Chapter 4

Conclusion

My dissertation research has focused on identifying host factors and molecular mechanisms that inhibit human L1 retrotransposition. In Chapter two, I expanded on previous observations that L1 RNA is subject to splicing (Belancio et al., 2006; Belancio et al., 2008), and demonstrated that splicing is detrimental to L1 retrotransposition in a cell culture assay. I discovered a new class of Spliced Integrated Retrotransposed Elements (SpIREs) and examined whether two predominant classes of SpIREs within the human genome are capable of retrotransposition. Intra-5'UTR SpIREs are spliced within their 5'UTR, whereas 5'UTR/ORF1 SpIREs are spliced from the 5'UTR into the ORF1 coding sequence. I hypothesized that intra-5'UTR SpIREs and 5'UTR/ORF1 SpIREs could have different consequences for L1 retrotransposition. This led to my investigation, and subsequent demonstration, that downstream initiation codons of 5'UTR/ORF1 spliced isoforms could be utilized for ORF1 translation initiation. Additionally, I discussed the possibility that utilization of L1 splice sites change over evolutionary history and hypothesize that these changes arise as a consequence of an L1-host arms race.

I also wanted to identify host factors that restrict L1 propagation. Namely, I built upon previous observations that host factors recognize and transcriptionally silence recently retrotransposed engineered L1 reporter genes (L1-reporter) in a human embryonic carcinoma cell line (hEC) either during or immediately after retrotransposition (Garcia-Perez et al., 2010). Chapter three describes the development of a forward genetic screen aimed at identifying factors that might be involved in L1-reporter silencing events. In sum, I, and others in the lab, developed a CRISPR/Cas9 based gene knockout strategy and validated the efficacy of the strategy in hECs. Finally, I

performed preliminary experiments demonstrating that our strategy could lead to the discovery of new factors that silence integrated L1s in hECs

Identification of a new spliced L1 isoform and the history of L1 splicing

Previous publications demonstrated genomic L1s and transfected L1 RNAs are spliced and or prematurely polyadenylated (Belancio et al., 2006; Belancio et al., 2008; Belancio et al., 2010; Perepelitsa-Belancio and Deininger, 2003). Additional observations from our lab identified a polymorphic L1 in the human population that contained an intra-5'UTR deletion that appeared to be the result of splicing (Beck et al., 2010). One of my goals was to understand how splicing of L1 RNAs might change over evolutionary time. Additionally, I wanted determine how splicing might affect L1 retrotransposition.

I first wanted to catalogue and annotate Spliced Integrated Retrotransposed Elements (SpIREs) present in the human genome reference (HGR). Belancio and colleagues originally identified an intra-5'UTR spliced isoform ($SpIRE_{97/790}$), and a 5'UTR/ORF1 spliced isoform ($SpIRE_{97/976}$) that are present at four and ten copies, respectively in the HGR (Figure 2.1, 2.6) (Belancio et al., 2006). I inspected the HGR for the spliced polymorphic L1 identified in Beck et al., 2010 ($SpIRE_{97/622}$) and surprisingly found that it was present at 116 copies and was ten-fold more abundant than previously identified SpIREs (Figure 2.1, 2.6). Given the relative abundance of $SpIRE_{97/622}$, I sought to determine if I could observe this splicing event at the RNA level.

I transfected HeLa cells with a luciferase construct that contained a full-length L1 5'UTR driving expression of a FireFly luciferase gene (pPL5'UTRLuc, Figure 2.2). Using Northern blots, I was unable to observe a splicing event consistent with $SpIRE_{97/622}$ splicing. Indeed, RT-PCR experiments using RNAs from HeLa cells transfected with pPL5'UTRLuc failed to produce any products consistent with $SpIRE_{97/622}$ splicing (Figure 2.2). Surprisingly, those same experiments revealed products consistent with the $SpIRE_{97/790}$ splicing event (Figure 2.2). Thus, experiments in cell culture suggested that RNAs derived from pPL5'UTRLUC predominantly gave rise to splicing events consistent with $SpIRE_{97/790}$, which is only present at 4 copies in the human genome, but not spliced

RNAs consistent with $SpIRE_{97/622}$, which is present at 116 copies in the human genome. Fortuitously, a recent publication presented data that supported the hypothesis that changes in the L1 5'UTR, driven by host factor repression, could explain changes in L1 splicing events.

In the introductory chapter of my thesis, I described how the L1 5'UTR has changed over evolutionary time. I now hypothesize that changes within the L1 5'UTR also have affected the landscape of L1 RNA splicing. Experiments in cell culture demonstrated that the KRAB zinc-finger gene ZNF93 could inhibit L1 retrotransposition of the older L1PA3 subfamily, but not the younger L1PA2, or human specific L1 (L1Hs) subfamilies. A 129 bp deletion within the L1PA3 5'UTR deleted the ZNF93 DNA (Jacobs et al., 2014). The deletion subsequently allowed the L1PA2 5'UTR to escape ZNF93 repression and amplify in the genome (Jacobs et al., 2014) (Figure 2.9, 4.1). Strikingly, the $SpIRE_{97/790}$ that I identified in my RNA analyses, are only found in the L1PA2 and L1Hs subfamilies, while $SpIRE_{97/622}$ is found in the L1PA6-L1Hs subfamilies (Figure 2.6, 4.1) Based on *in silico* prediction programs, the 129-nucleotide deletion likely resulted in the juxtaposition of a splice acceptor sequence near a polypyrimidine tract, thereby favoring splicing consistent with $SpIRE_{97/790}$ over $SpIRE_{97/622}$ (Figure 2.9, 4.1).

Genomic evidence also supports this hypothesis. Since L1PA3, the number of subfamily specific full-length genomic L1s has been declining. There are 1548 full-length L1PA3 elements and only 1088 full-length L1PA2 elements, a decrease of roughly 33% (Figure, 2.6) (Boissinot and Furano, 2001; Boissinot et al., 2004; Khan et al., 2006; Witherspoon et al., 2006). However, the number of $SpIRE_{97/622}$ drops from 53 to 10 in the same time span, a decrease of ~80%. Thus, the genomic record, coupled with the in *vitro* data I uncovered, suggests that as a new L1 5'UTR evolves to escape host factor repression, it can be subject to new, deleterious splicing events.

It is surprising that potentially deleterious sequences would be retained in L1. The splice donor that is utilized in all observed reactions is likely conserved because it resides within a RUNX3 transcription factor-binding site (Figure 2.6) (Yang et al., 2003). Indeed, I showed that a single point mutation that both disrupts the splice donor sequence as well as abolishes the RUNX3 site drastically reduces L1 transcript levels

(Figure 2.2). The reason for conservation of splice acceptor sites within the 5' UTR is less clear, but they may reside within undefined transcription factor binding sites important for 5'UTR function. The splice acceptor site within ORF1 encompasses positions 2 and 3 of codon 66; thus, any mutations to the second codon position would lead to an amino acid substitution, which may disrupt ORF1p function.

L1 splicing negatively impacts L1 retrotransposition

I next wanted to determine the effect of splicing on L1 retrotransposition. I created an L1 retrotransposition construct driven by a 5'UTR that contained a deletion consistent with the $SpIRE_{97/622}$ splicing event and tested its ability to retrotranspose in a cell culture retrotransposition assay (Figure 2.3, 2.4) (Moran et al., 1996; Wei et al., 2001). Notably, the deletion resulted in an L1 that retrotransposed ~11-fold less efficiently than an L1 driven by a full-length 5'UTR (Figure 2.3, 2.4). However, addition of a constitutive promoter augmenting the internally deleted 5'UTR largely rescued retrotransposition, suggesting that if intra-5'UTR spliced L1s fortuitously retrotranspose downstream of an endogenous promoter they could undergo subsequent rounds of retrotransposition (Figure 2.4, 2.5). However, it is notable, that any subsequent retrotransposition events occurring by this scenario would likely lack a promoter. Thus, these events would only be capable of arithmetic amplification. In sum, my data suggest that virtually all $SpIRE_{97/622}$ sequences represent evolutionary 'dead' ends for L1, and cannot give rise to subsequent retrotransposition events.

Since intra-5'UTR splicing drastically hinders L1 retrotransposition efficiency, I next wanted to test the effect of 5'UTR/ORF1 splicing on L1 retrotransposition. ORF1p is necessary for L1 retrotransposition, so I hypothesized that this deletion would render the L1 defective for retrotransposition (Moran et al., 1996). I created an L1 retrotransposition construct containing a deletion consistent with the $SpIRE_{97/976}$ splicing event and tested its ability to retrotranspose in a cell culture retrotransposition assay (Figure 2.3, 2.4) (Wei et al., 2001). Reproducibly, the 5'UTR/ORF1 deletion resulted in an L1 that was unable to retrotranspose (Figure 2.4). Western blot analyses indicated that 5'UTR/ORF1 deletion construct translate two truncated forms of ORF1p that are apparently unable to support retrotransposition (Figure 2.3, 2.4). If ORF1p is supplied in

trans from a construct co-transfected with the deficient 5'UTR/ORF1 L1 construct, a small level of retotransposition is rescued (*trans*-complementation) (Figure 2.4) (Wei et al., 2001). However, *trans*-complementation is much less efficient than the canonical retrotransposition mechanism of *cis*-preference (Wei et al., 2001). The inefficiency of *trans*-complementation likely explains the observation that there are only 11 5'UTR/ORF1 SpIREs in the human genome compared to 120 intra-5'UTR SpIREs (Figure 2.6).

Splicing of L1 RNA may simply be a by-product of a canonical cellular process occasionally acting on an L1 molecule. However, Northern blots from a panel of human tissues demonstrates that the vast majority of L1 RNAs expressed in somatic tissues are heavily processed, likely by either pre-mature polyadenylation and/or splicing (Belancio et al., 2010). Thus, I hypothesized a number of scenarios whereby L1 RNA splicing could potentially serve in a more 'active' role of L1 retrotransposition repression.

A 10,000-foot view of the splicing reaction

Splicing is a complex and dynamic process that includes dozens of proteins and RNAs that act, in concert, to ensure proper processing of cellular transcripts (Wahl et al., 2009). The spliceosomal ribonucleoprotein particle (RNP) is composed of small nuclear RNAs (snRNAs) as well as other RNA binding proteins (Wahl et al., 2009). Thus, splicing relies on both *cis*-acting sequences at exon/intron borders and *trans* acting factors, such as heterogenous nuclear proteins and their cognate RNAs (hnRNPs), as well as serine-arginine rich splicing factors (SRSFs), which regulate exon inclusion or exclusion in a cell specific manner.

Aside from the highly conserved splice donor sequence, *cis*-acting sequences such as the splice acceptor and branchpoint sequence that bind *trans*-acting factors are difficult to predict based on sequence composition alone. Splicing of normal genic introns is dependent on *trans*-acting factors that properly fold intronic mRNA and catalyze the reaction (Wahl et al., 2009). SRSF proteins typically bind exonic splicing enhancers (ESEs) to promote exon inclusion, whereas hnRNPs typically bind exonic splice silencers (ESSs) to promote exon exclusion and facilitate splice site selection (Chiou et al., 2013; Han et al., 2010; Long and Caceres, 2009; Pandit et al., 2013).

These factors often act antagonistically to one another and it is the sum of multiple *trans*-acting factors that determine splice site usage and exon choice (Han et al., 2010; Long and Caceres, 2009).

Could Splicing of L1 mRNA act as a transposon sensor in human cells?

Due to the complexity of splicing and the requirement for a high-fidelity product, the cell has evolved mechanisms to prevent or stop splicing if the reaction stalls or is not completed (Koodathingal and Staley, 2013; Semlow and Staley, 2012). Stalled or slowed spliceosomes can signal an RNAi response in the yeast *Cryptoccocus neoformans* (Dumesic et al., 2013). Data from Dumseic et al., 2013 suggest that this process is utilized to target *C. neoformans* transposons that are spliced at a slower rate than normal genic mRNAs. Though this mechanism relies on an RNA-dependent RNA polymerase (RDRP), a similar mechanism could be utilized for L1 independent of RDRP activity. In *vivo* and in *vitro* studies demonstrate L1 siRNAs can originate from antisense transcription of the L1 5'UTR and are postulated to target L1 transcripts in human cultured cells (Soifer et al., 2005; Yang and Kazazian, 2006). If an L1-derived siRNA pairs with an L1 RNA that has a stalled splicing reaction, it could serve as a signal to recruit proteins involved in RNAi-mediated genomic silencing or L1 transcript degradation, an idea that has been previously suggested (Dumesic and Madhani, 2013).

Additionally, in HeLa cells, it has been demonstrated that Argonaute proteins associate with the splicing machinery and members of the serine/arginine rich (SRSF) family of splicing factors (Ameyar-Zazoua et al., 2012). Argonaute proteins were demonstrated to regulate alternative exon inclusion at the *CD44* gene by inducing local heterochromatin changes and slowing PolII elongation (Ameyar-Zazoua et al., 2012). One hypothesis is that SRSF proteins associated with stalled L1 splicing recruit Argonaute proteins that remodel the local chromatin environment, potentially disrupting L1 transcription. Alternatively, stalled splicing could recruit Argonaute proteins (Huang and Zhu, 2014) that mediated destruction of the L1 RNA.

Stalled splicing, as a transposon sensor for human L1s, is highly speculative. Subsequent experiments should examine the context of splicing from RNAs expressed

from genomic L1s. These experiments are technically challenging analysis due to the nature of repeat DNA sequences. However, recent reports have used long RNA sequencing reads (>250 bp) to accurately identify spliced gene variants (Cho et al., 2014). Deep sequencing of genomic expressed L1 RNAs may yield insight into the constellation L1 splicing events in human cells. Experimenters could knock out potential candidate genes that may modulate L1 splicing and splicing-mediated silencing. Subsequent sequencing to determine the constellation of expressed L1 RNAs would begin to shed light on whether splicing of L1 RNAs recruits gene silencing, or RNA destruction machinery. L1 RNA-IP followed by tandem-mass spectrometry should also be included to identify which proteins are bound to nascent L1 RNA. Isolation of cytoplasmic and nuclear RNA fractions could shed light on identifying proteins that interact with L1 RNA in specific compartments.

Which splicing factors interact with L1 and what is their effect on retrotransposition?

