFFs. Presumably, these cells have a way of keeping histone transcripts translationally inactive. Fluorescence *in situ* hybridization studies could indicate whether these mRNAs are sequestered to a particular part of cell, whether in the nucleus within foci similar to histone locus bodies or processing bodies found in the cytoplasm. Either way, it seems likely that the transcripts are also kept in a reversibly silenced, or readily activated, state. One of the defining features of pluripotency is the capacity for unlimited differentiation. To carry out such lineage progression, the genome must be re-packaged to facilitate the new gene expression requirements of the cell (Meshorer and Misteli, 2006) and this likely requires an influx of histone proteins. Perhaps additional histone mRNAs are kept "on deck" in iPS cells in the event that differentiation signals are encountered.

Although we find that differences in histone protein abundance are not as dramatic as the mRNA data would suggest, the ~10-fold increase seen in iPS cells (Figure 4.7) raises several questions regarding the regulation of these proteins. It is important to note that Coomassie and western blot analyses merely reflect the overall abundance of histones and do not give information as to whether the histones detected were associated with DNA or in unbound populations. The nuclear environment of stem cells is characterized as having a slightly larger proportion of unbound histones than differentiated cells (Meshorer et al., 2006). We found that iPS cells express ~10-fold more histone protein per cell than HFFs (Figure 4.7 and Figure 4.13) even though the number of cells in S phase was only increased 2-fold (Figure 4.11). Assuming the chromatin structure of the iPS cells used in this study is consistent with previously published descriptions, it is likely that iPS cells have a much larger proportion of unbound histones than previously appreciated. Unlike histone mRNAs, histone proteins persist throughout the cell cycle. It would be interesting to determine whether the unbound histones are restricted to cells

synthesizing DNA or if they are somehow stored in a discrete location outside of S phase. In yeast, excess histone protein results in mitotic chromosome loss (Meeks-Wagner and Hartwell, 1986) and cytotoxicity (Gunjan and Verreault, 2003; Gunjan et al., 1999) so the high abundance of histones in iPS cells is intriguing.

In iPS cells where there is an increased abundance of unbound histones, it may be necessary for histone chaperones to maintain interactions longer rather than immediately delivering histones to nucleosomes. NASP, a protein containing three histone-binding domains, is able to interact with histones (Batova and O'Rand, 1996) in a cell cycle-regulated manner (Richardson et al., 2000). More specifically, NASP is a chaperone for linker histone H1 to assist chromatin formation. NASP shares sequences homology with the N1/N2 family of histone chaperone proteins that preferentially associate with H3 and H4 (Wang et al., 2008). Nucleoplasmin family members NPM2/NPM3 preferentially bind histones H2A and H2B (Laskey et al., 1993; Ramos et al., 2010). These proteins function to escort histone proteins to nuclear locations where they perform their roles in supporting chromatin architecture. Co-immunoprecipitations of these factors with histones could indicate whether their associations are more persistent in stem cells versus HFFs. Even if excess histones are not bound to chaperones more frequently in pluripotent cells, co-localization of histones with markers for various cytoplasmic and nuclear foci could give clues as to a mechanism of bulk histone storage.

The amount of histone proteins in the nucleus is closely monitored to assure proper levels are maintained. Imbalanced expression of histones can lead to compromised genomic integrity and chromosomal instability (Williamson and Pinto, 2012). It has been previously shown in dividing starfish embryos that the histone/DNA ratio changes through development (Shabalkin, 1996). Although these embryos experience increased histone expression during the first four cleavages, this examples provides evidence that histone mRNAs are developmentally regulated in addition to the regulation conferred by the cell cycle.

Overall, our observations indicating that replication-dependent histone mRNAs are significantly stabilized in iPS cells further emphasizes the importance of histones to pluripotent gene expression. Although it seems counter intuitive for a cell type with less compact chromatin to express more histones, the plasticity demonstrated by stem cells provides a possible explanation for the dramatic increase compared to differentiated cells. Further investigation to identify mechanisms responsible for differential stability of histone mRNAs within and between cell lines might provide additional targets for optimizing nuclear reprogramming.

Chapter 5: C2H2 zinc finger protein mRNAs are stabilized in iPS cells

5.1 Introduction

The second class of mRNAs that showed altered stability in our global analysis were those encoding C2H2 Zinc Finger (ZNF) proteins. As C2H2 ZNFs represent the largest class of transcription factors in eukaryotes, their expression is highly likely to influence the establishment and maintenance of the pluripotent state making their regulation a high priority for further study. Genes containing C2H2 ZNF domains are highly transcribed during early embryogenesis relative to the expression of all transcription factors but decrease through later stages in organisms ranging from *Drosophila* to *Xenopus* (Adryan and Teichmann, 2010; Schep and Adryan, 2013). Correlating well with this observation, our global analysis data indicated that C2H2 ZNF mRNAs were more stable in iPS cells as compared to HFFs (discussed in Chapter 3). While mechanisms regulating C2H2 ZNF mRNA stability are not well characterized, recent studies have identified two features of ZNF mRNAs that may contribute to this regulation. These features as well as further analysis of the stabilization of C2H2 ZNF mRNAs in iPS cells are discussed here.

5.1.1 C2H2 ZNF genes are highly conserved

C2H2 ZNF genes are members of the ZNF gene superfamily that make up one of the largest families in the mammalian genome, second only to olfactory receptor genes. There are approximately ~800 ZNF genes organized in clusters across the human genome (Grimwood et al., 2004; Knight and Shimeld, 2001) and because of this, many studies have focused exclusively on their molecular evolution (Emerson and Thomas, 2009; Lorenz et al., 2010; Stubbs et al., 2011). Not much is known regarding the mechanisms driving ZNF gene duplication but their coevolution with retroelements suggests that these proteins may be involved in an “arms race” to prevent genome damage from retrotransposition

(Thomas and Schneider, 2011). Studies show that this family is still evolving and expanding in mammals (Tadepally et al., 2008).

The ZNF family of genes is divided into subfamilies based on their type of zinc finger fold and protein domains that enable ZNF proteins to interact with various molecules such as DNA, RNA, and other proteins. There are six fold groups in the ZNF family specified by domains, namely C2H2 (Miller et al., 1985), TAZ (De Guzman et al., 2000), Zn_2/Cys_6 (Carr et al., 1990), knuckle (Danielsen et al., 1989), treble clef (Grishin, 2001), and zinc ribbon (Qian et al., 1993). C2H2 ZNF genes make up the majority of the ZNF family and are the best characterized.

C2H2, or Krüppel-like domains, are homologous to tandem repeat ZNF domains identified in the *Drosophila melanogaster* transcription factor Krüppel (Schuh et al., 1986). These domains are considered to be the "classical" zinc fingers where structural motifs of repeating pairs of cysteine and histidine residues create folds stabilized by one or more zinc ions (Miller et al., 1985). C2H2 domains occur in two or more tandem repeats separated by a highly conserved linker region that is important for DNA-binding (Foster et al., 1997; Pabo et al., 2001) and mitotic regulation (Dovat et al., 2002; Rizkallah et al., 2011; Figure 5.1). In addition to providing binding strength, these zinc finger folds also confer sequence specificity (Nardelli et al., 1991; Thukral et al., 1992). Although the C2H2 domain has the ability to also interact with RNA and proteins, its function in sequence-specific DNA-binding is the most apparent; most characterized C2H2 ZNF genes encode transcription factors, making up 40% of all transcription factor genes (Messina et al., 2004). Effector functions are carried out through additional domains including KRAB (Krüppel-Associated Box; Bellefroid et al., 1991), BTB/POZ (ZBTB; Zollman et al., 1994), SCAN (ZSCAN; Williams et al., 1995), and SET (Jenuwein et al., 1998) domains.

5.1.2 Many C2H2 ZNF genes contain the transcription repressor KRAB domain

ZNF genes containing the KRAB domain make up the largest subfamily of C2H2 ZNFs with 423 genes encoding 671 distinct KRAB ZNF proteins in humans (Ding et al., 2009; Huntley et al., 2006) and this subset of C2H2 ZNF genes was strongly over-represented in our set of mRNAs that were stabilized in iPS cells (discussed in Chapter 3). Figure 5.1 depicts a C2H2 ZNF protein with a KRAB domain at its typical location near the N-terminus. This prevalent C2H2 ZNF protein domain spanning 75 amino acids is highly charged and divided into two regions, boxes A and B (Bellefroid et al., 1991). Upon binding of the ZNF domain to DNA, the A box mediates transcriptional repression (Witzgall et al., 1994) through a direct interaction with KAP1 (KRAB-Associated Protein 1), also known as TRIM28 or TIF1β (Friedman et al., 1996). KAP1 in turn recruits chromatin modifiers such as heterochromatin protein CBX5 (Chromobox Homolog 5; Lechner et al., 2000) and histone methyltransferase SETDB1 (SET Domain, Bifurcated 1; Schultz et al., 2002) to silence transcription promoter regions. KAP1 and KRAB ZNFs have the ability to mediate transcriptional repression in regions up to 15 kb away through heterochromatin spreading characterized by loss of H3 acetylation, increased H3 methylation, and decreased RNA Pol II recruitment (Groner et al., 2010).

Although the predominant function of the KRAB domain is transcriptional repression, ZNF proteins bearing this region in addition to other domains are not limited to repressor activity. For example, identification of targets regulated by ZNF263, which contains KRAB and SCAN domains, revealed that this transcription factor has the ability to activate or repress transcription (Frietze et al., 2010). Further, not all KRAB ZNF proteins interact with KAP1 to elicit repression. Several SCAN-KRAB ZNF proteins exhibit KAP1-independent repressor activity, alluding to an additional mechanism by which transcriptional silencing is achieved (Itokawa et al., 2009).

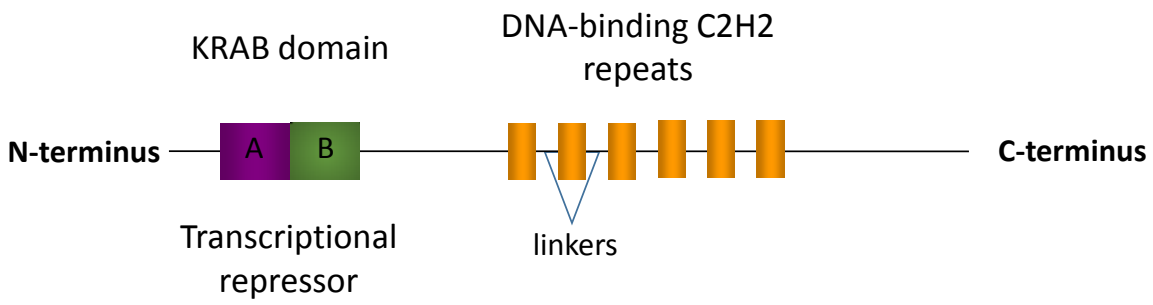


Figure 5.1 KRAB ZNF proteins have tandem C2H2 DNA-binding domains and an upstream transcription repressor KRAB domain. Repeating C2H2 domains are separated by linker domains that facilitate DNA-binding. The KRAB domain is made of A and B boxes that confer transcriptional repression.

5.1.3 Expression of C2H2 ZNF mRNAs is developmentally-regulated

As noted above, in several species C2H2 ZNF expression is widespread during early embryogenesis but declines as development progresses (Adryan and Teichmann, 2010; Schep and Adryan, 2013). This is consistent with the repressive role of C2H2 ZNFs as pluripotent cell types are characterized by a euchromatic genome structure requiring repression of lineage-specific genes (Efroni et al., 2008; Gaspar-Maia et al., 2011). Co-repressor KAP1 is also required for silencing differentiation-specific genes and establishing DNA methylation patterns in stem cells (Hu et al., 2009; Quenneville et al., 2012). Relatively few C2H2 ZNF transcription factors have been individually characterized but several have been identified with roles in embryogenesis such as ZFP281, a repressor for pluripotency (Fidalgo et al., 2011) and ZFP57, a regulator of imprinted genes (Quenneville et al., 2011). Given the importance of transcriptional silencing during pluripotency (Gaspar-Maia et al., 2011), there are likely several more C2H2 ZNF genes required for an undifferentiated state that remain to be characterized.

Several mechanisms involved in coordinated regulation of C2H2 ZNF genes during of embryonic development have been uncovered. First, four miRNA families have been identified that target sequences within the region encoding the repeating C2H2 domains to initiate deadenylation and translational repression of the mRNA (Huang et al., 2010; Schnall-Levin et al., 2011). Although miRNAs

are known for binding sequences in the 3'UTR of mRNAs, their interactions are not restricted to non-coding portions (Duursma et al., 2008; Forman et al., 2008). MicroRNA-181a (miR-181a) was the first miRNA demonstrated to target the C2H2 ZNF ORF region (Huang et al., 2010) but additional members of the miR-181 family (miR-181b-d) and miRNA families miR-23, miR-188, and miR-199 have since been shown to share this ability (Schnall-Levin et al., 2011). These miRNAs bind target sequences within C2H2 and linker domains. As these domains are often present in tandem repeats, multiple binding sites provide potential for increased miRNA activity. It is possible that this mechanism contributes to the differential stability of C2H2 ZNF mRNAs we observed in iPS cells compared to HFFs.

Of the four miRNA families that target ZNF transcripts, several members have previously been linked with various important development and differentiation processes. MicroRNA-23 has been implicated in regulating neuronal differentiation of human teratocarcinoma cells (Kawasaki and Taira, 2003), neovascularization (Zhou et al., 2011), and neurogenesis of the embryonic spinal cord (Farrell et al., 2011). Studies involving miR-199 show a role in organ development (Mungunsukh and Day, 2013). Members of the miR-181 family are important for myoblast differentiation (Naguibneva et al., 2006) and in human leukemia cells, miR-181 disrupts the LIN28/let-7 feedback circuit to facilitate cell differentiation (Li et al., 2012). As each of these miRNAs also targets other transcripts, it remains to be seen whether any of these events occur through regulation of ZNFs.

The second interesting feature of C2H2 ZNF genes that may allow for their coordinated regulation is that the mRNAs they encode are bimorphic. Bimorphic transcripts are defined as those having a significant population of unadenylated mRNAs in addition to the standard polyadenylated population. ZNF transcripts were over-represented among unadenylated mRNAs detected in both embryonic stem cells and HeLa cells, although more so in the stem cells (Yang et al., 2011). Poly(A) tail shortening could impact export from the nucleus or translation of these mRNAs as well as their

turnover. While further investigation is needed, these data suggest that the poly(A) tail status of ZNF mRNAs may be differentially regulated as a means of gene expression control.

As introduced in Chapter 3, global analysis of mRNA decay rates in iPS cells revealed significant stabilization of C2H2 ZNF mRNAs compared to HFFs. Considering the role of C2H2 ZNF genes as transcription regulators and their importance for maintaining pluripotent gene expression profiles, we sought to further investigate this class of transcripts to provide clues as to a mechanism of stabilization. We hypothesized that this class of transcripts may be coordinately regulated by families of ZNF ORF-targeting miRNAs and that these miRNAs may be down-regulated in iPS cells to allow for ZNF mRNA stabilization. The results of this investigation are presented here[4].

5.2 C2H2 ZNF protein mRNAs are significantly stabilized in iPS cells

Analysis of functional domains enriched in the mRNAs stabilized in iPS cells compared to HFFs found that KRAB (P-value = 1.2×10^{-23}) and C2H2 (P-value = 7.4×10^{-18}) zinc finger domains were significantly over-represented (discussed in Chapter 3). Within our dataset of significantly stabilized mRNAs, C2H2 ZNF mRNAs made up ~20% (118 of 548) of the transcripts. This was approximately twice as many as would be predicted since 532 C2H2 ZNF mRNAs were represented in the dataset of 5,481 mRNAs (10% of the dataset). The stability of all C2H2 ZNF mRNAs and all KRAB ZNF mRNAs were then compared to half-life changes for all ZNF mRNAs (including those containing ZBTB, SCAN, and other non-C2H2 domains) and all 5,481 mRNAs in the dataset (Figure 5.2). Overall, ZNF transcripts exhibited stabilization (P-value = 4.2×10^{-6}) but not to the extent seen for ZNF family transcripts bearing the C2H2 (P-value = 5.5×10^{-13}) and especially KRAB (P-value $< 2.2 \times 10^{-16}$) domains. Based on these results, we found

[4] Many of the results presented in Chapter 5 appeared in: Neff *et al*. Global analysis reveals multiple pathways for unique regulation of mRNA decay in induced pluripotent stem cells. Genome Research. Vol. 22, No. 8, pg. 1457-1467, August 2012.

that there was significant stabilization of C2H2 and KRAB ZNF mRNAs over and above the general trend of all ZNF transcripts.

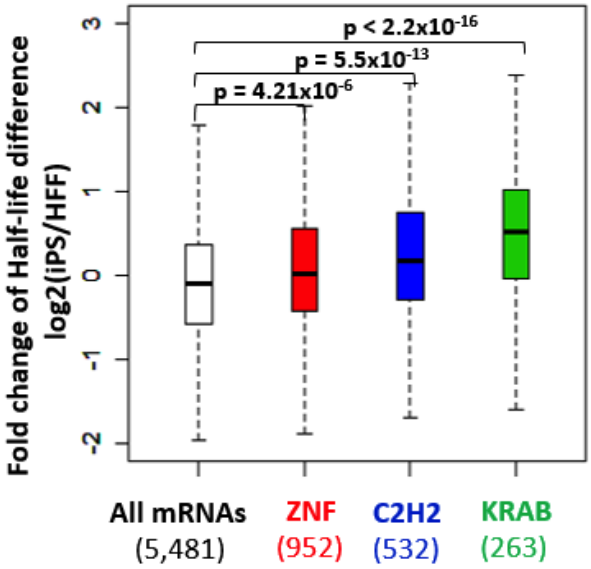


Figure 5.2 C2H2 ZNF and KRAB ZNF mRNAs are significantly stabilized in iPS cells compared to HFFs. Fold change in half-life between HFF and iPS cells for different subsets of transcripts assessed from microarray data.

We then went on to validate the stabilization of three C2H2 ZNF mRNAs, *ZNF43* (19 C2H2 zinc fingers plus KRAB domain), *ZNF134* (11 C2H2 zinc fingers), and *ZNF627* (10 C2H2 zinc fingers plus KRAB domain) by qRT-PCR where abundances at each time point were normalized to expression of *GAPDH* mRNA (Figure 5.3). To date, the function of ZNF134 has not been characterized. Nucleotide polymorphisms of ZNF627 are associated with myocardial infarction (Horne et al., 2007; Koch et al., 2011). Interestingly, ZNF43 is a transcription repressor important for maintenance of an undifferentiated state in Ewing sarcoma cells (González-Lamuño et al., 2002) and may also function in maintenance of pluripotency. Our microarray data indicated that HFF and iPS cells exhibit half-lives of 344 minutes and 645 minutes for *ZNF43* mRNA, 185 minutes and 895 minutes for *ZNF134* mRNA, and

119 minutes and 391 minutes for *ZNF627* mRNA. Although our qRT-PCR measured half-lives are shorter than those determined by microarray, we found that all three mRNAs exhibit increased half-lives in iPS cells compared to HFFs as expected. Within this small subset there was no apparent correlation between the number of C2H2 or KRAB domains contained within the mRNA and the level of stabilization seen in iPS cells. Specifically, *ZNF134* and *ZNF627* had very similar half-lives in both cell lines despite the presence of a KRAB domain in *ZNF627*, indicating that the KRAB domain was not required for stabilization. Further, *ZNF43* mRNA exhibited stabilization to a lesser degree than *ZNF134* and *ZNF627* mRNAs despite having a KRAB domain and more C2H2 zinc finger domains than the other two transcripts. Stabilization of *ZNF43* mRNA in iPS cells was also verified in an independent set of matched cell lines (A. Jalkanen personal communication). Although it is not clear what mechanisms are responsible for stabilization of these mRNAs in iPS cells, the half-lives measured here suggest that factors other than domain representation may influence their stability.

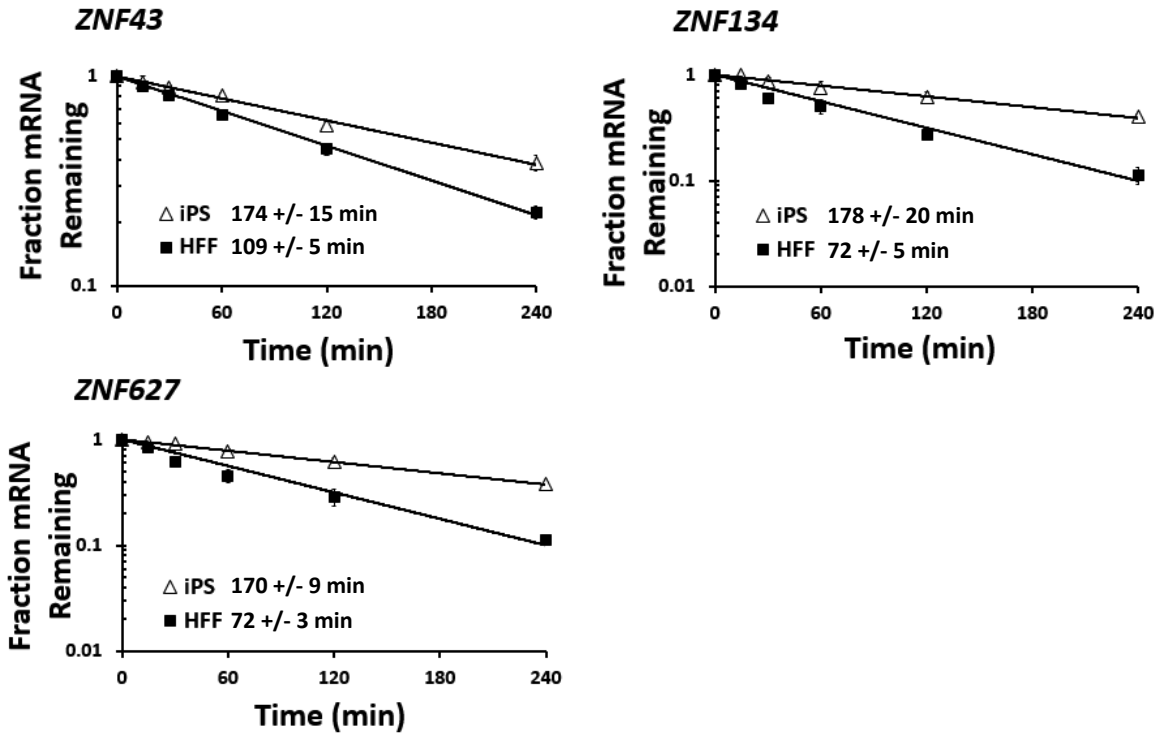


Figure 5.3 *ZNF43*, *ZNF134*, and *ZNF627* mRNAs are stabilized in iPS cells compared to HFFs. Half-lives were assessed in HFF and iPS cells for each mRNA following actinomycin D treatment. Abundance levels were measured for each indicated mRNA by qRT-PCR at each time point and normalized to levels of *GAPDH* mRNA. The standard deviations were derived from three independent replicates. Half-lives for each mRNA are denoted to the right of the cell-line keys.

5.3 ORF-targeting miRNAs have decreased abundance in iPS cells

As discussed in Chapter 3, hexamer analysis to identify putative *cis*-acting elements revealed an enrichment of sequences in the 5'UTR, ORF, and 3'UTR of stable and unstable mRNAs in iPS cells. Upon further investigation we found that several hexamers identified as over-represented in the ORF of stable iPS cell mRNAs were complementary to seed sequences of miRNA families previously implicated in C2H2-domain targeting (Huang et al., 2010; Schnall-Levin et al., 2011). Specifically, Figure 5.4 shows these hexamer sequences and their heat map representation of significance score. The significance of a hexamer was calculated by the Fisher's exact test comparing its frequencies in the most stable and least stable mRNA sets. A significance score was assigned to each hexamer equaling $-\log_{10}$(P-value)*s, where

s equals -1 if the hexamer was more common in the least stable transcripts, and s equals 1 if otherwise. Six of these over-represented ORF hexamers demonstrated sequence overlap with binding sites for miRNA families miR-23, miR-181, and miR-188. A putative binding site for miR-199, another family that targets C2H2 ZNF mRNAs, was not identified among the top enriched hexamers. As we saw from the heat map tiles correlating hexamer sequences with stability, these *cis*-elements were strongly associated with mRNA stabilization in iPS cells (yellow) and some sequences were also enriched in the least stable HFF mRNAs (e.g. GAAUGU, UGUGGG, and AUGUGA; bright blue).

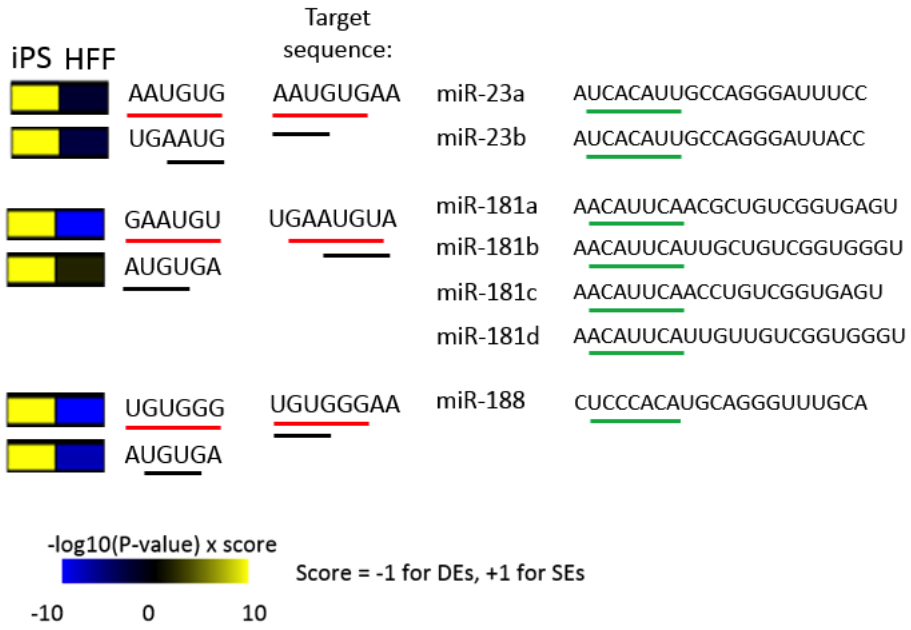


Figure 5.4 Hexamer sequences similar to seed sequences of ZNF ORF-targeting miRNAs are over-represented in stable iPS cell mRNAs compared to HFFs. Alignment of hexamer *cis*-elements identified in stabilized iPS cell mRNAs showing overlap with ZNF ORF-targeting miRNA target sequences. Hexamers over-represented in stable mRNAs (SE) are depicted in yellow while hexamers over-represented in unstable mRNAs (DE) are in blue. Sequences of hexamers matching miRNA target sequences are underlined in red and black. miRNA seed sequences are underlined in green.

For each of the three target sequences where overlap was detected, the corresponding hexamer sequences were either a complete (red underscore) or incomplete (black underscore) match. Based on

the heat map tiles in Figure 5.4, hexamers with the best match to the target sequence were slightly more likely to be found in mRNAs exhibiting a larger degree of differential stability than their incomplete-binding counterparts. However, perfect binding of the entire length of miRNAs to their target transcripts leading to endonucleolytic cleavage and degradation is rarely used in humans (Bracken et al., 2011; Karginov et al., 2010; Shin et al., 2010). Instead, imperfect binding of miRNAs to targets leads to deadenylation and translational silencing (Pillai et al., 2005). It is also important to note that more than one of these putative ORF miRNA binding sites is likely present within a single mRNA so a true correlation between degree of stabilization and miRNA binding sites would require much more detailed analysis to delineate contributions of each *cis*-element.

To further investigate the association of miRNA target sequences with C2H2 ZNF stabilization, we looked at the representation of miRNA binding sites within stabilized C2H2 ZNF mRNAs compared to C2H2 ZNF mRNAs that were not stabilized in iPS cells. By counting the number of target sequences for each miRNA family, we identified the presence of miR-23, miR-181, miR-188, and miR-199 binding sites within all C2H2 ZNF mRNAs that half-lives were determined for in both cell lines. Bioinformatic analysis highlighted the enrichment of these miRNA binding sites in the ORF and 3'UTR and also showed that stabilized C2H2 ZNF mRNAs contain more miRNA binding sites than those not stabilized (Figure 5.5). This positively correlated miRNA binding sites for miR-23, miR-181, miR-188, and miR-199 with transcript stabilization in iPS cells. The trend of these data suggested that the presence of miRNA binding sites in C2H2 ZNF mRNAs may play a role in their stabilization in iPS cells although other factors are also likely to contribute.

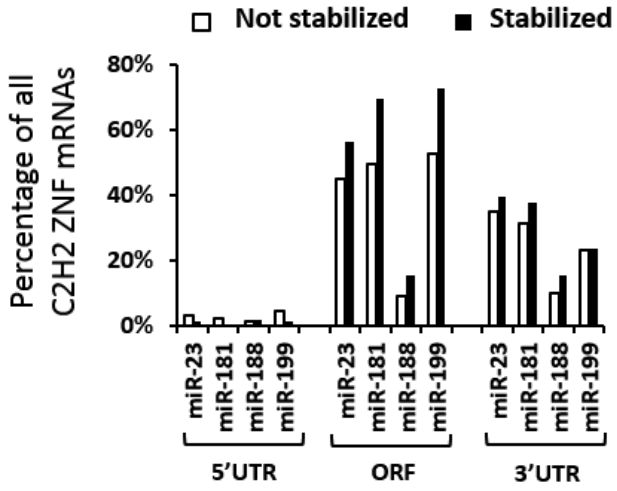


Figure 5.5 C2H2 ZNF mRNAs containing target sites for ORF-targeting miRNAs tend to be stabilized in iPS cells rather than not stabilized compared to HFFs. The percentage of C2H2 ZNF mRNAs that contained miRNA binding sites in the 5UTR, ORF, and 3'UTR and were stabilized as compared to C2H2 ZNF mRNAs that were not stabilized.

Given the enrichment of hexamers complementary to ORF-targeting miRNA seed sequences in the ORF of stabilized C2H2 ZNF mRNAs, we hypothesized that lower abundance of these miRNA families in iPS cells might contribute to their regulation. We therefore next sought to measure the abundance of miRNAs miR-23a, miR-23b, miR-181a, miR-181b, miR-181d, miR-188-3p, and miR-199a relative to *5S* rRNA by qRT-PCR (Shi and Chiang, 2005). Briefly, total RNA was collected from untreated cells and 3' elongated with >150 adenosine nucleotides using Poly(A) Polymerase (PAP). An oligo(dT) adapter containing a short stretch of 12 thymidine nucleotides and unique sequence was then annealed to all transcripts to facilitate reverse transcription and cDNA synthesis. Samples were then amplified in reactions containing a specific forward primer to detect individual miRNAs and a universal reverse primer that bound the non-oligo(dT) portion of the adapter. Oligos for miRNAs were selected based on validated results from a previous study comparing various methods to measure miRNA abundance (Git et al., 2010). A specific primer for miR-181c was not listed within the validated primer sets of the Git et al. study therefore this miRNA was not included in our qRT-PCR analysis.

