universität wien

DISSERTATION

Titel der Dissertation

„Impact of inverted SINEs on gene expression“

Verfasserin

Mansoureh Tajaddod

angestrebter akademischer Grad

Doctor of Philosophy (PhD)

Wien, February 2015

Studienkennzahl lt. Studienblatt:	A 094 490
Dissertationsgebiet li Studienblatt:	Molekulare Biologie
Betreut von:	Ao. Univ. Prof. Dr. Michael F. Jantsch

I want to dedicate my thesis to my parents, who constantly support me with their love.

Table of Contents

Abstract

Alu elements are conserved, ~300 nucleotide long repeats that belong to the SINE family of retrotransposons found abundantly in primate genomes. Alu elements are enriched in gene-rich regions, where they are located within non coding segments of transcripts, such as introns and untranslated regions. SINEs can have a dramatic impact on the transcriptome by several means such as repressing global transcription by impairing polymerase II activity, affecting splicing or by triggering Staufen mediated decay. In addition, we and also other groups showed that SINE elements in inverted orientation (*i*SINE) could control gene expression. As underlying mechanisms, RNA-editing, sequence specific degradation or translational control has been discussed.

In this project we determined the impact of *i*SINE located in 3' UTRs on gene expression. We demonstrated that the presence of *i*SINE in 3' UTRs can repress gene expression and reduce mRNA levels. Using knock-out cells we showed that the reduced gene expression does not rely on known double-stranded RNA binding proteins such as DICER, Staufen, PKR, or ADARs. In addition *i*SINE mediated gene expression is sequence independent.

The reduced RNA levels measured for RNAs containing *i*SINE is not correlated with an increase in mRNA decay. In addition, the length or position of poly-adenylation, which could alter translational efficiency or mRNA stability, has not been affected. Chromatin immunoprecipitation data indicate that the distribution of Pol-II is different in *i*SINE-containing mRNA and control mRNA. It therefore appears that *i*SINEs can modulate the rate of mRNA transcription. Thus, besides the previously reported nuclear retention and translational repression induced by *i*SINE, transcription seems regulated by a third Alu-triggered mechanism.

Abstrakt

Alu Elemente sind konservierte, ~300 Nukleotide lange Short Interspersed Elements (SINE), welche die häufigsten Retrotransposons in Primatengenomen darstellen. Alu Elemente finden sich gehäuft in genreichen Regionen. Dort sind sie in nicht-kodierenden Bereichen von Transkripten, wie etwa Introns und nicht-translatierten Bereichen lokalisiert. SINEs können einen dramatischen Einfluss auf das Transkriptom haben. Etwa indem sie durch Inhibition der Polymerase II global die Transkription reprimieren, mRNA Spleißen oder *Staufen mediated decay* beeinflussen. Außerdem wurde in unserem, und anderen Laboren gezeigt, dass SINE Elemente in tandemartiger, invertierter Orientierung (*i*SINE) die Genexpression kontrollieren können. Als zugrundeliegende Mechanismen wurde RNA Editierung, sequenzspezifische Degradierung und translationale Kontrolle diskutiert.

In diesem Projekt haben wir den Einfluss von in 3'UTRs liegenden iSINEs auf die Genexpression bestimmt. Wir haben gezeigt, dass das Vorhandensein von iSINEs in 3'UTRs Genexpression reduzieren kann und zu verminderten RNA Mengen führt. Unter Verwendung von *Knockout* Zellen haben wir demonstriert, dass die reduzierte Genexpression nicht durch RNA-Doppelstrang bindende Proteine, wie etwa DICER, STAUFEN, PKR oder ADAR bedingt ist. *i*SINE-vermittelte Reduktion von Genexpression ist zudem sequenzunabhängig.

Die für *i*SINE Transkripte beobachtete Reduktion von RNA Levels, ist nicht auf einem Anstieg des mRNA Abbaus zurückzuführen. Zudem ist die Länge oder Position der Polyadenylierung, welche Translation oder mRNA Stabilitiät beeinflussen könnte, nicht beeinträchtigt. Unsere Daten der Chromatinimmunopräzipitation (ChIP) zeigen, dass sich die Polymerase II Verteilung für *i*SINE beinhaltende RNAs von Kontroll-RNAs unterscheidet. Offenbar können *i*SINEs die Transkriptionsrate beeinflussen. Zusätzlich zu den bereits dokumentierten Mechanismen der „*nuclear retention*" und der Beeinflussung der Translation, könnte also Transkription durch einen dritten Alu-vermittelten Mechanismus beeinflusst sein.

Introduction

Introduction

Transposable elements

The vast majority of genomes of higher eukaryotes are occupied by transposable elements (TE). In humans, nearly half of the genome is derived from mobile genetic elements (Figure 1a) (Britten and Kohne 1968; Lander, Consortium et al. 2001). TEs are separated into two major classes; DNA transposons and retrotransposons. DNA transposons, which constitute ~3% of the human genome, employ a cut and paste mechanism to transfer their genetic information. These elements seem to be not currently active in the human genome (Pace and Feschotte 2007). In contrast, retrotransposons can amplify themselves by a copy and paste mechanism, producing an RNA intermediate that is reverse transcribed and inserted at new genomic locations.

Retrotransposable elements are divided in two subgroups based on the presence or absence of long terminal repeats (LTR). LTR retrotransposons occupy 8% of the human genome and their current activity is very low in human. In contrast, the majority of transposable elements in the human genome are resulted from non-LTR retrotransposons, which are presently the most active elements (Figure 1a). Non-LTR retrotransposons are divided in three subgroups including Alu elements of the family of short interspersed elements (SINEs), long interspersed element (LINE-1, L1) and SVA elements which consist of a SINE region, a variable number tandem repeat (VNTR) region and an Alu-like region. These elements collectively made approximately one-third of the human genome (Figure 1a). Alu elements are the most abundant retrotransposons in terms of number. There are more than one million Alu copies in the human genome as a result of their activity over the past million years (Belancio, Hedges et al. 2008; Cordaux and Batzer 2009).

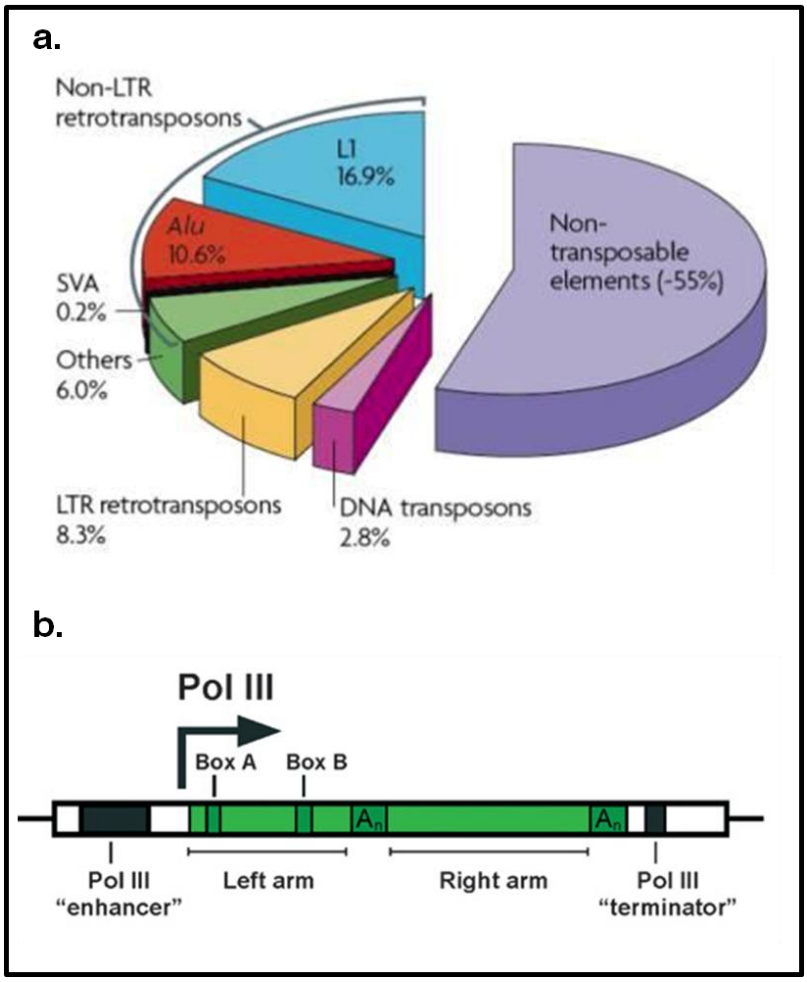

Figure 1. The distribution of transposable elements in the human genome. a) Transposable elements in the human genome occupy around 45% of the human genome and the majority of them are non-long terminal repeat (LTR) retrotransposons such as LINE-1 (L1), Alu and SVA elements (Cordaux and Batzer 2009). **b)** Alu elements consist of two monomers linked by a poly-A rich region. It has two pol-III promoters (box A and box B) located in the first monomer and each monomer ends with and poly A tail (Figure adapted from (Hasler 2007).

Impact of transposable elements on the genome shape and its evolution

Transposable elements had a huge impact on the eukaryote genome evolution. Through transposition and recombination, TEs have changed the structure and

subsequently functions of genomes during evolution. As an example, insertion of TEs in the coding region could disrupt or alter the function of proteins (Cordaux and Batzer 2009). In addition, in plants it has been shown that TEs could mobilize large DNA fragments in the genome and lead to genome rearrangement and translocation (Wicker, Zimmermann et al. 2005). TE could also make a high variation in the genome size. The DNA content of haploid genome, which is called C value, highly varies among highly related organisms with comparable complexity (Rosbash, Ford et al. 1974). This variation is explained by the differences in the abundance of TEs, mainly retrotransposable elements, in these organisms (Fedoroff 2012).

TEs are vastly accumulated in the eukaryotic genomes in comparison with prokaryotes and this accumulation is believed to be the result of the complex epigenetic silencing machinery present in the eukaryotes. The silencing mechanisms can reduce the harmful effect of transposons, allowing them to be tolerated in the genome and drive their evolution. Indeed, it has been proposed that a balance between epigenetic and transposition has led to the evolution of the modern eukaryotic genomes (Burns and Boeke 2012; Fedoroff 2012). In addition, TEs are still active in the genome and Alu is the most proliferative ones. It is estimated, that a de novo Alu insertion occurs per 20 human births (Burns and Boeke 2012). Therefore, due to the genomic tolerance to transposable elements and their current activity, these elements are dynamically involved in genome evolution and inter-species variation.

Structure of SINEs

Alu elements, are the most abundant primate specific SINEs and occupy 10% of the human genome (Cordaux and Batzer 2009). The typical length of an Alu element is ~300 bp and it presents in a dimeric structure. They are formed by the fusion of two monomers derived from the *7SL* RNA gene (Kriegs, Churakov et al. 2007).The monomers are held together by an A-rich linker region (Figure 1b). They contain two internal RNA polymerase III (RNAPIII) promoter (A and B boxes) which are CG or CpG rich. These elements end with an oligo(dA)-rich tail of variable length for each monomer

(Batzer and Deininger 2002). Since *Alu* elements do not have RNAPIII termination signals, their transcription extends into the downstream flanking sequence (typically a run of four or more consecutive thymines) until a terminator is found (Matthew S. Comeaux 2009).

Based on the time of retrotranspositon, Alu elements are classified in three subfamilies. The oldest group is called the Alu J subfamily, the intermediate age is known as Alu S elements (Jurka and Milosavljevic 1991; Jurka and Smith 1988), while the youngest subfamily is called the Alu Y elements (Batzer et al. 1996).

B1 elements are the Alu equivalents in the mouse genome. These elements also belong to the SINE family and occupy 2.7% of the mouse genome. Same as Alu elements, they are derived from the *7SL* RNA gene, but they have a monomeric structure and are only around 130bp long (Berger and Strub 2011).

Mechanism of transposition

As mentioned, SINEs are amplified in the genome by retrotransposition. SINEs do not encode any functional protein for their autonomous mobilization, whereas they depend on LINE-encoded proteins with reverse transcriptase and endonuclease activities (Moran, Holmes et al. 1996). The LINE-encoded reverse transcriptase recognizes the LINE RNA and then it cleaves the target genomic DNA at the consensus sequence TTAAAA (Feng, Moran et al. 1996). The generated 3'OH is used to prime the reverse transcription of the LINE RNA and subsequently the second strand of the target DNA is cleaved and used to prime second-strand synthesis (Luan, Korman et al. 1993). This mechanism, called target-primed reverse transcription (TPRT), is also used by Alu elements for their transposition (Dewannieux, Esnault et al. 2003).

Effect of SINEs on mRNA life

During evolution, SINEs have accumulated in the genomes of higher eukaryotes. For example in humans, 10% of the genome is occupied by Alu elements. The vast majority of SINEs were inactivated by mutation and only a minor fraction is still transcribed. Since SINEs are preferentially located in gene-rich regions, even though most of them are transcriptionaly inactive, they are able to interfere in different steps of the mRNA life. SINEs in noncoding regions of mRNAs such as the 5'UTR, the 3'UTR or introns, which were transcribed along the mRNA by pol-II, potentially interfere with mRNA processing, translation or mRNA degradation. In addition, SINEs located in the proximity of genes, could interfere with the regulation of mRNA transcription. In the following part we will discuss about the effects of SINEs on the mRNA life, with particular focus on the Alu elements.

Transcriptional effect

As mentioned, SINEs could interfere with different steps of an mRNA life-cycle. Regulation of transcription is the first step, which could be influenced by SINEs. It was reported that SINEs in the proximity of genes could act as a transcriptional biding site, enhancer or silencer (Ichiyanagi 2013). In addition, the pol-III transcribed SINEs, which is also called free SINEs, could block pol-II transcription by their direct interaction with the pol-II enzyme (Mariner, Walters et al. 2008).

Epigenetic regulation

Alu elements have 25 CpG sites inside and outside the pol-III promoters, which is the essential region for methylation. Whereas Alu elements are highly methylated in somatic cells, they are hypomethylated in spermatogenic cells, oocytes, and embryos in early development. Therefore, SINE-derived RNAs are very low or undetectable in somatic comparing to spermatogenic cells, oocytes, and embryos. (Hellmannblumberg,

Hintz et al. 1993; Xie, Wang et al. 2009; Ichiyanagi 2013). In addition, the mouse B1 methylation, which is high in somatic cells and low in germ cells, is highly correlated with their expression (Ichiyanagi, Li et al. 2011).

Methylation not only can inactivate pol-III dependent transcription of SINE RNAs, it can also interfere in the expression of nearby genes. Generally, the density of B1 elements is very low in pol-II promoter regions. However it has been shown that B1 elements are at a higher density in testis-specific promoters compared with other promoters. Since SINEs are highly methylated in somatic cells, one could speculate that B1 elements are involved in the suppression of testis-specific genes in somatic cells (Figure 2) (Ichiyanagi, Li et al. 2011; Ichiyanagi 2013).

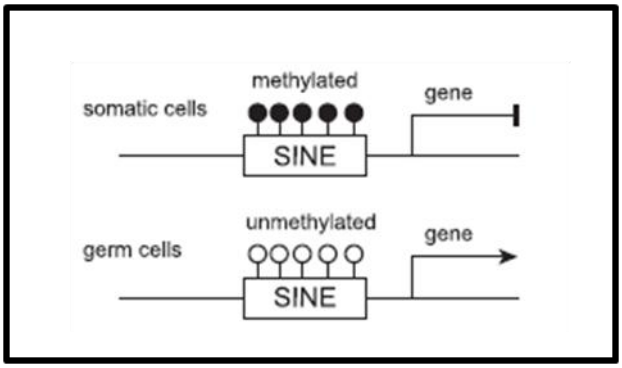

Figure 2. SINE methylation-mediated gene regulation. SINEs located near promoters could be methylated which regulates the expression of nearby genes in a tissue specific manner (Figure adapted from (Ichiyanagi 2013))

*i*SINEs impair pol-II transcription

Generally, SINE RNAs are accumulated at very low levels in cells, but upon heat shock and other cellular stresses SINE RNA levels are increased (Liu, Chu et al. 1995). Rapid increase in the expression proposes that they could have a function under stress conditions. It has been shown that SINE RNA can suppress the expression of a subset of genes during heat shock. The molecular mechanism of this repression has not been completely clarified, but it has been shown that Alu and B2 RNAs (another SINE in mouse) directly and tightly bind to polymerase II on the promoter and can thereby

repress transcription (Figure 3) (Allen, Von Kaenel et al. 2004; Mariner, Walters et al. 2008). Binding of the noncoding RNA to the transcriptional machinery, blocks the completion of the transcriptional closed complex formation at the promoter and leads to RNA pol-II stalling (Allen, Von Kaenel et al. 2004; Mariner, Walters et al. 2008; Yakovchuk, Goodrich et al. 2009). In a recent study, it has been shown that the proper interaction of pol-II with the DNA promoter is blocked through B2 RNA binding to the DNA cleft or the active site of pol-II. In the presence of the SINE RNA, RNA pol-II does not have normal contact with DNA, however, it is still kept at the promoter through its interaction with general transcription factors. As soon as, the SINE RNA has been removed by RNase I, the interaction of pol-II with the core promoter is restored and the transcriptional machinery is activated (Ponicsan, Houel et al. 2013). These observations propose a new mechanism of gene specific transcriptional regulation by SINE elements.

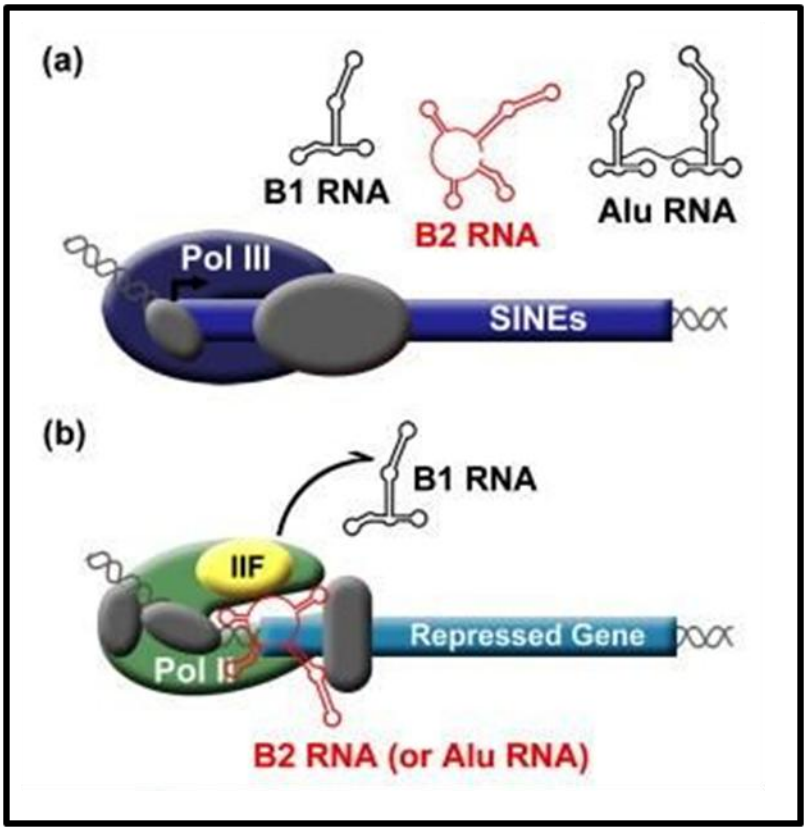

Figure 3. SINE RNAs regulate the gene expression in response to heart shock. a) Upon stress, the expression of SINE RNAs is increased **b)** and subsequently, Alu and B2 RNAs block

gene expression through binding to pol-II at the promoter (Figure adapted from (Ponicsan, Kugel et al. 2010)).

mRNA processing effect

A to I RNA Editing in Alu elements

RNA editing describes the alteration of RNA by nucleotide modifications, insertions or deletions. The most common editing in metazoa is the deamination of adenosines to inosines. Inosines are recognized as guanosines by the translational and splicing machineries. Therefore, editing can affect splicing or lead to a codon exchange (Sommer, Kohler et al. 1991; Rueter, Dawson et al. 1999). This modification is catalyzed by the Adenosine Deaminases Acting on RNA (ADAR) family of proteins. ADARs are double-stranded RNA binding proteins and essential enzymes in mammals. There are two active ADAR enzymes in the human genome, ADAR1 and ADAR2 (Hundley and Bass 2010).

More than 90% of editing in the human transcriptome occurs in the Alu elements that are embedded and cotranscribed with the mRNA (Levanon, Eisenberg et al. 2004). Notably, the Alu elements are not distributed randomly in the genome, they are more frequently located in gene rich regions (Versteeg, van Schaik et al. 2003). It has also been shown that editing is favored more when two Alu elements in the genes are located in inverted orientation and their distance is shorter than 2 kb. This observation suggests that these Alu elements could make double-stranded structures which is the substrate for ADARs (Figure 4) (Athanasiadis, Rich et al. 2004).

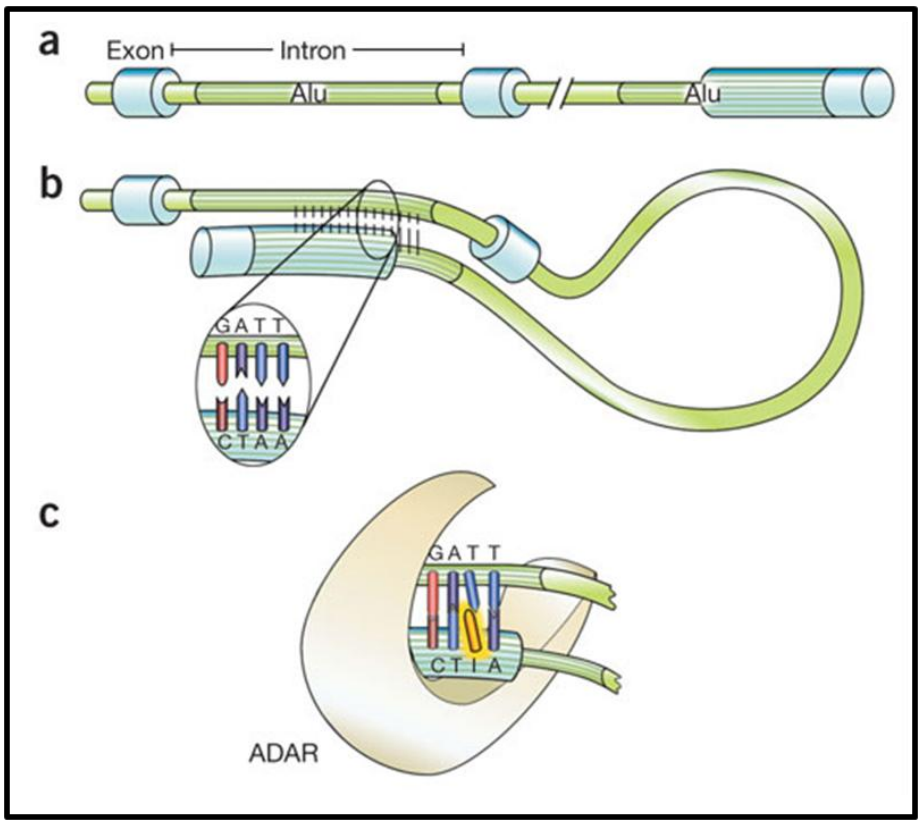

Figure 4. Editing of inverted Alu elements. a) The gene contains two Alus in inverted orientation. **b)** When the distance between two Alus is less than 2kb, they could form a double-stranded RNA structure. **c)** Double-stranded RNA is recognized by ADAR family and edited in several positions (Figure adapted from (Levanon, Eisenberg et al. 2004)).

The role of editing in Alu elements is not well understood. Recently, it has been shown that adjacent inverted Alus in the genes could function as inducer of editing by recruiting of ADAR (Daniel, Silberberg et al. 2014). Alus could form a double stranded structure and prepare a suitable binding site for ADAR enzymes. Therefore they may recruit editing enzymes and subsequently induce the editing at a single site located several hundred nucleotides from the Alu elements in the surrounding transcript (Figure 5). Since Alu elements are primate specific and editing is higher in humans when compared with non-primates, this observation suggested that Alu elements could contribute to primate and human genome evolution trough induction and regulation of site selective RNA editing (Daniel, Silberberg et al. 2014).

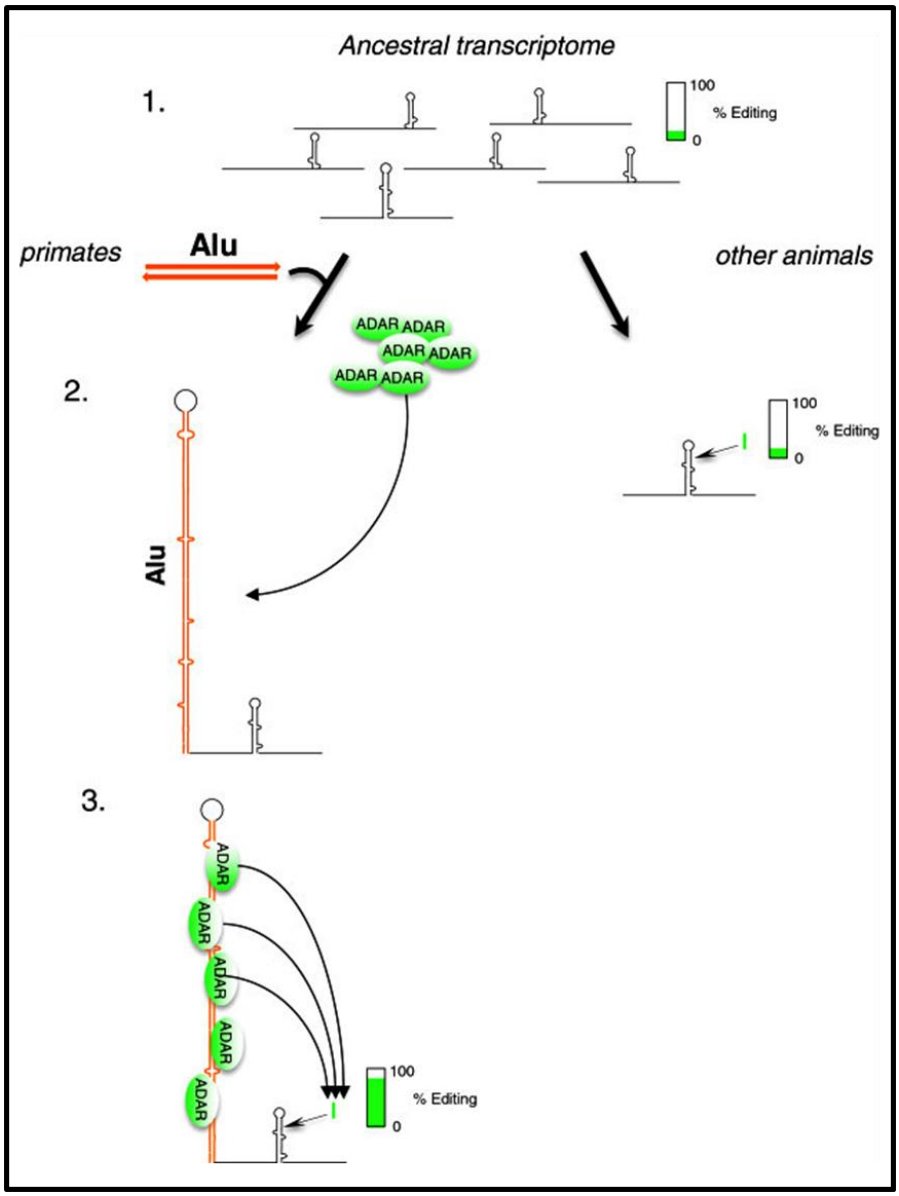

Figure 5. Induction of adenosine to inosine by inverted Alu elements in primates. 1) The editing levels in most metazoa are low. **2)** Alu elements inserted in primate genomes and Alus which have an inverted partner in the nearby, could make double stranded structure and recruit ADAR enzymes. **3)** The accumulation of ADARs may lead to higher levels of editing at a single site near the Alus insertion. (Figure adapted from (Daniel et al, 2014))

Alternative splicing

Alternative splicing (AS) is a mechanism which produces several variants of proteins from a single gene and it is the major source of protein diversity in the genome. Alternative splicing happens in 95% of human genes (Modrek and Lee 2002; Pan, Shai et al. 2008). In higher eukaryotes, AS is more abundant than in lower eukaryotes. In addition, the number of genes and exons that undergo AS, is larger in vertebrates than in invertebrates, which shows the importance of AS in evolution (Keren, Lev-Maor et al. 2010).

