Lehigh University
Lehigh Preserve

Theses and Dissertations

2012

Alu Repetitive Elements Transcribed by RNA Polymerase III are A-to-I RNA Editing Targets

Dylan Edward Dupuis
Lehigh University

Follow this and additional works at: http://preserve.lehigh.edu/etd

Recommended Citation
Dupuis, Dylan Edward, "Alu Repetitive Elements Transcribed by RNA Polymerase III are A-to-I RNA Editing Targets" (2012). *Theses and Dissertations.* Paper 1105.

This Dissertation is brought to you for free and open access by Lehigh Preserve. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of Lehigh Preserve. For more information, please contact preserve@lehigh.edu.

Alu Repetitive Elements Transcribed By RNA Polymerase III Are A-to-I RNA Editing Targets

by

Dylan E. Dupuis

A Dissertation

Presented to the Graduate and Research Committee

of Lehigh University

in Candidacy for the Degree of

Doctor of Philosophy

in

The Department of Biological Sciences

Lehigh University

May 2012

Copyright - 2012
Dylan E. Dupuis

Approved and recommended for acceptance as a dissertation in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Dylan E. Dupuis
Alu Repetitive Elements Transcribed By RNA Polymerase III Are A-to-I RNA Editing Targets

_____
Defense Date

_____
Michael Kuchka, PhD
Dissertation Director

_____
Approved Date

Committee Members:

_____
Linda Lowe-Krentz, PhD

_____
Vassie Ware, PhD

_____
Daniel Lopresti, PhD

Acknowledgements

My time at Lehigh University has truly been an inspirational experience. I would like to thank the wonderful people who have made my time at Lehigh University so memorable and successful. I would like to thank the Department of Biological Sciences for accepting me into the Molecular and Cellular Biology program, and for supporting me over the last year and a half of my graduate career. I would also like to thank Stefan Maas for giving me the opportunity to work in his lab. Mike Kuchka deserves a very special debt of gratitude for his willingness to step in as my new advisor and head of my graduate committee. In addition, he has been sensational in helping me finish my degree. I would also like to thank the other members of my committee, Linda Lowe-Krentz, Vassie Ware, and Dan Lopresti, for all of your help, patience, and willingness to take on such an important role in my own graduate career. I would also like to thank Lynn Cassimeris for loaning me some space in her lab while I finished the last pieces of my graduate research. Renee Roberts also has my thanks for her help with some of my experiments. I would also like to thank the wonderful graduate student community in the Department of Biological Sciences. Especially, Tina Godfried-Sie, Bruce Carney, Joe Leese, Matt Fischl, Mike Kearse, Kevin Tong, Sue Baker, Bill Coleman, Ryan Wynne, Jen Snekser, Jenny Gumm, Vance Imhoff, Victoria Caruso, John Fong, Josh Slee, and the rest of BOGs.

I would like to thank my family for all of their love and support. To my parents for their unquestioned faith in myself, and to my two brothers for always being there. Finally I would like to thank the two most important people in my life, my beautiful wife

Melanie and our wonderful daughter Eliana. Thank you for giving me all the love, support, inspiration and motivation I could ever ask for and more.

Table of Contents

Title Page I

Copyright II

Signature Page III

Acknowledgements IV

Table of Contents VI

List of Abbreviations IX

List of Figures XI

Abstract 1

Chapter 1- Introduction to A-to-I RNA Editing and Alu Repetitive Elements 2

- A-to-I RNA Editing is Catalyzed by a Conserved Family of Proteins 4
- ADAR1 and ADAR2 Expression and Regulation 4
- ADAR Site Selectivity and Editing Efficiency are Separate Events 7
- A-to-I RNA Editing Influences Protein Function by Introducing Non-Synonymous Codon Changes 10
- Coordination of Editing and Splicing 12
- The effects of A-to-I RNA Editing on Alternative Splicing 15
- A-to-I RNA Editing in Untranslated Gene Regions 16
- miRNA Processing and Function is Impacted by A-to-I RNA Editing 18
- ADAR1-p150 and ADAR2 Function in Innate Immune Response to Viral Infection 20
- Use of High-Throughput Technology to Study A-to-I RNA 22

Alu Elements Have Heavily Impacted Human Evolution 25

Regulation of Alu Expression 29

Alu Retrotransposition Occurs Via the L1 Reverse Transcriptase 33

Alu Elements Regulate Gene Expression at the Genomic Level In *Cis* 40

Embedded Alu elements Regulate mRNA In *Cis* 43

Alu Elements Regulate Gene Expression In *Trans* 47

Figures 51

Chapter 2 – RNA Polymerase III Transcripts Are A-to-I RNA Editing Targets 56

Introduction 57

Methods and Materials 60

Results 62

Discussion 66

Figures 71

Chapter 3 – Transcriptionally Active Alu Elements Are A-to-I RNA Editing Targets 74

Introduction 75

Methods and Materials 77

Results 79

Discussion 83

Figures 86

Chapter 4 – Transcriptionally Active Alu elements Undergo Basal Level A-to-I RNA Editing *In Vivo* 91

Introduction 92

Methods and Materials 94

Results 96

Discussion 101

Figures 108

Chapter 5 – Closing Remarks and Future Directions 115

Implications of A-to-I RNA editing on Alu Retrotransposition and Human Genome Evolution 123

A-to-I RNA Editing and Alu Derived lncRNA 126

Future Directions 128

References 133

Curriculum Vitae 157

List of Abbreviations

5-HTC2 – Serotonin Receptor

A-to-I RNA Editing – Adenosine to Inosine RNA editing

ADAR – Adenosine Deaminase Acting on RNA

CTD – C-Terminal Domain

dsRNA - double stranded RNA

ECS – Editing Complementary Site

GluR – Glutamate Receptor

hESC – Human Embryonic Stem Cells

lncRNA – long non-coding RNA

miRNA – micro RNA

NARF - Nuclear Prelamin A Recognition Factor

piRNA – piwi linked RNA

Pol - Polymerase

pre-miRNA – Precursor miRNA

pri-miRNA – Primary miRNA

RAX-R – Alu right arm excluding the poly-A regions

RISC – RNA Induce Silencing Complex

RNP – Ribonucleoprotein

scAlu – Small Cytoplasmic Alu

SINE – Short Interspersed Nuclear Element

SNP – Single Nucleotide Polymorphism

SRP – Signal Recognition Particle

ssRNA – single stranded RNA

TPRT- Target Primed Reverse Transcription

UTR – Untranslated Region

List of Figures

Figure 1.1 – A-to-I RNA editing reaction

Figure 1.2 – A-to-I RNA editing is catalyzed by the ADAR family of enzymes

Figure 1.3 – A-to-I RNA editing occurs in dsRNA

Figure 1.4 – The General Structure of an Alu Element

Figure 1.5 – Alu Elements Retrotranspose Via an RNA Intermediate

Figure 2.2 – The R/G site from GluR-B is an A-to-I RNA editing target when expressed by RNA Pol III

Figure 2.3 – The influence of the RNA Pol II CTD on ADAR function

Figure 3.1 – Design and expression of the Alu expression vectors

Figure 3.2 – Summary of A-to-I RNA editing observed in Alu RNA

Figure 3.3 – Editing Analysis of inverted Alu repeats expressed by RNA Pol II and RNA Pol III

Figure 4.1 – Flow chart of Alu selection screen

Figure 4.2 – Genomic structure surrounding two transcriptionally active Alu elements

Figure 4.3 – Basal level A-to-I RNA editing is observed in Alu-93

Figure 4.4 – A-to-I RNA editing analysis of Alu-93 by nucleotide position

Figure 4.5 – Sequence alignment of Alu-93 cDNA, gDNA, and the AluY family consensus sequence

Figure 4.6 – A model of A-to-I RNA editing during Alu retrotransposition

Abstract

A-to-I RNA editing is a post-transcriptional modification catalyzed by the ADAR enzyme family that targets dsRNA. A major group of editing targets in humans is the primate specific Alu repetitive elements. Due to their high copy number they are often present as inverted repeats embedded within larger mRNA transcripts expressed by RNA Pol II. Alu elements can also be transcribed by RNA Pol III and are termed transcriptionally active. These Alu transcripts are able to expand in the human genome via retrotransposition, and also have functional properties that can impact global gene expression. RNA Pol III transcripts including transcriptionally active Alu elements have not previously been shown to undergo ADAR directed A-to-I RNA editing.

This dissertation details an investigation of A-to-I RNA editing in RNA Pol III transcripts and in transcriptionally active Alu elements. Using an RNA Pol III specific promoter, a known editing target, the R/G site from glutamate receptor B, is shown to undergo ADAR catalyzed A-to-I RNA editing. This editing target demonstrates differences in editing efficiency depending on both the type of polymerase responsible for transcription as well as ADAR1 or ADAR2 activity. This same strategy is used to compare editing in transcriptionally active Alu elements, showing that promiscuous editing takes place regardless of polymerase or ADAR type. A-to-I RNA editing only occurred in the presence of inverted Alu repeats. Finally, transcriptionally active Alu elements are investigated *in vivo* and show evidence of basal level editing. The ramifications of A-to-I RNA editing in transcriptionally Alu elements include impacting human genome evolution and the functionality of Alu derived lncRNA, while providing insights into mechanisms regulating ADAR activity.

Chapter 1

Introduction to A-to-I RNA Editing and Alu Repetitive Elements

A-to-I RNA editing is a post-transcriptional reaction that converts an adenosine nucleotide to inosine within double-stranded RNA (dsRNA) (figure 1.1). The reaction is catalyzed by the adenosine deaminase acting on RNA (ADAR) family of proteins (figure 1.2) (Lehman and Bass 2000). This occurs via a deamination reaction targeting the amine group on carbon-6 in the purine ring. The resulting change can have drastic effects on the final RNA transcript because inosine is interpreted as a guanosine by the cellular machinery. This can result in codon or splice site changes (Athanasiadis et al. 2004, Gommans et al. 2008, Gommans et al. 2008, Lev-Maor et al. 2007). In addition to an effective sequence change in the RNA transcript, A-to-I RNA editing can influence other RNA targeted mechanisms, such as RNA nuclear export, miRNA regulation, and innate immune response to viral infection (Chen et al. 2008, Chen and Carmichael 2009, Kawahara et al. 2008, Taylor et al. 2005).

A major target of A-to-I RNA editing in humans are Alu repetitive elements (Athanasiadis et al. 2004, Peng et al. 2012). Due to their high copy number in the genome, inverted Alu repeats are often expressed embedded within mRNA transcripts (Sela et al. 2007). High Alu sequence conservation allows for formation of highly base paired structures, generating A-to-I RNA editing substrates that can be edited at multiple locations. Alu elements can also be expressed independently via an RNA Pol III dependent mechanism. Alu elements expressed in such a way have functional activity and are also able to retrotranspose (Britten et al. 1988, Hasler and Strub 2006, Mariner et al. 2008). In addition to their role as A-to-I RNA editing targets, Alu elements have other *cis*-acting properties which impact gene expression. Alu retrotransposition has enabled Alu elements to have a significant impact on human genome evolution.

Here I review the fields of both A-to-I RNA editing and Alu elements.

A-to-I RNA Editing is Catalyzed by a Conserved Family of Proteins

A-to-I RNA editing is a post-transcriptional reaction that converts adenosine to inosine within dsRNA (figure 1.3). The reaction is catalyzed the Adenosine Deaminase acting on RNA (ADAR) family of proteins. ADARs are conserved across metazoans (figure 1.2). Two important model organisms used to study A-to-I RNA editing are *C. elegans* and *Drosophila*. *C. elegans* has two ADAR homologs, Adr-1, and Adr-2. While either Adr-1 or Adr-2 deletions reduce editing activity and impair chemotaxis, only Adr-2 has a functional deaminase domain (Tonkin et al. 2002). On the other hand, *Drosophila* encode only one ADAR enzyme (Petschek et al. 1996). Both model species have been important in determining the role ADAR plays both during development and as an RNA regulatory protein.

Vertebrates encode three ADAR enzymes ADAR1, ADAR2, and ADAR3. ADAR3 is expressed exclusively in the brain and located in the nucleus, suggesting a role in neural activity or development (Melcher et al. 1996). Like its vertebrate homologs, ADAR3 has multiple dsRNA binding domains, as well as an arginine rich region near the N-terminus that enables single-stranded RNA (ssRNA) binding (Chen et al. 2000). However, its deaminase domain has not been shown to be active (Chen et al. 2000). In addition, ADAR3 knockout mice display no phenotypic abnormalities (Nishikura 2010). Any further discussion of vertebrate ADAR activity will focus solely on ADAR1 and ADAR2.

ADAR1 and ADAR2 Expression and Regulation

ADAR1 is expressed via two independent mechanisms. The first is via a constitutively active promoter that allows for the expression of a short form of ADAR1, called ADAR1-p110. ADAR1-p110 contains three dsRNA binding domains located near the N-terminus, and a deaminase domain located near the C-terminus. ADAR1-p110 is expressed in most tissues, but is most active in the brain (Gan et al. 2006). ADAR1-p110 is localized entirely within the nucleus.

The other mechanism of ADAR1 expression is by an interferon induced promoter (George and Samuel 1999). Expression from this promoter yields the long form of ADAR1, called ADAR1-p150. In addition to the dsRNA binding domains and deaminase domain present in ADAR1-p110, ADAR1-p150 also has two Z-DNA-like binding domains. These Z-DNA-like binding domains contain a nuclear export signal that allows for ADAR1-p150 nuclear export via an interaction with the nuclear exportin CRM1 (Poulsen et al. 2001). This enables ADAR1-p150 to localize to both the cytoplasm and nucleus. While no specific role for the Z-DNA binding domain has been identified, editing efficiency decreases if this domain is mutated (Herbert and Rich 2001, Koeris et al. 2005).

ADAR1 has been shown to be vital to development, as ADAR1 knockout mice are embryonic lethal (Wang et al. 2000). The reason for this is still unknown since phenotype onset has not been correlated with lack of editing in a specific gene in ADAR1 knockout mice. Rather, it may be a combination of reduction in editing in multiple genes as well as loss of other regulatory functions associated with ADAR1. These other regulatory functions include nuclear retention of edited RNA or changes in gene

expression when ADAR1 is present (Chen and Carmichael 2009, Chen et al. 2008, Osenberb et al. 2010). These will be discussed later in further detail.

Unlike ADAR1, ADAR2 has only 2 dsRNA binding domains in addition to its deaminase domain (Xu et al. 2006). Inositol hexaphosphate is also present in the ADAR2 core, and is required for ADAR2 activity (Macbeth et al. 2005). ADAR2 is expressed in most tissues with the highest activity in the brain (Gan et al. 2006). In addition, ADAR2 is only localized to the nucleus. ADAR2 is able to edit its own pre-mRNA (Dawson et al. 2004). This is an important regulatory target as the edited ADAR2 pre-mRNA results in the creation of an alternatively spliced ADAR2 transcript, which expresses a non-functional ADAR2. Impairing the ability of ADAR2 to self-edit results in global increase in ADAR2 directed editing in mice (Feng et al. 2006). Thus, the ability of ADAR2 to edit its own pre-mRNA serves as a self-regulatory feedback loop.

ADAR2 knockout mice are born with a severe epileptic phenotype and die within two weeks of birth (Higuchi et al. 2000). However, unlike ADAR1, the reason for this phenotype is attributed entirely to lack of editing at a specific site in the mRNA encoding the Glutamate Receptor (GluR), a neural receptor highly involved in excitatory neural transmission. This edited site is called the Q/R site since A-to-I RNA editing results in a glutamine to arginine codon change. Replacing the edited adenosine with a guanosine at the genomic level in ADAR2 knockout mice completely rescues this phenotype (Higuchi et al. 2000). While ADAR2 can edit other sites, the necessity of editing at the Q/R site may have led to the optimization of ADAR2 editing at this site in particular (Lai et al.

1997). This is supported by the observation that the Q/R site is edited to near 100% in normal brain tissue, and is almost exclusively edited by ADAR2.

Regulation of both ADAR1 and ADAR2 may be controlled by modulating localization between the nucleus and nucleolus. SUMO-1 is an ubiquitin like protein that can be conjugated to select proteins. SUMOylation leads to protein sequestration to the nucleolus, the region of the nucleus involved in ribosomal processing (Desterro et al. 2005). While ADAR1 is targeted by SUMOylation, and SUMOylated ADAR1 is present in the nucleus, nucleolar sequestration occurs independent of SUMOylation, as ADAR1 SUMOylation resistant mutants still localize to the nucleolus. However, SUMOylation leads to a reduction of ADAR1 activity. Similarly, ADAR2 can localize to the nucleolus. Like ADAR1, this occurs independent of SUMOylation since ADAR2 lacks a consensus SUMOylation site. Rather, ADAR2 localization to the nucleolus depends on binding to rRNA molecules that can form dsRNA (Sansam et al. 2003). In addition to sequestration, rRNA binding leads to a reduction in ADAR2 editing activity. ADAR2 editing activity is further reduced by binding to snoRNA within the nucleolus (Vitali et al. 2005). snoRNAs are localized in the nucleolus and are involved in post-transcriptional regulation of rRNA.

ADAR Site Selectivity and Editing Efficiency are Separate Events

ADARs can target either a single adenosine or multiple adenosines (promiscuous editing) within a single RNA molecule for A-to-I RNA editing (Athanasiadis et al. 2004, Sommer et al. 1991). In addition, the frequency of editing (editing efficiency) can vary between individual adenosines within a single RNA molecule. In a broad sense, editing site selectivity and efficiency are inversely correlated (Gommans et al. 2008).

Promiscuously edited RNA generally display lower editing efficiency compared to generally higher editing efficiency in RNA molecules edited at a single site.

It is largely unknown why only certain adenosines but not others are edited within dsRNA. Though ADAR1 and ADAR2 share many editing substrates, they have different editing efficiencies (Kallman et al. 2003, Lai et al. 1997, Lehman and Bass 2000). In addition, their site selectivity can differ in promiscuously edited RNA molecules. Site selectivity and editing efficiency are dependent on the coordination of ADARs' binding and catalytic domains as well as the sequence and secondary structure of the substrate itself. Both ADAR1 and ADAR2 are able to form homodimers (Valente and Nishikura 2007). This can occur independent of RNA binding but is necessary for A-to-I RNA editing. While heterodimerization has not been demonstrated, ADAR1 is able to form dimers between ADAR1-p110 and ADAR1-p150.

The dsRNA binding domains within both ADAR1 and ADAR2 display different dsRNA binding affinities (Brooks et al. 1998). The three dsRNA binding domains of ADAR1 are not equivalent in substrate binding. In both humans and *Xenopus*, the dsRNA binding domain located closest to the deaminase domain, domain 3, was demonstrated to be most important for substrate binding. The middle dsRNA binding domain was shown to be the least necessary for substrate binding in humans, but second most important for substrate binding in X*enopus*. Meanwhile, the dsRNA binding domain located closest to the N-terminus had the second most significant effect on substrate binding in humans but was least important in *Xenopus*. This reversal in domain necessity may be due to either species specific differences in ADAR1, or to differences in the binding kinetics associated with the specific substrate used in each study.

ADAR2 displays different binding affinities between its dsRNA binding domains and different binding affinities for different editing targets. Deletion studies indicate a higher involvement in substrate binding by the dsRNA binding domain located closest to the deaminase domain (Ohman et al. 2000, Xu et al. 2006). Structural analysis of both dsRNA binding domains indicates not just secondary structure, but also the specific sequence of the dsRNA itself plays a role in substrate binding (Stefl et al. 2010).

Both editing site selectivity and editing efficiency are dependent on the sequence of the editing substrate (Athanasiadis et al. 2004, Kawahara et al. 2008, Lehman and Bass 2000). Extensive analysis of editing targets revealed preferences for specific nucleotides both 5' and 3' to the edited site. The -2 and -1 positions relative to the editing site show nucleotide preferences of U > C > G > A and C > U > A > G, respectively. Positions +1 and +2 relative to the editing site showed preferences of G > C > U > A and C > G > U > A, respectively. In addition, there is a preference for cytosine, but a negative preference for either adenosine, guanosine, or uracil directly opposite the editing site within the dsRNA structure (Kallman et al. 2003, Wong et al. 2001).

The tertiary structure of the editing target is important, especially in determining patterns of editing (Enstero et al. 2009, Kleinberger and Eisenberg 2011). While extensive dsRNA structure is necessary for substrate binding by ADAR, the location of adenosines within the dsRNA is important for editing site selection. In promiscuously edited RNA, recognition of a principal editing site promotes editing of other adenosines (Enstero et al. 2009). Editing of these other adenosines is limited by both their distance from this principal site, as well as location of those edited sites to the same side of the

helix as the principal site. Editing of additional sites beyond this principal site is hypothesized to occur in a single direction.

A-to-I RNA Editing Influences Protein Function by Introducing Non-Synonymous Codon Changes

The earliest targets of A-to-I RNA editing were discovered serendipitously in the brain. The earliest known targets encoded for neural receptors and each resulted in a codon change that impacted protein function. GluR-B contains two editing sites called the Q/R site and the R/G site (Higuchi et al. 1993, Lomeli et al. 1994). They are named thus because editing results in a glutamine to arginine change and an arginine to glycine codon change, respectively. The glutamate receptor itself is a calcium channel located at the synapse of excitatory neurons. Activation by binding to glutamate causes the calcium channel to open, allowing for a flow of Ca^{+2} down its concentration gradient across the membrane. This leads to membrane depolarization and subsequent activation of the neuron. The Q/R site is located within exon 11 (Sommer et al. 1991). The secondary structure is formed between an exonic sequence and an editing complementary site (ECS) located within intron 11 (Higuchi et al. 1993). Editing at this site is significant because the amino acid encoded at the editing site is located within the Ca^{+2} channel. In the case of the unedited target, both hydrophobicity and the amino acid structure of glutamine at this location causes the ion channel to be constitutively open, allowing for an unchecked flow of Ca^{+2} through the channel. Editing results in the presence of an arginine at this location, preventing the flow of ions in the absence of glutamate binding (Sommer et al. 1991). As already discussed, the necessity for editing at this particular site has been

demonstrated in ADAR2 knockout mice, where lack of editing leads to a severe neurological phenotype and death within 2 weeks of birth (Higuchi et al. 2000).

The other editing site present in GluR is the R/G site. This editing site is located in the last codon of exon 13 (Lomeli et al. 1994). The dsRNA hairpin structure overlaps the exon/intron junction and is 81 nucleotides in total length. Though not as physiologically significant as the Q/R site, editing at the R/G site is involved in recovery following membrane depolarization, where edited channels display significantly faster recovery rates over unedited variants.

Another important A-to-I RNA editing target in which editing impacts the codon sequence is the serotonin receptor (5–HTC2) (Burns et al. 1997). 5-HTC2 is a member of the G-protein-coupled superfamily of receptors. Activation by ligand binding in neuronal synapses causes 5-HTC2 to interact with the G-proteins G_q or $G_{12/13}$ to initiate a signal cascade that leads to production of inositolphospate and diacyglycerol (Price et al. 2001). In addition to the functional relevance, what makes this editing target interesting is that 5-HTC2 contains five editing sites spaced across 14 nucleotides that influence the sequence of three codons. A-to-I RNA editing can result in 32 mRNA variants resulting in 24 protein isoforms. Editing reduces G-protein coupling and leads to a reduction in agonist-receptor sensitivity and selectivity (Berg et al. 2001, da Silva et al. 2010, Price et al. 2001). While editing may provide a mechanism to regulate receptor sensitivity, mice engineered to exclusively express the fully edited 5-HTC2 isoform display phenotypes consistent with Prader-Willi Syndrome (Morabito et al. 2010).

A-to-I RNA editing as a modulator of protein activity was displayed by comparing differences in editing rates in K^+ channels between different species of

octopus (Garrett and Rosenthal 2012). Editing causes an I/V change in codon 321, resulting in a protein channel with faster closing rates. The amount of editing observed depended on the water temperature of their natural environment. A near linear negative correlation was found between observed editing levels and water temperature. This indicates A-to-I RNA editing can provide a mechanism of post-transcriptional adaptation to changes in environmental cues.

Coordination of Editing and Splicing

Protein coding transcripts undergo a series of co-transcriptional processing steps including the addition of a 5' cap, alternative splicing, and poly-adenylation (Buratowski 2009, Hirose and Okhuma 2007). Each of these steps is coordinated by RNA Pol II, the polymerase responsible for expression of protein coding genes. A key structural feature of RNA Pol II is the C-terminal domain (CTD). In humans, the CTD is made up of 52 heptad repeats with consensus amino acid sequence Tyr-Ser-Pro-Thr-Ser-Pro-Ser (Fong and Bentley 2001). This region is highly involved in coordination of co-transcriptional processing events, and is required for efficient transcription initiation. In general, the heptads 1-25 of the CTD are involved in transcription initiation and the switch from initiation to elongation, while heptads 26-52 are involved in RNA processing activities (Buratowski 2009, Hirose and Okhuma 2007).

A-to-I RNA editing occurs in double stranded regions formed between intramolecular complementary sequences. However, complementary regions often straddle exon/intron junctions. Because of this, editing must precede splicing (Ryman et al. 2007). To test this relationship, RNA Pol II expression vectors have been developed

that vary in the structure of their CTD and are also resistant to the RNA Pol II inhibitor α-amanitin (Fong and Bentley 2001).

An interesting comparison can be made between the R/G site and the Q/R site in GluR-B. The R/G site forms a hairpin structure with a dsRNA region 38 nucleotides in length with the complementary regions separated by 5 nucleotides (Lomeli et al. 1994). In addition, the editing site is situated 2 nucleotides upstream of the intron/exon junction. The complementary regions that allow for editing at the Q/R site are separated by a greater distance. In addition, the exon/intron junction is not situated within the dsRNA region, but rather is located in the loop region separating the complementary sites (Lai et al. 1997). These two editing sites differ in terms of the position of the exon/intron junction relative to the dsRNA region and the editing site.

Using editing and splicing variants of the R/G editing site, editing was shown to inhibit splicing (Schoft et al. 2007). Either editing incompetent sequences or the addition of RNA Helicase A, which disrupted ADAR2 binding, increased splicing (Bratt and Ohman 2003). RNA Pol II CTD deletion did not affect A-to-I RNA editing in splice-incompetent variants. However, R/G editing targets with enhanced splicing capabilities indicated that efficient editing was dependent on the presence of the CTD when splicing is enabled (Ryman et al. 2007). This indicates that editing and splicing compete with each other when an active exon/intron junction is present within the dsRNA region. Also, these experiments show that the CTD is important in ensuring that editing precedes splicing in such scenarios.

Similar studies using the Q/R site produced different results. Here, editing and splicing are in less competition with each other since the splice site is present in the loop

rather than within the dsRNA region. Editing in splice competent constructs was near 100% in the presence of wild-type CTD. Also, transient transfection of ADAR2 caused splicing to increase (Ryman et al. 2007, Schoft et al. 2007). This effect by ADAR2 on splicing was shown to be editing dependent. Meanwhile, RNA Pol II CTD deletion resulted in a reduction in editing but an increase in splicing (Ryman et al. 2007). In this case, the CTD delays splicing until editing can occur, indicating a role by the CTD in increasing editing efficiency.

This was supported by an investigation into the self-editing of ADAR2 mRNA (Laurencikiene et al. 2006). In this case, editing results in the creation of an alternatively spliced exon that yields an inactive ADAR2 protein isoform. Editing and thus alternative splicing were shown to be dependent on the presence of the CTD. However, this could be rescued with the presence of either the first half or second half of the CTD. This indicates low regional specificity in ADAR2 to associate with the CTD.

Taken all together, the CTD is necessary to ensure proper editing and splicing when the complementary regions are separated by exon/intron junctions. In addition, A-to-I RNA editing can either enhance or inhibit splicing (Bratt and Ohman 2003, Laurencikiene et al. 2006, Ryman et al. 2007). Differences in this behavior may be dependent on both the location of the exon/intron junction relative to the dsRNA structure, as well the distance between the complementary regions. In cases where the complementary regions are more distantly separated, the CTD may be acting to stall splicing to allow A-to-I RNA editing to take place (Ryman et al. 2007). This stalling may be necessary as it has been shown that editing efficiency decreases as the distance between complementary sequences increases (Athanasiadis et al. 2004).

The effects of A-to-I RNA Editing on Alternative Splicing

Alternative splicing is the discretionary inclusion of individual exons into mature mRNA. This is in contrast to constitutive splicing in which individual exons are always included in the mature mRNA (McManus and Gravely 2011). Alternative splicing can take several different forms, such as mutually exclusive splicing where alternative splicing results in an either/or scenario. Such exons are often called flip-flop exons. Another type occurs with the selective inclusion of an alternatively spliced exon that does not influence the splicing of other exons into the mature mRNA. In most cases, alternative splicing is associated with a change either in function, cellular localization, or regulation of the resulting protein.

A-to-I RNA editing can directly influence alternative splicing by creating or destroying the 5'-GU or AG-3' splice sites present on the ends of intronic sequences. This is illustrated by the self-targeting capability of ADAR2 to edit its own transcript (Dawson et al. 2004). In this case, editing results in the creation of an –AI 3' dinucleotide that functions as an AG-3' dinucleotide. This results in an alternatively spliced mRNA that encodes a non-functional protein.

A more prevalent example of A-to-I RNA editing directly influencing alternative splicing is seen in Alu repetitive elements. Alu elements are primate specific repetitive elements present in over 1 million copies in the human genome. Due to both their high copy number and high sequence conservation, it is very common for Alu elements to be embedded within mRNA molecules (Grover et al. 2003, Sela et al. 2007, Tsirigos and Rigoutsos 2009, Urrutia et al. 2008). It is also common for inverted Alu pairs to be expressed within the same transcript such that they can fold back on each other creating

dsRNA. The commonality of this scenario makes Alu elements major A-to-I RNA editing targets (Athanasiadis et al. 2004, Kim et al. 2007). Alu elements will be discussed in more specific detail in later sections.

Alu elements were first shown to be alternatively spliced as a result of A-to-I RNA editing in the human nuclear prelamin A recognition factor (NARF) (Lev-Maor 2007). Here, inverted Alu repeats in intron 7 were edited, creating a 3' splice site in a sense oriented Alu element located downstream of its inverted partner. Editing created 5 potential splice sites that were each splice competent. Since the function of NARF is yet to be clearly defined, the impact of these splice variants could not be determined. However, primate genome comparisons indicate that the head to head insertion of this inverted Alu pair occurred 58-40 million years ago and that mutations over the course of 34-50 million years eventually enabled A-to-I RNA editing to direct the inclusion of this particular Alu exon (Moller-Krull et al. 2008). No functional domain properties have yet been attributed polypeptides derived from exonized Alu sequences.

A-to-I RNA Editing in Untranslated Gene Regions

Protein coding genes are transcribed with untranslated regions (UTRs) at both their 5' and 3' ends in addition to intronic sequences. 5'UTRs and intronic sequences are enriched for Alu elements relative to translated regions (Tsirigos and Rigoutsos 2009). As such, UTRs harbor major targets for promiscuous A-to-I RNA editing (Kim et al. 2007, Peng et al. 2012). High levels of editing across multiple sites can significantly impact mRNA localization and ultimately gene expression.

Inverted Alu pairs were identified in the 3'UTRs of 333 genes (Chen et al. 2008). GFP reporter genes containing inverted Alu pairs were shown to be down

regulated, but not reporter genes that carried only a single Alu element or no Alu elements in their 3'UTR. This down regulation was shown to be caused by nuclear retention of mRNA with promiscuously edited 3' UTRs (Chen et al. 2008). Also, ADAR1 knockdown in human embryonic stem cells (hESCs) resulted in an increase in RNA from genes involved in development and differentiation, but a decrease in RNA from genes involved in nucleic acid and protein metabolism (Osenberg et al. 2010). Inosine containing mRNA has been shown to have a reduced ability to undergo translation (Scadden 2007). This suggests an editing dependent role by ADAR1 in down regulating gene expression.

Alternatively, transiently transcribed, but not chromosomally located, reporter genes displayed an increase in RNA levels dependent on binding and editing by ADAR1 but not ADAR2 (Gommans and Maas 2008). Also, in both *C. elegans* and humans, mRNA with promiscuously edited 3'UTRs were shown to be bound by multiple ribosomes and was translated (Hundley et al. 2008). This indicates that editing dependent regulation of promiscuously edited RNA can have different effects on gene expression.

The dichotomy may be due to a combination of the manner of expression as well as the system used in these studies. Translation inhibition was observed *in vitro* when dsRNA was engineered with inosine already present, as well as in HeLa cells when dsRNA contained high levels of GU base pairs (Scadden 2007). Such molecules were shown to be present in stress-granule like complexes that led to down regulation of translation. dsRNA molecules engineered to contain high amounts on inosine were

shown to be targeted for cleavage by the Tudor-SN subunit of RNA-induced Silencing Complex (RISC) (Scadden 2005).

Additionally, cell type specific structures are also involved in regulating promiscuously edited mRNA. NEAT1 is a non-protein coding gene that expresses a long non-coding RNA molecule (lncRNA). NEAT1 lncRNA behaves as a scaffold for the formation of paraspeckles. Paraspeckles are localized to the nucleus and house paraspeckle proteins PSP1 and p54. These proteins act to sequester promiscuously edited mRNA into paraspeckles and away from nuclear exportins (Chen et al. 2008, Clemson et al. 2007, Zhang and Carmichael 2001). Interestingly, NEAT1, and thus paraspeckles, are absent from certain cell types such as embryonic stem cells (Chen and Carmichael 2009). Cells lacking paraspeckles display a reduced retention of promiscuously editing mRNA in the nucleus.

miRNA Processing and Function is Impacted by A-to-I RNA Editing

miRNAs are small RNA molecules, 21 to 25 nucleotides in length, which can regulate expression of specific target mRNA transcripts through an RNA silencing pathway. miRNAs are derived from a primary miRNA transcript (pri-miRNA) that may contain multiple miRNA sequences. pri-miRNAs fold into a double stranded RNA structure that is processed by the RNA endonuclease Drosha into a ~60 nucleotide precursor miRNA (pre-miRNA) (Lee et al. 2002, Lee et al. 2003). pre-miRNA is then transported to the cytoplasm where it is further processed into mature double stranded miRNA by the enzyme Dicer (Bernstein et al. 2001, Lee et al. 2002). One of the two mature miRNA strands is then bound by the protein Argonaute and incorporated into the RNA Induced Silencing Complex (RISC) (Hutvagner 2005). RISC/miRNA complexes

regulate translation of specific target mRNA molecules based on complementary base pairing between the single stranded miRNA and the target sequence located within the 3' UTR of the mRNA. Key to this regulation is a part of the miRNA termed the seed sequence. This seed sequence is found in the 5' region of the miRNA and spans nucleotides 2-7. Base pairing of only the seed sequence to the target RNA has been shown to be sufficient for down regulation of expression, regardless of base pairing between the rest of the miRNA (Brennecke et al. 2005, Doench and Sharp 2004).

Due to their inherent stem-loop structure, miRNAs are targets of A-to-I RNA editing. miRNAs can be targeted by A-to-I RNA editing at all stages of miRNA maturation (Blow et al. 2006, Luciano et al. 2001, Kawahara et al. 2005, Kawahara et al. 2008). In the nucleus, A-to-I RNA editing typically, but not always, has an inhibitory role in pri-miRNA processing by Drosha. Processing by Drosha in some miRNAs was unaffected by editing, while in a minority of pre-miRNA, A-to-I RNA editing actually increased processing (Kawahara et al. 2008). Interestingly, edited pri-miRNA can be degraded by Tudor-SN, indicating a post transcriptional mechanism of regulating pri-miRNA concentrations (Yang et al. 2006).

Just as with processing by Drosha, pre-miRNA processing by Dicer is also influenced by A-to-I RNA editing (Kawahara et al. 2007, Kawahara et al. 2008). However, editing does not always inhibit pre-miRNA processing. The location of the edited nucleotide relative to the Dicer cleavage site influences pre-miRNA processing (Kawahara et al. 2008).