Splicing associated proteins are logical candidates to study with regard to their effect on L1 biology and retrotransposition. A number of recent reports have identified splicing factors, that co-immunoprecipitate with ORF1p and ORF2p (Goodier et al., 2013; Moldovan, 2015; Moldovan and Moran, 2015; Peddigari et al., 2013; Taylor et al., 2013). Of the 15 proteins identified in Goodier et al. 2013 that inhibit retrotransposition at greater than 75%, just over half (8/15) are associated with splicing and mRNA processing. There is ample evidence that suggests hnRNPL effectively inhibits L1 retrotransposition in cell culture assays when overexpressed (Goodier et al., 2013), and permits increased retrotransposition when knocked down (Moldovan, 2015). However, not all splicing factors inhibit retrotransposition, as overexpression of hnRNPQ, hnRNPK, and hnRNPR had little to no effect on L1 retrotransposition in a cell culture assay (Goodier et al., 2013). It remains possible that knockdown of factors could affect L1 RNA splicing. In Appendix one, I demonstrated that overexpressing the splicing factors SRSF1 and SRSF7 decrease efficiency of L1 retrotransposition in a cell culture assay to ~20%, and ~45% when compared to controls (Figure A2.A). The inhibitory mechanism of SRSF proteins on L1 retrotransposition is unknown, but they represent attractive candidates for future investigation.

Could splicing be beneficial for L1 retrotransposition?

It is remotely possible that in certain contexts, L1 RNA splicing is beneficial for retrotransposition. Splicing has been demonstrated to aid in the recruitment of activating transcription factors (Furger et al., 2002). Furthermore, a 5' splice donor sequence induces increased rates of transcription even in the *absence* of splicing the transcript (Damgaard et al., 2008). Given this information, it is possible that in addition to retaining the RUNX3 binding site (Figure 2.6), the 5'UTR splice donor sequence has been conserved to increase local transcription at the L1 locus. However, utilization of the splice donor in splicing reactions weakens the promoter of subsequent retrotransposition events, and thus is not likely to be useful.

Alternatively, splicing could aid in promoting L1 RNAs transition to the cytoplasm. It has been shown that splicing promotes export of mRNA to the cytoplasm (Valencia et al., 2008). Though splicing of the L1 is detrimental, it may be that splicing the first RNA molecules that emanate from a locus recruits export factors that are then utilized for export of future transcripts transcribed from that locus. Though this is unlikely, it is not impossible.

How might L1 prevent splicing?

As I demonstrated, L1 RNA splicing is detrimental to L1 retrotransposition. Thus, it seems likely selection has driven mutations in the L1 sequence to limit splicing events. There are two main possibilities and both involve the L1 RNA. As the 5'UTR encodes both a sense and anti-sense promoter, dsRNA pairing between the sense and anti-sense transcripts is possible. As the antisense transcript largely spans the 5'UTR SpIRE "intron" it is a likely candidate to bind back to the sense L1 transcript and prevent splicing, or splicing-mediated inhibition by RNAi factors. Indeed this sort of sense/antisense pairing has been shown to prevent inclusion of alternatively spliced exons at the TRalpha gene (Hastings et al., 2000). When expressed, an antisense transcript overlapping with the final intron/exon boundary of the TRalpha gene prevented inclusion of that exon (Hastings et al., 2000; Hastings and Krainer, 2001). Importantly, those investigators demonstrated that alternative splicing was not due to transcriptional interference, but results from the antisense transcript selectively

interfering with splicing in a sequence dependent manner, suggesting that complimentary base pairing of these transcripts was responsible for excluding the final exon (Hastings et al., 2000; Hastings and Krainer, 2001).

More likely, secondary conformation of L1 RNA may be sufficient to prevent splicing. The RUNX3 transcription factor binding site is a potent transcriptional activator (Yang et al., 2003). Mutations within the splice donor sequence residing within the RUNX3 site results in a dramatic decrease of L1 promoter activity and retrotransposition (Yang et al., 2003) (Figure 2.2). Secondary L1 RNA structure(s) that sequester the splice donor thus preventing its utilization is likely the most efficient mechanism to evade splicing. *In silico* folding programs demonstrate that the splice donor is bound in secondary conformation. As mentioned above, secondary structure is likely affected by *trans*-acting factors that bind to *cis*-acting sequences in the RNA. For example, altered secondary structures of the self-splicing pre-rRNA intron in *Tetrahymena thermophiles* disrupts efficient splicing. However, in humans, it is unclear what precise role secondary structures play in ultimately producing the mature mRNA (Buratti and Baralle, 2004). Experimental evidence is required to determine what role secondary structure plays in L1 RNA processing.

Development of a screen to identify factors that repress genomic L1 expression

Chapter 3 and Appendix 2 describe my efforts to uncover new cellular factors that affect genomic expression of L1s. Numerous studies have attempted to identify proteins that interact with L1 at the level of the RNP (Dai et al., 2012; Goodier, 2016; Goodier et al., 2010; Goodier et al., 2015; Goodier et al., 2007; Hancks et al., 2011; Moldovan, 2015; Taylor et al., 2013; Zhang et al., 2014). These studies have been very successful and produced a plethora of candidate and verified genes that negatively, and sometimes positively influence L1 retrotransposition. Proteins that modulate expression of human L1 from a genomic context, however, require further study. The difficulty in studying the expression genomic L1s stems largely from the fact that L1s are present at ~500,000 different locations in the human genome. In order to begin to address the question of how genomic L1s are regulated, we built on the observation that the human embryonic carcinoma (hEC) cell line, PA-1, silences a reporter gene delivered by L1

retrotransposition (L1-reporter) (Garcia-Perez et al., 2010). My experiments are carried out with the assumption that L1 reporter gene silencing may be a proxy for silencing the entire L1 element.

I conducted a pilot study in hECs that employed a forward genetic screen to attempt to identify potential factors involved in L1-reporter gene silencing. The schematic of the screen is detailed in Chapter 3, but I will briefly describe its design here (Figure 3.1). We aimed to knock out ~19,000 genes in the PA-1 genome by taking advantage of a recently developed CRISPR/Cas9 genome editing platform (Ran et al., 2013; Sanjana et al., 2014; Shalem et al., 2014). This platform allows us to knock out a single gene in a single PA-1 cell via a lentiviral transduction that delivers a plasmid carrying the necessary components for gene knockout (Cas9 gene, single guide sequence, and puromycin resistance gene) (Figure 3.1). After knockout, we subjected the population of edited PA-1 cells to retrotransposition assays using an L1 that, upon retrotransposition and integration into the genome, expresses a protein conferring resistance to the drug G418 (Figure 3.3) (Moran et al., 1996). We reasoned that edited PA-1 cells resistant to G418 were lacking a gene product that may mediate L1-reporter gene silencing. Cells containing a knockout of a gene that had no effect on L1-reporter silencing would not be able to grow in the presence of G418. Thus, by using drug selection, we devised a method to potentially identify genes that are involved in L1-reporter gene silencing.

It is crucial to determine the kinetics of successful genome editing and gene knockout. I tested the efficiency of editing in PA-1 cells 7, 21, and 31 days post-transduction (Figure 3.1). I reasoned that knocking out genes essential for cell viability would result in the loss of guide sequences targeting those genes over time. Thus, I used sgRNA dropout as a proxy for editing efficiency. The data suggested guides targeting ~5,000 genes were negatively selected by 21 days post-transduction, representing putative essential genes. This number is likely an overestimate, as we lacked the sequencing depth to precisely determine guide drop out (Figure 3.2). Moreover, recently published reports suggest that ~10% of genes (~2,000) in surveyed cell types are essential for viability (Blomen et al., 2015; Hart et al., 2014; Hart et al., 2015; Wang et al., 2015). In the future, it will be necessary to increase our sequencing

coverage to more accurately define the efficiency of editing and sgRNA dropout in our population of cells.

Recently numerous groups have used CRISPR/Cas9 or other methods to identify essential genes in different cell types (Blomen et al., 2015; Hart et al., 2014; Hart et al., 2015; Wang et al., 2015). Roughly 2,610, and 2,676 genes were demonstrated to be essential for cell viability in a glioblastoma cell line (GBM) and a HeLa cell line, respectively (Hart et al., 2015). Interestingly, only 1,463 genes are common to both the GBM and HeLa cell lines, suggesting some heterogeneity in gene essentiality across cell types (Hart et al., 2015). I compared the list of essential genes generated in GBM and HeLa cells to the list of putatively essential genes in PA-1 cells. Of the ~5,000 putative essential genes identified in PA-1 cells, only 1,153 genes were demonstrated to be essential in GBM cells, and 1,081 were demonstrated to be essential in HeLa cells. Thus, despite the fact that the set of putative essential genes in PA-1 cells is likely inflated, we could use this knockout platform to identify essential genes specific to PA-1 cells.

Despite our sequencing experiments being underpowered, I did have evidence to suggest that PA-1 cells were edited by 21 days post-transduction (Figure 3.2). As this was a pilot study, I determined that by 21 days post-transduction, editing of PA-1 cells was sufficient to continue with our investigation (Figure 3.2, Figure 3.3). Using the day 21 edited PA-1 cell population; I next performed L1 retrotransposition assays in an attempt to identify candidate genes that silence an L1-reporter gene (Figure 3.3).

Identification and validation of candidate genes

One difficulty with performing a screen is prioritizing candidate genes. We decided on a straightforward prioritization method that resulted in 188 potential candidates (see Chapter three for our prioritization methods). Once candidate genes were identified, the next obstacle is determining which of those candidates to validate. The most logical approach is to first investigate the highest confidence genes as well as those that may have previously been indicated to influence the biological process in question. The two candidates at the top of our list are *Neurofirbomin 2* (*NF2*) and *Exportin 7* (*XPO7)* (Figure 3.3). *NF2* will be explored further later in this section.

The XPO7 protein has previously been demonstrated to mediate nuclear export of histone H2A and H3A variants, and loss of XPO7 led to a failure of chromatin condensation in erythroid cells (Hattangadi et al., 2014). It is unclear how loss of *XPO7* could result in depression of L1-reporter silencing in PA-1 cells. Evidence suggests histone deacetylases play a role in L1-reporter gene silencing (Garcia-Perez et al., 2010). One hypothesis is that loss of XPO7 in PA-1 cells prevents certain histone variants from entering the nucleus. Some histone variants may be involved with chromatin compaction and transcriptional silencing (Bonisch and Hake, 2012). Thus, it is plausible that some histone variants play a role in L1-reporter silencing, and preventing those variants from entering the nucleus results in an inability to silence an integrated L1 reporter.

Neurofirbomin 2 was the top candidate gene identified in my screen, so I decided to perform preliminary validation experiments. Somewhat to my surprise, orthogonal experiments in which I knocked out *NF2* in PA-1 cells and subsequently subjected those cells to retrotransposition assays resulted in the expression of an L1-reporter (Chapter 3, Figure 3.3). Additional experiments are necessary to validate the knockdown, but evidence suggests *NF2* may play a direct or indirect role in silencing an L1-reporter in PA-1 cells. Importantly, simply knocking out *NF2* does not result in drug resistance (Figure 3.3).

It is unclear how *NF2* might silence an L1-reporter in PA-1 cells. In humans, *NF2* is generally considered to be a tumor suppressor gene and loss of *NF2* results in the formation of tumors composed of Schwann cells (schwannomas) on cranial and peripheral nerves (Welling et al., 2007). Schwann cells act by myelinating and thus insulating the axons of peripheral nerves, which aids in the propagation of electrical impulses down the nerve fiber (Salzer, 2015). *NF2* is associated with a number of signaling pathways that generally control rates of cellular division and cellular size and has shown to be present in both the cytoplasm and nucleus of the cell (Cooper and Giancotti, 2014; Petrilli and Fernandez-Valle, 2016).

It is unclear if *NF2* directly regulates the cell cycle; however, it has been documented that expression of the *NF2* protein, MERLIN, in primary endothelial cells

results in cell cycle arrest (Beltrami et al., 2013; Li et al., 2010). Given that NF2 associates with cytoskeletal proteins, its effect on L1-reporter silencing may be indirect (Cooper and Giancotti, 2014; Petrilli and Fernandez-Valle, 2016). It might be possible that loss of *NF2*-mediated cell cycle control induces a stress response pathway (Figure 4.2). Some evidence suggests artificially inducing stress in HeLa cells leads to increased expression of endogenous L1s (Teneng et al., 2007).

In the nucleus, NF2 can bind to the ubiquitin ligase CRL4 which blocks CLR4 enzymatic activity (Cooper and Giancotti, 2014; Petrilli and Fernandez-Valle, 2016). CLR4 promotes expression of oncogenes including some that are associated with chromatin remodeling (Cooper and Giancotti, 2014; Petrilli and Fernandez-Valle, 2016). It is possible that loss of *NF2*, and thus unchecked CLR4 activity, results in an altered transcriptional profile that ultimately results in derepression of silencing an L1-reporter in PA-1 cells. Regardless of how *NF2* may be involved in silencing of an L1-reporter in PA-1 cells, the preliminary validation experiments suggest that our strategy can work (Figure 3.4).

Other candidate genes that warrant validation experiments

A number of additional candidate genes should also be considered for validation experiments. *Tumor protein 53* (*TP53*) was identified in our screen as a potential candidate (Figure 3.3). *TP53* has previously been implicated in negatively affecting L1 retrotransposition. Recently it was demonstrated that knockout of *tp53* in zebrafish resulted in an increase of engineered L1 retrotransposition (Wylie et al., 2016a; Wylie et al., 2016b). These same embryos showed an increase of endogenous ORF1p expression and also a loss of the repressive H3K9me3 chromatin modification at the 5'UTRs of endogenous L1s (Wylie et al., 2016a; Wylie et al., 2016b) (Figure 4.2). Additionally, human colon cancer tissues with *TP53* inactivating mutations expressed increased levels of endogenous, human specific L1s compared to colon cancers with intact *TP53* (Wylie et al., 2016a; Wylie et al., 2016b) Given these results, the potential role of *TP53*-mediated L1-reporter gene silencing in PA-1 cells warrants further investigation.

In chapter one I described the role of Krab-Zinc Finger (KZNF) protein-mediated L1 repression. This screen also identified a number of zinc finger candidate genes, including *ZNF177*, *ZNF320*, and *ZNF768* (Figure 3.3). Literature searches yielded no insights into the function of *ZNF320* or *ZNF768.* Previous reports have demonstrated *ZNF177* promoter hypermethylation in various cancers including endometrial cancer (Chen et al., 2015), lung cancer (Diaz-Lagares et al., 2016), and hepatocellular carcinomas (Kuo et al., 2014). Those studies identified hypermethylation of *ZNF177* as a biomarker for cancer and did not speculate on the effect *ZNF177* hypermethylation may have on the cancer phenotype. Given that increased L1 expression and retrotransposition has been associated with potentially driving cancer progression (see Chapter one), I propose a wild speculation: ZNF177 could actively represses transcription of genomic L1s, hypermethylation of the *ZNF177* promoter reduces cellular levels of the ZNF177 protein, thus relaxing restriction of L1 expression (Figure 4.2). Perhaps in some cancers, hypermethylation of *ZNF177* results in increased levels of L1 expression and L1-mediated mutagenesis. Clearly, that hypothesis is highly speculative, but the fact that other ZNF proteins have been shown to negatively effect L1 expression and retrotransposition warrants further investigation into these candidate genes.