It has been shown that Alu elements can affect mRNA splicing mainly through a mechanism called Alu exonization. Exonization describes the inclusion of intronic sequences into the mRNA. It has been shown that 5% of alternatively spliced exons were derived from Alu elements (Sorek, Ast et al. 2002). There are two main mechanisms which lead to the creation of splice sites and the exonization of Alu elements; mutations, which happened during the primate evolution, and RNA editing. There are 9 potential 5' splice sites and 14 potential 3' splice sites in the consensus Alu sequence. Therefore, a few mutations in the 3' or 5' splice sites are enough to create a new exon (Figure 6a). Most (19 out of 23) of the potential splice sites are located in the minus strand. Therefore, Alu elements inserted in antisense orientation in genes have more chances to be exonized. It also has been shown that 85% of Alu containing exons are derived from antisense Alu elements and the right arm of Alu is exonized in most of the cases (Sorek, Ast et al. 2002; Sela, Mersch et al. 2007; Keren, Lev-Maor et al. 2010).

Another mechanism that can lead to the exonization of Alu elements is RNA editing. As mentioned before, two close Alu elements in an intronic region could form a double stranded structure and become edited. Editing could then create a new splice site (figure 6b). For example, nuclear prelamin A recognition factor (NARF) has an Alu-exon which is regulated by editing. RNA editing can create a 3' splice site and also eliminates a premature stop codon within the Alu-exon (Lev-Maor, Sorek et al. 2007).

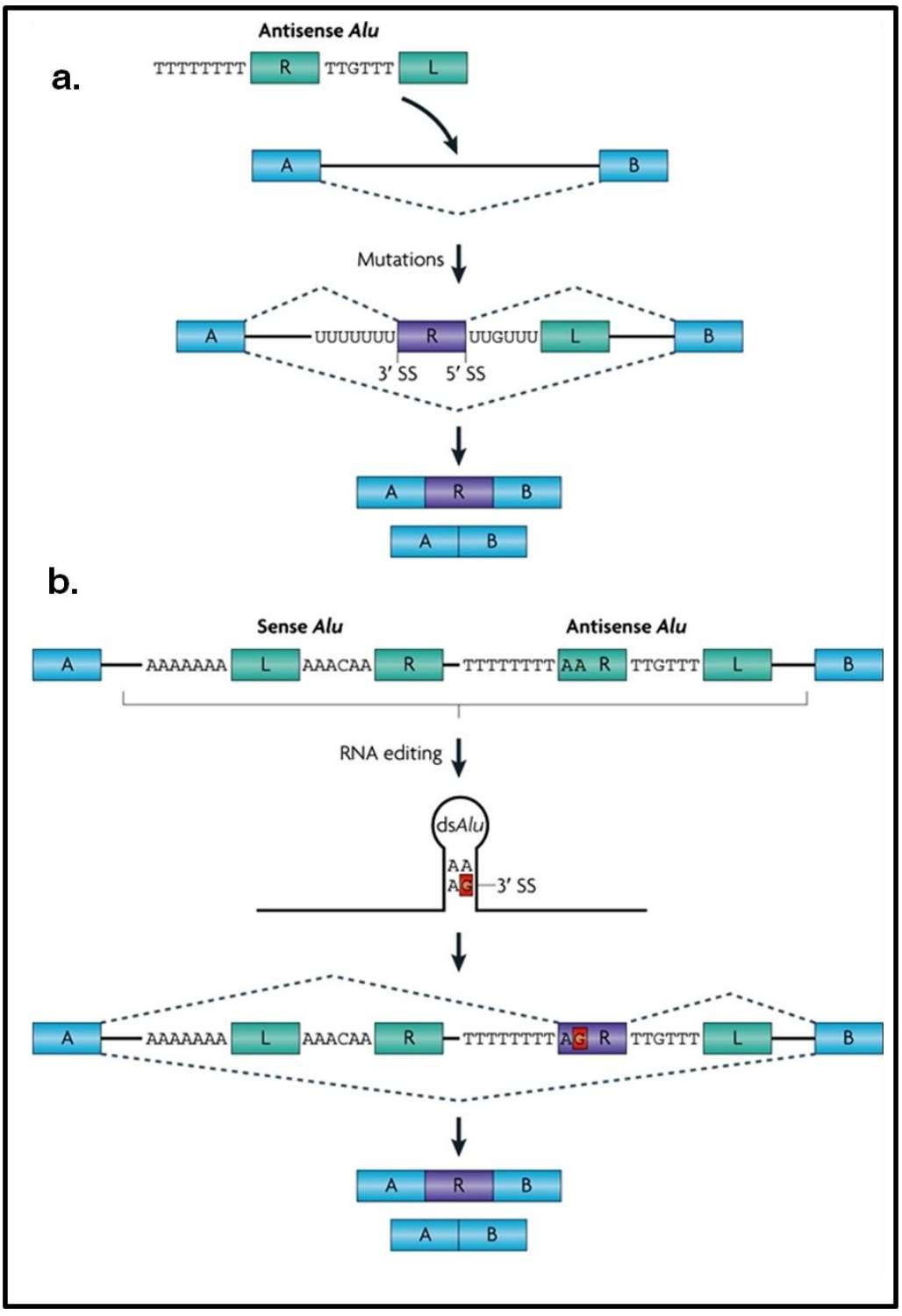

Figure 6. Two mechanisms of Alu exonizaion. a) Mutations in Alu elements inserted in introns could introduce new 3' and 5' splice sites and lead to Alu exonization. In most of the cases exonization occurs in the right arm of Alu elements which are inserted in antisence orientation. **b)** Editing could also lead to Alu exonization. Inverted Alu elements, which are

located in intron in close distances, could make double stranded structure. Then, they can be edited. New splice sites could be created by editing and subsequently Alu elements can be exonized (Figure adapted from (Keren, Lev-Maor et al. 2010)). (Blue box: exon, green box: Alu arm, purple box: spliced alu arm, dashed line: alternative splicing).

Translational effect

Free Alu and translational inhibition

Free Alu RNAs, which are pol-III transcripts, can affect mRNA translation by different means. Alu RNAs could form secondary structure similar to their ancestor, 7SL RNA. Consequently they are able to bind to the SRP 9/14 proteins of the signal recognition particle (SRP). The binding has been confirmed *in vivo* and *in vitro*. In primates, the expression of SRP9/14 is 20-fold more than other SRP proteins, therefore, there is a large pool of free SRP9/14 available for binding to Alu RNA. It has been shown that Alu RNA in complex with SRP9/14 could inhibit global translation by interfering with translation initiation. The exact molecular mechanism of this inhibition is not clear yet (Hasler and Strub 2006).

It has been shown that Alu RNA is able to modulate protein kinase R (PKR) activity. PKR is a double-stranded RNA binding protein that inhibits protein synthesis via eIF2α phosphorylation and becomes activated by double stranded RNA, cytokine, growth factors and cell stress signals. PKR is known as an important protein in antiviral defenses (Garcia, Meurs et al. 2007). Alu RNA at low concentration can activate PKR. However, high concentration of Alu RNA can inhibit PKR activity and therefore increase protein synthesis (Chu, Ballard et al. 1998; Williams 1999).

Embedded Alu in mRNAs

As mentioned before, Alu elements are not distributed randomly in the genome, they are more concentrated in gene rich regions such as 5'UTRs, 3'UTRs, and introns (Versteeg, van Schaik et al. 2003). Therefore, these elements could be transcribed by

pol-II along with the surrounding mRNA. Since 3'UTRs and 5'UTRs are important elements for mRNA translation, Alus in these regions could modulate translation. There are several studies showing that Alu elements in the UTRs regulate translational initiation (Hasler and Strub 2007). One example is the Alu element in the 5'UTR of BRCA1 gene. BRCA1 is a DNA repair protein which is less expressed in ovarian and breast cancer (Magdinier, Ribieras et al. 1998; Zheng, Luo et al. 2000). There are two forms of the BRCA1 mRNA, which have different patterns of expression. In normal mammary tissue, the mRNA with a short 5'UTR is expressed, whereas in breast cancer tissue the mRNA with a long 5'UTR is expressed (Xu, Brown et al. 1995; Sobczak and Krzyzosiak 2002). The translational efficiency of the mRNA with the long 5'UTR is 10 times lower than the short one. This translational deficiency is due to an Alu element in the 5'UTR that is located upstream of the start codon. The Alu element forms a stable secondary structure in the 5'UTR and inhibits translational initiation and subsequently reduces protein levels (Sobczak and Krzyzosiak 2002).

Another example includes the Alu elements in the 3'UTR that interfere with translation *via* the interaction with PKR. It has been shown that during mitosis, inverted Alu elements in the 3'UTR of the mRNA, which are able to form a double-stranded RNA structure, bind to PKR. In addition, it was shown that PKR is phosphorylated and subsequently activated in the early phase of mitosis. Therefore, it was proposed that in the early phase of mitosis, PKR is activated by binding to an inverted Alu RNA in the 3'UTR of mRNA, which are more accessible because of the disintegration of the nuclear envelope (Kim, Lee et al. 2014). Activated PKR inhibits global translation during mitosis via eIf2α phosphorylation and regulates mitotic progression by phosphorylation of c-Jun N-terminal kinase (JNK) (Figure 7).

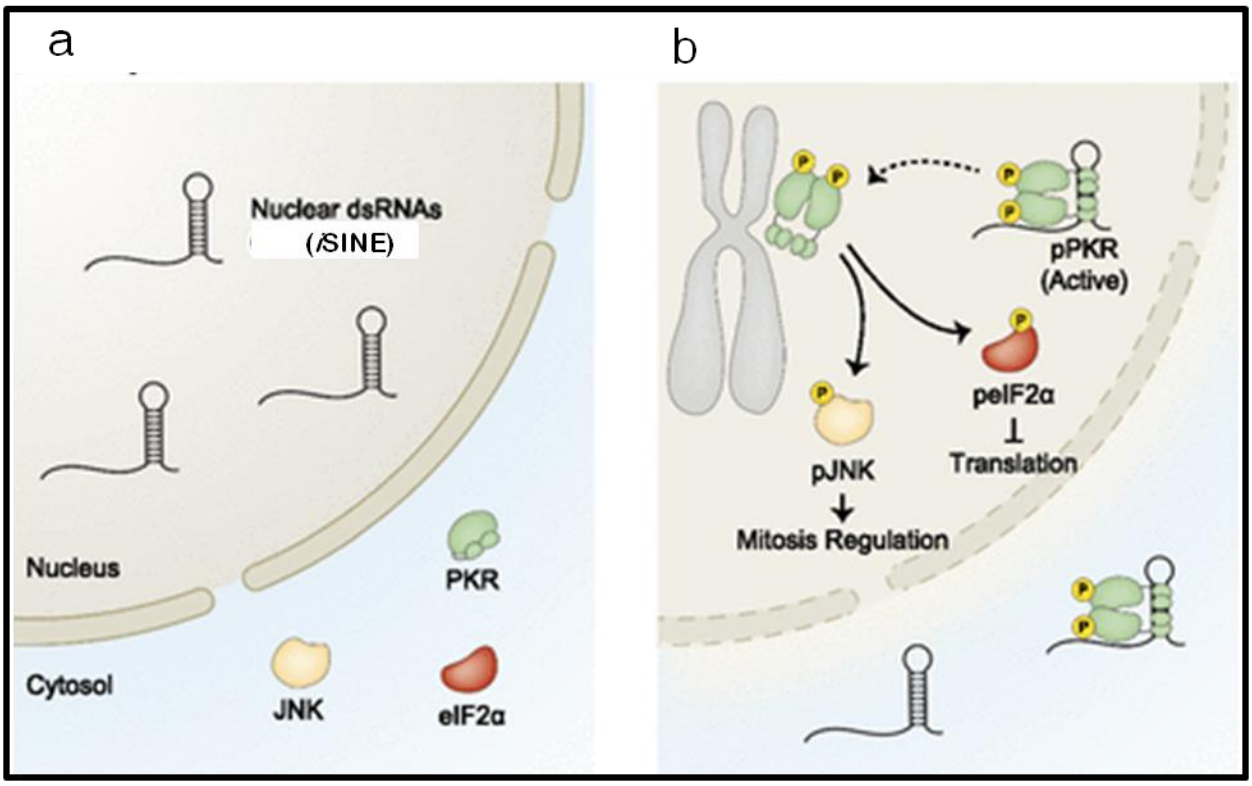

Figure 7. Proposed model for cell cycle regulation by PKR. a) In interphase, inverted SINE (*i*SINE) containing mRNAs are less accessible and retained in the nucleus, therefore, could not activate PKR. **b)** In mitosis, *i*SINE RNAs can interact with PKR and activate it. Activation of PKR leads to phosphorylation of eIF2α and JNK to regulate translation and cell cycle progression. (Taken from (Kim et al., 2014))

mRNA degradation

SINEs and Staufen1 interaction

Inverted SINEs in the 3'UTR of mRNAs are introduced as one of the main classes of endogenous Staufen1 (Stau1) targets. These elements are able to bind several Staufen1 proteins simultaneously and this multimerization could stabilize Stau1 binding (Ricci, Kucukural et al. 2014). In addition, it has been shown that intermolecular base paired SINEs, which are formed between two mRNAs or an mRNA and a long non coding RNA, can bind to Stau1 (Gong and Maquat 2011; Gong, Tang et al. 2013).

Staufen1 is a highly conserved double stranded RNA binding protein which is express in most tissues. In mammals, it has four double stranded RNA binding domains and only two of them are necessary for RNA binding (dsRBD3 and dsRBD4) (Park and Maquat 2013). Stau1 proteins are involved in several post-transcriptional processes such as mRNA transport to neuronal dendrites, translational regulation via ribosome interaction and translation-dependent mRNA degradation known as Staufen1 mediated mRNA decay (SMD) (Kohrmann, Luo et al. 1999; Dugre-Brisson, Elvira et al. 2005; Kim, Furic et al. 2005).

Staufen1 mediated mRNA decay (SMD) is an mRNA degradation mechanism triggered by binding of Stau1 to the 3'UTR of a target mRNA. A well characterized human Stau1 binding site (SBS) forms a perfect 19 base pair long double-stranded structure with a 100 nucleotide apex in the 3'UTR of ADP-ribosylation factor 1 (ARF1) (Kim, Furic et al. 2007). During SMD, Stau1 bind to the SBS and activates UPF1 by phosphorylation, then UPF1 triggers translational repression and mRNA degradation. Recently it has been shown that SBSs can also be formed by intermolecular base pairing between Alus embedded in two different RNA molecules. Single Alus in one mRNA can base pair with an Alu in a lncRNA or another mRNA. Subsequently, the intermolecular SBS can be recognized by Stau1 and trigger the SMD pathway (figure 8) (Gong and Maquat 2011; Gong, Tang et al. 2013).

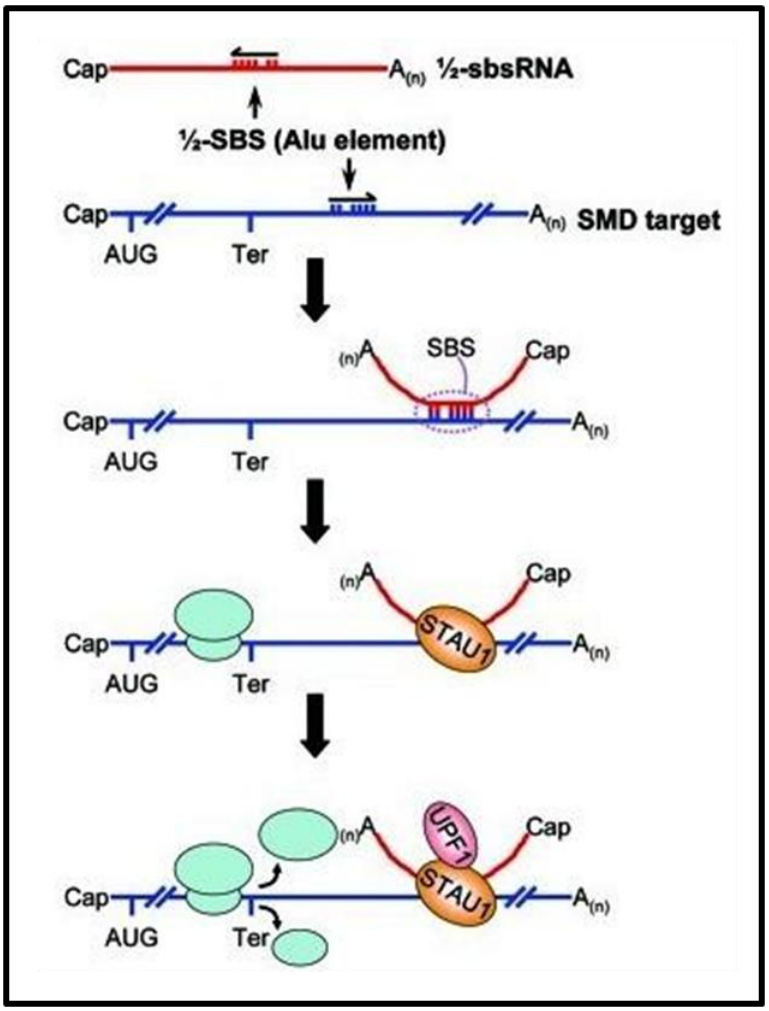

Figure 8. The proposed model for SMD triggered by intermolecular base-pairing. One Alu in a lncRNA (1/2 SBS RNA) is base-paired with another Alu in the mRNA. Stau1 recognizes double-stranded RNA and binds to it. Then Stau1 triggers UPF1 activation and subsequent translational inhibition and mRNA degradation (Figure adapted from (Gong and Maquat 2011)).

Stau1 is not only involved in inverted SINEs (*i*SINE) containing mRNA degradation, it also could regulate nuclear transport of target mRNAs. In two studies, It has been shown that Stau1 increases the nucleocytoplasmic export of the mRNA by binding to *i*SINE 3'UTR (Elbarbary, Li et al. 2013; Ricci, Kucukural et al. 2014). In addition, Elbarbary et al., also showed that Stau1 enhances mRNA translation by binding to *i*SINE containing reporter mRNAs. They proposed that PKR binds to the cytoplasmic

*i*SINE containing mRNA and subsequently it will be activated and it will repress mRNA translation. In the presence of Stau1, Stau1 binds to *i*SINE and by precluding PKR binding, it enhances the mRNA translation (Figure 9)(Elbarbary, Li et al. 2013). In contrast, in a transcriptome-wide analysis, it has been shown that ribosome association of *i*SINE containing mRNAs did not change upon Stau1 up regulation. Therefore, the Stau1 effect on *i*SINE translation still requires further investigation (Ricci, Kucukural et al. 2014).

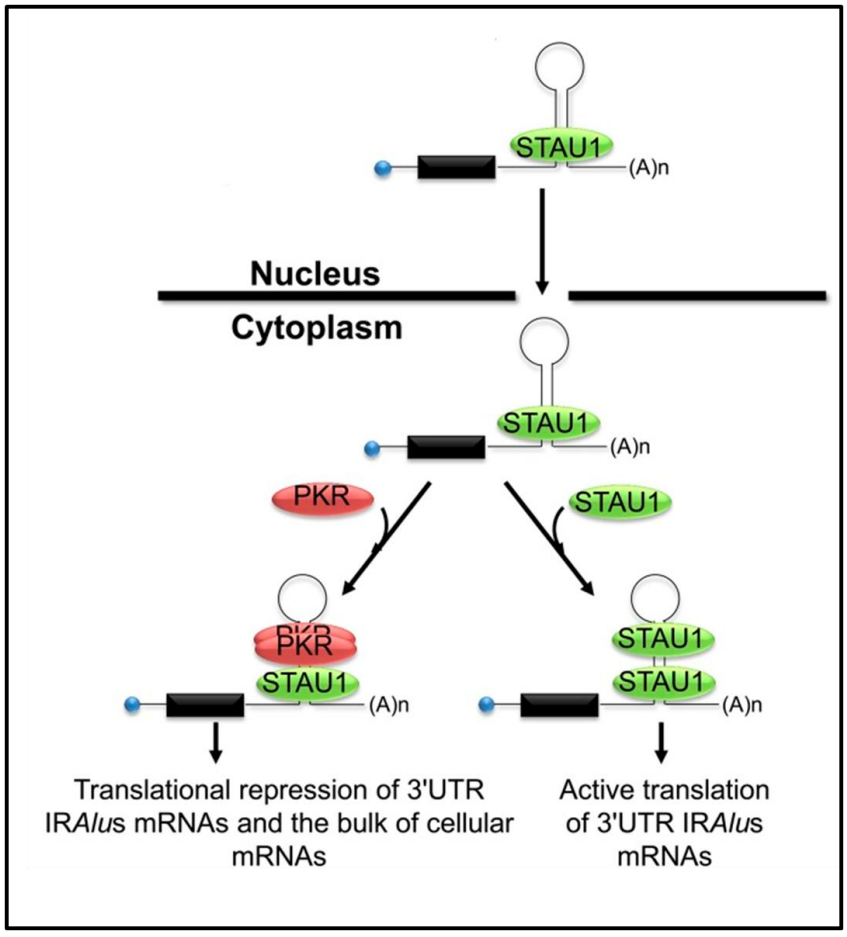

Figure 9. Proposed model for how Stau1 enhanced *i*SINE containing mRNA translation. Stau1 can increase *i*SINE mRNA nucleo-cytoplasmic export. In addition, Stau1 competes with PKR for binding to *i*SINEs. PKR binding and activation could lead to translational inhibition. Therefore, Stau1 by precluding PKR binding could keep the translation active (Figure adapted from (Elbarbary, Li et al. 2013)).

iSINE regulate the gene expression

As I mentioned before, SINEs can influence the mRNA faith by regulating the mRNA transcription, transport, translation or degradation. In my PhD studies, I focused on the effect of inverted SINE (*i*SINE), which are located in the mRNA 3'UTR, on the gene expression.

In our lab and also in two other labs, we have shown that *i*SINE in the 3' UTR could reduce the gene expression (Chen, DeCerbo et al. 2008; Capshew, Dusenbury et al. 2012). In addition, our bioinformatics analysis shows that there are 235 protein coding genes in the human genome with *i*SINE in the 3'UTR. Therefore, *i*SINEs could act as regulatory elements that modulate the gene expression in the genome. There are two major proposed mechanisms for *i*SINE mediated gene repression that we discuss in the following part.

Nuclear retention

Chen et al., have shown that *i*SINEs in the 3'UTR of mRNAs strongly repress gene expression. Further, the *i*SINE are highly edited and co localized with $p54^{nrb}$ in the nucleus (Chen, DeCerbo et al. 2008). $p54^{nrb}$ is an RNA-binding protein which is located in the nucleus and is involved in several nuclear process such as transcriptional termination and mRNA splicing (Kameoka, Duque et al. 2004; Kaneko, Rozenblatt-Rosen et al. 2007). It has been shown that $p54^{nrb}$ has a high preference for inosine-containing RNAs (Zhang and Carmichael 2001). It was therefore proposed that *i*SINEs in the mRNA 3'UTR are highly edited and prepare a binding platform for $p54^{nrb}$, which prevents their export to the cytoplasm and therefore, suppressing gene expression (Figure 10) (Chen, DeCerbo et al. 2008).

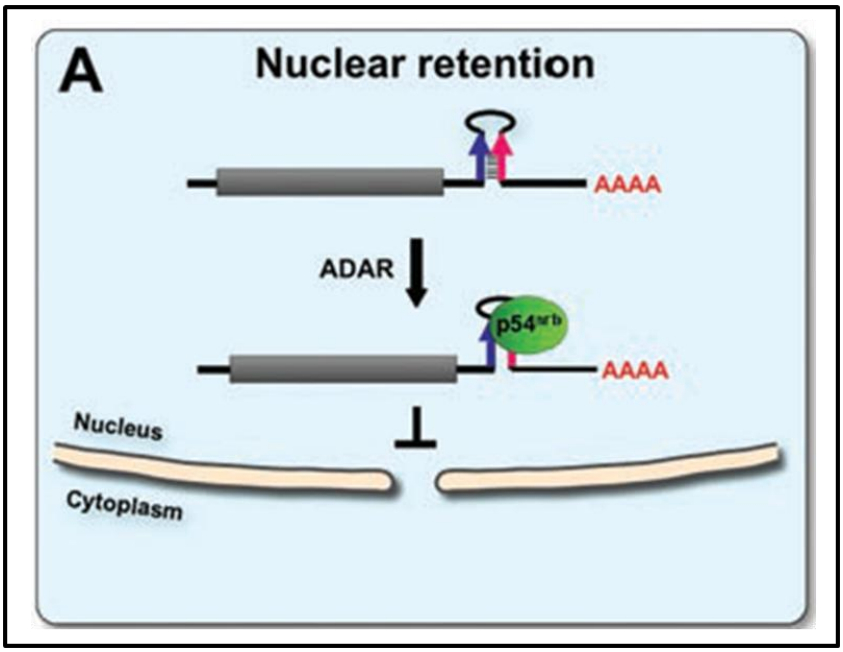

Figure 10. The proposed model for *i*SINE mediated nuclear retention. Highly edited *i*SINE in 3'UTR of mRNA are in association with $p54^{nrb}$, which prevents their export to the cytoplasm and leads to nuclear retention (Figure adapted from (Chen and Carmichael 2008)).