Finally, A-to-I RNA editing has the ability impact RNA targeting by miRNAs. While the miRNA seed sequence is key to targeting and down regulating mRNA

containing a complementary sequence, its length is only 7 nucleotides. A single editing event can significantly impact gene targeting. Indeed, edited miRNAs not only show less preference in regulating transcripts which they original targeted, but they can also target new transcripts for regulation that were not originally targeted by that particular miRNA (Blow et al. 2006, Kawahara et al. 2007). A recent study of A-to-I RNA editing in the human brain revealed 47 of 209 pri-miRNAs are edited, and further predicted that about 16% of human pri-miRNAs are A-to-I RNA editing targets (Kawahara et al. 2008). Similar to this, editing at miRNA target sites within Alu repetitive elements can significantly change their targeting by Alu directed miRNAs (Borchert et al. 2009).

ADAR1-p150 and ADAR2 Function in Innate Immune Response to Viral Infection

RNA viruses are a class of viruses in which genetic information is stored as RNA (Pond et al. 2012). Retroviruses, such as human immune-deficiency virus type 1 (HIV-1), encode a retrotranscriptase which allows for integration of viral genetic material into the host genome. Non-retroviruses, such as the measles virus, encode for RNA-dependent RNA Polymerase that is necessary for viral growth. In the case of the measles virus, viral growth occurs entirely in the cytoplasm.

The viral life cycle has three rate limiting steps (Pond et al. 2012, George et al. 2009, Samuel 2011). First, viral infection begins with attachment by the virus to the cell membrane, followed soon thereafter by injection of viral genetic material into the cell. The second rate limiting step is viral proliferation inside the host cell. In the case of retroviruses, proliferation involves the translation of viral RNA by the host ribosome to produce either the reverse transcriptase or the RNA-dependent RNA Polymerase. This enables integration of retroviral genetic material into the host cell's genome, or copying

of viral RNA. Integration enables production of additional viral RNA molecules, translation of viral proteins, and repackaging of viral RNA into newly formed protein coats. The final rate limiting step is the release of newly formed virus from the cell enabling infection of other cells.

The ADAR1 gene has an alternate promoter that is interferon dependent, and produces ADAR1-p150 (George and Samuel 1999). ADAR1-p150 contains a nuclear export signal that allows for it to be expressed in both the nucleus and the cytoplasm (Poulsen et al. 2001). Localization to the cytoplasm allows for ADAR1-p150 to target double-stranded viral RNA. Indeed, HIV-1, hepatitis C virus, and measles virus RNA have been shown to undergo A-to-I RNA editing (Doria et al. 2009, Suspene et al. 2010, Taylor et al. 2005). In the case of hepatitis C virus, ADAR1 significantly decreased viral growth via an editing dependent mechanism (Taylor et al. 2005). However, some viruses have adapted to take advantage of A-to-I RNA editing. In the case of hepatitis delta virus, viral growth increases following A-to-I RNA editing (Polson et al. 1996).

Though ADAR1-p150 is an editing enzyme expressed in response to viral infection, it can influence viral infection via editing independent mechanisms (George et al. 2009). ADAR1-p150 displayed an interesting dichotomy in response to measles virus infection. ADAR1 knockdown in mouse embryonic fibroblasts led to a decrease in the rate of infection by measles virus. However, viral proliferation following infection by measles virus increased in the presence of ADAR1 (Ward et al. 2010). This is due to the effect ADAR1-p150 has on the protein phosphor-kinase R (PKR). PKR is an interferon induced dsRNA binding protein which upon binding to dsRNA dimerizes and autophosphorylates. This activation results in expression of other antiviral factors and can

lead to initiation of apoptosis. ADAR1-p150 reduces PKR activation and PKR induced apoptosis (Toth et al. 2009). This may occur either through substrate competition between ADAR1-p150 and PKR or could be due to the ability of ADAR1-p150 to interact with PKR (Nie et al. 2006). These editing independent effects promoting viral growth were also shown to occur in T-lymphocytes following infection by HIV-1 (Doria et al. 2009). Thus while ADAR1-p150 may aid in reducing initial viral infectivity, it can have an inhibitory effect on cell response to viral infection.

Another interesting study found ADAR2 can also influence infection. ADAR2 overexpression resulted in an increase in HIV-1 proteins in the cytoplasm and an increase in viral release from the cell (Doria et al. 2011). This was shown to be due to a decrease in PKR activity. In addition, ADAR2 knockdown in Jurkat cells resulted in a decrease in HIV-1 viral replication. However, while viral proliferation increased following infection, the rate of infectivity was unaffected. A similar influence on viral growth by ADAR2 was observed in ADAR2 knockout mouse fibroblast cells infected with polyomavirus (George and Samuel 2011). Once again, viral growth, but not infection rate was influenced by ADAR2.

Use of High-Throughput Technology to Study A-to-I RNA Editing

The first A-to-I RNA editing targets were found in the early 1990's in neuronal receptors by serendipity (Burns et al. 1997, Sommer et al. 1991). Since that time, greater focus has been placed in identifying novel editing sites that impact gene function. Expanded use of high throughput technology has greatly increased the search capacity and accuracy of this endeavor. These methods can be broadly divided into two general strategies: search of existing gene databases and use of next generation sequencing.

Though each method has its own set of advantages and disadvantages, their application has greatly expanded beyond the identification of novel recoding events in protein coding genes. High throughput approaches to studying A-to-I RNA editing has shown that A-to-I RNA editing is prevalent. Indeed, A-to-I RNA editing is now implicated in having a significant potential to greatly expand heterogeneity in humans (Athanasiadis et al. 2004, Kim et al. 2004, Li et al. 2011, Paz-Yaacov et al. 2010, Peng et al. 2012).

Improved computational algorithms as well as investigation of unique databases have greatly expanded the number of potential recoding events that are created by A-to-I RNA editing. One example of this was the use of the single nucleotide polymorphism (SNP) database to identify A-to-I RNA editing sites that had previously been mislabeled as SNPs. This screen used a system of filters to scan the SNP database for A/G discrepancies between the consensus genomic sequence and the SNP (Gommans et al. 2008). This screen identified 554 potential editing sites, three of which, one site in SRp25 and two sites in IGFBP7, were identified as novel editing sites that resulted in a codon changes. A second example was the development of the RNA Editing Dataflow System (REDS) program (Maas et al. 2011). This program aligned the expressed sequence tag (EST) database to the genomic database and used a filter system to identify high probability A-to-I RNA editing sites that result in non-synonymous codon changes. Several of these novel sites were verified as bona fide editing sites.

Though a database driven approach can be very cost effective upfront, these high throughput methods require specific verification of each potential editing site. This verification process can be especially costly in terms of time spent. This issue of time can be largely avoided with the use of next generation sequencing. Development of

RNAseq technology has enabled high throughput sequencing on an individual transcript basis.

The use of RNAseq technology can best be demonstrated by the 1000 Genome Project (Li et al. 2011). This project seeks to compare the genomes and transcriptomes of a large group of individuals to catalog and understand genetic diversity across a population (The 1000 Genome Consortium, 2011). A 2011 study by the 1000 Genome Project used high throughput sequencing to sequence Human B cell RNA from 27 different individuals (Li et al. 2011). These RNA sequences were then aligned to the DNA sequence of each individual, and genomic and RNA sequences were compared for nucleotide mismatches. Researchers identified 28,766 total mismatches spread across 10,210 exons and 4741 total genes. A/G mismatches, indicative of A-to-I RNA editing, represented 23% of the mismatches and were the most common type. These A/G mismatches were enriched in coding regions and 3'UTRs. Interestingly, the number of A/G mismatches varied between individuals, indicating an editing dependent mechanism in increasing variability.

A similar study investigated A/G mismatches between genomic and cDNA sequences derived from tissues broadly sampled from a single person (Peng et al. 2012). 22,288 total mismatches were identified, with A/G mismatches representing about 93%. Of these mismatches, an overwhelming majority (~90%) were located in Alu elements. Also, A/G mismatches were enriched in intronic and 3'UTR sequences, which is consistent with Alu distributions within protein coding genes (Sela et al. 2007). These results were consistent with earlier studies indicating Alu elements as major A-to-I RNA

editing targets (Athanasiadis et al. 2004, Kim et al. 2007, Paz et al. 2007, Paz-Yaacov et al. 2010)

A third interesting study sought to correlate A-to-I RNA editing in Alu elements with primate evolution (Paz-Yaacov et al. 2010). This study compared human, chimpanzee, and rhesus monkey for differences in editing levels in the brain. The highest amount of editing was observed in humans while chimpanzee and rhesus monkey did not significantly differ. This difference was not due to differences in ADAR expression. Rather, variation in the presence of nearby inverted Alu elements appeared to play a significant role. Humans and chimpanzees share a majority of their Alu elements. In addition, since their divergence, humans and chimpanzees have had 5530 and 1642 new Alu insertions, respectively (CSAC 2005, Mills et al. 2006). Since the divergence between humans and chimpanzees, new Alu insertions were found to be enriched in genes associated with neuronal development and function (Paz-Yaacov et al. 2010). Increased transcriptome diversity was associated with these new Alu insertions.

Alu Elements Have Heavily Impacted Human Evolution

The human genome is 3.4 billion base pairs in length. The initial draft of the human genome estimated about 50% of the human genome to be comprised of repetitive elements (IHGSC 2001). However, a study using improved sequencing methods combined with different computational and statistical analysis has estimated that repetitive elements may comprise up to two-thirds of the human genome by mass (de Koning 2011). Alu elements compose a major fraction of the repetitive DNA content. Alu elements are primate specific repetitive elements in the short interspersed element (SINE) family of repetitive elements. They are about 300 nt in length and are present in

over 1 million copies in the human genome (Britten et al. 1988). Altogether, Alu elements make up about10% of the human genome by mass (IHGSC 2001).

An Alu element can be divided into a left arm and a right arm separated by a central poly-A track (Figure 1.4) (Batzer et al. 1996). The two armed structure evolved from two individual monomers separately derived from the 7SL RNA of the signal recognition particle (SRP). The fusion of these two monomers about 65 million years ago coincides with the evolutionary divergence of primates from other mammals (Quentin 1992). The left monomer contains putative A-Box and B-Box promoter elements for RNA Polymerase III. Though similar in sequence the right monomer lacks these promoter elements, but does contain an additional 31 nucleotide insert that is not present in the left arm. The right arm ends in a poly-A tail that can be comparatively longer than the central poly-A-track. The overall length and genetic sequence of this poly-A tail is highly variable among individual Alu elements (Roy-Engel et al. 2002).

Alu elements have been divided into three major families: AluJ, AluS, and AluY (Batzer et al. 1996, Britten et al. 1989), that are differentiated based on conserved discrepancies in each family's consensus sequence. Each of these Alu families can be further characterized into different subfamilies based on additional sequence discrepancies. For example, the AluY family can be further categorized as AluYa5, AluYa8, and AluYb8 among others (Batzer et al. 1996). Currently, Alu element retrotransposition is only seen in the AluY subfamilies. Conservation of the AluY sequence and the observation that not all members of the AluY subfamily are able to retrotranspose has led to the hypothesis that genetic sources for Alu retrotransposition are produced by a relatively limited number of Alu elements called master genes (Britten et

al. 1988, Deininger and Slagel 1998). The presence of the various AluY subfamilies supports the idea that there are multiple master genes that are currently responsible for the expansion of Alu elements (Deininger and Slagel 1988, Styles and Brookfield 2007). The commonality of AluY subfamily sequence between humans and chimpanzees suggests that the currently active master genes are shared between both species (Britten et al. 1989, Styles and Brookfield 2007, CSAC 2005, Deininger and Slagel 1988, Han et al. 2005). This means that the AluY subfamily arose prior to the divergence between humans and chimpanzees. Currently, Alu elements are expanding at rate of about 1 new heritable insertion per 20 births, (Cordaux et al. 2006).

Alu elements provide an interesting insight into the primate, and thus, human evolution. Primate genomes can be compared in two ways with regards to Alu elements. The first is through comparative analysis of Alu insertions (Liu et al. 2009, Mills et al. 2006, Paz-Yaacov et al. 2010, Xing et al. 2007). Simply stated more closely related species have more Alu insertions in common. The second is through analysis of sequence divergence in Alu elements shared between species (Britten 2010, Liu et al. 2009, Paz-Yaacov et al. 2010). This provides insight into the relative age of the Alu element since the more recent the insertion, the higher the sequence conservation between species (Brookfield and Styles 2007, Han et al. 2005, Liu et al. 2009, Xing et al. 2007). It is also demonstrates a selective pressure to maintain the Alu consensus sequence, indicating an important role for Alu elements (Liu et al. 2009). These kinds of analyses have been important in understanding the evolutionary relationship between different primate species and the relationship between Alu retrotransposition and primate evolution (CSAC 2005, Mills et al. 2006 Xing et al. 2007). Perhaps one of the most

striking evolutionary correlations is that of increased Alu expansion correlating with increased brain size in humans (Britten 2010, Paz-Yaacov et al. 2010).

Similar studies have been used to compare Alu insertion polymorphisms to better ascertain linear relationships within a species. Such forensic analysis has been performed in both humans and non-humans alike (Garcia-Obreggon 2007, Li et al. 2009, Xing et al. 2007). These studies aid in understanding human migration patterns out of Africa as well as selective forces impacting human evolution at the genomic level.

One interesting effect Alu elements have had on primate genome evolution is through their influence on chromosome recombination and DNA damage repair. During meiosis, homologous chromosomes align with each other, and entire genomic regions can crossover to the other chromosome. This process of recombination is very important in maintaining genotype diversity within a population. AluY elements were shown to have increased recombination rates in areas immediately surrounding the Alu loci. The levels observed were consistent with other recognized recombination hotspots present in the genome (Witherspoon et al. 2009). Alu directed gene rearrangements are implicated in gene deletion/insertion events in primates that can accelerate genome divergences between species (Han et al. 2007). Alu-directed gene rearrangements have been implicated as a cause of Hunter's disease and von Hippel-Lindau disease (Casarin et al. 2006, Ricci et al. 2003). This indicates the potential consequences of Alu directed insertion/deletion events that can occur during recombination. In addition, insertion/deletion events may arise as a consequence of DNA damage repair. Alu elements and L1 repetitive elements have both been implicated in mediating double strand break repair through a retrotransposition like process (Wallace et al. 2010, Srikanta

et al. 2008, Srikanta 2009). This process is a further source of Alu mediated insertion/deletion events in the genome. These processes represent a minor mechanism for repetitive element retrotransposition (Suzuki et al. 2009).

Alu elements have played a significant role in role in shaping the human genome. However, their impact is not limited to their role in shaping genomic structure. Alu elements can impact gene expression in *cis* at both the genomic and RNA levels, and can regulate gene expression in *trans* as functional non-coding RNA molecules. Embedded Alu elements have been shown to be major targets of A-to-I RNA editing. However, A-to-I RNA editing has not been demonstrated in transcriptionally active Alu elements. This dissertation shows that transcriptionally active Alu elements can undergo A-to-I RNA editing. By understanding Alu retrotransposition and Alu function, the potential impact A-to-I RNA editing has on these processes can be better understood.

Regulation of Alu Expression

Alu elements can be transcribed by two distinct mechanisms. The first is as embedded Alu elements present within a larger transcript. A common example of this is mRNA, which often contain Alu elements within their UTR's, but also express exonized Alu elements within the open reading frame (Sela et al. 2007). Such genes are transcribed by RNA Pol II. The other mechanism of Alu expression is via RNA Pol III (Dieci et al. 2007). RNA Pol III is responsible for expression of non-protein coding genes such as tRNA, the U6 RNA of the spliceosome, and the 7SL RNA of the signal recognition particle (SRP). Alu elements expressed in such a way are termed transcribed Alu elements, or transcriptionally active Alu elements (Li and Schmid 2001). Since Alu

elements have A-box and B-box promoter elements they are classified as Class II genes when expressed by RNA Pol III (Dieci et al. 2007).

Transcribed Alu elements are distinguished from embedded Alu elements at the RNA level based on a number of characteristics. Chief among these are the transcription start and termination sites. In transcriptionally active Alu elements, transcription begins immediately at the 5' end of the Alu sequence. This is in contrast to embedded Alu elements, which can have 5' flanking sequences that can be traced to the transcription start site of the host gene, potentially located kilobases away.

Transcribed Alu elements also contain 3' flanking sequences. While 3' flanking sequences cannot be used to differentiate transcriptionally active Alu elements from embedded Alu elements, the mechanism of termination can. RNA Pol II and RNA Pol III employ different mechanisms of transcription termination. RNA Pol III recognizes a specific termination sequence encoded as four consecutive deoxythymidines in mammals, 5 in lower eukaryotes (Chu et al. 1995, Matsuzaki et al. 1994). On the other hand, termination during transcription by RNA Pol II is poorly understood in eukaryotic systems. Rather than relying on a specific consensus termination site, RNA Pol II termination is more processing dependent. For example, full length mRNA transcripts contain poly-adenylation signals encoded as AAUAAA, though this sequence alone is insufficient to induce transcription termination. This is supported by the presence of multiple poly-adenylation signals within some genes (Lee et al. 2008). Interestingly, transposable elements have been shown to contain such poly-adenylation signals (Borodulina and Kramerov 2008). These signals were shown to be active in B2 elements,

a rodent specific SINE family related to Alu elements. Whether they are active in transcribed Alu elements is yet to be shown.

Expression of transcriptionally active Alu elements begins with the binding of RNA Pol III specific transcription factors TFIIIC at both the A-box and B-box promoters. TFIIIC recruits a second transcription factor, TFIIIB to the region near the transcription start site (Ishiguro et al. 2002). TFIIIB is a multimer comprised of three subunits: TATA-binding protein, Brf1, and Bdp1. Following TFIIIB binding to DNA, TFIIIC is released, and RNA Pol III is recruited. The RNA Pol III core enzyme interacts with TFIIIB via an associated subunit complex comprised of three proteins, RPC3/6/7 (Kenneth et al. 2008). The association between TFIIIB and RPC3/6/7 is necessary for RNA Pol III directed transcription initiation, but is not required for transcription elongation. Transcription continues until reaching an RNA Pol III termination site.

RNA Pol III transcription termination occurs at specific sites encoded as four consecutive thymines and is mediated by the transcription termination factor La (Goodier and Maraia 1998, Maraia et al. 1994). However, transcription termination efficiency varies between different genes and is dependent on other factors outside of the consensus termination sequence (Chu et al. 1995). Short palindromic sequences upstream of the termination site as well as G/C nucleotides immediately flanking the termination site can improve termination efficiency (Chu et al. 1995, Chu et al. 1997, Gunnery et al. 1999).

It is important to note that the RNA Pol III termination sequence is not necessarily located immediately adjacent to the end of the 3’ poly-A region. The intervening space between the end of the Alu transcript and the termination site is often, but not always, unique genetic sequence. This unique sequence can be used to annotate individual Alu

transcripts to a specific region of the genome. This 3' unique sequence can also harbor other functional RNA molecules. Prominent examples are miRNA, which can be processed into fully functional miRNA derived from transcriptionally active Alu elements (Borchert et al. 2006, Gu et al. 2009).

Transcriptionally active Alu elements can be regulated at the genomic level by multiple factors. First, Alu elements contain high CpG content. These CpG dinucleotides can be methylated, resulting in localized reduction in Alu transcription (Kochanek et al. 1995, Liu et al. 1994, Muiznieks and Doerfler 1994). Expression of transcriptionally active Alu elements is typically repressed in healthy tissue. However, Alu expression increases following various kinds of cell stress, including viral infection, heat shock, and exposure to DNA damaging reagents (Li and Schmid 2001, Liu et al. 1995, Panning and Smiley 1995, Rudin and Thompson 2001). Alu RNA interacts with RNA Pol II during cell stress to repress transcription in *trans* (Mariner et al. 2008). CpG methylation and Alu RNA functional activity will be discussed in greater detail in later sections.

Alu expression is regulated post transcriptionally via several different mechanisms. One of the more prominent mechanisms is via regulation by piwi-interacting RNA (piRNA). piRNA are short RNA molecules that bind to complementary sequences and regulate expression using a mechanism similar to miRNA. A recent study found that older Alu elements have larger numbers of piRNAs that map to their specific sequence. This indicates positive selection for piRNA as a mechanism of regulating Alu expression (Lukic and Chen 2011).

Transcriptionally active Alu elements are also be regulated by DICER1 (Kaneko et al. 2011). Geographic atrophy, an age related degenerative disease, occurs via retinal pigmented epithelium (RPE) cell degeneration. This degeneration can occur due to both a loss of DICER1 and increased abundance of Alu transcripts. In healthy RPE, knockdown of DICER1, but not other enzymes involved in miRNA processing and activity, resulted in an increase in Alu RNA and onset of RPE cell degeneration. This indicated that Alu transcripts are regulated by DICER1 in a miRNA-independent mechanism, and that loss of Alu regulation may promote disease onset.

Transcribed Alu elements can also undergo post-transcriptional modification. A key post-transcriptional modification results in the formation of scAlu (Shaikh et al. 1997). scAlu transcripts are comprised of only the left arm of the Alu element, but are generated from full length Alu transcripts via 3'-end processing in the central-A track. scAlu elements are important due to their ability to associate with the SRP9/14 subunit of the SRP (Hasler and Strub 2006). This scAlu/SRP9/14 complex inhibits translation initiation. They also associate with RNA Pol II during heat shock (Mariner et al. 2008). Finally, increased processing of younger Alu elements into scAlu relative to older Alu elements is implicated as a potential retrotransposition regulatory mechanism (Sarrowa et al. 1997).

Alu Retrotransposition Occurs Via the L1 Reverse Transcriptase

Alu elements retrotranspose via an RNA intermediate transcribed by RNA Pol III (figure 1.5).The estimated frequency of inherited Alu retrotransposition is once every 20 generations (Cordaux et al. 2006). Alu elements retrotranspose via a mechanism called target-primed reverse transcription (TPRT). TPRT uses the genomic insertion site as a

primer for reverse transcription. In this mechanism, reverse transcription and retrotransposition occur simultaneously. Alu retrotransposition is dependent on key structures in the Alu element as well as the L1 repetitive element and is regulated by a myriad array of mechanisms.

Alu elements do not encode their own retrotransposition machinery. Instead, they rely on L1 for their own expansion in the genome. L1 elements are members of the long interspersed element (LINE) family of transposons. L1 elements are also protein coding genes, encoding an mRNA with two open reading frames, ORF1 and ORF2, which produce proteins ORF1p and ORF2p (Alisch et al. 2006). Each ORF is highly conserved, and are functionally interchangeable with homologs from other species without significantly impacting retrotransposition (Feng et al. 1996, Januszyk et al. 2007, Wagstaff et al. 2011, Weichenrieder et al. 2004). ORF1p encodes an RNA binding protein with RNA chaperone activity dependent on a coiled-coil domain (Callahan et al. 2011, Janusyzk et al. 2007). While ORF1p is required for L1 retrotransposition, it is not necessary for Alu retrotransposition, though its presence does enhance the rate of Alu retrotransposition (Kroutter et al. 2009, Martin et al. 2005, Wallace et al. 2008). ORF2 encodes both an endonuclease and a reverse transcriptase, and is by itself sufficient for Alu retrotransposition (Feng et al. 1996, Mathias et al. 1991). Though ORF2 shows a strong preference for binding its own RNA, poly-adenylated RNA molecules, both coding and non-coding, can replace the bound L1 mRNA via a template switching mechanism (Esnault et al. 2000, Kroutter et al. 2009, Kulpa and Moran 2005, Wei et al. 2001). This helps facilitate retrotransposition of the newly bound RNA transcript.

Alu retrotransposition via L1 mediated TPRT begins with Alu transcription by RNA Pol III. Alu transcripts are then exported to the cytoplasm where they interact with the SRP9/14 subunit of the SRP (Andrews and Kole 1987, Hsu et al. 1995). Alu bound SRP9/14 can associate with the ribosome. Though the mechanism is not fully understood, this is believed to be important in aiding Alu transcript recruitment to ORF2p following translation of the L1 protein. Typically, ORF2p preferentially binds its own encoding mRNA (Kulpa and Moran 2006, Wei et al. 2001). However, Alu elements replace the L1 mRNA via a template switching mechanism involving the Alu poly-A tail forming a new Alu RNP complex with ORF2p. Though not required, ORF1p can also associate with this RNP. The RNP is then imported into the nucleus. TPRT is initiated by the endonuclease activity of ORF2p opening the DNA strand in the consensus sequence TTAAAA (Feng et al. 1996). These A/T rich regions prime reverse transcription in the Alu poly-A tail, which proceeds in 3' to 5' direction (Kroutter et al. 2009). It is unclear whether the opposite DNA strand is cut before, during or after retrotransposition. In any case, the use of genomic DNA sequence as a primer results in target site duplications at the 5' and 3' ends of the newly retrotransposed Alu element due to the presence of overhanging ends formed during endonuclease cleavage and reverse transcription priming (Kojima 2010).

The A-box and B-box promoter elements are RNA Pol III specific promoter elements located within the left arm of the Alu element. The process of Alu element expansion is such that the promoter elements remain intact following retrotransposition. However, not all Alu elements have maintained functional promoters. Over time, the promoters of many Alu elements, especially from the AluJ and AluS families, have lost

functionality due to accumulation of mutations (Comeaux et al. 2009). This limits the number of Alu elements that are able to retrotranspose to those with intact, functional promoter elements.

Development of a marked-Alu expression vector greatly expanded research into structural features necessary for Alu retrotransposition (Dewainnieux et al. 2003). These marked Alu elements can be tracked in the genome and also provide neomycin resistance, enabling quantitative measurement of Alu retrotransposition rates. Two structures investigated were the length and heterogeneity of the Alu poly-A tail (Comeaux et al. 2009). A minimum length of 20 uninterrupted adenosines was shown to be a lower limit for retrotransposition using the marked Alu system. Longer stretches of uninterrupted A-tail length did not show a difference in retrotransposition rates. However, an interruption of just a single non-adenosine nucleotide within the A-tail was sufficient to significantly reduce retrotransposition rates. The importance of the poly-A tail length in Alu retrotransposition is supported by genomic analysis that revealed that the average poly-A tail length of AluJ and AluS family elements have a shorter mean distribution compared to the AluY family (Roy-Engel et al. 2006). In addition, newly inserted disease causing Alu elements have poly-A tail length distributions larger than 40 nucleotides, with a mean length of 77 nucleotides. This study indicates that the number of master Alu elements responsible for retrotransposition is limited to Alu elements with intact poly-A tails of specific minimal length.

Regions flanking the poly-A tail also influence retrotransposition rates (Comeaux et al. 2009). Random mutations in the Alu right arm reduce retrotransposition, indicating the right arm is involved in helping to facilitate retrotransposition. The Alu right arm has

previously been shown to be involved in formation of the Alu/SRP9/14 complex, highlighting the role that SRP9/14 plays in Alu retrotransposition (Hasler and Strub 2006). In addition, the 3' flanking sequence reduced Alu retrotransposition in a length dependent manner. This indicates that longer 3' flanking regions may either inhibit the interaction between Alu elements and L1 reverse transcriptase, or inhibit the insertion mechanism.

Alu elements are thought to have gone through altering periods of high and low retrotransposition rates (Han et al. 2005, Kim et al. 2004). One proposed mechanism for the changes in retrotransposition rates is the coevolution of both the Alu sequence and the SRP9/14 subunit of the SRP (Liu et al. 1994, Sarrowa et al. 1997). Alu elements retrotranspose via an RNA Pol III transcribed Alu RNA (Bennett et al. 2008, Sarrowa et al. 1997). Though not precisely understood, the SRP9/14 is thought to recruit Alu transcripts to the L1 retrotransposition machinery. The human SRP9/14 subunit binds Alu elements more efficiently than mouse SRP9/14 (Bovia et al. 1997). Closer inspection of Alu right arm binding specificity revealed that SRP9/14 preferentially bound AluSx > AluY > AluYa5. Reasons for this type of preference were attributed to changes in the Alu right arm in younger Alu elements causing a shift in Alu secondary structure that is less favored for SRP9/14 binding. In addition to full length Alu RNA, the SRP9/14 has been shown to bind small cytoplasmic Alu (scAlu) RNA, and may help promote the processing of Alu RNA into scAlu RNA (Aleman et al. 2000, Batzer et al. 1996, Bovia et al. 1995, Bovia et al. 1997, Esnault et al. 2000, Hasler and Strub 2006, Hsu et al. 1995, Liu et al. 1994). Transfection of AluYa5 resulted in relatively high processing of AluYa5 elements into scAlu elements as compared to AluSx and AluY,

each of which produced little to no scAlu. This is supported by research showing that younger Alu elements are more highly processed into scAlu (Liu et al. 1994). This indicates that Alu retrotransposition may be partly limited by the reduced ability of retrotransposably active Alu elements to interact with SRP9/14, coupled with an increase in processing of Alu RNA in scAlu RNA.

Both L1 and Alu retrotransposition is inhibited by the APOBEC3 family of proteins (Bogerd et al. 2006, Chiu et al. 2006). APOBEC3 genes are members of the cytidine deaminase family. APOBEC3 proteins are single stranded RNA binding proteins that can catalyze C-to-U RNA editing via a deamination reaction. These genes have previously been shown to play a key role in the innate immune response. APOBEC3 enzymes can also influence Alu retrotransposition by binding to and sequestering Alu RNA away from L1 reverse transcriptase. This regulatory mechanism behaves independent of C-to-U editing. Interestingly, some of the APOBEC3 enzymes are expressed during early stages of embryo development, indicating an adapted role for these enzymes in regulating retrotransposition.

Though Alu elements and L1 elements both retrotranspose using the same machinery, there are differences in terms of integration site selection. An investigation of G/C content within a 50 bp window of the insertion site displayed a bias for L1 insertion into A/T rich sites, but a bias for G/C rich sites for Alu insertion. This preference was lost when the size of the genomic window was expanded to larger sizes (Gasior et al. 2006). Comparison of insertion preference into imprinted regions showed no preference for L1 content, but Alu content was significantly lower compared to non-imprinted regions (Greally 2002). Similarly, Alu elements were retained in regions

flanking housekeeping genes but not in genes expressed in a tissue specific manner (Eller et al. 2007, Grover et al. 2003). This indicates an inhibitory effect on Alu retrotransposition by heterochromatin, a type of high density DNA packaging that is associated with low levels of gene expression.

A majority of new Alu retrotranspositions are considered to have a neutral evolutionary impact (Cordaux et al. 2006). However, insertion site selectivity indicates a potential impact on gene expression. Alu elements are enriched in oncogenes compared to tumor suppressor genes, suggesting a role in cancer onset (Zhang et al. 2011). Also, development of high throughput methods to track Alu insertions revealed 13,692 polymorphic Alu insertions in the human brain (Baillie et al. 2011). This indicates that retrotransposition may provide a mechanism of increasing Alu related gene regulatory mechanisms during somatic tissue development.

Though retrotransposition rates in somatic cells may differ in a cell type specific manner, inherited Alu retrotranspositions must occur in either human embryonic stem cells (hESCs) or germ line cells. Indirect evidence of Alu retrotransposition has been observed in hESCs (Macia et al. 2011). First, Alu element expression was found to be enriched in the AluY family. Second, Alu elements expressed by hESCs were subcloned into the marked Alu expression vector and were assayed for retrotransposition activity. The AluY subfamily was shown to be retrotranspositionally competent in hESCs. This study, however, did not identify newly inserted Alu elements derived from endogenous Alu transcripts.

Evidence for retrotransposition in human germ line cells is even more fleeting. L1 elements have been shown to be retrotranspositionally active during meiosis in a case

study (Brouha et al. 2002). In addition, L1 is expressed in human oocytes and is retrotranspositionally active in germ line cells and embryos from transgenic mouse and rat models (Georgiou et al. 2009, Kano et al. 2009). Similar studies have not been performed concerning Alu retrotransposition. However, Alu elements are hypomethylated in human male germ line cells, indicating an increased likelihood of Alu expression (Hellmann-Blumberg et al. 1993, Kim et al. 2007).

Alu Elements Regulate Gene Expression at the Genomic Level in *Cis*

Alu elements can impact gene expression in *cis* and in *trans* at both the genomic and RNA levels. *Cis* regulation of gene expression is accomplished by Alu elements through epigenetic regulation and expansion of promoter binding sites in the genome. These influences have been selected for since Alu elements are enriched in promoter regions (Grover et al. 2003, Polak and Domany 2006, Tsirigos and Rigoutsos 2009). Epigenetic regulation is the regulation of gene expression through histone or DNA modifications that can influence DNA packaging and gene expression. Alu elements can influence nucleosome position and are sites for CpG dinucleotide methylation (Bettecken et al. 2011, Rodriguez et al. 2007, Tanaka et al. 2010).

Nucleosomes are DNA packing units comprised of 146-147 bp of DNA wound around a histone octamer (Lugar et al. 1997). Neighboring nucleosomes are separated by linker DNA that can vary in length, and are further wound around each other to create a tightly packaged chromatin structure. Chromatin can be described in two broad packaging types. Euchromatin is a loose packing associated with active gene regions, and is characterized by unmethylated DNA and specific combinations of histone mono-methylation and histone acetylation. Heterochromatin is tightly packaged DNA

associated with low gene expression, key chromosome structural elements such as the centromere or telomere, and imprinted gene regions. Heterochromatin is most typically marked by DNA methylation at CpG dinucleotides and deacylated, dimethylated histones.

Nucleosome positioning is influenced by Alu elements. Though nucleosomes are relatively evenly distributed along the chromosome, they do show sequence specific positional preference along the DNA helix. Nucleosome mapping in primates displayed enriched associations with the Alu left arm and enriched, though slightly lower, associations with the Alu right arm (Tanaka et al. 2010). In addition, in unexpressed genes, nucleosome positioning to Alu elements was significantly higher than to non-Alu sequences. This nucleosome positioning is mediated by the high presence of CpG dinucleotides located within Alu elements (Bettecken et al. 2011). CpG dinucleotides were shown to be strongly preferred nucleosome positioning sites.

DNA methylation is an enzymatically catalyzed epigenetic regulatory process that is impacted by Alu elements (Muiznieks and Doerfler 1994). DNA methylation occurs at CpG dinucleotides present within CpG islands and results in decreased expression of nearby genes. CpG islands are regions of CpG rich sequence that can be kilobases in length. Due to their high CpG content, Alu elements often constitute CpG islands (Cho et al. 2007, Kang et al. 2006). Interestingly, older Alu elements are more highly methylated than younger Alu elements, suggesting a functional adaptation for older Alu elements (Rodriguez et al. 2007). Investigations of Alu methylation patterns show tissue specific methylation patterns, and a dynamic process of differential methylation throughout development (Hellmann-Blumberg et al. 1993, Kim et al. 1994, Rubin et al.

1994, Xie et al. 2011). Furthermore, aberrant Alu methylation patterns have been displayed in multiple types of cancer (Cho et al. 2007, Gao et al. 2012, Rodriguez et al. 2007, Wang et al. 2011, Xie et al. 2010).

In addition to epigenetic regulation, Alu elements can impact gene expression through their ability to distribute transcription factor binding sites throughout the genome (Polak and Domany 2006). Alu elements contain many transcription factor binding sites for both RNA Pol II and RNA Pol III (Bolotin et al. 2011, Komiyama et al. 2010, Laperriere et al. 2007, Polak and Domany 2006, Zhou et al. 2002). Interestingly, a loss in expression by RNA Pol III is correlated with a gain in RNA Pol II transcription factor binding sites (Shankar et al. 2004). Alu retrotransposition has been suggested as a significant means of expanding RNA Pol II regulatory sites in the genome (Polak and Domany 2006, Shankar et al. 2004).