As shown in Figure 5.6, each of the miRNAs measured were expressed at lower levels in iPS cells compared to HFFs. The most dramatic differences were seen for miR-23a/miR-23b and miR-199a whose expression is ~17-fold and ~50-fold lower, respectively in iPS cells compared to HFFs. These results are not surprising as pluripotent cell types express only a subset of differentially regulated miRNAs and lower levels overall compared to differentiated cells (Lakshmipathy et al., 2007). Nevertheless, the reduced abundance of ORF-targeting miRNAs in iPS cells correlated well with the increased mRNA stability of C2H2 ZNF transcripts.

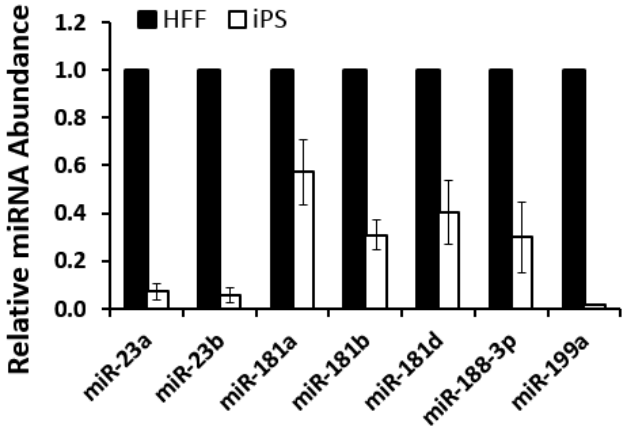


Figure 5.6 C2H2 ZNF ORF-targeting miRNAs have decreased abundance in iPS cells relative to HFFs. Levels of ZNF-targeting miRNAs in iPS cells were measured by qRT-PCR, normalized to the abundance of *5s* rRNA, and relative to levels in HFFs. Standard deviations represent three replicates.

In order to determine which of the four miRNA families might make the largest overall contribution to regulation of ZNF mRNA stability, we compared the abundances of these families to each other. We found that of the C2H2 ZNF ORF-targeting miRNAs measured here, miR-23a and miR-23b were the most abundant in both cell lines (Figure 5.7). In HFFs, miR-199a was expressed at ~10% the level of miR-23a while the remaining miRNAs were expressed at ~1% relatively (Figure 5.7A). In iPS cells, members of the miR-181 family were expressed at 5-10% the level of miR-23a while miR-199a and miR-

188-3p were expressed at a very low abundance measuring ~3% of miR-23a levels (Figure 5.7B). As expression of individual miRNAs within tissues is known to vary (Liang et al., 2007), the differences in abundance seen here within cell lines are not unexpected. Although we could not draw definitive conclusions regarding the functional relevance of these differences in miRNA abundance, we noted that miR-23a is expressed at several-fold higher levels when compared to expression of other C2H2 ZNF ORF-targeting miRNAs in both cell lines and may therefore have a larger impact than the other miRNAs.

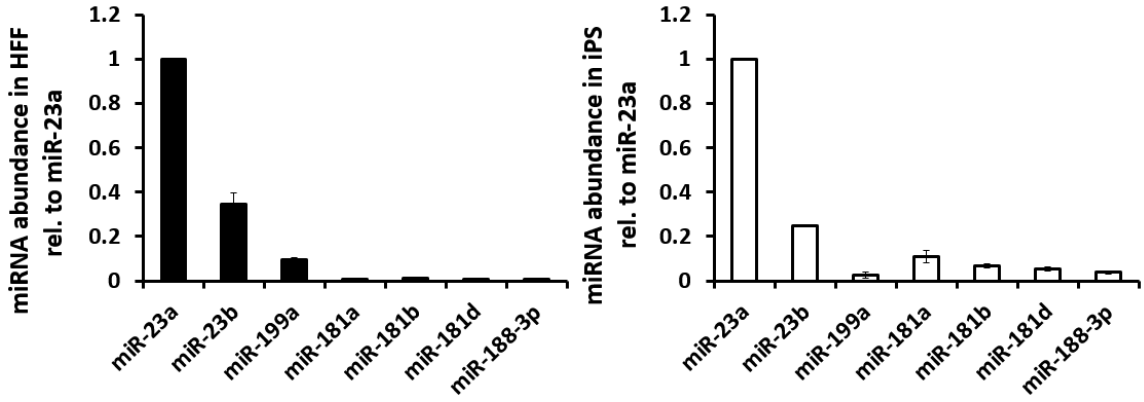


Figure 5.7 ZNF ORF-targeting miRNAs are differentially expressed within each cell line. Abundance of ZNF-targeting miRNAs in (A) HFF and (B) iPS cells as measured by qRT-PCR, normalized to levels of *5s* rRNA, and relative to the abundance of miR-23a. Standard deviations represent three replicates.

Overall, our analyses have indicated that C2H2 ZNF mRNAs, many of which include the KRAB transcriptional repression domain, are significantly stabilized in iPS cells compared to HFFs. Further, we showed that hexamer sequences enriched in the coding region of stable iPS mRNAs potentially serve as binding sites for miRNA families miR-23, miR-181, miR-188, and miR-199. Finally, we also demonstrated that ORF-targeting miRNAs are expressed at much lower levels in iPS cells, possibly contributing to the marked stabilization of C2H2 ZNF mRNAs.

5.4 Discussion

Given that C2H2 ZNF transcription factors have elevated expression during early embryogenesis but decreasing expression during development (Adryan and Teichmann, 2010; Schep and Adryan, 2013), we were intrigued to find that C2H2 ZNF mRNAs were more stable in iPS cells as compared to HFFs. This family of transcripts typically encodes transcription repressors, factors important for suppressing differentiation genes in pluripotent cell types (Efroni et al., 2008). The large size of the C2H2 ZNF family means that they potentially can have a significant impact on gene expression profiles either by acting redundantly on a large scale to regulate a single pathway, or through multiple coordinated regulatory events with each ZNF targeting a different set of genes. In addition, the sheer number of ZNF transcripts within the cell means that they have the potential to "soak up" regulatory factors and indirectly effect the expression of other genes. Examination of the literature uncovered two unusual features of ZNF mRNAs which might contribute to this regulation leading to stabilization in iPS cells.

The open chromatin structure of stem cells supports the unique features characteristic of pluripotent cell types (discussed in Chapter 4). The loosely bound nature of the stem cell genome does not provide physical repression gained through heterochromatin compaction, therefore transcriptional repressors inhibit expression of lineage-specific genes (Efroni et al., 2008; Gaspar-Maia et al., 2011; Lessard and Crabtree, 2010). Although we cannot say at this time whether stabilization of ZNF mRNAs results in increased production of ZNF proteins, it is likely that these differences in decay contribute to the distinct expression pattern of C2H2 ZNF transcription factors during embryogenesis. Additionally, a global increase in the availability of these transcriptional repressors might be important for maintaining genome integrity by preventing transcription and transposition of retroelements. The KAP1 protein is essential for controlling endogenous retroviruses in mouse ES cells (Rowe et al., 2010) and as KRAB-ZNFs are the major proteins that recruit KAP1, it is easy to imagine how elevated ZNF abundance might be required for effective retrovirus repression.

Further investigation of C2H2 ZNF mRNA stabilization was based on observations that ZNF mRNAs can be targeted by miRNAs that recognize the tandem-repeated C2H2 zinc finger regions (Huang et al., 2010; Schnall-Levin et al., 2011) and that ZNF mRNAs often have short poly(A) tails (Yang et al., 2011). This led us to the hypothesis that ZNF mRNAs may be targeted for deadenylation and decay by miRNAs but to a lesser extent in iPS cells than in HFFs. In support, we found that all four miRNA families that target ZNF mRNAs are less abundant in iPS cells which might explain the reduced decay of these mRNAs in this cell type. Although we did not assess the poly(A) tail length of ZNF mRNAs in HFF and iPS cells, we might expect the tails to be even shorter in HFFs than in iPS cells if the short poly(A) tail is generated though miRNA-induced deadenylation. While the Yang et al., 2011 study found ZNF mRNAs to have short poly(A) tails in ES cells compared to HeLa cells, genetically matched cell types were not investigated. Therefore, determination of poly(A) tail lengths in both cell types would be necessary to validate this possibility. Alternatively, it is possible that these are independent means of regulation and ZNF transcripts have constitutively short poly(A) tails. In this case, the mechanism by which the miRNAs target the ZNF transcripts for decay would be unclear as the canonical mechanism of action is for miRNAs to induce deadenylation which in turn leads to decay. If ZNF transcripts are already predominantly deadenylated, the miRNAs would have to recruit other components of the decay machinery in order to induce decay.

Overall, the stabilization of C2H2 ZNF mRNAs in iPS cells contributes to our understanding of how these genes achieve elevated expression in early development. Although many follow up studies remain, we have shown that miRNA families that target these transcripts have decreased expression in iPS cells and may explain their increased stability. As many of these transcription factors bear a repressor KRAB domain, it seems likely that stabilization of C2H2 ZNF mRNAs supports the overall enhanced expression of these genes in embryogenesis.

Chapter 6: mRNAs with C-rich elements in their 3'UTR are destabilized and poly(C)-binding proteins are differentially expressed in iPS cells

6.1 Introduction

As stated in Chapter 3, a third class of mRNAs identified as differentially regulated in iPS cells compared to HFFs contained C-rich elements in their 3'UTRs and were less stable in iPS cells. C-rich sequence motifs are important determinants of mRNA stability (Kong et al., 2006; Lindquist et al., 2000; Weiss and Liebhaber, 1995) and also regulate alternative polyadenylation (Ji et al., 2013a), cytoplasmic polyadenylation (Ji et al., 2013a; Paillard et al., 2000), and translational silencing (Ostareck-Lederer and Ostareck, 2004; Reimann et al., 2002). The enrichment of C-rich sequences in unstable iPS cell mRNAs was striking for several reasons including (i) C-rich elements were previously associated with very abundant and stable transcripts such as those encoding α-globin (Wang et al., 1995; Weiss and Liebhaber, 1995) and collagen I (Lindquist et al., 2000), (ii) C-rich elements were not enriched in unstable transcripts in any other global analyses to date (Dolken et al., 2008; Friedel et al., 2009; Lam et al., 2001; Lee et al., 2010; Raghavan et al., 2002; Schwanhäusser et al., 2011, 2013; Sharova et al., 2009; Thomsen et al., 2010; Vlasova et al., 2008; Yang et al., 2003), suggesting that this regulation mechanism could be specific to stem cells, and (iii) several proteins known to interact with C-rich elements are also linked with modulation of mRNA decay (Gherzi et al., 2004; Kiledjian et al., 1995; Kosinski et al., 2003; Makeyev and Liebhaber, 2000).

We hypothesized that RNA-Binding Proteins (RBPs) that recognize C-rich elements may be differentially expressed in HFF and iPS cells and that such factors might contribute to coordinated regulation of genes important for pluripotency. We therefore investigated the expression of four Poly(C)-Binding Proteins (PCBPs) in HFF and iPS cells with the goal of determining whether these

proteins were involved in the destabilization of mRNAs with 3'UTR C-rich elements in iPS cells. As an introduction to these studies, the role of C-rich elements and the PCBP family are discussed below.

6.1.1 C-rich cis*-elements are important for mRNA metabolism*

C-rich motifs are over-represented in the 3'UTRs of human mRNAs, often close to polyadenylation sites (Louie et al., 2003). These are one of several identified motifs that influence mRNA stability, but are among the least understood. Additional roles in regulating polyadenylation and translation illustrate the importance of further characterizing these *cis*-acting sequence elements. For example, short stretches of C residues in the 3'UTR are associated with stabilizing effects resulting in longer transcript half-lives, as demonstrated by the human *α-globin* mRNA (Wang et al., 1995; Weiss and Liebhaber, 1995). RNA-binding proteins interact with this motif to form a complex that protects the poly(A) tail of the transcript from shortening (Kiledjian et al., 1995; Waggoner and Liebhaber, 2003). C-rich motifs lying 30 - 40 nucleotides upstream of poly(A) sites enhance cleavage and polyadenylation reactions; important events in 3' end processing (Hu et al., 2005; Louie et al., 2003). In another mechanism, binding of *trans*-acting factors to C-rich Differentiation Control Elements (DICEs) in the 3'UTR blocks ribosome assembly on the initiation codon to impose translational silencing (Ostareck-Lederer and Ostareck, 2004; Reimann et al., 2002). DICEs can also coordinately regulate transcription, splicing, and translation (Meng et al., 2007), further highlighting the importance of C-rich elements to establishing expression profiles.

6.1.2 Poly(C)-binding proteins are important trans-acting factors

To date, the majority of *trans*-acting factors that bind C-rich elements are members of the Heterogeneous Nuclear Ribonucleoprotein (hnRNP) K Homology (KH) domain superfamily, namely HNRNP K (Siomi et al., 1993) and PCBP family members PCBP1-4 (HNRNP E1-4; Kiledjian et al., 1995;

Makeyev and Liebhaber, 2000). Polypyrimidine Tract Binding Protein (PTBP1/HNRNP I; Kosinski et al., 2003) also binds C-rich elements to influence stability but its predominant function is as a splicing regulator (Fred et al., 2006; Sawicka et al., 2008). NOVA1 and NOVA2 proteins also contain KH domains and are important for mRNA metabolism although their expression is restricted to neurons (Buckanovich and Darnell, 1997; Yang et al., 1998). We focused on the PCBP family as candidates for regulating the stability of C-rich element-containing mRNAs in HFF and iPS cells.

PCBP1-4 are encoded at separate loci but share a similar protein structure characterized by the presence of two KH domains (KH1 and KH2) in close proximity at the N-terminus followed by a more distant KH domain (KH3) at the C-terminus (Makeyev and Liebhaber, 2002). These conserved domains serve as nucleic acid binding regions and are separated by less conserved hinge domains. In PCBP1 and PBCP2, a Nuclear Localization Signal (NLS I) is located in the hinge region between KH2 and KH3. PCBP2 has a second signal, NLS II, embedded within KH3 (Chkheidze and Liebhaber, 2003). Although each member is able to bind C-rich sequences, the extent of their functional redundancy is not well understood (Makeyev and Liebhaber, 2002). These proteins are involved in an array of processes including transcription, splicing, polyadenylation, translation, and mRNA decay. The expression of PCBPs is regulated to alter these processes when changes in gene expression are necessary, such as during development or differentiation (Paillard et al., 2000; Radford et al., 2008; Salomonis et al., 2009). The ability of these proteins to function in a coordinated manner that influences multiple mRNA processes at once (Meng et al., 2007) makes them intriguing candidates for modulating the stability of C-rich 3'UTR mRNAs in iPS cells.

6.1.2.1 Expression of PCBPs is complex and differentially regulated

The PCBP family of four genes is estimated to encode at least eight distinct proteins (Figure 6.1). *PCBP1* is intronless and encodes a single protein. Interestingly, intronless genes are frequently

important for proliferation and development in human cells (Grzybowska, 2012). The remaining three PCBP genes contain multiple introns and undergo alternative splicing. *PCBP2* mRNA produces an alternative splice variant excluding exon 8a corresponding to a 31 amino acid region between the KH2 and KH3 domains (Funke et al., 1996). Two *PCBP3* mRNA isoforms result from alternative splicing of exon 8 (RefSeq) and *PCBP4* mRNA is expressed in its full-length form (*PCBP4/MCG10*) and as two splice variants (*PCBP4B/MCG10as* and *PCBP4C*). *PCBP4B* mRNA excludes exon 9 within the KH2 domain (Zhu and Chen, 2000) while *PCBP4C* excludes a portion of the last exon, exon 15, encoding a region of KH3 (Makeyev and Liebhaber, 2000). Although individual functions and relative contributions have not been assigned to alternative splice variants, it seems likely that changes in KH RNA-recognition domains would affect their RNA-binding capabilities. Distinct patterns of *PCBP4* alternative splicing are evident during stem cell differentiation into cardiac precursor cells (Salomonis et al., 2009) and possibly indicate that the roles of PCBP splicing isoforms are different from those of their full-length counterparts.

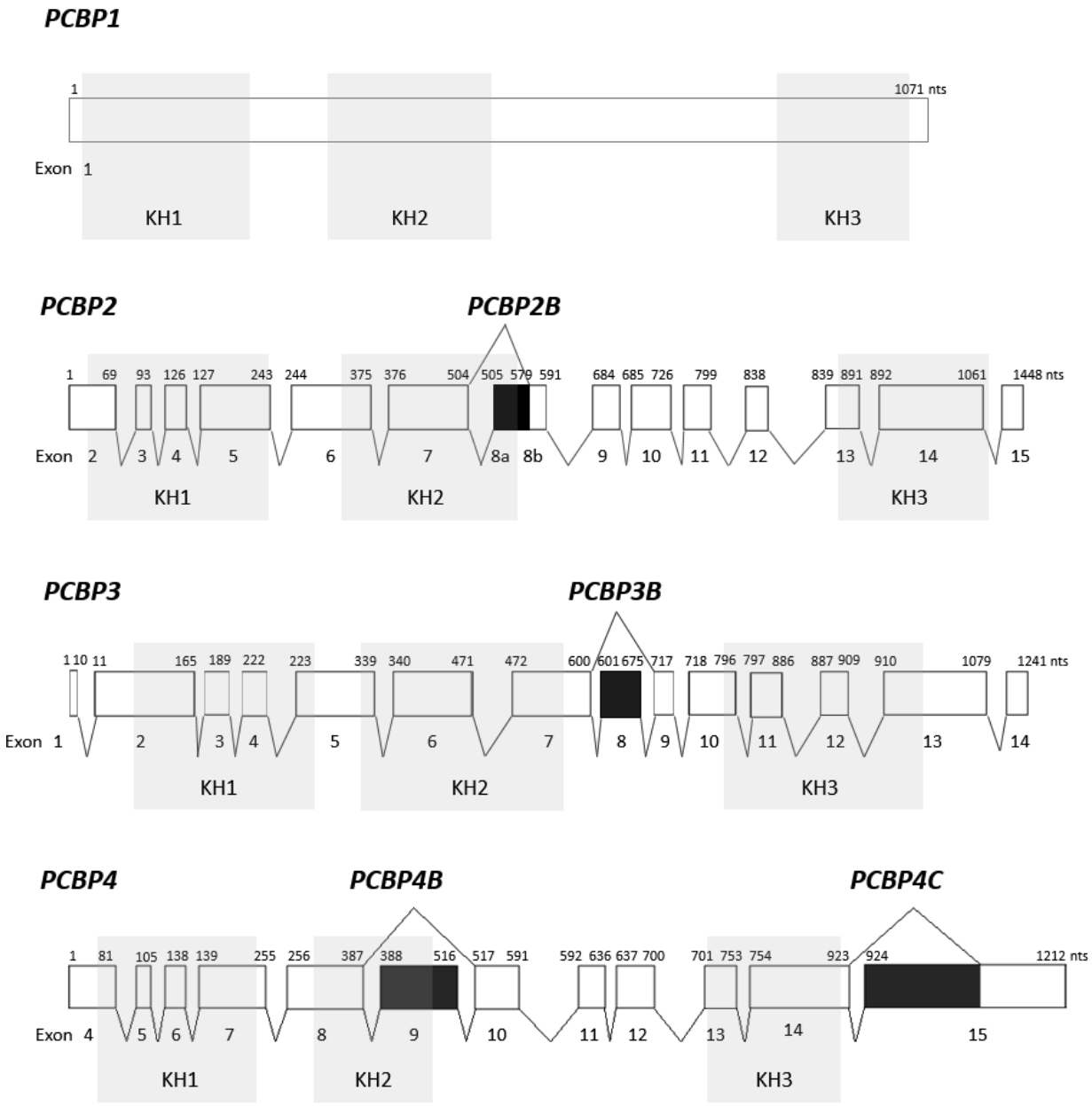


Figure 6.1 The PCBP family encodes eight proteins. Alternative splicing maps depict the ORFs of PCBP mRNA isoforms that generate distinct protein variants. The numbers above the exons indicate nucleotides while the numbers below indicate the exon number. The lighter shaded boxes indicate regions encoding KH domains and the darker shaded boxes indicate portions of mRNA alternatively spliced.

Post-translational events can alter the function or localization of PCBPs. Phosphorylation status influences the ability of PCBPs to bind specific DNA and RNA sequences (Leffers et al., 1995), thereby exchanging roles between transcription, splicing, and translation (Meng et al., 2007). Upon TGF-β-

activation, PCBP1 is phosphorylated and released from structural TGF-β-Activated Translation (BAT) elements in the 3'UTR of mRNAs important for cellular differentiation, permitting their translation (Chaudhury et al., 2010). In another example, phosphorylation of PCBP1 not only disrupts protein-RNA interactions but also alters protein localization, impacting endothelial Nitric Oxide Synthase (*NOS3*) mRNA stability under hypoxic conditions (Ho et al., 2013).

Shuttling of PCBPs allows them to influence both nuclear and cytoplasmic events (Chkheidze and Liebhaber, 2003; Makeyev and Liebhaber, 2002). PCBP1 is predominantly nuclear, moving in and out of nuclear speckles where splicing processes occur (Berry et al., 2006; Chkheidze and Liebhaber, 2003). PCBP2 is also nuclear in localization although its truncated splice isoform localizes to both the nucleus and cytoplasm (Chkheidze and Liebhaber, 2003). In HeLa cells, PCBP3 and PCBP4 are restricted to the cytoplasm (Chkheidze and Liebhaber, 2003). Although PCBP3 is not predicted to contain a NLS, the protein can bind DNA at promoter regions (Choi et al., 2007), suggesting that it may enter the nucleus under certain circumstances. It remains to be seen whether enrichment within defined cellular locations are indicative of specialized functions performed by each PCBP.

6.1.2.2 PCBPs have diverse functions

PCBPs have been implicated in a number of mRNA processing events including transcription, splicing, polyadenylation, translation, and mRNA stability. Although PCBP1 and PCBP2 are better characterized than the more divergent PCBP3 and PCBP4, all four proteins have demonstrated an influence on mRNA metabolism through C-rich *cis*-acting elements, some of which are redundant. For example, PCBPs can positively or negatively regulate transcription of the Mouse μ-Opioid Receptor (MOR) gene by binding a poly(C) element in the promoter region (Kim et al., 2005). PCBP1 and PCBP2 activate transcription (Choi et al., 2008) while binding by PCBP3 results in transcriptional repression (Choi et al., 2007). However, transcriptional regulation by PCBPs has only been demonstrated for a few

genes (Choi et al., 2009). Notably, PCBP1 is required for maintaining a transcriptionally-silent state in mouse oocytes although the mechanism is unknown and could be direct through modulation of transcription or indirect by affecting expression of other genes (Xia et al., 2012). As discussed in Chapter 5, differentiated cells are less transcriptionally active than pluripotent cells and PCBP1, or other PCBPs, may function to maintain this status.

There is also evidence that PCBPs have a role in alternative splicing. PCBP3 is an activator of *tau* mRNA splicing in association with early-onset dementia (Wang et al., 2010b) while PCBP1 negatively regulates alternative splicing of the *CD44* mRNA in HepG2 cells (Zhang et al., 2010). Again, relatively few mRNAs have been identified as targets of PCBP-regulated alternative splicing.

An intriguing function of PCBP regulation is the ability to influence polyadenylation. Transcripts are targeted to enhance 3' end processing through an interaction between PCBPs and the 3' processing complex (Ji et al., 2011). Binding of PCBPs to mRNAs can also influence alternative poly(A) site usage (Ji et al., 2013b). In *Xenopus* embryos, a poly(C) tract in the 3'UTR of maternal mRNAs is bound by the PCBP2 homologue to mediate cytoplasmic polyadenylation during early development (Paillard et al., 2000). Further, human *α-globin* mRNA, which is stabilized by PCBPs in erythroid cells, undergoes cytoplasmic polyadenylation in *Xenopus* embryos but not in a mature cell environment (Vishnu et al., 2011), indicating that this function of PCBPs may be developmentally regulated.

PCBPs also have a role in translational regulation. *In vivo* experiments show that PCBPs stay bound to *α-globin* mRNAs during active translation (Ji et al., 2003). PCBP1 binds a triple GCCCAG motif in the 5'UTR of *PRL-3* mRNA encoding a gene associated with metastasis and represses translation (Cloke et al., 2010) by physically inhibiting ribosome assembly (Wang et al., 2010a). PCBP1 and PCBP2 also bind the Internal Ribosome Entry Site (IRES) in the 5'UTR of transcripts made by several viruses to influence replication and translation including poliovirus (Gamarnik and Andino, 2000), coxsackievirus

(Zell et al., 2008), human papilloma virus (Collier et al., 1998), and hepatitis C virus (Spångberg and Schwartz, 1999).

The most studied role of PCBPs is their regulation of mRNA stability. Binding to 3'UTR C-rich elements allows PCBPs to increase mRNA half-life by competing with destabilizing factors such as miRNAs (Ho et al., 2013) and through a direct interaction with PABP to prevent deadenylation (Wang et al., 1999) and endonucleolytic cleavage (Wang and Kiledjian, 2000). As mentioned previously in Section 6.1.1, the stabilizing effects of C-rich elements in the 3'UTR of human *α-globin* mRNAs has been extensively studied and revealed that PCBPs are essential in this regulation. Additional support is found wherein PCBP1 binds the 3'UTR of androgen receptor mRNA in human endometrial stromal cells to regulate expression (Cloke et al., 2010) and *NOS3* mRNA to confer stabilization (Ho et al., 2013). Also noteworthy is the role of PCBP1 in stabilization of folate receptor (*FR-α*) mRNAs in response to folate deficiency (Tang et al., 2011). Folate is essential for proliferation of neural stem cells (Sato et al., 2006) and levels must be maintained to ensure proper neural tube formation (Chen et al., 2012). In another example, PCBP2 binds and stabilizes collagen I (*COL1A1)* mRNA during hepatic cell activation (Lindquist et al., 2000). There is also evidence that PCBPs destabilize mRNAs wherein PCBP4 binds the 3'UTR of cyclin-dependent kinase inhibitor *p21* mRNA (also known as *CDKN1A*) to modulate steady-state and induced levels by negatively regulating transcript stability (Scoumanne et al., 2011). p21 is a regulator of cell cycle progression (Czerniak et al., 1987; Jung et al., 2010) and also regulates expression of pluripotency factor SOX2 (Marqués-Torrejón et al., 2013) to suppress pluripotency (Hong et al., 2009).

Although much remains to be learned regarding the involvement of PCBPs in pluripotency, the ability of these proteins to regulate mRNA stability through 3'UTR *cis*-acting elements and the anti-proliferative role of PCBP4 as a tumor suppressor (Zhu and Chen, 2000) led us to hypothesize that this family is involved in the regulation of mRNAs containing 3'UTR C-rich elements that are destabilized in

iPS cells compared to HFFs. The results of experiments to elucidate the expression patterns of these proteins in differentiated and pluripotent cells are presented below[5].

6.2 C-rich hexamers are over-represented in 3'UTRs of destabilized mRNAs in iPS cells

The analysis presented in Chapter 3 revealed that C-rich hexamers were over-represented in the 3'UTRs of 151 transcripts out of 548 destabilized iPS cell mRNAs (Appendix A4). The top 20 hexamer sequences showed significant enrichment (up to P-value = 8.49×10^{-14}) in destabilized iPS mRNAs and are summarized in Table 6.1. We saw that the vast majority of these putative *cis*-elements were rich in cytosine residues while adenines and guanines were noticeably sparse. C-rich 3'UTR elements are frequently linked with transcript stabilization (Scoumanne et al., 2011; Waggoner and Liebhaber, 2003) and their overwhelming representation seen here warranted further investigation of this class of mRNAs.

[5] Some of the results presented in Chapter 6 appeared in: Neff et al. Global analysis reveals multiple pathways for unique regulation of mRNA decay in induced pluripotent stem cells. Genome Research. Vol. 22, No. 8, pg. 1457-1467, August 2012.

Table 6.1 List of hexamer sequences over-represented in the 3'UTR of mRNAs destabilized in iPS cells. The p-value represents the level of significance of each hexamer.

hexamer	p-value
CCCCUC	8.49E-14
CUCCCC	1.54E-13
CCCCCU	1.61E-12
CCCUGC	1.17E-10
CACCCC	1.76E-10
CCCCUG	2.75E-10
GCCCCC	8.34E-10
GCCCCU	1.25E-09
CCCACC	2.47E-09
CCCUCC	2.70E-09
CCACCC	6.10E-09
AGCCCC	5.22E-08
CUGCCC	5.30E-08
UUUUUU	8.60E-08
CCCCAC	5.32E-07
CCCCUU	1.33E-06
CCCCCA	1.45E-06
UCCCCU	1.45E-06
UAUAUA	1.77E-06
CCGUAC	2.51E-06

6.3 Destabilized mRNAs encode factors associated with transcription, chromatin, and development

To determine whether the mRNAs bearing C-rich 3'UTR elements had any functional similarities, the annotation program DAVID (Huang et al., 2009) was used to identify Gene Ontology (GO) terms associated with this mRNA set. As shown in Table 6.2, GO terms with the greatest enrichment described functions of transcription regulators, chromatin modifiers, and factors important for embryonic development. As discussed in Chapters 4 and 5, stem cell gene expression requires unique regulation of chromatin structure and controlled transcription of pluripotent- and lineage-specific genes. Therefore, it was exciting to find that this class of differentially regulated mRNAs also encodes factors with roles in these functions. Interestingly, we also saw enrichment of GO terms associated with neural tube development and tube closure. This over-representation is consistent with the hypothesis that PCBPs

may be involved in modulating stability of these transcripts as PCBP1 regulates the stability of folate receptor mRNAs in neural stem cells to prevent neural tube defects (Chen et al., 2012; Sato et al., 2006).

Table 6.2 List of Gene Ontology terms associated with mRNAs destabilized in iPS cells that bear 3'UTR C-rich elements. The p-value represents the level of significance of each term.