Alteration of Translational efficiency

Capshew et al., have also shown that not all but some of the *i*SINE containing genes show altered gene expression. They reported that *i*SINE containing mRNA are transported to the cytoplasm in equal levels as mRNA lacking *i*SINEs. In addition, the presence of double-stranded structure in the 3'UTR reduces the translational efficiency and the observed effect is independent of editing. Therefore, it was proposed that secondary structures in the 3'UTR could function as translational control elements (Capshew, Dusenbury et al. 2012).

Specific aims of this dissertation: Deciphering the molecular mechanism(s) of gene silencing by *i*SINEs

Despite that two phenomena have been suggested to regulate *i*SINE-mediated gene repression, the underlying molecular mechanisms of this repression are still unclear. Preliminary data from our lab showed that *i*SINE could also reduce the gene expression independent of ADARs. Therefore, it is possible that *i*SINEs modulate gene expression through a different mechanism, based on cell type, cell phase and other unknown factors. Therefore, during my PhD studies I was involved in deciphering the molecular mechanism(s) of gene silencing by *i*SINEs.

Since not all *i*SINE containing 3'UTRs reduce gene expression to the same level, in the first step we aimed to identify important features in the 3'UTRs that can modulate the extent of repression, such as, SINEs configuration, SINEs location and the extent and length of the double-strandedness.

Moreover, we investigated the influence of RNA-binding proteins on *i*SINE mediated gene repression. The proteins proposed to interact with SINEs, such as Staufen1, PKR or ADARs were tested for their impact on the observed phenomenon.

Furthermore, we evaluated whether transcription, RNA-localization, translation, or RNA stability was affected by the presence of *i*SINEs. This unbiased approach should allow gaining insight on the potential cellular process by which *i*SINEs can regulate gene expression.

Lastly, we explored the distribution and expression of *i*SINE containing RNAs at a genome-wide level (in collaboration with the research group of Ivo Hofacker). From this approach, it is expected that regulatory elements such as *i*SINEs would not be tolerated in all regions and types of genes. Therefore, the distribution of *i*SINEs and the global expression patterns of genes containing these elements should give insight on the selection of these elements and their potential impact on gene regulation.

Material and Methods

Material and Methods

Construction of renilla and firefly reporter constructs

The 3' UTRs of *Inadl* or *Nicn1* were cloned downstream of the open reading frame of the renilla luciferase in pHRL-TK (Promega, Madison, WI) in the XbaI site. The firefly luciferase expressing plasmid pGL3 was used as a reference plasmid (Promega, Madison, WI). Alternatively, the 3'UTRs of interest were cloned downstream of firefly luciferase between XbaI and NotI into pmirGLO that expresses renilla luciferase from the same plasmid as a reference (Promega, Madison, WI). The *Znf708* 3'UTR also was inserted downstream of firefly between SacI and SalI into the pmirGLO.

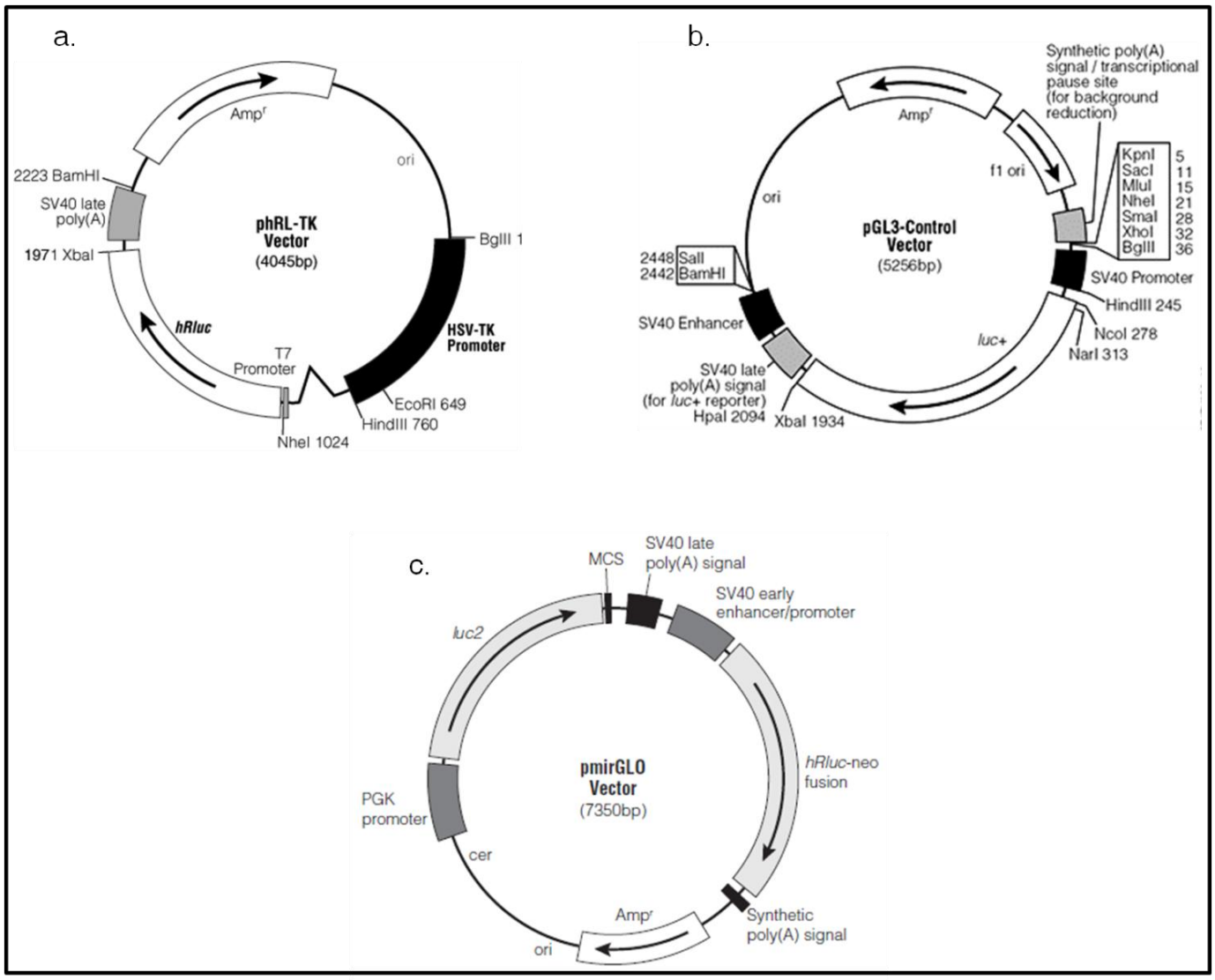

Figure 1. Schematic representation of the reporter containing vectors. a) phRL-TK contain the renilla reporter gene. **b)** PGL3 express the firefly **c)** In pmirGLO both reporter genes, renilla and firefly, are expressed from same backbone (Promega, Madison, WI).

Dual luciferase assay

To determine the luciferase reporter expression, cells were transfected in 24-well microtiter plates using Nanofectin reagent (PAA, Pasching, Austria) or jetPEI reagent (Pplyplus transfection) following the manufacturers instructions. For transfecting the mouse embryonic fibroblast (MEF) Nanofectamin was used (PAA, Pasching, Austria).

After 6hrs of transfection, cells were washed and incubated for 24 or 48 hrs prior to lysis and luciferase measurements. For luciferase (renilla and firefly) the dual luciferase assay (Promega, Madison, WI) was used. The readings for the experimental luciferase were normalized to readings for the reference construct. Experiments were done in at least 3 biological replicates.

Mouse embryonic fibroblast isolation and culture

To obtain mouse embryonic fibroblasts (MEFs) of different genetic backgrounds, mice heterozygous for ADAR1$^{+/-}$, ADAR2$^{+/-}$, or both ADAR1$^{+/-}$/2$^{+/-}$were intercrossed. Embryos were isolated from gravid mothers at day E11.5 (ADAR1) or E14. Embryos were genotyped and sex matched by PCR. Homozygous and wild type female embryos were homogenized with a syringae, cells were filtered through a cell strainer and cultured in DMEM supplemented with 20% FCS, gentamycin, penicillin, and streptomycin. Cells were cultured up to 8 passages and used for transfection reporter assays. To get the Myd88$^{-/-}$ MEFs, the same procedure was followed. The Stau1$^{-/-}$ MEFs have been gifted from Michael Kibler Lab and PKR$^{-/-}$ from Caetano Reis e Sousa lab.

Differentiation of embryonic stem cells

Mouse embryonic stem cells were differentiated using spontaneous differentiation of embryoid bodies. ESC were trypsinized and re-suspended in differentiation media (DMEM supplemented with 20% FBS, 0.1 mM β-mercaptoethanol, and pen/strep) at 5 x 10^4 cells/ml. Embryoid bodies were formed using the hanging drop method. For this, 300 cells were placed in a drop on the lid of a tissue culture dish. The dish was filled with PBS and cells were kept hanging at 37° C for 2 days. The newly formed EBs were transferred to gelatin coated dishes and left for 7 days for spontaneous differentiation. Embryoid bodies were trypsinized and used for further experiments.

RNA extraction

To determine RNA levels, cell lysates prepared for the dual luciferase assay were used immediately after lysis. Lysates were purified using the Quiagen RNAeasy mini kit (Quiagen, Hilden, Germany). After purification, an extra round of DNAse I and DpnI digestion was included, to avoid plasmid DNA contamination. DNAseI and DpnI was heat inactivated and the RNA was precipitated with ethanol,prior to RT-PCR or qPCR.

RT PCR and qPCR

cDNA synthesis was done with random hexamers and RevertAid RNAseH minus mMuLV reverse transcriptase following the manufacturers instructions (Fermentas, Lithuania). As a control, MOCK reactions without RTase were set up. For qPCR a GoTaq qPCR master mix was used (Promega, Madison, WI) on a BioRad iQ5 cycler (BioRad, Hercules, CA). At least three biological and two technical replicates were done for each qPCR assay.

Relative differences in RNA levels were determined by using the delta delta CT method (Hellemans, Mortier et al. 2007).

To determine RNA levels of firefly and renilla luciferase by qPCR the following primer pairs were used.

Renilla:

forward: AGATCATGCGGAAACTGGAG

reverse: CGAAGGTAGGCGTTGTAGTTG

Firefly in pHRLTK:

forward: GTGCCAACCCTATTCTCCTTC

reverse: CCGCTTCCCCGACTTCCTTAG

Firefly in pmirGLO:

forward: GGCTGAATACAAACCATCGG

reverse: CGCTCGTTGTAGATGTCGTTAG

Western blotting

To determine the phosphorylation status of eIF2-alpha, transfected MEFs were resuspended in 2x SDS running buffer and lysed by sonication. The cell extract was run on 10%SDS Laemmli gels (Ausubel F.M. and Janssen 1987), blotted onto nitrocellulose and detected with a phospho-specific eIF2-alpha specific antibody (ABCAM, Cambridge, UK) and a secondary horseradish peroxidase-coupled antibody. Signals were detected using a Pierce super-signal detection kit (Pierce, Rockford, IL). Membranes were subsequently stripped and total eIF2-alpha was detected using a pan-eIF2-alpha antibody. Levels of phosphorylation were calculated by dividing the amount of phospho eIF2-alpha by the amount of total eIF2-alpha.

Poly A length detection

To detect the length of the poly A tail, we used a Splint-mediated PolyA Tail measurement assay (Shattuckeidens, Mcclure et al.) (Minasaki, Rudel et al. 2014). In the following part, the protocol of the aforementioned method is explained in detail.

1. Extract the total RNA by triFAST and treat the purified RNA with DNaseI.

2. Clean up the RNA with Phenol:Chloroform:Isoamylalcohol (PCI) and Chloroform:Isoamylalcohol (CI). Precipitate RNA with ethanol.

3. Equally split samples into two tubes (RNase H+ and RNase H-), 1-2 ug each.

4. Heat the RNA sample at 95C for 1 min and immediately transfer tubes to ice water.

5. Add 2 ul oligo-dT (100uM). Rack, spin and incubate at RT for 20 min.

7. Add 5 ul 10x RNase H buffer and 1 ul RNase inhibitor (Ribolock / Thermo-Fermentas) to both tubes

8. Add **1 ul RNase H to RNase H+ tube** and **1ul DEPC-H2O to RNase H- tube**. Incubate them at 37C for 1 hour (hr).

9. Clean up RNA with PCI and CI. Precipitate RNA with ethanol. Resuspend the RNA in 5 ul DEPC water.

10. Per reaction, prepare at RT a master mix of 3 ul RNA anchor oligo (10uM) (MJ2440) and 2 ul DNA splint oligo (10uM) (MJ4073)

MJ2440: P-UCGUAUGCCGUCUUCUGCUUGUidT

MJ4073: AGAAGACGGCATACGATTTTTTTTT

11. Transfer 5ul RNA to PCR tubes on ice and add 5 ul of the RNA anchor / DNA splint mix to the RNA sample on ice. For annealing the total RNA, RNA anchor and DNA splint in a PCR machine follow the PCR program.

70C for 5 min

60C for 5 min

42C for 5 min

25C for 5 min

15C forever

12. in each tube add 6 ul DEPC-H2O + 1 ul RNase inhibitor (Ribolock / Thermo-Fermentas) + 2 ul 10X T4 RNA ligase buffer + 1 ul T4 RNA ligase 2 (Rnl2)

13. Incubate the samples at 15C o/n.

14 digest the splint DNA with 8-10 ul DNase I (fermentas, 10u/ul). Incubate at 37C for 30-60 min.

15. Clean up RNA with PCI and CI. Precipitate RNA with ethanol. Resuspend the RNA in 20ul DEPC water.

16. Heat the RNA sample at 95C for 30 sec, and then transfer immediately to ice water.

17. Prepare RT reaction master mix as following and mix with ligated RNA sample.

26 ul DEPC-H2O

1 ul RT primer (10 uM, MJ2438)

1 ul 10 mM dNTP mix

10 ul 5X RT buffer

1 ul RNase inhibitor (Ribolock / Thermo-Fermentas)

1 ul RTase enzyme (RT+) or DEPC-H2O (RT-)

MJ2438: CAAGCAGAAGACGGCATACGA

18. Incubate the samples in PCR machine at

50C for 30 min

85C for 5 min

20C forever

19. Do the **Polymerase Chain Reaction (PCR)**

20. The PCR master mix consisting of

17ul H2O

2.5 ul 5X dream taq PCR buffer

1 ul 10mM dNTP mix

2 ul Forward primer (0.1 ug/ul, gene-specific outer primer)

2 ul Reverse primer (0.1 ug/ul, MJ2438)

0.5 ul dream taq polymerase enzyme (fermentas)

2ul cDNA

21. Incubate in PCR machine at

i. 98C for 2 min

ii. 98C for 25 sec

iii. 60C for 25 sec

iv. 72C for 20 sec

v. Back to ii, repeat 17 cycles

vi. 72C for 1 minute as a last extension

22. Prepare 2nd PCR master mix which is identical to the 1st PCR mix in Step 1. Add 1 ul of 1st PCR product instead of cDNA and run the same PCR program.

Actinomycin treatment

For detecting the rate of mRNA degradation, the mRNA transcription has been blocked by Actinomycin D (10 ug/ul) and the mRNA was collected in the regular time interval (0h, 1h, 2h). The mRNA amount was quantified by qPCR. RNA extraction, cDNA synthesis and qPCR were done as explained before.

PolII chromatin immuneprecipitation (ChIP)

To check the rate of mRNA synthesis, we compared the polII occupancy along the reporter gene using ChIP. In the following part, the method is explained in detail.

Crosslinking

- Start with 3 * 10cm dishes for each IP and crosslinking by adding 189μl of formaldehyde (37%) per 7ml of medium directly to the cell culture dish with adhesive cells (final concentration of formaldehyde 1%)
- Incubate at 37°C or gently rotate at room temperature (RT) for 10 minutes
- Stop the crosslinking reaction by adding 1 ml of 2.5M glycine
- Incubate for 5min at RT by gentle rotation
- Wash twice with 10ml 1x PBS
- Add another 5ml of 1xPBS for cell harvesting

Harvesting

- Harvest cells by scraping and transfer PBS/Cell suspension into Falcon tube on ice
- Centrifuge for 5min at 1200rpm at 4°C
- Discard supernatant (SNT), resuspend cell pellet with 5ml WASH1 buffer (+inhibitors)* *
- Incubate 10min on ice
- Centrifuge 5min at 1200rpm at 4°C
- Discard SNT, and re-suspend cell pellet with 5ml WASH2 buffer(+inhibitors)* *
- Incubate 10min on ice
- Centrifuge 5min at 1200rpm at 4°C
- Resuspend cell pellet in 0.7 to 1ml of **freshly prepared** LYSIS Buffer (+inhibitors)* *

** For protease inhibitor (PI) use one tablet of Complete (Roche) dissolved in 1ml of water (50x stock solution). Phosphatase inhibitors (50x stock solution): sodium orthovanadate (final concentration 100 μM, activated for 5 min at 95^{o}C), β-glycerophosphate (final concentration 10 mM), sodium fluoride (final concentration 10 mM), sodium molybdate (final concentration 10 μM)

Lysis Buffer	10ml
RT 20% SDS	0.5ml
0.5M EDTA	0.2ml
1M Tris pH 8.1	0.5ml
50x PI tablet	0.2ml
100x Na Butyrate	100µl
Phosphates inhibitors	0.2ml

Chromatin Sharing using Bioruptor (Diagenode)

- Sonicator: Bioruptor
- Conditions: Power max, Time 30sce on and 30sce off.
- Use 2ml eppendorf tubs with 500 – 700 µl of crosslinked chromatin

- Sonicate crosslinked chromatin 30x (for transfected cells)
- Centrifuge 2x 10min at 4°C maximal speed to get rid of insoluble cellular parts each time collecting SNT into a new tube.

Measurement of Chromatin Concentration

- Measure 1µl of sonicated chromatin with Qubet when using Lysis Buffer as Blank (determine DNA concentration)
- Use 25-200µg of shared chromatin per IP (volume should not be bigger than 200µl) and 1/10 of the amount for INPUT

At this point chromatin can be aliquoted, snap frozen in liquid N_2 and stored at -80 for few months.

Setting of the IP

- Dilute shared chromatin 5-10x with Dilution Buffer (+PI +Butyrate)
- Take 1/10 of amount of chromatin taken for IP as input and keep at 4°C till elution step on next day

- Add 4µg per IP of polII antibodies ab5408
- Incubate on the rotor ON at 4°C to allowed

Parallel to this prepare and block beads for IP.

Preparation and blocking of Magnetic beads

Use Dynabeads Pan mouse IgG from life technology.

- Take 25µl of Magnetic beads per IP
- Put on magnetic rack and wash three times in Dilution Buffer (+PI +Butyrate)
- Re-suspend beads in 90µl of dilution buffer and 10µl of BSA (10mg/ml) per IP
- Block beads by rolling over night at 4°C

Harvesting of chromatin

- Distribute blocked beads to new 1.5ml safe lock eppendorfs, 100µl/IP.
- Spin down shortly (1s), 4°C and separate them from blocking solution with help of magnetic rack. Discard SNT.
- Add chromatin/antibody solution to the beads, vortex briefly and roll for 3-5 hours at 4°C (this will allow binding of chromatin/antibody complex to the magnetic beads)

Washing of Chromatin IPs

- Spin down (1s) chromatin/antibody/ magnetic beads suspension, 4°C, separate magnetic beads from solution with help of magnetic rack 4°C and discard SNT.
- Wash magnetic beads with washing buffers in following order:
    - 1x RIPA
    - 1x High Salt
    - 1x LiCl
    - 1x TE

- After adding 1ml of washing buffer to each IP vortex briefly to re-suspend beads and rotate for 10 min at 4°C
- Spin chromatin/antibody/ magnetic beads suspension down for 1 s at 4°C, separate magnetic beads from solution with help of magnetic rack and discard SNT.
- Add 1ml of next washing buffer
- During washing steps prepare elution buffer

Elution Buffer	**5ml**
20% SDS	0.5ml
1M $NaCO_3$ (in 2ml H_2O)	0.5ml
DTT	50µl

Elution of chromatin from magnetic beads

- After last wash with TE spin chromatin/antibody/ magnetic beads suspension down for 1 s, 4°C, separate magnetic beads from solution with help of magnetic rack 4°C and discard SNT.
- Add 400µl of Elution buffer to each IP
- Incubate at RT for 30min while shaking strongly

At this stage do not forget to join inputs that were prepared day before!!!

- Fill input with elution buffer up to 400µl.

- Spin down shortly to collect magnetic beads and separate them from eluted chromatin with help of magnetic rack
- Carefully transfer SNT into a new safe lock 1.5ml eppendorfs

Reverse crosslinking of eluted chromatin

- Add 16μl 5M NaCl to all samples (including inputs)
- Incubate 6 hours to overnight at 65°C on heating block at 300rpm

Proteinase K digestion and DNA precipitation

- Add 8μl 0.5M EDTA + 16μl 1M Tris pH6.5 + 2μl Proteinase K
- Shake 1 hour at 55°C
- Transfer the contaent to new tube
- Add 600μl PCI
- Centrifuge at 14.000xg for 5 min
- Transfer aqueous phase to the new Eppi
- Add 1ml ice cold 96% ETOH + 40μl 3M sodium acetate, pH 5.2 + 1μl glycogen
- Incubate at least 30min at -20°C
- Centrifuge 30min at 4°C full speed
- Wash with 1ml 70% ETOH
- Air dry pellet and dissolve in 200μl H_2O
- Incubate 30min at RT
- Do the Real time PCR

Primers

Part A

forward: TCTGCGGCTTTCGGGTCGTG

reverse: CCGCTGGCGATCTCGTGCAAG

Part B

forward: GGCGGCAAGATCGCCGTGTA

reverse (*NICN1*): GGCCCTATGGGCACACCCAC

reverse (*5'+3'S ZNF708*): ACCATGCCCGGCCCTACCTC

Part C

forward: CGGCCGCTTCGAGCAGACAT

reverse: TCACTGCATTCTAGTTGTGGTTTGTCC

ChIP buffers

WASH I	**400ml**
10% Triton X-100	10ml
0.5M EDTA	8ml
0.5M EGTA	0.4ml
1M HEPES	4ml

WASH II	**400ml**
4M NaCl	20ml
0.5M EDTA	0.8ml
0.5M EGTA	0.4ml
1M HEPES pH7.9	4ml

Dilution Buffer	**400ml**
4M NaCl	16.704ml
1M Tris pH 8.1	6.688ml
0.5M EDTA	0.96ml
10% Triton X-100	44ml
20% SDS	0.2ml

RIPA	**400ml**
4M NaCl	15ml
1M Tris pH 8.0	20ml
20% SDS	2ml
10% NaDOC (fresh)	20ml
10% NP40	40ml

High Salt	**400ml**
4M NaCl	50ml
1M Tris pH 8.0	20ml
20% SDS	2ml
10% NP40	40ml

LiCl Wash	**400ml**
1M LiCl	100ml
1M Tris pH 8.0	20ml
10% NaDOC (fresh)	20ml
10% NP40	40ml

TE	**400ml**
1M Tris pH 8.0	4ml
0.5M EDTA	0.8ml

Results

Results

*i*SINEs modulate gene expression

Generation of iSINE-containing UTRs in reporter genes

To determine the impact of inverted SINEs on gene expression we focused on two different 3' UTRs that harbor two Alu elements in inverted orientation. On one hand, we picked the 3' UTR of the *nicolin* (*Nicn1*) gene that had already been proven to interfere with gene expression when fused to a green fluorescent protein reporter (Chen, DeCerbo et al. 2008). *Nicn1* is a protein coding gene which is only present in mammals. The function of NICN1 is not known, but the protein is localized in the nucleus and shows tissue-specific expression (Backofen, Jacob et al. 2002).

On the other hand, we picked the 3' UTR of the InaD-like gene (*Inadl*). INADL is a hydrophobic protein that mediates clustering of membrane proteins via the PDZ domain. The PDZ domain is a protein binding module contributing to the formation of macromolecular complexes in membranes (Philipp and Flockerzi 1997).

The 3' UTR of *Nicn1* contains an Alu Sp1 and an Alu Sp2 element in a tail to tail configuration spaced only 70 bp apart. The two Alu Sp elements are 81% identical in their sequences. The 3' UTR of *Inadl* contains an Alu Sx and an Alu Sg element in a head to head configuration. Both Alus are 79% identical in sequences and spaced about 1000 bp apart. The Alu elements containing UTRs were cloned downstream of the open reading frame (ORF) of renilla luciferase in the phRL-TK vector which is driven by the moderately active thymidine kinase promoter, while the reference firefly luciferase is driven from a SV40 promoter. The 3' UTR of *Nicn1* was cloned so that the first Alu was located about 190 bp downstream of the luciferase ORF while the Alu elements in *Inadl* were located about 350 bp downstream of the luciferase stop codon (Figure1). The *i*SINEs in both 3' UTRs are believed to form a basepairing interaction as

they had been reported to be heavily edited by ADARs (Athanasiadis, Rich et al. 2004; Levanon, Eisenberg et al. 2004).