I am confident that this pilot study laid the groundwork for future investigation. The list of genes uncovered in this preliminary study yielded some intriguing candidates that warrant additional investigation. Additionally, preliminary validation of the top ranked gene, *NF2*, suggests that the scheme we developed has the potential to identify new candidate genes that may be involved in L1-reporter silencing. These preliminary data warrant rigorous testing and additional validation of *NF2,* as well as other attractive candidate genes.

Concluding Remarks

My thesis has focused on dissecting the interaction between L1 and the human host. Primarily, I strove to understand how host factors negatively influence retrotransposition and how L1 has “escaped” repression from these factors. My work has led to a better understanding of how of post-transcriptional splicing inhibits L1 retrotransposition and hobbles continued L1 propagation. I also demonstrated that in

the absence of the canonical ORF1 methionine, downstream ORF1 methionine codons could be used for translation initiation, but these truncated proteins are non-functional protein. I demonstrated the feasibility of using a CRISPR/Cas9 genome wide screen in the PA-1 cell line and optimized experimental conditions to maximize its efficiency. I then used that screen in an attempt to identify novel proteins that silence an integrated L1-reporter. To my knowledge, no one has attempted to identify factors that silence expression of genomic L1s at this scale. Preliminary experiments potentially validated the top candidate gene as well as produced additional candidates that warrant follow-up. Like those experimentalists before me, I have stood on the shoulders of giants, and contributed new knowledge of human biology.

Figure 4.1: Evolution of the 5'UTR and its affect on splicing with the 5'UTR. This figure shows how splicing within the 5'UTR has changed over time due to deletions within the 5'UTR. The right axis denotes each L1 subfamily and in brackets is the approximate time (in million years ago (MYA)) that they were amplifying. The grey bar is the 5'UTR and black boxes within are sequences that are deleted in subsequent subfamilies. Black boxes are in the correct relative position. The splice donor (SD, red) and splice acceptors (SA, green) mentioned in Chapter 2 and 4 are indicated at their relative positions within the 5'UTR. The light color of the SA indicates its presence, but that it is not utilized in a splicing reaction. Of note is the transition from the L1PA3 to L1PA2 subfamily that resulted in the usage of a very conserved, but previously unused SA sequence.

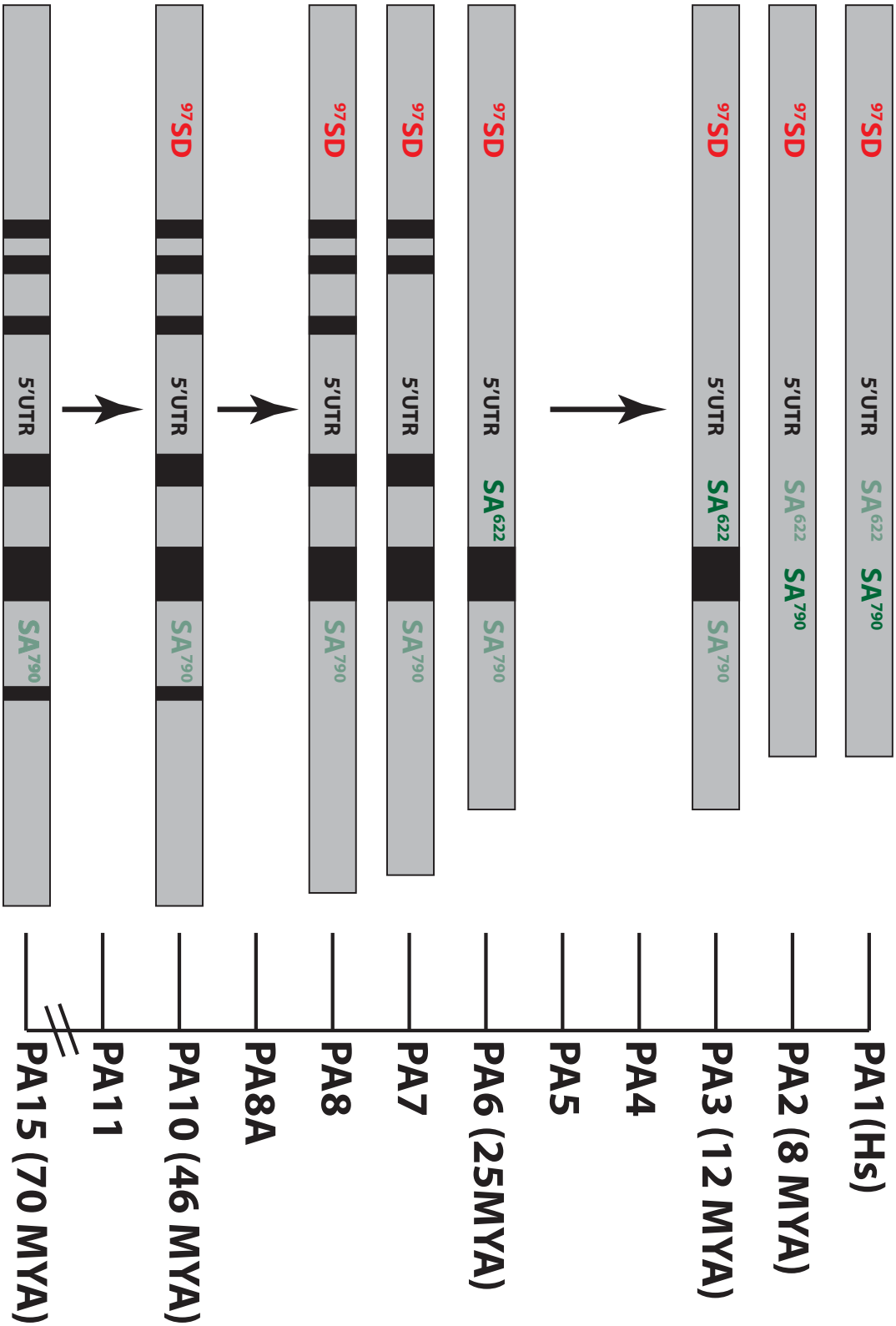

Figure 4.1: Evolution of the 5'UTR and its affect on splicing with the 5'UTR.

Figure 4.2: Model of how different factors might affect LINE-1 silencing in PA-1 cells.

This model details the two stages of silencing an L1 in PA-1 cells and the point at which candidate genes may influence that. The left half depicts the initiation of silencing. The black lines indicate genomic DNA. The wavy blue and yellow line indicates an L1 RNA that is being reverse transcribed by the ORF2p molecule an inserted via TPRT (blue circle). This structure is an intermediate during L1 target-site primed reverse transcription (See Chapter 1 and Figure 1.5). Samples of genes that may mediate the initiation of silencing are listed above (black box) and what cellular process they could be involved in concerning L1 silencing (red box). The right half of the model depicts a factor (black box) that may be associated with maintaining silencing of L1 and what cellular process it could be involved in (red box). L1 is shown integrated into genomic DNA, yellow and blue indicate ORF1 and ORF2 respectively).

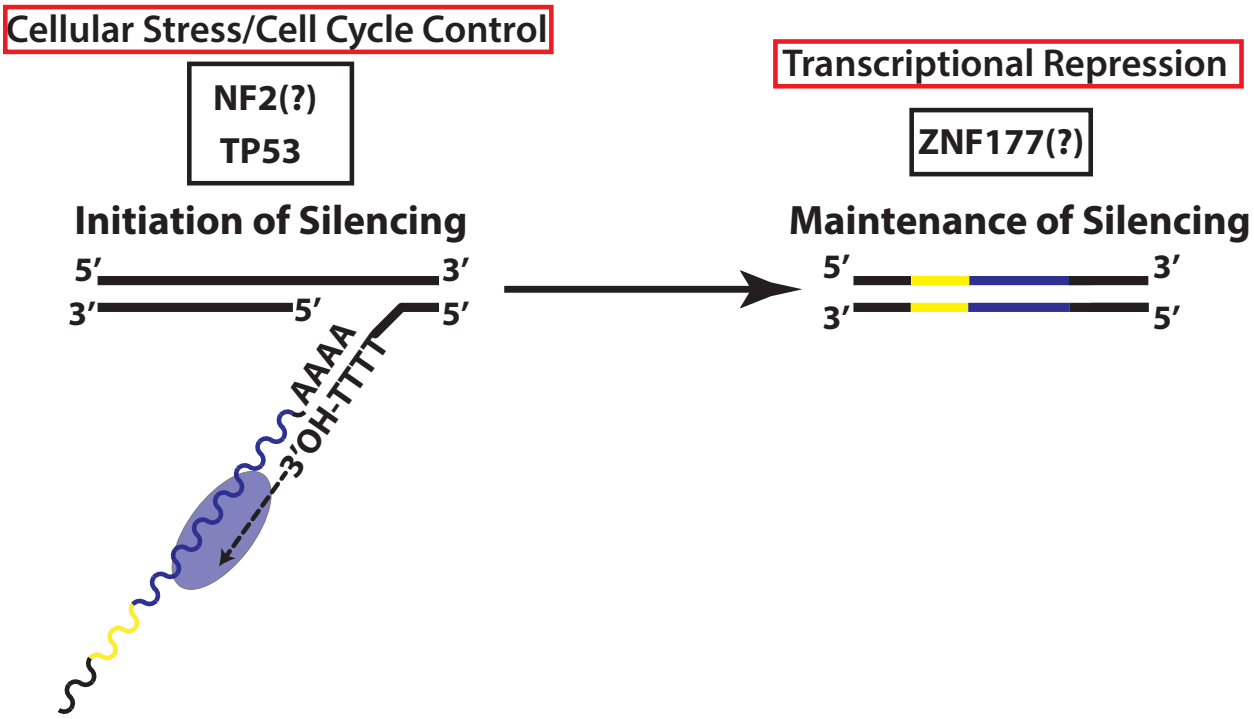

Figure 4.2: Model of how different factors might affect LINE-1 silencing in PA-1 cells.

References

Ameyar-Zazoua, M., Rachez, C., Souidi, M., Robin, P., Fritsch, L., Young, R., Morozova, N., Fenouil, R., Descostes, N., Andrau, J.C*., et al.* (2012). Argonaute proteins couple chromatin silencing to alternative splicing. Nature structural & molecular biology *19*, 998-1004.

Beck, C.R., Collier, P., Macfarlane, C., Malig, M., Kidd, J.M., Eichler, E.E., Badge, R.M., and Moran, J.V. (2010). LINE-1 retrotransposition activity in human genomes. Cell *141*, 1159-1170.

Belancio, V.P., Hedges, D.J., and Deininger, P. (2006). LINE-1 RNA splicing and influences on mammalian gene expression. Nucleic acids research *34*, 1512-1521.

Belancio, V.P., Roy-Engel, A.M., and Deininger, P. (2008). The impact of multiple splice sites in human L1 elements. Gene *411*, 38-45.

Belancio, V.P., Roy-Engel, A.M., Pochampally, R.R., and Deininger, P. (2010). Somatic expression of LINE-1 elements in human tissues. Nucleic acids research *38*, 3909-3922.

Beltrami, S., Kim, R., and Gordon, J. (2013). Neurofibromatosis type 2 protein, NF2: an uncoventional cell cycle regulator. Anticancer research *33*, 1-11.

Blomen, V.A., Majek, P., Jae, L.T., Bigenzahn, J.W., Nieuwenhuis, J., Staring, J., Sacco, R., van Diemen, F.R., Olk, N., Stukalov, A*., et al.* (2015). Gene essentiality and synthetic lethality in haploid human cells. Science *350*, 1092-1096.

Boissinot, S., and Furano, A.V. (2001). Adaptive evolution in LINE-1 retrotransposons. Molecular biology and evolution *18*, 2186-2194.

Boissinot, S., Roos, C., and Furano, A.V. (2004). Different rates of LINE-1 (L1) retrotransposon amplification and evolution in New World monkeys. Journal of molecular evolution *58*, 122-130.

Bonisch, C., and Hake, S.B. (2012). Histone H2A variants in nucleosomes and chromatin: more or less stable? Nucleic acids research *40*, 10719-10741.

Buratti, E., and Baralle, F.E. (2004). Influence of RNA secondary structure on the pre-mRNA splicing process. Molecular and cellular biology *24*, 10505-10514.

Chen, Y.C., Tsao, C.M., Kuo, C.C., Yu, M.H., Lin, Y.W., Yang, C.Y., Li, H.J., Yan, M.D., Wang, T.J., Chou, Y.C*., et al.* (2015). Quantitative DNA methylation analysis of selected genes in endometrial carcinogenesis. Taiwanese journal of obstetrics & gynecology *54*, 572-579.

Chiou, N.T., Shankarling, G., and Lynch, K.W. (2013). hnRNP L and hnRNP A1 induce extended U1 snRNA interactions with an exon to repress spliceosome assembly. Molecular cell *49*, 972-982.

Cho, H., Davis, J., Li, X., Smith, K.S., Battle, A., and Montgomery, S.B. (2014). High-resolution transcriptome analysis with long-read RNA sequencing. PloS one *9*, e108095.

Cooper, J., and Giancotti, F.G. (2014). Molecular insights into NF2/Merlin tumor suppressor function. Febs Lett *588*, 2743-2752.

Dai, L.X., Taylor, M.S., O'Donnell, K.A., and Boeke, J.D. (2012). Poly(A) Binding Protein C1 Is Essential for Efficient L1 Retrotransposition and Affects L1 RNP Formation. Molecular and cellular biology *32*, 4323-4336.

Damgaard, C.K., Kahns, S., Lykke-Andersen, S., Nielsen, A.L., Jensen, T.H., and Kjems, J. (2008). A 5' splice site enhances the recruitment of basal transcription initiation factors in vivo. Molecular cell *29*, 271-278.

Diaz-Lagares, A., Mendez-Gonzalez, J., Hervas, D., Saigi, M., Pajares, M.J., Garcia, D., Crujerias, A.B., Pio, R., Montuenga, L.M., Zulueta, J.*, et al.* (2016). A Novel Epigenetic Signature for Early Diagnosis in Lung Cancer. Clinical cancer research : an official journal of the American Association for Cancer Research *22*, 3361-3371.

Dumesic, P.A., and Madhani, H.D. (2013). The spliceosome as a transposon sensor. Rna Biol *10*, 1653-1660.

Dumesic, P.A., Natarajan, P., Chen, C., Drinnenberg, I.A., Schiller, B.J., Thompson, J., Moresco, J.J., Yates, J.R., 3rd, Bartel, D.P., and Madhani, H.D. (2013). Stalled spliceosomes are a signal for RNAi-mediated genome defense. Cell *152*, 957-968.