One interesting Alu associated transcription factor binding site is for the heat shock protein HSF (Pandey et al. 2011). Genes that are differentially expressed during heat shock were enriched in their promoter and UTR regions for Alu elements containing HSF binding sites. Expression of these genes increases following heat shock, and is dependent on the conservation of the HSF binding site. In addition, genes down-regulated during heat shock contain transcriptionally active Alu elements in their 5' end that are in the antisense orientation relative to the direction of expression of the downregulated gene. These antisense oriented Alu elements harbor a high amount of HSF binding sites that promote Alu transcription by RNA Pol III following heat shock. Expression of these Alu elements following heat shock results in the down-regulation of

the host gene. Though it is unknown precisely how this mechanism works, transcriptional interference and RNAi have been hypothesized as possible mechanisms.

One thing that must be discussed is the relationship between Alu retrotransposition and gene expression. Alu elements are enriched in promoter regions of housekeeping genes (Grover et al. 2003, Tsirigos and Rigoutsos 2009, Urrutia et al. 2008). However, housekeeping genes are defined as being broadly expressed by most tissues since they are often necessary for normal cell function. Transcription of these genes is not regulated by CpG methylation. Rather, lack of DNA methylation in these genes is hypothesized to help promote Alu retrotransposition to these regions by maintaining DNA in an euchromatin state (Urrutia et al. 2008).

Embedded Alu elements Regulate mRNA In *Cis*

Embedded Alu elements can significantly impact post-transcriptional gene expression. The influence of A-to-I RNA editing in mediating alternative splicing and nuclear retention in genes containing inverted Alu repeats has already been discussed, and so will not be detailed here. Editing independent mechanisms of gene regulation are also prevalent and include alternative splicing, alternative poly-adenylation, and RNA degradation (Amit et al. 2007, Borchert et al. 2009, Chen et al. 2008, Gong and Maquat 2011, Lee et al. 2008, Lev-Maor et al. 2003, Lin et al. 2008, Sela et al. 2007). These processes are highlighted by Alu insertion polymorphisms that result in disease phenotypes (Apoil et al. 2007, Eller et al. 2007, Ganguly et al. 2003, Gu et al. 2007, Mustajoki et al. 1999, Tighe et al. 2002). The combination of these functions and high presence of embedded Alu elements in protein coding genes create a case for Alu elements having a dramatic impact on expanding protein coding potential.

Alu exonization can arise by a number of different mechanisms. The most significant is through A-to-I RNA editing which can generate new splice sites. However, other mechanisms such as point mutations can also result in Alu exonization capabilities (Amit et al. 2007, Lev-Maor et al. 2003). Though Alu exons can be constitutively spliced, the prevalence of alternatively spliced Alu exons is more prevalent (Lin et al. 2008, Sela et al. 2007). Interestingly, older AluJ elements displayed higher exon inclusion levels than the AluS family elements even though the AluJ family is present in a lower copy number within the human genome (Lin et al. 2008). This indicates that the extended presence of AluJ in the primate genome has allowed for evolutionary adaptation to their presence. This is supported by the presence of key structures, such as splice site enhancers, that are present in highly exonized Alu elements (Schwartze et al. 2009). Unsurprisingly, Alu exonization is differentially regulated in different tissues and is associated with cancer phenotypes (Amit et al. 2007, Lin et al. 2008).

Exonized Alu elements may impart limited functionality to resulting proteins. Alu exons are both more prevalent and have higher inclusion levels in 5' and 3' UTRs relative to translated regions (Sela et al. 2007, Shen et al. 2011). Alu elements also have low coding probabilities based on bioinformatics screens, and to date have no identified protein coding function attributed to their sequence (Piriyapongsa et al. 2007, Shen et al. 2011).

However, Alu exons can significantly impact protein function through other mechanisms. A bioinformatics screen indicated alternative splicing of Alu elements often results in either a frame shift (24%) or a premature stop codon (61%) (Sela et al. 2007). In a case study of a Finnish family with hereditary acute intermittent porphyria,

an Alu insertion polymorphism was responsible for the presence of an antisense oriented AluYa5 in exon 5 of porphobilinogen deaminase (PBGD). This Alu insertion created a premature stop codon and resulted in an almost complete loss of PBGD mRNA (Mustajoki et al. 1999). One possible explanation for the mRNA depletion is nonsense mediated decay (NMD), a mechanism that specifically targets mRNA molecules that undergo premature translation termination (Maquat 2005). This mechanism may be at work in additional examples. An Alu exon in the 5'UTR in NOSIP and ZNF81 resulted in reduced translation efficiency, while an Alu exon in the 5'UTR of ZNF808 increased translation efficiency (Shen et al. 2011). Also, in a case study, immunodeficiency syndrome with high-IgM was caused by an Alu retrotransposition in exon1 of CD40LG that resulted in the complete loss of translation (Apoil et al. 2007).

Intronically located Alu elements can also influence the splicing of non-Alu exons. Alu elements can harbor splice silencers (Schwartz et al. 2009). Case studies have revealed several diseases that are caused by a newly retrotransposed Alu element into an intron resulting in the exclusion of a constitutively spliced exon located downstream (Eller et al. 2007, Ganguly et al. 2003, Gu et al. 2007, Tighe et al. 2002).

Embedded Alu elements located in the 3'UTR can also impact transcriptome diversity due to the presence of alternate poly-adenylation signals. Poly-adenylation signals are encoded as AATAAA. This motif is common in Alu elements in three locations: the 3' end of the left arm, the central-A track, and the poly-A tail (Chen et al. 2008). Each of these signals can function as alternate poly-adenylation sites. In addition, 3'UTRs can harbor multiple Alu elements, further increasing their impact on poly-adenylation (Chen and Carmichael 2008, Sela et al. 2007). However, alternate poly-

adenylation sites within Alu elements are selected against since these signals are less conserved as their distance away from the 3' end within the 3'UTR increases (Lee et al. 2008).

Embedded Alu elements can also be responsible for reduction in mRNA transcript levels. One mechanism is through Alu directed miRNA regulation (Borchert et al. 2009, Lehnert et al. 2009, Smalheiser and Torvik 2006). Alu directed miRNAs target the most conserved regions of Alu elements. Interestingly, a primate specific 100 kb region located in chromosome 19 (C19MC), is home to 43 pri-miRNA, most of which have seed sequences that map to Alu elements (Lehnert et al. 2009). C19MC is also enriched for the presence of Alu elements. The high number of miRNAs and Alu elements are believed to have co-evolved via repeated Alu mediated gene duplication events (Zhang et al. 2008). Another note of interest is that some of these miRNAs are expressed by RNA Pol III via Alu elements located upstream (Borchert et al. 2006).

A second mechanism of Alu directed regulation is by base pairing between embedded Alu sequences and Alu derived lncRNA. These lncRNAs interact with embedded Alu elements located in the 3'UTR of mRNA, forming an RNA/lncRNA duplex. This duplex binds to the dsRNA binding protein Staufen1, which facilitates mRNA decay (Gong and Maquat 2011).

Though Staufen1 has only been shown to target RNA/lncRNA duplexes that form in 3'UTRs, a similar lncRNA dependent mechanism may facilitate down regulation of mRNA transcript levels when Alu elements are present in intronic and exonic regions located upstream of the 3'UTR. As already discussed, heat shock factor HSF is involved in expressing a lncRNA derived from an antisense oriented Alu element present in a gene

that is down regulated following cell stress. The lncRNA is responsible for reduction in expression of the host gene (Pandey et al. 2011). In a 2007 study, differences in mRNA transcript levels were investigated for genes heterozygous for a specific intronically located AluY element. AluY insertions were found to correlate with lower levels of primary mRNA transcripts, though this was found to be tissue dependent (Lebedev et al. 2007). This mechanism may be fairly common as 87% of alternatively spliced Alu elements are in the antisense orientation relative to the direction of transcription of the surrounding gene (Sela et al. 2007). This may allow for lncRNA derived from these specific Alu elements to specifically target their embedded antisense transcript.

Alu Elements Regulate Gene Expression in *Trans*

lncRNAs are long-noncoding RNA molecules greater than 200 nts in length that have a wide range of functions. These functions include but are not limited to X chromosome inactivation, imprinting, epigenetic reprogramming, and other *trans*-mediated functions (Cabili et al. 2011, Chen and Carmichael 2008, Loewer 2010). Transcriptionally active Alu elements express Alu RNA molecules that display cell stress response-like behavior, regulating both transcription and translation in *trans*. An important note is that Alu expression can increase following various types of cell stress (Li and Schmid 2001, Liu et al. 1995, Panning and Smiley 1995, Rudin and Thompson 2001).

Alu RNA was shown to inhibit transcription during heat shock by interacting directly with RNA Pol II (Mariner et al. 2008). This transcription inhibition was rescued by treatment with antisense Alu RNA, which knocked down Alu transcript levels. Interestingly, this transcription repression was not universal. Transcription of

housekeeping genes, but not an 18S rRNA control, decreased while hsp70 transcription increased. Alu RNA can simultaneously associate with up to two RNA Pol II complexes at promoter regions of genes that are transcriptionally repressed following heat shock. Both the Alu left and right arms were shown to be able to separately associate with RNA Pol II. Additionally, scAlu, the processed left arm of Alu RNA, can also associate with RNA Pol II. However, only the Alu right arm, and not the Alu left arm or scAlu, could repress transcription by RNA Pol II. This is due to the presence of a structurally conserved repression domain in the 3' end of the right arm that is absent from both the Alu left arm and scAlu. Alu elements bind to RNA Pol II pre-initiation complexes and repress transcription initiation.

Like Alu elements, the murine specific transposable element B1 is also derived from the 7SL RNA. B1 elements share secondary structure homology with the Alu left arm and were shown to interact with RNA Pol II in a fashion similar to both the Alu left arm and scAlu (Mariner et al. 2008). TFIIF facilitated the dissociation of B1 and scAlu from RNA Pol II (Wagner et al. 2010). This function was inhibited by introducing the repression domain from the Alu right arm into the B1 elements leading to transcriptional repression. In addition, B1 elements repressed transcription, but only in the absence of TFIIF. Transcription factors may modulate the Alu/RNA Pol II interaction at some but not all promoter regions following heat shock.

Alu elements also play a role in regulating translation. This is supported by a study showing increased reporter gene expression due to an increase in Alu expression. The Alu right arm was shown to be responsible for this process (Rubin et al. 2002). Though Alu RNA displayed a positive influence on translation, Alu bound to SRP9/14

subunit of the SRP had an inhibitory effect on translation initiation *in vitro*, though these RNP molecules did not affect translation elongation (Hasler and Strub 2006). This SRP9/14 dependent inhibition was observed when bound to either scAlu or Alu right arm, indicating no preference for either Alu arm in SRP9/14 dependent regulation. In addition, polysome profiles revealed a decrease in the level of monosomes and polysomes present on a reporter transcript upon addition of Alu/SRP9/14, but an increase upon addition of only Alu RNA. This supports the variation in translational levels in the presence of Alu RNA and Alu/SRP9/14. Interestingly, it is the Alu domain of the SRP that is involved in delaying translation elongation following translocation signal sequence recognition by SRP54 (Huck et al. 2004, Wild et al. 2004).

Conclusion

A-to-I RNA editing is highly prevalent in Alu repetitive elements embedded in protein coding genes expressed by RNA Pol II. However Alu elements can also be expressed by RNA Pol III through their own internally located promoters. Neither this group of transcriptionally active Alu elements nor RNA Pol III transcripts, in general, have previously been shown to undergo ADAR catalyzed A-to-I RNA editing.

The following dissertation seeks to understand the relationship between A-to-I RNA editing and transcriptionally active Alu elements. The following three chapters describe the endeavor taken to demonstrate A-to-I RNA editing in Alu RNA expressed by RNA Pol III. Using a known editing target, the R/G site from GluR-B, expressed under an RNA Pol III promoter, RNA Pol III transcripts are shown to undergo A-to-I RNA editing by ADAR. This also demonstrates that A-to-I RNA editing can occur independent of the RNA Pol II CTD, a region highly involved in regulating and

coordinating post-transcriptional RNA modification. Using a similar strategy, inverted Alu repeats, but not single Alu elements, expressed by RNA Pol III are shown to undergo promiscuous A-to-I RNA editing. In addition, editing in these transcripts is revealed to be dependent on the formation of dsRNA between inverted Alu pairs. Finally, an *in vivo* investigation of A-to-I RNA editing in transcriptionally active Alu elements shows evidence of basal level editing in human brain tissue.

The closing chapter discusses the potential significance that A-to-I editing in transcriptionally active Alu elements could have on human evolution, Alu retrotransposition, and the different functional properties that Alu elements have been shown to have. This chapter concludes by detailing future experiments that could be undertaken to expand our understanding of the impact of A-to-I RNA editing on transcriptionally active Alu elements.

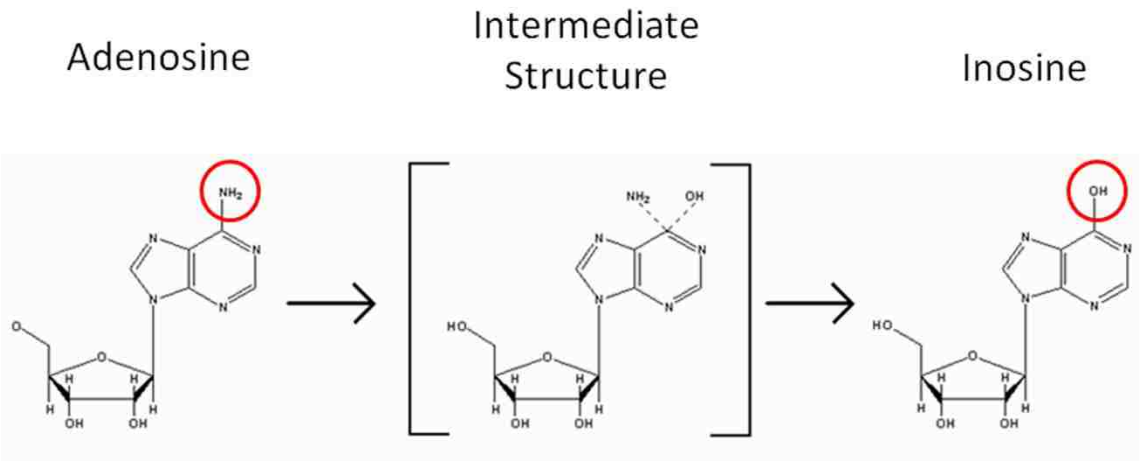
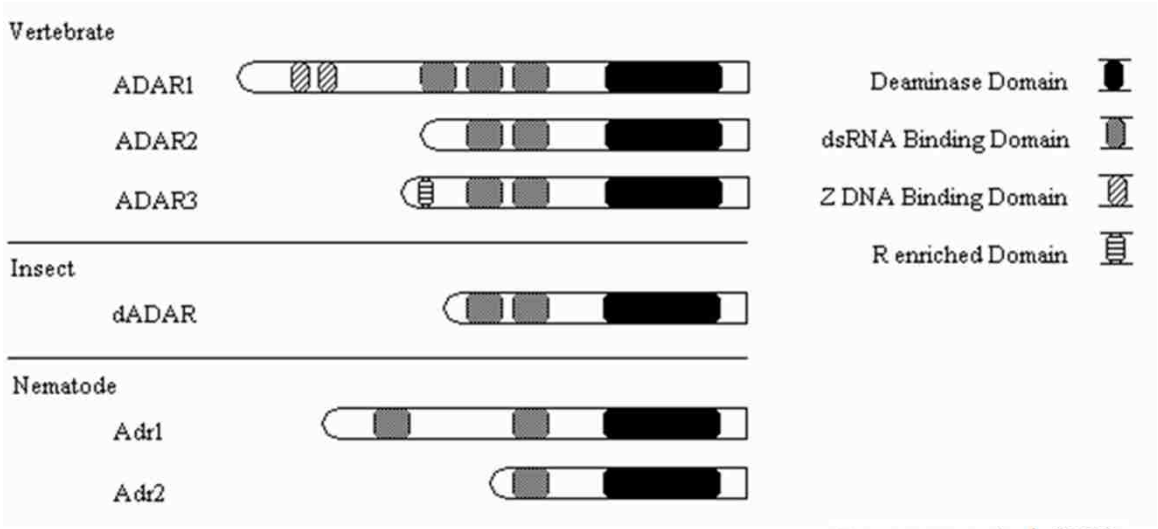


Figure 1.1- A-to-I RNA editing reaction. The amine group attached to carbon-6 in the purine ring is targeted for deamination resulting in the presence of a hydroxyl group. This hydroxyl group imparts the base pairing characteristics of guanosine on the edited nucleotide, causing inosine to be interpreted as guanosine by much of the cellular machinery.



Gommans et al. 2008

Figure 1.2 - A-to-I RNA Editing is Catalyzed by the ADAR Family of Enzymes. Adenosine deaminases are conserved across metazoans. In general they share two functional domains, dsRNA binding domains near the N-terminus, and a catalytic deaminase domain near the C-terminus.

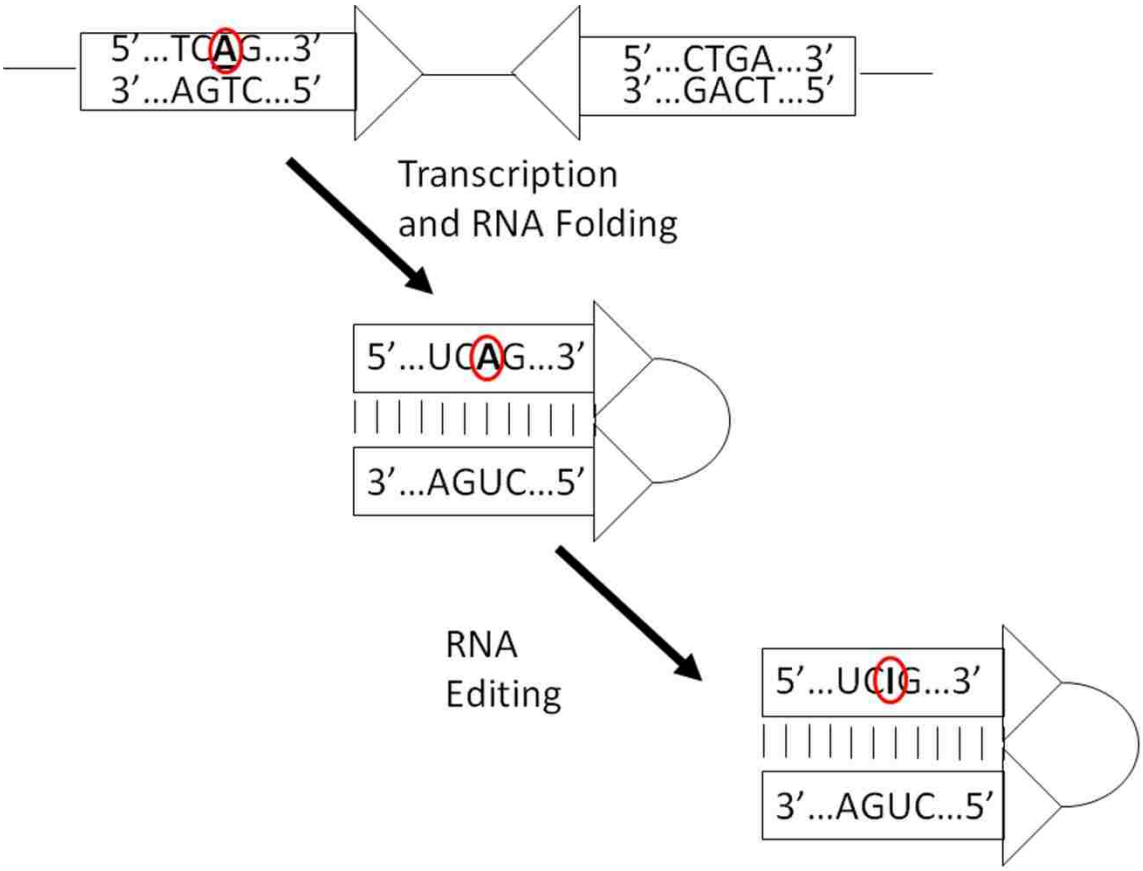


Figure 1.3 – A-to-I RNA editing occurs in dsRNA. Complementary regions present within the RNA transcript base pair with each other, creating a dsRNA molecule. Individual adenosines present within the dsRNA region are targeted by ADAR for deamination.

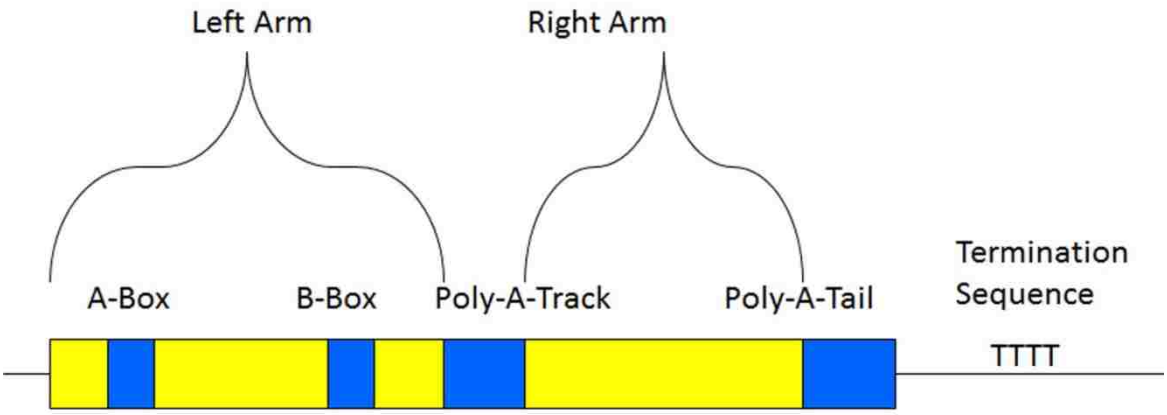


Figure 1.4 - The General Structure of an Alu Element. Alu elements have two arms separated by a central-A track. The left arm contains A-Box and B-Box RNA Pol III promoter elements. The Alu element ends in a poly-A tail at the 3'-end.

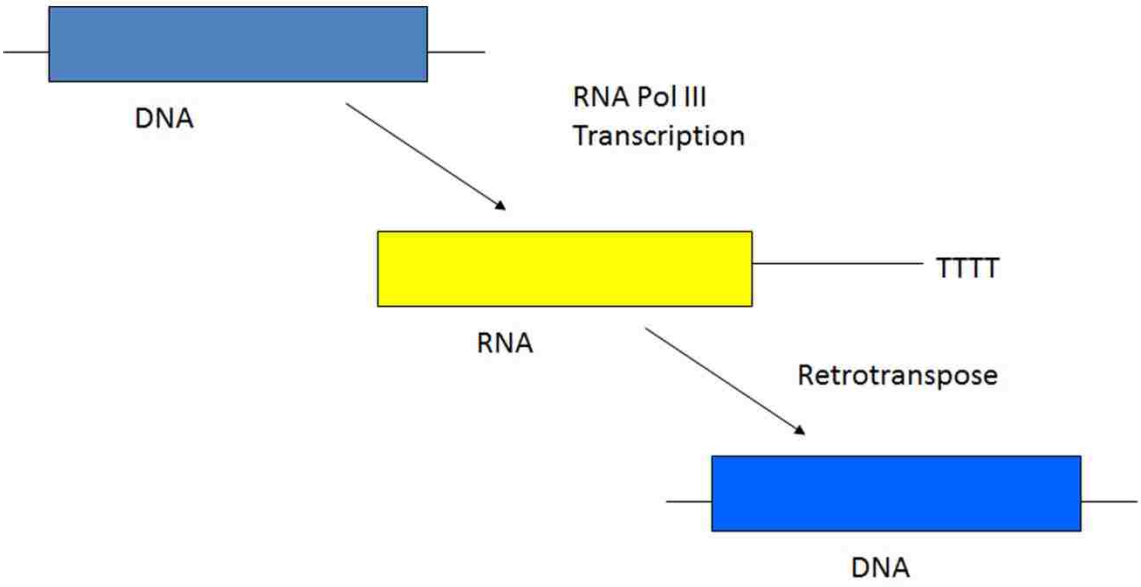


Figure 1.5 – Alu Elements Retrotranspose Via an RNA Intermediate. This RNA intermediate is transcribed by RNA Pol III. Transcription continues until reaching a transcription termination site that can be located at varying distances away from the Alu 3'-end. However, only the Alu element undergoes retrotransposition.

Chapter 2

RNA Polymerase III Transcripts Are A-to-I RNA Editing Targets

Introduction

A-to-I RNA editing is the conversion of adenosine to inosine via a deamination reaction performed by the ADAR family of enzymes. This is significant in terms of gene expression because inosine shares base-pairing characteristics with guanosine, and is thus interpreted as guanosine by much of the cellular machinery. This allows A-to-I RNA editing to influence the coding sequence and alternative splicing of mRNAs (Sommer et al. 1991, Lev-Maor et al. 2007).

One example of an A-to-I RNA editing target is the R/G editing site in glutamate receptor B (GluR-B) mRNA. This editing site is located in the last codon of exon 13 and forms a hairpin structure with a complementary region located immediately within intron 13 (Lomeli et al. 1994). Editing results in an arginine, R, to glycine, G, change in the codon. Proteins with the edited amino acid display a reduced recovery time in neurons following membrane depolarization.

Co-transcriptional processing, including 5' capping, alternative splicing, and polyadenylation, are coordinated by a region of RNA Pol II termed the C-Terminal Domain (CTD) (Buratowski 2009, Fong and Bentley 2000, Hirose and Okhuma 2007). This region is composed of 52 heptad repeats that can undergo reversible phosphorylation during transcription. The CTD can be roughly divided into two regions based on processes they are associated with. Heptads 1-25 are involved in transcription initiation and the switch from initiation to elongation, while heptads 27-52 are involved in coordination of co-transcriptional RNA processing (Buratowski 2009, Fong and Bentley 2000).

The CTD has previously been shown to be involved in coordinating ADAR2 directed A-to-I RNA editing with splicing (Bratt and Ohman 2003, Laurencikiene et al. 2006, Ryman et al. 2007, Schoft et al. 2007). Proper coordination of these events is necessary since dsRNA structure necessary for editing can often be formed between exonic and intronic regions (Bratt and Ohman 2003, Schoft et al. 2007). A previous study investigating coordination of editing and splicing at the R/G site of GluR-B indicated that the presence of the RNA Pol II CTD does not affect editing when the splice site is inactive (Ryman et al. 2007). However, the CTD is necessary for efficient editing when splicing is enabled. That same study investigated coordination of editing and splicing in the Q/R editing site in GluR-B. Whereas the R/G editing site is located 2 nucleotides upstream of the end of exon 13, the Q/R site is located more centrally within exon 11 (Higuchi et al. 1993, Lomeli et al. 1994). The regions necessary for dsRNA formation to allow editing at the Q/R site are more distantly separated relative to the R/G site. In this case the CTD increased editing by ADAR2 while showing an overall enhancement of splicing dependent on editing at the Q/R site (Ryman et al. 2007). A third study investigated a self-targeted editing site within ADAR2 mRNA that leads to creation of an alternatively spliced exon (Laurencikiene et al. 2006). In this case, editing and thus inclusion of the alternative exon behaves as a feedback inhibitory mechanism regulating ADAR2 function. This study indicated that full CTD deletion, rather than a partial CTD deletion, was necessary to significantly impair alternative splicing. Together these studies indicate that ADAR2 and the spliceosome can compete for shared targets. Also, CTD function becomes more important in ensuring that editing precedes splicing as

the distance between editing complementary regions increases across an exon/intron junction.

RNA Pol III is a polymerase responsible for expression of non-protein coding housekeeping genes, such as tRNA, the U6 RNA of the spliceosome, and the 7SL RNA of the signal recognition particle, among others (Dieci et al. 2007, Nikitina et al. 2011). Promoter elements for RNA Pol III can be located both upstream of the transcription start site and internally within the gene (Roy et al. 2000). This is in contrast to RNA Pol II promoter elements which are mostly located upstream of the transcription start site (Baumann et al. 2010). In addition, RNA Pol III lacks a CTD-like structure that is present in RNA Pol II (Nikitina et al. 2011). While RNA Pol III genes undergo their own set of post-transcriptional processes, such as nucleotide modification in tRNA, they are wholly different from those processes observed in transcripts expressed by RNA Pol II (Su and Randau 2011). To be clear, tRNA can undergo A-to-I RNA editing by an ADAR homolog called adenosine deaminase acting on tRNA, ADAT. This enzyme specifically targets tRNA for deamination.

To date, A-to-I RNA editing by ADAR has not been shown in RNA Pol III transcripts. Also, the role of the RNA Pol II CTD in regulating ADAR1 activity has not been investigated. Here I use a known editing target, the R/G site in GluR-B, to investigate the ability of RNA Pol III transcripts to be edited by ADAR1 and ADAR2. I also use RNA Pol II CTD variants to study how the CTD effects ADAR1 editing. I show that RNA Pol III transcripts can be edited, and that editing efficiency in these transcripts differs between ADAR1 and ADAR2. Also, the partial deletion of RNA Pol II CTD increases A-to-I RNA editing efficiency by ADAR1 but not ADAR2.

Methods and Materials

Plasmid Construction

The RNA Pol III expression backbone, pBS-U6-Term, was built using pBS as a base. The U6 promoter was amplified by PCR from human genomic DNA using primers U6-F and U6-R. These primers had a BamHI and PstI restriction sites, respectively, that were used to ligate the sequence into the vector. The termination sequence was generated by annealing together Terminator-F-Oligo and Terminator-R-Oligo. These were ligated into the vector using HindIII and XhoI restriction sites. The use of the U6 promoter and an RNA Pol III terminator for construction of RNA Pol III expression vectors has previously been described (Chu et al. 1995, Roy et al. 2000). To make pBS-U6-RG, the R/G editing site from GluR-B was amplified by PCR using primers RG-F-PstI and RG-R-HindIII. These contain PstI and HindIII restriction sites, respectively, which were used to ligate the target sequence into the vector.

RNA Pol II expression constructs, pCI-CMV-RG, used a pCI backbone. The target sequence was amplified by PCR from pBS-U6-RG using primers RG-F-EcoRI and RPBS. The sequence was cleaved with EcoRI and KpnI which was used for the ligation.

Plasmids pAT7RpbWT and pAT7Rpb1-25 were kindly donated by David Bentley from University of Colorado, Denver.

Cell Culture and Transfections

HeLa cells were grown to 50% confluence in 6 well plates in growth media containing minimal essential media, 10% fetal bovine serum, and 1x antimycotic solution. Cells were grown at normal growth conditions at 37° C with 5.5% CO_2 and 88% humidity. Transfections were performed in 3.5 cm dishes using Superfect reagent

according to the manufacturer's protocol. 1 ug of either pBS-U6-RG or pCI-CMV-RG and 1ug of pEX-ADAR1, pEX-ADAR2 was used in each reaction. Cells were incubated for 3 hours at normal growth conditions following addition of the transfection mixture. Each well was then washed with PBS, then received 3 mL growth media and incubated at normal growth conditions for 48 hours. Transfections using pAT7-RpbWT and pAT7-Rpb1-25 were performed using a modified protocol with Superfect reagent. 1.25 ug pAT7-RpbWT or pAT7-Rpb1-25 was used with 0.75 ug target sequence and 0.75 ug ADAR expression vectors. 50% additional Superfect reagent was used with each transfection. All other steps were performed according to manufacturer's protocol. Cells were grown for 30 hours, then alpha-amanitin was added to growth media at 3.5 mg per mL. Cells were then grown for 24-48 hours at normal growth conditions.

RNA Isolation and RT-PCR

RNA was isolated using TRIzol Reagent according to the manufacturer's protocol. RNA was then treated with Turbo DNase in the presence of RNAsin and incubated for 1 hour at 37°C to remove contaminating plasmid DNA. Reactions were stopped and purified by phenol-chloroform reactions followed by ethanol precipitation. RNA purity was tested by PCR. DNase treatments were repeated as necessary.

1 to 3 ug total RNA was reverse transcribed using Superscript III reverse transcriptase with random hexamers according to the manufacturer's protocol. PCR was performed using Taq polymerase. For amplification of the R/G sequence expressed by the U6 promoter, forward primer RG-F-PstI was used. Sequences expressed by CMV promoter were amplified using forward primer PCI-1D. T-Site-R-KpnI was used as the

reverse primer in all PCR amplifications. Reactions were gel purified and sequenced by Sanger sequencing.

Statistical Evaluation

All transfections were, at minimum, performed in triplicate with each RT-PCR product sequenced three times. The averages and standard deviations of each experimental condition were compared using Student's T-Test with a p-value of p=0.05.

Results

ADAR Directed A-to-I RNA Editing Occurs in RNA Pol III Transcripts

To determine if RNA Pol III transcripts can undergo A-to-I RNA editing, a plasmid was developed to express a known editing target by RNA Pol III. The R/G editing site and complementary editing sequence from GluR-B was cloned into a pBS vector downstream of a RNA Pol III promoter (U6) and upstream of an RNA Pol III termination site. This generates a dsRNA molecule 171 nt in length with a double stranded hairpin 38 nt long (figure 2.1A). The R/G editing sequence plus the termination sequence were then amplified by PCR and cloned into a pCI vector downstream of an RNA Pol II promoter (CMV) to be used as a positive control (figure 2.1B).

HeLa cells were transiently transfected with either the RNA Pol II or RNA Pol III expression vector along with either a constitutively active ADAR1 or ADAR2 expression vector, or in the presence of endogenous ADAR activity. RNA was isolated using TRIzol Reagent, purified with DNase and then amplified by RT-PCR. Sanger sequencing showed that the RNA Pol III transcribed R/G editing site underwent A-to-I RNA editing (figure 2.1C).

A-to-I RNA editing of RNA Pol III transcripts was significantly higher (p=0.05) when ADAR2 was overexpressed (50.4% +/- 3.4), in comparison to editing when ADAR1 was overexpressed (20.7% +/- 1.2) or in the presence of endogenous ADAR activity, (23.9% +/- 3.6). Editing of RNA Pol III transcripts was not significantly different when ADAR1 was overexpressed, relative to editing under endogenous ADAR activity. This indicates A-to-I RNA editing at the R/G site is catalyzed by ADAR2, but not ADAR1, when the editing target is transcribed by RNA Pol III.

The RNA Pol II expressed editing target was most efficiently edited when ADAR2 was overexpressed (93.3% +/- 1.3) (figure 2.1D). This was significantly higher (p=0.05) than the amount of editing observed when ADAR1 was overexpressed (63.9% +/- 2.2) or in the presence of endogenous ADAR activity (5.5% +/- 0.67). Editing by ADAR1 was also statistically higher than editing in the presence of endogenous ADAR activity.

Comparing editing efficiency based on promoter type showed additional differences in editing when ADAR1 or ADAR2 was overexpressed (figure 2.1E). The editing target transcribed by RNA Pol II was more efficiently edited (p=0.05) than its RNA Pol III counterpart when either ADAR1 or ADAR2 was overexpressed. However, editing in the presence of endogenous ADAR activity was significantly higher (p=0.05) in RNA Pol III transcripts. This suggests that transiently transfected ADAR1 or ADAR2 genes are differentially regulated relative to endogenous ADAR enzymes.

Transient transfection of HeLa cells with a known A-to-I RNA editing target demonstrates that genes expressed by RNA Pol III can undergo ADAR directed A-to-I

RNA editing. However, editing efficiency varies based on both the promoter type as well as the type of ADAR being overexpressed.

ADAR Type but not Promoter Type Influences Target Expression

One possible explanation for why editing was higher when ADAR2 was overexpressed compared to ADAR1 overexpression is due to differences in expression levels of the target gene. ADAR1, but not ADAR2, overexpression has previously been shown to cause an increase in RNA levels of transiently transfected reporter genes (Gommans and Maas 2008). Higher expression of the editing target when ADAR1 is overexpressed would shift the substrate-to-enzyme ratio. This would result in the observation of lower editing efficiency when ADAR1 is overexpressed relative to when ADAR2 is overexpressed.