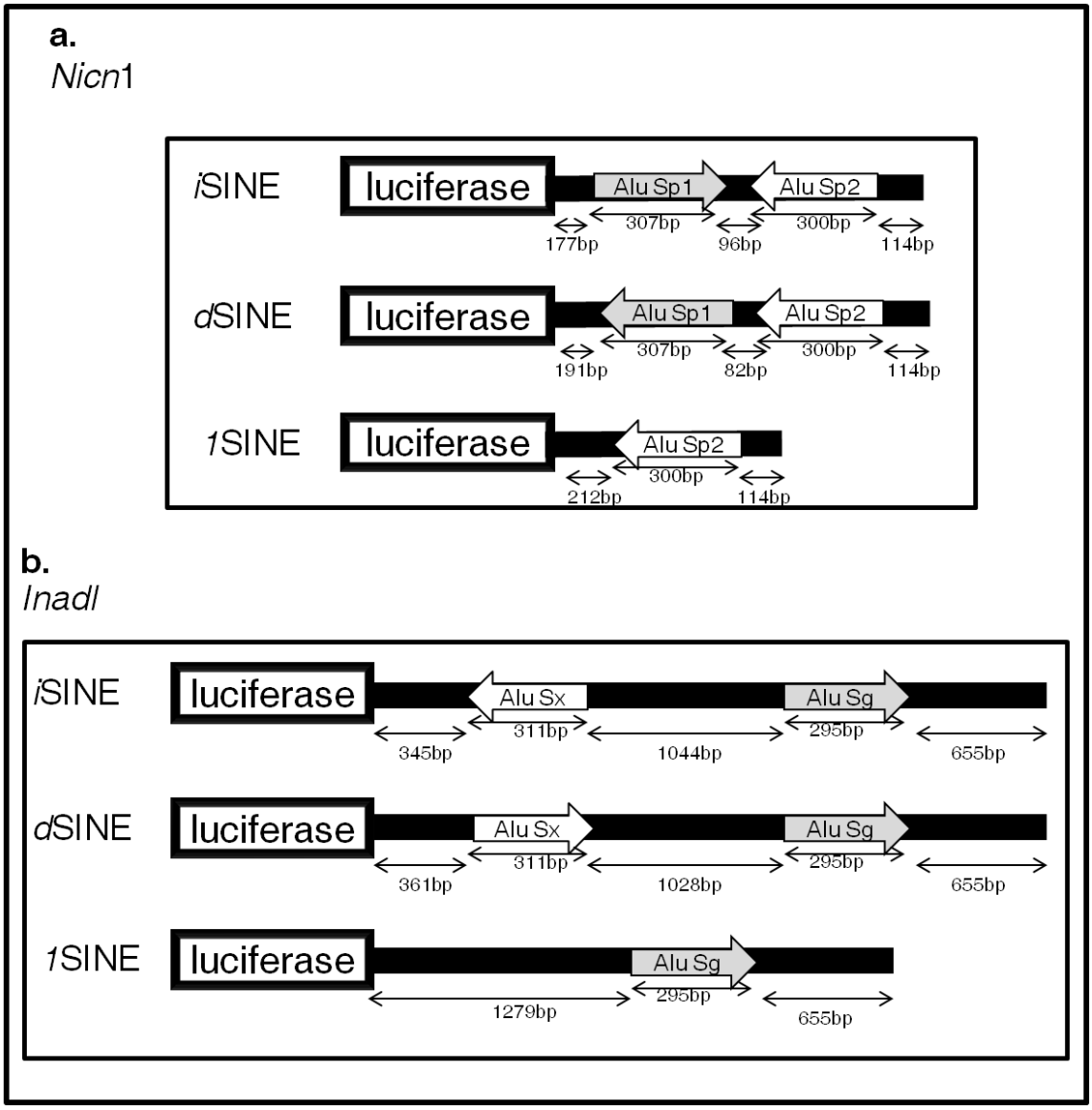

Figure1: Two inverted SINE (*i*SINE) containing 3'UTRs were cloned downstream of a reporter gene. a) The *i*SINE containing 3'UTR of the *Nicn1* gene was inserted downstream of the renilla luciferase ORF. SINEs are located in a head to head configuration in the 3'UTR. As a control, one of the Alu elements is flipped to make a duplication of SINE (*d*SINE). In addition, Alu Sp1 was removed to maintain one SINE in the constructs (*1*SINE). **b)** The 3'UTR of *Inadl* gene, which contains an *i*SINE in a tail to tail configuration, was inserted downstream of the

renilla luciferase gene. Two control vectors, *d*SINE and *1*SINE, were generated by inverting or deleting one of the two Alu elements in the *Inadl i*SINE containing construct.

iSINEs can form double-stranded structures

To determine whether the cloned *i*SINEs can form double-stranded structures in the context of our reporter constructs and in the cell lines used, the reporter constructs were transfected into mouse embryonic fibroblasts (MEFs) and the editing status was determined by sequencing of the cDNAs. The inverted SINEs became edited in MEFs suggesting that the predicted double-stranded structures are also formed by the RNAs expressed from the reporter constructs (Figure 2).

Moreover, editing was more pronounced in the *Nicn1* sequencing traces, where 20 sites were edited above 50%. In contrast, editing rates in the *Inadl* 3' UTR only reached a maximum of 36%. Efficiency of editing was measured by dividing A and G peak height by G peak. This finding is consistent with the idea that the two more closely spaced SINEs in *Nicn1* are more likely to form a double-stranded structure than the more distantly spaced SINEs in *Inadl*.

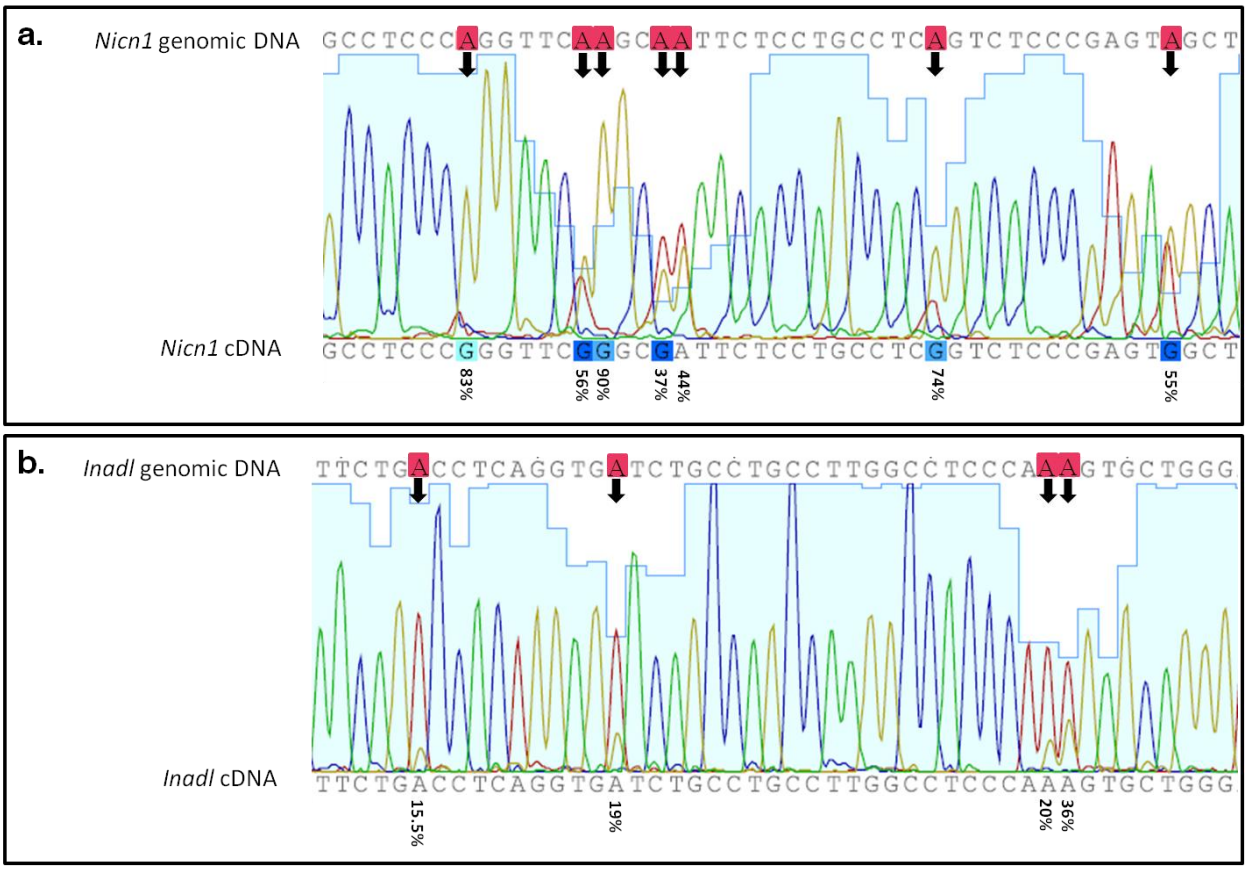

Figure 2: Alu elements in *i*SINE containing constructs are edited, indicating that *i*SINEs can form double-stranded structures. a) *Nicn1* and **b)** *Inadl i*SINE containing constructs were transfected in wtMEFs, total RNA was extracted and cDNA was synthesized. cDNA sequencing indicated that the *Nicn1 i*SINE was edited with more than 50% efficiency. *Inadl i*SINE was edited with less than 36% efficiency. Efficiency of editing was measured by dividing the sum of A and G peak heights by the G peak height. The edited adenosine is marked by an arrow and the amount of editing efficiency is shown under the editing peak. Yellow peak resembles G and red peak A residues.

iSINEs reduce gene expression

Next, we tested the impact of these *i*SINEs on reporter gene expression. As a control, one of the two SINEs was inverted giving rise to a duplication of SINEs (*d*SINE). As an additional control, one of the two SINEs was removed, leaving a single SINE in the

construct (*1*SINE) (Figure 1). The constructs were cloned in the phRL-TK vector downstream of renilla luciferase which is driven by the moderately active thymidine kinase promoter while the reference firefly luciferase is driven from a SV40 promoter.

In wild-type mouse embryonic fibroblasts (MEFs), the *i*SINEs of *Nicn1* and *Inadl* led to a reduction in gene expression when compared to that of a *1*SINE construct. Gene expression was reduced by 50% to 30% depending on the 3' UTR used (Figure 3a, b). In contrast, *d*SINEs show a variable reduction of the reporter gene expression.

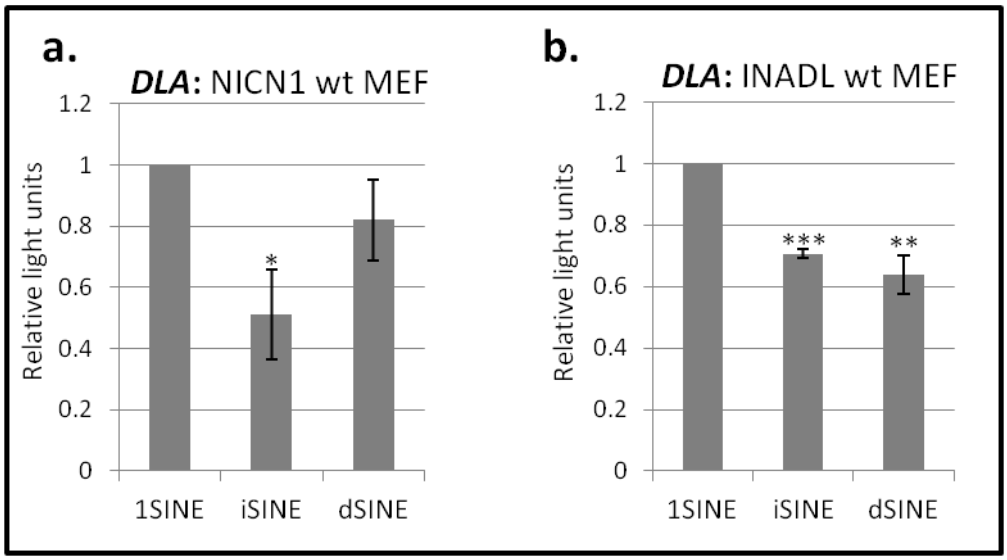

Figure 3: *i*SINEs in 3'UTR repress reporter gene expression. Dual luciferase assays (DLA) of reporters harboring a single SINE (*1*SINE), two SINEs of inverted orientation (*i*SINE) or two SINEs in direct orientation (*d*SINE) were done in wild-type MEFs. The different SINE configurations were derived from the *nicolin1* (*Nicn1*) or the *inaD-like* (*Inadl*) gene, transfected in wtMEFs and the expression of reporter genes was detected 24-48 h after transfection **a)** In wtMEFs the *i*SINE of *Nicn1* leads to a significant reduction in reporter gene expression while the *d*SINE construct only leads to a marginal reduction in reporter gene expression. **b)** In contrast, both *i*SINE and *d*SINE of *Inadl* significantly reduce gene expression. The experiments have been done in triplicate. Asterisks indicate p-values calculated with student's T-ters * $p<0.05$, ** $p<0.005$ and *** $p<0.0005$.

Screen for factors influencing gene silencing by *i*SINEs

Sequence configurations that influence the regulatory effect of iSINEs

The *Nicn1 i*SINE led to a 50% reduction in gene expression while, the *Inadl i*SINE showed only 30% reduction. To understand the different effects on gene expression, we aimed at understanding differences within the 3'UTRs that can modulate the effect of inverted SINEs on gene expression. We therefore altered several features in SINE-containing constructs. We changed the SINE configuration, the probability of forming double-stranded structures and the relative distance of SINEs to regulatory elements in the messenger RNA such as the polyadenylation signal or the stop codon. To investigate the importance of these features, we decided to create a new *i*SINE containing construct that allowed us to alter parameters in a coordinated manner. We selected the 3'UTR of *Znf708* and cloned it downstream of the open reading frame (ORF) of firefly luciferase in the pmirGLO vector. This vector simultaneously expresses renilla and firefly luciferase to allow easy quantification of changes in gene expression using a dual luciferase assay. The *Znf708* 3' UTR harbors an Alu Sc and an Alu Sg element in its natural configuration that are spaced 160 bp apart in a head to head configuration and exhibit 77% sequence identity. The first Alu is located 1026 bp downstream of the stop codon and the second one is located 500 bp upstream of the polyadenylation site (Figure 4a).

First we checked the importance of SINE orientations on the reduction of gene expression. To create all possible SINE orientations, the two SINEs in the *Znf708* 3'UTR containing vector were flipped individually. Therefore, in addition to the endogenous head to head (forward, +; followed by reverse, c) configuration (*i*SINE(+,c)), we created an *i*SINE construct with the Alus in tail to tail orientation (*i*SINE(c,+)), one construct with duplicated SINEs in tail to head (*d*SINE(c,c)) and another one with head to tail configuration (*d*SINE(+,+)). Also, the second SINE was removed leaving a single SINE in forward (*1*SINE(+)) or reverse (*1*SINE(c)) orientation (Figure 4b). To compare the effect on expression levels, SINE harboring constructs were transfected in human osteosarcoma (U2OS) cells. The results showed that *i*SINE-

containing 3'UTRs in either configuration, (+,c) and (c,+) led to a 20% to 27% reduction in gene expression (Figure 4c), which indicates that the configuration of SINEs does not contribute in the strength of reduction.

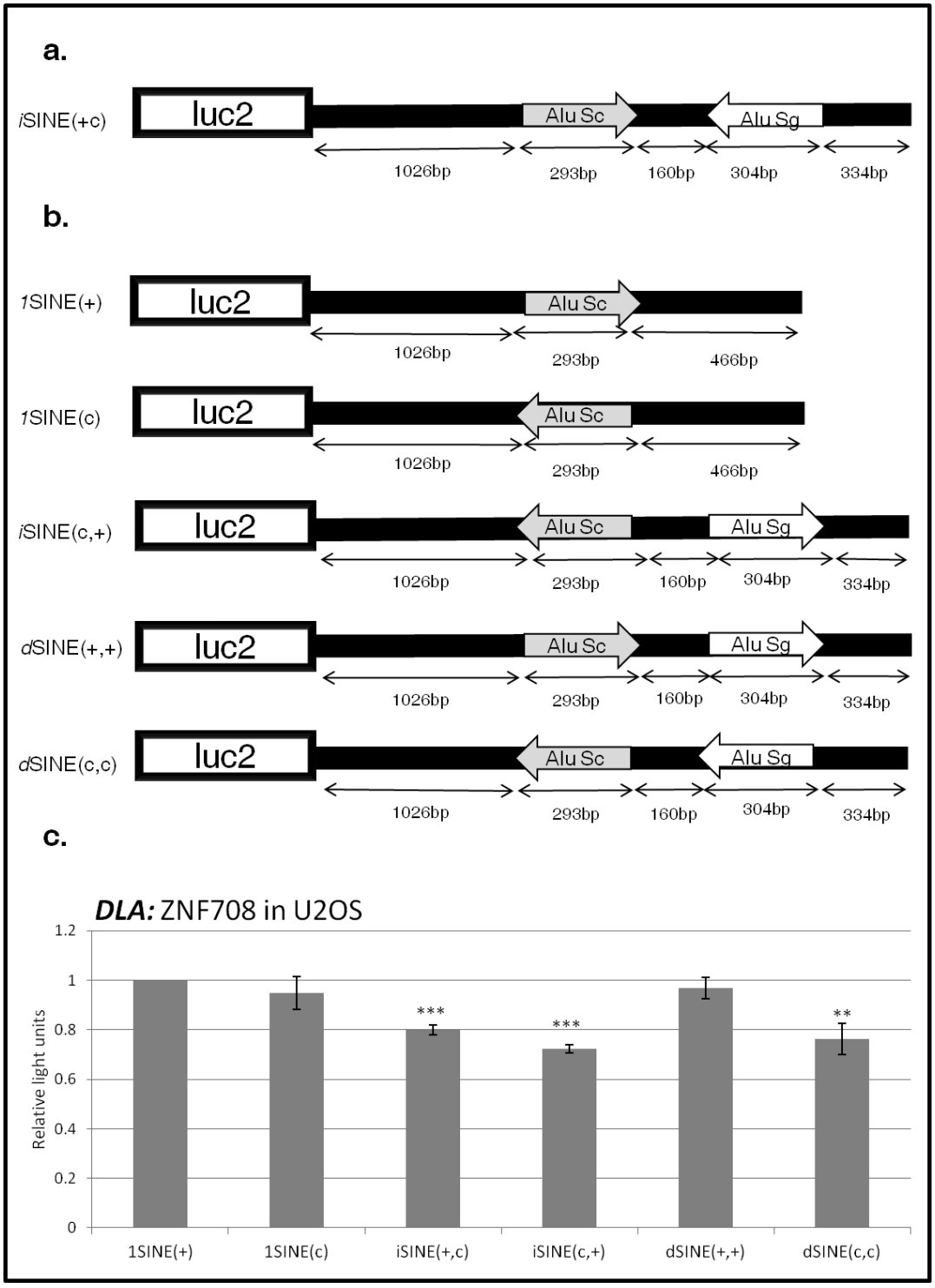

Figure 4. SINE configuration created in *Znf708* 3'UTR, tested in U2OS cells. a) The *Znf708* 3' UTR was inserted downstream of the firefly reporter gene in the pmirGLo vector. The 3'UTR harbors an Alu Sc and Alu Sg element that are spaced 160 bp apart in a head to head configuration and show 77% sequence identity. The first Alu is located 1026 bp downstream of the stop codon and the second one is around 500 bp upstream of the polyadenylation site**. b)** The first or the second SINE in *Znf708* 3**'**UTR was removed to generate *1*SINE+ and *1*SINEc, respectively. Also individual SINEs were flipped individually to generate all possible SINE configurations, *d*SINE(+,+) and *d*SINE(c,c). **c)** SINE containing constructs were transfected in U2OS and the expression of the reporter gene was measured after 24 h. *i*SINE (c,+) showed the strongest reduction by 27%. *i*SINE(+,c) and *d*SINE(c,c) also showed 20% and 23% reduction, respectively. Asterisks indicate p-values calculated with student's T-test * $p<0.05$, ** $p<0.005$ and *** $p<0.0005$. (+: Sense and c: Antisense)

Another difference between *Nicn1* and *Inadl* 3'UTR is the relative location of SINEs in the 3'UTRs. Capshew, et al., have shown that the distance between SINEs and the stop codon can influence the regulatory function of *i*SINEs. When the distance is reduced to 65 nucleotides, the effect of the SINEs is diminished (Capshew, Dusenbury et al. 2012). Therefore, we shortened the SINE containing 3'UTRs (Figure 5a). In a first step, the sequence between the stop codon and the first SINE was shortened to 500 bp. However, shortening the distance between the stop codon and the first SINE did not alter the reduction in gene expression (Figure 5b). This *i*SINE containing construct led to 30-35% reduction in gene expression. Next, the distance between the stop codon and the SINE was further reduced to about 50 bp. Also the distance between the second SINE and the poly-A signal was shortened. Consequently SINEs were located 50 bp downstream of the stop codon and 150 bp upstream of the poly-A site (Figure 5a). In the new set of shortened constructs similar results were obtained. *i*SINE(+,c) reduce gene expression only moderately while the *i*SINE(c,+) construct showed a 25% reduction (Figure 5c). Based on this data, it appeared that the location of *i*SINEs in the 3'UTR has little effect on the extent by which *i*SINEs reduce gene expression.

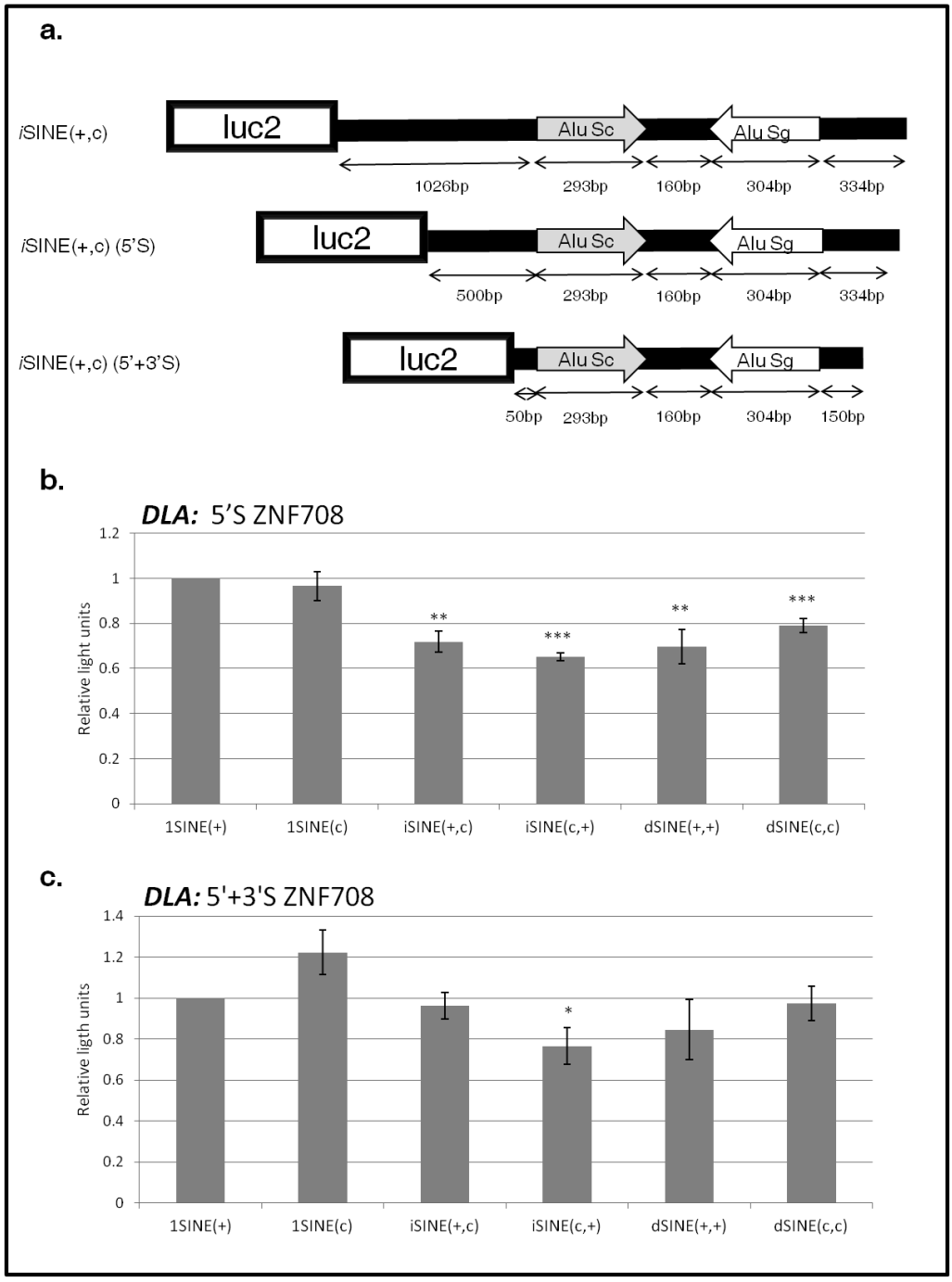

Figure 5. Shortening of the *Znf708* 3'UTR does not affect *i*SINE-mediated reduction of gene expression. a) The *Znf708* 3'UTR sequence was shortened to obtain two sets of SINE containing constructs of different length. The constructs were termed 5' shortened (5'S) and 5' and 3' shortened (5'+3'S). The 5'S constructs were made by removing 500 bp upstream of the SINEs. In order to create the 5'+3'S 3'UTR constructs, the whole sequence between the stop codon and the first Alu were shortened to 50 bp. In addition, 150 bp were removed downstream

of the SINEs. Again all possible SINE configurations were made for both shortened constructs. **b)** U2OS cells were transfected with the various reporter constructs and the expression was measured after 24 hrs. After shortening, the strength of reduction by *i*SINEs did not increase and they led to 30-35% reduction in gene expression. **c)** In the 5'+3'S *Znf708* containing constructs, only *i*SINE(c,+) showed a significant reduction of 25%. Asterisks indicate p-values calculated with student's T-test * $p<0.05$, ** $p<0.005$ and *** $p<0.0005$.

Another factor by which *Nicn1* and *Inadl* 3'UTRs differ, is the extent of double-strandedness. Our editing assay showed that editing of the *Nicn1 i*SINE is more prominent than *Inadl i*SINEs, which could be due to a higher probability of making double-stranded structures. To check the importance of double-strandedness, we made constructs which harbored identical SINEs in an inverted orientation (Figure 6a). This constructs could form a perfect double stranded structure. Interestingly, our data showed that perfect, inverted (*pi)*SINE(+,c) located in either at their natural 3' UTR context or in the context of the shortened, 5'+3'S *Znf708* 3'UTR led to a stronger reduction (around 35%) when compared to the natural Alu Sc Alu Sg *i*SINE(+,c) (Figure 6b, c). Therefore, increasing the extent of double-strandedness more efficiently represses gene expression.