Furger, A., O'Sullivan, J.M., Binnie, A., Lee, B.A., and Proudfoot, N.J. (2002). Promoter proximal splice sites enhance transcription. Genes & development *16*, 2792-2799.

Garcia-Perez, J.L., Morell, M., Scheys, J.O., Kulpa, D.A., Morell, S., Carter, C.C., Hammer, G.D., Collins, K.L., O'Shea, K.S., Menendez, P.*, et al.* (2010). Epigenetic silencing of engineered L1 retrotransposition events in human embryonic carcinoma cells. Nature *466*, 769-773.

Goodier, J.L. (2016). Restricting retrotransposons: a review. Mobile DNA *7*, 16.

Goodier, J.L., Cheung, L.E., and Kazazian, H.H., Jr. (2013). Mapping the LINE1 ORF1 protein interactome reveals associated inhibitors of human retrotransposition. Nucleic acids research *41*, 7401-7419.

Goodier, J.L., Mandal, P.K., Zhang, L., and Kazazian, H.H., Jr. (2010). Discrete subcellular partitioning of human retrotransposon RNAs despite a common mechanism of genome insertion. Human molecular genetics *19*, 1712-1725.

Goodier, J.L., Pereira, G.C., Cheung, L.E., Rose, R.J., and Kazazian, H.H., Jr. (2015). The Broad-Spectrum Antiviral Protein ZAP Restricts Human Retrotransposition. PLoS genetics *11*, e1005252.

Goodier, J.L., Zhang, L., Vetter, M.R., and Kazazian, H.H., Jr. (2007). LINE-1 ORF1 protein localizes in stress granules with other RNA-binding proteins, including components of RNA interference RNA-induced silencing complex. Molecular and cellular biology *27*, 6469-6483.

Han, S.P., Tang, Y.H., and Smith, R. (2010). Functional diversity of the hnRNPs: past, present and perspectives. Biochem J *430*, 379-392.

Hancks, D.C., Goodier, J.L., Mandal, P.K., Cheung, L.E., and Kazazian, H.H., Jr. (2011). Retrotransposition of marked SVA elements by human L1s in cultured cells. Human molecular genetics *20*, 3386-3400.

Hart, T., Brown, K.R., Sircoulomb, F., Rottapel, R., and Moffat, J. (2014). Measuring error rates in genomic perturbation screens: gold standards for human functional genomics. Molecular systems biology *10*, 733.

Hart, T., Chandrashekhar, M., Aregger, M., Steinhart, Z., Brown, K.R., MacLeod, G., Mis, M., Zimmermann, M., Fradet-Turcotte, A., Sun, S.*, et al.* (2015). High-Resolution CRISPR Screens Reveal Fitness Genes and Genotype-Specific Cancer Liabilities. Cell *163*, 1515-1526.

Hastings, M.L., Ingle, H.A., Lazar, M.A., and Munroe, S.H. (2000). Post-transcriptional regulation of thyroid hormone receptor expression by cis-acting sequences and a naturally occurring antisense RNA. The Journal of biological chemistry *275*, 11507-11513.

Hastings, M.L., and Krainer, A.R. (2001). Pre-mRNA splicing in the new millennium. Current opinion in cell biology *13*, 302-309.

Hattangadi, S.M., Martinez-Morilla, S., Patterson, H.C., Shi, J., Burke, K., Avila-Figueroa, A., Venkatesan, S., Wang, J., Paulsen, K., Gorlich, D.*, et al.* (2014). Histones to the cytosol: exportin 7 is essential for normal terminal erythroid nuclear maturation. Blood *124*, 1931-1940.

Huang, C.F., and Zhu, J.K. (2014). RNA Splicing Factors and RNA-Directed DNA Methylation. Biology *3*, 243-254.

Jacobs, F.M., Greenberg, D., Nguyen, N., Haeussler, M., Ewing, A.D., Katzman, S., Paten, B., Salama, S.R., and Haussler, D. (2014). An evolutionary arms race between KRAB zinc-finger genes ZNF91/93 and SVA/L1 retrotransposons. Nature.

Khan, H., Smit, A., and Boissinot, S. (2006). Molecular evolution and tempo of amplification of human LINE-1 retrotransposons since the origin of primates. Genome research *16*, 78-87.

Koodathingal, P., and Staley, J.P. (2013). Splicing fidelity: DEAD/H-box ATPases as molecular clocks. Rna Biol *10*, 1073-1079.

Kuo, C.C., Lin, C.Y., Shih, Y.L., Hsieh, C.B., Lin, P.Y., Guan, S.B., Hsieh, M.S., Lai, H.C., Chen, C.J., and Lin, Y.W. (2014). Frequent methylation of HOXA9 gene in tumor tissues and plasma samples from human hepatocellular carcinomas. Clinical chemistry and laboratory medicine *52*, 1235-1245.

Li, W., You, L., Cooper, J., Schiavon, G., Pepe-Caprio, A., Zhou, L., Ishii, R., Giovannini, M., Hanemann, C.O., Long, S.B.*, et al.* (2010). Merlin/NF2 suppresses tumorigenesis by inhibiting the E3 ubiquitin ligase CRL4(DCAF1) in the nucleus. Cell *140*, 477-490.

Long, J.C., and Caceres, J.F. (2009). The SR protein family of splicing factors: master regulators of gene expression. Biochem J *417*, 15-27.

Moldovan, J.B. (2015). Identification of Cellular Host Factors That Associate With LINE-1 ORF1p and the Effect of the Zinc Finger Antiviral Protein Zap on LINE-1 Retrotransposition.

Moldovan, J.B., and Moran, J.V. (2015). The Zinc-Finger Antiviral Protein ZAP Inhibits LINE and Alu Retrotransposition. PLoS genetics *11*, e1005121.

Moran, J.V., Holmes, S.E., Naas, T.P., DeBerardinis, R.J., Boeke, J.D., and Kazazian, H.H., Jr. (1996). High frequency retrotransposition in cultured mammalian cells. Cell *87*, 917-927.

Pandit, S., Zhou, Y., Shiue, L., Coutinho-Mansfield, G., Li, H., Qiu, J., Huang, J., Yeo, G.W., Ares, M., Jr., and Fu, X.D. (2013). Genome-wide analysis reveals SR protein cooperation and competition in regulated splicing. Molecular cell *50*, 223-235.

Peddigari, S., Li, P.W., Rabe, J.L., and Martin, S.L. (2013). hnRNPL and nucleolin bind LINE-1 RNA and function as host factors to modulate retrotransposition. Nucleic acids research *41*, 575-585.

Perepelitsa-Belancio, V., and Deininger, P. (2003). RNA truncation by premature polyadenylation attenuates human mobile element activity. Nature genetics *35*, 363-366.

Petrilli, A.M., and Fernandez-Valle, C. (2016). Role of Merlin/NF2 inactivation in tumor biology. Oncogene *35*, 537-548.

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. Nat Protoc *8*, 2281-2308.

Salzer, J.L. (2015). Schwann cell myelination. Cold Spring Harbor perspectives in biology *7*, a020529.

Sanjana, N.E., Shalem, O., and Zhang, F. (2014). Improved vectors and genome-wide libraries for CRISPR screening. Nature methods *11*, 783-784.

Semlow, D.R., and Staley, J.P. (2012). Staying on message: ensuring fidelity in pre-mRNA splicing. Trends in biochemical sciences *37*, 263-273.

Shalem, O., Sanjana, N.E., Hartenian, E., Shi, X., Scott, D.A., Mikkelsen, T.S., Heckl, D., Ebert, B.L., Root, D.E., Doench, J.G.*, et al.* (2014). Genome-scale CRISPR-Cas9 knockout screening in human cells. Science *343*, 84-87.

Soifer, H.S., Zaragoza, A., Peyvan, M., Behlke, M.A., and Rossi, J.J. (2005). A potential role for RNA interference in controlling the activity of the human LINE-1 retrotransposon. Nucleic acids research *33*, 846-856.

Taylor, M.S., Lacava, J., Mita, P., Molloy, K.R., Huang, C.R., Li, D., Adney, E.M., Jiang, H., Burns, K.H., Chait, B.T.*, et al.* (2013). Affinity proteomics reveals human host factors implicated in discrete stages of LINE-1 retrotransposition. Cell *155*, 1034-1048.

Teneng, I., Stribinskis, V., and Ramos, K.S. (2007). Context-specific regulation of LINE-1. Genes to cells : devoted to molecular & cellular mechanisms *12*, 1101-1110.

Valencia, P., Dias, A.P., and Reed, R. (2008). Splicing promotes rapid and efficient mRNA export in mammalian cells. Proceedings of the National Academy of Sciences of the United States of America *105*, 3386-3391.

Wahl, M.C., Will, C.L., and Luhrmann, R. (2009). The spliceosome: design principles of a dynamic RNP machine. Cell *136*, 701-718.

Wang, T., Birsoy, K., Hughes, N.W., Krupczak, K.M., Post, Y., Wei, J.J., Lander, E.S., and Sabatini, D.M. (2015). Identification and characterization of essential genes in the human genome. Science *350*, 1096-1101.

Wei, W., Gilbert, N., Ooi, S.L., Lawler, J.F., Ostertag, E.M., Kazazian, H.H., Boeke, J.D., and Moran, J.V. (2001). Human L1 retrotransposition: cis preference versus trans complementation. Molecular and cellular biology *21*, 1429-1439.

Welling, D.B., Packer, M.D., and Chang, L.S. (2007). Molecular studies of vestibular schwannomas: a review. Current opinion in otolaryngology & head and neck surgery *15*, 341-346.

Witherspoon, D.J., Marchani, E.E., Watkins, W.S., Ostler, C.T., Wooding, S.P., Anders, B.A., Fowlkes, J.D., Boissinot, S., Furano, A.V., Ray, D.A.*, et al.* (2006). Human population genetic structure and diversity inferred from polymorphic L1(LINE-1) and Alu insertions. Hum Hered *62*, 30-46.

Wylie, A., Jones, A.E., and Abrams, J.M. (2016a). p53 in the game of transposons. BioEssays : news and reviews in molecular, cellular and developmental biology *38*, 1111-1116.

Wylie, A., Jones, A.E., D'Brot, A., Lu, W.J., Kurtz, P., Moran, J.V., Rakheja, D., Chen, K.S., Hammer, R.E., Comerford, S.A.*, et al.* (2016b). p53 genes function to restrain mobile elements. Genes & development *30*, 64-77.

Yang, N., and Kazazian, H.H., Jr. (2006). L1 retrotransposition is suppressed by endogenously encoded small interfering RNAs in human cultured cells. Nature structural & molecular biology *13*, 763-771.

Yang, N., Zhang, L., Zhang, Y., and Kazazian, H.H., Jr. (2003). An important role for RUNX3 in human L1 transcription and retrotransposition. Nucleic acids research *31*, 4929-4940.

Zhang, A., Dong, B., Doucet, A.J., Moldovan, J.B., Moran, J.V., and Silverman, R.H. (2014). RNase L restricts the mobility of engineered retrotransposons in cultured human cells. Nucleic acids research *42*, 3803-3820.

Appendix 1

LINE-1 RNP associated splicing factors

Appendix 1 addresses my investigation of splicing factors that associate with L1 ORF1p. More than twenty splicing factors have been reported to associate with the L1 ORF1p by immunoprecipitation followed by tandem mass spectrometry experiments and immunofluorescent co-localization studies (Goodier et al., 2013; Moldovan and Moran, 2015; Taylor et al., 2013). In cell culture based retrotransposition assays, Goodier *et al.,* 2013 demonstrated that overexpression of eight splicing factors dramatically decreased retrotransposition of an engineered human L1 in 293T cells (Goodier et al., 2013).

One splicing factor that has been demonstrated to associate with L1 ORF1p is the heterogenous ribonucleoprotein L (hnRNPL) (Goodier et al., 2013; Moldovan, 2015; Moldovan and Moran, 2015; Peddigari et al., 2013). The hnRNPL splicing factor is an RNA binding protein that affects alternative splicing as well as poly-adenylation site selection in RNAs derived from protein coding genes (Hung et al., 2008). RNA pull-down experiments demonstrated hnRNPL physically associated with L1 RNA in mouse cells (Peddigari et al., 2013). Additionally, siRNA mediated knockdown of hnRNPL in HeLa-JVM cells resulted in an increase in engineered L1 retrotransposition (Moldovan, 2015). By comparison, overexpression of hnRNPL resulted in a decrease in engineered L1 retrotransposition in human 293T cells (Goodier et al., 2013) Importantly, not all splicing proteins inhibit retrotransposition. For example, overexpression of other hnRNP proteins like hnRNPQ, hnRNPK, and hnRNPR, had little affect on engineered L1 retrotransposition in human 293T cells (Goodier et al., 2013) These data suggest that some, but not all, splicing factors interact with L1 RNA and may modulate L1 retrotransposition (Goodier et al., 2013).

Other splicing factors, like the family of serine/arginine rich splicing factors (SRSF), also interact with L1 ORF1p (Goodier et al., 2013; Moldovan and Moran, 2015). Among the SRSF family of proteins, ORF1p co-immunoprecipitated with SRSF1, SRSF7, and SRSF10 (Goodier et al., 2013; Moldovan and Moran, 2015), which are involved in both constitutive and alternative mRNA splicing (Twyffels et al., 2011). In addition, SRSF proteins are also found to co-localize with cytoplasmic stress granules (SGs), which sequester mRNA (Twyffels et al., 2011). Indeed, co-expression experiments indicate that SRSF1 and L1 ORF1p can co-localize with SG markers in 293T cells (Goodier et al., 2013).

We and others have previously demonstrated that spliced L1 mRNAs are present in the human genome [*e.g.*, see Chapter 2 and (Belancio et al., 2006; Belancio et al., 2008; Belancio et al., 2010)]. We also determined that splicing-mediated deletions of L1 RNA negatively affect retrotransposition of an engineered human L1 in a cell culture based assay (Chapter 2). Thus, I sought to determine whether the overexpression of the splicing factors hnRNPL, hnRNPQ, SRSF1, and SRSF7, affected retrotransposition of an engineered L1 in HeLa-JVM cells.

Results

Overexpression of the splicing factors hnRNPL, SRSF1, and SRSF7 repress L1 reporter retrotransposition.