In order to determine if observed differences in A-to-I RNA editing were due to differences in the expression level of the target sequence, qRT-PCR was performed (figure 2.1F). Expression levels relative to GAPDH were normalized to target gene expression by RNA Pol III in the presence of overexpressed ADAR1. Expression levels were significantly lower ($p=0.05$) when ADAR1 was overexpressed relative to when ADAR2 was overexpressed. This was observed under both RNA Pol II and RNA Pol III expression, indicating that differences in editing of RNA Pol II and RNA Pol III expressed targets were not due to differences in target site expression.

Furthermore, a comparison of expression levels between the two types of polymerases when either ADAR1 or ADAR2 was overexpressed showed that expression levels were not significantly different. Thus, while ADAR overexpression but not promoter type influenced transcription, higher RNA levels were paired with higher levels

of editing. This means that differences observed in editing were not influenced by differences in RNA levels.

RNA Pol II CTD Influences A-to-I RNA Editing by ADAR1 but not ADAR2

A key structural difference between RNA Pol II and RNA Pol III is the CTD in RNA Pol II. The C-terminal half of the CTD has previously been shown to regulate post-transcriptional RNA processing, such as alternative splicing and the addition of the 5' cap (Fong and Bentley 2000). In addition, the CTD has previously been shown to be important in coordinating A-to-I RNA editing by ADAR2 with alternative splicing (Laurencikiene et al. 2006, Ryman et al. 2007). In order to determine if A-to-I RNA editing by ADAR1 is influenced by the CTD, two different RNA Pol II expression plasmids were used that express either a full length CTD (pAT7Rpb-WT) or a truncated CTD (pAT7Rpb-1-25) that lacks heptad repeats 26-52 (Fong and Bentley 2000). The expressed RNA polymerases are resistant to an RNA Pol II inhibitor, α-amanitin. This allows for knock down of endogenous RNA Pol II activity following transient transfection.

Editing analysis indicated that A-to-I RNA editing was unaffected by differences in CTD length when ADAR2 was overexpressed (figure 2.2). This was consistent with previous studies showing the CTD does not influence ADAR2 activity at the R/G site when splicing is inactive (Ryman et al. 2007). However, full length CTD reduced A-to-I RNA editing ($p=0.05$) when ADAR1 was over expressed and in the presence of endogenous ADAR activity. This implies that the repression in ADAR1 activity by full length CTD, relative to the truncated CTD, is the reason for the difference in editing efficiency in the presence of endogenous ADAR activity. Also, the reduction of editing

in the case of ADAR1 overexpression was significant at p=0.05 but not at p=0.025. This indicates that the CTD influence on ADAR1 activity may be only marginal, and would potentially not be significant with larger sample sizes. These experiments also show that ADAR editing activity can also occur post-transcriptionally rather than only co-transcriptionally.

Discussion

ADAR directed A-to-I RNA editing in transcripts expressed by RNA Pol III has not been previously shown. Here a known editing target, the R/G site from GluR-B, was expressed under either an RNA Pol II or RNA Pol III promoter. Editing was evaluated in the presence of either ADAR1 or ADAR2 overexpression, or in the presence of endogenous ADAR activity. Editing efficiency was influenced by both polymerase and ADAR type. Furthermore, editing in the presence of either ADAR1 or endogenous ADAR activity, but not ADAR2, increased when the CTD of RNA Pol II was truncated. The effect on ADAR1 was significant at p=0.05 but not p=0.025. The lack of difference in ADAR2 editing by partial CTD deletion is consistent with previous studies of ADAR2 that showed no influence by the CTD on editing at the R/G site when splicing was disabled (Ryman et al. 2007). In addition, in editing targets where splicing is enabled, ADAR2 activity is only effected by complete deletion of the CTD (Laurencikiene et al. 2006).

My study shows that RNA Pol III transcripts can be A-to-I RNA editing targets. It also shows that ADAR1 and ADAR2 have different editing efficiencies when targeting the R/G site. This supports previous work that demonstrated ADAR dependent differences in editing efficiency at the R/G site (Lai et al. 1997). My work expands on

this by showing that these ADAR dependent differences in editing efficiency are present in RNA Pol III transcripts as well. Also interesting was that ADAR1 overexpression did not increase A-to-I RNA editing in RNA Pol III transcripts beyond editing levels observed in the presence of endogenous ADAR activity. This indicates that ADAR1 does not target the R/G editing site under RNA Pol III expression. In order to determine if ADAR1 directed editing of all RNA Pol III transcripts is similar, investigation of additional targets will be needed.

Target gene expression levels were examined to try to account for differences in editing. Since editing was lower when ADAR1 was overexpressed relative to ADAR2 overexpression, it is possible that the difference could be due to differences in target gene expression. However, while less editing was seen in the presence of ADAR1, less of the target RNA was present. This indicates that editing levels were not influenced by levels of target gene expression. This runs counter to a previous study that showed ADAR1 overexpression causes an increase in RNA of transiently transfected editing reporter genes (Gommans and Maas 2008). This effect was not observed when the reporter gene was genomically integrated. In addition, ADAR1 knockdown in hESCs resulted in a decrease in RNA levels in genes associated with nucleic acid and protein metabolism, but an increase in RNA levels in genes associated with neurogenesis and cell differentiation (Osenberg et al. 2010). The basis for these differing effects on RNA levels caused by ADAR1 overexpression remains unknown.

Another interesting observation from the RNA Pol II/Pol III comparison was that editing efficiency was lower in RNA Pol III transcripts, but only when ADAR1 or ADAR2 was overexpressed. This dichotomy presents some interesting scenarios as to

what may be occurring. A plausible hypothesis is that endogenous ADARs are typically sequestered away from RNA Pol II transcripts but not from RNA Pol III. However, overexpression of ADAR overloads this regulatory mechanism, allowing for greater access to RNA Pol II transcripts. The nucleolus is a potential site of ADAR localization that may play this role (Desterro et al. 2005, Sansam et al. 2003, Vitali et al. 2005). Both ADAR1 and ADAR2 have been shown to be able to localize to the nucleolus. ADAR1, but not ADAR2, can be SUMOylated, a post-translational modification associated with nucleolar localization (Desterro et al. 2005). Though nucleolar sequestration was shown to be SUMO-independent, SUMOylation did inhibit ADAR1 editing activity. In addition, ADAR2 sequestration to the nucleolus was shown to decrease editing activity (Sansam et al. 2003, Vitali et al. 2005). Thus, sequestration may be one potential mechanism regulating ADAR activity.

The CTD of RNA Pol II has previously been shown to act as both a scaffold and coordinator of post-transcriptional processing of mRNA (Buratowski 2009, Fong and Bentley 2000, Hirose and Ohkuma 2007). While crosslinking experiments have been unsuccessful in demonstrating an association between the CTD and either ADAR1 or ADAR2, CTD deletion studies demonstrated the role the CTD plays in coordinating editing with splicing (Laurencikiene et al. 2006, Nishikura 2010, Ryman et al. 2007). This indicates at least a transient relationship between RNA Pol II and ADAR.

Previous research has shown that the CTD of RNA Pol II influences A-to-I RNA editing. For the R/G site from GluR-B, the CTD was only necessary for highly efficient editing if splicing was enabled (Ryman et al. 2007). In addition, removal of only half of the CTD was insufficient in influencing A-to-I RNA editing by ADAR2 (Laurencikiene

et al. 2006). This was supported by CTD-deletion experiments done here that showed no difference in A-to-I RNA editing in the presence of either wild-type or truncated CTD when ADAR2 was overexpressed. This was not surprising given the fact that the exon 13 splice site is inactive in the R/G target gene used for this study. Taken together, this shows that ADAR2 can act independently from the co-transcriptional regulatory activity of the RNA Pol II CTD. This is supported by the presence of A-to-I RNA editing in transcripts expressed by RNA Pol III, which lacks a CTD.

The influence of the CTD on ADAR1 has not previously been demonstrated. ADAR1 activity was significantly lower in the presence of wild-type CTD compared to the truncated CTD. However, this difference was only significant at $p=0.05$. This indicates that this difference may disappear with larger sample sizes. However, if this difference truly is significant, this would be counter to previous studies of ADAR2 activity which indicated that a partial CTD deletion was insufficient to influence A-to-I RNA editing when splicing near the editing site was enabled (Laurencikiene et al. 2006). Only full CTD deletion influenced editing by ADAR2, but only in splicing competent editing targets (Laurencikiene et al. 2006, Ryman et al. 2007). This would also indicate that ADAR1 and ADAR2 have different sensitivities to regulation by the CTD, and that heptads 27-52 of the CTD inhibit ADAR1 activity. Furthermore, the higher amount of A-to-I RNA editing in the presence of endogenous ADAR activity due to partial CTD deletion would be attributed to an increase in ADAR1 activity. This conclusion was reached with the observation that ADAR1 but not ADAR2 activity increased with the partial CTD deletion.

A-to-I RNA editing has not been previously shown in RNA Pol III transcripts. Using a known editing target expressed under an RNA Pol III promoter, this species of RNA was shown to be targeted by A-to-I RNA editing. Comparisons of editing efficiency demonstrated that although RNA Pol III transcripts can undergo A-to-I RNA editing, editing efficiency was lower than when the same target was expressed by RNA Pol II. In addition, editing in the presence of ADAR2 was higher than in the presence of ADAR1. Finally, partial deletion of the RNA Pol II CTD showed that editing occurs post-transcriptionally rather than only co-transcriptionally.

Primer	Primer Sequence
U6-F	ACGAGGATCCAAGGTCGGGCAGGAAGAGGGCC
U6-R	GTAACTGCAGGTGTTTCGTCCTTTCCACAAG.
Termination-F-Oligo	AGCTTGACCATGGATTACATATGATCGACAGACTAGGTACGACGATACTAGCGTTTTGAC
Termination-R-Oligo	TCGAGTCAAAACGCTAGTATCGTCGTACCTAGTCTGTCGATCATATGTAATCCATGGTCA
RG-F-PstI	ATATCTGCAGCCTGGATTCCAAAGGCTATGGC
RG-R-HindIII	ATATAAGCTTAAGATACATCAGGGTAGGTGGG.
RG-F-EcoRI	AAGAATTCCCTCGATTCCAAAGGCTATGGC
RPBS	TTGTAATACGACTCACTATAGG
PCI-1D	CACAACAGTCTCGAACTTAAGC
T-Site-R-KpnI	CTAGAGGTACCGTATCGTCGTACCTAGTCTGTCG

Table 1- Table of Primers

.

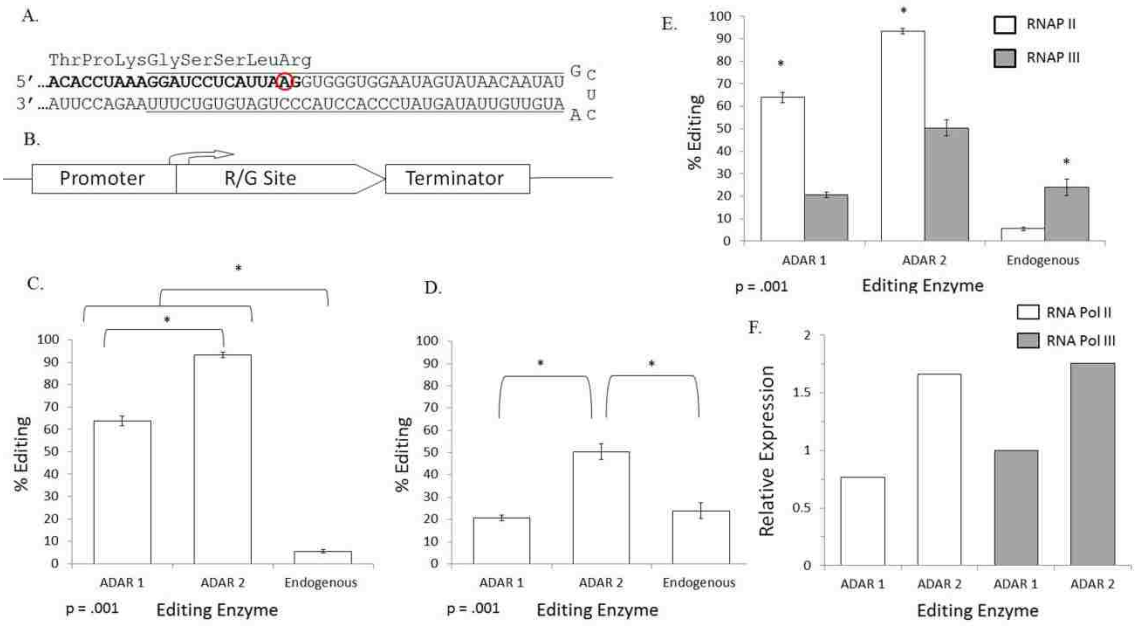


Figure 2.1 – The R/G site from GluR-B is an A-to-I RNA editing target when expressed by RNA Pol III. A) The R/G editing site from GluR-B is displayed. The editing site (red circle) is in the last codon of exon 13 (bold text). The hairpin forms a dsRNA region (Lines) 38 nucleotides in length with a 5 nucleotide loop. B) The general structure of the R/G expression vector. Expression vectors varied by promoter used to express the editing target by either RNA Pol II (CMV) or RNA Pol III (U6). C) Comparison of A-to-I RNA editing in RNA Pol III transcripts in the presence of overexpressed ADAR1, ADAR2 or endogenous ADAR activity. D) Comparison of A-to-I RNA editing in RNA Pol II transcripts in the presence of overexpressed ADAR1, ADAR2 or endogenous ADAR activity. E) Polymerase dependent comparison of editing based on ADAR1 or ADAR2 overexpression, or endogenous ADAR activity. F) qRT-PCR results relative to GAPDH were normalized to expression by RNA Pol III in the presence of overexpressed ADAR1. This displays expression levels of the R/G editing target based on both promoter type and ADAR overexpression.

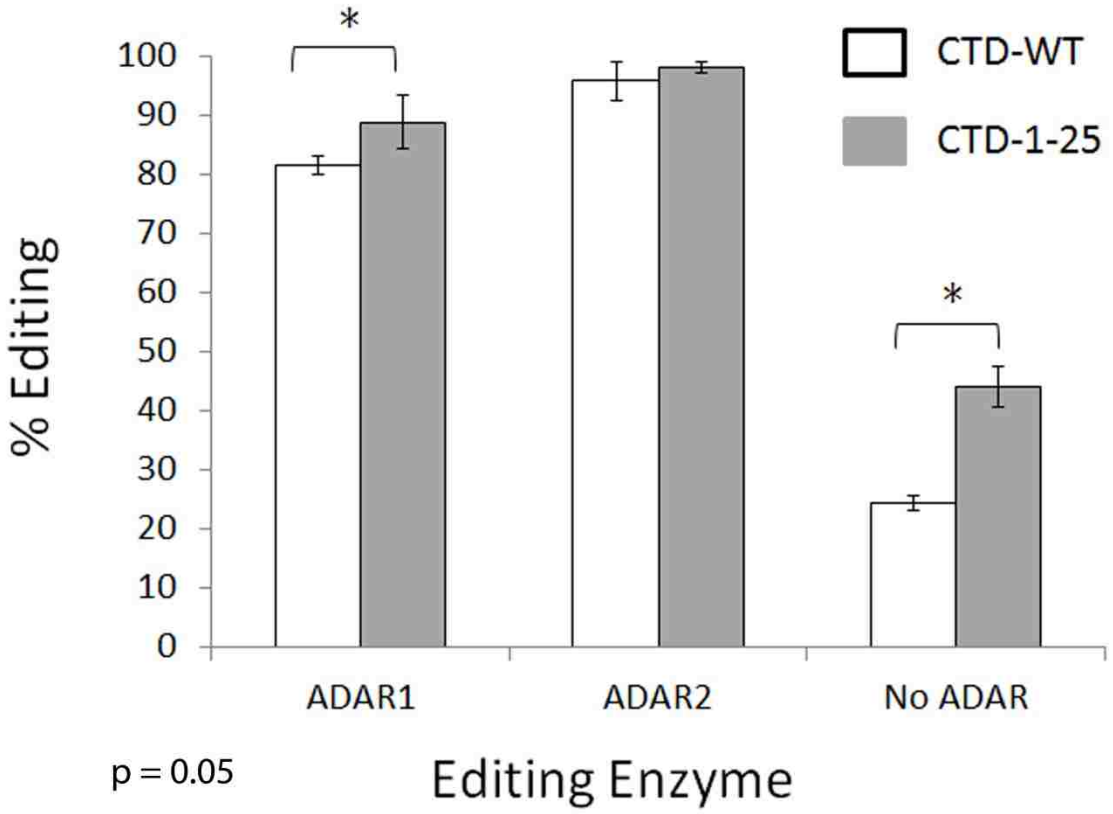


Figure 2.2. The influence of the RNA Pol II CTD on ADAR function. RNA Pol II containing full length (CTD-WT) or a truncated CTD (CTD-1-25) was used to express the R/G editing target in HeLa cells. Editing was evaluated in the presence of overexpressed ADAR1 or ADAR2, or in the presence of endogenous ADAR activity.

Chapter 3

Transcriptionally Active Alu Elements Are A-to-I RNA Editing Targets

Introduction

A-to-I RNA editing is a post-transcriptional modification that targets double stranded RNA molecules. Alu elements are primate specific repetitive elements present in over 1 million copies in the human genome. Due to their high copy number, Alu elements are often expressed as inverted pairs within pre-mRNA, allowing formation of a double stranded structure that is permissive for ADAR binding and subsequent editing. As such, Alu elements constitute major editing targets when embedded within pre-mRNA molecules (Athanasiadis et al. 2004, Peng et al. 1012). Not only are Alu elements highly targeted as editing substrates, but they can be edited at multiple adenosines (Athanasiadis et al. 2004, Osenberg et al. 2010).

These editing events can have important regulatory implications on gene expression. Editing within an embedded Alu element present in intron 8 of nuclear prelamin-A recognition factor, NARF, results in the creation of an alternatively spliced Alu exon (Lev-Maor et al. 2007). Within the 3' UTR, promiscuous editing in Alu elements can result in a decrease in translation via a nuclear retention mechanism that targets highly edited RNA substrates, and can lead to cleavage by Tudor-SN (Chen et al. 2008, Chen and Carmichael 2009, Scadden 2005, Scadden 2007). Furthermore, editing in Alu elements within the 3' UTR can both create and destroy miRNA target sites (Borchert et al. 2009).

In addition to being expressed embedded within mRNA molecules, Alu elements can also be expressed via their own internally located promoter elements (Chu et al. 1995, Li and Schmid 2001, Roy et al. 2000). These A-box and B-box promoter elements

allow for expression via RNA Pol III (Kenneth et al. 2008). Alu elements expressed in such a way are termed transcriptionally active Alu elements, or transcribed Alu elements (Li and Schmid 2001). Transcriptionally active Alu elements are the source for Alu retrotransposition (Bennett et al. 2008, Britten et al. 1988, Dewannieux et al. 2003). Additionally, transcribed Alu elements have been shown to have cell stress response-like behaviors (Mariner et al. 2008, Wagner et al. 2010). They associate with RNA Pol II following heat shock, leading to decreased transcription of non-cell-stress response related genes. Transcribed Alu elements have also been shown to associate with the ribosome both independently or in complex with SRP9/14 subunit of the SRP (Hasler and Strub 2006, Rubin et al. 2002). These associations result in altered translation rates. The role of Alu elements during cell stress is supported by the finding that Alu expression increases following various types of cell stress (Li and Schmid 2001, Liu et al. 1995, Panning and Smiley 1995, Rudin and Thompson 2001).

Editing in inverted Alu pairs is prevalent when embedded within mRNA (Athanasiadis et al. 2004, Peng et al. 1012, Osenberg et al. 2010). However, A-to-I RNA editing in transcriptionally active Alu elements has not been demonstrated. A-to-I RNA editing may be inhibited by factors such as the type of polymerase responsible for transcription and an inability to transcribe the inverted Alu element for a number of reasons. Chief among these is the commonality of RNA Pol III transcription terminators, which are typically encoded as four or five consecutive thymine nucleotides (Chu et al. 1995, Chu et al. 1997, Gunnery et al. 1999). Although RNA Pol III termination sites are not 100% efficient, their presence between two inverted Alu elements could limit dsRNA formation by preventing inclusion of the inverted Alu element within RNA Pol III

transcripts. In addition, Alu elements contain two poly-A regions, a central poly-A track and a 3' poly-A tail (Batzer et al. 1996). When inverted, these regions are encoded as poly-T regions and could potentially behave as RNA Pol III termination sites, further limiting dsRNA formation.

To date, all published research on A-to-I RNA editing within Alu elements has focused on embedded Alu elements within mRNA and pre-mRNA. Here, I use Alu expression vectors to investigate transcribed Alu elements for their potential to undergo A-to-I RNA editing. In addition, I conduct a brief survey of structural characteristics necessary for editing to occur in these transcripts. Key amongst these structures are potential termination sites within the inverted Alu element. I show that Alu elements expressed by RNA Pol III can undergo A-to-I RNA editing, and that editing is dependent on the presence of an inverted Alu element within the RNA molecule.

Methods and Materials

Plasmid Construction

Alu elements expressed by RNA Pol III used the RNA Pol III expression backbone pBS-U6-Term. The Alu element with gene accession number AF344193 was used as the target sequence, and was amplified by PCR using nested primers (Li and Schmid 2001). The first PCR used forward primer TarAlu-F1, the second PCR used forward primer TarAlu-F-PstI, which contains a PstI restriction site (Table 3.1). Reverse primer TarAlu-R-HindIII containing a HindIII restriction site was used in both reactions. These restriction sites were used to ligate the sequence in pBS-U6-Term to make pBS-U6-Alu-Term. The Inverted Alu element was amplified by PCR using nested primers. The first PCR used forward primer InvAlu-F1 and reverse primer InvAlu-R1. The

second PCR used reverse primer InvAlu-R-KpnI, containing a KpnI restriction site. Forward primers InvAlu-F2T and InvAlu-F2dT were used for amplifying the inverted Alu element for vector pBS-U6-Alu-Term-ulA and pBS-U6-Alu-Term-ulAΔT, respectively. These contained XhoI restriction sites. To make pBS-U6-Alu-ulA and pBS-Alu-ulAΔT, the termination sequence was removed by two rounds of site-directed mutagenesis. The first SDM reaction used primers SDMF1 and SDMR1. The second SDM used primers SDMF2 and SDMR2.

RNA Pol II expression vectors used plasmid pCI. pBS-U6-Alu-Term, pBS-U6-Alu-Term-ulA, and pBS-U6-Alu-Term-ulAΔT were used as PCR templates to make the insert for pCI-Alu, pCI-Alu-ulA and pCI-Alu-ulAΔT, respectively. Forward primer PBSAlu-F-Eco and reverse primer RPBS were used. EcoRI and KpnI restriction sites were used to ligate the sequence into pCI downstream of the CMV promoter.

Cell Culture and Transfections

HeLa cells were grown to 50% confluence on 6 well plates in growth media containing minimal essential media, 10% fetal bovine serum, and 1x antimycotic solution. Cells were grown at normal growth conditions at 37° C with 5.0% CO_2 and 88% humidity. Transfections were performed in 3.5 cm wells using Superfect reagent according to the manufacturer's protocol. Transfections used 1 ug of the target sequence expression vector and 1ug of pEX-ADAR1, pEX-ADAR2, or no additional plasmid per well. Transfection proceeded for 3 hours at normal growth conditions. Cells were washed with 1x PBS, then incubated at normal growth conditions for 36-48 hours.

RNA Preparation and RT-PCR

RNA was isolated using TRIzol Reagent according to the manufacturer's protocol. 1 mL of TRIzol Reagent was used per well. Isolated RNA was treated with Turbo DNase for 1 hour at 37°C to remove contaminating DNA. Reactions were stopped and purified by phenol-chloroform purification followed by ethanol precipitation. RNA purity was tested by PCR. DNase treatments were repeated as necessary.

RNA was reverse transcribed using Superscript III reverse transcriptase and primed random hexamers. Between 1 and 3 ug of RNA was used per reaction. Alu elements expressed by pBS-U6-Alu and pCI-Alu expression vectors were amplified using forward primer 93-1D-Eco or PCI-1D, respectively. Reverse primer T-Site-R-Kpn was used in each reaction. Inverted Alu pairs expressed by RNA Pol III or RNA Pol II were amplified using forward primers 93-1D-Eco or PCI-1D or, respectively and reverse primer I-Site-R-Kpn. All PCR amplicons were purified by gel electrophoresis and were sequenced using Sanger sequencing method.

Subcloning and Blue/White Colony Screen

PCR amplicons were digested with EcoRI and KpnI and were ligated into pBS vector. Ligations were then transformed into z-competent cells and plated on LB-ampicillin plates containing x-gal for blue/white colony selection. Plates were grown overnight at 37°C. White colonies were selected and inoculated into LB media containing 100 ug/mL ampicillin and grown overnight at 37°C. Plasmids were then isolated using Qiagen Miniprep kit, and inserts were sequenced by Sanger sequencing using a primer for the T7 promoter.

Results

A single Alu element is insufficient for A-to-I RNA editing

A single Alu transcript is able to take on a relatively conserved secondary structure that contains some double stranded characteristics (Bennett et al. 2008, Hasler and Strub 2006, Huck et al. 2004, Mariner et al. 2008). However, comprehensive analysis of A-to-I RNA editing in an RNA transcript containing a single Alu element has not been published. To determine if a single Alu element is sufficient to behave as an A-to-I RNA editing substrate, expression vectors were designed to express a single Alu element under either an RNA Pol II (CMV) or RNA Pol III (U6) promoter (figure 3.1A,B.). These expression vectors were termed pCI-Alu and pBS-U6-Alu-Term, respectively. These were co-transfected into HeLa cells with either constitutively active ADAR1 or ADAR2 genes, or no additional ADAR.

RT-PCR products were sequenced by Sanger sequencing. Initial electropherograms did not reveal any evidence of A-to-I RNA editing (Figure 3.2). Since editing analysis by Sanger sequencing has a lower sensitivity threshold between 10% to 15% editing, all amplicons were subcloned into pBS vector for blue/white colony screening and miniprep extraction to allow for single sequence analysis (Luciano et al. 2004). 100 subcloned sequences of each sample were individually sequenced. No editing was observed regardless of either promoter type or presence of additional ADAR1 or ADAR2. This shows that a single Alu element is not targeted by A-to-I RNA editing.

Development of inverted Alu expression vectors

Inverted Alu elements embedded within mRNA are commonly edited at multiple adenosines, and represent major A-to-I RNA editing targets in the human transcriptome (Athanasiadis et al. 2004, Peng et al. 2012). To determine if inverted Alu pairs expressed by RNA Pol III can be targets of A-to-I RNA editing, four different RNA Pol III Alu

expression vectors were designed (Figure 3.1A). Using pBS-U6-Alu-Term as a starting point, an Alu element was ligated into the vector downstream of the termination site in an inverted orientation relative to the direction of gene expression. There were two variations of the inverted Alu element that were inserted into the vector. RNA Pol III termination sites are encoded as four consecutive thymines (Chu et al. 1995, Chu et al. 1997, Gunnery et al. 1999). When inverted, the poly-A tail of the inverted Alu element would be encoded as a run of thymines, and could potentially behave as a termination site (Batzer et al. 1996). Due to this possibility, the inverted Alu element either included or did not include (ΔT), the poly-A-tail. These vectors were termed pBS-U6-Alu-Term-ulA and pBS-U6-Alu-Term-ulAΔT, respectively.

Due to the design strategy used to construct these vectors, an RNA Pol III termination site was present between the two inverted Alu elements. Though RNA Pol III termination sites are not always 100% efficient, the efficiency of this termination site had not been tested. In order to eliminate the impact this transcription termination site could have on expression, site directed mutagenesis was used to destroy the termination site. These expression vectors were termed pBS-U6-Alu-ulA and pBS-U6-Alu-ulAΔT.

Vectors pBS-U6-Alu-Term-ulA and pBS-U6-Alu-Term-ulAΔT were used as templates to amplify the inverted Alu pairs by PCR. The amplified Alu pairs were then subcloned into pCI vector downstream of the CMV promoter to generate comparable Alu expression vectors dependent on RNA Pol II transcription (figure 3.1B). These were named pCI-Alu-ulA and pCI-Alu-ulAΔT. Only two of these expression vectors were made because RNA Pol III termination sites are not recognized as termination sites by RNA Pol II.

Inverted Alu elements expressed by RNA Pol III undergo A-to-I RNA Editing

The inverted Alu expression vectors were co-transfected into HeLa cells with expression vectors overexpressing either ADAR1 or ADAR2, or were expressed in the presence of endogenous ADAR activity. Both pBS-U6-Alu-Term-ulA and pBS-U6-Alu-Term-ulAΔT failed to transcribe beyond the termination site (figure 3.1C). This was confirmed by comparison of RT-PCR products using reverse primers located either immediately upstream (T-Site-R-Kpn) or immediately downstream (I-Site-R-Kpn) of the termination site. Sanger sequencing and single transcript analysis showed no evidence of A-to-I RNA editing (figure 3.2). Lack of editing here was not unexpected given the absence of the inverted Alu element within the RNA transcripts.

The absence of the termination site in pBS-U6-Alu-ulA and pBS-Alu-ulAΔT allowed for expression of longer transcripts (figure 3.1c). Sanger sequencing was used to observe A-to-I RNA editing in targets expressed by either vector (figure 3.2 and figure 3.3). The number of edited sites in pBS-U6-Alu-ulA in the presence of ADAR1, ADAR2 or endogenous ADAR was 24, 23, and 18, respectively, out of 42 total adenosines. For pBS-Alu-ulAΔT, the number of edited sites was 20, 23, and 15 edited adenosines out of 42 total. The number of total edited sites in each of the RNA Pol III expression constructs was comparable to positive controls expressed by pCI-Alu-ulA, 21, 26, 18, and pCI-Alu-ulAΔT, 20, 25, 18. While there were some differences in the total number of edited sites observed, it is possible that the 15% editing threshold used to differentiate G-peaks from background noise in the electropherogram could have led to an underrepresentation of the actual number of editing sites. This may be especially true

when observing editing in the presence of endogenous ADAR activity, where observed editing events were often very near to this minimum threshold value.

This shows that Alu elements expressed by RNA Pol III undergo promiscuous A-to-I RNA editing when an inverted Alu element is present in the transcript. Also, the amount of editing was comparable to that observed in RNA Pol II expressed inverted Alu pairs.

Discussion

The goal of these experiments was to first determine if Alu elements expressed by RNA Pol III could undergo A-to-I RNA editing, and to then determine structural characteristics necessary for editing to occur in these situations. This was done using HeLa cells transfected with vectors that express Alu elements either by RNA Pol II or RNA Pol III.

Although a single Alu element is able to take on a secondary structure containing some double stranded regions, a single Alu element alone is insufficient to enable A-to-I RNA editing. This was observed regardless of the type of polymerase responsible for transcription as well as whether ADAR1 or ADAR2 was overexpressed. In addition, the pBS-U6-Alu-Term-ulA and pBS-Alu-Term-ulAΔT constructs, which also do not express the inverted Alu element, showed no editing. This supports the hypothesis that a single Alu element is insufficient to form an A-to-I RNA editing substrate.

Expressed inverted Alu pairs undergo A-to-I RNA editing at multiple adenosines when expressed by RNA Pol III. This was observed in the presence of overexpressed ADAR1, ADAR2 or endogenous ADAR activity. Promiscuous editing within inverted Alu elements expressed by RNA Pol III was similar to levels observed under RNA Pol II

expression. Both results were consistent with previous research investigating editing within inverted Alu elements embedded within mRNA (Athanasiadis et al. 2004, Peng et al. 2012, Osenberg et al. 2010). These results indicate that it is the structure of the editing substrate rather than the transcriptional mechanism that is important for the development of promiscuous A-to-I RNA editing patterns.

The design of the inverted Alu expression vectors had two purposes. The first was to enable the study of structural elements necessary for A-to-I RNA editing in transcriptionally active Alu elements. The other was to evaluate specific sequences in the inverted Alu element that may potentially behave as RNA Pol III termination sites. RNA Pol III termination sequences are encoded as a minimum of four consecutive thymine nucleotides (Chu et al. 1995, Chu et al. 1997, Gunnery et al. 1999). Alu elements contain two poly-A regions that are encoded as poly-T regions when inverted, potentially behaving as RNA Pol III termination sites (Batzer et al. 1996). Importantly, neither the central poly-A track nor the poly-A tail of the inverted Alu element behaved as a transcription termination site. This can be inferred from the presence of edited nucleotides in the target Alu element expressed by U6-Alu-ulA and pBS-Alu-ulAΔT, in conjunction with the analysis of editing within a single Alu element demonstrating that the presence of the inverted Alu element within the RNA molecule was necessary for editing. This is significant because it indicates a greater likelihood for expression of the entire inverted Alu repeat when the forward Alu element is transcribed by RNA Pol III. This increases the probability of A-to-I RNA editing within transcriptionally active Alu elements expressed *in vivo*.

The observation that Alu elements expressed by RNA Pol III can be highly edited at multiple locations may be significant in relation to a number of mechanisms. These include retrotransposition, the known roles of transcriptionally active Alu RNA, and regulation of promiscuously edited RNA molecules. In a broad sense, Alu retrotransposition and Alu RNA functional activities are dependent on both the Alu sequence and secondary structure (Bennett et al. 2008, Hasler and Strub 2006, Mariner et al. 2008, Rubin et al. 2002, Wagner et al. 2010). A-to-I RNA editing is a mechanism that changes the RNA sequence, and alters the secondary structure present within the RNA molecule. Editing in Alu RNA could potentially behave as a regulatory mechanism impacting the different processes that Alu RNA is involved in. For example, A-to-I RNA editing could affect Alu retrotransposition rates by changing the binding affinity of Alu elements to SRP9/14, an interaction thought to be necessary for Alu retrotransposition (Bennett et al. 2008, Sarrowa et al. 1997). Additionally, A-to-I RNA editing would change the Alu sequence prior to retrotransposition, thereby altering the sequence of the newly transposed Alu element from its parent gene. Such changes could influence other downstream mechanisms such as DNA methylation at CpG dinucleotides present within the Alu element. These relationships between transcriptionally active Alu elements and A-to-I RNA editing will be explored in more detail in the closing chapter of this dissertation.

.