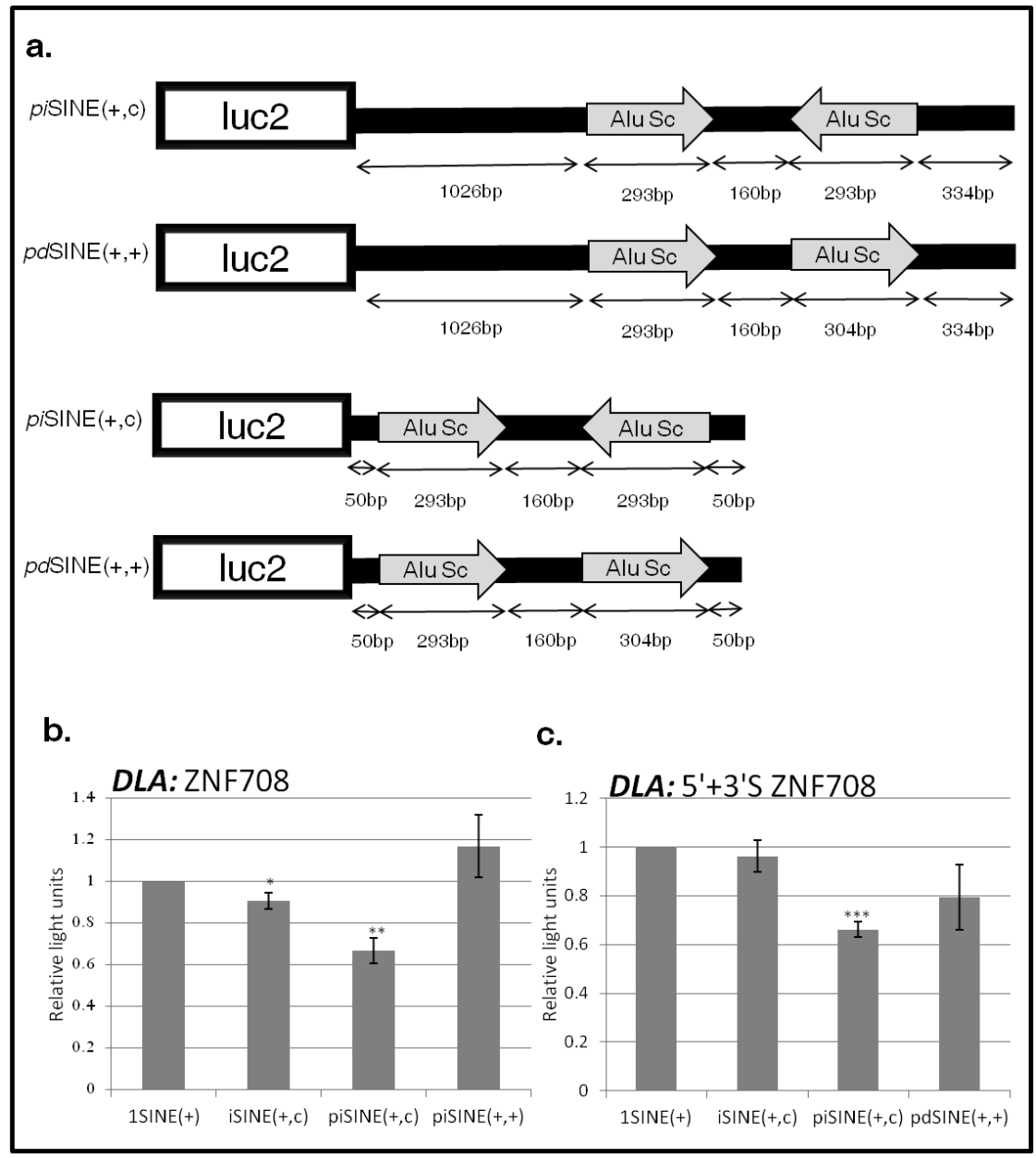

Figure 6. Increasing the extent of double strandedness can influence *i*SINE-dependent gene reduction. a) In the native and truncated 5'+3'S *Znf708*-containing constructs the Alu Sg was replaced by Alu Sc to make perfect *i*SINE(+,c) (*pi*SINE) and perfect *d*SINE(+,+) (*pd*SINE) containing constructs. SINEs in the *pi*SINE harboring the reporter mRNA are able to make perfect double-stranded structures. **b, c)** *pi*SINE(+,c) and *pd*SINE(+,+) constructs were transfected in U2OS cells and 24h later cells were lysed and dual luciferase assay was performed. The native and the 5'+3'S *Znf708* context *pi*SINE(+,c) reduce gene expression to 66%, whereas the *i*SINE(+,c) leads to only 5-10% reduction. Asterisks indicate p-values calculated with student's T-test * $p<0.05$, ** $p<0.005$ and *** $p<0.0005$.

iSINEs repress gene expression independent of A-to-I editing

Human Alu elements have been show to be highly edited (Philipp and Flockerzi 1997; Levanon, Eisenberg et al. 2004). In addition, in 2008 it was proposed that *i*SINEs in 3'UTRs are edited and retained in the nucleus (Chen, DeCerbo et al. 2008). To check whether ADARs could interfere with the observed gene reduction, different *i*SINE containing constructs were transfected into editing deficient cells. A similar effect was seen in cells with and without editing activity. In MEFs lacking both, ADAR1 and ADAR2 a significant reduction was seen for the reporter constructs carrying the *i*SINEs of *Nicn1* or *Inadl* (Figure 7c, d). This observation indicates that the reduction of gene expression by *i*SINEs is independent of editing.

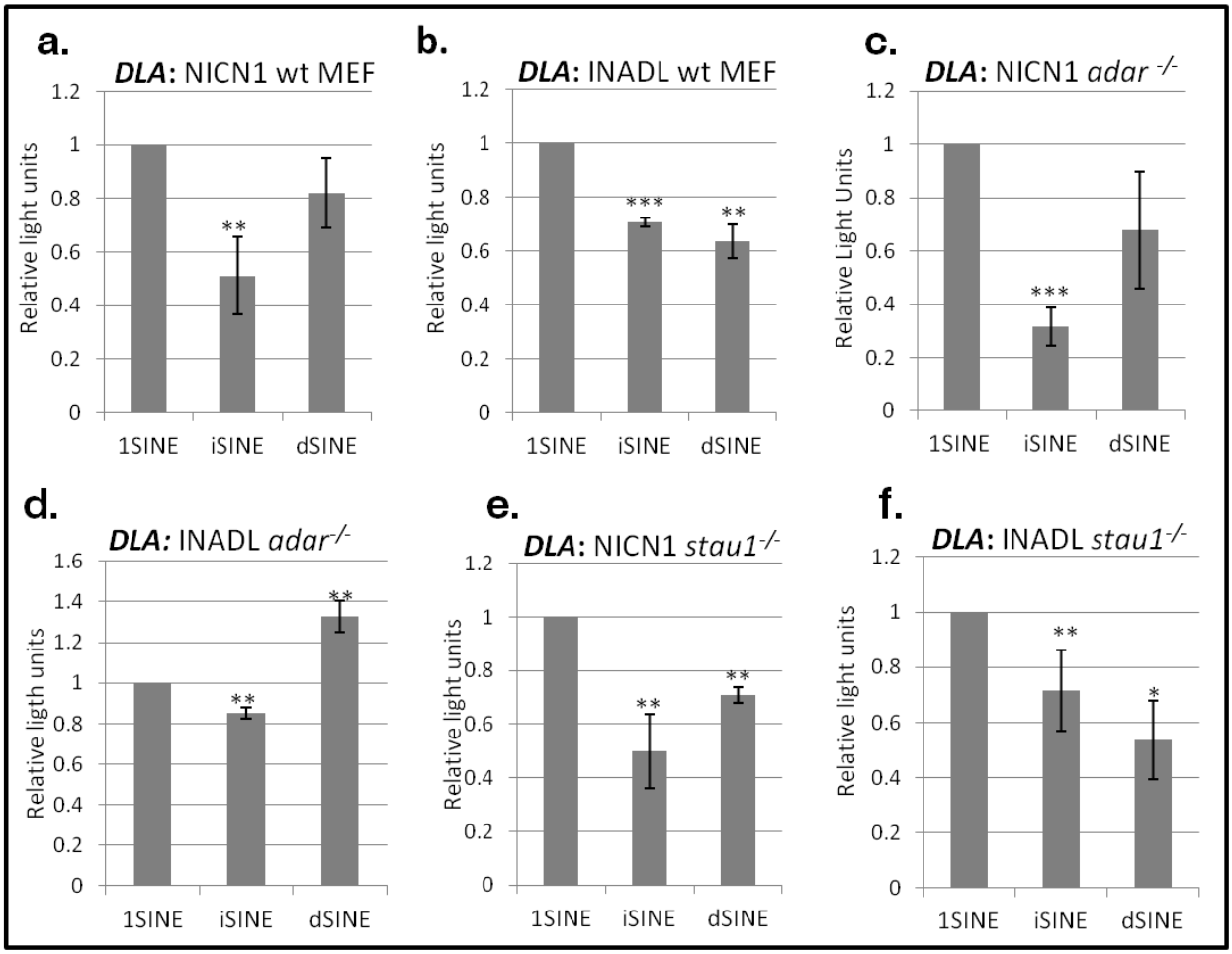

Figure 7. *i*SINE mediated gene repression is not mediated by RNA editing or the SMD pathway. SINE containing constructs derived from *Nicn1* and *Inadl* 3'UTR were transfected in cells lacking editing activity or STAU1. After 48h the expression of reporter genes were measured. **a**) *Nicn1* and **b)** *Inadl i*SINE containing reporter genes in wild type MEFs show a 50% and 30% reduction in gene expression, respectively. **c)** In MEFs derived from *adar1*$^{-/-}$, *adar2*$^{-/-}$ embryos (*adar*$^{-/-}$) the *i*SINE of *Nicn1* strongly reduces gene expression by almost 70%. **d)** In *Inadl i*SINE showed a reduction to 20% in *adars*$^{-/-}$ cells. **e)** *Nicn1* and **f)** *Inadl i*SINE showed a significant reduction of 50% and 70% in *stau1*$^{-/-}$ cells, respectively. Asterisks indicate p-values calculated with student's T-test * $p<0.05$, ** $p<0.005$ and *** $p<0.0005$.

Staufen1 independent effects of SINEs

Alu elements interfere with gene expression by triggering Staufen mediated RNA-decay (SMD) (Kim, Furic et al. 2007; Gong and Maquat 2011; Gong, Tang et al. 2013).

To test whether the repression of gene expression might be regulated by SMD, we generated MEFs from *Staufen1* deficient mice (Vessey, Macchi et al. 2008). *i*SINE constructs of both *Nicn1* and *Inadl* were transfected into *stau1*$^{-/-}$ fibroblasts. Despite the absence of Staufen1, the *iSINEs* showed a significant reduction in gene expression, comparable to that reduction observed in wt MEFs (Figure 7e, f).

iSINEs and dSINEs repress gene expression in mouse and human cells

Staufen1 mediated gene repression by SINEs had been shown to occur in human cell lines where Alu-bearing long non-coding (lnc) RNAs basepair with Alu elements located in 3' UTRs of mRNAs (Gong and Maquat 2011). However, the SINEs used in the constructs harboring *Nicn1* or *Inadl* 3'UTRs were of human origin while the *stau1* and *adar* deficient cell lines were mouse cells.

We already showed that *i*SINEs derived from the mouse *Znf708* 3'UTR led to a reduction in gene expression in human U2OS cells. Therefore, to determine whether the

*i*SINE containing UTRs of *Nicn1* or *Inadl* would also affect the reporter gene expression in human cells, reporter constructs were transfected in human osteosarcoma (U2OS) cells. We obtained that in these human cells, the *i*SINEs of *Nicn1* and *Inadl* led to 32% and 45% reduction in gene expression, respectively (Figure 8a, b). Like in MEFs a tandem (*d*SINE) arrangement of SINEs in *Inadl* also led to a reduction in gene expression (Figure 8b). These results confirmed that SINEs originating from primate Alus can interfere with gene expression in mouse and human cells.

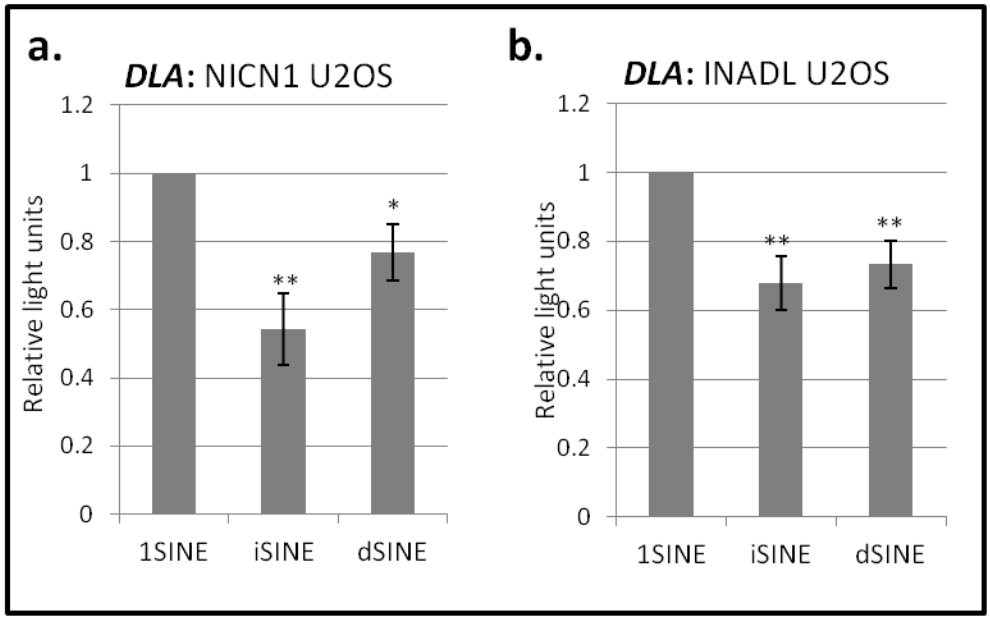

Figure 8. *i*SINE and *d*SINE mediate gene repression in a species-independent manner. The constructs derived from **a)** *Nicn1* or **b)** *Inadl* were transfected in human U2OS cells and the expression level was detected after 48hrs. In both cases, *i*SINE 3'UTR led to a reduction in gene expression by 45% and 32% in *Nicn1* and *Inadl*, respectively. The reduction in human U2OS cells mimics the effect observed in mouse embryonic fibroblasts. Asterisks indicate p-values calculated with student's T-test * $p<0.05$, ** $p<0.005$ and *** $p<0.0005$.

Next, we tested whether also mouse SINEs, derived from B1 elements would interfere with gene expression. B1 elements are half the length of Alu elements and are much less conserved than their primate counterparts. Consequently, the double-stranded regions formed between two antiparallel B1 elements are shorter and less extensive

(Neeman, Levanon et al. 2006). We therefore compared the impact on gene expression of two inverted human Alu elements and two inverted mouse B1 elements. To do this, the 3' UTR of the human *Znf708* gene was cloned downstream of the firefly luciferase in pmirGLO. *Znf708* harbors two Alu elements in inverted orientation. For comparison, the human Alu elements were replaced with two B1 elements of the mouse *car5b* gene (Figure 9a). Since the two B1 elements in *car5b* are only partially complementary, replacement of the Alu elements with the two different B1 elements of *car5b* only mildly affected luciferase expression (Figure 9b). Our pervious data showed that the extent of double-strandedness can influence gene expression. Consequently, the second B1 element was replaced by the first B1 element, and thus duplicated in inverted orientation (Figure 9a). The resulting formation of a short but perfect double-stranded structure led to a reduction in gene expression by 30% (*pi*SINE in Figure 9b), while the corresponding *pd*SINE had almost no influence on luciferase expression (Figure 9b). This suggests that the extent of complementarily and thus base-pairing influences the strength by which gene expression of reporter genes is repressed. Taken together our data indicates that regulation of gene expression by *i*SINEs is apparently sequence- and species independent, and is a common phenomenon, at least between rodents and primates.

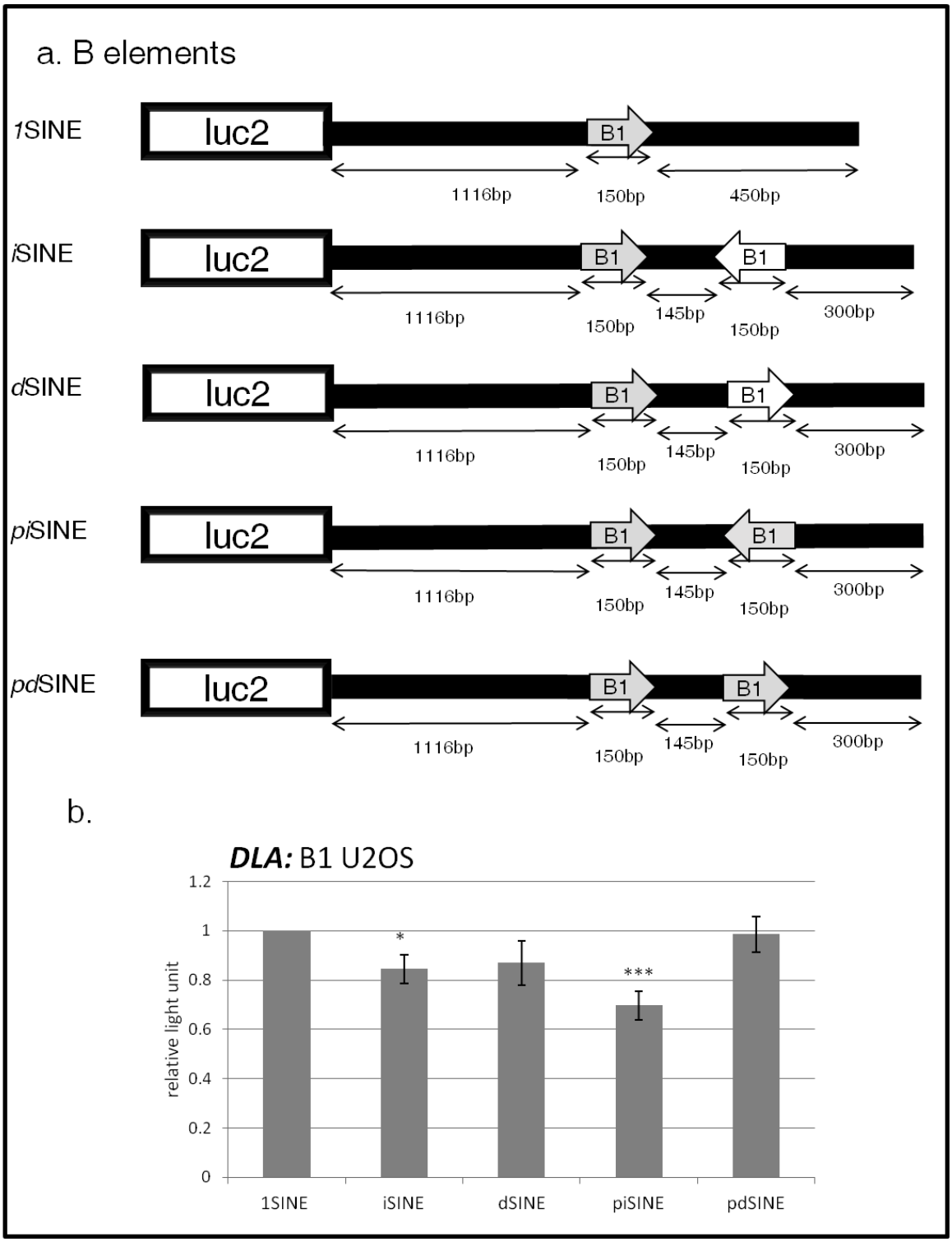

Figure 9. Mouse *i*SINEs can reduce gene expression in human cells. **a)** To generate *i*SINE containing constructs derived from B1 elements, we used the human *Znf708* 3'UTR as a starting construct. The *Znf708* Alu elements were replaced by B1 elements of the mouse *car5b*

gene. As a control, the second B1 was removed or inverted to generate *1*SINE and *d*SINE constructs. In addition, the second B1 element was replaced by the first B1 element in both possible orientations, which led to the creation of a perfect *d*SINE (*pd*SINE) and perfect *i*SINE (*pi*SINE) that allows the formation of a perfect double-stranded structure. **b)** The new constructs were transfected into U2OS cells and dual luciferase assay was performed after 24hrs. Inverted B1 elements mildly reduce gene expression by 15%. However, *pi*SINE led to a reduction in gene expression by 30%. Asterisks indicate p-values calculated with student's T-test * $p<0.05$, ** $p<0.005$ and *** $p<0.0005$.

Sequence independent effects of inverted repeats on gene expression

Since both human- and mouse-derived SINEs seemed to repress luciferase activity in a complementarity-dependent manner we wondered how an unrelated double-stranded sequence would influence the reporter gene expression. We therefore generated an artificial double-stranded region. To do so, we duplicated and inverted a region of the 3' UTR of chicken pyruvate kinase (PK) and fused this 3'UTR to the luciferase reporter ORF. This 3'UTR has been widely used as it was shown to stabilize RNA and allow efficient translation of reporter constructs (Jantsch and Gall 1992; Peculis and Gall 1992). A stretch of about 285 bp was duplicated to mimic the average length of an *i*SINE (Figure 10). However, in contrast to naturally occurring basepairing *i*SINEs, such inverted artificial repeat forms a perfect double-stranded RNA and has less propensity to form internal structures than SINEs, which are shown to form internal stem-loops (Berger and Strub 2011). Additionally, truncations of the double-stranded inverted region were made by shortening the duplicated region to 150 or 50 nucleotides, respectively (Figure 10).

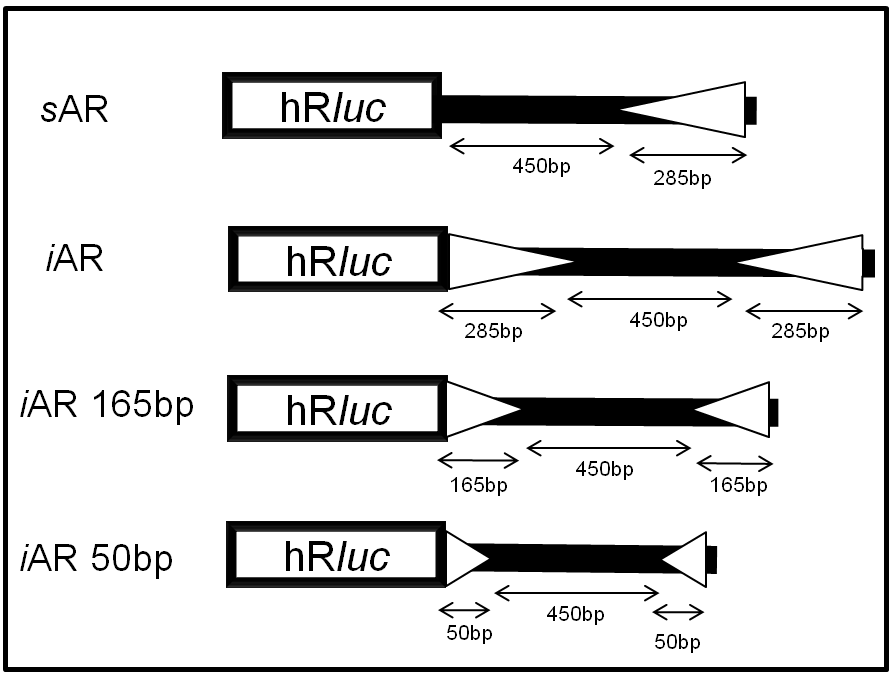

Figure 10. 3' UTRs with artificial double-stranded structures. The pyruvate kinase (PK) 3'UTR was inserted downstream of the reporter gene and used as a control (single artificial repeat, *s*AR). 285 nucleotides of the normal PK 3'UTR were duplicated and inserted next to the 3' end of UTR in inverted orientation (inverted artificial repeat, *i*AR). To test the importance of the length of the double-stranded structure, we made two constructs containing a 165 bp or 50 bp long double stranded RNA.

When expressed in wt MEFs, the construct carrying the inverted artificial repeat showed a strong (>30 fold) reduction in gene expression when compared to the normal PK 3'UTR (*s*AR) (Figure 11 a). Expression of the same construct in editing deficient *adar1*$^{-/-}$, *adar2*$^{-/-}$ MEFs also led to a strong repression of gene expression, again indicating that editing is not involved in this process (Figure 11b). We next tested whether reduction of the gene expression depends on the length of the double stranded region. Shortening of the artificial repeat to 165 nucleotides still led to strong reduction in the reporter gene expression. The shortening to a 50 bp long double stranded region also reduced the gene expression, however, the reduction was weaker than the one observed for the longer double-stranded structures (Figure 11c). In addition, our data showed that *i*AR strongly reduces the gene expression in both mouse and human cells. Nevertheless,

the observed difference in the strength of the reduction between cell types suggests that the cell type could play a role in this phenomenon (compare figure 11b and 11c).

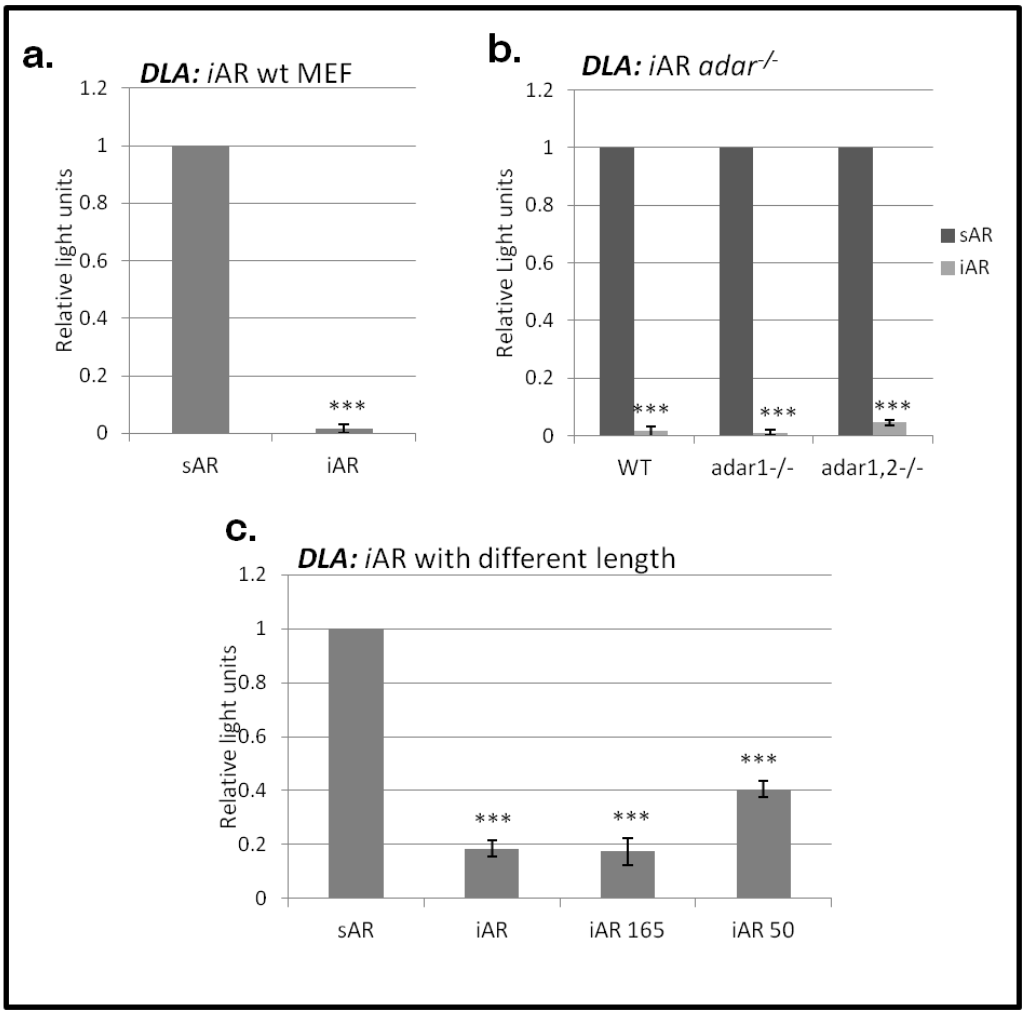

Figure 11. Reduction of gene expression caused by *i*SINEs seems sequence independent. Constructs with an artificial repeat in the 3'UTR were transfected in cells and a dual luciferase assay was done after 24-48 hrs. **a)** *i*AR containing constructs led to a strong reduction in gene expression (around 30 fold). **b)** In *adars* deficient MEF cells, the *i*AR was still able to reduce the gene expression indicating that such reduction is independent of ADARs. **c)** Artificial double-stranded constructs with different length were transfected in Hela cells and a dual luciferase assay was done after 24 hrs. *i*AR containing reporter gene showed more than 5 fold reduction which is a weaker effect in comparison with wt MEF. A 165 bp *i*AR led to a reduction in gene expression of more than 5 fold, but the 50 bp long *i*AR reduced the expression

of reporter gene only by 50%. Asterisks indicate p-values calculated with student's T-test * $p<0.05$, ** $p<0.005$ and *** $p<0.0005$.

iSINEs and artificial double-stranded RNAs fail to activate PKR

The observed reduction of gene expression clearly depended on the extent and length of the double-stranded RNA structures formed. Double-stranded RNA can act as a strong trigger to activate the double-stranded RNA dependent kinase PKR. PKR activation leads to phosphorylation of the translation factor eIF2-alpha thus eliciting a global repression of translation (Williams 1999; Elbarbary, Li et al. 2013). The fact that in the dual luciferase assay only the reporter carrying the *i*SINE in its 3'UTR was affected but not the reference reporter already argued against a general repression of translation.