We first sought to determine if the overexpression of hnRNPL, hnRNPQ, SRSF1, or SRSF7 affect L1 retrotransposition. To accomplish this goal, we utilized a transient cell culture L1 retrotransposition assay (Moran et al., 1996; Wei et al., 2001). The L1 retrotransposition assay relies on an episomal, engineered L1 expression construct containing a retrotransposition indicator cassette within the L1 3'UTR. The retrotransposition indicator cassette consists of an anti-sense copy of the neomycin phosphotransferase gene (*mneoI)* (Freeman et al., 1994; Kopera et al., 2011; Moran et al., 1996). The coding sequence of *mneoI* is interrupted by either a γ−globin intron or a *Tetrahymena thermophilus* self-splicing group I intron (Esnault et al., 2002; Moran et al., 1996) that resides in the same transcriptional orientation as L1. This arrangement ensures that a functional neomycin phosphotransferase gene will only be activated

upon L1 retrotransposition into genomic DNA (Freeman et al., 1994; Kopera et al., 2011; Moran et al., 1996). Retrotransposition of *mneoI* confers resistance to the drug G418. Counting the resultant drug resistant foci provide a quantitative readout of L1 retrotransposition.

HeLa-JVM cells were transiently co-transfected with a retrotransposition-competent L1 (RC-L1; pJM101/L1.3) and a cDNA plasmid that expresses *hnRNPL*, *hnRNPQ*, *SRSF1-ASF1* (full-length isoform), *SRSF1-ASF3* (short isoform lacking the final coding exon), or *SRSF7*. Retrotransposition efficiency of each co-transfection was normalized to empty vector (pCEP4) co-transfected with the RC-L1 (Moldovan and Moran, 2015; Richardson et al., 2014).

To control for off-target effects of cDNA overexpression, HeLa-JVM cells were co-transfected with the pCEP4 empty vector, or a cDNA expression plasmid that constitutively expresses the *neomycin phosphotransferase* gene (pCDNA3). This control tests for potentially toxic side effects of cDNA overexpression and is crucial to determine if cDNA overexpression affects L1 retrotransposition or cell viability/growth (Moldovan and Moran, 2015). Resultant G418-resistant colonies provide a quantitative readout of the effect of cDNA overexpression on cell viability in the presence of G418 and was used normalize retrotransposition efficiency for all assays.

Co-transfection of an RC-L1 and phnRNPL led to a ~75% decrease in L1 retrotransposition of RNAs when compared to control levels (Figure A2.A), which is in agreement with other reports (Goodier et al., 2013; Moldovan and Moran, 2015; Peddigari et al., 2013). Similarly, co-transfection of an RC-L1 and either pSRSF1-ASF1 or pSRSF-ASF3 led to an ~80% or ~70%, reduction in L1 retrotransposition when compared to controls (Figure A2.A). By comparison, co-transfection of an RC-L1 and a pSRSF7 led to a ~55% decrease in L1 retrotransposition when compared to controls (Figure A2.A). In agreement with previous studies, co-transfection of an RC-L1 and phnRNPQ led to a ~10% decrease in L1 retrotransposition when compared to control levels (Figure A1.1). As a positive control, we observed that co-transfection of an RC-L1 and an *APOBEC3* cDNA expression plasmid (pk_A3A) led to a ~95% reduction in L1 retrotransposition (Figure A1.A) (Bogerd et al., 2006; Moldovan and Moran, 2015; Muckenfuss et al., 2006; Richardson et al., 2014).

The L1 retrotransposition assay relies on splicing of the intron in the *mneoI* cassette (Moran et al., 1996). It is possible that overexpression of splicing factors disrupts the splicing of the intron from the *mneoI* cassette, leading to a false positive result. To examine this possibility, we utilized an RC-L1 that contains a self-splicing group I intron (pL1.3*neo*^{III}) within the *mneoI* cassette (Esnault et al., 2002; Garcia-Perez et al., 2007). Co-transfection of pL1.3*neo*^{III} with the phnRNPL, phnRNPQ, pSRSF1-ASF1, pSRSF1-ASF3, or pSRSF7 cDNA expression plasmids led to a slightly greater decrease in L1 retrotransposition (Figure A2.B) when compared to similar co-transfection experiments conducted with pJM101/L1.3 (Figure A1.B). Once again, control experiments revealed that co-transfection of pL1.3*neo*^{III} and pk_A3A drastically reduced L1 retrotransposition (Figure A1.B).

Previous reports have demonstrated that overexpression of some cDNAs affects the accumulation of various L1 RNA species in transfected cells (Moldovan, 2015; Moldovan and Moran, 2015). In an effort to determine how hnRNPL reduces RC-L1 retrotransposition, we co-transfected HeLa-JVM cells with pJM101/L1.3, phnRNPL, phnRNPQ, or with the empty pCEP4 vector. Polyadenylated RNAs were collected from whole cell lysates of transfected cells, separated on a denaturing agarose gel, transferred to a nylon membrane, and analyzed by northern blot (see methods). To detect L1 RNAs, we generated a strand-specific riboprobe complementary to nucleotides +7-99 of the L1 5'UTR (Belancio et al., 2006; Moldovan, 2015; Moldovan and Moran, 2015) (Figure A1C).

Full-length RNAs derived from pJM101/L1.3 are expected to be over six kilobases (kb) in length (Moldovan, 2015; Moldovan and Moran, 2015). Indeed, we readily detected full-length L1 RNA in cells that were co-transfected with pJM101/L1.3 and pCEP4 or the phnRNPQ (Figure A1.C, black arrow). In contrast, we observed a stark reduction in full-length L1 RNA in cells co-transfected with pJM101/L1.3 and phnRNPL (Figure A1.C, black arrow). Additionally, we observed a reduction in RNAs ranging from ~1.5 kb to ~3 kb in phnRNPL co-transfected cells compared to those transfected with pCEP4 or phnRNPQ (Figure A1.C, red bar). Finally, in cells transfected with pJM101/L1.3 and phnRNPL, an RNA species just above ~3 kb is present which is not present in other transfection conditions (Figure A1.C, red arrow). Untransfected

HeLa-JVM did not express detectable levels of full-length L1 RNA and served as a negative control (Figure A1.C). Additional control experiments demonstrated that overexpression of cDNAs did not affect levels of β-actin RNA (Figure A1.C).

Discussion

The above results suggest that overexpression of hnRNPL in HeLa-JVM cells represses retrotransposition of RC-L1 derived RNAs, in part, by reducing full-length L1 RNA levels. As endogenous L1 RNAs are spliced and subject to premature poly-addenylation, we hypothesize that hnRNPL overexpression may also modulate alternative splicing and or pre-mature poly-addenylation of L1 RNA (Chapter2; (Belancio et al., 2006; Belancio et al., 2008; Perepelitsa-Belancio and Deininger, 2003).

The mechanism of how hnRNPL overexpression decreases full-length transcript of L1 RNA is not known. Previous reports suggest hnRNPL can affect poly(A) site selection in mRNAs (Hung et al., 2008). We propose that this process may be a contributing mechanism to our observation that overexpression of hnRNPL decreases levels of full-length L1 RNA from a transfected L1. The presence of a ~3.5 kb RNA species in phnRNPL co transfected cells which was not present in our other experimental conditions could be indicative of a change in poly(A) splice site selection, leading to a decrease in full-length RNA (Figure A1.C). We also observed globally decreased signals of RNA in phnRNPL transfected cells, thus, it is possible overexpression of phnRNPL could also decrease absolute levels of cellular L1 RNA. Further experiments are needed to test the validity of these hypotheses.

It is unclear how overexpression of SRSF proteins inhibits engineered L1 retrotransposition. The SRSF proteins may act by directly modulating alternative splicing of L1 RNA, resulting in potentially retrotransposition-deficient RNAs. Alternatively, stalled or slowed spliceosomes have been demonstrated to signal an RNAi response in the yeast *Cryptoccocus neoformans* (Dumesic et al., 2013). RNAi-based mechanisms of gene silencing rely on short (21-31 nt) RNA sequences that are loaded onto Argonaute proteins forming an RNA-induced silencing complex (RISC). In HeLa cells, Argonaute proteins have been observed to associate with the SRSF 1, 3, 7, and 10 proteins (Ameyar-Zazoua et al., 2012). One hypothesis is that splicing L1 RNA is inefficient, resulting in stalled or slowed spliceosomes. Inefficient splicing of L1 RNA

at genomic loci could potentially recruit RNAi machinery, via SRSF proteins, that would then lead to RNAi induced transcriptional repression (Huang and Zhu, 2014). This hypothesis is highly speculative and requires further investigation.

SRSF1 also has been observed to co-localize with ORF1p in the cytoplasm (Goodier et al., 2013), and SRSF1 is postulated to play a role in sequestering mRNAs to cytoplasmic stress granules (SGs) (Twyffels et al., 2011). Additionally, ORF1p has been demonstrated to associate with SG markers in cell culture and it has been suggested that the ORF1p-SG association may be a mechanism to inhibit L1 retrotransposition (Goodier et al., 2013; Moldovan and Moran, 2015). Thus, SG markers, ORF1p, and the SRSF1 protein have all been demonstrated to co-localize to similar cytoplasmic foci; however, to our knowledge no experiments have demonstrated all three simultaneously co-localizing.

Given co-localization experiments, we favor the hypothesis that SRSF-mediated repression of L1 retrotransposition in cell culture assays is independent of splicing; however, we cannot rule out that possibility. It is possible that SRSF proteins interact with cytoplasmic L1 ORF1p molecules and facilitate their sequestration into SGs. Sequestration could subsequently prevent completion of the retrotransposition cycle. Additional experimentation is necessary to determine where in the retrotransposition cycle the SRSF proteins may act, but we propose investigating their role in the cytoplasm is the most logical place to begin.

Methods

Cell Culture

HeLa-JVM cells were cultured in high glucose Dulbecco's Modified Eagle Medium (DMEM) lacking pyruvate (Invitrogen). DMEM was supplemented with 10% fetal bovine calf serum (FBS) and 1X penicillin/streptomycin/glutamine (Invitrogen) to create DMEM-complete medium as described previously (Moran et al., 1996). HeLa-JVM cells were grown in a humidified tissue culture incubator (Thermo Scientific, Waltham MA) at 37^{o}C in the presence of 7% CO_2.

Plasmids Used

pJM101/L1.3: contains a full-length version of L1.3 in the pCEP4 backbone. The 3'UTR of L1.3 contains the *mneoI* retrotransposition indicator cassette (Dombroski et al., 1993; Moran et al., 1996; Sassaman et al., 1997).

pCEP4: mammalian fosmid-based expression vector obtained from ThermoFisher Scientific.

pCDNA3: mammalian expression plasmid that expresses the neomycin phosphotransferase resistance gene and was obtained from Invitrogen.

phnRNPL: plasmid expressing hnRNPL and was obtained from Open Biosystems (6174088).

phnRNPQ: plasmid expressing hnRNPQ and was obtained from Open Biosystems (5495201).

pSRSF1-ASF1: plasmid expressing canonical isoform (ASF1) (MHS6278-20275537) and was obtained from ThermoFisher Scientific.

pSRSF1-ASF3: plasmid containing a shorter isoform (ASF3) lacking the final exon (MHS1010-202800565) and obtained from ThermoFisher Scientific.

pSRSF7: plasmid expressing SRSF7 (MHS6278-202755457) and obtained from ThermoFisher Scientifc.

pL1.3neo^{III}: mammalian fosmid-based expression vector that is similar to pJM101/L1.3 except contains a *Tetrahymena thermophillus* self splicing group I intron in the *mneoI* cassette (Esnault et al., 2002).

Retrotransposition Assay

To monitor L1 retrotransposition, we used a modified version of a previously described transient transfection assay (Kopera et al., 2016; Moran et al., 1996; Wei et al., 2001). Briefly, 2×10^3 HeLa-JVM cells were plated in each well of a 6-well tissue culture dish (BD Biosciences). Approximately 24 hours post-plating, each well was co-transfected using a mixture containing 100 µl Opti-MEM® (Life Technologies), 3 µl FuGENE6 (Promega) transfection reagent, and .5 µg pJM101/L1.3 plasmid DNA or pL1.3*NEO^{III}* and .5 µg of a cDNA expression plasmid, or.5 µg of the pCEP4 empty vector plasmid. Approximately 24 hours post-transfection, the media was replaced with DMEM-

complete medium to stop the transfection. Seventy-two hours post-transfection DMEM-complete medium supplemented with 400 µg/mL of G418 (Life Technologies) was added to cells to select for retrotransposition events. The medium was replaced daily. Fourteen days post transfection the resultant G418-resistant foci were washed with 1X Phosphate-Buffered Saline (PBS), fixed for 10 minutes at room temperature in a 1X PBS solution containing 2% paraformaldehyde (Sigma Aldrich) and 0.4% glutaraldehyde (Sigma Aldrich), then stained with a 0.1% crystal violet solution for 30 minutes at room temperature to visualize the G418-resistant foci. To control for toxicity and off target effects, HeLa-JVM cells were plated at a density of $2X10^3$ cells/well in each well of a 6-well tissue culture dish (BD Biosciences). Transfections were performed exactly as described above except 0.5 µg of pCDNA3 plasmid DNA was co-transfected with 0.5 µg of cDNA expression plasmid or 0.5 µg of the pCEP4 empty vector plasmid. The pcDNA3 control assay was treated with G418 as described above for the retrotransposition assay. Three biological replicates were performed for each retrotransposition assay. Images were acquired by scanning the plates using a CanoScan. Error bars on all retrotransposition assays represent standard deviation of technical triplicates from the indicated experiment. Three biological replicates were performed for each assay.

RNA isolation

Briefly, $8X10^6$ HeLa-JVM cells were plated into a T-175 Falcon tissue culture flask (BD Biosciences). Approximately twenty-four hours post-plating, each flask was co-transfected using a mixture containing 1 mL of Opti-MEM® (Life Technologies), 120µl of the FuGene HD transfection reagent (Promega), and 10 µg of pJM101/L1.3 plasmid DNA and 10 µg cDNA plasmid DNA or 10 µg pCEP4 empty vector. The tissue culture medium was changed 24 hours post-transfection. Forty-eight hours post transfection the cells were washed in ice-cold 1X phosphate buffered saline (PBS) and collected in 2ml of PBS using a cell scraper (BD Falcon). Cells were subsequently transferred to a 15 mL conical tube (BD Biosciences), and centrifuged at 3000 x g for 5 minutes at 4^oC. Cell pellets were frozen at -20^oC overnight. The frozen pellets were thawed and total RNA was prepared using the TRIzol reagent following the protocol provided by the

manufacturer (Life Technologies). Poly(A) RNAs then were isolated from the total RNAs using a Oligotex mRNA Midi Kit (Qiagen), suspended in UltraPure™ DNase/RNase-Free distilled water (Thermo Fisher Scientific), and quantified using a NanoDrop 1000 spectrophotometer (Thermo Fisher Scientific).