Primer	Primer Sequence
TarAlu-F1	GAC TTCTAAAGGA AGCAGCATGG
TarAlu-F-PstI	ATCTGCAG GGCCGGGTGCAGTGGCTCACTCC
TarAlu-R-HindIII	ATAAGCTT GCACTTTTTGTTTTTTTTGAGA
InvAlu-F1	ATCTGCAG GGCCGGGTGC AGTGGCTCACTCC
InvAlu-F2T	ATCTCGAGGCATCTCGATATCTGAGACATGTTTTTTTTTTCT TAAATGTATCC
InvAlu-F2dT	ATCTCGAGGCATCTCGATATCTGAGACATGGAGACGGAGT CTCGCAGTCACC
InvAlu-R1	GAATTGTATTGAGAATGCACTGG
InvAlu-R-KpnI	ATGGTACCGAATTGTATTGAGAATGCACTGG
SDMF1	CGACGATACTAGCGATTTGACTCGAGGCATCTCG
SDMR1	CGAGATGCCTCGAGTCAAATCGCTAGTATCGTCG
SDMF2	CGACGATACTAGCGAGTTGACTCGAG
SDMR2	CGAGATGCCTCGAGTCAACTCGCTAG
PBSALU-F-Eco	ATGAATTCGGCCGGGTGCAGTGGCTCACTCC
RPBS	TTGTAATACGACTCACTATAGG
93-1D	GCACTTTGGGAGGCTGAGGCAAGTAG
93-1D-Eco	GGAATTCGCACTTTGGGAGGCTGAGGCAAGTAG
PCI-1D	CACAACAGTCTCGAACTTAAGC
T-Site-R-KpnI	CTAGAGGTACCGTATCGTCGTACCTAGTCTGTCG
I-Site-R-KpnI	ATGGTACCCATGTCTCAGATATCGAGATGC

Table 3.1- Table of Primers

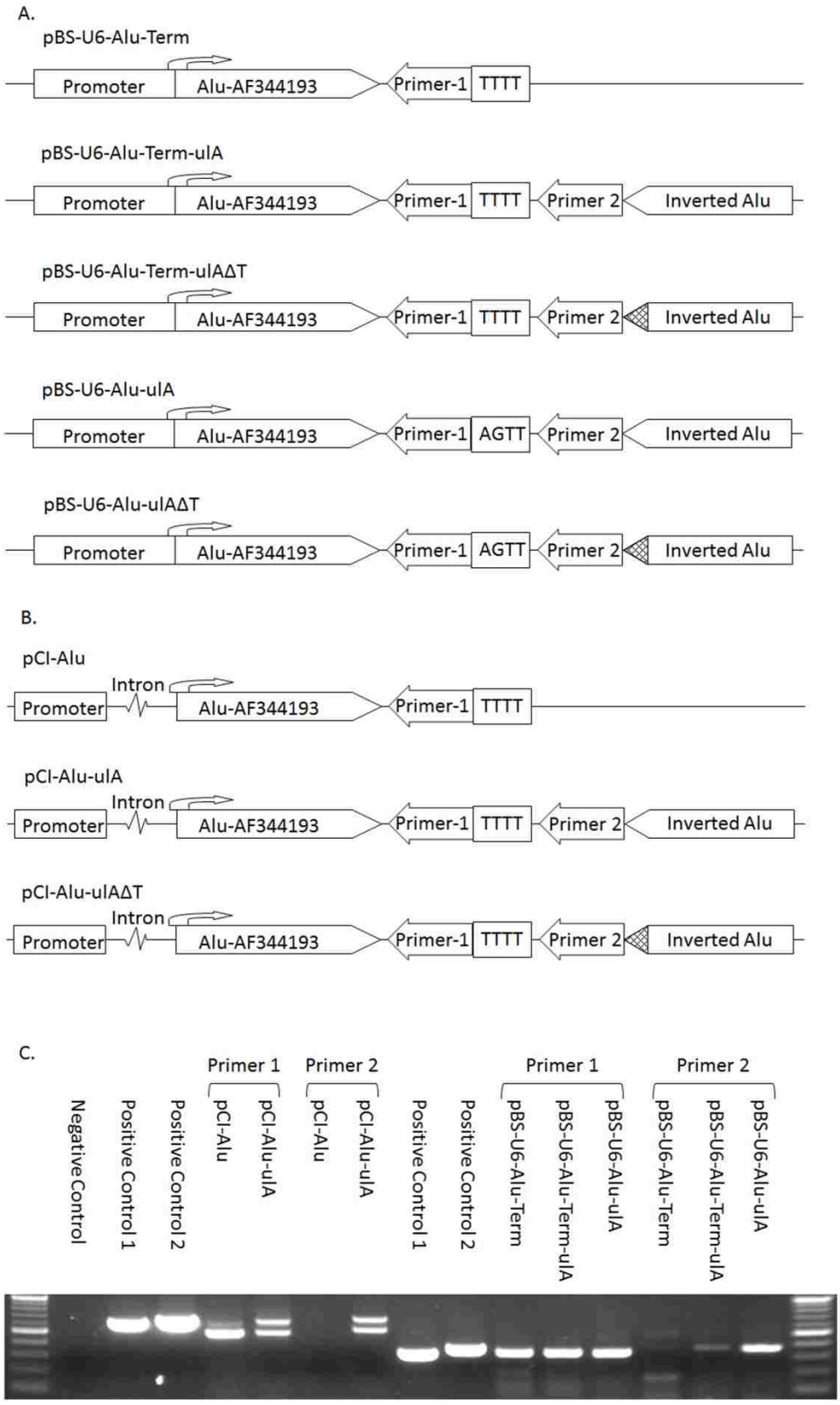

A.
pBS-U6-Alu-Term
Promoter
Alu-AF344193
Primer-1
TTTT
pBS-U6-Alu-Term-ulA
Primer 2
Inverted Alu
pBS-U6-Alu-Term-ulAΔT
pBS-U6-Alu-ulA
AGTT
pBS-U6-Alu-ulAΔT
B.
pCI-Alu
Intron
pCI-Alu-ulA
pCI-Alu-ulAΔT
C.
Primer 1
Primer 2
Negative Control
Positive Control 1
Positive Control 2
pCI-Alu
pCI-Alu-ulA
pBS-U6-Alu-Term
pBS-U6-Alu-Term-ulA
pBS-U6-Alu-ulA

Figure 3.1 – Design and Expression of Alu Expression Vectors. A) RNA Pol III expression vectors were built on a pBS vector backbone and used a U6 promoter to drive expression of the target Alu sequence. The RNA Pol III specific termination site is encoded by four consecutive thymines, and is preceded by the primer site for T-Site-R-Kpn. Inverted Alu pairs have a second primer site for primer I-Site-R-Kpn located between the termination site and the inverted Alu element. B) RNA Pol II expression vectors were built on a pCI vector backbone and used a CMV promoter to drive expression. A short intronic sequence is located upstream of the target Alu. The termination site for RNA Pol III transcription is present but is not recognized by RNA Pol II. Two of the RNA Pol III expression vectors and one of the RNA Pol II expression vectors have had the poly-A tail region removed (ΔT) from the inverted Alu element. This region is encoded as poly-T when inverted and potentially behaves as an RNA Pol III termination site. C) RT-PCR of RNA expressed by the different expression vectors. The presence of the termination site inhibits expression of longer transcripts by RNA Pol III. pCI-Alu-ulA and pBS-U6-Alu-Term-ulA plasmids were used as positive controls.

	ADAR 1	ADAR2	No ADAR
pCI-Alu	No	No	No
pBS-U6-Alu-Term	No	No	No
pCI-Alu-ulA	Edited 21 of 42 Adenosines	Edited 26 of 42 Adenosines	Edited 18 of 42 Adenosines
pCI-ALU-ulAΔA	Edited 20 of 42 Adenosines	Edited 25 of 42 Adenosines	Edited 18 of 42 Adenosines
pBS-U6-Alu-TERM-ulA	No	No	No
pBS-U6-ALU-TERM-ulAΔA	No	No	No
pBS-U6-ALU-ulA	Edited 24 of 42 Adenosines	Edited 23 of 42 Adenosines	Edited 18 of 42 Adenosines
pBS-U6-ALU-ulAΔA	Edited 20 of 42 Adenosines	Edited 23 of 42 Adenosines	Edited 15 of 42 Adenosines

Figure 3.2 – Summary of A-to-I RNA editing observed in Alu RNA. 42 total adenosines were evaluated in each Alu element in the presence of overexpressed ADAR1, ADAR2, or endogenous ADAR activity. Electropherograms were initially inspected for A/G double peaks indicative of A-to-I RNA editing. Alu RNA expressed by pBS-U6-Alu-Term, pCI-Alu, pBS-U6-Alu-Term-ulA, and pBS-U6-Alu-Term-ulAΔT were also inspected on a single transcript basis in 100 subcloned sequences each.

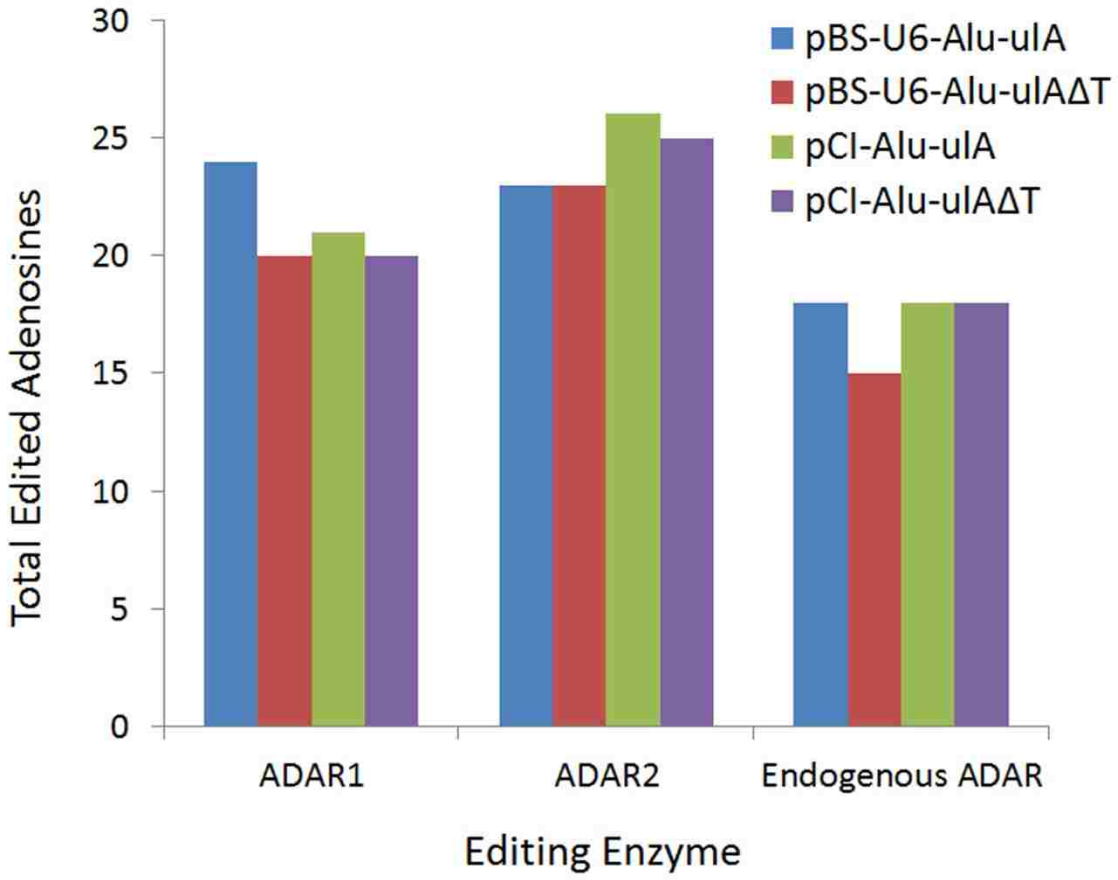


Figure 3.3 – Editing Analysis of Inverted Alu repeats expressed by RNA Pol II and RNA Pol III. 42 adenosine nucleotides in the target sequence were inspected for the presence of A/G double peaks in their electropherogram. A 15% lower threshold was used to differentiate A-to-I RNA editing from background.

Chapter 4

Transcriptionally Active Alu Elements Undergo Basal Level A-to-I RNA Editing *In Vivo*

Introduction

A-to-I RNA editing is a post transcriptional RNA modification that targets dsRNA. The reaction is catalyzed by the ADAR family of proteins, and the resulting inosine is interpreted as guanosine by other cellular machinery. Editing displays both spatial and temporal differences within an organism (Athanasiadis et al. 2004, Gan et al. 2006, Kawahara et al. 2003, Nicholas et al. 2010). In general, ADAR activity is higher in more developmentally complex tissues (Athanasiadis et al. 2004, Gan et al. 2006). In humans, A-to-I RNA editing is most active in the brain. Temporal differences in editing are also present, and can vary on a gene to gene basis within the same tissue (Nicholas et al. 2010). For example, Gabra-3 encodes for the alpha-3 subunit of the GABA receptor in the brain. In human cortex, editing levels in Gabra-3 are unaffected by age. This is in contrast to editing observed in CYFIP2, which is involved in neuronal development and maintenance. CYFIP2 displays an age dependent decrease in A-to-I RNA editing in human cortex, demonstrating a gene specific temporal influence on editing activity.

Alu elements have been shown to be highly targeted for A-to-I RNA editing when embedded within larger mRNA molecules (Athanasiadis et al. 2004, Peng et al. 2012). The high editing levels in Alu elements is due in large part to their high copy number in the human genome (Athanasiadis et al. 2004, IHGHC 2001, Sela et al. 2007). Their commonality often results in two oppositely oriented Alu elements being expressed within a single transcript. This allows dsRNA to form, enabling ADAR binding and subsequent editing.

Alu elements are composed of a left and right arm separated by a central poly-A track and ending in a poly-A tail (Batzer et al. 1996). The presence of A-box and B-box

RNA Pol III promoter elements in the Alu left arm enables Alu transcription by RNA Pol III (Chu et al. 1995, Dieci et al. 2007, Li and Schmid 2001, Roy et al. 2000). These Alu elements are termed transcriptionally active Alu elements or transcribed Alu elements (Li and Schmid 2001). Their expression increases following viral infection, onset of cancer or disease, or following cell exposure to DNA damaging reagents (Li and Schmid 2001, Liu et al. 1995, Panning and Smiley 1995, Rudin and Thompson 2001). Indeed, transcribed Alu elements expressed in response to cell stress associate with both RNA Pol II and the ribosome, thereby influencing expression of target genes in *trans* (Hasler and Strub 2006, Mariner et al. 2008, Rubin et al. 2002, Wagner et al. 2009).

Transcriptionally active Alu elements have also been implicated as the source of newly retrotransposed Alu elements (Bennett et al. 2008, Britten et al. 1988, Dewainnieux et al. 2003). Due to their evolutionary relationship the 7SL RNA from the SRP, they are able to associate with the SRP9/14 SRP subunit (Bennett et al. 2008, Huck et al. 2004,Sarrowa et al. 1997). SRP9/14 helps facilitate Alu retrotransposition by shuttling Alu elements to the ribosome, where they can then complex with ORF2p. ORF2p is a reverse transcriptase encoded by the L1 repetitive element and is responsible for Alu retrotransposition (Feng et al. 1996, Mathias et al. 1991). Though ORF2p typically binds its own mRNA, transcriptionally active Alu elements are able to supplant L1 mRNA via a template switching mechanism, enabling Alu retrotransposition (Esnault et al. 2000, Kroutter et al. 2009, Kulpa and Moran 2005, Wei et al. 2001).

Alu elements are highly involved in epigenetic regulation. Alu elements are enriched for CpG dinucleotides, and as such often constitute CpG islands (Cho et al. 2007, Kang et al. 2006). CpG islands can be targeted for DNA methylation, a catalyzed

base modification that results in reduced expression of nearby genes. In addition, nucleosome mapping showed that nucleosomes preferentially associate with Alu elements, and that this preference is mediated by interactions between the histone core complex and CpG dinucleotides present in Alu elements (Bettecken et al. 2011, Tanaka et al. 2010). Since inosine is interpreted as guanosine, A-to-I RNA editing in retrotransposably active Alu RNA has the potential to generate new CpG dinucleotides prior to Alu retrotransposition.

Here, I investigate transcriptionally active Alu elements for evidence of A-to-I RNA editing in humans *in vivo.* A filter system was employed to screen for transcriptionally active Alu elements that are likely to be editing targets. RNA derived from human brain and spleen from multiple individuals was probed for editing within the selected Alu elements. One brain sample showed evidence of basal level A-to-I RNA editing. Editing at a single site within the Alu element was preferentially edited relative to other adenosines within the Alu element. Also, editing resulted in the generation of CpG dinucleotides in the Alu RNA that were absent from the genomic sequence. A-to-I RNA editing was not observed in this Alu element in other tissue samples, nor was it observed in other Alu elements investigated. I hypothesize that A-to-I RNA editing could have a significant impact on the sequence of Alu elements that are able to retrotranspose.

Methods and Materials

DNA and RNA

DNA and RNA used for *in vivo* analysis was obtained from Biochain Inc. as a matched set product called Dr. P Set. Dr. P Sets derived from three brain tissue donors and one

spleen donor were used for the analysis. RNA was treated with DNaseI to ensure complete removal of genomic DNA.

PCR and RT-PCR

3 ug of RNA was reverse transcribed using Superscript III reverse transcriptase according to manufacturer's protocol. For each selected Alu element PCR was performed using successive PCR reactions. Primary and nested primer sets were designed for Alu-344192, Alu344193 Alu 344196, Alu517a, and Alu 517b (Table 3). Samples were purified by phenol/chloroform reaction followed by ethanol precipitation. PCR amplicons were gel purified on a 4% agarose gel using QIAEX II gel extraction kit, and were sequenced by the Sanger sequencing method.

Subcloning, Blue/White Colony Screen, and Sequence Analysis

PCR amplicons were cut using either EcoRI and KpnI (Alu-93 and Alu -96) or PstI and BamHI (Alu-92, Alu517a, Alu517b) and were ligated into pBS vectors using T4 ligase. Samples were then transformed into bacteria and spread on agar plates containing 100 ng/mL Ampicillin and 4 uL/mL Gal-X solution for blue/white colony screening. White colonies were selected to inoculate LB media containing 100ng/mL Ampicillin and were incubated overnight at 37°C. Plasmids were isolated using Qiagen Mini-prep kits, and were sequenced by Sanger sequencing method using a T7 primer. Sequences were aligned and compared using Seqman.

Statistical Evaluation

The comparison of the number of mismatches by position was performed using a X^2-Test with a p-value of $p = 0.001$. The expected value for this test was determined using the formula described by Luciano et al., 2001, for calculating an number of

mismatches, or errors, to be expected based the experimental procedure. This error rate uses the equation E=A*B + A*C*D, where E is the expected error, A is the number of adenosine nucleotides, B is the error rate of the reverse transcriptase, C is the error rate of Taq Polymerase, and D is the number of PCR cycles (Luciano et al. 2001). For comparison of mismatches by position, X^2-Test was also used with a p-value of p = 0.005. The expected frequency was determined by averaging the total number of A/G mismatches across all adenosines in which A/G mismatches had occurred.

Results

Alu Identification

A literature search identified 87 Alu elements previously shown to be transcriptionally active (Li and Schmid 2001, Gu et al. 2009, Borchert et al. 2006). These Alu elements were inspected using the UCSC genome browser. They were then subjected to a series of manual filters in order to reduce the number of Alu elements investigated to those most likely to undergo A-to-I RNA editing, and to eliminate Alu elements that would be experimentally difficult to study (figure 4.1). The first filter removed Alu elements if they shared a genomic region with a known gene that is in the sense orientation relative to the direction of Alu transcription. This filter ensured that any editing observed would be derived from RNA Pol III transcripts and not from the embedded Alu sequence. A second filter searched sequences 3kb downstream of the Alu element, and removed those that did not contain an inverted Alu element within that region. This filter removed Alu elements that are unable to form the dsRNA secondary structure that is necessary for A-to-I RNA editing to occur. A final filter eliminated any Alu elements that lacked unique DNA sequence between the inverted Alu elements. Due

to the repetitive nature of Alu elements, this filter was necessary to ensure amplification of the specific target Alu element using an RT-PCR protocol.

This filter process narrowed the number of potentially edited Alu elements down to five. Two Alu elements, Alu-517a and Alu-517b, were located upstream of, and are responsible for the expression of, miRNAs 517a and 517b (Borchert et al. 2006). The other 3 Alu elements, with the gene accession numbers AF314492 (Alu-92), AF314493 (Alu-93), and 314496 (Alu-96), were previously identified by cDNA cloning and primer extension analysis (Li and Schmid 2001).

Alu-93 is from the AluY family and has an inverted Alu element from the AluSx family located downstream and separated by 727 nucleotides (figure 4.2A). Alu-96 is from the AluS family and has an inverted Alu element from the AluS family located 1543 nucleotides downstream (figure 4.2B). These Alu elements were amplified by RT-PCR using nested primers designed such that the reverse primer was located downstream of the first canonical RNA Pol III termination site, encoded as four consecutive thymines. In total, 3 brain tissue samples and 1 spleen sample were used as RNA sources for editing analysis. Donor information was unavailable for the first brain sample. All other donors were males ages 34 and 82 for the other two brain tissue samples, and age 74 for the spleen sample. Only Alu-93 and Alu-96 were successfully amplified by RT-PCR. Alu-92, Alu-517a, and Alu517b were not successfully amplified using either standard or modified RT-PCR procedures.

Alu-93 undergoes basal level A-to-I RNA Editing

Amplification of Alu-93 yielded a PCR product 375 nucleotides in length. The genomic sequence of Alu-93 is 179 nucleotides in length between the beginning of the

central-A track and the end of the poly-A tail. 57 total adenosine nucleotides are present, with 14 adenosines located in the central-A track, 30 adenosines in the right-arm excluding the two poly-A regions (RAX-A), and 13 adenosines present in the poly-A tail. Due to primer design and difficulties derived from inefficient sequencing of the central-A track, only the right arm of Alu-93 was inspected for A-to-I RNA editing.

Sanger sequencing of the initial RT-PCR product of Alu-93 did not reveal any evidence of A-to-I RNA editing. Since Sanger sequencing cannot accurately measure editing below 10%, a single transcript analysis protocol was used to gain more detailed information. In total, 196 subcloned sequences from brain donor-1 were analyzed, and total nucleotide mismatches (N/N) between gDNA and cDNA were recorded (figure 4.3). 39 A/G mismatches were observed in total, with 30 A/G mismatches located in the RAX-A region, and 9 A/G mismatches located in the poly-A tail.

Total A/G mismatches in the RAX-A region was significantly higher ($p=0.001$) than the calculated error rate of 3.9 A/N mismatches. This error rate was calculated based on the error rates of the reverse transcriptase and Taq polymerase, as well as the number of PCR cycles used during RT-PCR amplification (Luciano et al. 2001). The total number of A/G mismatches was also significantly higher ($p=0.001$) than those observed in other N/N mismatch combinations as well as the total A/G mismatches ($p=0.001$) observed in a control sample of 50 subcloned sequences derived from amplified genomic DNA (gDNA). The A/G mismatches in this cDNA sample were the only type of mismatch that significantly differed from either the expected error rate or the gDNA control.

Given that these A/G mismatches are significantly higher than the expected error rate as well as the external control, using gDNA, and internal control, comparing other N/N mismatches, it is hypothesized that these A/G mismatches are the result of A-to-I RNA editing. However, since the total number of edited nucleotides is so low relative to the total sample size (less than 1% editing), this will be termed basal level editing.

A-to-I RNA Editing is not Randomly Distributed

A-to-I RNA editing was tabulated by nucleotide position relative to the start of the central-A track in order to determine if any of these nucleotides are preferentially edited (figure 4.4). The first nucleotide in the central-A track was used as the +1 position since the entire left arm of Alu-93 was not amplified. Using this strategy, the RAX-A region begins at position 17 and ends at position 165. The poly-A tail spans positions 166 to 179. Of the 30 adenosines present in RAX-A, 12 adenosines had an A/G mismatch in at least one of the subcloned RT-PCR amplicons. Only one of these nucleotides, position 146, had more than 2 A/G mismatches. This position had 15 mismatches, which was significantly higher ($p=0.05$) than would be expected based on random distribution. Interestingly, this site has an adenosine immediately upstream and 2 nucleotides downstream, both of which are unfavorable for editing. The poly-A tail was edited at 6 out of 13 adenosines. Editing was not enriched in this region relative to RAX-A. In addition, this region did not have any nucleotides that were preferentially edited.

A sequence alignment was performed between the AluY consensus sequence, the Alu-93 genomic sequence, and the cumulatively edited Alu-93 cDNA sequence (figure 4.5). Since inosine is interpreted as guanosine by much of the cellular machinery, edited adenosines were replaced by guanosines in the cDNA sequence in order to gain a better

understanding of the structural ramifications of A-to-I RNA editing in Alu-93. There were two notes of interest in relation to Alu retrotransposition and the ability of CpG dinucleotides present in Alu elements to undergo DNA methylation. First, A-to-I RNA editing generates five new CpG dinucleotides, including one in the poly-A tail, that were not present in the gDNA sequence. 1 of the newly created CpG dinucleotides located in the RAX-A region restores a CpG dinucleotide to the Alu-93 sequence that was lost relative to the AluY consensus sequence. The other three newly formed CpG dinucleotides in the RAX-A region represent novel CpG dinucleotides that are absent from the AluY family consensus sequence. Such editing effects prior to retrotransposition provide a mechanism of maintaining high CpG dinucleotide content in Alu elements as they retrotranspose.

Analysis of A-to-I RNA editing by position indicates that editing is not randomly distributed and that a single nucleotide within Alu-93 is preferentially edited. In addition, A-to-I RNA editing may play a role in maintaining, and possibly even creating new, CpG dinucleotides prior to retrotransposition.

Alu-96 undergoes C-to-U but not A-to-I RNA editing

Analysis of A-to-I RNA editing in additional brain samples and human spleen provided different results. In 2 additional brain tissue samples, A/G mismatches in both Alu-93 and Alu-96 did not display any evidence of A-to-I RNA editing. Sequencing of the matched gDNA from one of the brain samples revealed the presence of two SNPs in the Alu-93 sequence. Analysis of RNA derived from human spleen did not show any evidence of A-to-I RNA editing in either Alu element. However, C-to-U editing was prevalent in Alu-96. Alu elements have previously been shown to be C-to-U editing

targets, and the family of enzymes responsible, APOBEC, has been shown to be active in spleen (Bogerd et al. 2006, Chiu et al. 2006Greeve et al. 1998, Tan et al. 2009). Thus A-to-I RNA editing was limited in the selected Alu elements based on analysis in human brain and spleen.

Discussion

This study investigated A-to-I RNA editing in transcriptionally active Alu elements *in vivo.* Using a filtering process, Alu elements previously described as being transcriptionally active were selected based on their likelihood of being edited. Only one Alu element, Alu-93 had any indication of A-to-I RNA editing. Other Alu elements selected for analysis either could not be amplified (Alu-92, Alu517a, Alu517b) or did not show any evidence of A-to-I RNA editing (Alu-96). Editing in Alu-93 was observed at basal levels in only one of the tested brain samples. Editing in this sample was not randomly distributed, with a significant preference for editing at one specific adenosine. In addition, A-to-I RNA editing has the potential to create new CpG dinucleotides or restore CpG dinucleotides that had been lost relative to that Alu element's family consensus sequence This is especially significant in Alu elements that are able to retrotranspose. Testing of additional brain samples as well as spleen did not reveal any evidence of A-to-I RNA editing in either Alu-93 or Alu-96, however, C-to-U editing was observed in spleen in Alu-96. Brain and spleen were chosen as tissue sources due to their relatively high levels of editing activity (Athanasiadis et al. 2004, Gan et al. 2006).

A-to-I RNA editing of the two analyzed transcriptionally active Alu elements was determined to be at basal levels. Comparison of A/G mismatches to internal controls, an external control, and a calculated error rate indicated that A-to-I RNA editing was

occurring in Alu-93 transcripts in one of the brain tissue samples. This error rate has previously been described to as an estimatation of false positives that may arise during RT-PCR (Luciano et al. 2004). This error rate is also an overestimate of A/G mismatches as this error calculation is an estimate all A/N mismatches, not just A/G mismatches. Accuracy of this error rate is demonstrated by comparison of other types of N/N mismatches as well as by comparison to A/G mismatches that arise in the external control of sequenced gDNA PCR inserts. Neither control demonstrated a significant difference from the estimated error rate. The total amount of editing observed was below 1% for the population of adenosines inspected within Alu-93. This level of editing is consistent with basal level editing.

Such basal level editing may be more prevalent than previously anticipated . A 2011 study published by the Maas and Lopresti groups used a bioinformatics screen to identify A-to-I RNA editing throughout the transcriptome (Maas et al. 2011). High numbers of adenosines edited at low levels were predicted based on the computational alignment between the human genome and EST databases. This was supported by a 2009 study that identified widespread low level editing in the human transcriptome (Li et al. 2009). Gommans et al. 2009, previously hypothesized basal level editing as a selectively advantageous mechanism. They proposed that cells use A-to-I RNA editing at minimal evolutionary cost to randomly probe RNA transcripts at low levels in search of unintentional benefits.

Analysis of A-to-I RNA editing by position indicated a single nucleotide position (position-146) was preferentially edited relative to other adenosines in the Alu RNA sequence. What makes this site additionally more interesting is the presence of adenosine

nucleotides in the -1 and +2 positions relative to the editing site. An adenosine nucleotide at either position is generally considered unfavorable for A-to-I RNA editing (Athanasiadis et al. 2004, Kawahara et al. 2008, Lehman and Bass 2000). In addition, editing analysis of 500 embedded inverted Alu repeats indicated that this site is not preferentially edited (Athanasiadis et al. 2004). The higher editing levels observed at position-146 is suggestive of two different mechanisms. First, editing in transcriptionally active Alu elements, and more specifically Alu-93, may have slightly different site-selectivity when compared to editing in embedded Alu elements. Second, it has been previously proposed that editing begins at a single editing nucleation site (Enstero et al. 2009). Promiscuous editing then proceeds from there in a unidirectional manner. Higher levels of editing at position-146 is suggestive of this nucleotide behaving as a nucleation site that then fails to initiate promiscuous editing at additional nucleotides.

A-to-I RNA editing provides a potentially interesting link between retrotransposition and DNA methylation of Alu elements. DNA methylation targets CpG dinucleotides present within CpG islands. This leads to a conformational change in DNA packaging, resulting in a reduction of gene expression in the surrounding genomic region. Due to their high CpG content, Alu elements often constitute CpG islands (Cho et al. 2007, Kang et al. 2006). This is supported by their increased presence near the promoter regions of protein coding genes (Grover et al. 2003, Polak and Domany 2006, Tsirigos and Rigoutsos 2009). As Alu elements retrotranspose, they are potentially introducing new DNA methylation sites into the genome. This greatly expands their regulatory impact on gene expression.

A-to-I RNA editing may target Alu transcripts prior to retrotransposition (figure 4.6). Since inosine is interpreted as guanosine, editing prior to retrotransposition would generate an A-to-G change from the parent Alu sequence to the newly retrotransposed Alu element. This would act to preserve high levels of CpG dinucleotides that may have been lost, and may additionally introduce new CpG dinucleotides that are not present in the Alu consensus sequence. This was supported by the observation of editing introducing 5 new CpG dinucleotides in the cumulative construction of editing sites in Alu-93 that were not present in the Alu-93 gDNA sequence (figure 4.5). The prevalence of such a process may be high since it has been previously demonstrated that cytosine is strongly favored at the -1 nucleotide position relative to A-to-I RNA editing sites (Athanasiadis et al. 2004, Kawahara et al. 2008, Lehman and Bass 2000). Such a process would greatly increase the impact A-to-I RNA editing has had on the evolution of the human genome.

Additionally, A-to-I RNA editing may inhibit Alu retrotransposition (figure 4.6). ADAR activity has previously been shown to inhibit miRNA processing by binding to and sequestering pri- and pre-miRNA away from Drosha and Dicer, respectively (Kawahara et al. 2005, Kawahara et al. 2008, Yang et al. 2006). Editing of miRNA may influence the secondary structure, and thus binding kinetics to Drosha and Dicer. ADAR activity may regulate Alu retrotransposition in a similar fashion, first by sequestering Alu elements away from L1 reverse transcriptase, then by changing the RNA sequence and secondary structure through A-to-I RNA editing. A key region necessary for retrotransposition is the poly-A tail (Comeaux et al. 2009, Dewainnieux and Heidmann 2005). This region is involved in the template switching mechanism and in TPRT and as

such requires a certain minimal length and homogeneity to enable efficient retrotransposition. A-to-I RNA editing provides a mechanism of disrupting the continuity of the poly-A tail, and thus may negatively impact Alu retrotransposition.

The appearance of C-to-U editing in the spleen sample was also not surprising. C-to-U editing is catalyzed by the APOBEC family of enzymes (Kim et al. 2010, Rausch et al. 2009, Tan et al. 2010). These enzymes are single stranded RNA binding proteins that can associate with transcripts derived from repetitive elements and lead to a decrease in retrotransposition activity (Bogerd et al. 2006, Chiu et al. 2006, Kim et al. 2010, Rausch et al. 2009, Tan et al. 2010). This effect has previously been demonstrated on Alu retrotransposition. APOBEC represent a mechanism of regulating transcriptionally active Alu elements.

A-to-I RNA editing itself may provide a mechanism of regulating other functions of transcriptionally active Alu elements. Transcribed Alu elements have previously been shown to regulate both transcription and translation in *trans* (Hasler and Strub 2006, Mariner et al. 2008, Rubin et al. 2002, Wagner et al. 2009). These mechanisms are mediated through specific regions of the Alu element that maintain certain secondary structure characteristics. A-to-I RNA editing could regulate these processes by altering the secondary structure of these specific regions, thereby impacting Alu directed regulation of transcription and translation.

The differences in editing levels between the different individuals could be due to a number of factors. The manner of death, the time between death and tissue sampling, the region of the brain that was sampled, as well as tissue storage conditions prior to RNA extraction were all unknown. The health and dietary habits of each donor prior to

death were also unknown. Region specific differences in editing may be the most significant of these factors influencing A-to-I RNA editing. Though the highest levels of A-to-I RNA editing is observed in the brain, editing levels differ between different sub-regions (Athanasiadis et al. 2004, Gan et al. 2006, Kawahara et al. 2003). For example, editing in the cerebellum is significantly higher than the rest of the brain (Athanasiadis et al. 2004). A-to-I RNA editing has also been shown to be lower in white matter than in grey matter due to decreased ADAR2 expression in white matter (Kawahara et al. 2003). In addition, other environmental factors may play a role. A 2012 study of A-to-I RNA editing in different species of octopus showed a near linear increase in A-to-I RNA editing as ocean temperature decreased (Garrett and Rosenthal 2012). Other environmental factors besides temperature may also influence A-to-I RNA editing.

Known information about the tissue donors was limited to age and gender. A-to-I RNA editing was observed only in Alu-93, albeit at a low level, and from only one donor. Information about this donor was unavailable. All known donors were male and ranged in age from 34 to 82 years old. Since editing has previously been shown to vary with respect to age in a target specific manner, it may be that, for unknown regulatory reasons, these Alu elements are no longer favorable editing targets in the male brain after a certain age threshold (Nicholas et al. 2009). In addition, recent results from the 1000 Genome Project show that A-to-I RNA editing activity can vary significantly between different individuals (Li et al. 2011).

Finally, of the over 1 million Alu elements present in the human genome, only two transcriptionally active Alu elements were sequenced. Though low levels of A-to-I RNA editing were observed in one instance, this may not be the case with other Alu

elements, and should not be extrapolated as a description of A-to-I RNA editing in all transcriptionally active Alu elements. Only with the sampling of a wider number of transcriptionally active Alu elements could more definitive conclusions be drawn.

This investigation of A-to-I RNA editing in transcriptionally active Alu elements expressed *in vivo* showed A-to-I RNA editing in one Alu element. Though editing was determined to be at a basal level, analysis of editing in this particular Alu element revealed evidence of preferential editing at a specific adenosine. In addition, editing demonstrated the ability to generate new CpG dinucleotides, should this Alu transcript retrotranspose.