Nonetheless, to test for potential PKR activation more directly, we performed western blots with a phospho-specific eIF2-α antibody and a pan-eIF2-α antibody on human U2OS cells transfected with the *Inadl i*SINE, *d*SINE, and *1*SINE construct. Densitometric scanning of the signals obtained from these blots indicated that the double-stranded RNA expressed from the reporter construct failed to activate PKR (Figure 12a) (This part of work was done by my colleague, Florian Huber).

In addition, we took advantage of mouse genetics and tested for the presence of *i*SINE-mediated repression in MEFs derived from mice lacking PKR activity (Abraham, Stojdl et al. 1999). In agreement with the western-blot experiments, *Nicn1 i*SINE constructs still showed reduced expression in *pkr*$^{-/-}$ cells when compared to constructs carrying a single SINE (Figure 12d). Surprisingly, even the *i*ARs with long, extended double-stranded RNA structures still showed a strong, at least 10-fold reduction in renilla expression in PKR deficient cells, just like in the corresponding wt MEFs, again indicating that PKR was not activated by the RNA produced from this construct (Figure 12e).

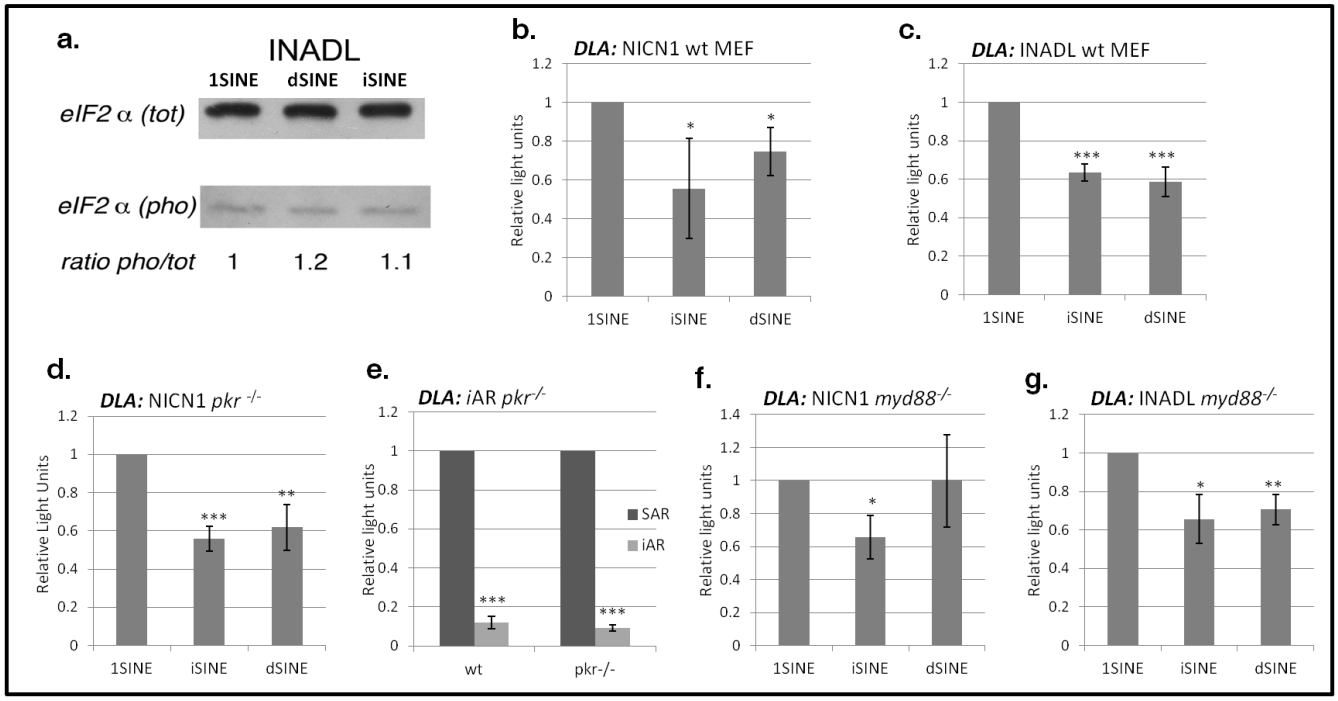

Figure 12. *i*SINEs do not activate double stranded RNA activated kinase PKR or signal through MyD88. *Inadl*, *Nicn* and constructs containing artificial double stranded regions were transfected in PKR or MyD88 deficient MEFs. **a)** A western blot of total cell lysates from cells transfected with the *Inadl* construct showed that phosphorylation of eIF2α did not increase in the presence of *i*SINE. Activation of PKR leads to eIF2α phosphorylation (This part of work has been done by my colleague, Florian Huber). Constructs harboring **b)** *Nicn1* and **c)** *Inadl* 3' UTRs transfected in wt MEF as control show a reduction in gene expression by 30-50%. **d)** DLA results showed *Nicn1 i*SINE and *d*SINE also reduced expression of the reporter gene in the absence of PKR by 40%. **e)** Perfect double stranded RNA reduced gene expression to 10% (as strongly as the reduction observed in wt MEFs). In *myd88*$^{-/-}$ MEF cells, **f)** *Nicn1* and **g)** *Inadl* *i*SINEs reduced the expression of reporter genes by 35%. Asterisks indicate p-values calculated with student's T-test * $p<0.05$, ** $p<0.005$ and *** $p<0.0005$.

iSINEs do not signal through MyD88

In 2012, Tarallo et al., showed that due to Dicer deficiency, Alu RNA will accumulate in the retinal pigmented epithelium. This leads to activation of the innate immune pathway and subsequently contributes to vision loss in Age Related Macular Degeneration

(AMD) (Tarallo, Hirano et al. 2012). Therefore, we tested the contribution of the innate immune pathway to *i*SINE mediated gene reduction. For this purpose, the expression of reporter genes was measured in *Myeloid differentiation primary response gene (88)* (MyD88) deficient MEFs. MyD88, a TLR adaptor to activate transcription factor NF-κB, is one of the key proteins in this cascade (Tarallo, Hirano et al. 2012). Our results showed that *i*SINE containing constructs are still able to reduce gene expression in *myd88*$^{-/-}$ cells, indicating that *i*SINEs embedded in the 3'UTR do not signal through MyD88 (figure 12f, g).

iSINE-containing RNAs show no nuclear retention

Previous studies had shown that mRNAs harboring SINEs in inverted orientation might get trapped in the nucleus upon editing, therefore preventing their efficient translation in the cytoplasm (Chen, DeCerbo et al. 2008). While we had already shown that the effect described here does not depend on RNA-editing, we still wanted to test whether the RNAs transcribed from our firefly reporter constructs would accumulate in the nucleus or become exported to the cytoplasm. To address this point, my colleague, Konstantin Licht, performed fluorescent in-situ hybridization using a firefly antisense probe to detect the RNA transcribed from single *1*SINE harboring constructs, duplicated *d*SINE and inverted *i*SINE constructs.

At the same time, the encoded protein was being detected using an antibody directed against the luciferase protein (Figure 13). This experiment showed clearly, that the firefly protein can be detected for all three constructs and that the RNA is readily exported to the cytoplasm. Thus, our data is in agreement with previous experiments that had reported efficient export of edited *i*SINE containing RNAs from the nucleus to the cytoplasm (Hundley, Krauchuk et al. 2008) (This work has been done by my colleague, Konstantin Licht).

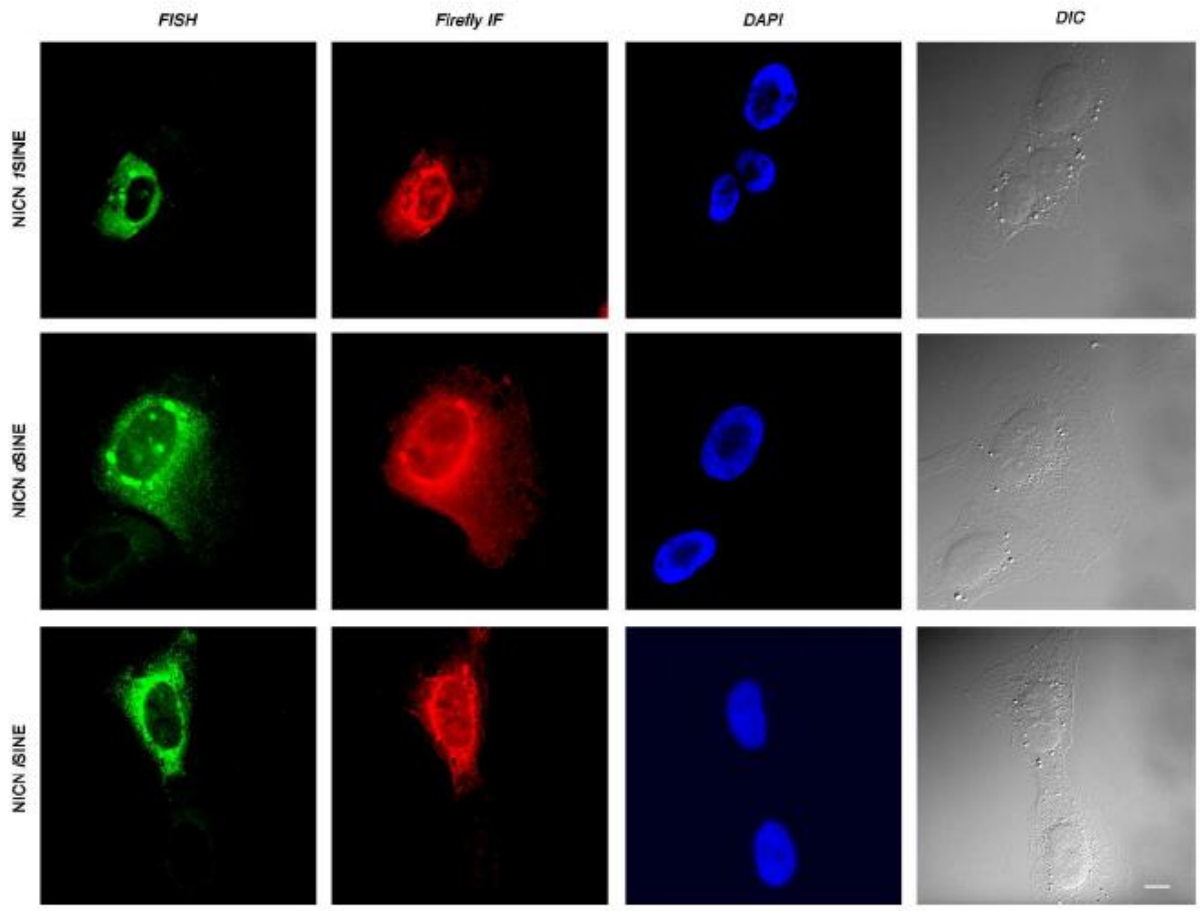

Figure 13. *i*SINEs do not lead to nuclear retention of RNAs. U2OS cells transfected with pmiRGLO harboring the *Nicn1 1*SINE, *d*SINE, or *i*SINE UTR downstream of the firefly luciferase ORF were stained for the expression of firefly luciferase. (red channel, Firefly IF). Localization of the RNA was determined by fluorescent in situ hybridization with a probe directed against firefly luciferase (green channel, FISH). While the FISH signal was generally weaker in cells transfected with the *i*SINE construct, no sign for nuclear accumulation of the RNA could be observed. Cell nuclei are stained with DAPI (blue channel) while the total cell is visualized by differential interference contrast (DIC). Scale bar=10μm.

Impact of SINEs on RNA levels

To gain more insight of the mechanism by which *i*SINEs lead to the observed reduction in gene expression, we evaluated the mRNA levels of the different *I*SINEs configurations. We compared the relative RNA levels of SINE harboring reporter mRNAs to control mRNAs using qPCR. Interestingly, RNA levels of a reporter harboring the *Nicn1 i*SINE were reduced in mouse cells when compared to single Alu-containing mRNAs (Figure 14a). The *d*SINE harboring mRNAs did not show any significant reduction in mRNA levels which suggests that *i*SINE and *d*SINE containing 3'UTRs may trigger two different mechanisms (It should be noted that the RNA samples for all experiments, except 5'+3'S *Znf708*, were taken from the same cells used for the luciferase assays shown before). Also the 3' UTR of the *Inadl* gene led to a reduction in mRNA levels in wt MEFs that perfectly matched to the reduction observed in luciferase activity (figure 14b). Here, the *d*SINE containing *Inadl* construct led to a reduction in RNA levels that was comparable to that observed for the luciferase activity (compare Figures 3b and 14b).

A reduction in the RNA levels could not only be observed in MEFs but also in human U2OS cells. The *Nicn1 i*SINE led to a reduction in mRNA levels by 20%, which is slightly weaker than the reduction observed in protein levels (compare figures 8a and 14c). *Inadl* SINEs containing constructs in U2OS cells followed the same trend as in wt MEFs (Figure 14d). This data indicates that also the reduction of the RNA levels expressed from constructs harboring different *i*SINEs in their UTRs is not species-specific.

Somewhat surprising was the finding that also the artificial, inverted repeat, derived from an inversion of the pyruvate kinase gene, showed a strong reduction in RNA levels that mimicked the reduction in luciferase expression (Figure 14e). In addition, perfect *i*SINEs from the short *Znf708* reduce mRNA expression (Figure 14f). The reduction in mRNA levels in all different *i*SINE and artificial inverted repeat containing reporter genes, suggests that a reduction in RNA levels might be the primary cause for the observed decrease in the reporter gene expression.

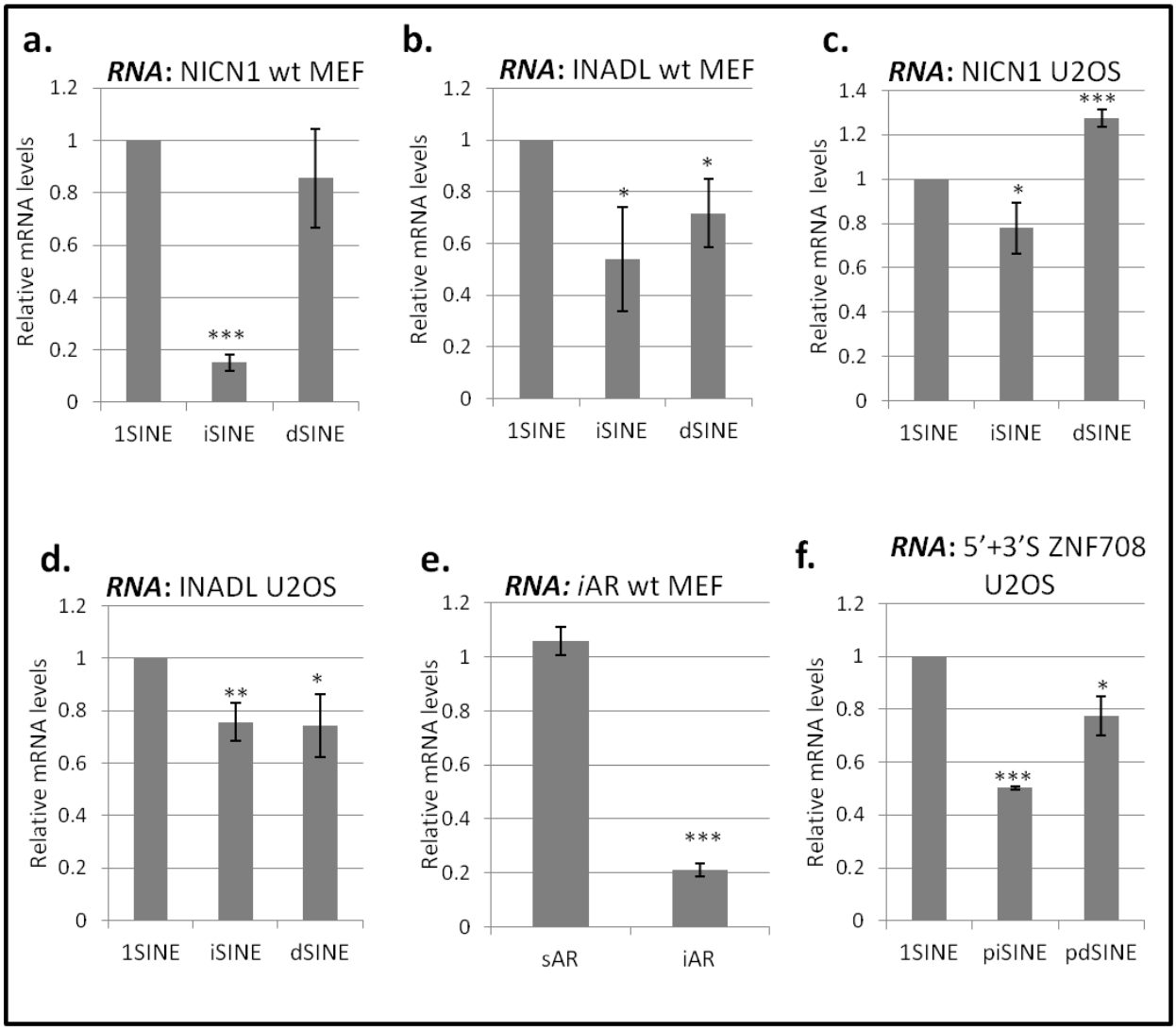

Figure 14. *i*SINEs lead to a reduction in RNA levels. Total RNA was extracted from transfected cells and cDNA was synthesized. Then, RNA levels were measured via qPCR of total cDNA. mRNA levels of luciferase-coding regions were normalized to the control reporter gene. The *i*SINE carrying **a)** *Nicn1* and **b)** *Inadl* constructs showed a significant reduction in mRNA levels by 85% and 47%, respectively. Only the *Inadl d*SINE construct showed a reduction in RNA levels. In human U2OS cells, **c)** *Nicn1* and **d)** *Inadl i*SINEs containing mRNAs showed a 20% reduction. **e)** Inverted artificial repeats (*i*AR) show a reduction in RNA levels by 80%, mimicking the reduction in luciferase activity. **f)** The 5'+3'S *ZNF708 pi*SINE which contains perfect SINEs, also led to a reduction in mRNA levels by 50%. Asterisks indicate p-values calculated with student's T-test * $p<0.05$, ** $p<0.005$ and *** $p<0.0005$.

Reduced RNA levels are not due to DICER or DROSHA activity

Based on our previous data we knew that *i*SINEs are highly edited. Since a double stranded structure is a requirement for editing we conclude that *i*SINE containing 3'UTRs form double-stranded structures. In addition, we observed a reduction in RNA levels. Based on these results, the siRNA machinery could be a potential mechanism leading to the observed RNA reduction. Since the double strand specific nuclease DICER1 is a key component in this pathway, we tested the effect of *i*SINEs in *dicer1*$^{-/-}$ MEFs. The results showed that even in the absence of DICER, the reduction in gene expression induced by *i*SINEs remained unchanged (Figure 15a, b). Likewise, the reduction in gene expression induced by *i*ARs was not affected by the presence or absence of DICER (Figure 15c).

DROSHA is another nuclear enzyme that is involved in the siRNA machinery and might be responsible for the observed RNA reduction. Moreover, it has been shown that DROSHA can bind to free Alus and downregulate their expression (Heras, Macias et al. 2013). Therefore, we tested the effect of *i*SINES in *drosha* knock down cells, where *drosha* mRNA levels were reduced by 50% (figure 15d). We obtained that *i*SINE-mediated gene repression was not inhibited in *drosha* knock down cells. At this stage we cannot exclude that residual DROSHA levels are sufficient to reduce the *i*SINE RNAs. Nevertheless, the reduction in the amount of DROSHA was sufficient to observe a dramatic increase in retrotransposition activity as determined by a standardized transposition assay (Heras, Macias et al. 2013)(data not shown).

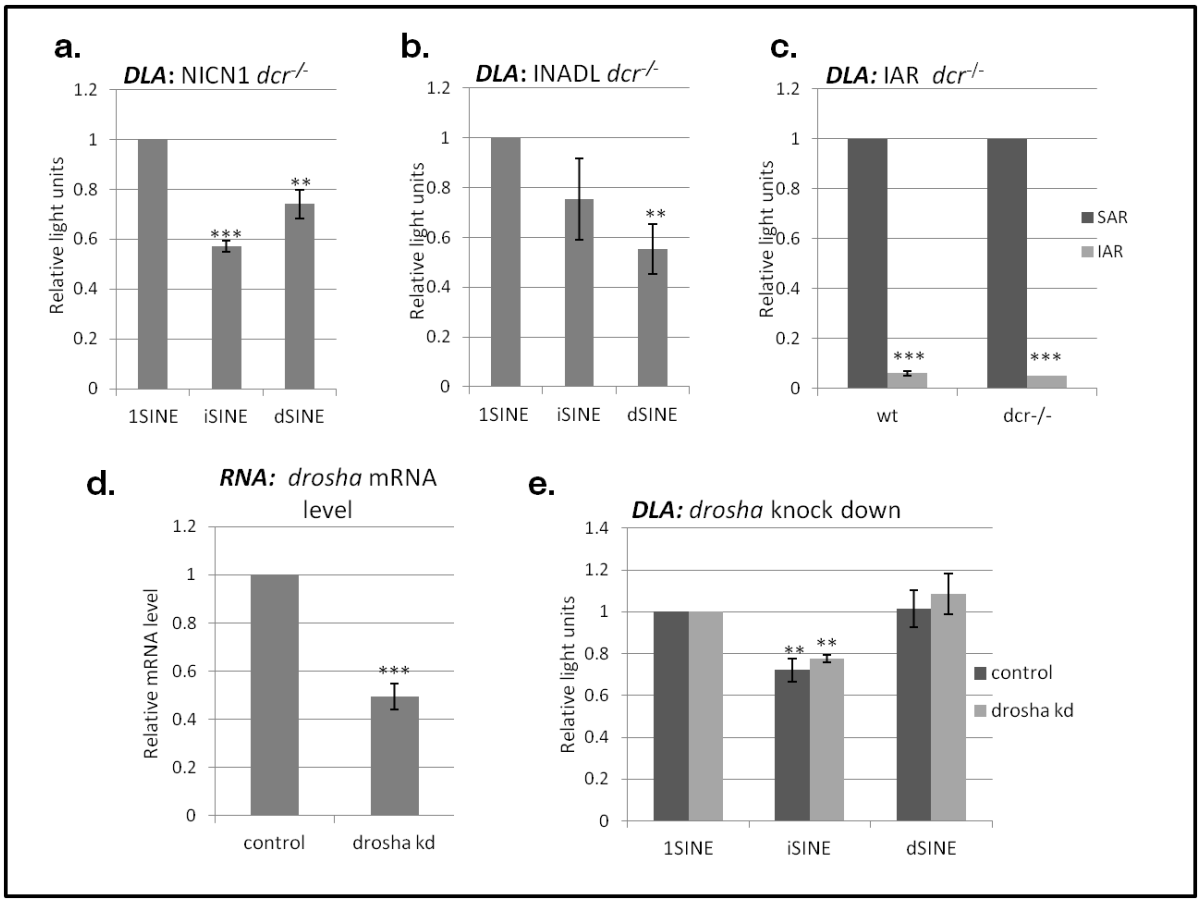

Figure 15. *i*SINE mediated gene repression is independent of DICER or DROSHA activity. Luciferease assays of *i*SINE containing constructs transfected in DICER or DROSHA deficient cells. **a, b)** In *dcr*$^{-/-}$ cells, the *Nicn1 i*SINE construct reduces gene expression by 40%, while in the *Inadl i*SINE construct the observed reduction in gene expression was not statistically significant. **c)** The *i*AR harboring reporter gene shows a 10 fold reduction in gene expression in the absence of DICER. **d)** Total RNA was extracted from *drosha* knock down cells and cDNA was synthesized. cDNA levels were determined by qPCR of total cDNA and normalized to the actin gene. The *Drosha* mRNA level was reduced by 50%. **e)** The *Nicn i*SINE construct reduced the expression of the reporter gene by 20% in the *drosha* knock down, similar to the reduction observed in control cells. Asterisks indicate p-values calculated with student's T-test * $p<0.05$, ** $p<0.005$ and *** $p<0.0005$.

Transcription, RNA-processing or RNA-stability?

Until now our experiments showed that the presence of inverted Alus leads to reduced RNA levels, which likely contribute to the observed reduction in luciferase protein production. Reduced RNA levels may be the result of reduced transcription, reduced or altered processing, or reduced RNA stability.

iSINE do not affect polyadenylation

Since *i*SINEs are located in 3'UTRs, they might affect poly-A polymerase and lead to a shortened poly-A tail. Shortened poly-A tails would also affect the stability of mRNAs and their translatability and might explain the observed reduction in mRNA and protein levels. Similarly, *i*SINEs might trigger the use of alternative polyadenylation sites. Different poly-A sites might also affect poly-A tail length and having an indirect effect on RNA-stability and translation. Based on these facts, we determined the position and length of poly-A tails in our expressed constructs using the splint-mediated poly-A tail measurement assay (Minasaki, Rudel et al. 2014) (Figure 16a).

The experiment showed that the poly-A signal used by *Nicn1 i*SINE containing mRNA is the same as control mRNA. In addition, the poly-A tail is of same length in *i*SINE, *d*SINE and *1*SINE containing mRNAs, in the range of 30-400 bp (Figure 16b). 5'+3'S *Znf708 pi*SINE, the second *i*SINE containing mRNA that we checked, showed the same pattern of polyadenylation as controls and the same poly-A signal is recognized in *i*SINE and the corresponding controls. The poly-A tails are equally long to the *1*SINE and *pd*SINE tails, in the range of 30-850 nucleotides (Figure 16c). To determine the position of the used poly-A signals, we performed sPAT assays on RNase H treated (RNase H+) RNAs. The resulting PCR products were sent for sequencing. Sequencing data showed that SINE-containing mRNAs utilized the vector-contained SV40 polyadenylation site located downstream of the inserted 3'UTRs (Figure 17a, b). Altogether, these data indicates that the *i*SINE-mediated reduction in gene expression is not due to shorter poly-A tails or alternative poly-A site usage.