Preparation of Northern Blot Probes

Briefly, oligonucleotide primers were used to PCR amplify portions of the L1.3 5'UTR (Moldovan and Moran, 2015)(L1.3 nucleotides 7-99. Notably, a T3 RNA polymerase promoter sequence was included on the reverse primer (L1.3 5'UTR 7-99 AS, underlined below). The resultant PCR products were separated on a 1% agarose gel and purified using QIAQuik gel extraction (Qiagen) following the manufacturers protocol. Strand-specific αP^{32}-UTP radiolabeled riboprobes were generated using the MAXIscript T3 system (Thermo Fisher Scientific).The labeling reaction was carried out at 37^{o}C using the following reaction conditions: 500ng of gel purified DNA template, 2 μL of transcription buffer supplied by the manufacturer, 1 μL each of unlabeled 10 mM ATP, CTP, GTP, 5 μL of αP^{32}-UTP (10 mCi/mL), and 2 μL of T3 RNA polymerase. The reaction components then were mixed and brought to a total volume of 20 μL using nuclease-free water in a 1.5 mL Eppendorf tube, which was incubated at 37^{o}C for 10 minutes in a heating block. Unincorporated nucleotides were subsequently depleted using the Ambion® NucAway™ Spin Columns (Thermo Fisher Scientific) following the protocol provided by the manufacturer. To generate a control β -actin riboprobe, the pTRI-β-actin-125-Human Antisense Control Template (Applied Biosystems) was used in T3 labeling reactions. Biological duplicates of each northern blot exhibited similar results.

Oligonucleotide sequences used to generate northern blot probes:

L1.3 5'UTR 7-99 Sense: 5'-GGAGCCAAGATGGCCGAATAGGAACAGCT-3'

L1.3 5'UTR 7-99 AS: 5'-AATTAACCCTCAAAGGGACCTCAGATGGAAATGCAG-3'

Northern Blots

Northern blot experiments were conducted using the NorthernMax-Gly Kit (Thermo Fisher Scientific) following the protocol provided by the manufacturer. Briefly, aliquots of poly(A) RNAs (2 μg) were incubated for 30 minutes at 50°C in Glyoxal Load Dye (containing DMSO and ethidium bromide) and then were separated on a 1.2% agarose gel for roughly 2 hours at 100V. The RNAs were transferred by capillary action to a Hybond-N nylon membrane (GE Healthcare) for four hours. After transfer membranes were UV cross-linked to the membrane using the Optimum Crosslink setting of a Stratalinker (Stratagene,). Membranes were then baked at 80°C for 15 minutes. Membranes were prehybridized for approximately four hours at 68°C in NorthernMax® Prehybridization/Hybridization Buffer (Thermo Fisher Scientific) and then were incubated overnight at 68°C with a strand specific RNA probe (final concentration of probe ~ 3×10^6 cpm/ml). The following day, the membranes were washed once with low stringency wash solution (2x saline sodium citrate (SSC), 0.1% sodium dodecylsulfate (SDS)) and then twice with high stringency wash solution (0.1x SSC, 0.1% SDS). The washed membranes were placed in a film cassette (Thermo Fisher Scientific, Autoradiography Cassette FBCA 57) and exposed to Amersham Hyperfilm ECL (GE Healthcare) overnight at -80°C. Films were developed using a JP-33 X-Ray Processor (JPI America Inc.).

Figure A1: Some splicing factors inhibit retrotransposition of an L1 reporter and overexpression of hnRNPL decreases full length L1 reporter transcript.
A) *Results from a representative retrotransposition assy*. The x-axis indicates the cDNA plasmid, or empty vector (pCEP4) co-transfected with pJM101/L1.3. The y-axis indicates retrotransposition efficiency (%). Relative retrotransposition efficiencies are normalized to pJM101/L1.3 co-transfected with pCEP4. Images and data are representative of one experiment and error bars indicate standard deviation of technical triplicates from the indicated data. B) *Results from a representative retrotransposition assay.* The same as (A) except the L1 retrotransposition plasmid is pL1.3*Neo*^{III}*,* which contains a *Tetrahymena thermophilus* self-splicing group I intron. Images and data are representative of one experiment and error bars indicate standard deviation of technical triplicates from the indicated data. C) *Results from a representative Northern blot.* Northern blot of RNA collected from transfected HeLa-JVM cells. Top panel, cells were co-transfected with pJM101/L1.3 and a cDNA expression plasmid (listed above) or empty vector (pCEP4). The final lane is RNA from untransfected HeLa-JVM cells (UTF). The riboprobe recognizes bp 7-99 of the L1 5'UTR (see methods). The black arrow indicates the predicted size of full-length L1 RNA. The red bar indicates a constellation of RNAs that increase in intensity when pJM101/L1.3 is cotransfected with phnRNPL (second lane), but not phnRNPQ (third lane) or pCEP4 (first lane). Panels below are actin loading controls. Size markers are indicated on the left. The experiment was performed twice with similar results.

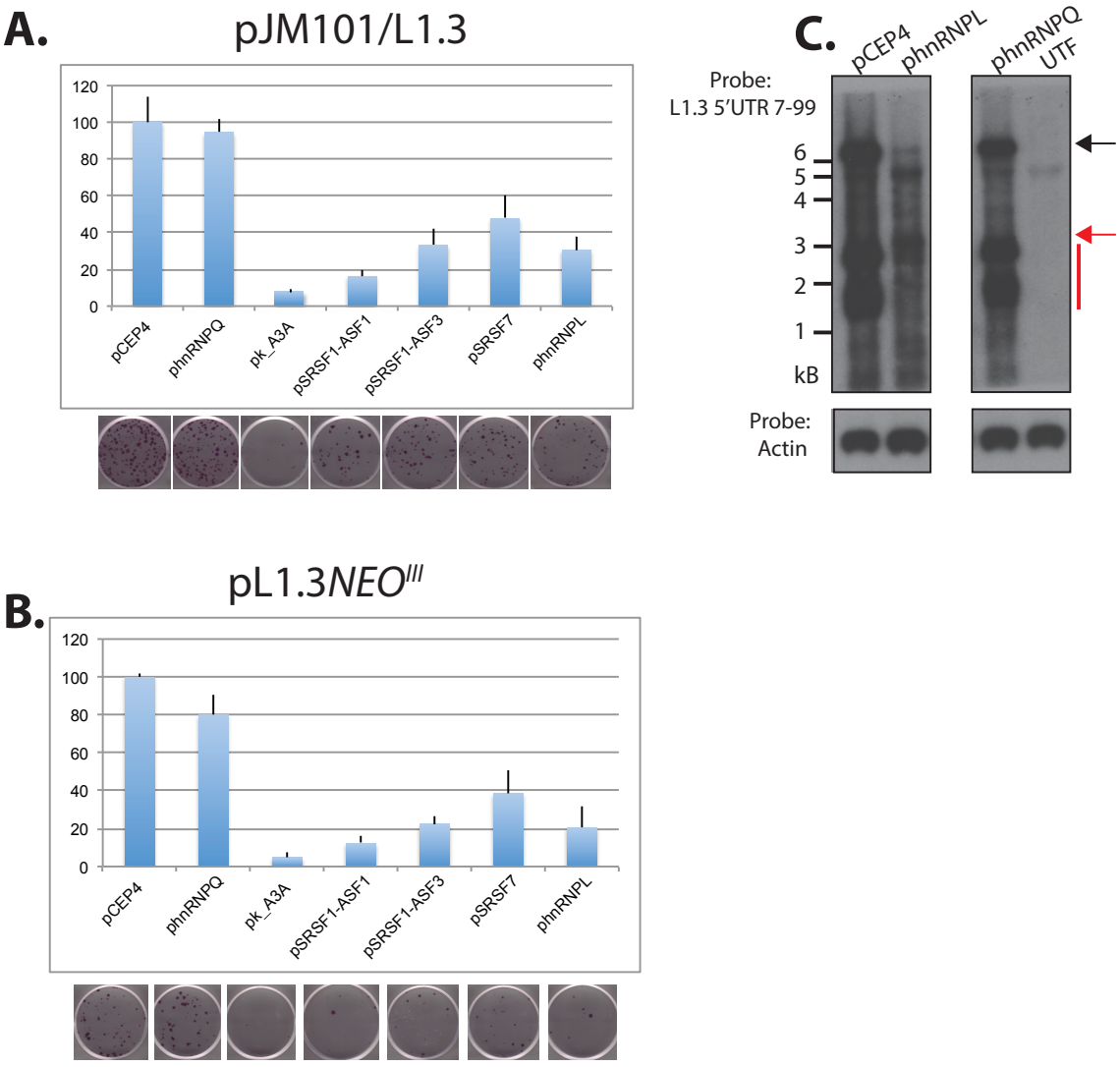

Figure A1: Some splicing factors inhibit retrotransposition of an L1 reporter and overexpression of hnRNPL decreases full length L1 reporter transcript.

References

Ameyar-Zazoua, M., Rachez, C., Souidi, M., Robin, P., Fritsch, L., Young, R., Morozova, N., Fenouil, R., Descostes, N., Andrau, J.C*., et al.* (2012). Argonaute proteins couple chromatin silencing to alternative splicing. Nature structural & molecular biology *19*, 998-1004.

Belancio, V.P., Hedges, D.J., and Deininger, P. (2006). LINE-1 RNA splicing and influences on mammalian gene expression. Nucleic acids research *34*, 1512-1521.

Belancio, V.P., Roy-Engel, A.M., and Deininger, P. (2008). The impact of multiple splice sites in human L1 elements. Gene *411*, 38-45.

Belancio, V.P., Roy-Engel, A.M., Pochampally, R.R., and Deininger, P. (2010). Somatic expression of LINE-1 elements in human tissues. Nucleic acids research *38*, 3909-3922.

Bogerd, H.P., Wiegand, H.L., Hulme, A.E., Garcia-Perez, J.L., O'Shea, K.S., Moran, J.V., and Cullen, B.R. (2006). Cellular inhibitors of long interspersed element 1 and Alu retrotransposition. Proceedings of the National Academy of Sciences of the United States of America *103*, 8780-8785.

Dombroski, B.A., Scott, A.F., and Kazazian, H.H., Jr. (1993). Two additional potential retrotransposons isolated from a human L1 subfamily that contains an active retrotransposable element. Proceedings of the National Academy of Sciences of the United States of America *90*, 6513-6517.

Dumesic, P.A., Natarajan, P., Chen, C., Drinnenberg, I.A., Schiller, B.J., Thompson, J., Moresco, J.J., Yates, J.R., 3rd, Bartel, D.P., and Madhani, H.D. (2013). Stalled spliceosomes are a signal for RNAi-mediated genome defense. Cell *152*, 957-968.

Esnault, C., Casella, J.F., and Heidmann, T. (2002). A Tetrahymena thermophila ribozyme-based indicator gene to detect transposition of marked retroelements in mammalian cells. Nucleic acids research *30*, e49.

Freeman, J.D., Goodchild, N.L., and Mager, D.L. (1994). A modified indicator gene for selection of retrotransposition events in mammalian cells. BioTechniques *17*, 46, 48-49, 52.

Garcia-Perez, J.L., Doucet, A.J., Bucheton, A., Moran, J.V., and Gilbert, N. (2007). Distinct mechanisms for trans-mediated mobilization of cellular RNAs by the LINE-1 reverse transcriptase. Genome research *17*, 602-611.

Goodier, J.L., Cheung, L.E., and Kazazian, H.H., Jr. (2013). Mapping the LINE1 ORF1 protein interactome reveals associated inhibitors of human retrotransposition. Nucleic acids research *41*, 7401-7419.

Huang, C.F., and Zhu, J.K. (2014). RNA Splicing Factors and RNA-Directed DNA Methylation. Biology *3*, 243-254.

Hung, L.H., Heiner, M., Hui, J., Schreiner, S., Benes, V., and Bindereif, A. (2008). Diverse roles of hnRNP L in mammalian mRNA processing: a combined microarray and RNAi analysis. Rna *14*, 284-296.

Kopera, H.C., Larson, P.A., Moldovan, J.B., Richardson, S.R., Liu, Y., and Moran, J.V. (2016). LINE-1 Cultured Cell Retrotransposition Assay. Methods in molecular biology *1400*, 139-156.

Kopera, H.C., Moldovan, J.B., Morrish, T.A., Garcia-Perez, J.L., and Moran, J.V. (2011). Similarities between long interspersed element-1 (LINE-1) reverse transcriptase and telomerase. Proceedings of the National Academy of Sciences of the United States of America *108*, 20345-20350.

Moldovan, J.B. (2015). Identification of Cellular Host Factors That Associate With LINE-1 ORF1p and the Effect of the Zinc Finger Antiviral Protein Zap on LINE-1 Retrotransposition.

Moldovan, J.B., and Moran, J.V. (2015). The Zinc-Finger Antiviral Protein ZAP Inhibits LINE and Alu Retrotransposition. PLoS genetics *11*, e1005121.

Moran, J.V., Holmes, S.E., Naas, T.P., DeBerardinis, R.J., Boeke, J.D., and Kazazian, H.H., Jr. (1996). High frequency retrotransposition in cultured mammalian cells. Cell *87*, 917-927.

Muckenfuss, H., Hamdorf, M., Held, U., Perkovic, M., Lower, J., Cichutek, K., Flory, E., Schumann, G.G., and Munk, C. (2006). APOBEC3 proteins inhibit human LINE-1 retrotransposition. The Journal of biological chemistry *281*, 22161-22172.

Peddigari, S., Li, P.W., Rabe, J.L., and Martin, S.L. (2013). hnRNPL and nucleolin bind LINE-1 RNA and function as host factors to modulate retrotransposition. Nucleic acids research *41*, 575-585.

Perepelitsa-Belancio, V., and Deininger, P. (2003). RNA truncation by premature polyadenylation attenuates human mobile element activity. Nature genetics *35*, 363-366.

Richardson, S.R., Narvaiza, I., Planegger, R.A., Weitzman, M.D., and Moran, J.V. (2014). APOBEC3A deaminates transiently exposed single-strand DNA during LINE-1 retrotransposition. Elife *3*, e02008.

Sassaman, D.M., Dombroski, B.A., Moran, J.V., Kimberland, M.L., Naas, T.P., DeBerardinis, R.J., Gabriel, A., Swergold, G.D., and Kazazian, H.H., Jr. (1997). Many human L1 elements are capable of retrotransposition. Nature genetics *16*, 37-43.

Taylor, M.S., Lacava, J., Mita, P., Molloy, K.R., Huang, C.R., Li, D., Adney, E.M., Jiang, H., Burns, K.H., Chait, B.T.*, et al.* (2013). Affinity proteomics reveals human host factors implicated in discrete stages of LINE-1 retrotransposition. Cell *155*, 1034-1048.

Twyffels, L., Gueydan, C., and Kruys, V. (2011). Shuttling SR proteins: more than splicing factors. Febs J *278*, 3246-3255.