Primer	Primer Sequence
92-1D	GAGGATCCCCAGCACTTTGGGAGGCCAAGGCG
92-2U	CTTCTTCTGGACTGTGAACCAG
92-3U	GTAACTGCAGCAATGATTCTCTTCTCCTTGCCAG
93-1D	ATGAATTCGCACTTTGGGAGGCTGAGGCAAGTAG
93-2U	GCTGAGACCTGTTCTATTCCACTAAC
93-3U	ATGGTACCCTCTGTGTCCCCAAGTTGTATAATTC
96-1D	ATGAATTCGCACTTTGGGAGACCAAAGCAGGAAG
96-2U	CATAGTGCAACCTGTGACATAGTCC
96-3U-	ATGGTACCCACCCTACCACTACTTCTAGGATTTC
Alu517a-1D	GAGGATCCCCAGCTACTCAGGAGGCTGACG
Alu517a-2U	GTAACTGCAGATCTTGGTGGAAGGTGCTATCC
Alu517b-1D	GAGGATCCGTGGATCACGAGGTCAGGAGATCG
Alu517b-2U	GTAACTGCAGCTTGAGCCCAAGAGATCTCGGTTGG

Table 3 – Table of Primers

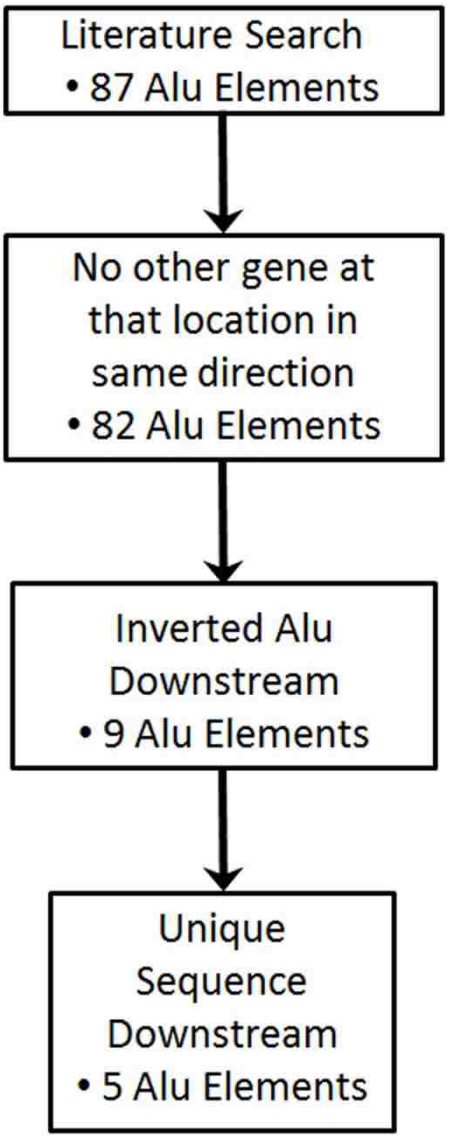


Figure 4.1 – Flow chart of the Alu Selection Screen. Following the initial identification of transcriptionally active Alu elements by literature search, Alu elements were manually examined using the UCSC Genome Browser. The filters were designed to only leave Alu elements that could be specifically targeted by RT-PCR, and that had a relatively high chance of being A-to-I RNA editing targets.

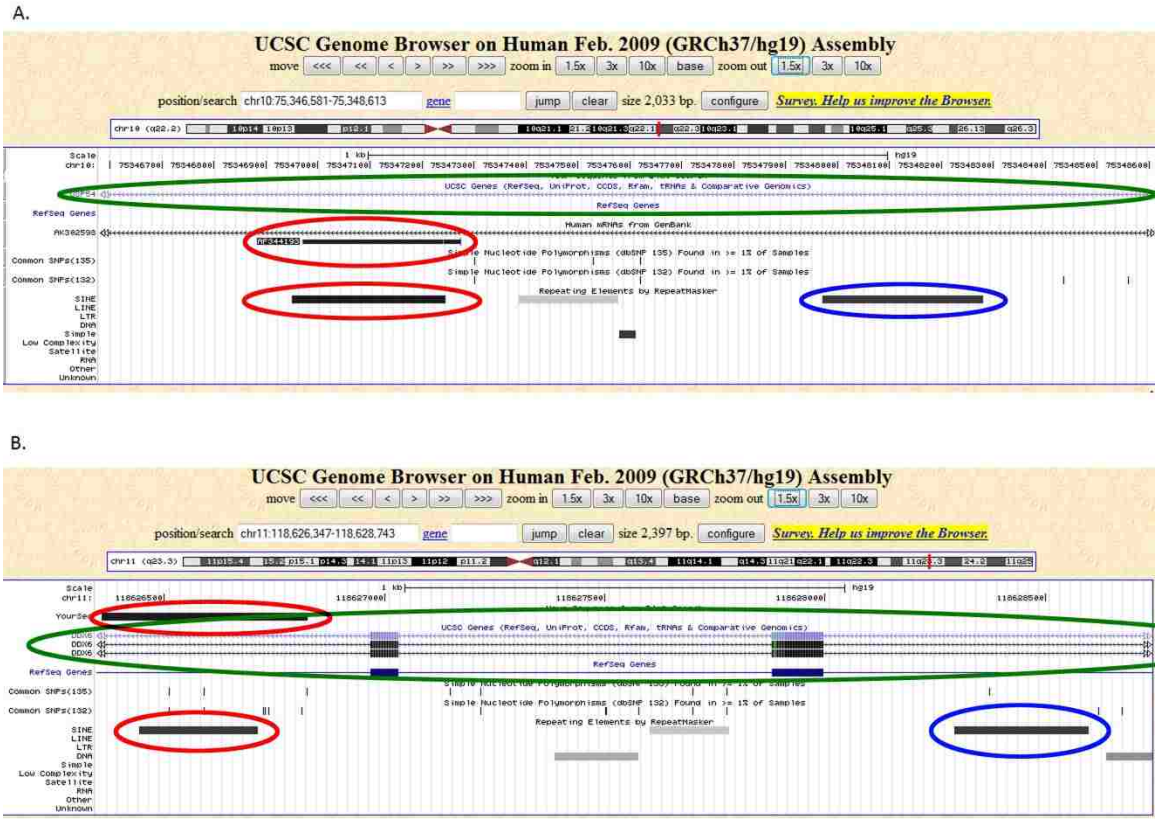


Figure 4.2 – Genomic Structure Surrounding Two Transcriptionally Active Alu Elements. The UCSC Genome browser displays A) Alu-93 and B) Alu-96 (red) along with the nearest inverted Alu element (blue) located downstream. Both Alu elements co-localize with a gene (green) that is expressed in the antisense direction relative to Alu expression.

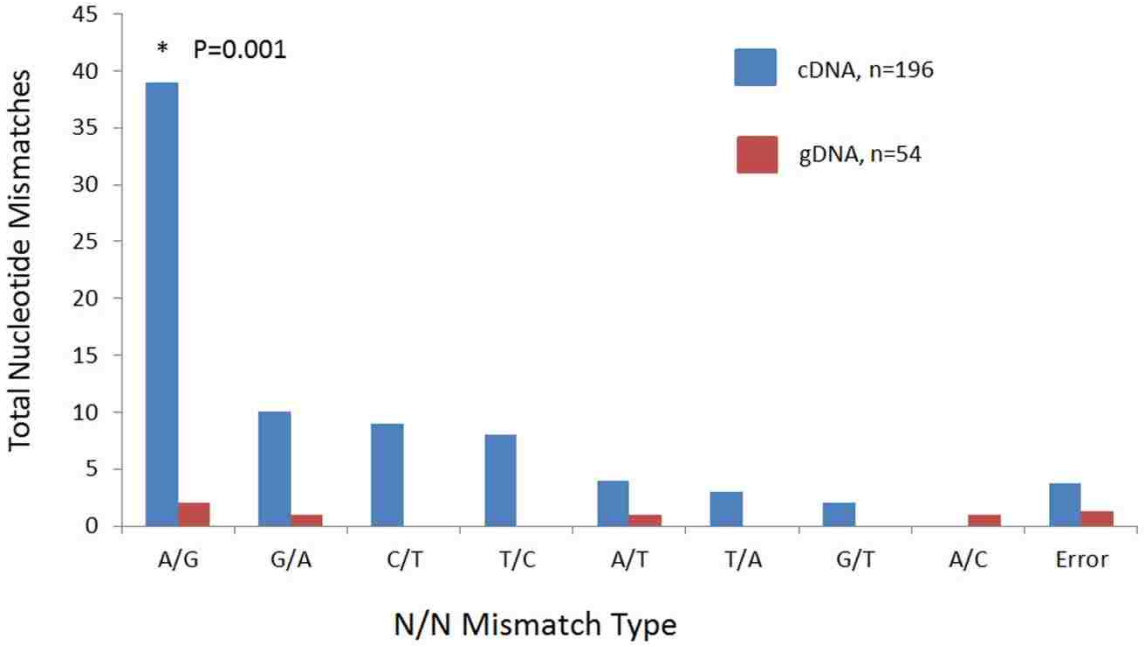


Figure 4.3 - Basal level A-to-I RNA editing is observed in Alu-93. RT-PCR amplicons from cDNA and PCR amplicons from gDNA were subcloned into pBS vector and sequenced using the Sanger sequencing method. All observed N/N mismatches between the genomic sequence and subcloned sequences were recorded and compared to an estimated number of A/N errors that are expected to occur during RT-PCR.

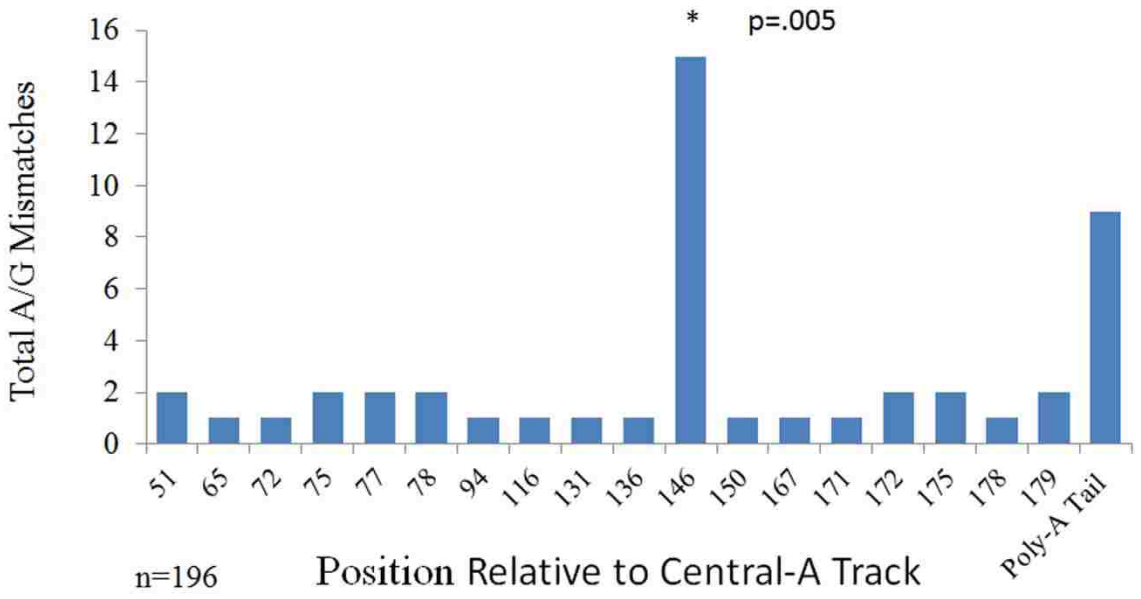


Figure 4.4 – A-to-I RNA editing analysis of Alu-93 by nucleotide position. Cumulative editing was plotted by nucleotide position. Only A-to-I RNA editing is displayed. Counting began at the first nucleotide of the central-A track. The central-A track spans nucleotides 1-16. RAX-A spans positions 17-165. The poly-A tail spans nucleotides 166-179.

```
CLUSTAL 2.0.8 multiple sequence alignment

AluY_consensus   AAAAA-TACAAAAAA-TTAGCCGGGGCGTGGTGGCGGGCGCCTGTAGTCCCAGCTACTCGG 58/178
Alu93_genomic    AAAAAACACAAAAAAATTAGCCGGGTGTGGTGGCGAGCACCTGTGGTCCCAGCTACTCGG 60/178
Alu93_cDNA       AAAAAACACAAAAAAATTAGCCGGGTGTGGTGGCGAGCACCTGTGGTCCCGGCTACTCGG 60/178
                 *****  ******** ********* ********* ** ***** ***** *********

AluY_consensus   GAGGCTAGAGGCAGGAGAATGGCGTGAACCCGGGAGGCGGAGCTTGCAGTGAGCCGAGAT 118/238
Alu93_genomic    GAGGAT-GAGGCAGGAGAATGGCGTGAACCTGGGAGGCGGAGCTTGCAGTGAGCCGAGAC 119/237
Alu93_cDNA       GAGGGT-GAGGCGGGGGGGTGGCGTGAACCTGGGGGCGGAGCTTGCAGTGAGCCGGGAC 119/237
                 **** * ***** ** *  ********** *** ********************* **

AluY_consensus   CGCGCCACTGCACTCCAGCAGTCCGGCCTGGGCGACAGAGCGAGACTCTGTCTCA----- 173/293
Alu93_genomic    TGCGCCACTGCACTCCAGC--------CTGGGCAACAGAGCGAGACTCCATCTCAAAAA 171/289
Alu93_cDNA       TGCGCCACTGCGCTCCGGC--------CTGGGCAGCAGGGCGAGACTCCATCTCAGAAAG 171/289
                 ********** **** **        ******  *** *** *****  ***** 

AluY_consensus   --------
Alu93_genomic    AACAAAAA 179/297
Alu93_cDNA       GACGAAAG 179/297
```

Figure 4.5 – Sequence Alignment of Alu-93 cDNA, gDNA, and AluY family consensus sequence. The fully edited cDNA sequence for Alu-93 containing all edited nucleotides is aligned to the Alu-93 genomic sequence and the AluY family consensus sequence. Edited nucleotides are underlined and boxed in red in order to highlight their location. CpG dinucleotides are highlighted within the family consensus sequence. Only the central-A track, Alu right arm, and poly-A tail are shown. Nucleotide positions at the end of the column are relative to the start of the central-A track (first number) and the start of the Alu element (second number).

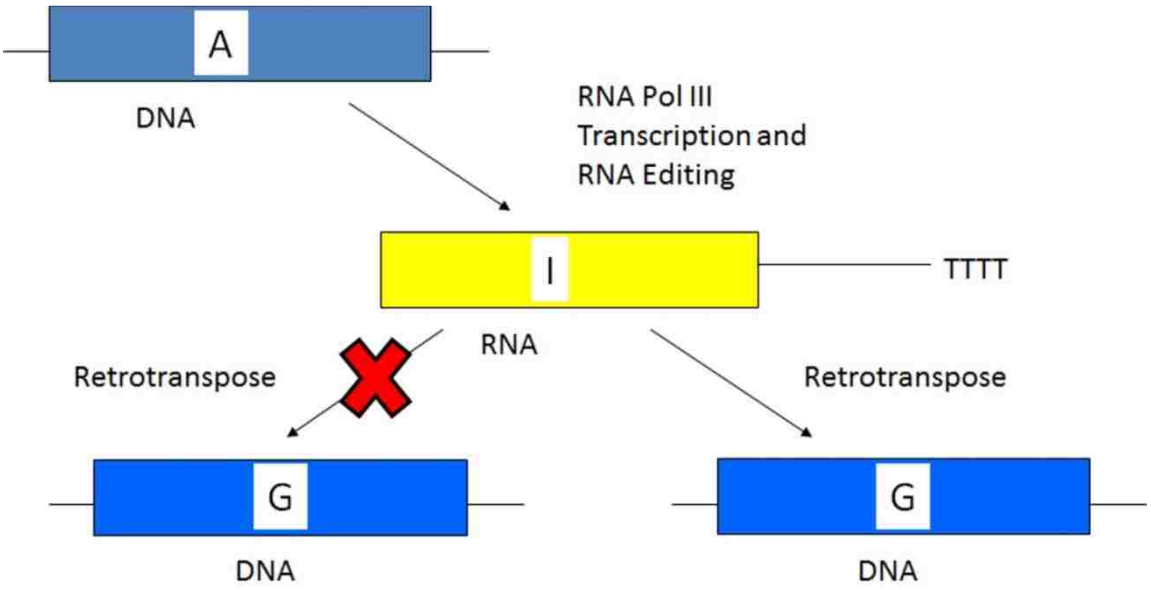


Figure 4.6 – A model of A-to-I RNA editing during Alu retrotransposition. A model of how A-to-I RNA editing in transcriptionally active Alu elements may impact retrotransposition. ADAR activity may alter the Alu sequence leading to introduction of guanosine in newly retrotransposed Alu elements. Alternatively, ADAR activity may influence the retrotransposition process.

Chapter 5

Closing Remarks And Future Directions

Alu elements are major A-to-I RNA editing targets within the human genome. However, analysis of editing within Alu elements has been limited to those present as inverted Alu repeats embedded within larger genes expressed by RNA Pol II. The goal of this dissertation was to expand our knowledge of known editing targets to include Alu elements expressed by RNA Pol III. The implications for editing in these targets include the impact of A-to-I RNA editing on human evolution, on Alu function as a lncRNA, and on regulation of A-to-I RNA editing.

The experimental strategy to studying A-to-I RNA editing in RNA Pol III transcripts had three key steps that each provided different information. The first step was the development of expression constructs that expressed a known non-Alu A-to-I RNA editing target under either an RNA Pol II or RNA Pol III promoter. This demonstrated that RNA Pol III transcripts undergo A-to-I RNA editing by ADARs. Comparison of A-to-I RNA editing between these constructs as well as in the presence of either overexpressed ADAR1 or ADAR2 showed that RNA Pol II expressed targets were more highly edited relative to RNA Pol III expressed editing targets. The R/G editing target was also more highly edited in the presence of ADAR2 compared to ADAR1 under both RNA Pol II and RNA Pol III expression. Finally, a partial deletion of RNA Pol II CTD increased A-to-I RNA editing by ADAR1, though this increase was marginal. Editing by ADAR2 was unaffected by the partial CTD deletion, but remained high (above 90% editing). These experiments indicate that ADARs are not dependent on an interaction with RNA Pol II to function and that a significant amount of editing can occur post-transcriptionally rather than co-transcriptionally.

The second step was to use the RNA Pol II and RNA Pol III expression constructs to better understand A-to-I RNA editing of Alu elements. Alu elements were promiscuously edited, but only if an inverted Alu element was present in the transcript. This was observed independent of the kind of polymerase responsible for transcription as well as overexpression of ADAR1 or ADAR2. Inverted Alu pairs expressed by RNA Pol III were promiscuously edited similar to editing observed in embedded Alu elements (Athanasiadis et al. 2004, Peng et al. 2012).

The final goal was to observe A-to-I RNA editing in transcriptionally active Alu elements expressed *in vivo*. This was done by first identifying individual Alu elements previously shown to be transcriptionally active. Next, using a filter process, these Alu elements were screened for the likelihood of forming dsRNA, and for feasibility of specific amplification. The selected Alu elements were then individually examined by RT-PCR and single transcript analysis using RNA derived from either human brain or spleen from different individuals. Only one brain tissue sample displayed evidence of A-to-I RNA editing. Though editing was occurred at a basal level, one nucleotide was preferentially edited.

These experiments provide interesting insights into different facets of A-to-I RNA editing. The experiments using the R/G site from GluR-B demonstrated that RNA Pol III transcripts can undergo ADAR directed A-to-I RNA editing. This was supported by the inverted Alu elements expressed by RNA Pol III being edited as well. While both targets were edited, these two experiments provide an interesting comparison since the R/G site provides a study of editing efficiency while Alu elements can be used to understand editing site selectivity.

Editing site selectivity is the targeting of ADARs to particle adenosines but not others throughout a dsRNA molecule, while editing efficiency is the amount of editing that takes place at an individual adenosine. In a broad sense, editing efficiency and editing site selectivity are separate events and are inversely correlated (Gommans et al. 2008, Kallman et al. 2003). The most efficiently edited RNA substrates, such as the Q/R site from GluR-B tend to be edited at a lower number of sites, while promiscuously edited sequences tend to display lower editing efficiency on a site-by-site basis. For example, the Q/R site from GluR-B is edited at only one adenosine, but is edited to almost 100% in human brain tissue, while a survey of embedded Alu elements showed that Alu elements can be edited at multiple sites and that editing efficiency ranged between less that 1% to greater than 70% at individual adenosines (Athanasiadis et al. 2004, Higuchi et al. 1993). These generalities were observed in both RNA Pol II and RNA Pol III transcripts, however, RNA Pol II transcripts were more efficiently edited.

Overexpression of ADAR1 in the R/G studies indicated that ADAR1 does not target the R/G site when it is expressed by RNA Pol III. This was not the case with inverted Alu repeats, in which promiscuous editing was observed regardless of the type of ADAR overexpressed. This indicates that the editing efficiency in RNA Pol III transcripts by ADAR1 is dependent on the ability of ADAR1 to target the editing substrate. It has previously been shown that RNA Pol II CTD increases editing efficiency by ADAR2 when the editing substrate contains a splice site (Bratt and Ohman 2003, Laurnecikiene et al. 2006, Ryman et al. 2007). It is possible that RNA Pol II helps to promote ADAR1 editing efficiency at the R/G site as well. Though partial deletion of the RNA Pol II CTD resulted in an increase in ADAR1 directed RNA editing, full CTD

deletion may be necessary to see larger changes in RNA editing efficiency. This increase was significant at p=0.05 but not p=0.025, indicating this change in editing may be marginal at best, or perhaps not significant with larger sample sizes. While the interaction between the CTD and ADAR1 has not previously been published, similar studies have been performed with ADAR2 (Laurnecikiene et al. 2006, Ryman et al. 2007). Both my results and previous work demonstrate that partial CTD deletion does not influence ADAR2 activity when a splice is absent from the double stranded region.

The full role of RNA Pol II in directing A-to-I RNA editing is difficult to define. Previous CTD deletion studies have demonstrated a role by the RNA Pol II CTD in coordinating editing by ADAR2 with splicing (Bratt and Ohman 2003, Laurnecikiene et al. 2006, Ryman et al. 2007). These studies have indicated differing effects by the CTD in regulating editing. This may depend on the location of the exon/intron junction relative to the dsRNA region. Further increasing the difficulty in understanding this coordinating role is the lack of data from ADAR and RNA Pol II co-IP studies that would demonstrate a direct interaction between the two (Nishikura 2010). This indicates that ADAR2 may be transiently associated with RNA Pol II during transcription rather than directly associated. This is in contrast to other co-transcriptional modifications, such as RNA splicing, in which partial deletion of heptads 26-52 impaired splicing efficiency, or RNA capping, in which capping enzymes have been shown to associate with heptads 26-52 of the RNA Pol II CTD (Fong and Bentley 2001).

Previous research has indicated that editing site selectivity and editing efficiency are influenced by a combination of the ability of ADAR to bind to its substrate and the localized sequence and structure surrounding the editing site (Kallman et al. 2003,

Ohman et al. 2000, Stefl et al. 2010, Xu et al. 2006). My results indicate that the type of polymerase responsible for transcription may also influence editing efficiency but not site selectivity. Editing at the R/G site was higher in RNA Pol II transcripts relative to RNA Pol III transcripts when ADAR1 or ADAR2 was overexpressed. This indicates either a coordinating role by RNA Pol II or a better ability by the ADAR enzymes to recognize RNA Pol II transcripts based on the presence of modifications, such as a 5' cap or poly-A tail, that are specific to RNA Pol II transcripts. The coordinative role of RNA Pol II is supported by the observation that ADAR1 overexpression increased editing in RNA Pol II transcripts but not RNA Pol III transcripts. However, promiscuous editing was observed in the Alu repeats, regardless of the manner of transcription or overexpression of ADAR1 or ADAR2. This indicates that editing efficiency may be partially dependent on the type of polymerase, factors involved in transcription, or presence of other RNA modifications, while editing site selectivity may be inherently regulated by the ADAR enzymes themselves, as well as the substrate sequence and secondary structure. An interesting experiment to test this hypothesis would be an *in vitro* titration experiment in which the RNA editing targets differ by the presence or absence of a 5' cap and poly-A tail, structures that are prevalent in RNA Pol II transcripts (Anderson 2005, Kapp and Loerrsh 2004).

Differences in editing between Alu elements expressed in cell culture versus transcriptionally active Alu elements studied *in vivo* provides another interesting comparison. While a single Alu element was not edited in cell culture, inverted Alu pairs were promiscuously edited when expressed by either RNA Pol II or RNA Pol III. However, transcriptionally active Alu elements were not promiscuously edited and

displayed basal level editing *in vivo*. This disparity is likely due to the manner of expression. The experiments performed in cell culture used an RNA Pol III promoter to drive expression of the inverted Alu pairs. The *in vivo* studies did not investigate either the level of transcription, or how often the nearest downstream inverted Alu element was present within the transcript. Inclusion levels were just one of several possible factors potentially impacting A-to-I RNA editing. The RT-PCR experiments performed here do not provide information about the quantity of RNA that is present. Quantitative RT-PCR and northern blot analysis could be used to address this issue.

Another potential factor is the selective degradation of highly edited transcriptionally active Alu elements *in vivo*. Tudor SN is a nuclease that targets inosine containing RNA molecules for degradation (Scadden 2005, Scadden 2007). RNA molecules containing higher amounts of inosine are more highly targeted. In addition, piRNA are short RNA molecules that selective target repetitive elements via a mechanism involving the RISC complex. piRNA may also be involved in reducing the quantity of Alu RNA. It is possible that low expression levels of the inverted Alu pair combined with selective degradation of promiscuously edited Alu elements combined to result in the observation of basal level editing in transcriptionally active Alu elements *in vivo*.

In addition, there were other factors in the *in vivo* study that could have contributed to the low editing levels that were observed. The tissue donors were all male, but were of different ages. There was also no information regarding the manner of death, the amount of time between death and tissue sampling or tissue sampling and RNA extraction, or the region of the brain sampled from each individual. In addition, this

study was limited by the starting pool of transcriptionally active Alu elements. Of the more than 1 million Alu elements present in the genome, only 87 were included in the Alu screen, and only 2 Alu elements reached the point of transcript analysis. So while basal level editing was observed in one instance, this should not be extrapolated to describe editing in all transcriptionally active Alu elements.

Implications of A-to-I RNA editing on Alu Retrotransposition and Human Genome Evolution

Transcriptionally active Alu elements are responsible for the expansion of Alu elements throughout the human genome (Bennett et al. 2008, Dewannieux et al. 2003). This has enabled Alu elements to play a significant role in shaping the human genome (IHGSC 2001). A-to-I RNA editing could impact the rate of Alu retrotransposition and can change the sequence that retrotransposes. This implicates A-to-I RNA editing as having a significant role in the evolution of the human genome.

Alu elements retrotranspose via an L1 directed retrotransposition process that uses Alu RNA as a template (Dewannieux et al. 2003). This process could be influenced by A-to-I RNA editing in a number of ways. The influence A-to-I RNA editing has on miRNA processing serves as a model of how editing could impact Alu retrotransposition. miRNA processing is inhibited by ADAR in both editing dependent and editing independent mechanisms (Kawahara et al. 2007, Kawahara et al. 2008, Yang et al. 2006). Editing dependent mechanisms rely on the effect editing has on the secondary structure of pri- and pre-miRNA. A-to-I RNA editing can destabilize these secondary by changing the base pairing within the dsRNA region, which affects miRNA processing. Editing

independent mechanisms rely on ADAR competing with Drosha and Dicer for substrates, reducing the amount of pri- and pre-miRNA available for processing.

Editing dependent influences on Alu retrotransposition rates would potentially influence the interaction between edited Alu elements and SRP9/14. Though the specific role of SRP9/14 during Alu retrotransposition is unclear, it is thought to be involved in helping to recruit Alu RNA to the ribosome as L1 Orf2 is translated. (Bennett et al. 2008). This enables Alu elements to recruit newly translated L1 Orf2 proteins for Alu retrotransposition. The Alu-SRP9/14 interaction is dependent on conserved secondary structure in the Alu left and right arms (Bennett et al. 2008, Hasler and Strub 2006, Huck et al. 2004, Sarrowa et al. 1997). Changes in the Alu left and right arm can alter the base pairing properties of Alu RNA, leading to impaired binding by SRP9/14 and a reduction in Alu retrotransposition (Bennett et al. 2008). Just as A-to-I RNA editing can influence miRNA binding by Drosha and Dicer, editing could also influence the Alu-SRP9/14 interaction by affecting the conserved Alu structures involved in the interaction with SRP9/14.

An editing independent mechanism in which ADAR binding could influence retrotransposition may be similar to the impact the APOBEC3 family of cytidine deaminases have on Alu retrotransposition. APOBEC3 can bind and sequester Alu RNA away from the retrotransposition machinery, thereby inhibiting Alu retrotransposition (Bogerd et al. 2006, Chiu et al. 2006). ADAR binding may have a similar sequestration impact. There are several mechanisms that act to repress Alu retrotransposition (Bogerd et al. 2006, Chiu et al. 2006, Kochanek et al. 1995, Liu et al. 1994, Lukic and Chen 2011, Muiznieks and Doerfler 1994). These include transcriptional regulation, piRNA directed

Alu RNA degradation, and sequestration by APOBEC3. ADAR binding to Alu transcripts may be an additional regulatory mechanism that aids in repressing Alu retrotransposition.

Even if A-to-I RNA editing does not impact Alu retrotransposition rates, editing may provide a mechanism of altering the Alu sequence that retrotransposes. For example, Alu elements have high C/G content and are enriched for CpG dinucleotides (Cho et al. 2007, Kang et al. 2006). Since inosine is interpreted as guanosine, A-to-I RNA editing provides a mechanism of generating new CpG dinucleotides by editing adenosines prior to retrotransposition. Since DNA methylation of CpG dinucleotides is an important mechanism of epigenetic regulation, the enriched presence of CpG dinucleotides in Alu elements allows for altered regulation of expression in the new genomic regions into which Alu elements retrotranspose (Baillie et al. 2011, Hellmann-Blumberg et al. 1993).

A-to-I RNA editing may also impact the identification of source genes, or master genes, responsible for Alu retrotransposition. Alu retrotransposition in humans is currently only seen in the AluY family (Britten et al. 1988, Deininger and Slagel 1998). This is supported by Alu insertion polymorphisms present among different individuals (Styles and Brookfield 2007, Cordaux et al. 2006, Liu et al. 2009, Xing et al. 2007). These Alu polymorphisms have intact RNA Pol III promoter elements in the left arm and long poly-A tails, two structures necessary for retrotransposition (Bennett et al. 2008, Dewannieux and Heidmann 2005). A majority of Alu elements in the human genome lack conservation of either one or both of these structures, indicating that they are unable to retrotranspose (Batzer et al. 1996, Bennett et al. 2008, Comeaux et al. 2009). This has

led to the hypothesis that Alu retrotransposons are derived from a relatively small group of master genes (Britten et al. 1988, Britten et al. 1989). The sequence variation among different Alu polymorphisms has led to the hypothesis that there are multiple master genes (Deininger and Slagel 1998). A-to-I RNA editing may affect the search for these master genes since editing could potentially alter the Alu sequence prior to retrotransposition. This increases the difficulty involved with identifying the specific master gene, since any guanosine present in the newly inserted Alu element could potentially be due to A-to-I RNA editing.

A-to-I RNA Editing and Alu Derived lncRNA

Another potential impact A-to-I RNA editing could have on transcriptionally active Alu elements is by regulating the functions of Alu derived lncRNA. lncRNA are functional non-protein coding RNA molecules longer than 200 nucleotides (Cabili et al. 2011, Loewer 2010). In the case of Alu RNA, this includes interactions with RNA Pol II following cell stress to selectively decrease transcription, association with the SRP9/14 and the ribosome to modulate translation activity, and regulation of gene expression through interactions between lncRNA and complementary RNA or DNA sequences (Hasler and Strub 2006, Mariner et al. 2008, Pandey et al. 2011, Rubin et al. 2002, Wagner et al. 2010). All of these functions share the characteristic of Alu elements behaving as global regulators of gene expression in *trans*.

Expression of transcriptionally active Alu elements increases following cell stress (Li and Schmid 2001, Liu et al. 1995, Panning and Smiley 1995, Rudin and Thompson 2001). Alu RNAs can then interact with RNA Pol II to decrease transcription of non-cell stress response genes (Mariner et al. 2008). The RNA Pol II transcription factor TFIIF is

involved in preventing the Alu-RNA Pol II interaction at promoter regions of genes whose expression levels do not decrease following cell stress (Wagner et al. 2009). A-to-I RNA editing could impact this regulatory process in two ways. First, A-to-I RNA editing may influence the Alu/RNA Pol II interaction by changing the Alu sequence and potentially the Alu secondary structure. Specifically, a conserved repression domain in the Alu right arm may be altered, affecting this repression functionality (Mariner et al. 2008,Wagner et al. 2009). Second, structural changes caused by A-to-I RNA editing could alter the ability of TFIIF to disrupt the Alu/RNA Pol II interaction. This could be by a mechanism by which editing alters the Alu-RNA Pol II interaction, thereby changing the influence TFIIF has on the Alu/RNA Pol II association.

A-to-I RNA editing may have a similar influence on the ability of Alu RNA to interact with both SRP9/14 and the ribosome to regulate translation in *trans*. Alu RNA by itself can interact with the ribosome, leading to an increase in reporter gene expression (Hasler and Strub 2006, Rubin et al. 2002). However, Alu RNA can also interact with SRP9/14 (Bovia et al. 1997, Hsu et al. 1995). This complex can then further interact with the ribosome leading to a decrease in translation initiation (Hasler and Strub 2006). While the Alu/SRP9/14 interaction can occur between either the Alu left or right arms, the Alu-ribosome interaction is dependent on the Alu right arm only. A-to-I RNA editing may participate in this regulatory process through both editing dependent and editing independent mechanisms.

In the editing dependent mechanism, A-to-I RNA editing would change the Alu sequence, thereby disrupting base pairing and dsRNA structures within Alu RNA. This could reduce the ability of the Alu right arm to interact with the ribosome. This would

decrease the activity of Alu RNA to enhance translation. While A-to-I RNA editing could also disrupt the formation of the Alu/SRP9/14 complex, this is less likely since each arm of the Alu RNA can participate in this interaction.

An editing independent mechanism would likely have a net inhibitory effect on translation. ADARs would compete with SRP9/14 and the ribosome for binding to Alu RNA. This would reduce the number of Alu/SRP9/14 and Alu/ribosome interactions by reducing the availability of Alu RNA. However, since Alu/SRP9/14 complex formation is achieved through binding to either of two regions in the Alu element versus one for the Alu/ribosome interaction, ADAR binding to Alu RNA may more significantly impact the Alu/ribosome interaction (Hasler and Strub 2006). This would result in a net decrease in translational activity when ADAR concentrations in the cell are high.

Finally, the ability of Alu RNA to regulate expression of complementary sequences could also be influenced by A-to-I RNA editing. Alu RNA expressed in an antisense direction relative to another gene enables binding between complementary sequences and can result in a decrease in the transcript level of the target gene (Gong and Maquat 2011, Pandey et al. 2011). Similar to the influence editing has on miRNA function, editing of transcriptionally active Alu elements could reduce complementarity to one gene while increasing complementarity to another, thereby altering which genes are regulated by the edited Alu RNA.

Future Directions

This dissertation has laid a foundation for future study by demonstrating that A-to-I RNA editing can target RNA Pol III transcripts in general, and more specifically transcriptionally active Alu elements. To further understand the significance of A-to-I

RNA editing in transcriptionally active Alu elements, I am proposing two general experimental ideas. The first is the use of a cell culture system to study cellular mechanisms regulated by A-to-I RNA editing of transcriptionally active Alu elements in order to understand how A-to-I RNA editing influences Alu retrotransposition. The second is the use of high throughput sequencing selectively targeting transcriptionally active Alu RNA. This would allow for analysis of A-to-I editing within transcriptionally active Alu elements on a large scale, allowing for a broader understanding of where and when A-to-I RNA editing targets these Alu transcripts.

The cell culture system would use stably transfected cells overexpressing either ADAR1 or ADAR2 in conjunction with both wild-type cells and the use of RNAi to knockdown ADAR activity. This would allow for an increased ability to study how ADAR impacts transcriptionally active Alu elements. This would be used in conjunction with RNAi to investigate factors, such as Tudor SN, that may influence the concentration of edited transcriptionally active Alu elements within the cell. In addition, these strategies could be used to gain a better understanding of the impact A-to-I RNA editing has in regulating the function of Alu derived lncRNA. These include Alu RNA interactions with RNA Pol II, SRP9/14, and the ribosome as well as the role of Alu derived lncRNA in STAU-1 mediated decay of complementary RNA (Gong and Maquat 2011, Hasler and Strub 2006, Mariner et al. 2008, Wagner et al. 2009). This cell culture system would take advantage of an array of cell types. This is necessary in order to highlight differences that may occur on a cell type by cell type basis as well as to understand the breadth of cell types in which these processes are impacted.