GO Term and Description		p-value
GO:0006350	transcription	4.68E-04
GO:0006357	regulation of transcription from RNA polymerase II promoter	4.80E-04
GO:0045449	regulation of transcription	6.61E-04
GO:0006325	chromatin organization	2.42E-03
GO:0021915	neural tube development	2.65E-03
GO:0043009	chordate embryonic development	2.94E-03
GO:0009792	embryonic development ending in birth or egg hatching	2.94E-03
GO:0016568	chromatin modification	3.13E-03
GO:0045893	positive regulation of transcription, DNA-dependent	3.14E-03
GO:0045944	positive regulation of transcription from RNA polymerase II promoter	3.29E-03
GO:0051254	positive regulation of RNA metabolic process	3.30E-03
GO:0010629	negative regulation of gene expression	4.19E-03
GO:0045941	positive regulation of transcription	6.67E-03
GO:0035113	embryonic appendage morphogenesis	6.74E-03
GO:0030326	embryonic limb morphogenesis	6.74E-03
GO:0016055	Wnt receptor signaling pathway	8.07E-03
GO:0010628	positive regulation of gene expression	8.46E-03
GO:0060606	tube closure	9.79E-03

We then went on to validate the destabilization of three mRNAs with C-rich elements in their 3'UTR. Although the stability of *DGCR8*, *DUSP7*, and *TOB2* mRNAs has not previously been studied, microarray data indicated that these transcript were destabilized in iPS cells and hexamer analysis identified the presence of 3'UTR C-rich elements. *DGCR8* mRNA encodes a component of the Microprocessor that is important for miRNA biogenesis (Gregory et al., 2004) and silencing self-renewal in stem cells (Wang et al., 2007). *DUSP7* mRNA encodes a phosphatase important for pluripotent regulation (Chappell et al., 2013) and *TOB2* mRNA encodes an anti-proliferative factor that inhibits

progression into S phase (Ikematsu et al., 1999). Destabilization of *DUSP7* mRNA in iPS cells was also verified in an independent set of matched cell lines (A. Jalkanen personal communication). Figure 6.2 below shows the 3'UTR sequence of each mRNA where C-rich hexamers containing five or six cytosine residues are highlighted in yellow. However, we also noted that several U-rich elements were present and may have contributed to the differential regulation of stability of these transcripts.

3'UTR of *DGCR8* mRNA (NM_022720.6)

GGGAGGUGGCACGGGCCAGGGCGCGGGGGCCGCCAGCCGCACUUCUGAGGAGACCAGCAGUCAUGCAUCGUGCACCACAGUGUCAGGCCUCCAACCCACGCUCCUUCCCUGUGGC
CAACCUGUGGGCCCGGCCUUAGGGUGGAGGCUUUAGUGUACAGGGACAGCCAUGGCCACACAGCACACAUGUGGAGCAGCGGCUCUCCCCUGGAAAGCUCCAGGCCUGAAUGGAUG
GACUCAGCGACUGCACCAGUGGCAGCUGGUGACUGUGGACAGUGGUGGACCCUGCUUCUGUGCACCUGCUGCAGGCUCUUUUUAUGAAGGCUUUCAUGAAUUUUAGUAUGUAAUA
CGCACUGACGACACAUGAUGCUUGGAUGACAGAUGAGAGGGGAUGGCUGAGUCCUGUGGCUGGCCCGUGAUGCCAGGUGGCCCAUGUGCCCAGGGCGCCUGCAGGGCUGCUACAG
GGACCUGGUCAGGAGGUGCACAUGGUGCCCUGCCCUCACCCACCCUCUGUGUUUCCCCUUCUUUGAAAAGGUAGAAGAGAAAGGAAUAUUUUAAACCUUUUUGGCUUAAACAGAA
UUUUAGCAUCAGAACUAGCUUUCUGGGAUUGGAGGCAAACCAUCAAGGUGGUCCCUCUCCAGUCUGGACACGAUGCCAGCAAGGAUGACGUCCUGCCACCUCCUGGAGUUACCCU
GGCCUCCUAGGGUCCCUUUUUCUGAUGAAGUCUUAAUUCCCUAAAAGCGCCUCUUUGGACACUGAGGCCCUCUCUGCCUUUCCUGGCCUCCGGCAACAGUUUUUUACAAAGAUUU
UUUGCAGUCGAGUCCAUAUGUCCACCCAUUGAUUUUUAAAGCUUUUGUGAUAUUUUAGCAUUUUGAAAGACUUUCACAGUGAGAGUAGAAGGUAGAUUUGGAAUCAUGCAUUUUA
GCAAGUGGACUUGUUGAAACAGGAAGCAAGGGCCUUCAGUGUAGCCCAUUCUUGAUCCAGAGCUGUUGCCUGUGACAGCGGUUUCUCUGGAUGUCAAAGGCAGCUGCCUGGUGCC
CAGCUUGCUUCUCGACUGGUGGCCCCUAUGGGUGGGUGUGCGAUGGAAAUGUGUUCCUGCCGGAGUCUGAGGCACCAGGGUGUGCUCAAAGGCUGGCCCCUGGUGGUGGACUGGCA
CCUGUGCAGAGUGCCGUGUGCUUGUGGUGCGCCAUCUGAAGCAAGAGUCCAGCGUUCUGCCGUGUCUGUCCCCCACCAUGCCCCCUACAGGCGGUACUGAUGGCGCUUUUUUUUU
UUUUUCUGUCAGGAAAACAAUGUUGGCCUGUGGGCCGCCCACAACAUAUCCUUCCCUCACUACCUGUGUGACCAAGGUUGGCUUCUGUUGACCUUUAAAAAAGAAACCCUCAACU
CAAAUUGCUAUAAUUAGACACUUGCUUCUGUCUUGCCUCCUGUCUGCAGCUGUGAAUAGUCAUUUGACUGUGACUGUUGCCCUUAGCCAGCCAGAUGCGCCUGUGAACCAAAGCU
UCGUGCACAUGUGUUCCCCUAAAGGUUGGGGAGCCUCGCUGUGUCUUGCUGUUCCCAGGCACCACCACAGCAGGUGCUGCCAUACUCUUGUGGUCUCUGUGCGCCCCCCCCCCCC
CCCCACCCGUCUGCCAAGCAUGGGUAUGAAUCGUGCACACAGCCAUGCUUCAAGGCCGGGGCAGGGGAGCCUGUGCUGAUGCCAUCCAGGGCACUGGGCUGUGCCUGGAAGGCGA
GCCUUGAUUGUCUGAACACAUAAAGCAAACUGUCCAGAAGGG

3'UTR of *DUSP7* mRNA (NM_001947.3)

AGGCCUGGUGCACGGGGGGCAUGGCACCAGGCCCCUGCUCGGCUCUCCACAGGGCUAGGUGGGAGAGCCCAAGCCCGCCACCUCUGGCCUGAGGAACCCCCAGAUGUCACCUGUG
CCCAGAGGCCCAGGCUGAUCGGUGUCGGAGCGCCCCUCACCAUCCUUGGGGGCAGGGCCCGCAGGCAAGGUCUCCCACUGCAGGGCUUGCUGGAGAGGCCUCGGCUCUUGGACAC
GUGGCUUUGGGCGUCCACCAGGGCCUCAUCCUGUCCAGGACGCUCCUUUCUGCUGACAGCCCAGCCAGUUUGGCUGUUUUUUAAAGACACAUCCACGGACCUGAGUUUACUUUUU
ACUUUUGGCAGGUAAAUCCAAGCUCCCUGGAGCACAAAGAGUGUUUGAGCUCUUCUUGAUUUUUCUUUUUUUUUUUUUUUUUUAACAAAAAGUGUUAUUUUCAGGCUACAUGCAA
CAGUGGAUUGUAUAACCCAGUAUUUCAUCCCUUUCCUGAUCCUGCAAGAGAGAGAAAUGUUCAGUUUUGUUUUUCGUUCAAUUUCAAAAAGCAAGUAUCGUGUUUGUUUUUGUUU
UUGUUUUUUAAUGGAAAAGACAGACGCCAUGUUGACCUUGACCAAAUGCGUUUUGCACAUGUGUGAAGCCUCUGUUUCCUCAGCGGGCCCUUCUUGAAGGGGUGGCUUACUGCUC
UUCAGGUAUUGGGACCCCCAGUCGUGCCGCCUCCCACCCUUGGCCCAGUCUGACUUAGCUCUGCAUUCGCCUGUGAUGACUGCUGCAAGUUGUGAUUGGUGGGCUGGAUGGUGGU
CUUUGCAUUUUCUGCCGUCUCAUCCAUCCCUCUGGUGGUCCCUGCAGACCUCCCUGAGGCAGAAGGUGGGGACCCUGGGGCCCUGCCGCUGUCCCCCAGGAGGUGGGCCGCAGGG
GGCUGCCAGGCAGUAGCAUUAGCCCUCUGGGGUCAGCCCCUAAGAAGGCCUGGGCUCUGAGCCCUUUUAUAUACUUAGCCACUACUUCUGUCUGUCUGUCUGUCUCUCUCUUC
CUCUCCCUCUCUCUCUUUCUCUCUCCCCUCUCUCUCUUUCUUUCUCUCUCCCCCCCUCCCUCUCUCUUCCUUUCCUCUCUCUUGUUGUAACUGGGAGUGGAGGCCCAGUGGC
UGGGGAGACAUUAGGUGGUGGGGCCCAGCCCGACCUCCAGGUUCUUCCUUCUCCCUAGCUGUUGCUUUGGUCUGGCCACUCCCAGCCCCCUUGUCCCCCUUGGAAGCUUGCCCUGC
CCUCAUCUUGCCCAUGCCUUCUACUGCCAGGAGACUUGCACCCAUUUCAACCCUAGGGCGGGGGCAAGUGGGGCAAGGAUGGACCAGCAGAAGGGGGGUAAGGCUCUGUUCACUU
CCCCCUGCCUCCACAGAACGAAGCCACGGAUUCCGUUAUCUUCCUCCAGUUUUGUUCCUUCUCCAGCCUCAGUUCCACCAGGUGUCAGGACUGCAUGGGGGCCUGGGGCAGGCAG
AGGAGUCAGGCCAGGGUCCCUGACGGAGCAGCACUCAGCAUGUGAGUGAGGCCACAGAAAAACUCUGCCCCACUGCUUCUUACCUCACGGGGGUGGCUUUCAGGGAUUCUUUAGC
GCAGCAGAUUAAAAUCUUGCCACAGUCGAGAAAUUGACAACAAGCUUCCAUGCUGUACAUGGUUCUCUUUUUCUCUCUUUUAUUUUUAAAAAGAAAACCCAGAAAGAUGUACCAG
AUUUGUGUAAAUGAGGGUAUGCCAGAAGGUGGCCAGUUUUGCUUUAUGAUCUUAUGAAGGAAGAUUUGUGACCCUACGUAUAUAUAUACACACACAUACAUAUAUAUAUAUCC
CGAACCAACAACGGGACUUUGUUUAUAUUGCAAAUAAAUAUUAUUUUUUCUUU

3'UTR of *TOB2* mRNA (NM_016272.3)

CCAUCUACCUGCCCGUGGGGCCAGGAGCACCCAAGACCACAGAAAAGAGAAAGGAAAGGCCAAAAAAAAGAGGAAAAGAAAAAUGUACAAAAAGAUUUUCGAUCUUCUCACUCUU
ACUUCUAGAAAAAAGAUCCAGCCCAGGCACGGAAUGGGAGGGCCCUGCAAAGCACCCAAGUGUGUUUCACUCACCCCUUACCGCUCGCCCACCGGCCUGUGAUUUAUUUGGUUGG
UUUGGGUUUUAUAUUUCUUUUUCUUUUUUUUUUUUCUUACAUGUAGUUUUCUAUAUAACACUCUUUAAUGAGUGGCUUCCUGUGCUAAGAAUGGUCCAGAUGUGAACUGCUGCUCC
CUGCUCCUCCCUUCCUCCUUCACACUCCUAGUUUAGGAUUGAUGUGUCCAAGCUCAUAGGGAAGAAAGAGCUGCAGCUGGAGCCUGACCCUCCCUGGAACAGGGAGAGGCUGGCC
CGUGUCACUGCCUCUGUCACCCAGCUAUCCUCGCACUCAAAGCCAUCCAACCUCAGCAGGCCUUCUCGGGCCCUGCCACCUGAAUAGGUCUGACCCCAGUCCCCACUCCCUUUCA
GGCUGGGCCACAGGGAGUUGCUGGCUGGCCACUUGACACCCUCCCCCUGGAGCUGAUGUCUGUGACUACAGGGAGAUUAGCACUUCGUCUAGUGAAACUCCUUUCAUCUCUGUCC
UAUGGCCCCACCCCACCAUUCCCCUCUAGGCCCAGACCAUCAUCAUGGCCCUGGCCAUUGAUGUGUAUGGGCUUUCUUUCCCUCAUGCAGCCCCUUCUUCCCCAGGCUGGAUGUUG
UGGGUAGGGAGGGGUGUGUGUGUGUGUAUUUGGUUUGCUGUCCUUUUUUAAAGGAUUCCAAGCCAUGUGAAACUUCCCUUCUGGAUGUGAUUCUGGGUCGCAAGUCCUUAUUUAU
AUGUGAGGCUGGGGAAUGGGCUGGGGGUAUUGGCAGUCCUUUUGCAGGGCAGUGUGUGUGGUGGGGUGACACCGCUGUGGCUUAGCCCAAGACACUCCCAGAGGAAAACACUGCA
GAAGGAACUGGUUUGCAGACUGUGGAAGGAUCUGCAGUUUUGUUUUUGACCAAAAAAAUAAUAAUAAGUUAGCUCUGAAGGGCAGAGGGAAUACCCAAGCCCCUGAUGCCUAUGA
GAAGUCCCUGGACUUCAACCCUCCUGUUGUUUGGCCUUAGCCCAGAGGGAGCUGCUCACCUGAGCACCCUUGGGGGUGGGCAGAGAGGCAGGGUGGGAUUUUAGAGUUAGUGUCU
GUGCGGGGGCAGCCCUGAGCCUGGAGUUGAGACUUUGGGGUCUCUUAGUUUGGAGGUGUUGAGUGCAUUUGUGCCCCUGCCUGGUUGAGAGCUUCUUGGUACCUCUUGCCACCCC
UUCUCACUGCCCUGACCCAACCCCACUGGACCUUGAUGCUGCGAGGAGUGGUGUCCUGACGGACUCAGCACUCCCGCCUGAUGUAUUGGAUCAUAGGAGAGCACUUGCUCUCCUG
CCUCUGCCAGGAGAGGGCUUGUUCCUCCAACUCUAGGAGGCCAGGCAAGCAUGGACAGGAGCCAAGGGAGCAGGGUCAUUAACUUUUUCUUCUUUGCAAAGUGGGCACUUGGCAU
CAGGGUCCCAAUCACCAGAAAGCACCAAAGCCCCUGGCACCCCACCCACUCCAUCCUACCCAGGGACCCCAAGUAGGCAACUGUUAUGGCAGUGGGUCCAGCCCAGGCCAGCACU
GCCAGCCUCCUCUCCCUGCAGUAGGCACCAGCUCUACCUCCCCCGGCAGGCAAUGUCCUGGCUUCUCAGCCCAGCACCAUCUGUUCCCCUAGACUUCUCAGGGGCCAGCCCAGUC
UGGGCCACCCUUUGUUUCCCUCAUCCUCGGCUCCCACACAGGUGACAGACCCAGCAGAUAGCUUCUCUCUGGGAAAGGUUGGAUGCUGCCUUACAUCCCCUUCUAGCCCUCCCUCC
CAUCCACACACACAGGCACCCACCCACACCAGGUCGGCUUGUUUCUCACAUGUAGGGAGAGAGGGGAGACCAACCCCUUUGUGUCUUUUGAAAUACGAAGAAAAAUGUGUGUUCA
GGAGCAUGACUCCAGUGCUGCGCUCUUGGGCCCAGUUCAGUCUGUCUUGUCUCAAAUCUAGGCAUUUUUGCUUCAAUUUUAUUUUUUUUAAGAAUACAAAAACAGAAAUCUGCAC
UAAUUUACCUGGUUUCGUAGGAAAACUUUUUUUUAUUUUUUACAUUUUUUGGUGUCCGUUUGUAUUGAAUAAUUUGCUACAUUUGUAAAAUGUAAGAGGUAUAUAAUAUAUGUAU
AUUUCUAACGUAAAAAACAUAAUUUUUUUCUUUUCAAGAUUUUUUCUUAAAAAGAUGAGAGAAACAUAUUUUUUCAGGAAAAACAAACUUUAAAAAAAAAAGAGGAGAAAUAAAA
CCUUUUCUCCCCUUUCCCCAUCCUCUAUCUAUCCCUCUUUCCCAGGAACAAAUCAAAAGGUGGAUUAUCUUCUGAAGAAUGGAAACUGUUAGUCCAGAAUGAUGUGUUUUUCUCA
AUGCAGUGAGUUAUAGAUUCUCUAGUUUUCUCCCUAGGGAUGGGAAGGGGCAUUGAGGCAAGCCUGGAGAGGAGCCCGGGGAGCAGGGUCAUGAACUUUUUUCUUUAGUGAAGG
AGGAAUACAAUCAAGGGUUUUGUAUUCAGAAUGUUGUGCAAUAUUUUGGAAUGGGACAUUGGUGUGUUUAGAGAUUUUAGUUUAAAAACAAAACAAAAAGAUUGAUCAAAUCUGU
ACAGUUUCUAUUGUUCCAGAUUUUUUUAAGUUUGUAUUAAAAGCAUGAUACAUAAUAGCCUUC

Figure 6.2 The 3'UTR of *DGCR8*, *DUSP7*, and *TOB2* mRNAs contain several C-rich hexamer sequences. 3'UTR sequences were retrieved from GenBank and analyzed for stretches of five or six C residues allowing one mismatch (highlighted in yellow). Each row contains 115 nucleotides.

When measured by qRT-PCR, we found that all three of these mRNAs have shorter half-lives in iPS cells compared to HFFs, as predicted by microarray (Figure 6.3). *DGCR8* and *DUSP7* mRNAs exhibited a ~3-fold decrease of half-life in iPS cells while *TOB2* mRNA was destabilized ~2-fold. Our microarray data indicated that HFF and iPS cells exhibited half-lives of 410 minutes and 231 minutes for *DGCR8* mRNA, 371 minutes and 96 minutes for *DUSP7* mRNA, and 233 minutes and 113 minutes for *TOB2* mRNA, respectively. As these mRNAs are relatively stable in HFFs, especially *DGCR8* mRNA, half-life calculations relied on extrapolation and longer half-lives may not be as accurate as those within the 4 hour range of the experiment. Therefore, we were not able to reliably determine whether there was a correlation between the number of C-rich elements in each mRNA and the degree of stabilization. Regardless, we were able to confirm that mRNAs bearing C-rich 3'UTR hexamer sequences have shorter half-lives in iPS cells and several encode proteins important for pluripotent gene expression.

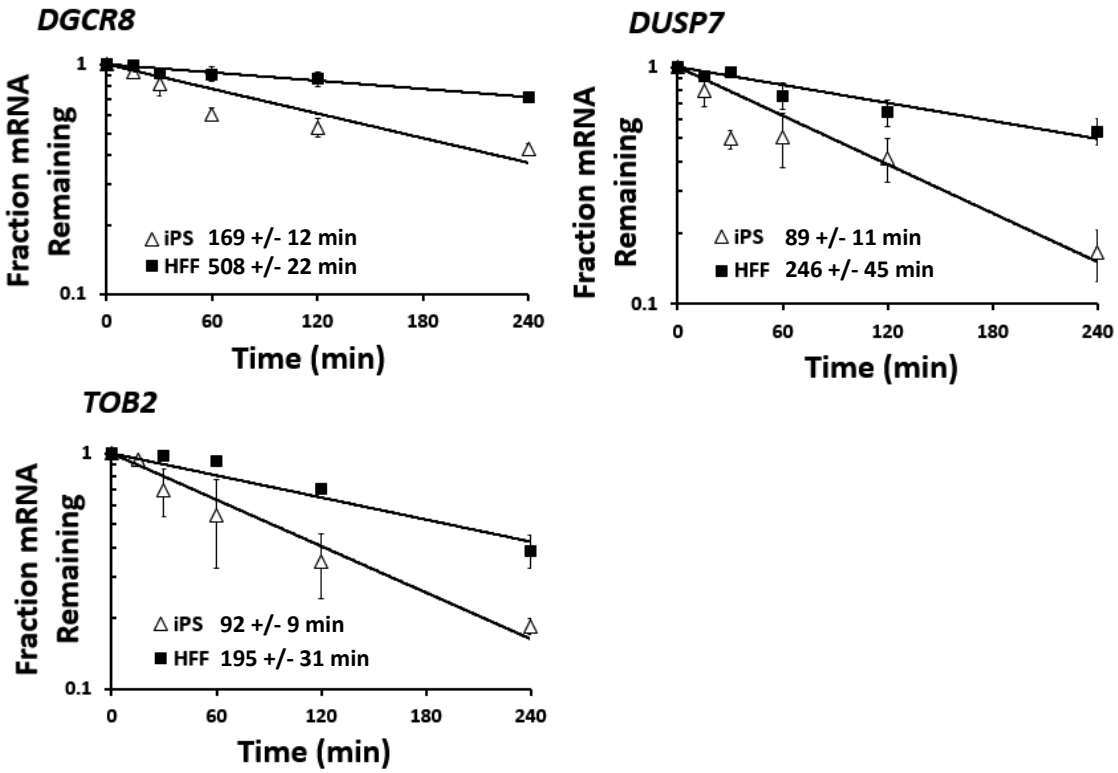


Figure 6.3 *DGCR8*, *DUSP7*, and *TOB2* mRNAs are destabilized in iPS cells compared to HFFs. Half-lives were assessed in HFF and iPS cells for each mRNA following actinomycin D treatment. mRNA levels were measured by qRT-PCR at each time point and normalized to levels of *GAPDH* mRNA. The standard deviations represent three replicates. Half-lives for each mRNA are denoted to the right of the cell-line keys.

Since poly(C)-binding proteins are known *trans*-acting factors that associate with C-rich elements to influence mRNA stability (Cloke et al., 2010; Kiledjian et al., 1995; Lindquist et al., 2000; Weiss and Liebhaber, 1995), we investigated their expression in HFF and iPS cells in hopes of determining whether they are differentially regulated in these two cell lines, possibly contributing to destabilization of mRNAs with 3'UTR C-rich elements in iPS cells.

6.4 Expression of poly(C)-binding proteins is differentially regulated in iPS cells

6.4.1 Poly(C)-binding proteins are differentially expressed

We first investigated the protein abundances of each PCBP by western blot analysis using whole cell lysates prepared from approximately equal numbers of HFF and iPS cells with antibodies specific for

PCBP1, PCBP2, PCBP3, and PCBP4C. Figure 6.4 shows a summary of these data where standard deviations were derived from quantification of three independent western blot experiments. We found that abundances of PCBP1 and PCBP2 were fairly similar between the cell lines while PCBP3 and PCBP4C are differentially expressed. Specifically, PCBP3 was expressed ~5-fold higher in iPS cells while expression of PCBP4C was increased ~26-fold in HFFs. This inverse regulation of PCBP protein abundance was intriguing and warranted further studies including analysis of PCBP mRNA expression.

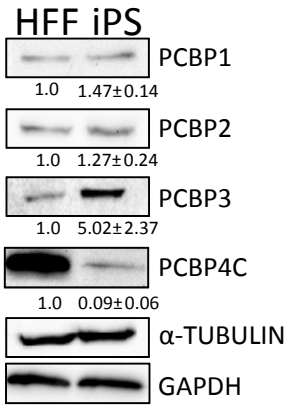


Figure 6.4 PCBPs are differentially expressed in HFF and iPS cells. Western blot analysis showing PCBP protein abundances in HFF and iPS cells. Equal numbers of cells and amounts of whole cell lysate were resolved on a 10% SDS-PAGE gel and transferred to polyvinylidene difluoride (PVDF) membranes. Quantification of iPS PCBP abundances was normalized to levels of either α-TUBULIN or GAPDH and relative to HFF. The standard deviations represent three independent replicates.

6.4.2 PCBP mRNA expression shows differential regulation

We wondered whether the different abundance of PCBP3 and PCBP4C was caused by changes in mRNA abundance or perhaps by post-translational events. The levels of PCBP mRNAs were measured in untreated HFF and iPS cells by qRT-PCR to help define relative PCBP mRNA expression. The primers were designed to constitutive exons and therefore detect all splice variants. As shown in Figure 6.5, *PCBP1* and *PCBP2* mRNA levels were not dramatically different in HFFs relative to expression in iPS cells, exhibiting ~2-fold increase, whereas *PCBP3* and *PCBP4* mRNAs appeared to have significantly different

expression levels mirroring the changes in proteins levels. There was a ~6-fold reduction in *PCBP3* mRNA abundance in HFFs while *PCBP4* mRNA expression was increased almost ~14-fold relative to iPS cells, suggesting that *PCBP3* and *PCBP4* mRNA expression may somehow be altered in pluripotency compared to the differentiated state. These results suggest that the changes in protein abundance are primarily due to altered mRNA levels (Figure 6.4).

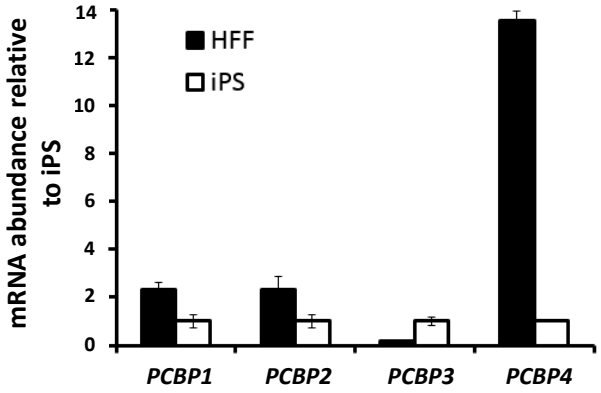


Figure 6.5 *PCBP1-4* mRNA abundances are differentially regulated in HFF and iPS cells. HFF and iPS cells were assessed for mRNA abundance of each PCBP. mRNA levels were measured by qRT-PCR, relative to the abundance in iPS cells and normalized to *GAPDH* mRNA. The standard deviations represent pipetting error from triplicate samples.

We also considered the relative abundance of each PCBP mRNA within each cell line compared to the level of *PCBP1* mRNA (Figure 6.6). Here, abundance values (normalized to *GAPDH* mRNA) were graphed relative to *PCBP1* mRNA abundance rather than relative to abundance levels in iPS cells (Figure 6.5). From Figure 6.6, we saw that *PCBP1* and *PCBP2* mRNAs were expressed within a ~2-fold range of one another while *PCBP3* and *PCBP4* mRNAs were expressed at dramatically lower levels in both cell lines. However, we cannot make assumptions about how these differences in relative PCBP mRNA abundance influence protein abundance and function.

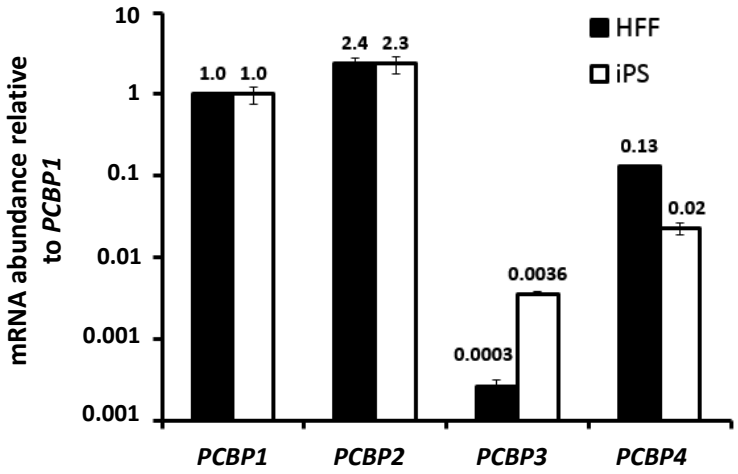


Figure 6.6 Levels of *PCBP1-4* mRNAs are differentially expressed in both HFF and iPS cells. mRNA abundance data used to generate Figure 6.5 was analyzed to determine levels of each mRNA relative to levels of *PCBP1* mRNA. The standard deviations represent pipetting error from triplicate samples.

6.4.3 PCBP4 *mRNA is differentially spliced*

As alternative splice isoforms of PCBP4 are differentially expressed in embryonic stem cells (Salomonis et al., 2009) and cancer cells (Pio et al., 2004, 2010), which in many ways resemble stem cells (Dreesen and Brivanlou, 2007; O'Brien et al., 2010; Pardal et al., 2005; Reya et al., 2001), we examined the splicing of *PCBP4* mRNA more closely. Analysis of *PCBP4* isoform expression by RT-PCR indicated that this transcript undergoes alternative splicing in both cell lines. Using cDNA made with equal amounts of RNA and PCR primers flanking the entire coding region of *PCBP4* mRNA (and both alternatively spliced introns), splice variants containing three different open reading frames were identified, all of which were more abundant in HFFs compared to iPS cells (Figure 6.7). This is consistent with qRT-PCR data that indicated *PCBP4* mRNA levels are much higher in HFFs (Figure 6.5). Sequence analysis performed on the HFF PCR products revealed that these bands corresponded to the full-length *PCBP4* isoform (*PCBP4*), an intermediate isoform excluding exon 9 (*PCBP4B*), and a smaller isoform excluding a large portion of exon 15 (*PCBP4C*).

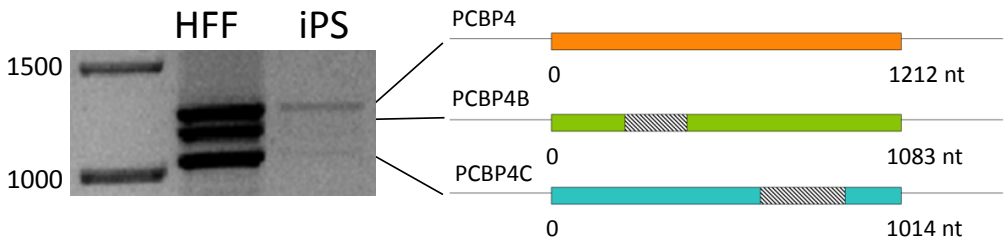


Figure 6.7 *PCBP4* mRNA is alternatively spliced to produce mRNAs with three different open reading frames. RT-PCR was performed using HFF and iPS cell cDNA and primers flanking the *PCBP4* mRNA coding region. Sequence analysis revealed excluded regions of each splice variant (right, crosshatched boxes).