Wei, W., Gilbert, N., Ooi, S.L., Lawler, J.F., Ostertag, E.M., Kazazian, H.H., Boeke, J.D., and Moran, J.V. (2001). Human L1 retrotransposition: cis preference versus trans complementation. Molecular and cellular biology *21*, 1429-1439.

Appendix 2

Methyltransferase methylates H2B and suppresses human retrotransposition

Appendix two addresses my contribution to work done in collaboration with Dr. Robert Schneider and Dr. Idan Cohen, which began at the Institut Génétique Biologie Moléculaire Cellulaire in Strausbourg and is now being completed at the Institute of Functional Epigenetics at the Helmholtz Center in Munich. The Schneider group discovered a novel histone 2B (H2B) methylation (H2Bme) in mammalian cells. Intriguingly, the Schneider laboratory demonstrated that the novel H2Bme appears to be enriched at genomic L1s, Alus, and other simple DNA repeats, but is not enriched at constitutive heterochromatin or protein coding genes (Performed in Schneider laboratory). The Schneider group also identified a putative methyltransferase as the enzymatic protein responsible for catalyzing the methylation on histone 2B *in vivo* (Performed in Schneider laboratory).

To further explore the biological role of the novel H2Bme, the Schneider laboratory created two HeLa cell lines that stably express an shRNA targeting the putative methyltransferase transcript (*i.e.*, KD2, and KD4) and a control HeLa cell line that expresses a non-targeting shRNA (*i.e.*, SCR). Additionally, western blots performed in the Schneider laboratory to detect novel H2Bme in KD2 and KD4 cells revealed a global decrease of novel H2Bme upon methyltransferase knockdown when compared to SCR cells. Moreover, chromatin immunoprecipitation (ChIP) of novel H2Bme followed by qPCR using primers targeting either the L1 5'UTR L1, or the 5' terminus of Alu, revealed a decrease of novel H2Bme at genomic L1 and Alu loci in KD2 and KD4 cells when compared to SCR cells (performed in Schneider laboratory). Finally, global expression of L1 and Alu RNA, as determined by qPCR, was increased in KD2 and KD4 cells when compared to SCR cells (performed in Schneider laboratory).

Given that knockdown of the putative methyltransferase resulted in increased expression of endogenous L1 RNA, we first asked if endogenous L1 ORF1p expression increased in KD2 and KD4 cells when compared to SCR cells. Next, we asked if

retrotransposition of a transiently transfected engineered L1 construct was elevated in KD2 and KD4 cells compared to SCR cells (Moran et al., 1996; Wei et al., 2001). Finally, we transiently transfected a separate HeLa cell line (HeLa-JVM) with siRNAs targeting the putative methyltransferase, or control siRNAs. We then asked if knockdown of the putative methyltransferase resulted in increased retrotransposition of a transiently transfected engineered L1 construct when compared to control cells.

Results

Western blot experiments were performed using whole cell lysates (WCL) derived from the SCR, KD2, and KD4 HeLa cell lines to determine the steady state levels of ORF1p. ORF1p was detected using a human specific ORF1p antibody (Moldovan and Moran, 2015). Previous publications have shown that certain HeLa cell lines express low levels of ORF1p (Moldovan and Moran, 2015). Consistent with previous findings, WCLs from SCR cells expressed low levels of ORF1p (Figure A2.A). In contrast, WCLs from both the KD2 and KD4 cell lines robustly expressed L1 ORF1p (Figure A2.A). The higher level of ORF1p expression in the KD4 cell line likely is due to more efficient knockdown of the putative methyltransferase compared to the KD2 cell line (Figure A2.A). Consistent with this notion, global levels of the novel H2Bme were lower in the KD4 cell line than the KD2 cell line when compared to the SCR cell line (Performed in Schneider laboratory). Thus, knockdown of the putative methyltransferase results in an increase of endogenous L1 RNA (Performed in Schneider laboratory) as well as L1 ORF1p (Figure A2.A).

Next, we sought to determine if putative methyltransferase knockdown affected L1 retrotransposition. To test L1 retrotransposition, we employed a transient cell culture-based L1 retrotransposition assay (Moran et al., 1996; Wei et al., 2001). The L1 retrotransposition assay utilizes an episomal, engineered L1 expression construct containing a retrotransposition indicator cassette within the L1 3'UTR. The retrotransposition indicator cassette consists of an anti-sense copy of either the neomycin phosphotransferase gene (*mneoI)* or the blasticidin deaminase gene (*mblastI)* (Freeman et al., 1994; Kopera et al., 2011; Moran et al., 1996). The coding sequence of *mneoI* and *mblastI* is interrupted by an intron that resides in the same transcriptional

orientation as L1. This arrangement ensures that functional neomycin phosphotransferase or blasticidin deaminase genes will only be activated upon L1 retrotransposition into genomic DNA (Freeman et al., 1994; Kopera et al., 2011; Moran et al., 1996). Retrotransposition of *mneoI* and *mblastI* confers resistance to the drugs G418, and blasticidin, respectively. The resultant drug resistant foci provide a quantitative readout of L1 retrotransposition.

The SCR, KD2, and KD4 HeLa cell lines were transiently transfected with either a retrotransposition-competent L1 (RC-L1; pJM101/L1.3) or a retrotransposition-deficient L1 (RD-L1; pJM105/L1.3) containing the *mneoI* retrotransposition indicator cassette (Moran et al., 1996; Wei et al., 2001). Retrotransposition of RNAs derived from pJM101/L1.3 was markedly more efficient in the KD2 and KD4 cell lines compared to the control SCR cell line (Figure A2.B, left panel). We observed an approximate two-fold increase in L1 retrotransposition in the KD4 cell line when compared to the SCR cell line (Figure A2.B, left panel) and a more modest (~25%) increase in L1 retrotransposition in the KD2 cell line when compared to the SCR cell line (Figure A2.B, left panel). The differences in the efficiency of putative methyltransferase knockdown between the KD2 and KD4 cell lines likely explain the observed differences in L1 retrotransposition efficiency. Control experiments revealed that RNAs derived from pJM105/L1.3 did not undergo retrotransposition in the SCR, KD2, or KD4 cell lines (Figure A2.B, left panel).

To verify that putative methyltransferase knockdown results in an increase in engineered L1 retrotransposition, we transfected the SCR, KD2, and KD4 cell lines with an RC-L1 (pJJ101/L1.3) or an RD-L1 (pJJ105/L1.3) containing the *mblastI* retrotransposition indicator cassette (Kopera et al., 2011). Consistent with the above data, retrotransposition of RNAs derived from pJJ101/L1.3 was more efficient in the KD2 and KD4 cell lines when compared to the SCR cell line. We observed an approximately two-fold increase in L1 retrotransposition in the KD4 cell line when compared to the SCR cell line and again, a more modest (~75%) increase in L1 retrotransposition in the KD2 cell line when compared to the SCR cell line (Figure A2.B, right panel). Control experiments demonstrated that RNAs derived from a pJJ105/L1.3 could not undergo retrotransposition in the SCR, KD2, or KD4 cell lines (Figure A2.B, right panel).

To recapitulate the L1 retrotransposition phenotype observed in the KD2 and KD4 cell lines, we used siRNAs to transiently knock down the putative methyltransferase in HeLa-JVM cells. HeLa-JVM cells were first transiently transfected with siRNAs targeting either the putative methyltransferase or a non-targeting control. The siRNA-transfected cells were then subject to retrotransposition assays and subsequently transfected with a RC-L1, pJJ101/L1.3 (see methods). In agreement with our previous data, siRNA-mediated knockdown of the putative methyltransferase in HeLa-JVM cells consistently resulted in an approximately 1.5-fold increase of L1 retrotransposition compared to HeLa-JVM cells transfected with the non-targeting control siRNA (Figure A2.C, left panel).

We next tested whether the putative methyltransferase knockdown enhances cellular growth, thereby explaining the increase in blasticidin-resistant foci in HeLa-JVM cells. We transfected siRNA treated cells with a plasmid that constitutively expresses the blasticidin deaminase gene (pCDNA6/TR). Notably, there was not a significant difference in the number of blasticidin resistant colonies between the putative methyltransferase siRNA and non-targeting control siRNA treated HeLa-JVM cells (Figure A2.C, right panel).

It remains unclear how the novel H2Bme affects the retrotransposition of an engineered L1; however, it is possible that histone proteins assemble on the transfected episomal L1 expression construct (Jeong et al., 1991). To test this hypothesis, the Schneider laboratory transfected HeLa cells with pJM101/L1.3 and performed ChIP using antibodies recognizing novel H2Bme followed with qPCR using primers targeting the *mneoI* cassette. These experiments revealed a decrease in the novel H2Bme on the *mneoI* cassette in KD2 and KD4 cells when compared to SCR cells (Performed in Schneider laboratory). These data suggest that nucleosomes containing novel H2Bme may be assembling on the transfected L1 reporter construct.

Discussion

Our data demonstrate a role for a novel H2Bme-mediated repression of genomic L1 expression. However, the mechanism of repression remains unknown. It is unlikely that a single histone modification is sufficient to repress expression of L1 (Garcia-Perez

et al., 2010). It may be that the novel H2Bme recruits other histone modifying proteins that chemically modify the nucleosome. Alternatively, the novel H2Bme may recruit a protein(s) that physically bind the nucleosome that repress genomic expression. Direct repression of genomic L1 loci has precedence. MecP2 recruitment at CpG islands in the L1 5'UTR has been demonstrated to repress L1 expression (Muotri et al., 2010; Yu et al., 2001).

Interestingly, the novel H2Bme modification appears to favor evolutionarily young L1s (*i.e.*, those L1 subfamilies that have amplified within the last ~50 million years) (Performed in Schneider laboratory). It is enticing to speculate that younger L1s are more likely to be full-length and transcriptionally active than L1s from older subfamilies (Boissinot et al., 2004a; Boissinot et al., 2004b; Khan et al., 2006; Lander et al., 2001). It also is possible that the putative methyltransferase targets L1 loci either during, or shortly after L1 integration to silence L1 expression. However, given that the novel H2Bme is retained on L1s that are tens of millions of years old, a more plausible explanation is that this novel H2Bme is a secondary modification deposited at L1 loci post-integration. Such a mechanism would rely on other processes to first recognize the newly integrated L1, and subsequently recruit the putative methyltransferase to deposit novel H2Bme as a more stable, "long-term" modification. Additional experiments will be necessary to determine the kinetics of novel H2Bme mediated repression of L1s.

Methods

Cell Culture

HeLa-JVM cells were cultured in high glucose Dulbecco's Modified Eagle Medium (DMEM) lacking pyruvate (Invitrogen). DMEM was also supplemented with 10% fetal bovine calf serum (FBS) (Invitrogen) and 1X penicillin/streptomycin/glutamine (Invitrogen) to create DMEM-complete medium (Moran et al., 1996). Control SCR, Methyltransferase KD2, and KD4 HeLa cells were cultured in DMEM-complete medium and supplemented with 3mg/ml of puromycin (Sigma Aldrich) to maintain selection of shRNA expressing cells. All cells were grown in a humidified tissue culture incubator (Thermo Scientific, Waltham MA) at 37^{o}C in the presence of 7% CO_2.

shRNA knockdown (note: method performed in the Schneider laboratory)

HeLa cells were plated at a density of $3.5x10^6$ cells. Cells were subsequently transfected with 10 mg of non-targeting control human GIPZ Lentiviral shRNAmir vectors (RHS4346), or 10 mg of the methyltransferase-targeting GIPZ Lentiviral shRNAmir (Thermo Scientific). Five shRNA vectors targeting different methyltransferase transcript positions were tested from the Thermo vectors set. Twenty-four hours post transfection, media was changed and cells were cultured in DMEM media supplemented with 10% FBS, 20U/ml penicillin/streptomycin, 0.4mM glutamine (GIBCO), and 3μg/ml puromycin (Life Technologies). Transfected cells were cultured for approximately 3 weeks. puromycin resistant cells were re-seeded and allowed to expand to 80% confluency. All cell lines were subsequently analyzed for the reduction of the putative methyltransferase and the novel H2B methylation. Putative methyltransferase knockdown was most efficient in cells lines transfected with shRNA expressing vectors #2 (yielding the KD2 cell line), and #4 (yielding the KD4 cell line). Both KD2 and KD4 showed over 90% reduction of methyltransferase expression levels and were further used in the different assays.

Western Blot measuring LINE-1 ORF1p levels in SCR, KD2, and KD4 cell lines

Cells were grown to confluence in T-175 flasks, washed with 1X phosphate-buffered saline (PBS) (Invitrogen), and collected in 2ml of PBS using a cell scraper (BD Falcon). Cells were pelleted by microcentrifuge at 14,000rpm for 30 minutes. Whole cell lysates were obtained by suspending and incubating the cell pellet in 1ml of a 10% glycerol, 20mM Tris pH 7.5, 150mM NaCl, 0.1% NP-40 (Sigma-Aldrich), and 1X protease inhibitor (Roche) lysis buffer for 30 minutes on ice. Lysates were then centrifuged at 14,000rpm and the supernatants collected. Protein concentration was determined by Bradford assay. Samples were incubated with 2X NuPAGE (Novex) reducing buffer (containing 1.75-3.25% lithium dodecyl sulfate and 50 mM dithiothreitol (DTT)) and incubated at 100^oC for 10 minutes. An aliquot (20 μg) of total protein was loaded onto 10% precast mini-PROTEAN TGX (BioRad) gels and run at 200V for 1 hour in 1X Tris/Glycine/SDS buffer (1-2% Glycine, 0.25-0.5% Tris-HCl, and 0.01-0.1% SDS, BioRad). Transfer was performed using the Trans-Blot® Turbo™ Mini PVDF Transfer

Packs (BioRad Laboratories) with the Trans-Blot® Turbo™ Transfer System (BioRad Laboratories) at 25V constant for 7 minutes. PVDF membranes were then cut at the 75kD marker (Precision Plus Protein Kaleidescope, BioRad). The top half of the blot was used as a loading control using eIF3 antibody SC-28858 (Santa Cruz Biotechnology) at a concentration of 1:1000. The bottom half of the blot was probed using an antibody against the N-terminus (amino acids 31-49; EQSWMENDFDELREEGFRR) of human ORF1p at a concentration of 1:10,000. Development and characterization of the ORF1p antibody was performed in the Moran laboratory. All antibodies and membranes were blocked in 5% dry low-fat milk (Kroger Co.) overnight. Blots were then washed 3X with PBS-0.1% Tween. ECL HRP-linked secondary antibodies (GE Healthcare) were diluted to a concentration of 1:5000 in 5% milk and applied to the membrane for 30 to 60 minutes. Following three washes with 1X PBS-0.1% Tween, The signals then were visualized using the SuperSignal™ WestFemto Chemiluminescent Substrate reagent (ThermoFisher Scientific) according to the protocol provided by the manufacturer. The membranes were then exposed to Amersham Hyperfilm ECL (GE Healtchare) for a time that spanned two seconds to five minutes and were developed using a JP-33 X-Ray Processor (JPI America Inc.).