Another important process that could be studied using these cells lines is Alu retrotransposition. The marked Alu expression vector developed by Thierry Heidmann's lab could be used to gain a better understanding of how A-to-I RNA editing influences Alu retrotransposition rates as well as the Alu sequence that retrotransposes (Comeaux et al. 2009, Dewannieux et al. 2003). This marked Alu expression vector would be modified to include an inverted Alu element located downstream of the marked Alu. Differences in the number of drug resistant colonies would be used to measure the influence A-to-I RNA editing has on the rate of retrotransposition. This would be compared to controls that either lack the inverted Alu element, or to cells in which ADAR function has been knocked down. Furthermore, the level of A-to-I RNA editing could be compared between marked Alu RNA and newly retrotransposed Alu elements to determine the relative rate at which the edited Alu sequence retrotransposes. This would be extremely beneficial in gaining an understanding of the total impact A-to-I RNA editing has on Alu retrotransposition.

The second future direction involves the use of high throughput sequencing to understand A-to-I RNA editing in transcriptionally active Alu elements *in vivo*. Development of high throughput sequencing that specifically targets Alu RNA depends on two things (Li and Schmidt 2001). First, due to the high sequence conservation present in the over 1 million Alu elements in the human genome there is a great difficulty involved in properly annotating individual Alu transcripts to their source gene. This problem is compounded because A-to-I RNA editing may additionally alter the Alu RNA sequence. There needs to be a mechanism in place that can specifically align individual Alu transcripts to their source gene in the human genome. Second, because Alu elements

can be expressed either by RNA Pol II or RNA Pol III, there needs to be a way to differentiate these two species of Alu RNA. Addressing these issues requires a combination of database analysis and selection of the proper high throughput sequencing method.

Transcriptionally active Alu elements are expressed by RNA Pol III using internal promoters that recruit RNA Pol III to the transcription start site (Ishiguro et al. 2002, Kenneth et al. 2008). Transcription continues until it reaches a transcription termination site encoded as four consecutive thymine nucleotides (Chu et al. 1995, Chu et al. 1997, Gunnery et al. 1999). The location of these termination sequences relative to the end of the Alu element varies on an individual Alu gene basis (Borchert et al. 2006, Gu et al. 2009, Li and Schmid 2001). The intervening region between the end of the Alu element and the transcription termination site is often a sequence that is unique to each individual Alu locus. These unique regions could be used to address the issue of aligning Alu elements to their source gene.

The second issue of differentiating Alu transcripts based on their method of transcription would take advantage of differences in transcription start sites (Dieci et al. 2007, Li and Schmid 2001). Alu elements expressed by RNA Pol II are embedded within larger RNA molecules with transcription start sites located upstream from the 5'end of the Alu sequence. RNA Pol III transcripts have start sites at the 5'end of the Alu element. The differences between these two transcription start sites could be used to distinguish RNA Pol II from RNA Pol III transcripts.

The final step necessary for sequencing would be to choose a high throughput strategy that provides sequence reads long enough to cover the span between the 5'

transcription start site and the 3' unique sequence. High throughput sequencing uses parallel sequencing of DNA that has been cut into uniformly sized fragments (Hall 2007, Schuster 2008). Due to differences in sequencing strategies, the length of these DNA fragments (read lengths) varies depending on the strategy. Some high throughput sequencing methods use strategies employing read lengths greater than 300 nucleotides. These methods include Roche's 454 pyrosequencing (300-500 nt read lengths), DNA nanoball sequencing (400-500 nt read lengths) by Complete Genomics, and Pacific Biosciences' single molecule real time sequencing (SMRT) (1500 nt read lengths) (Foquet et al. 2008, Margulies et al. 2005, Porreca 2010). Sequence reads would be aligned to the genome to identify nucleotide mismatches between gDNA and cDNA that arise due to A-to-I RNA editing. In addition, the number of sequence reads would be used to quantify expression on an individual Alu transcript basis.

This high throughput method could be used to study transcriptionally active Alu elements from different tissues as well as in cell culture. Two interesting cell populations to study would be embryonic stem cells and germ cells. This is due to the potential impact A-to-I RNA editing could have on the sequence of Alu elements able to retrotranspose. In order for newly retrotransposed Alu elements to be inherited by an individual's offspring, retrotransposition would need to occur in either embryonic stem cells that eventually differentiate into germ cells, or in germ cells themselves which eventually mature into either egg or sperm cells. Identification of A-to-I RNA editing in transcriptionally active Alu elements in either of these cell populations would raise the possibility that A-to-I RNA editing is impacting the Alu sequence and thus human genome evolution.

References

References

1. Agranat, L., Raitskin, O., Sperling, J., Sperling, R. 2008. The editing enzyme ADAR1 and the mRNA surveillance protein hUpf1 interact in the cell nucleus. PNAS, 105(13): 5028-5033.

2. Aleman, C., Roy-Engel, A.M., Shaikh, T.H., Deininger, P.L. 2000. Cis-acting influences on Alu RNA levels. Nucleic Acids Research, 28(23): 4755-4761.

3. Alisch, R.S., Garcia-Perez, J.L., Moutri, A.R., Gage, F.H., Moran, J.V. 2006. Unconventional translation of mammalian LINE-1 retrotransposons. Genes & Development, 20: 210-224.

4. Amit, M., Sela, N., Keren, H., Melamed, Z., Muler, I., Shomron, N., Izraeli, S., Ast, G. 2007. Biased exonization of transposed elements in duplicated genes: a lesson from the TIF-IA gene. BMC Molecular Biology, 8: 109.

5. Anderson, J.T. 2005. RNA Turnover: Unexpected Consequences of Being Tailed". Current Biology 15 (16): R635–8

6. Andrews, P.G., Kole, R. 1987. Alu RNA transcribed in vitro binds the 68-kDa subunit of the signal recognition particle. The Journal of Biological Chemistry, 262 (6): 2908-2912.

7. Apoil, P.A., Kuhlein, E., Robert, A., Rubie, H., Blancher, A. 2007. HIGM syndrome caused by insertion of an AluYb8 element in exon 1 of the CD40LG gene. Immunogenetics, 59: 17-23.

8. Athanasiadis, A., Rich, A., Maas, S. 2004. Widespread A-to-I RNA editing of Alu-containg mRNAs in the human transcriptome. PLOS Biology, 2: 2144-2157.

9. Baillie, J.K., Barnett, M.W., Upton, K.R., Gerhardt, D.J., Richmond, T.A., De Sapio, F., Brennan, P., Rizzu, P., Smith, S., Fell, M., Talbot, R.T., Gustincich, S., Freeman, T.G., Mattick, J.S., Hume, D.A., Heutink, P., Carninci, P., Jeddeloh, J.A., Faulkner, G.J. 2011. Somatic retrotransposition alters the genetic landscape of the human brain. Nature, 24: 534-537.

10. Batzer, M.A., Deininger, P.L., Hellmann-Blumberg, U., Jurka, J., Labuda, D., Rubin, C.M., Schmid, C.W., Zietkiewiez, E., Zuckerkandl, E. 1996. Standardized nomenclature for Alu repeats. J. Mol. Evol, 42: 3-6.

11. Baumann, M., Pontiller, J., Ernst, W. 2010. Structure and basal transcription complex of RNA polymerase II core promoters in the mammalian genome: an overview. Mol Biotechnol., 45(3): 241-247.

12. Bennett, E.A., Keller, H., Mills, R.E., Schmidt, S., Moran, J.V., Weichenreider, O., Devine, S.E. 2008. Active Alu retrotransposons in the human genome. Genome Research, 18: 1875-1883.

13. Berg, K.A., Cropper, J.D., Niswender, C.M., Sanders-Bush, E., Emeson, R.B., Clarke, W.P. 2001. RNA-editing of the 5-HT2C receptor alters agonist-receptor-effector coupling specificity. British Journal of Pharmacology, 134: 386-392.

14. Bernstein, E., Caudy, A.A., Hammond, S.M., Hannon, G.J. 2001. Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature, 409: 363-366.

15. Bettecken, T., Frenkel, Z.M., Trifonov, E.N. 2011. Human nucleosomes: special role of CG dinucleotides and Alu-nucleosomes. BMC Genomics, 12: 273.

16. Blow, M.J., Grocock, R.J., van Dongen, S., Enright, A.J., Dicks, E., Futreal, P.A., Wooster, R., Stratton, M.R. 2006. RNA editing of human microRNAs. Genome Biology, 7: R27.

17. Bogerd, H.P., Wiegand, H.L., Hulme, A.E., Garcia-Perez, J.L., O'Shea, K.S., Moran, J.V., Cullen, B.R. 2006. Cellular inhibitors of long interspersed element 1 and Alu retrotransposition. PNAS, 103 (23): 8780-8795.

18. Bolotin, E., Chellappa, K., Hwang-Verslues, W., Schnabl, J.M., Yang, C., Sladek, F.M. 2011. Nuclear receptor HNF4α binding sequences are widespread in Alu repeats. BMC Genomics, 12: 560.

19. Borchert, G.M., Gilmore, B.L., Spengler, R.M., Xing, Y., Lanier, W., Bhattacharya, D., Davidson, B.L. 2009. Adenosine deamination in human transcripts generates novel microRNA binding sites. Human Molecular Genetics, 18 (24): 4801-4807.

20. Borchert, G.M., Lanier, W., Davidson, B.L. 2006. RNA polymerase III transcribes human microRNAs. Nature Structural & Molecular Biology, 1-5.

21. Bovia F., Fornallaz, M., Leffers, H., Strub, K. 1995. The SRP9/14 subunit of the signal recognition particle (SRP) is present in more than 20-fold excess over SRP in primate cells and exists primarily free but also in complex with small cytoplasmic Alu RNAs. Molecular Biology of the Cell, 6: 471-484.

22. Bovia, F., Wolff, N., Ryser, S., Strub, K. 1997. The SRP9/14 subunit of the human signal recognition particle binds to a variety of Alu-like RNAs and with higher affinity than its mouse homolog. Nucleic Acids Research, 25 (2): 318-325.

23. Bratt, E., Ohman, M. 2003. Coordination of editing and splicing of glutamate receptor pre-miRNA. RNA, 9: 309-318.

24. Brennecke, J., Stark, A., Russell, R.B., Cohen, S.M. 2005. Principles of microRNA-target recognition. PLoS Biology, 3(3): 0404-0418.

25. Britten, R.J. 2010. Transposable elements have strongly affected human evolution. PNAS, 107 (46): 19945-19948.

26. Britten, R.J., Baron, W.F., Stout, D.B., Davidson, E.H. 1988. Sources and evolution of human Alu repeated sequences. Proc. Natl. Acad. Sci. USA, 85: 4770-4774.

27. Britten R.J., Stout, D.B., Davidson, E.H. 1989. The current source of human Alu retroposons is a conserved gene shared with old world monkeys. Proc. Natl. Acad. Sci. USA, 86: 3718-3722.

28. Brooks, R., Eckmann, C.R., Janstch, M.F. 1998. The double-stranded RNA-binding domains of xenopus laevis ADAR1 exhibit different RNA-binding behaviors. FEBS Lett. 434: 121-126.

29. Brouha, B., Meischl, C., Ostertag, E., de Boer, M., Zhang, Y., Neijens, H., Roos, D., Kazazian Jr., H.H. 2002. Evidence consistent with human L1 retrotransposition in maternal meiosis I. American Journal of Human Genetics, 71: 327-336.

30. Buratowski, S. 2009. Progression through the RNA polymerase II CTD cycle. Molecular Cell, 36: 541-546.

31. Burns, C.M., Chu, H., Rueter, S.M., Hutchinson, L.K., Canton, H., Sanders-Bush, E., Emeson, R.B. 1997. Regulation of serotonin-2C receptor G-protein coupling by RNA editing. Nature, 387: 303-308.

32. Cabili, M.N., Trapnell, C., Goff, L, Koziol, M., Tazon-Vega, B., Regev, A., Rinn, J.L. 2011. Integrative annotation of human large intergenic noncoding RNAs reveal global properties and specific subclasses. Genes & Development, 25: 1915-1927.

33. Callahan, K.E., Hickman, A.B., Jones, C.E., Ghirlando, R., Furan, A.V. 2011. Polymerization and nucleic acid-binding properties of human L1 ORF1 protein. Nucleic Acids Research, 40 (2): 813-827.

34. Casarin, A., Martella, M., Polli, R., Leonardi, E., Anesi, L., Murgia, A. 2006. Molecular characterization of large deletions in the von Hippel-Lindau (VHL) gene by quantitative real-time PCR: the hypothesis of an alu-mediated mechanism underlying VHL gene rearrangements. Mol. Diagn. Ther., 10 (4): 243-249.

35. Chen, C., Ara, T., Gautheret, D. 2008. Using Alu elements as polyadenylation sites: A case of retroposon exaptation. Mol Biol. Evol., 26 (2): 327-334.

36. Chen, C.X., Cho, D.S., Wang, Q., Lai, F., Carter, K.C., Nishikura, K. 2000. A third member of the RNA-specific adenosine deaminase gene family, ADAR3, contains both single- an double-stranded RNA binding domains. RNA, 6: 755-767.

37. Chen, L.L., Carmichael, G.G. 2009. Altered Nuclear Retention of mRNAs containing inverted repeats in human embryonic stem cells: functional role of a nuclear noncoding RNA. Molecular Cell, 35: 467-478.

38. Chen, L.L., DeCerbo, J.N., Carmichael, G.G. 2008. Alu element-mediated gene silencing. The EMBO Journal, 1-12.

39. Chimpanzee Sequencing and Analysis Consortium. 2005. Initial sequence of the chimpanzee genome and comparison with the human genome. Nature, 437 (7055): 69-87.

40. Chiu, Y.L., Witkowska, H.E., Hall, S.C., Santiago, M., Soros, V.B., Esnault, C., Heidmann, T., Greene, W.C. 2006. High-molecular-mass APOBEC3G complexes restrict Alu retrotransposition. PNAS, 103 (42): 15588-15593.

41. Cho, N., Kim, B-H., Choi, M., Yoo, E., Moon, K., Cho, Y-M., Kim, D., Kang, GH. 2007. Hypermethlation of CpG island loci and hypomethylation of LINE-1 and Ale repeats in prostate adenocarcinoma and their relationship to clinicopathological features. Journal of Pathology, 211: 269-277.

42. Chu, W.M., Ballard, R.E., Schmid, C.W. 1997. Palindromic sequences preceding the terminator increase polymerase III template activity. Nucleic Acids Research, 25 (11): 2077-2082.

43. Chu, W.M., Liu, W.M., Schmid, C.W. 1995. RNA polymerase III promoter and terminator elements affect Alu RNA expression. Nucleic Acids Research, 23(10): 1750-1757.

44. Clemson, C.M., Hutchinson, J.N., Sara, S.A., Ensminger, A.W., Fox, A.H., Chess, A., Lawrence, J.B. 2009. An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. Molecular Cell, 33: 1-10.

45. Comeaux, M.S., Roy-Engel, A.M., Hedges, D.J., Deininger, P.L. 2009. Diverse cis factors controlling Alu retrotransposition: What causes Alu elements to die? Genome Research, 19 (4): 545-555.

46. Cordaux, R., Hedges, D.J., Herke, S.W., Batzer, M.A. 2006. Estimating the retrotransposition rate of human Alu elements. Gene, 373: 134-137.

47. Cordaux, R., Lee, J., Dinoso, L., Batzer, M.A. 2006. Recently integrated Alu retrotransposons are essentially neutral residents of the human genome. Gene, 373: 138-144.

48. da Silva, U.B.O., Morabito, M.V., Canal, C.E., Airey, D.C., Emeson, R.B., Sanders-Bush, E. 2010. Impact of RNA editing on functions of the serotonin 2C receptor in vivo. Frontiers in neuroscience, 4: A26.

49. Dawson, T.R., Sansam, C.L., Emeson, R.B. 2004. Structure and sequence determinants required for the RNA editing of ADAR2 substrates. The Journal of Biological Chemistry, 279 (6): 4941-4951.

50. de Andrade, A., Wang, M., Bonaldo, M.F., Xie, H., Soares, M.B. 2011. Genetic and epigenetic variations contributed by Alu retrotransposition. BMC Genomics, 12: 617.

51. de Koning, A.P.J., Gu, W., Castoe, T.A., Batzer, M.A., Pollock, D.D. 2011. Repetitive elements may comprise over two-thirds of the human genome. PLoS Genetics, 12: e1002384.

52. de la Matta, M., Kornblihtt, A.R. 2006. RNA polymerase II C-terminal domain mediates regulation of alternative splicing by SRp20. Nature Structural & Molecular Biology, 13 (11): 973-980.

53. Deininger, P.L. and Slagel, V.K. 1988. Recently amplified Alu family members share a common parental Alu sequence. Molecular and Cellular Biology, 8: 4566-4569.

54. Desterro, J.M.P., Keegan, L.P., Jaffray, E., Hay, R.T., O'Connell, M.A., Carmo-Fonseca, M. 2005. SUMO-1 modification alters ADAR1 editing activity. Molecular Biology of the Cell, 16: 5115-5126

55. Dewannieux, M., Esnault, C., Heidmann, T. 2003. Line-mediated retrotransposition of marked Alu sequences. Nature Genetics, 35: 41-48.

56. Dewannieux, M. and Heidmann, T. 2005. Role of poly(A) tail length in Alu retrotransposition. Genomics, 86:378-381.

57. Dieci G, Fiorino G, Castelnuovo M, Teichmann M, Pagano A. 2007. The expanding RNA polymerase III transcriptome. Trends Genet. **23** (12): 614–22.

58. Doench, J.G. and Sharp, P.A. 2004. Specificity of microRNA target selection in translational repression. Genes & Development, 504-511.

59. Doria, M., Neri, F., Gallo, A., Farace, M.G., Michienzi, A. 2009. Editing of HIV-1 RNA by the double-stranded RNA deaminase ADAR1 stimulates viral infection. Nucleic Acids Research, 37 (17): 5848-5858.

60. Doria, M., Tomaselli, S., Neri, F., Ciafre, S.A., Farace, M.G., Michienzi, A., Gallo, A. 2011. ADAR2 editing enzyme is a novel human immunodeficiency virus-1 proviral factor. Journal of General Virology. 92: 1228-1232.

61. Doucet, A.J., Hulme, A.E., Sahinovic, E., Kulpa, D.A., Moldovan, J.B., Kopera, H.C., Athanikar, J.N., Hasnoui, M., Bucheton, A., Moran, J.V., Gilber, N. 2010. Characterization of LINE-1 ribonucleoprotein particles. PLoS Genetics, 6 (10): e1001150.

62. Eller, C.D., Regelson, M., Merriman, B., Nelson, S., Horvath, S., Marahrens, Y. 2009. Repetitive sequence environment distinguishes housekeeping genes. Gene, 390 (2): 153-165.

63. Enstero, M., Daniel, C., Wahlstedt, H., Major, F., Ohman, M. 2009. Recognition and coupling of A-to-I editing sites are determined by the tertiary structure of the RNA. Nucleic Acids Research, 37 (20): 6916-6926.

64. Esnault, C., Maestre, J., Heidmann, T. 2000. Human LINE retrotransposons generate processed pseudogenes. Nature Genetics, 24: 363.

65. Feng, Q., Moran, J.V., Kazazian, H.H., Boeke, J.D. 1996. Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. Cell, 87: 905-916.

66. Feng, Y., Sansam, C.L., Singh, M., Emeson, R.B. 2006. Altered RNA editing in mice lacking ADAR2 autoregulation. Molecular and Cellular Biology, 26 (2): 480-488.

67. Fong, N., Bentley, D.L. 2001. Capping, splicing, and 3' processing are independently stimulated by RNA polymerase II: different functions for different segments of the CTD. Genes and Development, 15: 1783-1795.

68. M. Foquet, M., Samiee, K.T., Kong, X., Chauduri, B.P., Lundquist, P.M., Turner, S.W., Freudenthal, J., Roitman, D.B. 2008. Improved fabrication of zero-mode waveguides for single-molecule detection. Journal of Applied Physics., 103: 034301-1-034301-9.

69. Gao, Y., Baccarelli, A., Shu, X.O., Ji, B.T., Yu, K., Tarantini, L., Yang, G., Li, H.L., Hou, L., Rothman, N., Zheng, W., Gao, Y.T., Chow, W.H. 2012. Blood leukocyte Alu and LINE-1 methylation and gastric cancer risk in the Shanghai Women's Health Study. Br J Cancer, 106 (3): 585-591.

70. Gan, Z., Zhao, L., Yang, L., Huang, P., Zhao, F., Li, W., Liu, Y. 2006. RNA editing by ADAR2 is metabolically regulated in pancreatic islets and β-cells. Journal of Biological Chemistry, 281(44): 33386-33394.

71. Ganguly, A., Dunbar, T., Chen, P., Godmilow, L., Ganguly, T. 2003. Exon skipping caused by an intronic insertion of a young Alu Yb9 element leads to severe hemophilia A. Human Genetics, 113 (4): 348-352.

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

73. Garcia-Obregon, S., Alfonso-Sanchez, M.A., Perez-Miranda, A.M., de Pancordo, M.M., Pena, J.A. 2007. Polymorphic Alu insertions and the genetic structure of Iberian Basques. J Hum Genet., 52 (4): 317-327.

74. Garrett, S., Rosenthal, J.J.C. 2012. RNA editing underlies temperature adaptation in K^+ channels from Polar Octopuses. Science, 335: 848-851.

75. Gasior, S.L., Preston, G., Hedges, D.J., Gilber, N., Moran, J.V., Deininger, P.L. 2006. Characterization of pre-insertion loci of de novo L1 insertions. Gene, 390 (1-2): 190-198.

76. George, C.X., Samuel, C.E. 2011. Host response to polyomavirus infection is modulated by RNA adenosine deaminase ADAR1 but not by ADAR2. Journal of Virology, 85 (16): 8338.

77. George, C.X., Samuel, C.E. 1999. Characterization of the 5'-flanking region of the human RNA-specific adenosine deaminase ADAR1 gene and identification of an interferon-inducible ADAR1 promoter. Gene, 229 (1-2): 203-213.

78. George, C.X., Li.Z., Okonski, K.M., Toth, A.M., Wang, Y., Samuel, C.E. 2009. Tipping the balance: Antagonism of PKR kinase and ADAR1 deaminase functions by virus gene products. J. Interferon Cytokine Res., 29 (9): 477-87.

79. Georgiou, I., Noutsopoulos, D., Dimitriadou, E., Markopoulos, G., Apergi, A., Lazaros, L., Vaxevanoglou, T., Pantos, K., Syrou, M., Tzavaras, T. 2009. Retrotransposon RNA expression and evidence for retrotransposition events in human oocytes. Human Molecular Genetics, 18 (7): 1221-1228.

80. Gommans, W.M., Dupuis, D.E., McCane, J.E., Tatalias, N.E., Maas, S. Diversifying exon code through A-to-I RNA editing, in: H.C. Smith (Ed), RNA and DNA Editing, Wiley & Sons, Inc., New York, 2008.

81. Gommans, W.M., Maas, S. 2008. Characterization of ADAR1-mediated modulation of gene expression. Biochemical and Biophysical Research Communications, 377: 170-175.

82. Gommans, W.M., Mullen, S.P., Maas, S. 2009. RNA editing: a driving force for adaptive evolution? BioEssays 9999: 1-9.

83. Gommans, W.M., Tatalias, N.E., Sie, C.P., Dupuis, D., Vendetti, N., Smith, L., Kaushal, R., Maas, S. 2008. Screening of human SNP database identifies recoding sites of A-to-I RNA editing. RNA, 14: 1-12.

84. Gong, C., Maquat, L.E. 2011. lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3' UTRs via Alu elements. Nature, 470: 284-288.

85. Goodier, J.L., Mandal.P.K., Zhang, L., Kazazian Jr., H.H. Discrete subcellular partitioning of human retrotransposon RNAs despite a common mechanism of genome insertion. Human Molecular Genetics, 19 (9): 2010.

86. Goodier, J.L., Maraia, R.J. 1998. Terminator-specific recycling of a B1-Alu transcription complex by RNA polymerase III is mediated by the RNA terminus-binding protein La. The Journal of Biological Chemistry, 273 (40): 26110-26116.

87. Goodier, J.L., Zhang, L., Vetter, M.R., Kazazian Jr., H.H. 2007. LINE-1 ORF1 protein localizes to stress granules with other RNA-binding proteins, including components of RNA interference RNA-Induced silencing complex. Molecular and Cellular Biology, 27 (18): 6469-6483.

88. Greally, J.M. 2002. Short interspersed transposable elements (SINEs) are excluded from imprinted regions in the human genome. PNAS, 99 (1): 327-332.

89. Greeve, J., Axelos, D., Welker, S., Schipper, M., Greten, H. 1998. Distinct promoters induce APOBEC-1 expression in rat liver and intestine. Arteroscler Thromb Vasc Biol., 18 (7): 1079-1092.

90. Grover, D., Majumder, P.P., Rao, C.B., Brahmachari, S.K., Mukerji, M. 2003. Nonrandom distribution of alu elements in genes of various functional categories: insight from analysis of human chromosomes 21 and 22. Mol. Biol. Evol., 20 (9): 1420-1424.

91. Gu, Y., Kodama, H., Watanabe, S., Kikuchi, N., Ishitsuka, I., Ozawa, H., Fujisawa, C., Shiga, K. 2007. The first reported case of Menkes disease caused by an Alu insertion mutation. Brain Development, 29 (2): 105-108.

92. Gu, T.J., Yi, X., Zhao, X.W., Zhao, Y., Yin, J.Q. 2009. Alu-directed transcriptional regulation of some novel miRNAs. BMC Genomics, 10: 563.

93. Gunnery, S., Ma, Y., Mathews, M.B. 1999. Termination sequence requirements vary among genes transcribed by RNA polymerase III. JMB, 286: 745-757.

94. Hall, N. 2007. Advanced sequencing technologies and their wider impact in microbiology. J. Exp. Biol. 210 (9): 1518–25.

95. Han, K., Lee, J., Meyer, T.J., Wang, J., Sen, S.K., Srikanta, D., Liang, P., Batzer, M.A. 2007. Alu recombination-mediated structural deletions in the chimpanzee genome. PLoS Genetics, 3 (10): e184.

96. Han, K., Xing, J., Wang, H., Hedges, D.J., Garber, R.K., Cordaux, R., Batzer, M.A. 2006. Under the genomic radar: the stealth model of Alu amplification. Genome Research, 15: 655-664.

97. Hasler, J., Strub, K. 2006. Alu RNP and Alu RNA regulate translation initiation *in vitro*. Nucl. Acids Research, 34: 2374-2385.

98. Hellmann-Blumberg, U., Macarthy Hintz, M.F., Gatewood, J.M., Schimd, C.W. 1993. Developmental differences in methylation of human Alu repeats. Molecular and Cellular Biology, 13(8): 4523-4530.

99. Herbert, A., Rich, A. 2001. The role of binding domains for dsRNA and Z-DNA in the in vivo editing of minimal substrates by ADAR1. PNAS, 98(21): 12132-12137.

100. Higuchi, M., Maas, S., Single, F.N., Hartner, J., Rozov, A., Burnashev, N., Fledmeyer, D., Sprengel, R., Seeburg, P.H. 2000. Point mutation in an AMPA receptor gene rescues lethality in mice deficient in the RNA-editing enzyme ADAR2. Nature. 406:78-81.

101. Higuchi, M., Single, F.N., Kohler, M., Sommer, B., Sprengel, R., Seeburg, P.H. 1993. RNA editing of AMPA receptor subunit GluR-B: A base-paired intron-exon structure determines position and efficiency.

102. Hirose, Y., Okhuma, Y. 2007. Phosphorylation of the C-terminal domain of RNA polymerase II plays central roles in the integrated events of eukaryotic gene expression. Journal of Biochemistry, 141: 601-608.

103. Hsu, K., Chang, D.Y., Maraia, R.J. 1995. Human signal recognition particle (SRP) Alu-associated protein also binds Alu interspersed repeat sequence RNAs. The Journal of Biological Chemistry, 270 (17): 10179-10186.

104. Huck, L., Scherrer, A., Terzi, L., Johnson, A.E., Bernstein, H.D., Cusack, S., Weichenrieder, O., Strub, K. 2004. Conserved tertiary base pairing ensures proper RNA folding and efficient assembly of the signal recognition particle Alu domain. Nucleic Acids Research, 32 (16): 4915-4924.

105. Hundley, H.A., Krauchuk, A.A., Bass, B.L. 2008. C. elegans and H. sapiens mRNAs with edited 3' UTRs are present on polysomes. RNA, 14: 1-11.

106. Hurst S.R., Hough, R.F., Aruscavage, P.J., Bass, B.L. 1995. Deamination of mammalian glutamate receptor RNA by xenopus dsRNA adenosine deaminase: similarities to in vivo RNA editing. RNA, 1(10): 1051-1060.

107. Hutvagner, G. 2005. Small RNA asymmetry in RNAi: function in RISC assembly and gene regulation. FEBS Letters, 579: 5850-5857.

108. International Human Genome Sequencing Consortium. 2001. Initial sequencing and analysis of the human genome. Nature, 409, 860-921.

109. Ishiguro, A., Kassavetis, G.A., Geiduschek, E.P. 2002. Essential roles of Bdp1, a subunit of RNA polymerase III initiation factor TFIIIB, in transcription and tRNA processing. Molecular and Cellular Biology, 22 (10): 3264-3275.

110. Januszyk, K., Li, P.W.L., Villareal, V., Branciforte, D., Wu, H., Xie, Y., Feigon, J., Loo, J.A., Martin, S.L., Clubb, R.T. 2007. Identification and solution structure of a highly conserved C-terminal domain within ORF1p required for retrotransposition of long interspersed nuclear element-1. Journal of Biological Chemistry, 282 (34): 24893-24904.

111. Kale, S.P., Moore. L., Deininger, P.L., Roy-Engel, A.M. 2005. Heaby metals stimulate human LINE-1 retrotransposition. Int. J. Environ. Res Public Health, 2(1): 14-23.

112. Kallmann, A.M., Sahlin, M., Ohman, M. 2003. ADAR2 A→I editing: site selectivity and editing efficiency are separate events. Nucleic Acids Research, 31(16): 4874-4881.

113. Kaneko, H., Dridi, S., Tarallo, V., Gelfand, B.D., Fowler, B.J., Cho, W.G., Kleinman, M.E., Ponicsan, S.L., Hauswirth, W.W., Chido, V.A., Kariko, K., Yoo, J.W., Lee, D.K., Hadziahmetovic, M., Song, Y., Misra, S., Chaudhuri, G., Buaas, F.W., Braun, R.E., Hinton, D.R., Zhang, Q., Grossniklaus, H.E., Provis, J.M., Madigan, M.C., Milam, A.H., Justice, N.L., Albuquerque, R.J.C., Blandford, A.D.,

Bogdanovich, S., Hirano, Y., Witta, J., Fuchs, E., Littman, D.R., Ambati, B.K., Rudin, C.M., Chong, M.M.W., Provost, P., Kugel, J.F., Goodrich, J.A., Dunaief, J.L., Baffi, J.Z., Ambati, J. 2010. DICER1 deficit induces Alu RNA toxicity in age-related macular degeneration. Nature, 471(7338): 325-330.

114. Kang, M-I., Rhyu, M-G., Kim, Y-H., Jung, Y-C., Hong, S-J., Cho, C-S., Kim, H-S. 2006. The length of CpG islands is associated with the distribution of Alu and L1 retroelements. Genomics, 87: 580-590.

115. Kano, H., Godoy, I., Courtney, C., Vetter, M.R., Gerton.G.L., Osterta, E.M., Kazazian Jr., H.H. 2009. L1 retrotransposition occurs mainly in embrygenesis and creates somatic mosaicism. Genes & Development, 23: 1303-1312.

116. Kapp, L.D., Lorsch, J.R. 2004. The molecular mechanics of eukaryotic translation. Annual Review of Biochemistry 73 (1): 657–704

117. Kawahara, Y., Ito, K., Sun, H., Kanazawa, I., Kwak, S. 2003. Low editing efficiency of GluR2 mRNA is associated with low relative abundance of ADAR2 mRNA in white matter of normal human brain. European Journal of Neuroscience, 18 (1): 22-23.

118. Kawahara, Y., Megraw, M., Kreider E., Iizasa, H., Valente, L., Hatzegerogiou, A.G., Nishikura, K. 2008. Frequency and fate of microRNA editing in human brain. Nucleic Acids Research, 36 (16): 5270-5280.

119. Kawahara, Y., Zinshteyn, B., Chendrimada, T.P., Shiekhattar, R., Nishikura, K. 2007. RNA editing of the microRNA-151 precursor blocks cleavage by the Dicer-TRBP complex. EMBO, 8 (8): 763-769.

120. Kawahara, Y., Zinshteyn, B., Sethupathy, P., Iizasa, H., Hatzigeorgiou, A.G., Nishikur, K. 2007. Redirection of silencing targets by adenosine-to-inosine editing of miRNAs. Science, 315: 1137-1140.

121. Kenneth, N.S., Marshall, L., White, R.J. 2008. Recruitment of RNA polymerase III in vivo. Nucleic Acids Research, 36 (11): 3757-3764.

122. Kim, D. D.Y., Kim, T. T.Y., Walsh, T., Kobayashi, Y., Matise, T. C., Buyske, S., Gabriel, A. 2007. Widespread RNA editing of embedded Alu elements in the human transcriptome. Genome Research, 14:1719-1725.

123. Kim, E.Y., Bhattacharya, T., Kuntsman, K., Swantek, P., Koning, F.A., Malim, M.H., Wolinsky, S.M. 2010. Human APOBEC3G-mediated editing can promote HIV-1 sequence diversification and accelerate adaptation to selective pressure. Journal of Virology, 84 (19): 10402-10405.

124. Kim, T.M., Chung, Y.J., Rhyu, M.G., Jung, M.H. 2007. Germline methylation patterns inferred from local nucleotide frequency of repetitive sequences in the human genome. Mamm Genome 18: 277-285.

125. Kim, T.M., Hong, S.J., Rhyu, M.G. 2004. Periodic explosive expansion of human retroelements associated with the evolution of the hominoid primate. J Korean Med. Sci., 19: 177-185.

126. Kleinberger, Y., Eisenberg, E. 2011. Large-scale analysis of structural, sequence and thermodynamic characteristics of A-to-I RNA editing sites in human Alu repeats. BMC Genomics, 11: 453.

127. Kochanek, S., Renz, D., Doerfler, W. 1995. Transcriptional silencing of human Alu sequences and inhibition of protein binding in the box B regulatory elements by 5'-CG-3' methylation. FEBS, 360: 115-120.

128. Koeris, M., Funke, L., Shrestha, J., Rich, A., Maas, S. 2005. Modulation of ADAR1 editing activity by Z-RNA in vitro. Nucleic Acids Research, 33(16): 5362-5370.

129. Kojima, K.K. 2010. Different integration site structures between L1 protein-mediated retrotransposition in cis and retrotransposition in trans. Mobile DNA, 1: 17.

130. Kroutter, E.N., Belancio, V.P., Wagstaff, B.J., Roy-Engel, A.M. 2009. The RNA Polymerase dictates ORF1 requirement and timing of LINE and SINE retrotransposition. PLoS Genetics, 5 (4): e1000458.

131. Kulpa, D.A., Moran, J.V. 2006. Cis-preferential LINE-1 reverse transcriptase activity in ribonucleoprotein particles. Nature Structural & Molecular Biology, 13 (7): 655-660.

132. Lai, F., Chen, C.X., Carter, K.C., Nishikura, K. 1997. Editing of glutamate receptor B subunit ion channel RNAs by four alternatively spliced DRADA2 double-stranded RNA adenosine deaminases. Molecular and Cellular Biology, 17 (5): 2413-2424.

133. Laperriere, D., Wang, T-T., White, J.H., Mader S. 2007. Widespread Alu repeat-driven expansion of consensus DR2 retinoic acid response elements during primate evolution. BMC Genomics, 8: 1-12.