To more carefully quantify splicing patterns for *PCBP4* mRNA in HFF and iPS cells, we used RT-PCR with primers designed to detect the inclusion of exons 9 or 15. As *PCBP4* mRNA levels are relatively low in iPS cells, more RNA was used in the iPS cell PCR reactions. Splicing patterns for both of these exons were then compared between cell lines. From Figure 6.8 below, we saw that there was little difference in the percentage of transcripts that excluded or included exon 9 which represents splicing to produce *PCBP4B* mRNA. In the lower panel, however, it was clear that HFFs have a higher propensity to exclude a portion of exon 15, indicating more *PCBP4C* transcripts are produced in HFFs than iPS cells relative to total amounts of *PCBP4* mRNA. This was consistent with findings that *PCBP4* alternative splicing is altered during differentiation of stem cells into cardiac precursor cells where many proteins favor truncation with differentiation (Salomonis et al., 2009).

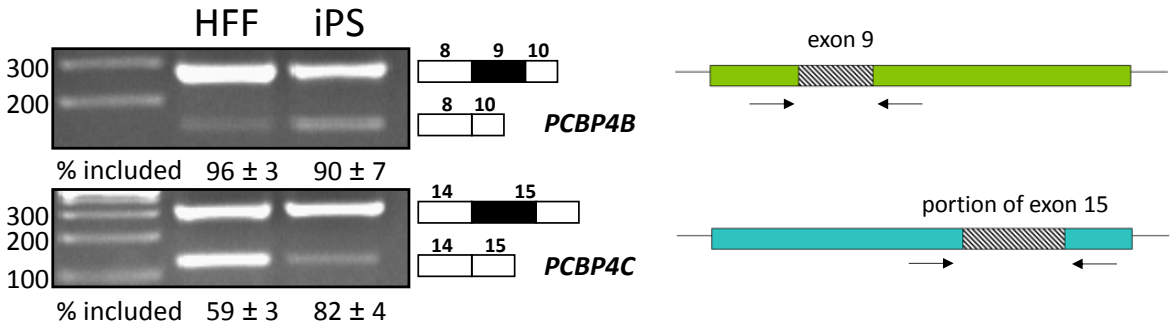


Figure 6.8 *PCBP4* mRNA is alternatively spliced to produce *PCBP4C* more frequently in HFFs compared to iPS cells. Semi-quantitative PCR was performed to determine *PCBP4* mRNA alternative splicing of exons 9 and 15 in HFF and iPS cells. Schematics to the right indicate location of forward and reverse primers used for PCR amplification (black arrows). Quantification of percent included based on density of band including the exon relative to the density of both bands combined. The standard deviation was derived from three independent PCR reactions using three independent sets of RNAs.

6.4.4 PCBP4 may regulate splice site selection at exon 15

Close examination of the mRNA sequence excluded from exon 15 to produce *PCBP4C* reveals that this region is C-rich (Table 6.3). Compared to the nucleotide representation for the coding region of *PCBP4* mRNA, the spliced portion showed over-representation of C residues. Specifically, the entire coding region of *PCBP4* mRNA was 33.2% C residues while the 197 nucleotide region of exon 15 excluded from *PCBP4C* mRNA was comprised of 46.2% cytosine. This part of the transcript encodes a proline-rich region and nuclear import and export signals (Scoumanne et al., 2011) and exclusion likely affects PCBP4C RNA-binding or other interactions relative to those of the full-length protein isoform.

Table 6.3 The portion of exon 15 excluded from *PCBP4C* mRNA is C-rich. Percent nucleotide representation of the full-length *PCBP4* mRNA isoform was compared to the percent nucleotide representation in sequences excluded from *PCBP4B* and *PCBP4C* mRNA isoforms.

	Percent nucleotide representation		
Nucleotide	Full-length (1212 bases)	Spliced exon 9 (129 bases)	Spliced exon 15 (197 bases)
A	19.7	17.1	13.7
U	19.2	22.5	19.8
C	33.2	28.7	46.2
G	27.9	31.8	20.3

The expression of *PCBP2* and *PCBP3* mRNA splice variants were not investigated in this study but it is worth mentioning that these mRNAs could also experience differential splicing between HFF and iPS cell lines (Figure 6.1). However, differential splicing during development has only been observed for PCBP4 to date (Salomonis et al., 2009).

6.4.5 PCBP4 protein isoform abundances support patterns of alternative splicing

Since our initial western blot analysis of PCBP4 abundance in HFF and iPS cells detected only isoform PCBP4C, we also investigated the abundance of the full-length isoform using another antibody. As shown in Figure 6.9 below, protein abundance of PCBP4 is similar between the two cell lines, despite reduced mRNA expression in iPS cells (Figures 6.5 and 6.7). Analysis of PCBP4 and PCBP4C isoforms by western blot did not allow comparison of relative abundances within a cell line due to variations in antibodies and detection efficiencies. Therefore, we were not able to estimate which PCBP4 protein isoform was more abundant in each cell line but it would be possible to do this in the future using recombinant proteins as standards.

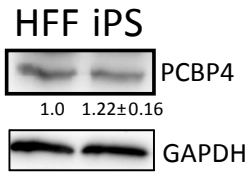


Figure 6.9 PCBP4 protein abundance is similar in both cell lines. Western blot analysis showing PCBP4 abundance in HFF and iPS cells. Equal numbers of cells and amounts of whole cell lysate were resolved on a 10% SDS-PAGE gel and transferred to PVDF membranes. Quantification of iPS PCBP4 abundance was normalized to levels of GAPDH and relative to HFF. The standard deviation was derived from three independent replicates.

Overall, we found that differential regulation of PCBP expression is evident through increased abundance of PCBP3 and decreased abundance of PCBP4C in iPS cells at the protein level. Further, *PCBP4C* mRNA is more abundant in HFFs than iPS cells.

6.4.6 Poly(C)-binding proteins may exhibit differential post-translational modifications

Phosphorylation of PCBP1 and PCBP2 greatly reduces the ability of these proteins to bind C-rich sequences (Leffers et al., 1995). For instance, PCBP1 bound to BAT elements in the 3'UTR of TGF-β-induced mRNAs is phosphorylated at serine 43 during Epithelial-Mesenchymal Transdifferentiation (EMT) causing the protein to release transcripts and allow translation (Chaudhury et al., 2010). It is possible that the post-translational status of PCBPs differs between HFF and iPS cells leading to differential regulation of target mRNAs. By first separating proteins according to pH and then resolving by molecular weight, western blot analysis was used to approximate the Isoelectric Point (pI) of each PCBP (Figure 6.10).

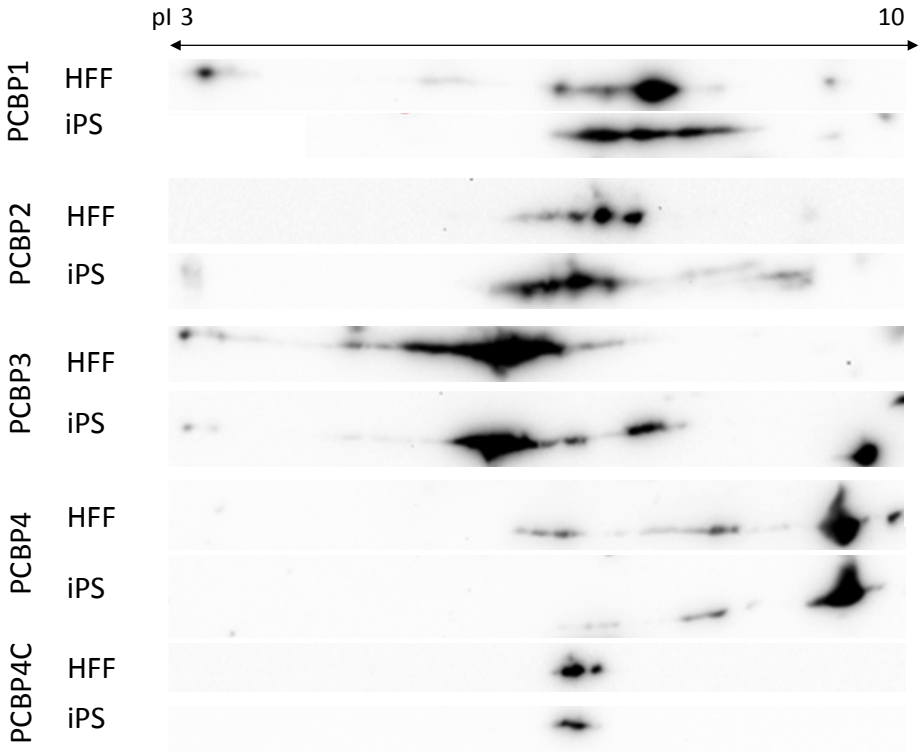


Figure 6.10 PCBPs demonstrate differences in pI in iPS cells compared to HFFs. 2D gel electrophoresis was performed using equal numbers of cells and amount of whole cell lysate. The first dimension was run on pH 1-10 gradient strips and the second dimension was run on Mini-Protean TGX gels. Proteins were transferred to PVDF membranes and antibody detection was perfomed using the ChemiDoc XRS+ System and ImageLab3.0 software.

Comparing the profile of spots, or foci, in each cell line, we noted several differences between HFF and iPS cells. First, PCBP1 immigrates predominantly in a single spot in HFFs but has both more basic and more acidic isoforms in iPS cells. The predicted unmodified or basal isoelectric point of PCBP1 is pI = 6.66 (PhosphoSitePlus; www.phosphosite.org) but without further analysis it is difficult to ascertain which spot corresponds to the unmodified protein. There was also a distinct species at the acidic pole in HFFs that is not present in iPS cells. However, this species is outside of the range predicted for PCBP1 proteins and may be non-specific.

The predicted pI of PCBP2 is pI = 6.33 (PhosphoSitePlus) and PCBP2 proteins in HFFs also appeared to be more basic in contrast to iPS cells where more intense species are enriched at the pI 3

end of the profile. This could indicate that the protein is more phosphorylated or otherwise modified in iPS cells.

When analyzing PCBP3, which has a predicted isoelectric point of pI = 8.22 (PhospoSitePlus), we found that these proteins in HFF and iPS cells are considerably more acidic than expected. However, few differences were noted when comparing spot profiles between cell lines.

PCBP4 has a predicted pI = 8.42 while that predicted for PCBP4C is pI = 6.55 (ExPASy; http://web.expasy.org/compute_pi). When both PCBP4 protein isoforms were analyzed, we found obvious differences in foci patterns. For PCBP4, a major species was evident near pI = 10. Additionally, comparing the fainter spots near the middle of the gradient, we saw that this protein is slightly more acidic in HFFs. The profiles of PCBP4C proteins also showed slight variability. Here, two spots indicated two predominant protein statuses with more PCBP4C proteins enriched at the more acidic focus. In iPS cells, the more basic focus was slightly more pronounced. Slight dissimilarities in GAPDH profiles were seen throughout 2D analysis of PCBP proteins which may be explained by sample and experiment variability (Appendix A4).

Although these analyses were not performed in triplicate, the slight differences in pI demonstrated by PCBPs may be a result of differences in post-translational modifications that occur in HFF and iPS cells to modulate the specific function of these proteins. Further analysis using a narrower pH gradient and treatment with phosphatases would be necessary to better characterize these pI changes. Although the analysis presented here does not provide an in-depth characterization of PCBP regulation in differentiated and pluripotent stem cells, it does support the possibility that differences in mRNA stability of transcripts with 3'UTR C-rich elements may be modulated through differential regulation of the PCBP family. However, a direct interaction of these proteins with the destabilized C-rich mRNAs must be established to confirm their involvement. Unfortunately, multiple attempts to

investigate RNA-binding of PCBPs in HFF and iPS cells were unsuccessful and could not be included in this study.

6.5 Discussion

Our global analysis of mRNA decay rates revealed that C-rich hexamer sequences were over-represented in the 3'UTR of mRNAs destabilized in iPS cells compared to HFFs. This was intriguing as global analyses have not previously implicated C-rich *cis*-acting elements in destabilization (Dolken et al., 2008; Friedel et al., 2009; Lam et al., 2001; Lee et al., 2010; Raghavan et al., 2002; Schwanhäusser et al., 2011, 2013; Sharova et al., 2009; Thomsen et al., 2010; Vlasova et al., 2008; Yang et al., 2003), although their involvement in mRNA stabilization is well-documented (Lindquist et al., 2000; Wang et al., 1995; Weiss and Liebhaber, 1995) and suggests that this phenomenon is stem cell-specific. In support, GO analysis of these destabilized mRNAs determined that they encode proteins with functions in transcription and chromatin remodeling (Table 6.2), both of which are differentially regulated during pluripotency (Efroni et al., 2008; Gaspar-Maia et al., 2011; Lessard and Crabtree, 2010). A common *cis*-acting element within functionally related mRNAs could allow coordinated regulation of stability through one or few *trans*-acting factors. We surmised that these transcripts are regulated through Poly(C)-Binding Proteins and that these PCBPs are differentially expressed in iPS cells compared to HFFs to modulate cell-specific functions.

Of the four PCBP members, PCBP4, an anti-proliferative factor (Zhu and Chen, 2000), demonstrated the most dramatic differential expression in iPS cells relative to HFFs. We found that *PCBP4C* mRNAs are generated by alternative splicing more frequently in HFFs (Figure 6.8) which also leads to increased abundance of this protein isoform compared to iPS cells (Figure 6.4). These data are consistent with previous reports that *PCBP4* mRNA exhibits stem-cell specific splicing patterns (Salomonis et al., 2009) and suggest that PCBP4C has functions separate from those of the full-length

isoform. It is of note that the spliced region of *PCBP4C* mRNA is C-rich, possibly indicating a mechanism for auto-regulation. Although we did not determine which PCBP4 isoform is more abundant in each cell line, the overall up-regulation of PCBP4 mRNA and protein levels in HFFs was intriguing. Regardless of whether PCBP4 regulates the mRNAs identified in our dataset of destabilized transcripts, down-regulation of this tumor suppressor could be important for establishing or maintaining pluripotency.

PCBP3 also demonstrated differential expression with increased protein and mRNA abundances in iPS cells compared to HFFs (Figures 6.4 and 6.5). As PCBP1 and PCBP2 levels were similar between the cell lines, the inversely regulated expression of PCBP3 and PCBP4 hints at a mechanism by which C-rich mRNAs are differentially stabilized. It is possible that these proteins target the same mRNAs with PCBP3 conferring mRNA destabilization and PCBP4 conferring stabilization. Alternatively, both proteins may act to stabilize transcripts wherein PCBP4 has greater stabilizing effects. Further experiments are necessary to determine whether the mRNAs bearing 3'UTR C-rich elements destabilized in iPS cells are directly associated with these proteins.

Post-translational modification of PCBPs also modulates their RNA-binding capabilities, as demonstrated by PCBP1 under hypoxic conditions (Ho et al., 2013). Here, phosphorylation of PCBP1 causes the RBP to release its bound mRNA, thereby exposing the transcript to miRNAs or other decay factors leading to degradation. Our 2D gel electrophoresis analysis indicated that PCBP1 has a decreased pI in iPS cells compared to HFFs (Figure 6.10), consistent with increased phosphorylation. This correlated well with our data in that phosphorylation of PCBP1 in iPS cells would prevent the protein from binding and stabilizing C-rich targets, resulting in destabilization relative to HFFs. Thus, although PCBP1 abundance is similar in both cell types, it may have altered function that contributes to differential regulation of mRNA stability.

Overall, we have shown that C-rich hexamer elements are over-represented in the 3'UTR of mRNAs destabilized in iPS cells and these transcripts encode proteins important for developmental

regulation of gene expression. Although additional investigation is required to elucidate the role of PCBPs in pluripotency, we have shown conclusively that these proteins are differentially expressed in HFF and iPS cells. As PCBP3 and PCBP4 are not well characterized, this study has provided additional insight to their expression and warrants more detailed follow-up analyses.

Concluding remarks

The goal of this project was to further our limited understanding of mRNA decay mechanisms that are essential for pluripotency. Although it was clear that post-transcriptional regulation would have an influence on stem cell gene expression, since it is a determinant of mRNA levels in all cell types, the nature of these influences are largely uncharacterized. Here, we showed that mRNA stability is differentially regulated for three important classes of mRNAs in iPS cells to support defining characteristics of pluripotent gene expression.

Marked stabilization of replication-dependent histone mRNAs was demonstrated for nearly all of the histone transcripts that half-lives were determined for in both HFF and iPS cells, and this resulted in increased levels of histone proteins. Genome hyperdynamics are required for stem cells to retain the self-renewal capability and histones play a fundamental role in establishing chromatin structure. We found that differential mRNA stability is a mechanism used in iPS cells, likely to support the histone demands of chromatin plasticity and this highlights the significance of regulated mRNA decay in establishing cell-specific expression patterns. To further elucidate differences in histone expression, future experiments should include determining where excess histone mRNAs and proteins are stored in iPS cells, since there must be mechanisms of bulk histone storage in place. It will also be important to perform a more detailed analysis of possible *cis*-elements within histone mRNAs in addition to their 3' stem loop that facilitate their differential regulation. This could also allow for the identification of other important factors involved in histone mRNA and protein metabolism.

We also found evidence for the contribution of mRNA decay to pluripotency in the pronounced stabilization of mRNAs encoding C2H2 ZNF transcription factors, many of which included the repressor KRAB domain. The expression of these transcription factors is developmentally regulated where abundances are increased during early embryogenesis. Our analysis demonstrated that this class of

mRNAs is coordinately regulated at the post-transcriptional level and it is possible that the longer C2H2 ZNF mRNA half-lives in iPS cells is needed for their increased protein abundances characteristic of early embryogenesis. Future efforts to better understand C2H2 ZNF expression in HFF and iPS cells should likely include determination of the length of poly(A) tails of these mRNAs, as our model merely suggests that the poly(A) tails would be shorter in HFFs. It will also be important to determine whether or not the stabilized C2H2 ZNF mRNAs are directly targeted by ORF-targeting miRNAs and whether stabilization of C2H2 ZNF mRNAs also leads to increased protein abundance.

As we hoped to identify novel regulatory mechanisms specific to pluripotency, we were intrigued to find C-rich elements over-represented in the 3'UTR of mRNAs destabilized in iPS cells. Although we were not able to directly demonstrate the functional importance of these sequence elements, we explored the PCBP family of proteins as possible stability factors that regulate their differential decay in iPS cells compared to HFFs. Further characterization of PCBPs in HFF and iPS cells should include determination of the mRNAs that are targeted by each and also determining whether PCBPs interact with other effector proteins to carry out additional roles important for gene expression. Future analyses remain to determine their exact roles in stem cell gene expression but our results revealed that PCBPs are differentially expressed in the two cell types, possibly contributing to destabilization of iPS cell mRNAs with C-rich 3'UTRs.

The data presented in this dissertation emphasize the importance of post-transcriptional control in pluripotency and provide a foundation for further characterization of the mechanisms involved. Although largely overlooked, regulated mRNA decay is clearly utilized during pluripotency to help maintain stem cell-specific gene expression patterns and identification of the factors involved could indicate targets for increasing reprogramming efficiency. Overall, our dataset and the subsequent analyses described here strengthen our understanding of pluripotent gene expression and the role of mRNA decay.

References

Adryan, B., and Teichmann, S.A. (2010). The developmental expression dynamics of *Drosophila melanogaster* transcription factors. Genome Biol. *11*, R40.

Alberts B., Bray D., Lewis J., Raff, M., Roberts, K., and Watson, J.D. (1994). Molecular Biology of the Cell. 3rd edition. New York: Garland Science. Available from: http://www.ncbi.nlm.nih.gov/books/NBK20684/

Alisch, R.S., Jin, P., Epstein, M., Caspary, T., and Warren, S.T. (2007). Argonaute2 is essential for mammalian gastrulation and proper mesoderm formation. PLoS Genet. *3*, e227.

Allard, P., Yang, Q., Marzluff, W.F., and Clarke, H.J. (2005). The stem-loop binding protein regulates translation of histone mRNA during mammalian oogenesis. Dev. Biol. *286*, 195–206.

Anderson, P. (2008). Post-transcriptional control of cytokine production. Nat. Immunol. *9*, 353–359.

Araki, Y., Takahashi, S., Kobayashi, T., Kajiho, H., Hoshino, S., and Katada, T. (2001). Ski7p G protein interacts with the exosome and the Ski complex for 3'-to-5' mRNA decay in yeast. EMBO J. *20*, 4684–4693.

Baggs, J.E., and Green, C.B. (2003). Nocturnin, a deadenylase in *Xenopus laevis* retina: a mechanism for posttranscriptional control of circadian-related mRNA. Curr. Biol. *13*, 189–198.

Bain, G., Kitchens, D., Yao, M., Huettner, J.E., and Gottlieb, D.I. (1995). Embryonic stem cells express neuronal properties in vitro. Dev. Biol. *168*, 342–357.

Van Bakel, H., Tsui, K., Gebbia, M., Mnaimneh, S., Hughes, T.R., and Nislow, C. (2013). A Compendium of Nucleosome and Transcript Profiles Reveals Determinants of Chromatin Architecture and Transcription. PLoS Genet. *9*, e1003479.

Bansal, N., Zhang, M., Bhaskar, A., Itotia, P., Lee, E., Shlyakhtenko, L.S., Lam, T.T., Fritz, A., Berezney, R., Lyubchenko, Y.L., Stafford, W.F., and Thapar, R. (2013). Assembly of the SLIP1-SLBP complex on histone mRNA requires heterodimerization and sequential binding of SLBP followed by SLIP1. Biochemistry. *52*, 520–536.

Bao, X., Zhu, X., Liao, B., Benda, C., Zhuang, Q., Pei, D., Qin, B., and Esteban, M.A. (2013). MicroRNAs in somatic cell reprogramming. Curr. Opin. Cell Biol. *25*, 208–214.

Barcaroli, D., Bongiorno-Borbone, L., Terrinoni, A., Hofmann, T.G., Rossi, M., Knight, R.A., Matera, A.G., Melino, G., and Laurenzi, V.D. (2006). FLASH is required for histone transcription and S-phase progression. Proc. Natl. Acad. Sci. *103*, 14808–14812.

Bardwell, V.J., Zarkower, D., Edmonds, M., and Wickens, M. (1990). The enzyme that adds poly(A) to mRNAs is a classical poly(A) polymerase. Mol. Cell. Biol. *10*, 846–849.

Bashkirov, V.I., Scherthan, H., Solinger, J.A., Buerstedde, J.M., and Heyer, W.D. (1997). A mouse cytoplasmic exoribonuclease (mXRN1p) with preference for G4 tetraplex substrates. J. Cell Biol. *136*, 761–773.

Batova, I., and O'Rand, M.G. (1996). Histone-binding domains in a human nuclear autoantigenic sperm protein. Biol. Reprod. *54*, 1238–1244.

Becker, K.A., Ghule, P.N., Therrien, J.A., Lian, J.B., Stein, J.L., van Wijnen, A.J., and Stein, G.S. (2006). Self-renewal of human embryonic stem cells is supported by a shortened G1 cell cycle phase. J. Cell. Physiol. *209*, 883–893.

Beese, L.S., and Steitz, T.A. (1991). Structural basis for the 3'-5' exonuclease activity of *Escherichia coli* DNA polymerase I: a two metal ion mechanism. EMBO J. *10*, 25–33.

Bellefroid, E.J., Poncelet, D.A., Lecocq, P.J., Revelant, O., and Martial, J.A. (1991). The evolutionarily conserved Krüppel-associated box domain defines a subfamily of eukaryotic multifingered proteins. Proc. Natl. Acad. Sci. *88*, 3608–3612.

Bergman, N., Moraes, K.C.M., Anderson, J.R., Zaric, B., Kambach, C., Schneider, R.J., Wilusz, C.J., and Wilusz, J. (2007). Lsm proteins bind and stabilize RNAs containing 5' poly(A) tracts. Nat. Struct. Mol. Biol. *14*, 824–831.

Bernstein, J.A., Khodursky, A.B., Lin, P.-H., Lin-Chao, S., and Cohen, S.N. (2002). Global analysis of mRNA decay and abundance in *Escherichia coli* at single-gene resolution using two-color fluorescent DNA microarrays. Proc. Natl. Acad. Sci. *99*, 9697–9702.

Berry, A.M., Flock, K.E., Loh, H.H., and Ko, J.L. (2006). Molecular basis of cellular localization of poly C binding protein 1 in neuronal cells. Biochem. Biophys. Res. Commun. *349*, 1378–1386.

Bianchin, C., Mauxion, F., Sentis, S., Séraphin, B., and Corbo, L. (2005). Conservation of the deadenylase activity of proteins of the Caf1 family in human. RNA. *11*, 487–494.

Bienz, M., Kubli, E., Kohli, J., deHenau, S., Huez, G., Marbaix, G., and Grosjean, H. (1981). Usage of the three termination codons in a single eukaryotic cell, the *Xenopus laevis* oocyte. Nucleic Acids Res. *9*, 3835–3850.

Bilger, A., Fox, C.A., Wahle, E., and Wickens, M. (1994). Nuclear polyadenylation factors recognize cytoplasmic polyadenylation elements. Genes Dev. *8*, 1106–1116.

Bilic, J., and Belmonte, J.C.I. (2012). Concise Review: Induced Pluripotent Stem Cells Versus Embryonic Stem Cells: Close Enough or Yet Too Far Apart? Stem Cells. *30*, 33–41.

Binda, O., Sevilla, A., LeRoy, G., Lemischka, I.R., Garcia, B.A., and Richard, S. (2013). SETD6 monomethylates H2AZ on lysine 7 and is required for the maintenance of embryonic stem cell self-renewal. Epigenetics Off. J. DNA Methylation Soc. *8*, 177–183.

Binder, H., Steiner, L., Przybilla, J., Rohlf, T., Prohaska, S., and Galle, J. (2013). Transcriptional regulation by histone modifications: towards a theory of chromatin re-organization during stem cell differentiation. Phys. Biol. *10*, 026006.

Blech-Hermoni, Y., Stillwagon, S.J., and Ladd, A.N. (2013). Diversity and conservation of CELF1 and CELF2 RNA and protein expression patterns during embryonic development. Dev. Dyn. *242*, 767–777.

Bleyman, M., and Woese, C. (1969). Transcriptional mapping. Introduction to the method and the use of actinomycin D as a transcriptional mapping agent. Proc. Natl. Acad. Sci. *63*, 532–539.

Boado, R.J., and Pardridge, W.M. (2002). Glucose deprivation and hypoxia increase the expression of the GLUT1 glucose transporter via a specific mRNA cis-acting regulatory element. J. Neurochem. *80*, 552–554.

Boeck, R., Tarun, S., Jr, Rieger, M., Deardorff, J.A., Müller-Auer, S., and Sachs, A.B. (1996). The yeast Pan2 protein is required for poly(A)-binding protein-stimulated poly(A)-nuclease activity. J. Biol. Chem. *271*, 432–438.

Bogenberger, J.M., and Laybourn, P.J. (2008). Human T Lymphotropic Virus Type 1 protein Tax reduces histone levels. Retrovirology. *5*, 9.

Bond, U.M., Yario, T.A., and Steitz, J.A. (1991). Multiple processing-defective mutations in a mammalian histone pre-mRNA are suppressed by compensatory changes in U7 RNA both in vivo and in vitro. Genes Dev. *5*, 1709–1722.

Boyer, L.A., and Sha, K. (2009). The chromatin signature of pluripotent cells. StemBook. Available http://www.stembook.org/

Bracken, C.P., Szubert, J.M., Mercer, T.R., Dinger, M.E., Thomson, D.W., Mattick, J.S., Michael, M.Z., and Goodall, G.J. (2011). Global analysis of the mammalian RNA degradome reveals widespread miRNA-dependent and miRNA-independent endonucleolytic cleavage. Nucleic Acids Res. *39*, 5658–5668.

Braun, J.E., Huntzinger, E., Fauser, M., and Izaurralde, E. (2011). GW182 proteins directly recruit cytoplasmic deadenylase complexes to miRNA targets. Mol. Cell. *44*, 120–133.

Bremer, K.A., Stevens, A., and Schoenberg, D.R. (2003). An endonuclease activity similar to *Xenopus* PMR1 catalyzes the degradation of normal and nonsense-containing human beta-globin mRNA in erythroid cells. RNA. *9*, 1157–1167.

Brennan, C.M., and Steitz, J.A. (2001). HuR and mRNA stability. Cell. Mol. Life Sci. *58*, 266–277.

Brouwer, R., Allmang, C., Raijmakers, R., van Aarssen, Y., Egberts, W.V., Petfalski, E., van Venrooij, W.J., Tollervey, D., and Pruijn, G.J. (2001). Three novel components of the human exosome. J. Biol. Chem. *276*, 6177–6184.

Buckanovich, R.J., and Darnell, R.B. (1997). The neuronal RNA binding protein Nova-1 recognizes specific RNA targets in vitro and in vivo. Mol. Cell. Biol. *17*, 3194–3201.

Burch, B.D., Godfrey, A.C., Gasdaska, P.Y., Salzler, H.R., Duronio, R.J., Marzluff, W.F., and Dominski, Z. (2011). Interaction between FLASH and Lsm11 is essential for histone pre-mRNA processing in vivo in *Drosophila*. RNA. *17*, 1132–1147.

Cakmakci, N.G., Lerner, R.S., Wagner, E.J., Zheng, L., and Marzluff, W.F. (2008). SLIP1, a Factor Required for Activation of Histone mRNA Translation by the Stem-Loop Binding Protein. Mol. Cell. Biol. *28*, 1182–1194.

Cao, K., Lailler, N., Zhang, Y., Kumar, A., Uppal, K., Liu, Z., Lee, E.K., Wu, H., Medrzycki, M., Pan, C., Ho, P.Y., Cooper, G.P., Dong, X., Bock, C., Bouhassira, E.E., and Fan, Y. (2013). High-Resolution Mapping of H1 Linker Histone Variants in Embryonic Stem Cells. PLoS Genet. *9*. e1003417.

Carr, M.D., Pastore, A., Gausepohl, H., Frank, R., and Roesch, P. (1990). NMR and molecular dynamics studies of the mKr2 zinc finger. Eur. J. Biochem. FEBS. *188*, 455–461.

Chappell, J., Sun, Y., Singh, A., and Dalton, S. (2013). MYC/MAX control ERK signaling and pluripotency by regulation of dual-specificity phosphatases 2 and 7. Genes Dev. *27*, 725–733.

Chaudhury, A., Hussey, G.S., Ray, P.S., Jin, G., Fox, P.L., and Howe, P.H. (2010). TGF-beta-mediated phosphorylation of hnRNP E1 induces EMT via transcript-selective translational induction of Dab2 and ILEI. Nat. Cell Biol. *12*, 286–293.