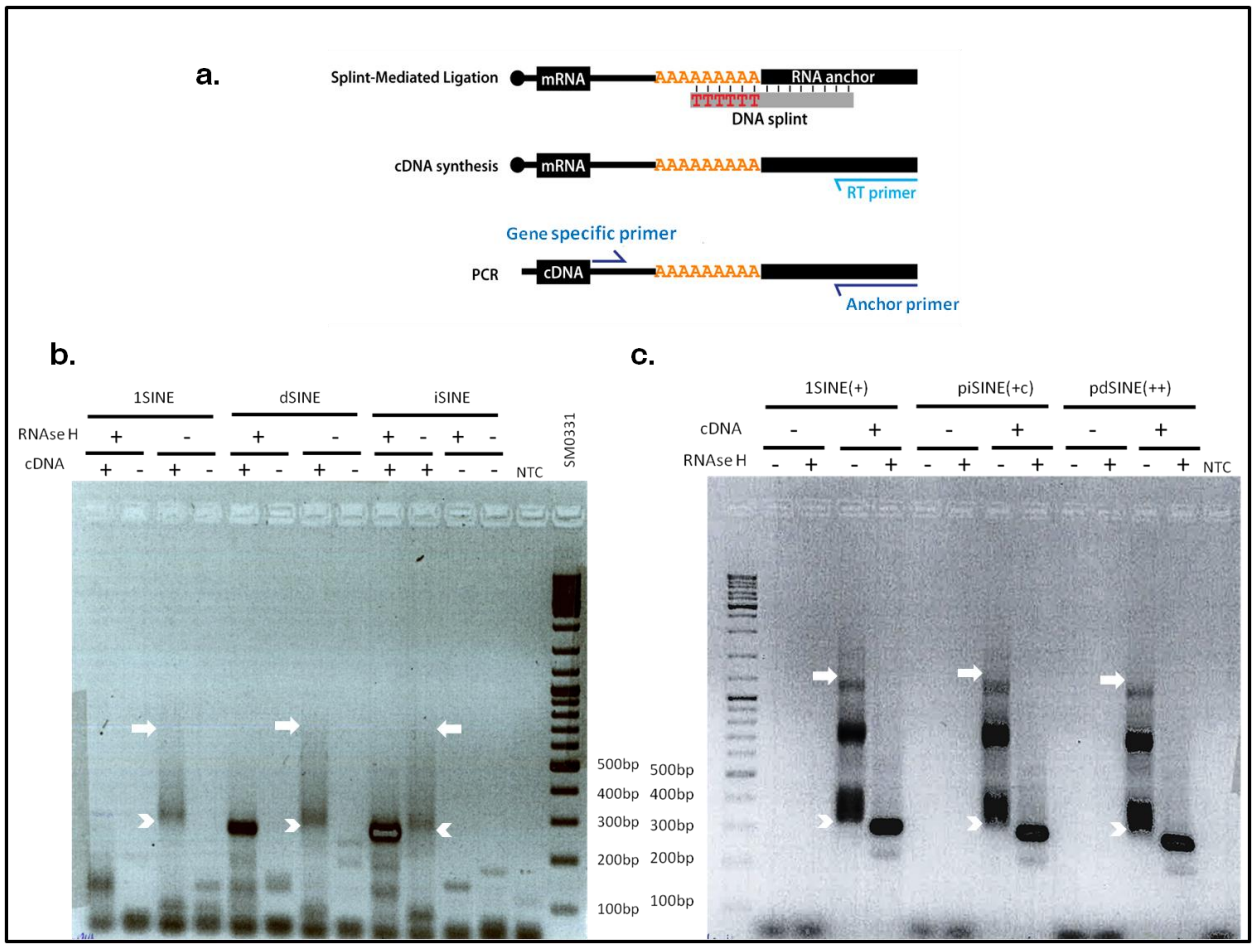

Figure 16. Length and position of the poly-A tail is identical in constructs harboring different SINE combinations. a) sPAT is a method for detecting poly-A length and polyadenlation sites in mRNAs. In this method total RNA is divided in two samples; one sample is treated with RNase H in the presence of oligo dT to remove the mRNA's poly-A tail and the other mRNA sample is kept intact. Then, an RNA adaptor is ligated to the mRNA poly-A tail with the help of a short, complementary splint DNA. An oligo complementary to the adaptor is used to synthesize a cDNA. Next, using a gene specific primer, the poly-A containing fragment is amplified. Presence or absence of RNase H and oligo-dT allows detection of the exact position of the poly-A site. **b)** The *Nicn1 i*SINE construct utilizes the same polyadenylation signal as the control mRNAs, *d*SINE and *1*SINE (compare RNase H+ lanes). In addition, the length of the poly-A tail in *i*SINE containing mRNA is as long as the one in the corresponding controls, in the range of 30-400 nucleotides (compare RNase H- lanes). **c)** Similarly, same poly-A site is used in all 5'+3'S *Znf708* SINE containing mRNAs (RNase H+ lanes). In addition, the *pi*SINE shows the same pattern of polyadenylation as the corresponding controls. Each band in RNAse H- lanes

resembles a prominent PCR product in the range of 30 to 800 nucleotides (RNAse H-lanes) (RNase H+/-: with or without RNase H, cDNA+/-: with or without reverse transcriptase, NTC: non template control in PCR reaction, white arrow and white arrow head show the range of poly-A tail)

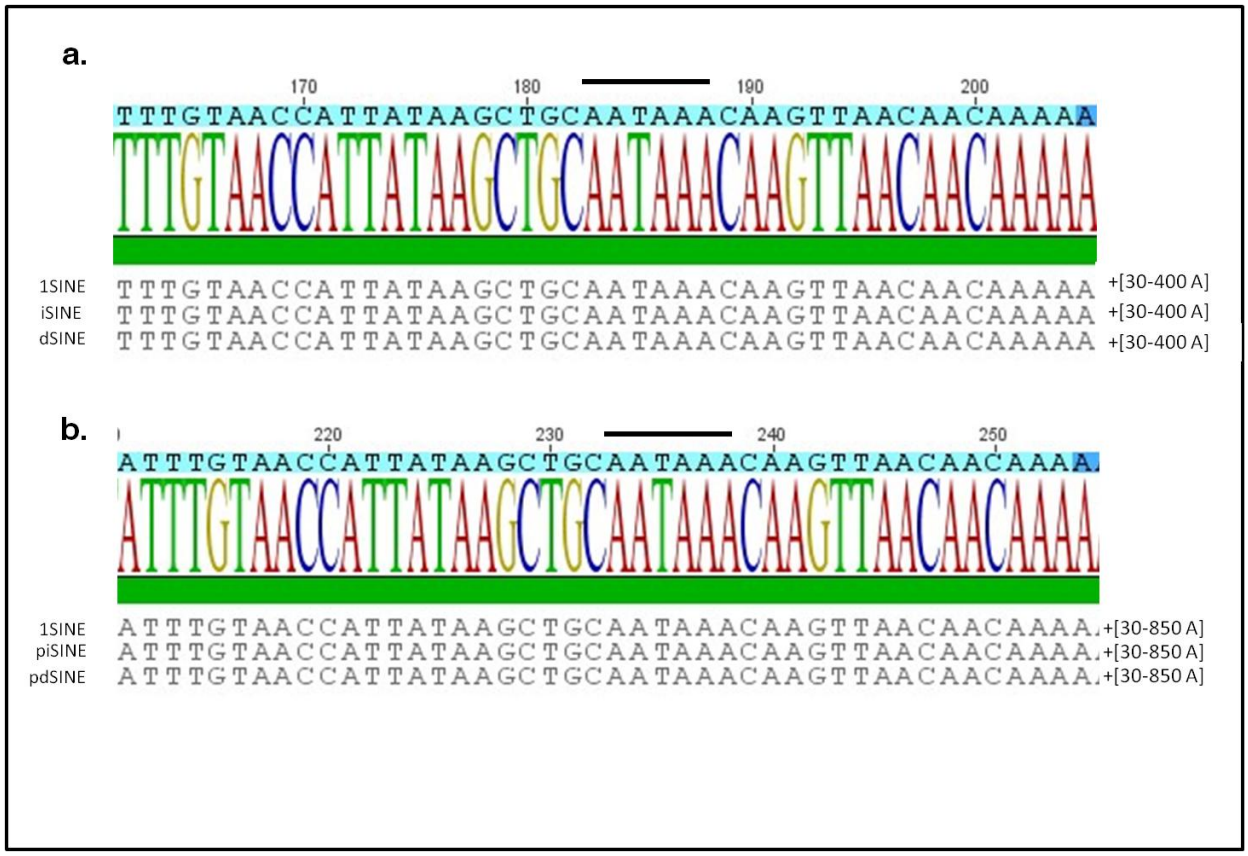

Figure 17. Polyadenylation site of poly A in *Nicn1* and *Znf708*. RNase H treated (RNase H+) PCR products from sPAT assay were extracted from gels and sequenced **a)** *Nicn1* and **b)** 5'+3'S *Znf708 pi*SINE used the SV40 polyadenylation signal of the vector, which was also used by the corresponding controls.

Stability of iSINE mRNA

Till now our data showed that *i*SINE in 3'UTR led to a reduction in mRNA levels. The observed reduction could be due to reduced mRNA stability and/or a low transcription rate. To test whether *i*SINE-containing RNAs have a reduced stability, we compared mRNA half-life of *i*SINE and a control RNA. For this purpose, mRNA transcription was blocked using Actinomycin D treatment. Subsequently, RNAs were collected in regular time intervals and mRNA levels were determined by real-time qPCR of cDNAs (Bensaude 2011).

Our data showed that the *Nicn1 i*SINE mRNA is degraded as fast as the *1*SINE mRNA (Figure 18). This observation suggested that the reduction of mRNA levels is not caused by low mRNA stability. In turn, this suggests that low transcriptional rates might be responsible for the observed reduced RNA levels and are consequently studied in the next step.

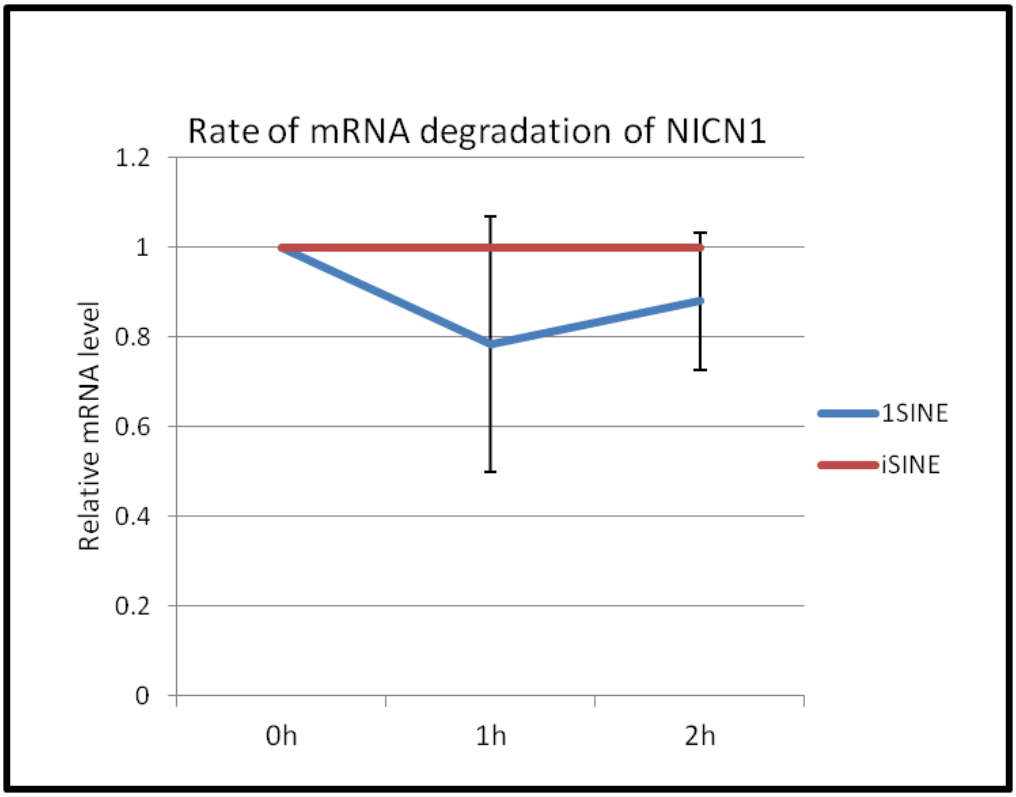

Figure 18, *i*SINE containing mRNAs are degraded as fast as *1*SINE mRNAs. mRNA transcription was blocked using Actinomycin D. Subsequently, mRNAs were collected 0,1 and 2 hrs after transcriptional inhibition and mRNA levels were determined by qPCR. The results showed *i*SINE containing mRNA is degraded as fast as the corresponding control mRNA, *1*SINE.

Transcriptional Rate

One possibility to explain reduced RNA-levels despite a comparable RNA stability could be different transcriptional rates of the constructs containing single or inverted SINEs. A reduction in transcriptional rates can be explained by a lower Pol-II occupancy. In order to test this possibility, we performed RNA Pol-II-chromatin immunoprecipitation assay (ChIP) to determine the Pol-II density along the DNA. Consequently, DNA coprecipitated with a Pol-II antibody was quantified by qPCR, using amplicons for separate regions of the reporter genes (Figure 19a). The data were always normalized to the first amplicon, which amplified a coding region near the start codon. Then the outcomes were compared to the *1*SINE containing reporter gene as a control.

The results of these ChIP experiments showed that Pol-II density was markedly decreased in the 3'UTR, near the poly-A signal in *i*SINE containing genes compared to the respective control (Figure 19b, c). The reduction in Pol-II density was observed for both *Nicn1 i*SINE and 5'+3'S *ZNF708 pi*SINE containing genes. Thus, the location of an *i*SINE in the 3'UTR seems to interfere in transcriptional elongation. This could be due to premature termination or transcriptional stalling. Understanding the exact mechanism underlying this phenomenon will need further investigation.

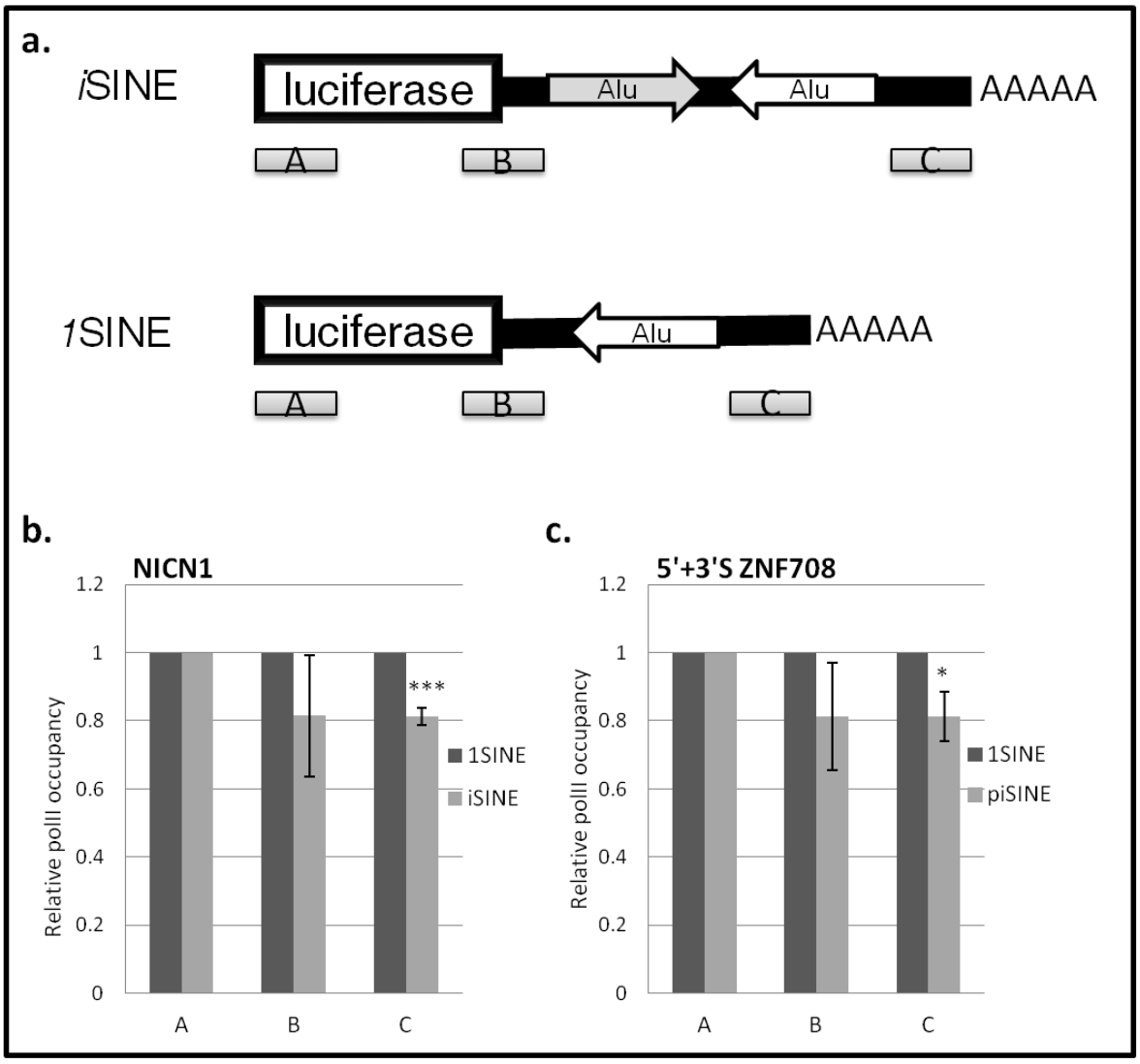

Figure 19, *i*SINEs can interfere with Pol-II occupancy. To compare the rate of mRNA transcription along the reporter gene, Pol-II density was measured using polymerase-II antibodies for chromatin immunoprecipitation (ChIP). **a)** The polII-immunoprecipitated DNA was analyzed by real time PCR. Three different regions of the reporter gene, including early coding region (A), stop codon (B) and near poly-A signal (C), were amplified. The amplicons are shown as grey boxes. tRNA phenylalanine was amplified as a negative control to check the efficiency of the pulldown (In addition, a non-specific IgG fraction was used as an unspecific antibody for pulldown). Reporters harboring the **b)** *Nicn1 i*SINE or the **c)** 5'+3'S *Znf708* with a perfect *i*SINE were transfected in U2OS cells and the Pol-II density along the genes was measured. The Pol-II occupancy is significantly reduced after the stop codon. The data was normalized to the first amplicon and then to the *1*SINE containing gene.

Bioinformatic analysis

Our data clearly showed that *i*SINEs can reduce mRNA levels, possibly by interfering with Pol-II activity. To study the effect of SINEs at a genome-wide scale, the expression levels of SINE containing transcripts were analyzed in 15 cell lines analyzed by the ENCODE consortium. This bioinformatic approach was performed in collaboration with the research group of Ivo Hofacker. The data showed that the expression levels of the *i*SINE containing transcripts are significantly lower than the *1*SINE containing transcripts (p-value: 1.53448e-12). Interestingly, the *i*SINE containing genes are also having a significantly lower expression than the *d*SINE containing genes (p-value: 3.29244e-08). This observation clearly shows that *i*SINE can regulate the gene expression of many transcripts and not only of the transcripts chosen in our assays (Figure 20).

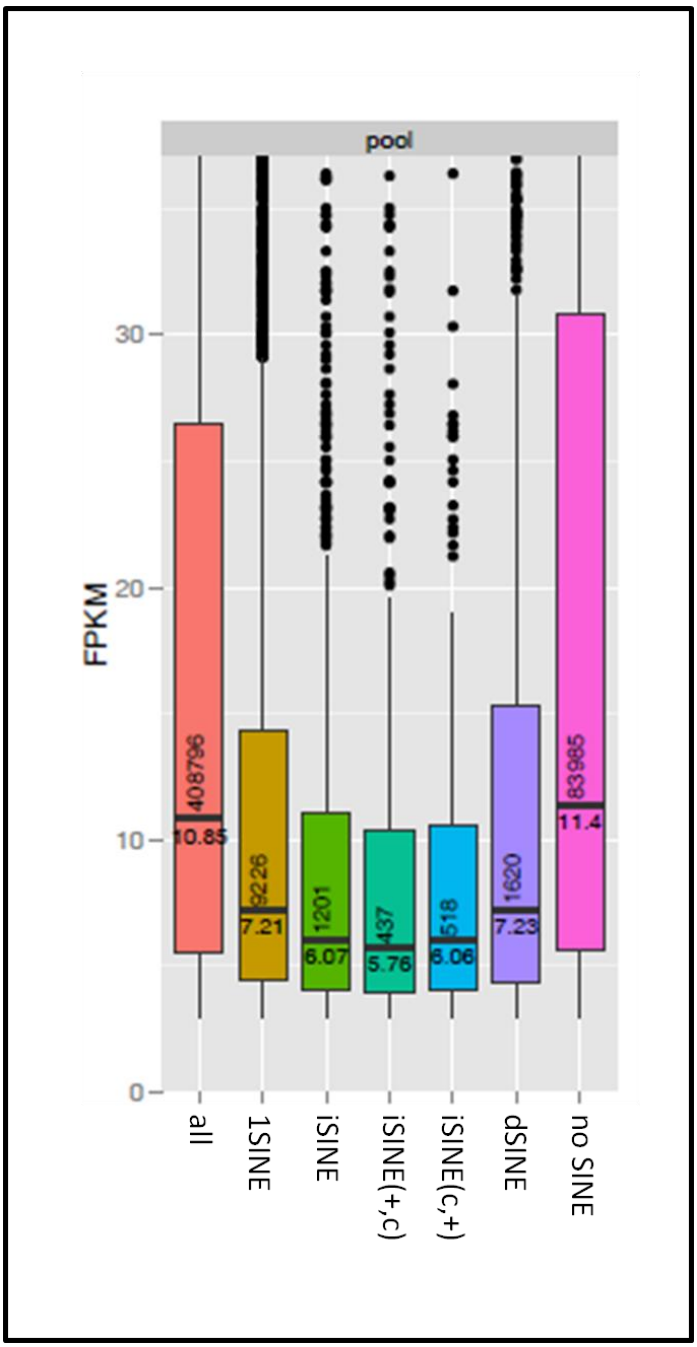

Figure 20. Expression levels of SINE-containing transcripts in 15 ENCODE cell lines (pooled). FPKM values were calculated for transcripts with *1*SINE, *d*SINE, *i*SINE in both configurations (+,c and c,+), pooled of transcripts without any SINEs (noSINE) and all transcripts (all). The expression levels of *i*SINE containing transcripts are significantly lower than *1*SINE and *d*SINE containing transcripts.

Discussion

Discussion

SINE elements are the numerically most abundant class of transposable elements in the genomes of higher metazoa (Britten and Kohne 1968; Lander, Consortium et al. 2001). They have a huge impact on the genomic landscape in these organisms and drive evolutionary development. Most importantly, SINEs are not distributed randomly in the genome, but are frequently found within genes and transcribed as a part of intronic sequences, UTRs, or even exons (Versteeg, van Schaik et al. 2003). For a long time, SINE elements have been viewed as junk DNA, but recently it has been shown that SINEs have an impact on gene expression. They can influence gene expression by modulating mRNA transcription, localization, translation and degradation (Gong and Maquat 2011; Ichiyanagi, Li et al. 2011; Ricci, Kucukural et al. 2014).

Multiple factors in 3' UTRs affect gene expression

Recent studies have demonstrated that the presence of inverted SINEs (*i*SINEs) in the transcribed regions of genes can have a significant influence on gene expression (Chen, DeCerbo et al. 2008; Gong and Maquat 2011; Capshew, Dusenbury et al. 2012). We also observed that *i*SINEs in 3'UTRs of mRNAs can reduce gene expression. However, not all *i*SINE containing 3'UTRs reduce gene expression with the same extent. To pinpoint the most important features in 3'UTRs that can modulate the extent of repression we compared several *i*SINE containing 3' UTRs.

A previous study has suggested that the relative position of *i*SINEs in the 3'UTR can influence the impact of SINEs on the gene expression (Capshew, Dusenbury et al. 2012). However shortening the distances between the *i*SINE and the stop codon or the poly-A signal in *Znf708* constructs did not completely diminish the effect of SINEs. Besides the location of SINEs, there may be other unknown factors in the 3'UTRs modulating the strength of the reduction. For instance, adjacent sequences could interfere with the folding of *i*SINEs into double-stranded structures. Moreover, protein

factors binding to flanking sequences may also interfere with gene expression and their deletion might show an effect in specific cases but not necessarily in a global manner. The observed effect upon deleting regions up- and downstream of the inverted SINEs as seen by the Hundley lab (Capshew, Dusenbury et al. 2012) may therefore be specific to the UTRs tested in their study. We also cannot exclude that differences in adjacent vector sequences such as poly A signals, promoters, reporters, or the cellular system can affect the impact of sequences in the 3' UTR on gene expression.

We next checked the influence of the extent of the double-strandedness by comparing different constructs harboring short mouse SINEs, longer human SINEs, identical SINEs from both species (100% double-stranded) and even artificial repeats of different length. Our data consistently shows that the length and extent of the predicted double-stranded RNA regions influence the strength of the repression in gene expression. In addition, a strong reduction caused by artificial repeats demonstrates that *i*SINE mediate reduction of gene expression is sequence independent. However the observed reduction induced by artificial repeats is higher when compared to the reduction induced by double-stranded structures formed by *i*SINEs. Several reasons may account for the different strengths in the reduction of gene expression. First, SINEs have the tendency to form double stranded structures themselves. Therefore, the folding kinetics of inverted SINEs will always be influenced by the folding of individual SINEs. Some 3' UTRS may even get caught in a folding trap and thus fail to form double stranded structures between the SINEs. Second, the homology of different SINEs lies between 70 and 80%, depending on the type of SINE investigated. Therefore, two SINEs, even if they belong to the same subfamily will rarely form a perfect double-stranded structure and hence will always contain bulges and unpaired regions. Lastly, we cannot exclude that sequence differences between basepaired Alu sequences and artificial vector sequences will recruit different factors and therefore trigger different cellular mechanisms that ultimately lead to differences in the reduction of gene expression. Nevertheless, our data show that the formation of a double-stranded structure is essential for *i*SINE mediate gene repression and the strength of the reduction is affected by the extent of the double-strandedness.