134. Laurencikiene, J., Kallman, A.M., Fong, N., Bentley, D.L., Ohman, M. 2005. RNA editing and alternative splicing: the importance of co-transcriptional coordination. EMBO, 7 (3): 303-307.

135. Lebedev, Y.B., Amosova, A.L., Mamedov, I.Z., Fisunov, G.Y., Sverdlov, E.D. 2007. Most recent AluY insertions in human gene introns reduce the content of the primary transcripts in a cell type specific manner. Gene, 390 (1-2): 122-129.

136. Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Radmark, O., Kim, S., Kim V.N. 2003. The nuclear RNase III Drosha initiates microRNA processing. Nature, 425: 415-419.

137. Lee, Y., Jeon, K., Lee, J-T., Kim, S., Kim, V.N. 2002. MicroRNA maturation: stepwise processing and subcellular localization. The EMBO, 21(17): 4663-4670.

138. Lee, J.Y., Ji, Z., Tian, B. 2008. Phylogenetic analysis of mRNA polyadenylation sites reveals a role of transposable elements in evolution of the 3'-end of genes. Nucleic Acids Research, 36 (17): 5581-5590.

139. Lehmann, K.A., Bass, B.L. 2000. Double-stranded RNA adenosine deaminases ADAR1 and ADAR2 have overlapping specificities. Biochemistry, 39: 12875-12884.

140. Lehnert, S., Van Loo, P., Thilakarathne, P.J., Marynen, P., Verbeke, G., Shuit, F.C. 2009. Evidence for co-evolution between human microRNAs and Alu-repeats. PLoS ONE, 4 (2): e4456.

141. Lev-Maor G., Sorek, R., Levanon, E.Y., Paz, N., Eizenberg, E., Ast, G. 2007. RNA-editing-mediated exon evolution. Genome Biology 8: R29.

142. Lev-Maor, G., Sorek, R., Shomron, N., Ast, G. 2003. The birth of an alternatively spliced exon: 3' splice-site selection in Alu exons. Science, 300: 1288-1291.

143. Levy, A., Schwartz, S., Ast, G. 2009. Large-scale discovery of insertion hotspots and preferential integration sites of human transposed elements. Nucleic Acids Research, 38 (5): 1515-1530.

144. Li, J., Han, K., Xing, J., Kim, H.S., Rogers, J., Ryder, O.A., Disotell, T., Yue, B., Batzer, M.A. 2009. Phylogeny of the macaques (Cercopithecidae: Macaca) based on Alu elements. Gene, 448 (2): 242-249.

145. Li, J.B., Levanon, E.Y., Yoon, J.K., Aach, J., Xie, B., Leproust, E., Zhang, K., Gao, Y., Church, G.M. 2009. Genome-wide identification of human RNA editing sites by parallel DNA capturing and sequencing. Science, 324: 1210-1213.

146. Li, M., Wang, I.X., Li, Y., Bruzel, A., Richards, A.L., Toung, J.M., Cheung, V.G. 2011. Widespread RNA and DNA sequence differences in the human transcriptome. Science, 333 (6038): 53-58.

147. Li, T.H., Schmid, C.W. 2001. Differential stress induction of individual Alu loci: implications for transcription and retrotransposition. Gene, 276: 135-141.

148. Lin, L., Shen, S., Tye, A., Cai, J.J., Jiang, P., Davidson, B.J., Xing, Y. 2008. Diverse splicing patterns of exonized Alu elements in human tissue. PLoS Genetics, 4 (10): e1000225.

149. Liu, G.E., Alkan, C., Jiang, L., Zhao, S., Eichler, E.E. 2009. Comparative analysis of Alu repeats in primate genomes. Genome Research, 19 (5): 876-885.

150. Liu, W.M., Chu, W.M., Choudary, P.V., Schmid, C.W. 1995. Cell stress and translational inhibitors transiently increase the abundance of mammalian SINE transcripts. Nucleic Acids Research, 23(10): 1758-1765.

151. Liu, W.M., Maraia, R.J., Rubin, C.M., Schmid, C.W. 1994. Alu transcripts: cytoplasmic localisation and regulation by DNA methylation. Nucleic Acids Research, 22(6): 1087-1095.

152. Loewer, S., Cabili, M.N., Guttman, M., Loh, Y.H., Thomas, K., Park, I.H., Garber, M., Curran, M., Onder, T., Agarwell, S., Manos, P.D., Datta, S., Lander, E.S., Schlaeger, T.M., Daley, G.Q., Rinn, J.L. 2010. Large intergenic non-coding RNA-RoR modulates reprogramming of human induced pluripotent stem cells. Nature Genetics, 42 (12): 1113-1117.

153. Lomeli, H., Mosbacher, J., Melcher, T., Hoger, T., Geiger, J.R.P., Kuner, T., Monyer, H., Higuchi, M., Bach, A., Seeburg, P.H. 1994. Control of kinetic properties of AMPA receptor channels by nuclear RNA editing. Science, 266:1709-1713.

154. Luciano, D.J., Mirsky, H., Vendetti, N.J., Maas, S. 2004. RNA editing of a miRNA precursor. RNA, 10: 1174-1177.

155. Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F., Richmond, T.J.1997. Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature 389 (6648): 251–60.

156. Lukic, S., Chen, K. 2011. Human piRNAs are under selection in Africans and repress transposable elements. Mol. Biol. Evol. 28 (11): 3061-3067.

157. Maas, S., Godfried Sie, C.P., Stoev, I., Dupuis, D.E., Latona, J., Porman, A.M., Evans, B., Rekawek, P., Kluempers, V., Mutter, M., Gommans, W.M., Lopresti, D. 2011. Genome-wide evaluation and discovery of vertebrate A-to-I RNA editing sites. Biochem Biophys Res Commun., 412 (3): 407-412.

158. Macbeth, M.R., Schubert, H.L., VanDemark, A.P., Lingam, A.T., Hill, C.P., Bass, B.L. 2005. Inositol hexakisphosphate is bound in the ADAR2 core and required for RNA editing. Science, 309: 1534-1539.

159. Macia, A., Munoz-Lopez, M., Cortes, J.L., Hastings, R.K., Morell, S., Lucena-Aguilar, G. Marchal, J.A., Badge, R.M., Garcia-Perez, J.L. 2011. Epigenetic control of retrotransposon expression in human embryonic stem cells. Molecular and Cellular Biology, 31(2): 300-316.

160. Maquat, L.E. 2005. Nonsense-mediated mRNA decay in mammals. Journal of Cell Science, 118 (9): 1773-1776.

161. Maraia, R.J. 1996. Transcription termination factor La is also an initiation factor for RNA polymerase III. Proc. Natl. Acad. Sci., 93: 3383-3387.

162. Maraia, R.J., Kenan, D.J., Keene, J.D. 1994. Eukaryotic transcription termination factor La mediates transcript release and facilitates reinitiation by RNA Pol III. Molecular and Cellular Biology, 14 (3): 2147-2158.

163. Marcucci, R., Brindle, J., Paro, S., Casadio, A., Hempel, S., Morrice, N., Bisso, A., Keegan, L.P., Del Sal, G., O'Connell, M.A. 2011. Pin1 and WWP2 regulate GluR2 Q/R site RNA editing by ADAR2 with opposing effects.

164. Margulies, M., Egholm, M., Altman, W.E., et al 2005. Genome sequencing in open microfabricated high density picoliter reactors. Nature 437 (7057): 376–80

165. Mariner, P.D., Walters, R.D., Espinoza, C.A., Drullinger, L.F., Wagner, S.D., Kugel, J.F., Goodrich, J.A. 2008. Human Alu mRNA is a modular transacting repressor of mRNA transcription during heat shock. Molecular Cell, 29: 499-509.

166. Martin, S.L., Cruceanu, M., Branciforte, D., Li, P.W.L., Kwok, S.C., Hodges, R.S., Williams, M.C. 2005. LINE-1 retrotransposition requires the nucleic acid chaperone activity of the ORF1 protein. Journal of Molecular Biology, 348: 549-561.

167. Mathias, S.L., Scott, A.F., Kazazian, H.H.J., Boeke, J.D., Gabriel, A. 1991. Reverse transcriptase encoded by a human transposable element. Science, 254: 1808-1810.

168. Matsuzaki, H., Kassavetis, G.A., Geiduschek, E.P. 1994. Analysis of RNA chain elongation and termination by *Saccharomyces cerevisiae* RNA polymerase III. Journal of Molecular Biology, 235: 1173-1192.

169. McManus, C.J., Graveley, B.R. 2011. RNA structure and the mechanisms of alternative splicing. Curr. Opin. Genet Dev., 21 (4): 373-379.

170. Melcher, T., Maas, S., Herb, A., Sprengel, R., Higuchi, M., Seeburg, P.H. 1996. RED2, a brain-specific member of the RNA-specific adenosine deaminase family. J. Biol. Chem, 271: 31795-31798.

171. Melcher, T., Maas, S., Herb, A., Sprengel, R., Seeburg, P.H., Higuchi, M. 1996. A mammalian RNA editing enzyme. Nature, 379: 460-464.

172. Mills, R.E., Bennett, E.A., Iskow, R.C., Luttig, C.T., Tsui, C., Pittard, W.S., Devine, S.E. 2006. Recently mobilized transposons in the human and chimpanzee genomes. The American Journal of Human Genetics, 78: 671-679.

173. Moller-Krull, M., Zemann, A., Roos, C., Brosius, J., Schmitz, J. 2008. Beyond DNA: RNA-editing and steps toward Alu-exonization in primates. Journal of Molecular Biology, 382 (3): 601-609.

174. Morabito, M.V., Abbas, A.I., Hood, J.L., Kesterson, R.A., Jacobs, M.M., Kump, D.S., Hachey, D.L., Roth, B.L., Emeson, R.B. 2010. Mice with altered serotonin 2C receptor RNA editing display characteristics of Prader-Willi Syndrome. Neurobiol Dis., 39 (2): 169-180.

175. Morse, D.P. and Bass, B.L. 1999. Long RNA hairpins that contain inosine are present in caenorhabditis elegans poly(A)+ RNA. Proc. Natl. Acad. Sci. USA, 96(11): 6048-6053.

176. Muiznieks, I., Doerfler, W. 1994. The topology of the promoter of RNA polymerase II- and III-transribed genes is modified by the methylation of 5'-CG-3' dinucleotides.

177. Mustajoki, S., Ahola, H., Mustajoki, P., Kauppinen, R. 1999. Insertion of Alu element responsible for acute intermittent poryphyria. Human Mutations, 13 (6): 431-438.

178. Nicholas, A., de Magalhaes, J.P., Kraytsberg, Y., Richfield, E.K., Levanon, E.Y., Khrapko, K. 2010. Age-related gene-specific changes of A-to-I mRNA editing in the human brain. Mech Ageing Dev., 131 (6): 445-447.

179. Nie, Y., Hammond, G.L., Yang, J.H. 2006. Double stranded RNA deaminase ADAR1 increases host susceptibility to virus infection. Journal of Virology, 81 (2): 917-923.

180. Nikitina, T.V., Tischenko, L.I., Schulz, W.A. 2011. Recent insights into regulation of transcription by RNA polymerase III and the cellular functions of its transcripts. Biol. Chem., 392 (5): 395-404.

181. Nishikura, K. 2010. Functions and regulation of RNA editing by ADAR deaminases. Annu. Rev. Biochem. 79: 321-349.

182. Ohman, M., Kallman, A.M., Bass, B.L. 2000. In vitro analysis of the binding of ADAR2 to the pre-mRNA encoding the GluR-B R/G site. RNA, 6: 687-697.

183. Osenberg, S., Yaacov, N.P., Safran, M., Moshkovitz, S., Shtrichman R., Sherf, O., Jacob-Hirsch, J., Keshet, G., Amariglio, N., Itskovitz-Eldor, J., Rechavi, G. 2010. Alu sequences in undifferentiated human embryonic stem cells display high levels of A-to-I RNA editing. PLoS ONE, 5 (6): e11173.

184. Pandey, R., Mandal, A.K., Jha, V., Mukerji, M. 2011. HSF binding in Alu repeats expands its involvement in stress through an antisense mechanism. Genome Biology, 12 (11): R117.

185. Panning, B. and Smiley, J.R. 1995. Activation of expression of multiple subfamilies of human Alu elements by adenovirus type 5 and herpes simplex virus type 1. JMB, 248:513-524.

186. Paz, N., Levanon, E.Y., Amariglio, N., Heimberger, A.B., Ram, Z., Constantini, S., Barbash, Z.S., Adamsky, K., Safran, M., Hirschberg, A., Krupsky, M., Ben-Dov, I., Cazacu, S., Mikkelsen, T., Brodie, C., Eisenberg, E., Rechavie, G. 2007. Altered adenosine-to-inosine RNA editing in human cancer.

187. Paz-Yaacov, N., Levanon, E.Y., Nevo, E., Kinar, Y., Harmelin, A., Jacob-Hirsch, J., Amariglio, N., Eisenberg, E., Rechavi, G. 2010. Adenosine-to-inosine RNA editing shapes transcriptome diversity in primates. PNAS. 107 (27): 12174-12179.

188. Peng, Z., Cheng, Y., Tan, B.C.M., Kang, L., Tian, Z., Zhu, Y., Zhang, W., Liang, Y., Hu, X., Tan, X., Guo, J., Dong, Z., Liang, Y., Bao, L., Wang, J. 2012. Comprehensive analysis of RNA-Seq data reveals extensive RNA editing in a human transcriptome. Nature Biotechnology, 30 (3): 253-260.

189. Petschek J.P., Mermer, M.J., Scheckelhoff, M.R., Simone, A.A., Vaughn, J.C. 1996. RNA editing in drosophilia 4f-rnp gene nuclear transcripts by multiple A-to-G conversions. JMB, 259: 885-890.

190. Pillai, R.S., Bhattacharyya, S.N., Artus, C.G., Zoller, T., Cougot, N., Basyuk, E., Bertrand, E., Filipowicz, W. 2005. Inhibition of translational initiation by Let-7 microRNA in human cells. Science, 309: 1573-1576.

191. Piriyapongsa, J., Ruledge, M.T., Patel, S., Borodovsky, M., Jordan, I.K. 2007. Evaluating the protein coding potential of exonized transposable element sequences. Biology Direct, 2:31.

192. Polak, P., Domany, E. 2006. Alu elements contain many binding sites for transcription factors and may play a role in regulation of developmental processes. BMC Genomics, 7: 1-15.

193. Polson, A.G., Bass, B.L., Casey, J.L. 1996. RNA editing of hepatitis delta virus antigenome by dsRNA-adenosine deaminase. Nature, 380 (6573): 454-456.

194. Pond, S.L., Murrell, B., Poon, A.F. 2012. Evolution of viral genomes: interplay between selection, recombination, and other forces. Methods Mol. Biol., 856: 239-72.

195. Porreca, J.G. 2010. Genome sequencing on nanoballs. Nature Biotechnology, 28: 43-44.

196. Poulsen, H., Nilsson, J., Damgaard, C.K., Egebjerg, J., Kjems, J. 2001. CRM1 mediates the export of ADAR1 through a nuclear export signal within the Z-DNA binding domains. Mol. Cell. Biol., 21: 7862-7871.

197. Price, R.D., Weiner, D.M., Chang, M.S.S., Sanders-Bush, E. 2001. RNA editing of the human serotonin 5-HT_{2C} receptor alters receptor-mediated activation of G_{13} protein. The Journal of Biological Chemistry, 276 (48): 44663-44668.

198. Quentin, Y. 1992. Fusion of a free left Alu monomer and a free right Alu monomer at the origin of the Alu family in the primate genomes. Nucleic Acids Research, 20(3): 487-493.

199. Rausch, J.W., Chelico, L., Goodman, M.F., Le Grice, S.F.J. 2009. Dissecting APOBEC3G substrate specificity by nucleoside analog interface. The Journal of Biological Chemistry, 284 (11): 7047-7058.

200. Ricci, V., Regis, S., Di Duca, M., Filocama, M. 2003. An Alu-mediated rearrangement as cause of exon skipping in Hunter Disease. Human Genetics, 112: 419-425.

201. Rodriguez, J., Vives, L., Jorda, M., Morales, C., Munoz, M., Vendrell, E., Peinado, M.A. 2008. Genome-wide tracking of unmethylated DNA Alu repeats in normal and cancer cells. Nucleic Acids Research, 36(3): 770-784.

202. Roy-Engel, A.M., Salem, A-H., Oyeniran, O.O., Deininger, L., Hedges, D.J., Kilroy, G.E., Batzer, M.A., Deininger, P.L. 2002. Active Alu element “A-tails”: size does matter. Genome Research, 12: 1333-1344.

203. Roy, A.M., West, N.C., Rao, A., Adhikari, P., Aleman, C., Barnes, A.P., Deininger, P.L. 2000. Upstream flanking sequences and transcription of SINEs. JMB, 302:17-25.

204. Rubin, C.M., Kimura, R.H. Schmid, C.W. 2002. Selective stimulation of translational expression by Alu RNA. Nucleic Acids Research, 30(14): 3253-3261.

205. Rubin, C.M., VandeVoort, C.A., Teplitz, R.L., Schmid, C.W. 1994. Alu repeated DNAs are differentially methylated in primate germ cells. Nucleic Acid Research, 22(23): 5121-5127.

206. Rudin, C.M., Thompson, C.B. 2001. Transcriptional activation of short interspersed elements by DNA damaging agents. Genes, Chromosomes & Cancer, 30:64-71.

207. Ryman, K., Fong, N., Bratt, E., Bentley, D.L., Ohman, M. 2007. The C-terminal domain of RNA Pol II helps ensure that editing precedes splicing of the GluR-B transcript. RNA, 13: 1071-1078.

208. Samuel, C.E. Adenosine deaminases acting on RNA (ADARs) are both antiviral and proviral.

209. Sansam, C.L., Wells, K.S., Emeson, R.B. 2003. Modulation of RNA editing by functional m nucleolar sequestration of ADAR2. PNAS, 100(24): 14018-14023.

210. Sarrowa, J., Chang, D.Y., Maraia, R.J. 1997. The decline in human Alu retrotransposition was accompanied by an asymmetric decrease in SRP9/14 binding to dimeric Alu RNA and increased expression of small cytoplasmic Alu RNA. Molecular and Cellular Biology, 17(3): 1144-1151.

211. Scadden, A.D.J. 2005. The RISC subunit Tudor-SN binds to hyper-edited double-stranded RNA and promotes its cleavage. Nature Structural & Molecular Biology, 6: 489-496.

212. Scadden, A.D.J. 2007. Inosine-containing dsRNA binds a stress-granule-like complex and downregulates gene expression in trans. Molecular Cell, 28: 491-500.

213. Schoft, V.K., Schopoff, S., Jantsch, M.F. 2007. Regulation of glutamate receptor B pre-mRNA splicing by RNA editing. Nucleic Acids Research, 35 (11): 3723-3732.

214. Schuster, S.,C. 2008. Next-generation sequencing transforms today's biology. Nat. Methods 5 (1): 16–8.

215. Schwartz, S., Gal-Mark, N., Kfir, N., Oren, R., Kim, E., Ast, G. Alu exonization events reveal features required for precise recognition of exons by the splicing machinery. PLoS Computational Biology, 5 (3): e1000300.

216. Sela, N., Mersch, B., Gal-Mark, N., Lev-Maor, G., Hotz-Wagenblatt, A., Ast, G. 2007. Comparative analysis of transposed elements' insertion within human and mouse genomes reveals Alu's unique role in shaping the human transcriptome. Genome Biology, 8: R127.

217. Shaikh, T.H., Roy, A.M., Kim, J., Batzer, M.A., Deininger, P.L. 1997. cDNAs derived from primary and small cytoplasmic Alu (scAlu) transcripts. JMB, 271: 222-234.

218. Shankar, R., Grover, D., Brahmachari, S.K., Mukerji, M. 2004. Evolution and distribution of RNA polymerase II regulatory sites from RNA polymerase III dependent mobile Alu elements. BMC Evolutionary Biology, 4: 37.

219. Shen, S., Lin, L., Cai, J.J., Jiang, P., Kenkel, E.J., Stroik, M.R., Sato, S., Davidson, B.L., Xing, Y. 2011. Widespread establishment and regulatory impact of Alu exons in human genes. PNAS, 108 (7): 2837-2842.

220. Smalheiser, NR., Torvik, VI. 2006. Alu element within human mRNAs are probable microRNA targets. TRENDS in Genetics, 22: 532-536.

221. Sommer B, Köhler M., Sprengel R., Seeburg PH. 1991. RNA editing in brain controls a determinant of ion flows in glutamate-gated channels. Cell, 67 (1): 11-9.

222. Srikanta, D., Sen, S.K., Conlin, E.M., Batzer, M.A. 2009. Internal priming: An opportunistic pathway for L1 & Alu retrotransposition in Hominins. Gene, 448 (2): 233-241.

223. Srikanta, D., Sen, S.K., Huang, C.T., Conlin, E.M., Rhodes, R.M., Batzer, M.A. 2008. An alternative pathway for Alu retrotransposition suggests a role in DNA double-strand break repair. Genomics, YGENO-08042.

224. Stefl, R., Oberstrass, F.C., Hood, J.L., Jourdan, M., Zimmermann, M., Skrisovska, L., Maris, C., Peng, L., Hofr, C., Emeson, R.B., Allain, F.H.T. 2010. The solution structure of the ADAR2 dsRBM-RNA complex reveals a sequence specific readout of the minor groove. Cell, 143, 225-237.

225. Steinberg, T.H., Mathews, D.E., Durbin, R.D., Burgess, R.R. 1990. Tagetitoxin: a new inhibitor of eukaryotic transcription by RNA Polymerase III. The Journal of Biological Chemistry, 265(1): 499-505.

226. Styles, P., Brookfield, J.F.Y. 2007. Analysis of the features and source gene composition of the AluYg6 subfamily of human retrotransposition. BMC Evolutionary Biology, 7: 102.

227. Su, A.A., Randau, L. 2011. A-to-I and C-to-U editing within transfer RNAs. Biochemistry (Mosc.), 76 (8): 932-937.

228. Suspene, R., Petit, V., Puyraimond-Zemmour, D., Aynaud, M.M., Henry, M., Guetard, D., Rusniok, C., Wain-Hobson, S., Vartanian, J.P. 2010. Double-stranded RNA adenosine deaminase ADAR-1-induced hypermutated genomes among inactivated seasonal influenza and live attenuated measles vaccines. Journal of Virology, 85 (5): 2458-2462.

229. Suzuki, J., Yamaguchi, K., Kajikawa, M., Ichiyanagi, K., Adachi, N., Koyama, H., Takedo, S., Okada, N. PLoS Genetics, 5 (4): e1000461.

230. Tan, L., Sarkis, P.T., Wang, T., Tian, C., Yu, X.F. 2009. Sole copy of Z2-type human cytidine deaminase APOBEC3H has inhibitory activity against retrotransposons and HIV-1. FASEB J., 23 (1): 279-287.

231. Tanaka, Y., Yamashita, R., Suzuki, Y., Nakai, K. 2010. Effects of Alu elements on global nucleosome positioning in the human genome. BMC Genomics. 11: 309.

232. Taylor, D.R., Puig, M., Darnell, M.E., Mihalik, K., Feinstone, S.M. 2005. New antiviral pathway that mediates hepatitis C virus replicon interferon sensitivity through ADAR1.

233. The 1000 Genome Consortium. 2011. A map of human genome variation from population-scale sequencing. Nature, 467: 1067-1073.

234. Tighe, P.J., Stevens, S.E., Dempsey, S., Le Deist, F., Rieux-Laucat, F., Edgar, J.D.M. 2002. Inactivation of the Fas gene by Alu insertion: retrotransposition in an intron causing splicing variation and autoimmune lymphoproliferative syndrome. Genes and Immunity, 3: 66-70.

235. Tonkin, L.A., Saccomanno, L., Morse, D.P., Brodigan, T., Krause, M., Bass, B.L. 2002. RNA editing by ADARs is important for normal behavior in Caenorhabditis elegans. EMBO J, 21 (22): 6025-6035.

236. Toth, A.M., Li, Z., Cattaneo, R., Samuel, C.E. 2009. RNA-specific adenosine deaminase ADAR1 suppresses measles virus-induced apoptosis and activation of protein kinase PKR. Journal of Biological Chemistry, 284 (43): 29350-29356.

237. Tsirigos, A., Rigoutsos, I. 2009. Alu and B1 repeats have been selectively retained in the upstream and intronic regions of genes of specific functional classes. PLoS Computational Biology, 5 (12): e1000610.

238. Urrutia, A.O., Ocana, L.B., Hurst, L.D. 2008. Do Alu repeats drive the evolution of the primate transcriptome? Genome Biology, 9 (2): R25.

239. Valente, L., Nishikura, K. 2007. RNA-binding independent dimerization of adenosine deaminases acting on RNA and dominant negative effects of nonfunctional subunits on dimer functions. The Journal of Biological Chemistry, 282(22): 16054-16061.

240. Vitali, P, Basyuk, E., Meur, E.L., Bertrand. E., Muscatelli, F., Cavaille, J., Huttenhofer, A. 2005. ADAR2-mediated editing of RNA substrates in the nucleolus is inhibited by C/D small nucleolar RNAs. JCB, 169(5): 745-753.

241. Wagner, S.D., Kugel, J.F., Goodrich, J.A. 2010. TFIIF facilitates dissociation of RNA polymerase II from non-coding RNAs that lack a repression domain. Molecular and Cellular Biology, 30 (1): 91-97.

242. Wagstaff, B.J., Barnerboi, M., Roy-Engel, A.M. 2011. Evolutionary conservation of the functional modularity of primate and murine LINE-1 elements. PLoS ONE, 6 (5): e19672.

243. Wahlstedt, H., Daniel, C., Enstero, M., Ohman, M. 2009. Large-scale mRNA sequencing determines global regulation of RNA editing during brain development. Genome Research, 19: 978-986.

244. Wallace, N., Wagstaff, B.J., Deininger, P.L., Roy-Engel, A.M. 2008. LINE-1 ORF1 protein enhances Alu SINE retrotransposition. Gene, 419:1-6.

245. Wang, Q., Khillian, J., Gadue, P., Nishikura, K. 2000. Requirement of the RNA editing deaminase ADAR1 gene for embryonic erythropoiesis. Science, 290 (5497): 1765-1768.

246. Wang, X., Fan, J., Liu, D., Fu, S., Ingvarsson, S., Chen, H. 2011. Spreading of Alu methylation to the promoter of the MLH1 gene in gastrointestinal cancer. PLoS ONE, 6 (10): e25913.

247. Ward, S.V., George, C.X., Welch, M.J., Liou, L.Y., Hahm, B., Lewicki, H., de la Torre, J.C., Samuel, C.E., Oldstone, M.B. 2011. RNA editing enzyme adenosine deaminase is a restriction factor for controlling measles virus replication that also is required for embryogenesis. PNAS, 108 (1): 331-336.

248. Wild, K., Halic, M., Sinning, I., Beckmann, R. 2004. SRP meets the ribosome. Nature Structural & Moleculare Biology, 11(11): 1049-1053.

249. Witherspoon, D.J., Watkins, W.S., Zhang, Y., Xing, J., Tolpinrud, W.L., Hedges, D.J., Batzer, M.A., Jorde, L.B. 2009. Alu repeats increase local recombinase rates. BMC Genomics, 10: 530.

250. Wong, S.K., Sato, S., Lazinski, D.W. 2001. Substrate recognition by ADAR1 and ADAR2. RNA, 7: 846-858.

251. Xie, H., Wang, M., Bonaldo, M.D.F., Rajaram, V., Stellpflug, W., Smith, C., Arndt, K., Goldman, S., Tomita, T., Soares, M.B. 2010. Epigenomic analysis of Alu repeats in human ependymomas. PNAS, 107(12): 6952-6957.

252. Xing, J., Witherspoon, D.J., Ray, D.A., Batzer, M.A., Jorde, L.B. 2007. Mobile DNA elements in primate and human evolution. Yearbook of Physical Anthropology 50: 2-19.

253. Xu, M., Wells, K.S., Emeson, R.B. 2006. Sustrate-dependent contribution of double-stranded RNA-binding motifs to ADAR2 function. Molecular Biology of the Cell, 17: 3211-3220.

254. Yang, W., Chendrimada, T.P., Wang, Q., Higuchi, M., Seeburg, P.H., Shiekhattar, R., Nishikura, K. 2006. Modulation of microRNA processing and expression through RNA editing by ADAR deaminases. Nature Structural & Molecular Biology, 13 (1): 13-21.

255. Zhang, Z., Carmichael, G.G. 2001. The fate of dsRNA in the nucleus: a p54nrb-containing complex mediates the nuclear retention of promiscuously A-to-I edited RNAs.

256. Zhang, W., Edwards, A., Fan, W., Deininger, P., Zhang, K. 2011. Alu distribution and mutation types of cancer genes. BMC Genomics, 12:157.

257. Zhang, R., Wang, Y.Q., Su, B. 2008. Molecular evolution of a primate-specific microRNA family. Mol. Biol. Evol., 25 (7): 1493-1502.

258. Zhou, Y-H., Zheng, J. B., Gu, X., Saunders, G.F., Yung, W-K. A. 2002. Novel PAX6 binding sites in the human genome and the role of repetitive elements in the evolution of gene regulation. Genome Research, 12: 1716-1722.

Curriculum Vitae

DYLAN E. DUPUIS

Curriculum Vitae

Contact Information

Work Address
Iacocca Hall
B-217
111 Research Drive
Bethlehem, PA 18015

Work Phone
908-644-9574

Work E-mail
ded205@lehigh.edu

Education

Lehigh University
Bethlehem, PA
Molecular Biology, Ph.D. program Expected Graduation: Spring 2012

Lycoming College
Williamsport, PA
B.S. in Biology and Chemistry GPA: 3.534 Graduated: May 2005

Honors and Awards

1. *Marjorie Nemes Fellowship* Lehigh University Biological Sciences Department
2. *Estes Fellowship* Lehigh University Summer Fellowship
3. *American Chemical Society* Susquehanna Valley Section Senior Chemistry Award
4. *Gamma Sigma Epsilon* Roy M. Sonntag Award
5. *Beta Beta Beta* National Biology Honors Society
6. *Gamma Sigma Epsilon* National Chemistry Honors Society- *Rho Delta* Chapter

Employment & Work Experience

Lehigh University August 2005 – Present
Bethlehem, PA
Graduate Student- Biological Sciences Department

- Research Assistant to Professor Stefan Maas, Ph.D., - Fall 2008
- Biosystems Dynamic Summer Institute (HHMI-BDSI)- Summer Collaboration with Professor Dan Lopresti, Ph.D. (Lehigh University Computer Science Department)

and Professor Ian Laurenzi, Ph.D. (Lehigh University Chemical Engineering Department)

- Teaching Assistant - Fall 2005-Spring 2008 and Spring 2009-Present
  - Elements in Biochemistry I (Bios 371) - Fall 2009, Fall 2010
  - Molecular Genetics Laboratory (Bios 346) - Spring 2009, Spring 2010
  - Central Nervous System & Behavior (Bios 272) - Spring 2008
  - Elements in Biochemistry II (Bios 372) - Spring 2007
  - Bio Core I: Cell and Molecular Lab (Bios 042) - Spring 2005, Spring 2006
  - Biochemistry Laboratory (Bios 377) - Fall 2005, Fall 2006, Fall 2007

Chanel Inc. June 2004 – August 2004
Piscataway, NJ
Internship – Analytical Laboratory / QA Laboratory

- Tested pilot products to determine shelf-life stability
- Tested and cataloged properties of raw materials used to make products
- Performed authenticity tests on possible counterfeit products

Publications

Dupuis, D.E., Roberts, R., Kuchka, M., Maas, S. Transcriptionally active alu elements are targets of A-to-I RNA editing. In preparation.

Maas, S., Godfried Sie, C.P., Stoev, I., **Dupuis, D.E.**, Latona, J., Porman, A., Evans, B., Rekawek, P., Kluempers, V., Mutter, M.A., Gommans, W.M., Lopresti, D. 2011. Genome-wide evaluation and discovery of vertebrate A-to-I RNA editing sites. *BBRC* 12 (3): 407-412.

Dylan Dupuis and Stefan Maas. 2010. Chapter 18: MiRNA Editing. *Methods in Molecular Biology: MicroRNAs and the Immune System.* Ed. Silvia Montecelli. Humana Press. Volume 667, Part 5. 267-279.

Willemijn M. Gommans, **Dylan E. Dupuis**, Jill E. McCane, Nick E. Tatalias, Stefan Maas. 2008. Diversifying Exon Code through A-to-I RNA Editing. *RNA and DNA Editing: Molecular Mechanisms and Their Integration into Biological Systems.* Ed. Harold C. Smith. John Wiley & Sons Inc.: pp 3-30.

Gommans WM, Tatalias NE, Sie CP, **Dupuis D**, Vendetti N, Smith L, Kaushal R, Maas S. 2008. Screening of human SNP database identifies recoding sites of A-to-I RNA editing. *RNA* 14(10): 2074-85

Invited Platform Presentations and Seminars

Dupuis D, Kuchka M, Maas S. A-to-I RNA editing in the RNA Pol III Transcriptome: A Search for Novel Targets of a Prevalent Modification. Lehigh University Biological Sciences Department. Marjorie Nemes Fellowship Invitee. April 29, 2011.

Dupuis D. Gordon-Keenan Research Conference. Discussion Leader. Bioinfomatic and evolutionary analysis of editing and modification. January 2011

Dupuis D, Maas S. A-to-I RNA Editing In Transcriptionally Active Alu Elements. Lehigh University Biological Sciences Graduate Symposium. April 2010.

Conference Abstracts And Poster Presentations

Dupuis D, Maas S. Investigating A-to-I RNA Editing in RNA Pol III Transcripts. Gordon Research Conference: RNA Editing. January 2011.

Dupuis D, Maas S. RNA Pol III Transcripts are A-to-I RNA Editing Targets. Lehigh University Biological Sciences Department Annual Graduate Student Open House. October 2010.

Dupuis D, Maas S. A-to-I RNA Editing Within Transcriptionally Active Alu Retrotransposons. Gordon Research Conference: RNA Editing. January 2009.

Dupuis D, Maas S. The Impact of A-to-I RNA Editing on Transcriptionally Active Alu Elements. Lehigh University Biological Sciences Department Annual Graduate Student Open House. October 2007.

Maas S, Lopresti D, Laurenzi I, **Dupuis D,** Strohmaeir M, Evans B, Latona J, Porman A, Rekawek P. Intergrating Computational and Experimental Methods in the Analysis of Molecular Complexity in Nature. HHMI- Biosystems Dynamics Summer Institute. August 2007.

Dupuis D, Maas S. Evidence of A-to-I RNA Editing in MCC2 mRNA. Lehigh University Biological Sciences Department Annual Graduate Student Open House. October 2006.

Professional Activities

Wrote supplemental online questions for *Cells 2nd Edition.* Chapter 2:Bioenergetics and Cellular Metabolism and Chapter 10: Chromatin and Chromosomes. Ed. Benjamin Lewin. Cell Press and Virtual Text.

References

Stefan W. Maas, Ph. D.
Associate Professor
Biological Sciences
Lehigh University
Iacocca Hall
Room B-217
111 Research Drive
Bethlehem, PA 18015
Phone:610-758-6276
Email: swm3@lehigh.edu

Michael Kuchka, Ph. D.
Associate Professor
Biological Sciences
Lehigh University
Iacocca Hall
Room B-217
111 Research Drive
Bethlehem, PA 18015

Phone: 610-758-3687
Email: mrk5@lehigh.edu

Daniel P. Lopresti, Ph. D.
Professor and Department Chair
Computer Science and Engineering
Lehigh University
19 Memorial Drive West
Bethlehem, PA 18015
Phone: 610-758-5782
Email: lopresti@cse.lehigh.edu