Chen, C.Y., Gherzi, R., Ong, S.E., Chan, E.L., Raijmakers, R., Pruijn, G.J., Stoecklin, G., Moroni, C., Mann, M., and Karin, M. (2001). AU binding proteins recruit the exosome to degrade ARE-containing mRNAs. Cell. *107*, 451–464.

Chen, Y., Wang, Z., Xie, Y., Guo, X., Tang, X., Wang, S., Yang, S., Chen, K., Niu, Y., and Ji, W. (2012). Folic acid deficiency inhibits neural rosette formation and neuronal differentiation from rhesus monkey embryonic stem cells. J. Neurosci. Res. *90*, 1382–1391.

Chkheidze, A.N., and Liebhaber, S.A. (2003). A Novel Set of Nuclear Localization Signals Determine Distributions of the α-CP RNA-Binding Proteins. Mol. Cell. Biol. *23*, 8405–8415.

Chlebowski, A., Tomecki, R., López, M.E.G., Séraphin, B., and Dziembowski, A. (2011). Catalytic properties of the eukaryotic exosome. Adv. Exp. Med. Biol. *702*, 63–78.

Choi, H.S., Kim, C.S., Hwang, C.K., Song, K.Y., Law, P.-Y., Wei, L.-N., and Loh, H.H. (2007). Novel function of the poly(C)-binding protein αCP3 as a transcriptional repressor of the mu opioid receptor gene. FASEB J. *21*, 3963–3973.

Choi, H.S., Song, K.Y., Hwang, C.K., Kim, C.S., Law, P.-Y., Wei, L.-N., and Loh, H.H. (2008). A proteomics approach for identification of single strand DNA-binding proteins involved in transcriptional regulation of mouse mu opioid receptor gene. Mol. Cell. Proteomics. *7*, 1517–1529.

Choi, H.S., Hwang, C.K., Song, K.Y., Law, P.-Y., Wei, L.-N., and Loh, H.H. (2009). Poly(C)-binding Proteins as Transcriptional Regulators of Gene Expression. Biochem. Biophys. Res. Commun. *380*, 431–436.

Claffey, K.P., Shih, S.C., Mullen, A., Dziennis, S., Cusick, J.L., Abrams, K.R., Lee, S.W., and Detmar, M. (1998). Identification of a human VPF/VEGF 5'ntranslated region mediating hypoxia-induced mRNA stability. Mol. Biol. Cell. *9*, 469–481.

Cloke, B., Shah, K., Kaneda, H., Lavery, S., Trew, G., Fusi, L., Higham, J., Dina, R.E., Ghaem-Maghami, S., Ellis, P., Brosens, J.J., and Christian, M. (2010). The Poly(C)-Binding Protein-1 Regulates Expression of the Androgen Receptor. Endocrinology. *151*, 3954–3964.

De Cola, A., Bongiorno-Borbone, L., Bianchi, E., Barcaroli, D., Carletti, E., Knight, R.A., Di Ilio, C., Melino, G., Sette, C., and De Laurenzi, V. (2012). FLASH is essential during early embryogenesis and cooperates with p73 to regulate histone gene transcription. Oncogene. *31*, 573–582.

Coles, L.S., Bartley, M.A., Bert, A., Hunter, J., Polyak, S., Diamond, P., Vadas, M.A., and Goodall, G.J. (2004). A multi-protein complex containing cold shock domain (Y-box) and polypyrimidine tract binding proteins forms on the vascular endothelial growth factor mRNA. Eur. J. Biochem. *271*, 648–660.

Collart, M.A. (2003). Global control of gene expression in yeast by the Ccr4-Not complex. Gene. *313*, 1–16.

Collart, M.A., and Panasenko, O.O. (2012). The Ccr4--not complex. Gene. *492*, 42–53.

Coller, J., and Parker, R. (2004). Eukaryotic mRNA decapping. Annu. Rev. Biochem. *73*, 861–890.

Coller, J.M., Tucker, M., Sheth, U., Valencia-Sanchez, M.A., and Parker, R. (2001). The DEAD box helicase, Dhh1p, functions in mRNA decapping and interacts with both the decapping and deadenylase complexes. RNA. *7*, 1717–1727.

Collier, B., Goobar-Larsson, L., Sokolowski, M., and Schwartz, S. (1998). Translational inhibition in vitro of human papillomavirus type 16 L2 mRNA mediated through interaction with heterogenous ribonucleoprotein K and poly(rC)-binding proteins 1 and 2. J. Biol. Chem. *273*, 22648–22656.

Copeland, P.R., and Wormington, M. (2001). The mechanism and regulation of deadenylation: identification and characterization of *Xenopus* PARN. RNA. *7*, 875–886.

De la Cruz, B.J., Prieto, S., and Scheffler, I.E. (2002). The role of the 5' untranslated region (UTR) in glucose-dependent mRNA decay. Yeast. *19*, 887–902.

Czerniak, B., Herz, F., Wersto, R.P., and Koss, L.G. (1987). Expression of Ha-ras oncogene p21 protein in relation to the cell cycle of cultured human tumor cells. Am. J. Pathol. *126*, 411–416.

Czyzyk-Krzeska, M.F., and Bendixen, A.C. (1999). Identification of the Poly(C) Binding Protein in the Complex Associated With the 3' Untranslated Region of Erythropoietin Messenger RNA. Blood. *93*, 2111–2120.

Danielsen, M., Hinck, L., and Ringold, G.M. (1989). Two amino acids within the knuckle of the first zinc finger specify DNA response element activation by the glucocorticoid receptor. Cell. *57*, 1131–1138.

Daugeron, M.C., Mauxion, F., and Séraphin, B. (2001). The yeast POP2 gene encodes a nuclease involved in mRNA deadenylation. Nucleic Acids Res. *29*, 2448–2455.

Dean, J.L., Wait, R., Mahtani, K.R., Sully, G., Clark, A.R., and Saklatvala, J. (2001). The 3' untranslated region of tumor necrosis factor alpha mRNA is a target of the mRNA-stabilizing factor HuR. Mol. Cell. Biol. *21*, 721–730.

Delgado-Olguin, P., and Recillas-Targa, F. (2011). Chromatin structure of pluripotent stem cells and induced pluripotent stem cells. Briefings Funct. Genomics. *10*, 37–49.

Deshpande, G., Calhoun, G., and Schedl, P. (2005). *Drosophila* argonaute-2 is required early in embryogenesis for the assembly of centric/centromeric heterochromatin, nuclear division, nuclear migration, and germ-cell formation. Genes Dev. *19*, 1680–1685.

Dibbens, J.A., Miller, D.L., Damert, A., Risau, W., Vadas, M.A., and Goodall, G.J. (1999). Hypoxic regulation of vascular endothelial growth factor mRNA stability requires the cooperation of multiple RNA elements. Mol. Biol. Cell. *10*, 907–919.

van Dijk, E., Le Hir, H., and Séraphin, B. (2003). DcpS can act in the 5'-3' mRNA decay pathway in addition to the 3'-5' pathway. Proc. Natl. Acad. Sci. *100*, 12081–12086.

Ding, G., Lorenz, P., Kreutzer, M., Li, Y., and Thiesen, H.-J. (2009). SysZNF: the C2H2 zinc finger gene database. Nucleic Acids Res. *37*, D267–273.

Dolken, L., Ruzsics, Z., Radle, B., Friedel, C.C., Zimmer, R., Mages, J., Hoffmann, R., Dickinson, P., Forster, T., Ghazal, P., and Koszinowski, U.H. (2008). High-resolution gene expression profiling for simultaneous kinetic parameter analysis of RNA synthesis and decay. RNA. *14*, 1959–1972.

Doma, M.K., and Parker, R. (2006). Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation. Nature. *440*, 561–564.

Dominski, Z., and Marzluff, W.F. (1999). Formation of the 3' end of histone mRNA. Gene. *239*, 1–14.

Dominski, Z., Erkmann, J.A., Yang, X., Sanchez, R., and Marzluff, W.F. (2002). A novel zinc finger protein is associated with U7 snRNP and interacts with the stem-loop binding protein in the histone pre-mRNP to stimulate 3?-end processing. Genes Dev. *16*, 58–71.

Dominski, Z., Yang, X., Kaygun, H., Dadlez, M., and Marzluff, W.F. (2003). A 3' exonuclease that specifically interacts with the 3' end of histone mRNA. Mol. Cell. *12*, 295–305.

Dominski, Z., Yang, X., and Marzluff, W.F. (2005). The polyadenylation factor CPSF-73 is involved in histone-pre-mRNA processing. Cell. *123*, 37–48.

Dori-Bachash, M., Shalem, O., Manor, Y.S., Pilpel, Y., and Tirosh, I. (2012). Widespread promoter-mediated coordination of transcription and mRNA degradation. Genome Biol. *13*, R114.

Dovat, S., Ronni, T., Russell, D., Ferrini, R., Cobb, B.S., and Smale, S.T. (2002). A common mechanism for mitotic inactivation of C2H2 zinc finger DNA-binding domains. Genes Dev. *16*, 2985–2990.

Dowell, R.D. (2011). The similarity of gene expression between human and mouse tissues. Genome Biol. *12*, 101.

Dreesen, O., and Brivanlou, A.H. (2007). Signaling Pathways in Cancer and Embryonic Stem Cells. Stem Cell Rev. *3*, 7–17.

Dunckley, T., and Parker, R. (1999). The DCP2 protein is required for mRNA decapping in *Saccharomyces cerevisiae* and contains a functional MutT motif. EMBO J. *18*, 5411–5422.

Dupressoir, A., Morel, A.P., Barbot, W., Loireau, M.P., Corbo, L., and Heidmann, T. (2001). Identification of four families of yCCR4- and Mg2+-dependent endonuclease-related proteins in higher eukaryotes, and characterization of orthologs of yCCR4 with a conserved leucine-rich repeat essential for hCAF1/hPOP2 binding. BMC Genomics. *2*, 9.

Duursma, A.M., Kedde, M., Schrier, M., Sage, C. le, and Agami, R. (2008). miR-148 targets human DNMT3b protein coding region. RNA. *14*, 872–877.

Dziembowski, A., Lorentzen, E., Conti, E., and Séraphin, B. (2007). A single subunit, Dis3, is essentially responsible for yeast exosome core activity. Nat. Struct. Mol. Biol. *14*, 15–22.

Eberle, A.B., Lykke-Andersen, S., Mühlemann, O., and Jensen, T.H. (2009). SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells. Nat. Struct. Mol. Biol. *16*, 49–55.

Efroni, S., Duttagupta, R., Cheng, J., Dehghani, H., Hoeppner, D.J., Dash, C., Bazett-Jones, D.P., Le Grice, S., McKay, R.D.G., Buetow, K.H., Gingeras, T.R., Misteli, T., and Meshorer, E. (2008). Global transcription in pluripotent embryonic stem cells. Cell Stem Cell. *2*, 437–447.

Emerson, R.O., and Thomas, J.H. (2009). Adaptive evolution in zinc finger transcription factors. PLoS Genet. *5*, e1000325.

Etienne, W., Meyer, M.H., Peppers, J., and Meyer, R.A., Jr (2004). Comparison of mRNA gene expression by RT-PCR and DNA microarray. BioTechniques. *36*, 618–620, 622, 624–626.

Fabian, M.R., Cieplak, M.K., Frank, F., Morita, M., Green, J., Srikumar, T., Nagar, B., Yamamoto, T., Raught, B., Duchaine, T.F., and Sonenberg, N. (2011). miRNA-mediated deadenylation is orchestrated by GW182 through two conserved motifs that interact with CCR4-NOT. Nat. Struct. Mol. Biol. *18*, 1211–1217.

Fabian, M.R., Frank, F., Rouya, C., Siddiqui, N., Lai, W.S., Karetnikov, A., Blackshear, P.J., Nagar, B., and Sonenberg, N. (2013). Structural basis for the recruitment of the human CCR4-NOT deadenylase complex by tristetraprolin. Nat. Struct. Mol. Biol. *20,* 735-739.

Farrell, B.C., Power, E.M., and Mc Dermott, K.W. (2011). Developmentally regulated expression of Sox9 and microRNAs 124, 128 and 23 in neuroepithelial stem cells in the developing spinal cord. Int. J. Dev. Neurosci. Off. J. Int. Soc. Dev. Neurosci. *29*, 31–36.

Fidalgo, M., Shekar, P.C., Ang, Y.-S., Fujiwara, Y., Orkin, S.H., and Wang, J. (2011). Zfp281 functions as a transcriptional repressor for pluripotency of mouse embryonic stem cells. Stem Cells. *29*, 1705–1716.

Fischer, N., and Weis, K. (2002). The DEAD box protein Dhh1 stimulates the decapping enzyme Dcp1. EMBO J. *21*, 2788–2797.

Fisher, C.L., and Fisher, A.G. (2011). Chromatin states in pluripotent, differentiated, and reprogrammed cells. Curr. Opin. Genet. Dev. *21*, 140–146.

Fluckiger, A.-C., Marcy, G., Marchand, M., Négre, D., Cosset, F.-L., Mitalipov, S., Wolf, D., Savatier, P., and Dehay, C. (2006). Cell cycle features of primate embryonic stem cells. Stem Cells. *24*, 547–556.

Forman, J.J., Legesse-Miller, A., and Coller, H.A. (2008). A search for conserved sequences in coding regions reveals that the let-7 microRNA targets Dicer within its coding sequence. Proc. Natl. Acad. Sci. *105*, 14879–14884.

Foster, M.P., Wuttke, D.S., Radhakrishnan, I., Case, D.A., Gottesfeld, J.M., and Wright, P.E. (1997). Domain packing and dynamics in the DNA complex of the N-terminal zinc fingers of TFIIIA. Nat. Struct. Biol. *4*, 605–608.

Fred, R.G., Tillmar, L., and Welsh, N. (2006). The role of PTB in insulin mRNA stability control. Curr. Diabetes Rev. *2*, 363–366.

Friedel, C.C., and Dölken, L. (2009). Metabolic tagging and purification of nascent RNA: implications for transcriptomics. Mol. Biosyst. *5*, 1271–1278.

Friedel, C.C., Dölken, L., Ruzsics, Z., Koszinowski, U.H., and Zimmer, R. (2009). Conserved principles of mammalian transcriptional regulation revealed by RNA half-life. Nucleic Acids Res. *37*, e115.

Friedman, J.R., Fredericks, W.J., Jensen, D.E., Speicher, D.W., Huang, X.P., Neilson, E.G., and Rauscher, F.J., 3rd (1996). KAP-1, a novel corepressor for the highly conserved KRAB repression domain. Genes Dev. *10*, 2067–2078.

Frietze, S., Lan, X., Jin, V.X., and Farnham, P.J. (2010). Genomic Targets of the KRAB and SCAN Domain-containing Zinc Finger Protein 263. J. Biol. Chem. *285*, 1393–1403.

Funke, B., Zuleger, B., Benavente, R., Schuster, T., Goller, M., Stévenin, J., and Horak, I. (1996). The mouse poly(C)-binding protein exists in multiple isoforms and interacts with several RNA-binding proteins. Nucleic Acids Res. *24*, 3821–3828.

Gallie, D.R. (1991). The cap and poly(A) tail function synergistically to regulate mRNA translational efficiency. Genes Dev. *5*, 2108–2116.

Gamarnik, A.V., and Andino, R. (2000). Interactions of viral protein 3CD and poly(rC) binding protein with the 5' untranslated region of the poliovirus genome. J. Virol. *74*, 2219–2226.

Gaspar-Maia, A., Alajem, A., Meshorer, E., and Ramalho-Santos, M. (2011). Open chromatin in pluripotency and reprogramming. Nat. Rev. Mol. Cell Biol. *12*, 36–47.

Gaspar-Maia, A., Qadeer, Z.A., Hasson, D., Ratnakumar, K., Leu, N.A., Leroy, G., Liu, S., Costanzi, C., Valle-Garcia, D., Schaniel, C., Lemischka, I., Garcia, B., and Pehrson, J.R. (2013). MacroH2A histone variants act as a barrier upon reprogramming towards pluripotency. Nat. Commun. *4*, 1565.

Georgiev, O., and Birnstiel, M.L. (1985). The conserved CAAGAAAGA spacer sequence is an essential element for the formation of 3' termini of the sea urchin H3 histone mRNA by RNA processing. EMBO J. *4*, 481–489.

Gerhold, D.L., Liu, F., Jiang, G., Li, Z., Xu, J., Lu, M., Sachs, J.R., Bagchi, A., Fridman, A., Holder, D.J., Doebber, T.W., Berger, J., Elbrecht, A., Moller, D.E., and Zhang, B.B. (2002). Gene expression profile of adipocyte differentiation and its regulation by peroxisome proliferator-activated receptor-gamma agonists. Endocrinology. *143*, 2106–2118.

Gherzi, R., Lee, K.-Y., Briata, P., Wegmüller, D., Moroni, C., Karin, M., and Chen, C.-Y. (2004). A KH domain RNA binding protein, KSRP, promotes ARE-directed mRNA turnover by recruiting the degradation machinery. Mol. Cell. *14*, 571–583.

Gherzi, R., Chen, C.-Y., Trabucchi, M., Ramos, A., and Briata, P. (2010). The role of KSRP in mRNA decay and microRNA precursor maturation. Wiley Interdiscip. Rev. RNA. *1*, 230–239.

Ghosh, D. (2000). High throughput and global approaches to gene expression. Comb. Chem. High Throughput Screen. *3*, 411–420.

Ghule, P.N., Dominski, Z., Yang, X., Marzluff, W.F., Becker, K.A., Harper, J.W., Lian, J.B., Stein, J.L., Wijnen, A.J. van, and Stein, G.S. (2008). Staged assembly of histone gene expression machinery at subnuclear foci in the abbreviated cell cycle of human embryonic stem cells. Proc. Natl. Acad. Sci. *105*, 16964–16969.

Gilbert, W.V., Zhou, K., Butler, T.K., and Doudna, J.A. (2007). Cap-independent translation is required for starvation-induced differentiation in yeast. Science. *317*, 1224–1227.

Gingerich, T.J., Feige, J.-J., and LaMarre, J. (2004). AU-rich elements and the control of gene expression through regulated mRNA stability. Anim. Heal. Res. Rev. *5*, 49–63.

Git, A., Dvinge, H., Salmon-Divon, M., Osborne, M., Kutter, C., Hadfield, J., Bertone, P., and Caldas, C. (2010). Systematic comparison of microarray profiling, real-time PCR, and next-generation sequencing technologies for measuring differential microRNA expression. RNA. *16*, 991–1006.

Godfrey, A.C., White, A.E., Tatomer, D.C., Marzluff, W.F., and Duronio, R.J. (2009). The *Drosophila* U7 snRNP proteins Lsm10 and Lsm11 are required for histone pre-mRNA processing and play an essential role in development. RNA. *15*, 1661–1672.

González-Lamuño, D., Loukili, N., García-Fuentes, M., and Thomson, T.M. (2002). Expression and regulation of the transcriptional repressor ZNF43 in Ewing sarcoma cells. Pediatr. Pathol. Mol. Med. *21*, 531–540.

Gratacós, F.M., and Brewer, G. (2010). The role of AUF1 in regulated mRNA decay. Wiley Interdiscip. Rev. RNA. *1*, 457–473.

Greber, B., Lehrach, H., and Adjaye, J. (2007). Silencing of core transcription factors in human EC cells highlights the importance of autocrine FGF signaling for self-renewal. BMC Dev. Biol. *7*, 46.

Gregory, R.I., Yan, K.-P., Amuthan, G., Chendrimada, T., Doratotaj, B., Cooch, N., and Shiekhattar, R. (2004). The Microprocessor complex mediates the genesis of microRNAs. Nature. *432*, 235–240.

Grewal, S.I.S., and Moazed, D. (2003). Heterochromatin and Epigenetic Control of Gene Expression. Science. *301*, 798–802.

Grimwood, J., Gordon, L.A., Olsen ,A., Terry, A., Schmutz, J., Lamerdin, J., Hellsten, U., Goodstein, D., Couronne, et al. (2004). The DNA sequence and biology of human chromosome 19. Nature. *428*, 529–535.

Grishin, N.V. (2001). Treble clef finger--a functionally diverse zinc-binding structural motif. Nucleic Acids Res. *29*, 1703–1714.

Groner, A.C., Meylan, S., Ciuffi, A., Zangger, N., Ambrosini, G., Dénervaud, N., Bucher, P., and Trono, D. (2010). KRAB-zinc finger proteins and KAP1 can mediate long-range transcriptional repression through heterochromatin spreading. PLoS Genet. *6*, e1000869.

Grönke, S., Bickmeyer, I., Wunderlich, R., Jäckle, H., and Kühnlein, R.P. (2009). Curled encodes the *Drosophila* homolog of the vertebrate circadian deadenylase Nocturnin. Genetics. *183*, 219–232.

Grzybowska, E.A. (2012). Human intronless genes: functional groups, associated diseases, evolution, and mRNA processing in absence of splicing. Biochem. Biophys. Res. Commun. *424*, 1–6.

Gueydan, C., Droogmans, L., Chalon, P., Huez, G., Caput, D., and Kruys, V. (1999). Identification of TIAR as a protein binding to the translational regulatory AU-rich element of tumor necrosis factor alpha mRNA. J. Biol. Chem. *274*, 2322–2326.

Guibert, S., and Weber, M. (2013). Functions of DNA methylation and hydroxymethylation in Mammalian development. Curr. Top. Dev. Biol. *104*, 47–83.

Gunjan, A., and Verreault, A. (2003). A Rad53 kinase-dependent surveillance mechanism that regulates histone protein levels in *S. cerevisiae*. Cell. *115*, 537–549.

Gunjan, A., Alexander, B.T., Sittman, D.B., and Brown, D.T. (1999). Effects of H1 histone variant overexpression on chromatin structure. J. Biol. Chem. *274*, 37950–37956.

De Guzman, R.N., Liu, H.Y., Martinez-Yamout, M., Dyson, H.J., and Wright, P.E. (2000). Solution structure of the TAZ2 (CH3) domain of the transcriptional adaptor protein CBP. J. Mol. Biol. *303*, 243–253.

Haimovich, G., Medina, D.A., Causse, S.Z., Garber, M., Millán-Zambrano, G., Barkai, O., Chávez, S., Pérez-Ortín, J.E., Darzacq, X., and Choder, M. (2013). Gene Expression Is Circular: Factors for mRNA Degradation Also Foster mRNA Synthesis. Cell. *153*, 1000–1011.

Halees, A.S., Hitti, E., Al-Saif, M., Mahmoud, L., Vlasova-St Louis, I.A., Beisang, D.J., Bohjanen, P.R., and Khabar, K. (2011). Global assessment of GU-rich regulatory content and function in the human transcriptome. RNA Biol. *8*, 681–691.

Hammond, S.M., Boettcher, S., Caudy, A.A., Kobayashi, R., and Hannon, G.J. (2001). Argonaute2, a link between genetic and biochemical analyses of RNAi. Science. *293*, 1146–1150.

Han, H., Irimia, M., Ross, P.J., Sung, H.-K., Alipanahi, B., David, L., Golipour, A., Gabut, M., Michael, I.P., Nachman, E.N., Wang, E., Trcka, D., Thompson, T., O'Hanlon, D., Slobodeniuc, V., Barbosa-Morias, N.L., Burge, C.B., Moffat, J., Frey, B.J., Nagy, A., Ellis, J., Wrana, J.L., and Blencowe, B.J. (2013). MBNL proteins repress ES-cell-specific alternative splicing and reprogramming. Nature. *498*, 241-245.

Han, J., Lee, Y., Yeom, K.-H., Kim, Y.-K., Jin, H., and Kim, V.N. (2004). The Drosha-DGCR8 complex in primary microRNA processing. Genes Dev. *18*, 3016–3027.

Happel, N., and Doenecke, D. (2009). Histone H1 and its isoforms: Contribution to chromatin structure and function. Gene. *431*, 1–12.

Harigaya, Y., and Parker, R. (2010). No-go decay: a quality control mechanism for RNA in translation. Wiley Interdiscip. Rev. RNA. *1*, 132–141.

Harigaya, Y., Jones, B.N., Muhlrad, D., Gross, J.D., and Parker, R. (2010). Identification and analysis of the interaction between Edc3 and Dcp2 in *Saccharomyces cerevisiae*. Mol. Cell. Biol. *30*, 1446–1456.

Harris, M.E., Bohni, R., Schneiderman, M.H., Ramamurthy, L., Schumperli, D., and Marzluff, W.F. (1991). Regulation of histone mRNA in the unperturbed cell cycle: evidence suggesting control at two posttranscriptional steps. Mol. Cell. Biol. *11*, 2416–2424.

Heintz, N., Sive, H.L., and Roeder, R.G. (1983). Regulation of human histone gene expression: kinetics of accumulation and changes in the rate of synthesis and in the half-lives of individual histone mRNAs during the HeLa cell cycle. Mol. Cell. Biol. *3*, 539–550.

Van Hemert, F.J., Jonk, L.J., and Destrée, O.H. (1992). Histone H1(0) mRNA and protein accumulate early during retinoic acid induced differentiation of synchronized embryonal carcinoma cells. Mol. Biol. Rep. *16*, 33–38.

Herrero, A.B., and Moreno, S. (2011). Lsm1 promotes genomic stability by controlling histone mRNA decay. EMBO J. *30*, 2008–2018.

Hindley, C., and Philpott, A. (2013). The cell cycle and pluripotency. Biochem. J. *451*, 135–143.

le Hir, H., Gatfield, D., Izaurralde, E., and Moore, M.J. (2001). The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. EMBO J. *20*, 4987–4997.

Ho, J.J.D., Robb, G.B., Tai, S.C., Turgeon, P.J., Mawji, I.A., Man, H.S.J., and Marsden, P.A. (2013). Active Stabilization of Human Endothelial Nitric Oxide Synthase mRNA by hnRNP E1 Protects against Antisense RNA and MicroRNAs. Mol. Cell. Biol. *33*, 2029–2046.

Hoefig, K.P., Rath, N., Heinz, G.A., Wolf, C., Dameris, J., Schepers, A., Kremmer, E., Ansel, K.M., and Heissmeyer, V. (2013). Eri1 degrades the stem-loop of oligouridylated histone mRNAs to induce replication-dependent decay. Nat. Struct. Mol. Biol. *20*, 73–81.

Holcik, M., and Liebhaber, S.A. (1997). Four highly stable eukaryotic mRNAs assemble 3' untranslated region RNA–protein complexes sharing *cis* and *trans* components. Proc. Natl. Acad. Sci. *94*, 2410–2414.

Holstege, F.C.P., Jennings, E.G., Wyrick, J.J., Lee, T.I., Hengartner, C.J., Green, M.R., Golub, T.R., Lander, E.S., and Young, R.A. (1998). Dissecting the Regulatory Circuitry of a Eukaryotic Genome. Cell. *95*, 717–728.

Hong, H., Takahashi, K., Ichisaka, T., Aoi, T., Kanagawa, O., Nakagawa, M., Okita, K., and Yamanaka, S. (2009). Suppression of Indued Pluripotent Stem Cell Generation by the p53-p21 Pathway. Nature. *460*, 1132–1135.

Horne, B.D., Carlquist, J.F., Muhlestein, J.B., Nicholas, Z.P., Anderson, J.L., and Intermountain Heart Collaborative Study Group (2007). Associations with myocardial infarction of six polymorphisms selected from a three-stage genome-wide association study. Am. Heart J. *154*, 969–975.

Houseley, J., LaCava, J., and Tollervey, D. (2006). RNA-quality control by the exosome. Nat. Rev. Mol. Cell Biol. *7*, 529–539.

Hsu, C.L., and Stevens, A. (1993). Yeast cells lacking 5'-->3' exoribonuclease 1 contain mRNA species that are poly(A) deficient and partially lack the 5' cap structure. Mol. Cell. Biol. *13*, 4826–4835.

Hu, G., Kim, J., Xu, Q., Leng, Y., Orkin, S.H., and Elledge, S.J. (2009). A genome-wide RNAi screen identifies a new transcriptional module required for self-renewal. Genes Dev. *23*, 837–848.

Hu, J., Lutz, C.S., Wilusz, J., and Tian, B. (2005). Bioinformatic identification of candidate cis-regulatory elements involved in human mRNA polyadenylation. RNA. *11*, 1485–1493.

Huang, L., and Wilkinson, M.F. (2012). Regulation of nonsense-mediated mRNA decay. Wiley Interdiscip. Rev. RNA. *3*, 807–828.

Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2009a). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. *4*, 44–57.

Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2009b). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. *37*, 1–13.

Huang, S., Wu, S., Ding, J., Lin, J., Wei, L., Gu, J., and He, X. (2010). MicroRNA-181a modulates gene expression of zinc finger family members by directly targeting their coding regions. Nucleic Acids Res. *38*, 7211–7218.

Hudson, B.P., Martinez-Yamout, M.A., Dyson, H.J., and Wright, P.E. (2004). Recognition of the mRNA AU-rich element by the zinc finger domain of TIS11d. Nat. Struct. Mol. Biol. *11*, 257–264.

Huntley, S., Baggott, D.M., Hamilton, A.T., Tran-Gyamfi, M., Yang, S., Kim, J., Gordon, L., Branscomb, E., and Stubbs, L. (2006). A comprehensive catalog of human KRAB-associated zinc finger genes: Insights into the evolutionary history of a large family of transcriptional repressors. Genome Res. *16*, 669–677.

Ikematsu, N., Yoshida, Y., Kawamura-Tsuzuku, J., Ohsugi, M., Onda, M., Hirai, M., Fujimoto, J., and Yamamoto, T. (1999). Tob2, a novel anti-proliferative Tob/BTG1 family member, associates with a component of the CCR4 transcriptional regulatory complex capable of binding cyclin-dependent kinases. Oncogene. *18*, 7432–7441.

Imamachi, N., Tani, H., Mizutani, R., Imamura, K., Irie, T., Suzuki, Y., and Akimitsu, N. (2013). BRIC-seq: A genome-wide approach for determining RNA stability in mammalian cells. Methods. S1046-2023(13)00263-6.

Imamura, M., Long, X., Nanda, V., and Miano, J.M. (2010). Expression and functional activity of four myocardin isoforms. Gene. *464*, 1–10.

Inoue, T., Sullivan, F.X., and Cech, T.R. (1986). New reactions of the ribosomal RNA precursor of Tetrahymena and the mechanism of self-splicing. J. Mol. Biol. *189*, 143–165.

Irie, S., and Sezaki, M. (1983). A quantitative determination of the relative amount of histones in polyacrylamide gel by silver stain. Anal. Biochem. *134*, 471–478.