Plasmids Used

pJM101/L1.3: contains a full-length version of L1.3 in the pCEP4 backbone. The 3'UTR of L1.3 contains the *mneoI* retrotransposition indicator cassette (Dombroski et al., 1993; Moran et al., 1996; Sassaman et al., 1997).

pJM105/L1.3: is the same as pJM101/L1.3 except it contains a mutation in the reverse transcriptase domain of ORF2, rendering the element retrotranspositionally inactive (Moran et al., 1996; Sassaman et al., 1997).

pJJ101/L1.3: contains a full-length version of L1.3 in the pCEP4 backbone. The 3'UTR of L1.3 contains the *blastI* retrotransposition indicator cassette (Kopera et al., 2011).

pJJ105/L1.3: is identical to pJJ101/L1.3 except it contains a mutation in the reverse transcriptase domain of ORF2, rendering the element retrotranspositionally inactive (Kopera et al., 2011)

pCDNA6: pcDNA6/TR: expresses the blasticidin deaminase resistance gene and was purchased from Invitrogen.

mCherry: pCEP4 based plasmid with the mCherry (Addgene) coding sequence cloned into the *HindIII* and *Bstz17I* restriction sites.

L1 Retrotransposition assays

To monitor LINE-1 retrotransposition, we used a modified version of a previously described transient transfection assay (Wei et al., 2000). Briefly, HeLa cells were plated in each well of a 6-well (BD Falcon) tissue culture dish at a density of $5X10^3$ cells/well for G418-based LINE-1 retrotransposition assays and at $1X10^4$ cells/well for blasticidin-based LINE-1 retrotransposition assays. Transfections for each plasmid condition were performed in 3 wells using a mixture containing 100 µl Opti-MEM® (Life Technologies), 3 µl FuGENE6 (Promega) transfection reagent, and 1 µg L1 plasmid DNA per well. Approximately 24 hours post-transfection, the media was replaced with DMEM-complete media to stop the transfection. Seventy-two hours post-transfection DMEM-complete medium containing 400 mg/ml G418 (GIBCO) or 10 mg/ml blasticidin (Calbiochem) were added to the cells. The media was replaced daily. Fourteen days post transfection G418 and blasticidin resistant foci were washed with 1X PBS, fixed for 10 minutes at room temperature in a solution of 2% formaldehyde and 0.2% glutaraldehyde in 1X PBS, and then stained with a 0.1% crystal violet solution for 30 minutes. The resultant G418- or blasticidin-resistant foci were subsequently counted to determine retrotransposition efficiency. Images were acquired by scanning the plates using a CanoScan. To determine transfection efficiencies, parallel 6-well dishes of SCR, KD2, and KD4 HeLa cell lines were co-transfected with 0.5 µg of the LINE-1 expression plasmid and 0.5 µg of a pCEP4 plasmid expressing mCherry. Seventy-two hours post-transfection, the transfected cells were subjected to fluorescence detection on an Accuri C6 Flow Cytometer (BD Biosciences) to determine the transfection efficiencies (*i.e.*, the percentage of mCherry-positive cells) for each experiment (Kopera et al., 2016). Three biological replicates were performed for each retrotransposition assay. Error bars on all retrotransposition assays represent standard deviation of technical triplicates from the indicated experiment.

siRNA mediated putative methyltransferase knockdown and control knockdown

HeLa-JVM cells were cultured as described above. siRNA-mediated depletion used ON-TARGET plus SMART pool siRNAs either targeting the putative human methyltransferase or a control non-targeting pool (Dharmacon, 001810). HeLa-JVM cells were plated in each well of a 6-well (BD Falcon) tissue culture dish at a density of $1X10^4$ cells/well. Each well was transfected with DharmaFECT Duo (T-2010) transfection reagent containing 200 pmol of the putative methyltransferase or control siRNA following the manufacturers recommendation. Approximately 24 hours post-transfection, the media was replaced with DMEM-complete media to stop the transfection. Forty-eight hours post transfection cells were washed with 1X PBS (Invitrogen) and fresh DMEM media was added. siRNA treated cells were subsequently transfected with pJJ101/L1.3 and the retrotransposition assay was carried out as described above. To control for potential off-target effects of methyltransferase knockdown; siRNA treated HeLa-JVM cells were plated in 6-well plates at $2X10^3$ cells/well and transiently transfected with 0.1µg of pcDNA6/TR (Invitrogen) plasmid that constitutively expresses the blasticidin deaminase gene. Transfection conditions and subsequent selection of pCDNA6/TR was carried out as described previously for the L1 retrotransposition assay. Three biological replicates were performed for each retrotransposition assay. Three biological replicates were performed for the Methyltransferase knockdown off target effect assay. Error bars on all assays represent standard deviation of technical triplicates from the indicated experiment.

Figure A2: Endogenous ORF1p expression and retrotransposition of an engineered L1 reporter construct are increased in putative methyltransferase knockdown HeLa cells.

A) *Representative ORF1p western blot. HeLa cell lines assayed are listed above the top panel*. Lysates were probed for expression of ORF1p (top panel) or eIF3 (loading control, bottom panel). Western blots were performed with fresh lysates three independent times yielding similar results. B) *Results from the L1 retrotransposition assay in stably expressing shRNA HeLa cell lines.* The x-axis indicates the cell line assayed and the y-axis indicates the relative retrotransposition efficiency (%). Relative retrotransposition efficiencies are normalized to the SCR cell line. A representative neomycin based retrotransposition assay (pJM101/L1.3, left panel) and blasticidin based retrotransposition assay (pJJ101/L1.3, right panel) are indicated. The pJM105/L1.3 and pJJ105/L1.3 transfections serve as negative retrotransposition controls. Images and data are representative of one experiment and error bars indicate standard deviation of technical triplicates from the indicated data. C) *Results from L1 retrotransposition assay in siRNA treated HeLa-JVM cells.* The x-axis indicates the siRNA target and the y-axis indicates the relative retrotransposition efficiency (%). Relative retrotransposition efficiencies are normalized to retrotransposition efficiency in the non-targeting control treated cells. A representative blasticidin based retrotransposition assay (pJJ101/L1.3, left panel) is indicated. A representative cell viability assay (pCDNA6/TR, right panel) is indicated. Images and data are representative of one experiment and error bars indicate standard deviation of technical triplicates from the indicated data.

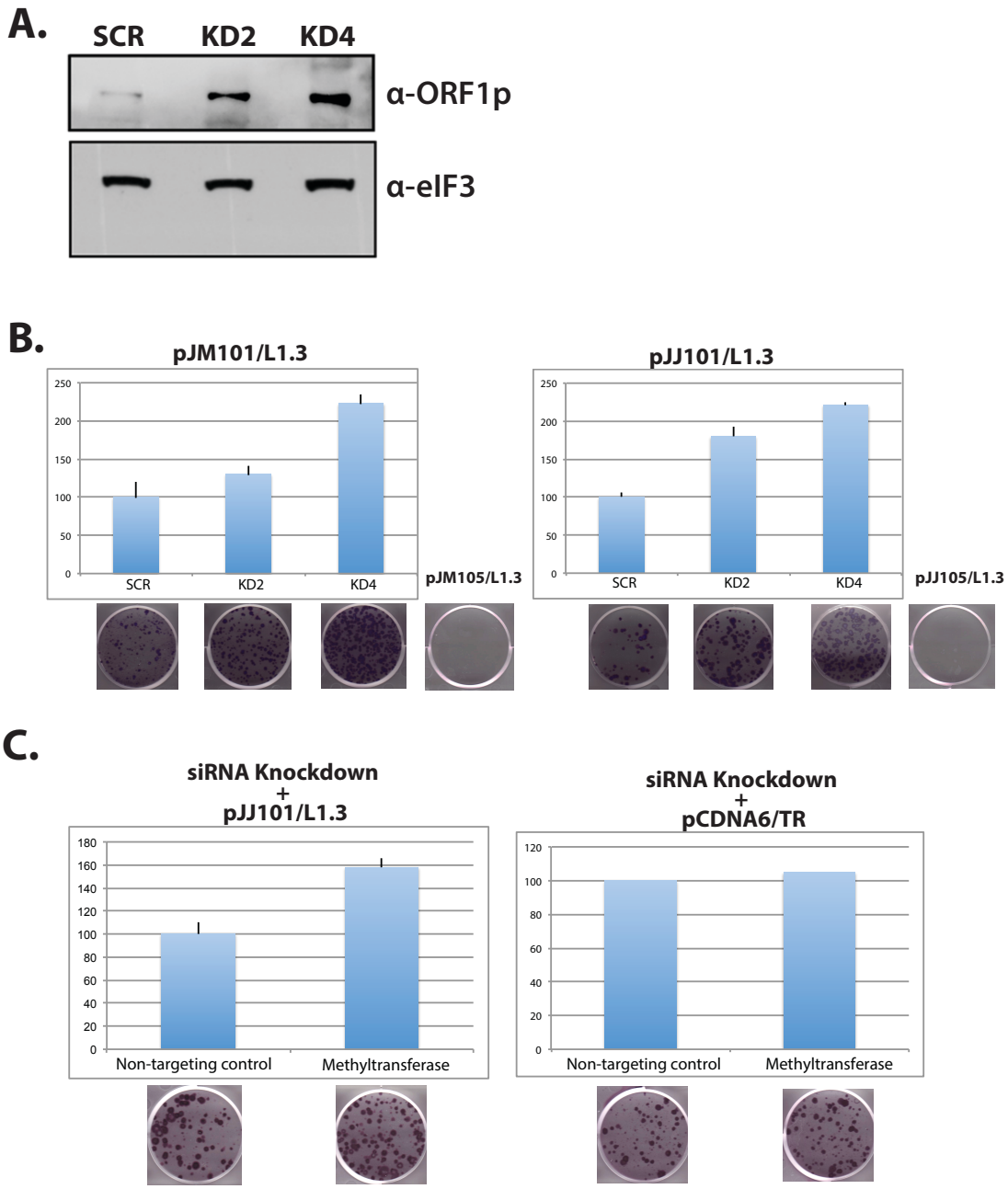

Figure A2: Endogenous ORF1p expression and retrotransposition of an engineered L1 reporter construct are increased in putative methyltransferase knockdown HeLa cells.

References

Boissinot, S., Entezam, A., Young, L., Munson, P.J., and Furano, A.V. (2004a). The insertional history of an active family of L1 retrotransposons in humans. Genome research *14*, 1221-1231.

Boissinot, S., Roos, C., and Furano, A.V. (2004b). Different rates of LINE-1 (L1) retrotransposon amplification and evolution in New World monkeys. Journal of molecular evolution *58*, 122-130.

Dombroski, B.A., Scott, A.F., and Kazazian, H.H., Jr. (1993). Two additional potential retrotransposons isolated from a human L1 subfamily that contains an active retrotransposable element. Proceedings of the National Academy of Sciences of the United States of America *90*, 6513-6517.

Freeman, J.D., Goodchild, N.L., and Mager, D.L. (1994). A modified indicator gene for selection of retrotransposition events in mammalian cells. BioTechniques *17*, 46, 48-49, 52.

Garcia-Perez, J.L., Morell, M., Scheys, J.O., Kulpa, D.A., Morell, S., Carter, C.C., Hammer, G.D., Collins, K.L., O'Shea, K.S., Menendez, P.*, et al.* (2010). Epigenetic silencing of engineered L1 retrotransposition events in human embryonic carcinoma cells. Nature *466*, 769-773.

Hulme, A.E., Bogerd, H.P., Cullen, B.R., and Moran, J.V. (2007). Selective inhibition of Alu retrotransposition by APOBEC3G. Gene *390*, 199-205.

Jeong, S.W., Lauderdale, J.D., and Stein, A. (1991). Chromatin assembly on plasmid DNA in vitro. Apparent spreading of nucleosome alignment from one region of pBR327 by histone H5. Journal of molecular biology *222*, 1131-1147.

Khan, H., Smit, A., and Boissinot, S. (2006). Molecular evolution and tempo of amplification of human LINE-1 retrotransposons since the origin of primates. Genome research *16*, 78-87.

Kopera, H.C., Larson, P.A., Moldovan, J.B., Richardson, S.R., Liu, Y., and Moran, J.V. (2016). LINE-1 Cultured Cell Retrotransposition Assay. Methods in molecular biology *1400*, 139-156.

Kopera, H.C., Moldovan, J.B., Morrish, T.A., Garcia-Perez, J.L., and Moran, J.V. (2011). Similarities between long interspersed element-1 (LINE-1) reverse transcriptase and telomerase. Proceedings of the National Academy of Sciences of the United States of America *108*, 20345-20350.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W.*, et al.* (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860-921.

Moldovan, J.B., and Moran, J.V. (2015). The Zinc-Finger Antiviral Protein ZAP Inhibits LINE and Alu Retrotransposition. PLoS genetics *11*, e1005121.

Moran, J.V., Holmes, S.E., Naas, T.P., DeBerardinis, R.J., Boeke, J.D., and Kazazian, H.H., Jr. (1996). High frequency retrotransposition in cultured mammalian cells. Cell *87*, 917-927.

Muotri, A.R., Marchetto, M.C., Coufal, N.G., Oefner, R., Yeo, G., Nakashima, K., and Gage, F.H. (2010). L1 retrotransposition in neurons is modulated by MeCP2. Nature *468*, 443-446.

Sassaman, D.M., Dombroski, B.A., Moran, J.V., Kimberland, M.L., Naas, T.P., DeBerardinis, R.J., Gabriel, A., Swergold, G.D., and Kazazian, H.H., Jr. (1997). Many human L1 elements are capable of retrotransposition. Nature genetics *16*, 37-43.

Wei, W., Gilbert, N., Ooi, S.L., Lawler, J.F., Ostertag, E.M., Kazazian, H.H., Boeke, J.D., and Moran, J.V. (2001). Human L1 retrotransposition: cis preference versus trans complementation. Molecular and cellular biology *21*, 1429-1439.

Wei, W., Morrish, T.A., Alisch, R.S., and Moran, J.V. (2000). A transient assay reveals that cultured human cells can accommodate multiple LINE-1 retrotransposition events. Anal Biochem *284*, 435-438.

Yu, F., Zingler, N., Schumann, G., and Stratling, W.H. (2001). Methyl-CpG-binding protein 2 represses LINE-1 expression and retrotransposition but not Alu transcription. Nucleic acids research *29*, 4493-4501.