Irion, S., Nostro, M.C., Kattman, S.J., and Keller, G.M. (2008). Directed differentiation of pluripotent stem cells: from developmental biology to therapeutic applications. Cold Spring Harb. Symp. Quant. Biol. *73*, 101–110.

Itokawa, Y., Yanagawa, T., Yamakawa, H., Watanabe, N., Koga, H., and Nagase, T. (2009). KAP1-independent transcriptional repression of SCAN-KRAB-containing zinc finger proteins. Biochem. Biophys. Res. Commun. *388*, 689–694.

Izaurralde, E., Lewis, J., McGuigan, C., Jankowska, M., Darzynkiewicz, E., and Mattaj, I.W. (1994). A nuclear cap binding protein complex involved in pre-mRNA splicing. Cell. *78*, 657–668.

Izaurralde, E., Lewis, J., Gamberi, C., Jarmolowski, A., McGuigan, C., and Mattaj, I.W. (1995). A cap-binding protein complex mediating U snRNA export. Nature. *376*, 709–712.

Jaeger, S., Martin, F., Rudinger-Thirion, J., Giegé, R., and Eriani, G. (2006). Binding of human SLBP on the 3'-UTR of histone precursor H4-12 mRNA induces structural rearrangements that enable U7 snRNA anchoring. Nucleic Acids Res. *34*, 4987–4995.

Jafarifar, F., Yao, P., Eswarappa, S.M., and Fox, P.L. (2011). Repression of VEGFA by CA-rich element-binding microRNAs is modulated by hnRNP L. EMBO J. *30*, 1324–1334.

Jenuwein, T., Laible, G., Dorn, R., and Reuter, G. (1998). SET domain proteins modulate chromatin domains in eu- and heterochromatin. Cell. Mol. Life Sci. *54*, 80–93.

Jeong, J., Adamson, L.K., Hatam, R., Greenhalgh, D.G., and Cho, K. (2003). Alterations in the expression and modification of histones in the liver after injury. Exp. Mol. Pathol. *75*, 256–264.

Jessen, B.A., and Stevens, G.J. (2002). Expression profiling during adipocyte differentiation of 3T3-L1 fibroblasts. Gene. *299*, 95–100.

Ji, X., Kong, J., and Liebhaber, S.A. (2003). In vivo association of the stability control protein alphaCP with actively translating mRNAs. Mol. Cell. Biol. *23*, 899–907.

Ji, X., Kong, J., and Liebhaber, S.A. (2011). An RNA-protein complex links enhanced nuclear 3' processing with cytoplasmic mRNA stabilization. EMBO J. *30*, 2622–2633.

Ji, X., Wan, J., Vishnu, M., Xing, Y., and Liebhaber, S.A. (2013). The Poly-C Binding Proteins, αCPs, Act as Global Regulators of Alternative Polyadenylation. Mol. Cell. Biol. *33,* 2560-2573.

Ji, Z., Lee, J.Y., Pan, Z., Jiang, B., and Tian, B. (2009). Progressive lengthening of 3' untranslated regions of mRNAs by alternative polyadenylation during mouse embryonic development. Proc. Natl. Acad. Sci. *106*, 7028–7033.

Jung, Y.-S., Qian, Y., and Chen, X. (2010). Examination of the expanding pathways for the regulation of p21 expression and activity. Cell. Signal. *22*, 1003–1012.

Kang, M.K., and Han, S.J. (2011). Post-transcriptional and post-translational regulation during mouse oocyte maturation. BMB Reports. *44*, 147–157.

Karginov, F.V., Cheloufi, S., Chong, M.M.W., Stark, A., Smith, A.D., and Hannon, G.J. (2010). Diverse endonucleolytic cleavage sites in the mammalian transcriptome depend upon microRNAs, Drosha, and additional nucleases. Mol. Cell. *38*, 781–788.

Kaspi, H., Chapnik, E., Levy, M., Beck, G., Hornstein, E., and Soen, Y. (2013). miR--290--295 Regulate Embryonic Stem Cell Differentiation Propensities by Repressing Pax6. Stem Cells. doi: 10.1002/stem.1465.

Kawasaki, H., and Taira, K. (2003). Hes1 is a target of microRNA-23 during retinoic-acid-induced neuronal differentiation of NT2 cells. Nature. *423*, 838–842.

Kaygun, H., and Marzluff, W.F. (2005a). Translation termination is involved in histone mRNA degradation when DNA replication is inhibited. Mol. Cell. Biol. *25*, 6879–6888.

Kaygun, H., and Marzluff, W.F. (2005b). Regulated degradation of replication-dependent histone mRNAs requires both ATR and Upf1. Nat. Struct. Mol. Biol. *12*, 794–800.

Kettenberger, H., Armache, K.-J., and Cramer, P. (2004). Complete RNA Polymerase II Elongation Complex Structure and Its Interactions with NTP and TFIIS. Mol. Cell. *16*, 955–965.

Khabar, K.S.A. (2005). The AU-rich transcriptome: more than interferons and cytokines, and its role in disease. J. Interf. Cytokine Res. *25*, 1–10.

Khanna, R., and Kiledjian, M. (2004). Poly(A)-binding-protein-mediated regulation of hDcp2 decapping in vitro. EMBO J. *23*, 1968–1976.

Kiledjian, M., Wang, X., and Liebhaber, S.A. (1995). Identification of two KH domain proteins in the alpha-globin mRNP stability complex. EMBO J. *14*, 4357–4364.

Kim, J.H., and Richter, J.D. (2006). Opposing polymerase-deadenylase activities regulate cytoplasmic polyadenylation. Mol. Cell. *24*, 173–183.

Kim, S.-S., Pandey, K.K., Choi, H.S., Kim, S.-Y., Law, P.-Y., Wei, L.-N., and Loh, H.H. (2005). Poly(C) binding protein family is a transcription factor in mu-opioid receptor gene expression. Mol. Pharmacol. *68*, 729–736.

Klauer, A.A., and van Hoof, A. (2012). Degradation of mRNAs that lack a stop codon: a decade of nonstop progress. Wiley Interdiscip. Rev. RNA. *3*, 649–660.

Knight, R.D., and Shimeld, S.M. (2001). Identification of conserved C2H2 zinc-finger gene families in the Bilateria. Genome Biol. *2*, research0016.1–research0016.8.

Koch, W., Hoppmann, P., Schömig, A., and Kastrati, A. (2011). Variations of specific non-candidate genes and risk of myocardial infarction: a replication study. Int. J. Cardiol. *147*, 38–41.

Koerber, R.T., Rhee, H.S., Jiang, C., and Pugh, B.F. (2009). Interaction of transcriptional regulators with specific nucleosomes across the *Saccharomyces* genome. Mol. Cell. *35*, 889–902.

Kong, J., Sumaroka, M., Eastmond, D.L., and Liebhaber, S.A. (2006). Shared stabilization functions of pyrimidine-rich determinants in the erythroid 15-lipoxygenase and alpha-globin mRNAs. Mol. Cell. Biol. *26*, 5603–5614.

Körner, C.G., and Wahle, E. (1997). Poly(A) tail shortening by a mammalian poly(A)-specific 3'-exoribonuclease. J. Biol. Chem. *272*, 10448–10456.

Koseoglu, M.M., Graves, L.M., and Marzluff, W.F. (2008). Phosphorylation of threonine 61 by cyclin a/Cdk1 triggers degradation of stem-loop binding protein at the end of S phase. Mol. Cell. Biol. *28*, 4469–4479.

Kosinski, P.A., Laughlin, J., Singh, K., and Covey, L.R. (2003). A complex containing polypyrimidine tract-binding protein is involved in regulating the stability of CD40 ligand (CD154) mRNA. J. Immunol. *170*, 979–988.

Krishnan, N., Lam, T.T., Fritz, A., Rempinski, D., O'Loughlin, K., Minderman, H., Berezney, R., Marzluff, W.F., and Thapar, R. (2012). The prolyl isomerase Pin1 targets stem-loop binding protein (SLBP) to dissociate the SLBP-histone mRNA complex linking histone mRNA decay with SLBP ubiquitination. Mol. Cell. Biol. *32*, 4306–4322.

Kudla, G., Lipinski, L., Caffin, F., Helwak, A., and Zylicz, M. (2006). High Guanine and Cytosine Content Increases mRNA Levels in Mammalian Cells. PLoS Biol. *4*, e180.

Kwon, S.C., Yi, H., Eichelbaum, K., Föhr, S., Fischer, B., You, K.T., Castello, A., Krijgsveld, J., Hentze, M.W., and Kim, V.N. (2013). The RNA-binding protein repertoire of embryonic stem cells. Nat. Struct. Mol. Biol. *20*, 1122-1130.

Lai, W.S., Parker, J.S., Grissom, S.F., Stumpo, D.J., and Blackshear, P.J. (2006). Novel mRNA targets for tristetraprolin (TTP) identified by global analysis of stabilized transcripts in TTP-deficient fibroblasts. Mol. Cell. Biol. *26*, 9196–9208.

Lakshmipathy, U., Love, B., Goff, L.A., Jörnsten, R., Graichen, R., Hart, R.P., and Chesnut, J.D. (2007). MicroRNA expression pattern of undifferentiated and differentiated human embryonic stem cells. Stem Cells Dev. *16*, 1003–1016.

Lam, L.T., Pickeral, O.K., Peng, A.C., Rosenwald, A., Hurt, E.M., Giltnane, J.M., Averett, L.M., Zhao, H., Davis, R.E., Sathyamoorthy, M., Wahl, L.M., Harris, E.D., Mikovits, J.A., Monks, A.P., Hollingshead, M.G., Sausville, E.A., and Staudt, L.M. (2001). Genomic-scale measurement of mRNA turnover and the mechanisms of action of the anti-cancer drug flavopiridol. Genome Biol. *2*, research0041.

Larson, J.L., and Yuan, G.-C. (2012). Chromatin states accurately classify cell differentiation stages. PloS One *7*, e31414.

Laskey, R.A., Mills, A.D., Philpott, A., Leno, G.H., Dilworth, S.M., and Dingwall, C. (1993). The role of nucleoplasmin in chromatin assembly and disassembly. Philos. Trans. R. Soc. Lond. B. Biol. Sci. *339*, 263–269.

Law, C., and Cheung, P. (2012). Histone variants and transcription regulation. Subcell. Biochem. *61*, 319–341.

Lebreton, A., Tomecki, R., Dziembowski, A., and Séraphin, B. (2008). Endonucleolytic RNA cleavage by a eukaryotic exosome. Nature. *456*, 993–996.

Lechner, M.S., Begg, G.E., Speicher, D.W., and Rauscher, F.J., 3rd (2000). Molecular determinants for targeting heterochromatin protein 1-mediated gene silencing: direct chromoshadow domain-KAP-1 corepressor interaction is essential. Mol. Cell. Biol. *20*, 6449–6465.

Lee, E.K., and Gorospe, M. (2011). Coding region: the neglected post-transcriptional code. RNA Biol. *8*, 44–48.

Lee, D.-H., Lim, M.-H., Youn, D.-Y., Jung, S.E., Ahn, Y.S., Tsujimoto, Y., and Lee, J.-H. (2009). hnRNP L binds to CA repeats in the 3'UTR of bcl-2 mRNA. Biochem. Biophys. Res. Commun. *382*, 583–587.

Lee, J.E., Lee, J.Y., Wilusz, J., Tian, B., and Wilusz, C.J. (2010). Systematic Analysis of Cis-Elements in Unstable mRNAs Demonstrates that CUGBP1 Is a Key Regulator of mRNA Decay in Muscle Cells. PLoS One. *5,* e11201.

Lee, J.E., Lee, J.Y., Trembly, J., Wilusz, J., Tian, B., and Wilusz, C.J. (2012). The PARN deadenylase targets a discrete set of mRNAs for decay and regulates cell motility in mouse myoblasts. PLoS Genet. *8*, e1002901.

Lee, J.-H., Jeon, M.-H., Seo, Y.-J., Lee, Y.-J., Ko, J.H., Tsujimoto, Y., and Lee, J.-H. (2004). CA Repeats in the 3'-Untranslated Region of bcl-2 mRNA Mediate Constitutive Decay of bcl-2 mRNA. J. Biol. Chem. *279*, 42758–42764.

Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Rådmark, O., Kim, S., and Kim, V.N. (2003). The nuclear RNase III Drosha initiates microRNA processing. Nature. *425*, 415–419.

Leffers, H., Dejgaard, K., and Celis, J.E. (1995). Characterisation of two major cellular poly(rC)-binding human proteins, each containing three K-homologous (KH) domains. Eur. J. Biochem. FEBS. *230*, 447–453.

Lennartsson, A., and Ekwall, K. (2009). Histone modification patterns and epigenetic codes. Biochim. Biophys. Acta. *1790*, 863–868.

Leonardo, T.R., Schultheisz, H.L., Loring, J.F., and Laurent, L.C. (2012). The functions of microRNAs in pluripotency and reprogramming. Nat. Cell Biol. *14*, 1114–1121.

Lessard, J.A., and Crabtree, G.R. (2010). Chromatin regulatory mechanisms in pluripotency. Annu. Rev. Cell Dev. Biol. *26*, 503–532.

Letunic, I., Doerks, T., and Bork, P. (2009). SMART 6: recent updates and new developments. Nucleic Acids Res. *37*, D229–232.

Letunic, I., Doerks, T., and Bork, P. (2012). SMART 7: recent updates to the protein domain annotation resource. Nucleic Acids Res. *40*, D302–305.

Levy, N.S., Chung, S., Furneaux, H., and Levy, A.P. (1998). Hypoxic Stabilization of Vascular Endothelial Growth Factor mRNA by the RNA-binding Protein HuR. J. Biol. Chem. *273*, 6417–6423.

Li, H., Chen, W., Zhou, Y., Abidi, P., Sharpe, O., Robinson, W.H., Kraemer, F.B., and Liu, J. (2009). Identification of mRNA binding proteins that regulate the stability of LDL receptor mRNA through AU-rich elements. J. Lipid Res. *50*, 820–831.

Li, X., Zhang, J., Gao, L., McClellan, S., Finan, M.A., Butler, T.W., Owen, L.B., Piazza, G.A., and Xi, Y. (2012). MiR-181 mediates cell differentiation by interrupting the Lin28 and let-7 feedback circuit. Cell Death Differ. *19*, 378–386.

Li, X.L., Blackford, J.A., and Hassel, B.A. (1998). RNase L mediates the antiviral effect of interferon through a selective reduction in viral RNA during encephalomyocarditis virus infection. J. Virol. *72*, 2752–2759.

Liang, M., Cowley, A.W., and Greene, A.S. (2004). High throughput gene expression profiling: a molecular approach to integrative physiology. J. Physiol. *554*, 22–30.

Liang, Y., Ridzon, D., Wong, L., and Chen, C. (2007). Characterization of microRNA expression profiles in normal human tissues. BMC Genomics. *8*, 166.

Lichtler, A.C., Sierra, F., Clark, S., Wells, J.R., Stein, J.L., and Stein, G.S. (1982). Multiple H4 histone mRNAs of HeLa cells are encoded in different genes. Nature. *298*, 195–198.

Lin, Q., Inselman, A., Han, X., Xu, H., Zhang, W., Handel, M.A., and Skoultchi, A.I. (2004). Reductions in linker histone levels are tolerated in developing spermatocytes but cause changes in specific gene expression. J. Biol. Chem. *279*, 23525–23535.

Lindquist, J.N., Kauschke, S.G., Stefanovic, B., Burchardt, E.R., and Brenner, D.A. (2000). Characterization of the interaction between alphaCP(2) and the 3'-untranslated region of collagen alpha1(I) mRNA. Nucleic Acids Res. *28*, 4306–4316.

Ling, J., Morley, S.J., Pain, V.M., Marzluff, W.F., and Gallie, D.R. (2002). The histone 3'-terminal stem-loop-binding protein enhances translation through a functional and physical interaction with eukaryotic initiation factor 4G (eIF4G) and eIF3. Mol. Cell. Biol. *22*, 7853–7867.

Ling, S.H.M., Qamra, R., and Song, H. (2011). Structural and functional insights into eukaryotic mRNA decapping. Wiley Interdiscip. Rev. RNA. *2*, 193–208.

Linz, B., Koloteva, N., Vasilescu, S., and McCarthy, J.E.G. (1997). Disruption of Ribosomal Scanning on the 5'-Untranslated Region, and Not Restriction of Translational Initiation *per se*, Modulates the Stability of Nonaberrant mRNAs in the Yeast *Saccharomyces cerevisiae*. J. Biol. Chem. *272*, 9131–9140.

Lister, R., Pelizzola, M., Dowen, R.H., Hawkins, R.D., Hon, G., Tonti-Filippini, J., Nery, J.R., Lee, L., Ye, Z., Ngo, Q.-M., Edsall, L., Antosiewicz-Bourget, J., Stewart, R., Ruotti, V., Millar, A.H., Thomson, J.A., Ren, B., Ecker, J.R. (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. Nature. *462*, 315–322.

Liu, H., Rodgers, N.D., Jiao, X., and Kiledjian, M. (2002). The scavenger mRNA decapping enzyme DcpS is a member of the HIT family of pyrophosphatases. EMBO J. *21*, 4699–4708.

Liu, J., Carmell, M.A., Rivas, F.V., Marsden, C.G., Thomson, J.M., Song, J.-J., Hammond, S.M., Joshua-Tor, L., and Hannon, G.J. (2004). Argonaute2 is the catalytic engine of mammalian RNAi. Science. *305*, 1437–1441.

Liu, J., Rivas, F.V., Wohlschlegel, J., Yates, J.R., 3rd, Parker, R., and Hannon, G.J. (2005). A role for the P-body component GW182 in microRNA function. Nat. Cell Biol. *7*, 1261–1266.

Liu, N., Han, H., and Lasko, P. (2009). Vasa promotes *Drosophila* germline stem cell differentiation by activating mei-P26 translation by directly interacting with a (U)-rich motif in its 3' UTR. Genes Dev. *23*, 2742–2752.

López de Silanes, I., Zhan, M., Lal, A., Yang, X., and Gorospe, M. (2004). Identification of a target RNA motif for RNA-binding protein HuR. Proc. Natl. Acad. Sci. *101*, 2987–2992.

Lorenz, P., Dietmann, S., Wilhelm, T., Koczan, D., Autran, S., Gad, S., Wen, G., Ding, G., Li, Y., Rousseau-Merck, M.-F., and Thiesen, H.J. (2010). The ancient mammalian KRAB zinc finger gene cluster on human chromosome 8q24.3 illustrates principles of C2H2 zinc finger evolution associated with unique expression profiles in human tissues. BMC Genomics. *11*, 206.

Loring, J.F., Porter, J.G., Seilhammer, J., Kaser, M.R., and Wesselschmidt, R. (2001). A gene expression profile of embryonic stem cells and embryonic stem cell-derived neurons. Restor. Neurol. Neurosci. *18*, 81–88.

Louie, E., Ott, J., and Majewski, J. (2003). Nucleotide Frequency Variation Across Human Genes. Genome Res. *13*, 2594–2601.

Lu, J., and Bushel, P.R. (2013). Dynamic expression of 3' UTRs revealed by Poisson hidden Markov modeling of RNA-Seq: Implications in gene expression profiling. Gene. *527*, 616–623.

Lu, K.P., Finn, G., Lee, T.H., and Nicholson, L.K. (2007). Prolyl cis-trans isomerization as a molecular timer. Nat. Chem. Biol. *3*, 619–629.

Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature. *389*, 251–260.

Lüningschrör, P., Stöcker, B., Kaltschmidt, B., and Kaltschmidt, C. (2012). miR-290 cluster modulates pluripotency by repressing canonical NF-κB signaling. Stem Cells. *30*, 655–664.

Lykke-Andersen, K., Gilchrist, M.J., Grabarek, J.B., Das, P., Miska, E., and Zernicka-Goetz, M. (2008). Maternal Argonaute 2 is essential for early mouse development at the maternal-zygotic transition. Mol. Biol. Cell. *19*, 4383–4392.

Ma, T., Tine, B.A.V., Wei, Y., Garrett, M.D., Nelson, D., Adams, P.D., Wang, J., Qin, J., Chow, L.T., and Harper, J.W. (2000). Cell cycle–regulated phosphorylation of p220NPAT by cyclin E/Cdk2 in Cajal bodies promotes histone gene transcription. Genes Dev. *14*, 2298–2313.

Machyna, M., Heyn, P., and Neugebauer, K.M. (2013). Cajal bodies: where form meets function. Wiley Interdiscip. Rev. RNA. *4*, 17–34.

Makeyev, A.V., and Liebhaber, S.A. (2000). Identification of two novel mammalian genes establishes a subfamily of KH-domain RNA-binding proteins. Genomics. *67*, 301–316.

Makeyev, A.V., and Liebhaber, S.A. (2002). The poly(C)-binding proteins: a multiplicity of functions and a search for mechanisms. RNA. *8*, 265–278.

Makino, D.L., Baumgärtner, M., and Conti, E. (2013). Crystal structure of an RNA-bound 11-subunit eukaryotic exosome complex. Nature. *495*, 70–75.

Malecki, M., Viegas, S.C., Carneiro, T., Golik, P., Dressaire, C., Ferreira, M.G., and Arraiano, C.M. (2013). The exoribonuclease Dis3L2 defines a novel eukaryotic RNA degradation pathway. EMBO J. *32*, 1842-1854.

Marino-Ramirez, L., Jordan, I.K., and Landsman, D. (2006). Multiple independent evolutionary solutions to core histone gene regulation. Genome Biol. *7*, R122.

Marqués-Torrejón, M.Á., Porlan, E., Banito, A., Gómez-Ibarlucea, E., Lopez-Contreras, A.J., Fernández-Capetillo, O., Vidal, A., Gil, J., Torres, J., and Fariñas, I. (2013). Cyclin-dependent kinase inhibitor p21 controls adult neural stem cell expansion by regulating Sox2 gene expression. Cell Stem Cell. *12*, 88–100.

Martemyanov, K.A., Krispel, C.M., Lishko, P.V., Burns, M.E., and Arshavsky, V.Y. (2008). Functional comparison of RGS9 splice isoforms in a living cell. Proc. Natl. Acad. Sci. *105*, 20988–20993.

Martin, F., Schaller, A., Eglite, S., Schümperli, D., and Müller, B. (1997). The gene for histone RNA hairpin binding protein is located on human chromosome 4 and encodes a novel type of RNA binding protein. EMBO J. *16*, 769–778.

Martin, F., Barends, S., Jaeger, S., Schaeffer, L., Prongidi-Fix, L., and Eriani, G. (2011). Cap-Assisted Internal Initiation of Translation of Histone H4. Mol. Cell. *41*, 197–209.

Marzluff, W.F., and Duronio, R.J. (2002). Histone mRNA expression: multiple levels of cell cycle regulation and important developmental consequences. Curr. Opin. Cell Biol. *14*, 692–699.

Marzluff, W.F., Gongidi, P., Woods, K.R., Jin, J., and Maltais, L.J. (2002). The Human and Mouse Replication-Dependent Histone Genes. Genomics. *80*, 487–498.

Marzluff, W.F., Wagner, E.J., and Duronio, R.J. (2008). Metabolism and regulation of canonical histone mRNAs: life without a poly(A) tail. Nat. Rev. Genet. *9*, 843–854.

Masuda, A., Andersen, H.S., Doktor, T.K., Okamoto, T., Ito, M., Andresen, B.S., and Ohno, K. (2012). CUGBP1 and MBNL1 preferentially bind to 3' UTRs and facilitate mRNA decay. Sci. Reports *2*. doi: 10.1038/srep00209

Matsui, T., Sasaki, A., Akazawa, N., Otani, H., and Bessho, Y. (2012). Celf1 regulation of dmrt2a is required for somite symmetry and left-right patterning during zebrafish development. Dev. *139*, 3553–3560.

Mattaj, I.W., and Englmeier, L. (1998). Nucleocytoplasmic transport: the soluble phase. Annu. Rev. Biochem. *67*, 265–306.

McCracken, S., Fong, N., Yankulov, K., Ballantyne, S., Pan, G., Greenblatt, J., Patterson, S.D., Wickens, M., and Bentley, D.L. (1997). The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. Nature. *385*, 357–361.

McGowan, C.H., and Russell, P. (1995). Cell cycle regulation of human WEE1. EMBO J. *14*, 2166–2175.

McPherson, J.P., Tamblyn, L., Elia, A., Migon, E., Shehabeldin, A., Matysiak-Zablocki, E., Lemmers, B., Salmena, L., Hakem, A., Fish, J., Kassam, F., Squire, J., Bruneau, B.G., Hande, M.P., and Hakem, R. (2004). Lats2/Kpm is required for embryonic development, proliferation control and genomic integrity. EMBO J. *23*, 3677–3688.

Medina, R., Buck, T., Zaidi, S.K., Miele-Chamberland, A., Lian, J.B., Stein, J.L., van Wijnen, A.J., and Stein, G.S. (2008). The histone gene cell cycle regulator HiNF-P is a unique zinc finger transcription factor with a novel conserved auxiliary DNA-binding motif. Biochemistry. *47*, 11415–11423.

Medina, R., Ghule, P.N., Cruzat, F., Barutcu, A.R., Montecino, M., Stein, J.L., van Wijnen, A.J., and Stein, G.S. (2012). Epigenetic control of cell cycle-dependent histone gene expression is a principal component of the abbreviated pluripotent cell cycle. Mol. Cell. Biol. *32*, 3860–3871.

Meeks-Wagner, D., and Hartwell, L.H. (1986). Normal stoichiometry of histone dimer sets is necessary for high fidelity of mitotic chromosome transmission. Cell. *44*, 43–52.

Meisner, N.-C., and Filipowicz, W. (2010). Properties of the regulatory RNA-binding protein HuR and its role in controlling miRNA repression. Adv. Exp. Med. Biol. *700*, 106–123.

Meister, G., Landthaler, M., Patkaniowska, A., Dorsett, Y., Teng, G., and Tuschl, T. (2004). Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. Mol. Cell. *15*, 185–197.

Meng, Q., Rayala, S.K., Gururaj, A.E., Talukder, A.H., O'Malley, B.W., and Kumar, R. (2007). Signaling-dependent and coordinated regulation of transcription, splicing, and translation resides in a single coregulator, PCBP1. Proc. Natl. Acad. Sci. *104*, 5866–5871.

Meshorer, E., and Misteli, T. (2006). Chromatin in pluripotent embryonic stem cells and differentiation. Nat. Rev. Mol. Cell Biol. *7*, 540–546.

Meshorer, E., Yellajoshula, D., George, E., Scambler, P.J., Brown, D.T., and Misteli, T. (2006). Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. Dev. Cell. *10*, 105–116.

Messina, D.N., Glasscock, J., Gish, W., and Lovett, M. (2004). An ORFeome-based analysis of human transcription factor genes and the construction of a microarray to interrogate their expression. Genome Res. *14*, 2041–2047.

Miele, A., Braastad, C.D., Holmes, W.F., Mitra, P., Medina, R., Xie, R., Zaidi, S.K., Ye, X., Wei, Y., Harper, J.W., van Wijnen, A.J., Stein, J.L., and Stein, G.S. (2005). HiNF-P directly links the cyclin E/CDK2/p220NPAT pathway to histone H4 gene regulation at the G1/S phase cell cycle transition. Mol. Cell. Biol. *25*, 6140–6153.

Millar, C.B. (2013). Organizing the genome with H2A histone variants. Biochem. J. *449*, 567–579.

Miller, M.A., and Olivas, W.M. (2011). Roles of Puf proteins in mRNA degradation and translation. Wiley Interdiscip. Rev. RNA. *2*, 471–492.

Miller, J., McLachlan, A.D., and Klug, A. (1985). Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus* oocytes. EMBO J. *4*, 1609–1614.

Milligan, L., Decourty, L., Saveanu, C., Rappsilber, J., Ceulemans, H., Jacquier, A., and Tollervey, D. (2008). A yeast exosome cofactor, Mpp6, functions in RNA surveillance and in the degradation of noncoding RNA transcripts. Mol. Cell. Biol. *28*, 5446–5457.

Mitchell, P., Petfalski, E., Shevchenko, A., Mann, M., and Tollervey, D. (1997). The exosome: a conserved eukaryotic RNA processing complex containing multiple 3'→5' exoribonucleases. Cell. *91*, 457–466.

Mittal, S., Aslam, A., Doidge, R., Medica, R., and Winkler, G.S. (2011). The Ccr4a (CNOT6) and Ccr4b (CNOT6L) deadenylase subunits of the human Ccr4-Not complex contribute to the prevention of cell death and senescence. Mol. Biol. Cell. *22*, 748–758.

Montell, C., Fisher, E.F., Caruthers, M.H., and Berk, A.J. (1983). Inhibition of RNA cleavage but not polyadenylation by a point mutation in mRNA 3' consensus sequence AAUAAA. Nature. *305*, 600–605.

Moraes, K.C.M., Wilusz, C.J., and Wilusz, J. (2006). CUG-BP binds to RNA substrates and recruits PARN deadenylase. RNA. *12*, 1084–1091.

Moran, J.L., Li, Y., Hill, A.A., Mounts, W.M., and Miller, C.P. (2002). Gene expression changes during mouse skeletal myoblast differentiation revealed by transcriptional profiling. Physiol. Genomics. *10*, 103–111.

Mott, N.N., and Pak, T.R. (2012). Characterisation of human oestrogen receptor beta (ERβ) splice variants in neuronal cells. J. Neuroendocrinol. *24*, 1311–1321.

Mowry, K.L., and Steitz, J.A. (1987). Identification of the human U7 snRNP as one of several factors involved in the 3' end maturation of histone premessenger RNAs. Science. *238*, 1682–1687.

Muhlrad, D., Decker, C.J., and Parker, R. (1994). Deadenylation of the unstable mRNA encoded by the yeast MFA2 gene leads to decapping followed by 5'-->3' digestion of the transcript. Genes Dev. *8*, 855–866.

Mullen, T.E., and Marzluff, W.F. (2008). Degradation of histone mRNA requires oligouridylation followed by decapping and simultaneous degradation of the mRNA both 5' to 3' and 3' to 5' Genes Dev. *22*, 50–65.

Munchel, S.E., Shultzaberger, R.K., Takizawa, N., and Weis, K. (2011). Dynamic profiling of mRNA turnover reveals gene-specific and system-wide regulation of mRNA decay. Mol. Biol. Cell. *22*, 2787–2795.

Mungunsukh, O., and Day, R.M. (2013). TGF-β1 Selectively Inhibits Hepatocyte Growth Factor Expression via an miR-199-dependent Posttranscriptional Mechanism. Mol. Biol. Cell. *24*, 2088-2097.

Murakami, K., Calero, G., Brown, C.R., Liu, X., Davis, R.E., Boeger, H., and Kornberg, R.D. (2013). Formation and fate of a complete 31-protein RNA polymerase II transcription preinitiation complex. J. Biol. Chem. *288*, 6325–6332.

Murthy, K.G., and Manley, J.L. (1992). Characterization of the multisubunit cleavage-polyadenylation specificity factor from calf thymus. J. Biol. Chem. *267*, 14804–14811.

Naguibneva, I., Ameyar-Zazoua, M., Polesskaya, A., Ait-Si-Ali, S., Groisman, R., Souidi, M., Cuvellier, S., and Harel-Bellan, A. (2006). The microRNA miR-181 targets the homeobox protein Hox-A11 during mammalian myoblast differentiation. Nat. Cell Biol. *8*, 278–284.

Nardelli, J., Gibson, T.J., Vesque, C., and Charnay, P. (1991). Base sequence discrimination by zinc-finger DNA-binding domains. Nature. *349*, 175–178.

Neilson, J.R., Zheng, G.X.Y., Burge, C.B., and Sharp, P.A. (2007). Dynamic regulation of miRNA expression in ordered stages of cellular development. Genes Dev. *21*, 578–589.

Nilsen, T.W. (2003). The spliceosome: the most complex macromolecular machine in the cell? BioEssays. *25*, 1147–1149.

Nizami, Z., Deryusheva, S., and Gall, J.G. (2010). The Cajal body and histone locus body. Cold Spring Harb. Perspect. Biol. *2*, a000653.

O'Brien, C.A., Kreso, A., and Jamieson, C.H.M. (2010). Cancer stem cells and self-renewal. Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res. *16*, 3113–3120.

Ostareck-Lederer, A., and Ostareck, D.H. (2004). Control of mRNA translation and stability in haematopoietic cells: the function of hnRNPs K and E1/E2. Biol. Cell Auspices Eur. Cell Biol. Organ. *96*, 407–411.

Pabo, C.O., Peisach, E., and Grant, R.A. (2001). Design and selection of novel Cys2His2 zinc finger proteins. Annu. Rev. Biochem. *70*, 313–340.

Paillard, L., Maniey, D., Lachaume, P., Legagneux, V., and Osborne, H.B. (2000). Identification of a C-rich element as a novel cytoplasmic polyadenylation element in *Xenopus* embryos. Mech. Dev. *93*, 117–125.

Palacios, I.M., Gatfield, D., St Johnston, D., and Izaurralde, E. (2004). An eIF4AIII-containing complex required for mRNA localization and nonsense-mediated mRNA decay. Nature. *427*, 753–757.

Pandey, N.B., and Marzluff, W.F. (1987). The stem-loop structure at the 3' end of histone mRNA is necessary and sufficient for regulation of histone mRNA stability. Mol. Cell. Biol. *7*, 4557–4559.

Pardal, R., Molofsky, A.V., He, S., and Morrison, S.J. (2005). Stem cell self-renewal and cancer cell proliferation are regulated by common networks that balance the activation of proto-oncogenes and tumor suppressors. Cold Spring Harb. Symp. Quant. Biol. *70*, 177–185.

Peng, S.S.-Y., Chen, C.-Y.A., Xu, N., and Shyu, A.-B. (1998). RNA stabilization by the AU-rich element binding protein, HuR, an ELAV protein. EMBO J. *17*, 3461–3470.

Pesole, G., Liuni, S., Grillo, G., and Saccone, C. (1997). Structural and compositional features of untranslated regions of eukaryotic mRNAs. Gene. *205*, 95–102.

Peter, M.E. (2009). Let-7 and miR-200 microRNAs: guardians against pluripotency and cancer progression. Cell Cycle. *8*, 843–852.

Pfaffl, M.W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. *29*, e45.

Pillai, R.S., Will, C.L., Lührmann, R., Schümperli, D., and Müller, B. (2001). Purified U7 snRNPs lack the Sm proteins D1 and D2 but contain Lsm10, a new 14 kDa Sm D1-like protein. EMBO J. *20*, 5470–5479.

Pillai, R.S., Bhattacharyya, S.N., Artus, C.G., Zoller, T., Cougot, N., Basyuk, E., Bertrand, E., and Filipowicz, W. (2005). Inhibition of translational initiation by Let-7 MicroRNA in human cells. Science. *309*, 1573–1576.

Pio, R., Zudaire, I., Pino, I., Castaño, Z., Zabalegui, N., Vicent, S., Garcia-Amigot, F., Odero, M.D., Lozano, M.D., Garcia-Foncillas, J., Calasanz, M.J., and Montuenga, L.M. (2004). αCP-4, Encoded by a Putative Tumor Suppressor Gene at 3p21, But Not Its Alternative Splice Variant αCP-4a, Is Underexpressed in Lung Cancer. Cancer Res. *64*, 4171–4179.

Pio, R., Blanco, D., Pajares, M.J., Aibar, E., Durany, O., Ezponda, T., Agorreta, J., Gomez-Roman, J., Anton, M.A., Rubio, A., Lozano, M.D., Lopez-Picazo, J.M., Subirada, F., Maes, T., and Montuenga, L.M. (2010). Development of a novel splice array platform and its application in the identification of alternative splice variants in lung cancer. BMC Genomics. *11*, 352.

Plath, K., and Lowry, W.E. (2011). Progress in understanding reprogramming to the induced pluripotent state. Nat. Rev. Genet. *12*, 253–265.

Pliss, A., Kuzmin, A.N., Kachynski, A.V., Jiang, H., Hu, Z., Ren, Y., Feng, J., and Prasad, P.N. (2013). Nucleolar molecular signature of pluripotent stem cells. Anal. Chem. *85*, 3545–3552.

Porciuncula, A., Zapata, N., Guruceaga, E., Agirre, X., Barajas, M., and Prosper, F. (2013). MicroRNA signatures of iPSCs and endoderm-derived tissues. Gene Expr. Patterns. *13*, 12–20.

Qian, X., Jeon, C., Yoon, H., Agarwal, K., and Weiss, M.A. (1993). Structure of a new nucleic-acid-binding motif in eukaryotic transcriptional elongation factor TFIIS. Nature. *365*, 277–279.

Qin, H., Blaschke, K., Wei, G., Ohi, Y., Blouin, L., Qi, Z., Yu, J., Yeh, R.-F., Hebrok, M., and Ramalho-Santos, M. (2012). Transcriptional analysis of pluripotency reveals the Hippo pathway as a barrier to reprogramming. Hum. Mol. Genet. *21*, 2054–2067.

Qiu, C., Ma, Y., Wang, J., Peng, S., and Huang, Y. (2010). Lin28-mediated post-transcriptional regulation of Oct4 expression in human embryonic stem cells. Nucleic Acids Res. *38*, 1240–1248.

Quenneville, S., Verde, G., Corsinotti, A., Kapopoulou, A., Jakobsson, J., Offner, S., Baglivo, I., Pedone, P.V., Grimaldi, G., Riccio, A., and Trono, D. (2011). In embryonic stem cells, ZFP57/KAP1 recognize a methylated hexanucleotide to affect chromatin and DNA methylation of imprinting control regions. Mol. Cell. *44*, 361–372.

Quenneville, S., Turelli, P., Bojkowska, K., Raclot, C., Offner, S., Kapopoulou, A., and Trono, D. (2012). The KRAB-ZFP/KAP1 system contributes to the early embryonic establishment of site-specific DNA methylation patterns maintained during development. Cell Reports. *2*, 766–773.

Query, C.C., Moore, M.J., and Sharp, P.A. (1994). Branch nucleophile selection in pre-mRNA splicing: evidence for the bulged duplex model. Genes Dev. *8*, 587–597.

Rabani, M., Levin, J.Z., Fan, L., Adiconis, X., Raychowdhury, R., Garber, M., Gnirke, A., Nusbaum, C., Hacohen, N., Friedman, N., Amit, I., and Regev, A. (2011). Metabolic labeling of RNA uncovers principles of RNA production and degradation dynamics in mammalian cells. Nat. Biotechnol. *29*, 436–442.

Radford, H.E., Meijer, H.A., and de Moor, C.H. (2008). Translational control by cytoplasmic polyadenylation in *Xenopus* oocytes. Biochim. Biophys. Acta. *1779*, 217–229.

Raghavan, A., and Bohjanen, P.R. (2004). Microarray-based analyses of mRNA decay in the regulation of mammalian gene expression. Brief. Funct. Genomic. Proteomic. *3*, 112–124.

Raghavan, A., Ogilvie, R.L., Reilly, C., Abelson, M.L., Raghavan, S., Vasdewani, J., Krathwohl, M., and Bohjanen, P.R. (2002). Genome-wide analysis of mRNA decay in resting and activated primary human T lymphocytes. Nucleic Acids Res. *30*, 5529–5538.

Rajala, T., Häkkinen, A., Healy, S., Yli-Harja, O., and Ribeiro, A.S. (2010). Effects of transcriptional pausing on gene expression dynamics. PLoS Comput. Biol. *6*, e1000704.

Ramirez, C.V., Vilela, C., Berthelot, K., and McCarthy, J.E.G. (2002). Modulation of eukaryotic mRNA stability via the cap-binding translation complex eIF4F. J. Mol. Biol. *318*, 951–962.

Ramos, I., Martín-Benito, J., Finn, R., Bretaña, L., Aloria, K., Arizmendi, J.M., Ausió, J., Muga, A., Valpuesta, J.M., and Prado, A. (2010). Nucleoplasmin binds histone H2A-H2B dimers through its distal face. J. Biol. Chem. *285*, 33771–33778.

Rattenbacher, B., Beisang, D., Wiesner, D.L., Jeschke, J.C., von Hohenberg, M., St Louis-Vlasova, I.A., and Bohjanen, P.R. (2010). Analysis of CUGBP1 targets identifies GU-repeat sequences that mediate rapid mRNA decay. Mol. Cell. Biol. *30*, 3970–3980.

Rehwinkel, J., Behm-Ansmant, I., Gatfield, D., and Izaurralde, E. (2005). A crucial role for GW182 and the DCP1:DCP2 decapping complex in miRNA-mediated gene silencing. RNA. *11*, 1640–1647.

Reimann, I., Huth, A., Thiele, H., and Thiele, B.-J. (2002). Suppression of 15-lipoxygenase synthesis by hnRNP E1 is dependent on repetitive nature of LOX mRNA 3'-UTR control element DICE. J. Mol. Biol. *315*, 965–974.

Reya, T., Morrison, S.J., Clarke, M.F., and Weissman, I.L. (2001). Stem cells, cancer, and cancer stem cells. Nature. *414*, 105–111.

Richards, M., Tan, S.-P., Tan, J.-H., Chan, W.-K., and Bongso, A. (2004). The transcriptome profile of human embryonic stem cells as defined by SAGE. Stem Cells. *22*, 51–64.

Richardson, R.T., Batova, I.N., Widgren, E.E., Zheng, L.X., Whitfield, M., Marzluff, W.F., and O'Rand, M.G. (2000). Characterization of the histone H1-binding protein, NASP, as a cell cycle-regulated somatic protein. J. Biol. Chem. *275*, 30378–30386.

Rizkallah, R., Alexander, K.E., and Hurt, M.M. (2011). Global mitotic phosphorylation of C2H2 zinc finger protein linker peptides. Cell Cycle. *10*, 3327–3336.

Rorbach, J., Nicholls, T.J.J., and Minczuk, M. (2011). PDE12 removes mitochondrial RNA poly(A) tails and controls translation in human mitochondria. Nucleic Acids Res. *39*, 7750–7763.

Ross, J. (1995). mRNA stability in mammalian cells. Microbiol. Rev. *59*, 423–450.

Rossetto, D., Avvakumov, N., and Côté, J. (2012). Histone phosphorylation: a chromatin modification involved in diverse nuclear events. Epigenetics Off. J. DNA Methylation Soc. *7*, 1098–1108.

Rowe, H.M., Jakobsson, J., Mesnard, D., Rougemont, J., Reynard, S., Aktas, T., Maillard, P.V., Layard-Liesching, H., Verp, S., Marquis, J., Spitz, F., Constam, D.B., and Trono, D. (2010). KAP1 controls endogenous retroviruses in embryonic stem cells. Nature. *463*, 237–240.

Saeki, H., Ohsumi, K., Aihara, H., Ito, T., Hirose, S., Ura, K., and Kaneda, Y. (2005). Linker histone variants control chromatin dynamics during early embryogenesis. Proc. Natl. Acad. Sci. *102*, 5697–5702.

Safer, B., Yang, L., Tolunay, H.E., and Anderson, W.F. (1985). Isolation of stable preinitiation, initiation, and elongation complexes from RNA polymerase II-directed transcription. Proc. Natl. Acad. Sci. *82*, 2632–2636.

Salomonis, N., Nelson, B., Vranizan, K., Pico, A.R., Hanspers, K., Kuchinsky, A., Ta, L., Mercola, M., and Conklin, B.R. (2009). Alternative Splicing in the Differentiation of Human Embryonic Stem Cells into Cardiac Precursors. PLoS Comput. Biol. *5*, e1000553.

Sandler, H., Kreth, J., Timmers, H.T.M., and Stoecklin, G. (2011). Not1 mediates recruitment of the deadenylase Caf1 to mRNAs targeted for degradation by tristetraprolin. Nucleic Acids Res. *39*, 4373–4386.

Sariban, E., Luebbers, R., and Kufe, D. (1988). Transcriptional and posttranscriptional control of c-fos gene expression in human monocytes. Mol. Cell. Biol. *8*, 340–346.

Sato, K., Kanno, J., Tominaga, T., Matsubara, Y., and Kure, S. (2006). De novo and salvage pathways of DNA synthesis in primary cultured neurall stem cells. Brain Res. *1071*, 24–33.

Savatier, P., Huang, S., Szekely, L., Wiman, K.G., and Samarut, J. (1994). Contrasting patterns of retinoblastoma protein expression in mouse embryonic stem cells and embryonic fibroblasts. Oncogene. *9*, 809–818.

Sawicka, K., Bushell, M., Spriggs, K.A., and Willis, A.E. (2008). Polypyrimidine-tract-binding protein: a multifunctional RNA-binding protein. Biochem. Soc. Trans. *36*, 641–647.

Sawicki, S.G., and Godman, G.C. (1971). On the differential cytotoxicity of actinomycin D. J. Cell Biol. *50*, 746–761.

Schaufele, F., Gilmartin, G.M., Bannwarth, W., and Birnstiel, M.L. (1986). Compensatory mutations suggest that base-pairing with a small nuclear RNA is required to form the 3' end of H3 messenger RNA. Nature. *323*, 777–781.

Schep, A.N., and Adryan, B. (2013). A Comparative Analysis of Transcription Factor Expression during Metazoan Embryonic Development. PloS One. *8*, e66826.

Schilders, G., Raijmakers, R., Raats, J.M.H., and Pruijn, G.J.M. (2005). MPP6 is an exosome-associated RNA-binding protein involved in 5.8S rRNA maturation. Nucleic Acids Res. *33*, 6795–6804.

Schmid, M., and Jensen, T.H. (2008). The exosome: a multipurpose RNA-decay machine. Trends Biochem. Sci. *33*, 501–510.

Schmid, M., and Jensen, T.H. (2010). Nuclear quality control of RNA polymerase II transcripts. Wiley Interdiscip. Rev. RNA. *1*, 474–485.

Schmidt, M.-J., West, S., and Norbury, C.J. (2011). The human cytoplasmic RNA terminal U-transferase ZCCHC11 targets histone mRNAs for degradation. RNA. *17*, 39–44.

Schnall-Levin, M., Rissland, O.S., Johnston, W.K., Perrimon, N., Bartel, D.P., and Berger, B. (2011). Unusually effective microRNA targeting within repeat-rich coding regions of mammalian mRNAs. Genome Res. *21*, 1395–1403.

Schneider, M.D., Najand, N., Chaker, S., Pare, J.M., Haskins, J., Hughes, S.C., Hobman, T.C., Locke, J., and Simmonds, A.J. (2006). Gawky is a component of cytoplasmic mRNA processing bodies required for early *Drosophila* development. J. Cell Biol. *174*, 349–358.

Schoenberg, D.R. (2011). Mechanisms of endonuclease-mediated mRNA decay. Wiley Interdiscip. Rev. RNA. *2*, 582–600.

Schoenberg, D.R., and Maquat, L.E. (2012). Regulation of cytoplasmic mRNA decay. Nat. Rev. Genet. *13*, 246–259.

Schuh, R., Aicher, W., Gaul, U., Côté, S., Preiss, A., Maier, D., Seifert, E., Nauber, U., Schröder, C., and Kemler, R. (1986). A conserved family of nuclear proteins containing structural elements of the finger protein encoded by Krüppel, a *Drosophila* segmentation gene. Cell. *47*, 1025–1032.

Schultz, D.C., Ayyanathan, K., Negorev, D., Maul, G.G., and Rauscher, F.J., 3rd (2002). SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. Genes Dev. *16*, 919–932.

Schultz, J., Milpetz, F., Bork, P., and Ponting, C.P. (1998). SMART, a simple modular architecture research tool: Identification of signaling domains. Proc. Natl. Acad. Sci. *95*, 5857–5864.

Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011). Global quantification of mammalian gene expression control. Nature. *473*, 337–342.

Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2013). Corrigendum: Global quantification of mammalian gene expression control. Nature. *495*, 126–127.

Scoumanne, A., Cho, S.J., Zhang, J., and Chen, X. (2011). The cyclin-dependent kinase inhibitor p21 is regulated by RNA-binding protein PCBP4 via mRNA stability. Nucleic Acids Res. *39*, 213–224.

Shabalkin, I.P. (1996). Changes in the histone/DNA ratio during the development of a multicellular organism. Bull. Exp. Biol. Med. *122*, 1251–1253.

Shalem, O., Groisman, B., Choder, M., Dahan, O., and Pilpel, Y. (2011). Transcriptome Kinetics Is Governed by a Genome-Wide Coupling of mRNA Production and Degradation: A Role for RNA Pol II. PLoS Genet. *7*, e1002273.

Sharova, L.V., Sharov, A.A., Nedorezov, T., Piao, Y., Shaik, N., and Ko, M.S.H. (2009). Database for mRNA half-life of 19 977 genes obtained by DNA microarray analysis of pluripotent and differentiating mouse embryonic stem cells. DNA Res. Int. J. Rapid Publ. Reports Genes Genomes. *16*, 45–58.

Sheets, M.D., and Wickens, M. (1989). Two phases in the addition of a poly(A) tail. Genes Dev. *3*, 1401–1412.

Shi, R., and Chiang, V.L. (2005). Facile means for quantifying microRNA expression by real-time PCR. BioTechniques. *39*, 519–525.

Shibayama, M., Ohno, S., Osaka, T., Sakamoto, R., Tokunaga, A., Nakatake, Y., Sato, M., and Yoshida, N. (2009). Polypyrimidine tract-binding protein is essential for early mouse development and embryonic stem cell proliferation. FEBS J. *276*, 6658–6668.

Shih, S.-C., and Claffey, K.P. (1999). Regulation of Human Vascular Endothelial Growth Factor mRNA Stability in Hypoxia by Heterogeneous Nuclear Ribonucleoprotein L. J. Biol. Chem. *274*, 1359–1365.

Shin, C., Nam, J.-W., Farh, K.K.-H., Chiang, H.R., Shkumatava, A., and Bartel, D.P. (2010). Expanding the MicroRNA Targeting Code: Functional Sites with Centered Pairing. Mol. Cell. *38*, 789–802.

Simon, E., Camier, S., and Séraphin, B. (2006). New insights into the control of mRNA decapping. Trends Biochem. Sci. *31*, 241–243.

Siomi, H., Matunis, M.J., Michael, W.M., and Dreyfuss, G. (1993a). The pre-mRNA binding K protein contains a novel evolutionarily conserved motif. Nucleic Acids Res. *21*, 1193–1198.

Siomi, H., Matunis, M.J., Michael, W.M., and Dreyfuss, G. (1993b). The pre-mRNA binding K protein contains a novel evolutionarily conserved motif. Nucleic Acids Res. *21*, 1193–1198.

Slager, H.G., Van Inzen, W., Freund, E., Van den Eijnden-Van Raaij, A.J., and Mummery, C.L. (1993). Transforming growth factor-beta in the early mouse embryo: implications for the regulation of muscle formation and implantation. Dev. Genet. *14*, 212–224.

Smith-Vikos, T., and Slack, F.J. (2012). MicroRNAs and their roles in aging. J. Cell Sci. *125*, 7–17.

Song, J., Saha, S., Gokulrangan, G., Tesar, P.J., and Ewing, R.M. (2012). DNA and chromatin modification networks distinguish stem cell pluripotent ground states. Mol. Cell. Proteomics. *11*, 1036–1047.

Song, M.-G., Bail, S., and Kiledjian, M. (2013). Multiple Nudix family proteins possess mRNA decapping activity. RNA. *19*, 390–399.

Spångberg, K., and Schwartz, S. (1999). Poly(C)-binding protein interacts with the hepatitis C virus 5' untranslated region. J. Gen. Virol. *80*, 1371–1376.

Spivakov, M., and Fisher, A.G. (2007). Epigenetic signatures of stem-cell identity. Nat. Rev. Genet. *8*, 263–271.

Steitz, T.A., and Steitz, J.A. (1993). A general two-metal-ion mechanism for catalytic RNA. Proc. Natl. Acad. Sci. *90*, 6498–6502.

Stoecklin, G., Lu, M., Rattenbacher, B., and Moroni, C. (2003). A constitutive decay element promotes tumor necrosis factor alpha mRNA degradation via an AU-rich element-independent pathway. Mol. Cell. Biol. *23*, 3506–3515.

Stubbs, L., Sun, Y., and Caetano-Anolles, D. (2011). Function and Evolution of C2H2 Zinc Finger Arrays. Subcell. Biochem. *52*, 75–94.

Su, W., Slepenkov, S.V., Slevin, M.K., Lyons, S.M., Ziemniak, M., Kowalska, J., Darzynkiewicz, E., Jemielity, J., Marzluff, W.F., and Rhoads, R.E. (2013). mRNAs containing the histone 3' stem-loop are degraded primarily by decapping mediated by oligouridylation of the 3' end. RNA. *19*, 1–16.

Sullivan, K.D., Mullen, T.E., Marzluff, W.F., and Wagner, E.J. (2009). Knockdown of SLBP results in nuclear retention of histone mRNA. RNA. *15*, 459–472.

Szenker, E., Ray-Gallet, D., and Almouzni, G. (2011). The double face of the histone variant H3.3. Cell Res. *21*, 421–434.

Tadepally, H.D., Burger, G., and Aubry, M. (2008). Evolution of C2H2-zinc finger genes and subfamilies in mammals: Species-specific duplication and loss of clusters, genes and effector domains. BMC Evol. Biol. *8*, 176.

Takagaki, Y., Ryner, L.C., and Manley, J.L. (1989). Four factors are required for 3'-end cleavage of pre-mRNAs. Genes Dev. *3*, 1711–1724.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. Cell. *131*, 861–872.

Tamm, I., Hand, R., and Caliguiri, L.A. (1976). Action of dichlorobenzimidazole riboside on RNA synthesis in L-929 and HeLa cells. J. Cell Biol. *69*, 229–240.

Tan, D., Marzluff, W.F., Dominski, Z., and Tong, L. (2013). Structure of histone mRNA stem-loop, human stem-loop binding protein, and 3'hExo ternary complex. Science. *339*, 318–321.

Tang, K., Breen, E.C., and Wagner, P.D. (2002). Hu protein R-mediated posttranscriptional regulation of VEGF expression in rat gastrocnemius muscle. Am. J. Physiol. Heart Circ. Physiol. *283*, H1497–1504.

Tang, Y.-S., Khan, R.A., Zhang, Y., Xiao, S., Wang, M., Hansen, D.K., Jayaram, H.N., and Antony, A.C. (2011). Incrimination of Heterogeneous Nuclear Ribonucleoprotein E1 (hnRNP-E1) as a Candidate Sensor of Physiological Folate Deficiency. J. Biol. Chem. *286*, 39100–39115.

Tani, H., Mizutani, R., Salam, K.A., Tano, K., Ijiri, K., Wakamatsu, A., Isogai, T., Suzuki, Y., and Akimitsu, N. (2012). Genome-wide determination of RNA stability reveals hundreds of short-lived noncoding transcripts in mammals. Genome Res. *22*, 947–956.

Temme, C., Zhang, L., Kremmer, E., Ihling, C., Chartier, A., Sinz, A., Simonelig, M., and Wahle, E. (2010). Subunits of the *Drosophila* CCR4-NOT complex and their roles in mRNA deadenylation. RNA. *16*, 1356–1370.

Terme, J.-M., Sesé, B., Millán-Ariño, L., Mayor, R., Izpisúa Belmonte, J.C., Barrero, M.J., and Jordan, A. (2011). Histone H1 variants are differentially expressed and incorporated into chromatin during differentiation and reprogramming to pluripotency. J. Biol. Chem. *286*, 35347–35357.

Tharun, S. (2009). Lsm1-7-Pat1 complex: a link between 3' and 5'-ends in mRNA decay? RNA Biol. *6*, 228–232.

Tharun, S., He, W., Mayes, A.E., Lennertz, P., Beggs, J.D., and Parker, R. (2000). Yeast Sm-like proteins function in mRNA decapping and decay. Nature. *404*, 515–518.

Thatcher, T.H., and Gorovsky, M.A. (1994). Phylogenetic analysis of the core histones H2A, H2B, H3, and H4. Nucleic Acids Res. *22*, 174–179.

Thomas, J.H., and Schneider, S. (2011). Coevolution of retroelements and tandem zinc finger genes. Genome Res. *21*, 1800–1812.

Thomsen, S., Anders, S., Janga, S.C., Huber, W., and Alonso, C.R. (2010). Genome-wide analysis of mRNA decay patterns during early *Drosophila* development. Genome Biol. *11*, R93.

Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., and Jones, J.M. (1998). Embryonic stem cell lines derived from human blastocysts. Science. *282*, 1145–1147.

Thukral, S.K., Morrison, M.L., and Young, E.T. (1992). Mutations in the zinc fingers of ADR1 that change the specificity of DNA binding and transactivation. Mol. Cell. Biol. *12*, 2784–2792.

Till, J.E., and McCulloch, E.A. (1961). A direct measurement of the radiation sensitivity of normal mouse bone marrow cells. Radiat. Res. *14*, 213–222.

Topisirovic, I., Svitkin, Y.V., Sonenberg, N., and Shatkin, A.J. (2011). Cap and cap-binding proteins in the control of gene expression. Wiley Interdiscip. Rev. RNA. *2*, 277–298.

Touriol, C., Morillon, A., Gensac, M.C., Prats, H., and Prats, A.C. (1999). Expression of human fibroblast growth factor 2 mRNA is post-transcriptionally controlled by a unique destabilizing element present in the 3'-untranslated region between alternative polyadenylation sites. J. Biol. Chem. *274*, 21402–21408.

Tripputi, P., Emanuel, B.S., Croce, C.M., Green, L.G., Stein, G.S., and Stein, J.L. (1986). Human histone genes map to multiple chromosomes. Proc. Natl. Acad. Sci. *83*, 3185–3188.

Tsanova, B., and van Hoof, A. (2010). Poring over exosome structure. EMBO Rep. *11*, 900–901.

Tucker, M., Valencia-Sanchez, M.A., Staples, R.R., Chen, J., Denis, C.L., and Parker, R. (2001). The transcription factor associated Ccr4 and Caf1 proteins are components of the major cytoplasmic mRNA deadenylase in *Saccharomyces cerevisiae*. Cell. *104*, 377–386.

Uchida, N., Hoshino, S.-I., and Katada, T. (2004). Identification of a human cytoplasmic poly(A) nuclease complex stimulated by poly(A)-binding protein. J. Biol. Chem. *279*, 1383–1391.

Valencia, P., Dias, A.P., and Reed, R. (2008). Splicing promotes rapid and efficient mRNA export in mammalian cells. Proc. Natl. Acad. Sci. *105*, 3386–3391.

Vardabasso, C., Hasson, D., Ratnakumar, K., Chung, C.-Y., Duarte, L.F., and Bernstein, E. (2013). Histone variants: emerging players in cancer biology. Cell. Mol. Life Sci. Epub ahead of print, PMID:23652611

Vilela, C., Ramirez, C.V., Linz, B., Rodrigues-Pousada, C., and McCarthy, J.E. (1999). Post-termination ribosome interactions with the 5'UTR modulate yeast mRNA stability. EMBO J. *18*, 3139–3152.

Vilela, C., Velasco, C., Ptushkina, M., and McCarthy, J.E. (2000). The eukaryotic mRNA decapping protein Dcp1 interacts physically and functionally with the eIF4F translation initiation complex. EMBO J. *19*, 4372–4382.

Vishnu, M.R., Sumaroka, M., Klein, P.S., and Liebhaber, S.A. (2011). The poly(rC)-binding protein alphaCP2 is a noncanonical factor in X. laevis cytoplasmic polyadenylation. RNA. *17*, 944–956.

Viswanathan, S.R., Daley, G.Q., and Gregory, R.I. (2008). Selective blockade of microRNA processing by Lin28. Science. *320*, 97–100.

Vlasova, I.A., and Bohjanen, P.R. (2008). Posttranscriptional regulation of gene networks by GU-rich elements and CELF proteins. RNA Biol. *5*, 201–207.

Vlasova, I.A., Tahoe, N.M., Fan, D., Larsson, O., Rattenbacher, B., Sternjohn, J.R., Vasdewani, J., Karypis, G., Reilly, C.S., Bitterman, P.B., and Bohjanen, P.R. (2008). Conserved GU-rich elements mediate mRNA decay by binding to CUG-binding protein 1. Mol. Cell. *29*, 263–270.

Vlasova-St Louis, I., Dickson, A.M., Bohjanen, P.R., and Wilusz, C.J. (2013). CELFish ways to modulate mRNA decay. Biochim. Biophys. Acta. *1829*, 695–707.

Waggoner, S.A., and Liebhaber, S.A. (2003). Regulation of alpha-globin mRNA stability. Exp. Biol. Med. *228*, 387–395.

Wahle, E., and Winkler, G.S. (2013). RNA decay machines: Deadenylation by the Ccr4–Not and Pan2–Pan3 complexes. Biochim. Biophys. Acta. *1829*, 561–570.

Wang, Z., and Kiledjian, M. (2000). The poly(A)-binding protein and an mRNA stability protein jointly regulate an endoribonuclease activity. Mol. Cell. Biol. *20*, 6334–6341.

Wang, H., Walsh, S.T.R., and Parthun, M.R. (2008). Expanded binding specificity of the human histone chaperone NASP. Nucleic Acids Res. *36*, 5763–5772.

Wang, H., Vardy, L.A., Tan, C.P., Loo, J.M., Guo, K., Li, J., Lim, S.G., Zhou, J., Chng, W.J., Ng, S.B., Li, H.X., and Zeng, Q. (2010). PCBP1 suppresses the translation of metastasis-associated PRL-3 phosphatase. Cancer Cell. *18*, 52–62.

Wang, T., Shi, S.-B., and Sha, H.-Y. (2013). MicroRNAs in regulation of pluripotency and somatic cell reprogramming: Small molecule with big impact. RNA Biol. *10*.

Wang, X., Kiledjian, M., Weiss, I.M., and Liebhaber, S.A. (1995). Detection and characterization of a 3' untranslated region ribonucleoprotein complex associated with human alpha-globin mRNA stability. Mol. Cell. Biol. *15*, 1769–1777.

Wang, Y., Liu, C.L., Storey, J.D., Tibshirani, R.J., Herschlag, D., and Brown, P.O. (2002). Precision and functional specificity in mRNA decay. Proc. Natl. Acad. Sci. *99*, 5860–5865.

Wang, Y., Barbacioru, C., Hyland, F., Xiao, W., Hunkapiller, K.L., Blake, J., Chan, F., Gonzalez, C., Zhang, L., and Samaha, R.R. (2006). Large scale real-time PCR validation on gene expression measurements from two commercial long-oligonucleotide microarrays. BMC Genomics. *7*, 59.

Wang, Y., Medvid, R., Melton, C., Jaenisch, R., and Blelloch, R. (2007a). DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal. Nat. Genet. *39*, 380–385.

Wang, Y., Gao, L., Tse, S.-W., and Andreadis, A. (2010b). Heterogeneous nuclear ribonucleoprotein E3 modestly activates splicing of tau exon 10 via its proximal downstream intron, a hotspot for frontotemporal dementia mutations. Gene. *451*, 23–31.

Wang, Z., Day, N., Trifillis, P., and Kiledjian, M. (1999). An mRNA stability complex functions with poly(A)-binding protein to stabilize mRNA in vitro. Mol. Cell. Biol. *19*, 4552–4560.

Wang, Z.F., Whitfield, M.L., Ingledue, T.C., 3rd, Dominski, Z., and Marzluff, W.F. (1996). The protein that binds the 3' end of histone mRNA: a novel RNA-binding protein required for histone pre-mRNA processing. Genes Dev. *10*, 3028–3040.

Warf, M.B., and Berglund, J.A. (2007). MBNL binds similar RNA structures in the CUG repeats of myotonic dystrophy and its pre-mRNA substrate cardiac troponin T. RNA. *13*, 2238–2251.

Weiss, I.M., and Liebhaber, S.A. (1995). Erythroid cell-specific mRNA stability elements in the alpha 2-globin 3' nontranslated region. Mol. Cell. Biol. *15*, 2457–2465.

Wells, S.E., Hillner, P.E., Vale, R.D., and Sachs, A.B. (1998). Circularization of mRNA by eukaryotic translation initiation factors. Mol. Cell. *2*, 135–140.

White, A.E., Burch, B.D., Yang, X.-C., Gasdaska, P.Y., Dominski, Z., Marzluff, W.F., and Duronio, R.J. (2011). *Drosophila* histone locus bodies form by hierarchical recruitment of components. J. Cell Biol. *193*, 677–694.

Whitelaw, E., and Proudfoot, N. (1986). Alpha-thalassaemia caused by a poly(A) site mutation reveals that transcriptional termination is linked to 3' end processing in the human alpha 2 globin gene. EMBO J. *5*, 2915–2922.

Whitfield, M.L., Zheng, L.-X., Baldwin, A., Ohta, T., Hurt, M.M., and Marzluff, W.F. (2000). Stem-Loop Binding Protein, the Protein That Binds the 3' End of Histone mRNA, Is Cell Cycle Regulated by Both Translational and Posttranslational Mechanisms. Mol. Cell. Biol. *20*, 4188–4198.

Whitfield, M.L., Kaygun, H., Erkmann, J.A., Townley-Tilson, W.H.D., Dominski, Z., and Marzluff, W.F. (2004). SLBP is associated with histone mRNA on polyribosomes as a component of the histone mRNP. Nucleic Acids Res. *32*, 4833–4842.

Williams, A.J., Khachigian, L.M., Shows, T., and Collins, T. (1995). Isolation and characterization of a novel zinc-finger protein with transcription repressor activity. J. Biol. Chem. *270*, 22143–22152.

Williams, R.L., Hilton, D.J., Pease, S., Willson, T.A., Stewart, C.L., Gearing, D.P., Wagner, E.F., Metcalf, D., Nicola, N.A., and Gough, N.M. (1988). Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. Nature. *336*, 684–687.

Williamson, W.D., and Pinto, I. (2012). Histones and genome integrity. Front. Biosci. *17*, 984–995.

Wilson, V.G., and Spillman, T. (1982). Changes in DNA binding by purified simian RNA polymerase II under transcribing and nontranscribing conditions. Biochim. Biophys. Acta. *699*, 232–240.

Witzgall, R., O'Leary, E., Leaf, A., Onaldi, D., and Bonventre, J.V. (1994). The Krüppel-associated box-A (KRAB-A) domain of zinc finger proteins mediates transcriptional repression. Proc. Natl. Acad. Sci. *91*, 4514–4518.

Wu, M., Reuter, M., Lilie, H., Liu, Y., Wahle, E., and Song, H. (2005). Structural insight into poly(A) binding and catalytic mechanism of human PARN. EMBO J. *24*, 4082–4093.

Xia, M., He, H., Wang, Y., Liu, M., Zhou, T., Lin, M., Zhou, Z., Huo, R., Zhou, Q., and Sha, J. (2012). PCBP1 is required for maintenance of the transcriptionally silent state in fully grown mouse oocytes. Cell Cycle. *11*, 2833–2842.

Xia, X., MacKay, V., Yao, X., Wu, J., Miura, F., Ito, T., and Morris, D.R. (2011). Translation Initiation: A Regulatory Role for Poly(A) Tracts in Front of the AUG Codon in *Saccharomyces cerevisiae*. Genetics. *189*, 469–478.

Xie, R., Medina, R., Zhang, Y., Hussain, S., Colby, J., Ghule, P., Sundararajan, S., Keeler, M., Liu, L.-J., van der Deen, M., Mitra, P., Lian, J.B., Rivera-Perez, J.A., Jones, S.N., Stein, J.L., van Wijnen, A.J., and Stein,

G.S. (2009). The histone gene activator HINFP is a nonredundant cyclin E/CDK2 effector during early embryonic cell cycles. Proc. Natl. Acad. Sci. *106*, 12359–12364.

Xue, Y., Ouyang, K., Huang, J., Zhou, Y., Ouyang, H., Li, H., Wang, G., Wu, Q., Wei, C., Bi, Y., Jiang, L., Cai, Z., Sun, H., Zhang, K., Zhang, Y., Chen, J., and Fu, X.D. (2013). Direct conversion of fibroblasts to neurons by reprogramming PTB-regulated microRNA circuits. Cell. *152*, 82–96.

Yang, E., Nimwegen, E. van, Zavolan, M., Rajewsky, N., Schroeder, M., Magnasco, M., and Darnell, J.E. (2003). Decay Rates of Human mRNAs: Correlation With Functional Characteristics and Sequence Attributes. Genome Res. *13*, 1863–1872.

Yang, L., Duff, M.O., Graveley, B.R., Carmichael, G.G., and Chen, L.-L. (2011). Genomewide characterization of non-polyadenylated RNAs. Genome Biol. *12*, R16.

Yang, X., Purdy, M., Marzluff, W.F., and Dominski, Z. (2006). Characterization of 3'hExo, a 3' exonuclease specifically interacting with the 3' end of histone mRNA. J. Biol. Chem. *281*, 30447–30454.

Yang, X., Torres, M.P., Marzluff, W.F., and Dominski, Z. (2009). Three proteins of the U7-specific Sm ring function as the molecular ruler to determine the site of 3'-end processing in mammalian histone pre-mRNA. Mol. Cell. Biol. *29*, 4045–4056.

Yang, X.-C., Sabath, I., Dębski, J., Kaus-Drobek, M., Dadlez, M., Marzluff, W.F., and Dominski, Z. (2013). A complex containing the CPSF73 endonuclease and other polyadenylation factors associates with U7 snRNP and is recruited to histone pre-mRNA for 3'-end processing. Mol. Cell. Biol. *33*, 28–37.

Yang, Y.Y.L., Yin, G.L., and Darnell, R.B. (1998). The neuronal RNA-binding protein Nova-2 is implicated as the autoantigen targeted in POMA patients with dementia. Proc. Natl. Acad. Sci. *95*, 13254–13259.

Zanier, K., Luyten, I., Crombie, C., Muller, B., Schumperli, D., Linge, J.P., Nilges, M., and Sattler, M. (2002). Structure of the histone mRNA hairpin required for cell cycle regulation of histone gene expression. RNA. *8*, 29–46.

Zarkower, D., Stephenson, P., Sheets, M., and Wickens, M. (1986). The AAUAAA sequence is required both for cleavage and for polyadenylation of simian virus 40 pre-mRNA in vitro. Mol. Cell. Biol. *6*, 2317–2323.

Zell, R., Ihle, Y., Seitz, S., Gündel, U., Wutzler, P., and Görlach, M. (2008). Poly(rC)-binding protein 2 interacts with the oligo(rC) tract of coxsackievirus B3. Biochem. Biophys. Res. Commun. *366*, 917–921.

Zhang, T., Huang, X.-H., Dong, L., Hu, D., Ge, C., Zhan, Y.-Q., Xu, W.-X., Yu, M., Li, W., Wang, X., Tang, L., Li, C.Y., and Yang, X.M. (2010). PCBP-1 regulates alternative splicing of the CD44 gene and inhibits invasion in human hepatoma cell line HepG2 cells. Mol. Cancer. *9*, 72.

Zhang, Y., Cooke, M., Panjwani, S., Cao, K., Krauth, B., Ho, P.-Y., Medrzycki, M., Berhe, D.T., Pan, C., McDevitt, T.C., and Fan, Y. (2012). Histone h1 depletion impairs embryonic stem cell differentiation. PLoS Genet. *8*, e1002691.

Zhao, Y., and Srivastava, D. (2007). A developmental view of microRNA function. Trends Biochem. Sci. *32*, 189–197.

Zheng, L., Dominski, Z., Yang, X.-C., Elms, P., Raska, C.S., Borchers, C.H., and Marzluff, W.F. (2003). Phosphorylation of stem-loop binding protein (SLBP) on two threonines triggers degradation of SLBP, the sole cell cycle-regulated factor required for regulation of histone mRNA processing, at the end of S phase. Mol. Cell. Biol. *23*, 1590–1601.

Zheng-Bradley, X., Rung, J., Parkinson, H., and Brazma, A. (2010). Large scale comparison of global gene expression patterns in human and mouse. Genome Biol. *11*, R124.

Zhou, Q., Gallagher, R., Ufret-Vincenty, R., Li, X., Olson, E.N., and Wang, S. (2011). Regulation of angiogenesis and choroidal neovascularization by members of microRNA-23/27/24 clusters. Proc. Natl. Acad. Sci. *108*, 8287–8292.

Zhu, J., and Chen, X. (2000). MCG10, a Novel p53 Target Gene That Encodes a KH Domain RNA-Binding Protein, Is Capable of Inducing Apoptosis and Cell Cycle Arrest in G2-M. Mol. Cell. Biol. *20*, 5602–5618.

Zollman, S., Godt, D., Privé, G.G., Couderc, J.L., and Laski, F.A. (1994). The BTB domain, found primarily in zinc finger proteins, defines an evolutionarily conserved family that includes several developmentally regulated genes in *Drosophila*. Proc. Natl. Acad. Sci. *91*, 10717–10721.

Zuo, Y., and Deutscher, M.P. (2001). Exoribonuclease superfamilies: structural analysis and phylogenetic distribution. Nucleic Acids Res. *29*, 1017–1026.

Appendices

Appendix A1. Selection of HFF and iPS cell sample replicates for microarray hybridization. *FOS* and *TUT1* mRNAs were used as short- and longer-lived test transcripts to determine half-life correlation between four HFF and iPS cell replicates (A, B, C, and D) where cells were treated with actinomycin d and collected at time points 0, 15, 30, 60, 120, and 240 minutes following treatment. The three replicates with the lowest standard deviation between mRNA abundances at each time point were selected for microarray hybridization.

Comparative analysis of test mRNA abundance in HFFs for selection of sample replicates.

FOS	**ABC**		**ABD**		**ACD**		**BCD**	
Time	**Expression**	**Std Dev**	**Expression**	**Std Dev**	**Expression**	**Std Dev**	**Expression**	**Std Dev**
0	1	0	1	0	1	0	1	0
15	0.8234	0.1498	0.7952	0.1399	0.7796	0.1180	0.8728	0.0665
30	0.5461	0.0423	0.5056	0.0877	0.5154	0.0899	0.4876	0.0634
60	0.2833	0.0046	0.3258	0.0698	0.3228	0.0724	0.3244	0.0712

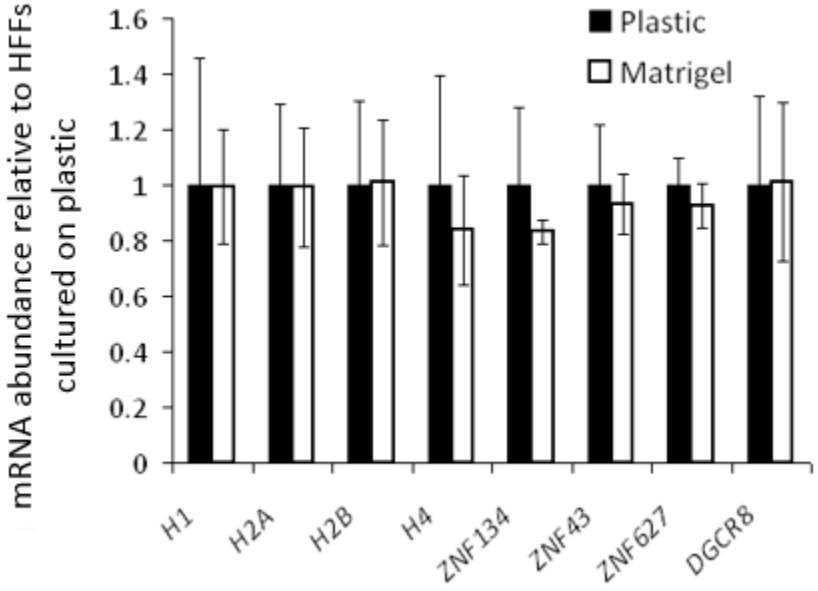
TUT1	**ABC**		**ABD**		**ACD**		**BCD**	
Time	**Expression**	**Std Dev**	**Expression**	**Std Dev**	**Expression**	**Std Dev**	**Expression**	**Std Dev**
0	1	0	1	0	1	0	1	0
60	0.6700	0.0225	0.8896	0.3629	0.8747	0.3756	0.8841	0.3681
120	0.5490	0.0263	0.5452	0.0328	0.5304	0.0307	0.5297	0.0294
240	0.2287	0.0367	0.2514	0.0701	0.2574	0.0694	0.2749	0.0455

Comparative analysis of test mRNA abundance in iPS cells for selection of sample replicates.

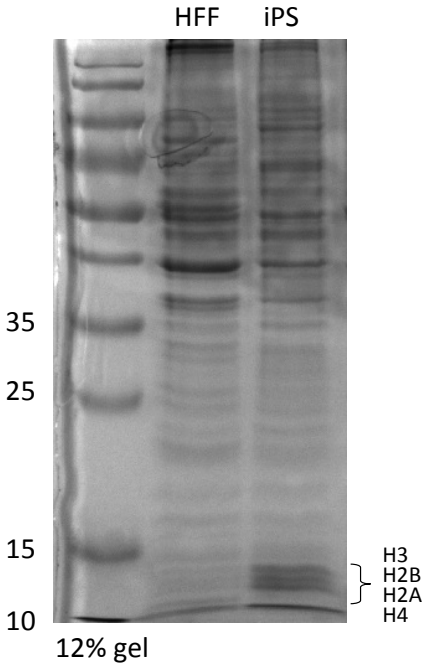
FOS	**ABC**		**ABD**		**ACD**		**BCD**	
Time	**Expression**	**Std Dev**	**Expression**	**Std Dev**	**Expression**	**Std Dev**	**Expression**	**Std Dev**
0	1	0	1	0	1	0	1	0
15	0.6928	0.0982	0.6952	0.0990	0.7378	0.0358	0.6735	0.0759
30	0.4833	0.0845	0.4124	0.0550	0.4682	0.1064	0.4468	0.1121
60	0.3994	0.0938	0.3663	0.0390	0.4254	0.0730	0.4138	0.0888

TUT1	**ABC**		**ABD**		**ACD**		**BCD**	
Time	**Expression**	**Std Dev**	**Expression**	**Std Dev**	**Expression**	**Std Dev**	**Expression**	**Std Dev**
0	1	0	1	0	1	0	1	0
60	0.9439	0.0934	0.8960	0.0942	0.8914	0.0862	0.9474	0.0873
120	0.8176	0.1824	0.7609	0.2381	0.6642	0.0879	0.7765	0.2326
240	0.4489	0.1999	0.4498	0.1990	0.3269	0.0281	0.4331	0.2120

Appendix A2. Culture of HFF cells on Matrigel has no effect on mRNA abundance. HFFs were cultured on 0.3 mg/ml Matrigel and mRNA levels were measured by qRT-PCR to determine that expression of test mRNAs were not affected compared to HFFs cultured on plastic.



Appendix A3. Abundances of core histone proteins were verified in an independent set of matched HFF and iPS cells. Coomassie staining was used to verify increased abundance of core histone proteins in iPS cells compared to HFFs. These results are consistent with the matched cell line used for experiments presented in this dissertation.



Appendix A4. 151 mRNAs destabilized in iPS cells compared to HFFs that bear C-rich 3'UTR elements. The P-value for each cell line indicates how well the time point abundance data fit to the exponential decay curve used to generate each half-life.

Transcript ID	Gene ID	Gene Symbol	iPS Half-life (min)	iPS P-value	HFF Half-life (min)	HFF P-value
7933204	11067	C10orf10	64	5.71E-05	517	3.64E-04
7998898	9074	CLDN6	194	6.84E-03	1455	3.40E-02
8133326	9883	POM121	53	6.93E-03	394	1.31E-02
8131996	9586	CREB5	253	2.56E-03	1733	1.87E-02
7955589	3164	NR4A1	189	4.46E-04	945	8.67E-03
8066266	9935	MAFB	299	2.36E-03	1395	4.36E-02
8058498	7855	FZD5	391	1.12E-03	1760	2.18E-02
8138842	10392	NOD1	462	1.06E-04	2119	1.30E-02
7913644	1870	E2F2	431	2.49E-03	1863	4.17E-02
7897469	199953	TMEM201	339	9.46E-03	1382	4.47E-02
7948211	6749	SSRP1	363	4.98E-06	1456	1.25E-02
8087833	1849	DUSP7	96	8.32E-04	371	2.41E-02
8037079	478	ATP1A3	189	2.08E-03	726	5.61E-03
8072229	4744	NEFH	270	3.85E-03	1048	1.60E-02
7903980	128346	C1orf162	401	6.55E-04	1514	4.52E-03
8026155	112939	NACC1	194	4.66E-03	721	4.44E-02
7906128	9673	SLC25A44	177	5.68E-04	644	1.23E-02
8046524	3239	HOXD13	172	1.29E-03	614	2.77E-02
7989657	53944	CSNK1G1	259	7.77E-05	902	9.47E-03
8068361	6526	SLC5A3	215	3.01E-04	719	4.34E-02
8069880	7074	TIAM1	438	3.94E-03	1449	1.79E-02
8173506	54821	ERCC6L	329	3.22E-03	1108	9.44E-05
8166447	139411	PTCHD1	529	2.26E-04	1691	2.85E-02
7945944	391	RHOG	135	1.67E-04	427	4.34E-02
7941565	64837	KLC2	216	7.05E-03	650	4.11E-02
8109179	133522	PPARGC1B	457	4.16E-03	1395	4.88E-02
8040725	56896	DPYSL5	605	3.11E-02	1838	2.17E-02
7972259	1638	DCT	690	3.10E-03	2031	1.48E-02
8119444	116113	FOXP4	263	1.44E-02	772	2.85E-02
8123658	63027	SLC22A23	427	2.50E-03	1210	2.78E-02
7930264	9148	NEURL	377	2.22E-02	1047	2.70E-02
7989073	283659	PRTG	308	1.13E-02	848	2.11E-03
8160459	1993	ELAVL2	680	8.14E-04	1913	4.21E-03
8158839	84726	BAT2L1	192	1.74E-03	535	1.14E-02
8087685	11070	TMEM115	168	5.25E-04	451	8.17E-03
7920852	22889	KIAA0907	300	6.37E-05	809	3.46E-02
8128620	57673	BEND3	318	9.99E-04	867	3.48E-03
8034263	1995	ELAVL3	582	4.54E-03	1515	1.64E-02

Appendix A4. Continued

Transcript ID	Gene ID	Gene Symbol	iPS Half-life (min)	iPS P-value	HFF Half-life (min)	HFF P-value
7901788	4774	NFIA	364	9.00E-04	962	1.86E-02
7994889	10847	SRCAP	122	1.97E-03	319	6.84E-04
8031441	84446	BRSK1	397	1.82E-03	1025	3.05E-02
7992789	51330	TNFRSF12A	201	1.03E-03	511	2.84E-02
8175947	3054	HCFC1	197	1.86E-02	501	1.90E-02
8058627	2066	ERBB4	706	1.36E-03	1812	1.19E-02
8063437	128553	TSHZ2	657	4.07E-03	1703	9.53E-03
8047565	150864	FAM117B	356	2.89E-03	925	1.21E-02
7984952	56905	C15orf39	110	4.08E-04	280	2.31E-04
8000167	23049	SMG1	212	7.28E-05	531	8.15E-03
8130408	26034	IPCEF1	844	3.80E-03	2095	2.74E-02
8125731	9278	ZBTB22	317	1.15E-03	775	4.66E-02
8118509	9374	PPT2	353	1.25E-04	863	4.44E-02
8055952	4929	NR4A2	327	1.17E-04	803	2.08E-03
7996185	4324	MMP15	405	5.65E-03	977	4.15E-02
8054281	164832	LONRF2	624	4.81E-02	1504	1.69E-02
8022420	162655	ZNF519	531	9.24E-03	1263	3.61E-03
8025828	3949	LDLR	121	1.15E-05	290	3.18E-04
8104035	8470	SORBS2	760	1.15E-02	1799	1.75E-02
8097857	84057	MND1	489	3.65E-02	1162	3.46E-02
7955873	3223	HOXC6	318	2.30E-03	755	2.15E-03
8166230	55787	TXLNG	307	5.43E-03	717	1.73E-02
8052689	200734	SPRED2	146	1.32E-04	341	9.85E-05
7962951	8085	MLL2	216	3.34E-03	497	1.14E-02
7980233	5228	PGF	313	2.05E-03	721	3.71E-03
8002692	463	ZFHX3	312	1.90E-03	723	1.72E-02
8137953	84629	TNRC18	313	5.03E-03	718	4.43E-02
8170921	55558	PLXNA3	292	1.00E-02	664	4.46E-02
8072796	164656	TMPRSS6	638	5.97E-03	1465	4.25E-02
7975626	91748	C14orf43	470	8.20E-03	1060	4.39E-02
8071044	100132288	LOC100132288	231	3.39E-04	509	1.15E-03
7986359	3480	IGF1R	276	1.13E-03	603	2.32E-02
8008139	65264	UBE2Z	290	5.51E-03	628	4.00E-02
7977761	6297	SALL2	357	1.80E-03	783	1.50E-03
8075126	4330	MN1	500	4.64E-03	1099	4.36E-02
8013776	55731	C17orf63	154	1.78E-03	333	1.31E-03
8016628	84687	PPP1R9B	305	3.24E-03	650	1.92E-02
8014794	782	CACNB1	341	1.72E-03	718	3.96E-02

Appendix A4. continued

Transcript ID	Gene ID	Gene Symbol	iPS Half-life (min)	iPS P-value	HFF Half-life (min)	HFF P-value
7940051	25921	ZDHHC5	308	3.59E-04	652	3.42E-02
8000638	23049	SMG1	306	2.59E-03	644	1.50E-02
7920000	23126	POGZ	157	1.84E-03	335	3.76E-03
7972713	1948	EFNB2	241	4.31E-05	513	2.32E-03
7949277	23130	ATG2A	294	7.84E-03	632	6.95E-03
8024391	84444	DOT1L	363	1.14E-02	749	1.98E-03
7918936	79679	VTCN1	1098	2.36E-02	2276	4.54E-02
7992347	23162	MAPK8IP3	323	4.88E-03	666	2.29E-02
8004671	23135	KDM6B	358	3.76E-03	746	7.66E-03
7999173	124402	FAM100A	193	7.58E-03	400	8.13E-03
7963698	11016	ATF7	383	1.56E-02	801	4.73E-02
8150036	23303	KIF13B	645	6.01E-04	1309	2.65E-02
7995739	2775	GNAO1	710	3.04E-02	1455	2.74E-02
7998211	8312	AXIN1	300	4.23E-04	611	1.73E-03
7972902	113622	ADPRHL1	532	9.09E-03	1076	6.72E-03
8169882	63035	BCORL1	216	1.30E-03	444	1.96E-03
8027956	9757	MLL4	275	2.04E-03	568	4.82E-03
8124726	7726	TRIM26	289	5.53E-04	592	1.24E-02
8053901	55654	TMEM127	268	2.94E-03	543	2.49E-02
8004404	284114	TMEM102	319	7.49E-04	648	1.85E-03
8108822	285613	RELL2	476	2.42E-02	979	4.72E-02
8025659	84971	ATG4D	300	3.32E-03	612	3.67E-03
8110430	23138	N4BP3	486	5.40E-03	981	3.96E-02
8162533	5727	PTCH1	253	2.37E-03	505	3.12E-03
8021047	26040	SETBP1	357	3.97E-04	714	5.15E-03
8001387	6299	SALL1	121	5.98E-04	243	1.06E-03
8112592	2297	FOXD1	158	2.96E-03	317	6.14E-03
8010188	57690	TNRC6C	258	1.17E-03	502	1.04E-02
8002969	4094	MAF	650	1.26E-02	1269	3.17E-02
7994655	79447	C16orf53	234	1.01E-03	459	1.40E-02
7949383	84447	SYVN1	227	2.85E-03	445	2.05E-02
8008566	252983	STXBP4	446	4.87E-03	880	4.96E-02
8171297	4281	MID1	440	1.40E-03	856	2.95E-02
7943282	55693	KDM4D	369	7.63E-05	719	1.15E-03
7904907	607	BCL9	332	3.47E-03	646	2.64E-02
8032518	148252	DIRAS1	428	2.04E-03	823	2.76E-02
7967331	79720	VPS37B	159	5.40E-04	303	4.80E-03
8178988	9278	ZBTB22	262	2.99E-04	505	1.47E-04

Appendix A4. continued

Transcript ID	Gene ID	Gene Symbol	iPS Half-life (min)	iPS P-value	HFF Half-life (min)	HFF P-value
8063097	63935	PCIF1	427	1.60E-02	817	2.29E-02
8009995	85451	UNK	188	5.82E-04	359	3.94E-04
8032347	57455	REXO1	282	8.33E-03	539	7.29E-03
7931469	170394	PWWP2B	490	3.56E-04	950	1.59E-02
8133331	9883	POM121	149	6.03E-03	285	1.25E-02
8019308	4097	MAFG	304	3.08E-04	578	4.35E-03
7946180	10612	TRIM3	510	1.06E-03	964	2.53E-02
8040927	29959	NRBP1	281	4.05E-04	525	2.38E-02
8040552	8648	NCOA1	382	3.07E-03	722	4.31E-02
8103745	9464	HAND2	501	8.30E-04	936	1.57E-02
7907370	26052	DNM3	877	1.31E-03	1665	2.17E-03
8019463	1453	CSNK1D	292	1.23E-03	547	2.33E-02
8023646	596	BCL2	681	1.16E-02	1284	4.41E-02
8141595	64599	GIGYF1	399	5.53E-03	735	1.11E-02
8108579	55374	TMCO6	475	3.41E-03	874	6.51E-03
7906728	84134	TOMM40L	485	7.00E-03	907	2.24E-02
8002333	23049	SMG1	330	1.97E-03	613	9.42E-03
8000217	23049	SMG1	346	7.21E-03	641	3.12E-02
8003249	79791	FBXO31	401	1.43E-04	749	2.19E-02
7947462	25841	ABTB2	258	6.82E-05	465	7.53E-05
8029856	2909	GRLF1	205	5.98E-05	369	9.97E-04
7980998	55727	BTBD7	292	3.90E-05	527	1.21E-02
7930537	6934	TCF7L2	241	1.13E-03	438	1.31E-02
7994985	9739	SETD1A	194	1.47E-03	350	9.45E-04
7985605	9640	ZNF592	159	1.18E-03	290	2.06E-04
8002370	6483	ST3GAL2	341	3.62E-03	617	3.96E-02
7961798	6660	SOX5	659	9.44E-04	1206	2.78E-02
8034722	5989	RFX1	284	8.06E-05	522	1.96E-03
8075468	50487	PLA2G3	634	1.85E-03	1160	3.45E-02
7959361	22877	MLXIP	178	1.01E-02	323	7.66E-03
8023377	51320	MEX3C	293	3.43E-03	535	3.82E-02
7939676	9776	KIAA0652	366	8.90E-03	668	3.97E-02
8013071	201163	FLCN	317	3.44E-04	578	7.73E-03
8018379	57513	CASKIN2	283	1.93E-03	511	7.50E-03
8153876	80728	ARHGAP39	390	3.72E-03	712	9.01E-03
8071314	54487	DGCR8	231	2.68E-05	410	1.53E-03
7905854	51043	ZBTB7B	267	2.03E-03	479	8.26E-03

Appendix A5. 2D gel electrophoresis analysis of PCBPs and GAPDH or α-TUBULIN in HFF and iPS cells. Analysis was performed using equal numbers of cells and amount of whole cell lysate. The first dimension was run on pH 1-10 gradient strips and the second dimension was run on Mini-Protean TGX gels and transferred to PVDF membranes. Antibody detection was performed using the ChemiDoc XRS+ System and ImageLab3.0 software.