In our study we show that *i*SINEs derived from the human genome reduce the gene expression in human and also in mouse cell lines. Interestingly, also mouse SINE elements in inverted orientation reduced the gene expression in human cells. We therefore conclude that *i*SINE mediated gene repression is a conserved mechanism. In agreement with this, is the observation that double-stranded 3'UTR structures formed by inverted repeats decrease the expression of endogenous *C. elegans* genes (Hundley, Krauchuk et al. 2008).

iSINE-mediated reduction of gene expression is ADAR independent

The mechanisms leading to reduced gene expression by *i*SINE containing UTRs are not well understood at this point. Inverted SINEs, for instance, are heavily edited since they likely form double-stranded structures which are recognized as substrates by members of the ADAR family of RNA-editing proteins (Athanasiadis, Rich et al. 2004; Levanon, Eisenberg et al. 2004). Therefore, inverted SINEs have been suggested to repress gene expression by nuclear binding and retention of edited, inosine-containing RNAs (Prasanth, Prasanth et al. 2005; Chen, DeCerbo et al. 2008). However, more recently it was shown that RNAs containing inverted SINEs can get exported to the cytoplasm but are translationally repressed, even upon knock-down of the RNA editing enzyme ADAR2 (Hundley, Krauchuk et al. 2008; Capshew, Dusenbury et al. 2012).

Our experiments presented here take these findings a step further using a clear genetic approach. We show that inverted SINEs can repress gene expression even in cells lacking both catalytically active ADAR enzymes thus demonstrating clearly that editing is not a prerequisite for this phenomenon. In addition, our fluorescent in-situ hybridization (FISH) data showed that *i*SINE containing mRNAs are equally distributed in the cell and do not show sub-cellular localization or nuclear retention. These data are in agreement with previous experiments that had reported efficient export of edited *i*SINE containing RNAs from the nucleus to the cytoplasm (Hundley, Krauchuk et al. 2008).

Double-stranded (ds) RNA binding proteins, Staufen1 and PKR, do not involve in the observed reduction

*i*SINEs represent a major group of binding targets for Staufen1 (Ricci, Kucukural et al. 2014). In addition, it has been shown that intermolecular *i*SINEs (a double-stranded RNA structure that is formed between two SINEs from two different RNAs) can trigger Staufen1 mediated decay (SMD) and subsequently block translation (Gong and Maquat 2011; Gong, Tang et al. 2013). Based on these findings, Staufen1 could be a candidate protein that is involved in *i*SINE-mediated reduction of gene expression. However, our data showed that Staufen1 has no effect on the expression of *i*SINE containing genes. It was also shown that overexpression of Staufen1 could increase protein production of *i*SINE containing genes (Elbarbary, Li et al. 2013) and they proposed that Staufen1–binding might mask double-stranded RNA structures to prevent translational inhibition mediated by PKR.

In agreement with our observations, an RNA-Seq experiment of Staufen1 associated RNAs did not show any significant changes in ribosomal association of *i*SINE containing mRNA in the presence or absence of Staufen1 and only a weak increase in nucleocytoplasmic export of the respective mRNAs was observed (Ricci et al., 2014). In general, no clear effect of Staufen-1 was found on gene expression of *i*SINE mRNAs in this latter study, indicating that even though *i*SINEs have been described as Staufen1 targets, the consequence of Staufen1 binding to *i*SINEs is subtle and requires further analysis.

PKR is another dsRNA binding protein, which has been shown to interact with *i*SINEs (Elbarbary, Li et al. 2013). In addition, it has been proposed that PKR is activated by binding to the *i*SINE containing 3'UTRs and blocking global translation in the early phase of mitosis (Kim, Lee et al. 2014). However, our experiments show that *i*SINEs are able to interfere with gene expression independent of PKR.

iSINEs reduce RNA levels

Interestingly, in our different assays we observed that the reduction in the reporter gene expression levels correlates well with the reduction in RNA levels induced by *i*SINEs. In addition, our analysis of ENCODE expression data showed that *i*SINE containing mRNAs have the lowest average expression levels in the transcriptome, lower than comparable RNAs carrying tandemly arranged SINEs (*d*SINE). Thus, besides the previously reported nuclear retention and translational repression induced by *i*SINEs, a third mechanism leading to a reduction in RNA-levels appears to exist.

An obvious pathway by which double-stranded RNA might be degraded would be the RNA*i* pathway. Nevertheless, even in the absence of DICER, the gene expression of *i*SINE containing UTRs is efficiently repressed and mRNA levels are reduced significantly. Likewise, we tested whether the nuclear enzyme Drosha might also be involved in this process. The use of lentiviral-delivered shRNA-mediated knockdown of Drosha led to a reduction of Drosha levels by 50%. However, this reduction in Drosha level did not inhibit the observed *i*SINE-mediated gene repression. At this stage we cannot exclude that residual Drosha levels are sufficient to reduce RNA levels of *i*SINE RNAs or whether Drosha is not involved in the described phenomenon.

To get some insight in the mechanism(s) leading to a reduction in RNA levels, we checked the *i*SINE dependent effect on polyadenylation. Since *i*SINEs are located in the 3'UTR of the mRNA, they could interfere with poly-A polymerase activity and lead to shorter poly-A tails or alternative poly-A site usage. In both cases, the mRNA stability could be affected. However our data indicate that *i*SINE do not affect the poly-A length nor the poly-A site usage. Alternatively, mRNA levels could be reduced because the mRNA half-life of *iSINE*-containing transcripts is changed. To put this hypothesis to test we used Actinomycin D treatment followed by RNA quantification. The experiments indicated that the RNA half-life is not altered between constructs harboring or lacking *i*SINEs.

Transcriptional interference of free Alu elements had been described previously (Mariner, Walters et al. 2008). Therefore we tested whether the transcriptional rate of

*i*SINE harboring constructs is reduced due to lower occupancy of Pol-II. We have evaluated this possibility by using Pol-II chromatin immunoprecipitation (ChIP). Our ChIP data showed that the presence of *i*SINEs leads to a reduction in Pol-II density distal to *i*SINEs in the 3'UTR. Pol-II density is reduced significantly near the poly-A signal. This could explain the observed reduction in mRNA levels and suggests that *i*SINE could modulate mRNA transcription. How *i*SINE and Pol-II interact and interfere in mRNA transcription is not clear and needs to be investigated further.

Nonetheless, it was shown that RNA stem loops in the promoter regions can regulate transcriptional elongation. The microprocessor complex, Drosha-Dgcr8, is recruited by the RNA stem loop and can initiate mRNA cleavage, which leads to uncapped transcripts. The uncapped RNA serves as signal for mRNA degradation (Wagschal, Rousset et al. 2012). Therefore, one could speculate that *i*SINEs trigger a similar pathway by recruiting an unknown dsRNA binding protein.

Moreover, it has also been shown that LINEs embedded in transcription units can repress transcription. LINE-mediated transcriptional repression is mediated by a block in transcriptional elongation (Han, Szak et al. 2004). It remains to be determined whether the observed reduction in RNA levels of *i*SINE containing constructs underlies a similar phenomenon.

Lastly, double-stranded structures in RNAs have also been shown to interact with the bacterial RNA polymerase exit channel, prolong RNA Polymerase pausing and consequently reduce the transcriptional elongation rate (Toulokhonov, Artsimovitch et al. 2001). Importantly, it has been observed that the crystal structure of RNA polymerase and Pol-II are similar in the region that interacts with the transcriptional bubble (Dangkulwanich, Ishibashi et al. 2014). Therefore, the long double-stranded structures formed by inverted SINEs might act in a comparable manner leading to an increase in Pol-II pausing. Clearly this hypothesis will need to be tested by different methods such as GRO-Seq which allows a detection of Pol-II density *in* vivo (Jonkers, Kwak et al. 2014).

Taken together, the recent studies have shown that *i*SINEs as well as single SINEs can affect gene expression by influencing mRNA transcription, transport and translation. Several RNA binding proteins such as ADAR1, p54nrb, Staufen1 and PKR have been shown to interact with *i*SINEs and subsequently affect mRNA modification, nuclear retention, mRNA transport and translational repression, respectively (Chen and Carmichael 2008; Elbarbary, Li et al. 2013; Daniel, Silberberg et al. 2014; Kim, Lee et al. 2014). Actually, *i*SINEs could act as regulatory elements in the mRNA by providing a double-stranded RNA structure, which serves as a platform for dsRNA binding proteins. Based on the availability of double-stranded RNA binding proteins in the cell, *i*SINEs might recruit them and trigger different mechanisms that subsequently modulate gene expression. Besides the previously reported *i*SINE regulatory effects, our study revealed a new role for *i*SINEs as a transcriptional modulator, which directly or indirectly interferes with Pol-II activity.

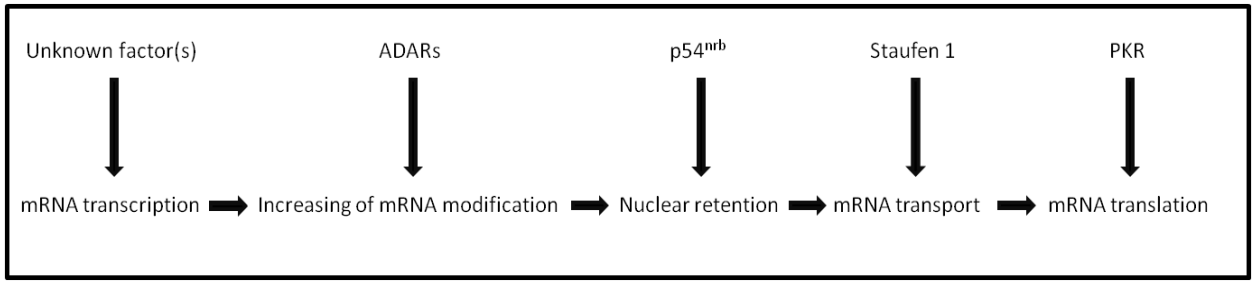

Figure1. The effects of *i*SINEs on the gene expression through the interaction with different double-stranded RNA binding proteins or unknown factors

Global effect of SINEs on gene expression

When analyzing the ENCODE expression data, the most dramatic differences in gene expression were observed between RNAs harboring and RNAs lacking any SINEs. The presence of a single SINE already reduces gene expression dramatically and this may be mediated by the above mentioned interference of SINE sequences with Pol-II activity (Ponicsan, Kugel et al. 2010). In our study we could also show that not only *i*SINEs but also *d*SINEs had a consistent but less dramatic effect on gene repression. While the

*Inadl d*SINE constructs consistently repressed gene expression at significant levels, the *Nicn d*SINE had only a very moderate effect on gene expression in MEFs. Consistently, the *d*SINE construct of *Nicn* had only a minor effect on RNA levels. Taken together, this suggests that while *i*SINE and *d*SINE constructs both show an effect on luciferase protein levels, the underlying molecular mechanisms might be different. Reduced RNA levels may be the primary effect of *i*SINEs while other steps in gene expression may be affected by *d*SINEs.

Our finding, that artificial, double-stranded RNAs can strongly repress gene expression independent of PKR and Dicer activity was surprising at first sight. However, a recent study nicely demonstrates that a transgene forming a perfect double-stranded structure in mouse also fails to trigger a significant PKR or Dicer response (Nejepinska, Malik et al. 2012). Thus, the mechanism leading to reduced RNA levels and subsequent reduced protein levels of perfect artificial inverted repeats harboring transgenes remains to be determined. It is clear, however, that neither RNA-editing nor PKR activation is involved in the observed repression in gene expression mediated by SINEs.

Supplementary Figures

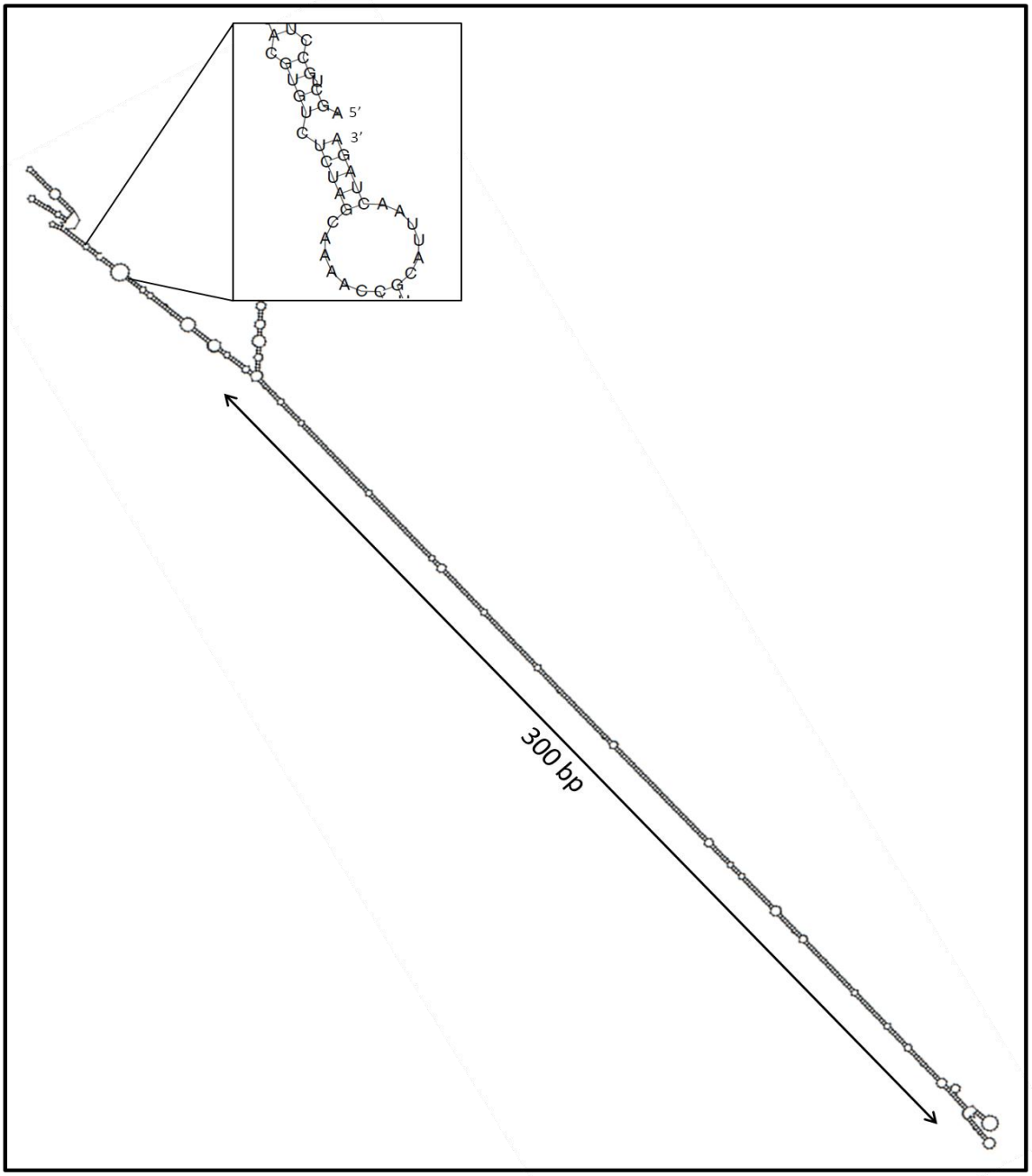

Figure 1. Secondary structure of *Nicn1 i*SINEs containing 3'UTR. *i*SINEs make a 300 bp double-stranded structure.

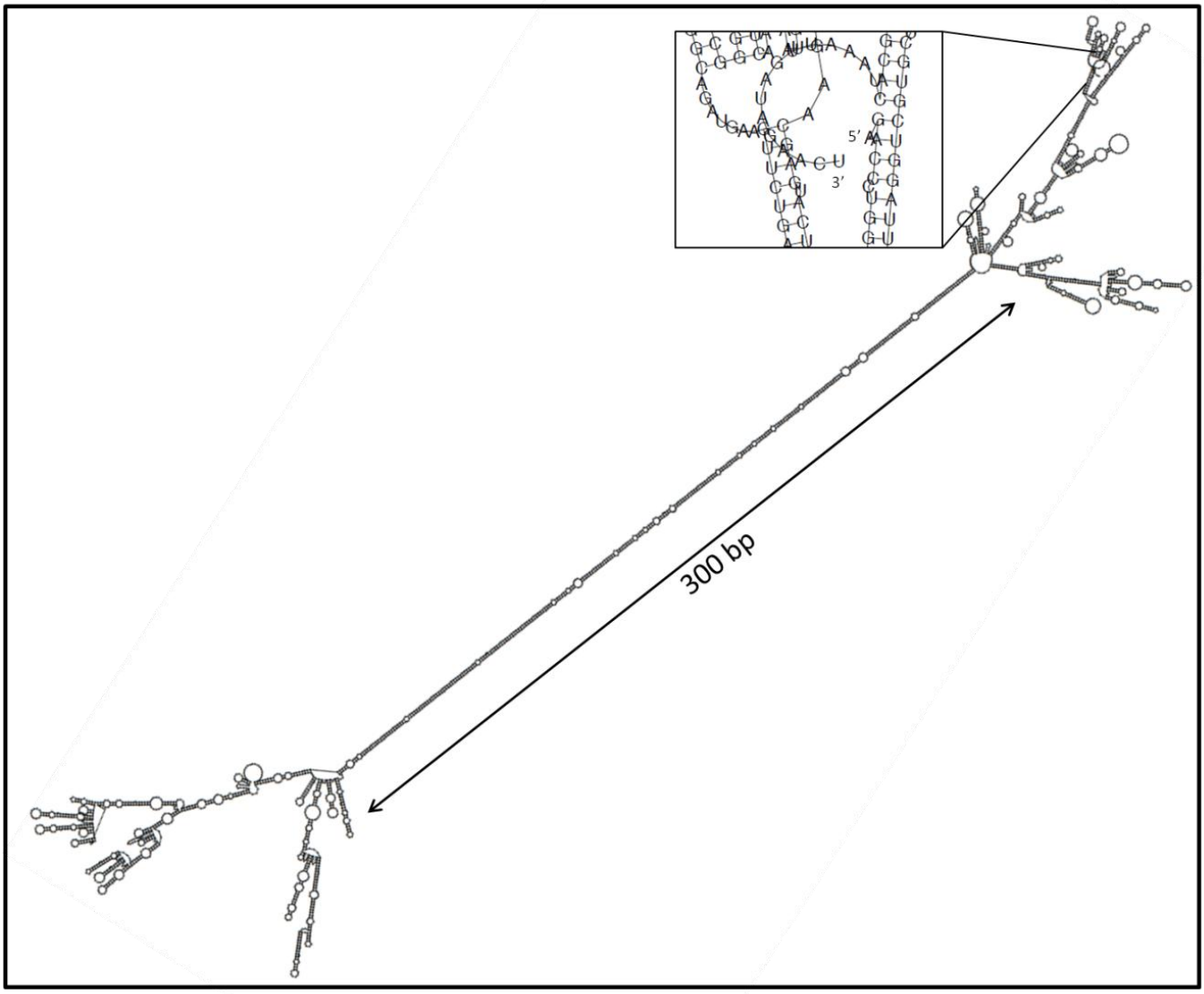

Figure 2. Secondary structure of InadI iSINE containing 3'UTR. *i*SINEs form a 300 bp double-stranded structure.

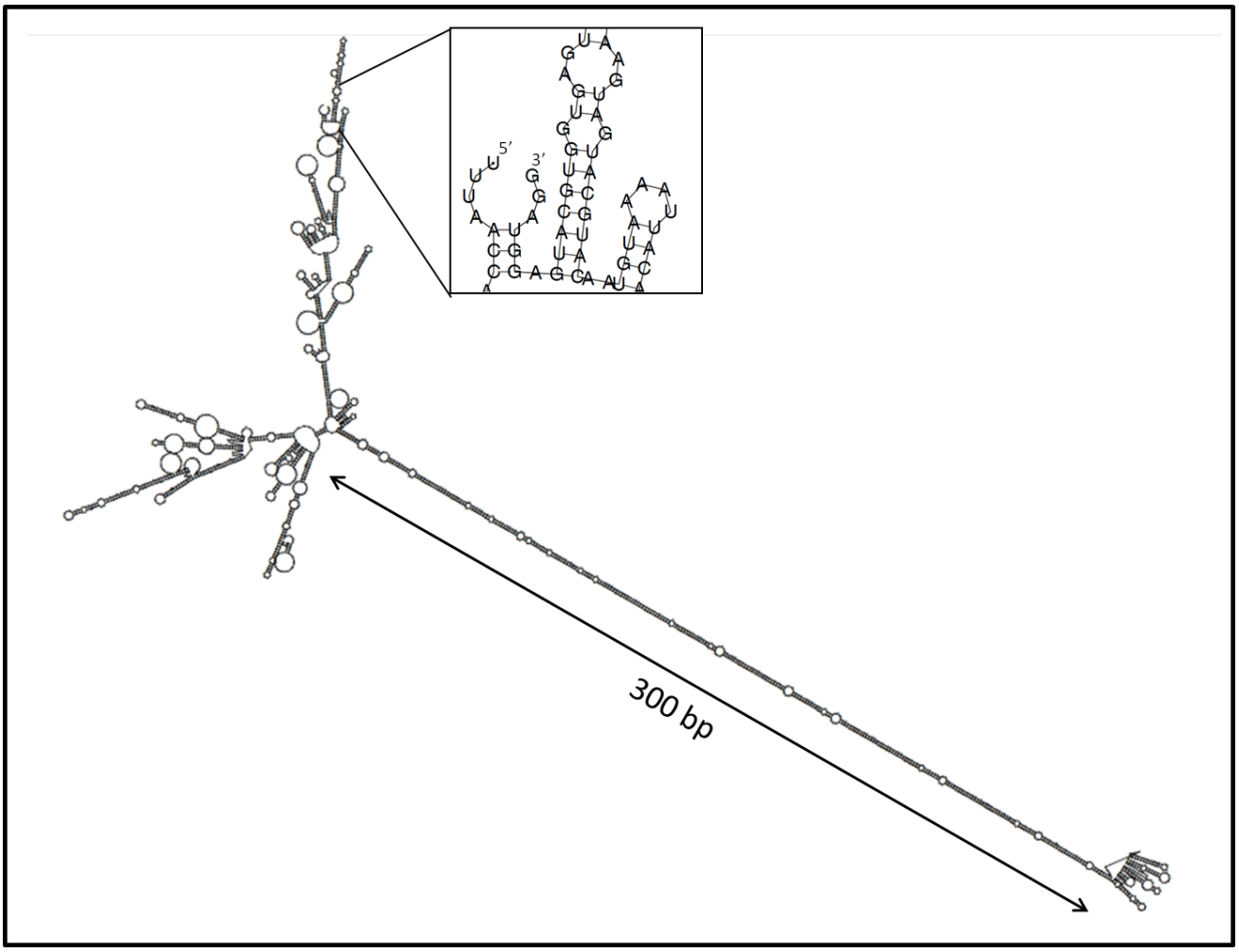

Figure 3. Secondary structure of *Znf708 i*SINE containing 3'UTR. *i*SINEs could make 300 bp double-stranded structure.

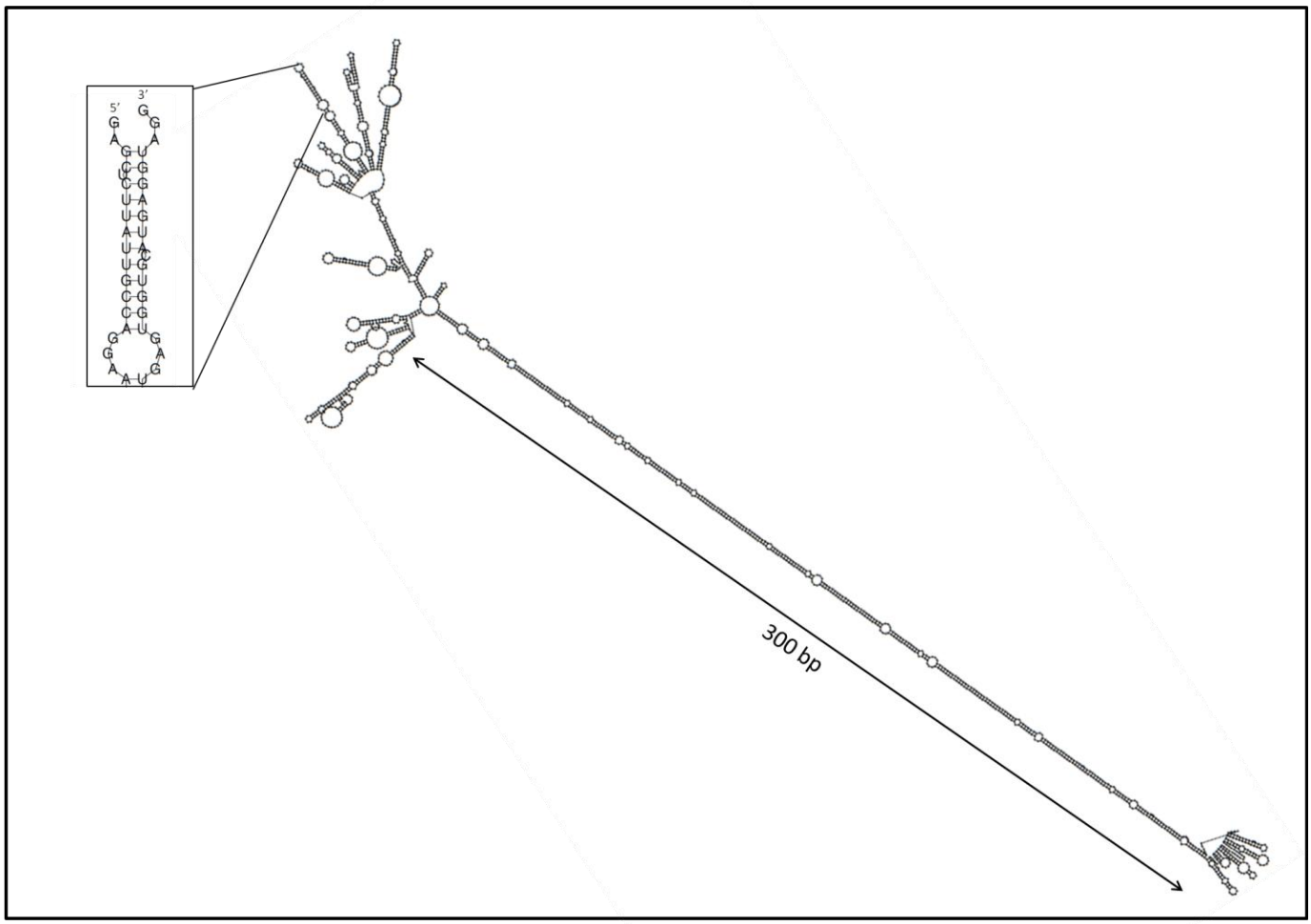

Figure 4. Secondary structure of 5'S *Znf708 i*SINE containing 3'UTR. The 5'S *Znf708* 3UTR is the shorter form of *Znf708* 3'UTR and it was made by removing the 500 nucleotides upstream of the *i*SINEs. *i*SINEs are able to make a 300 bp double-stranded structure.

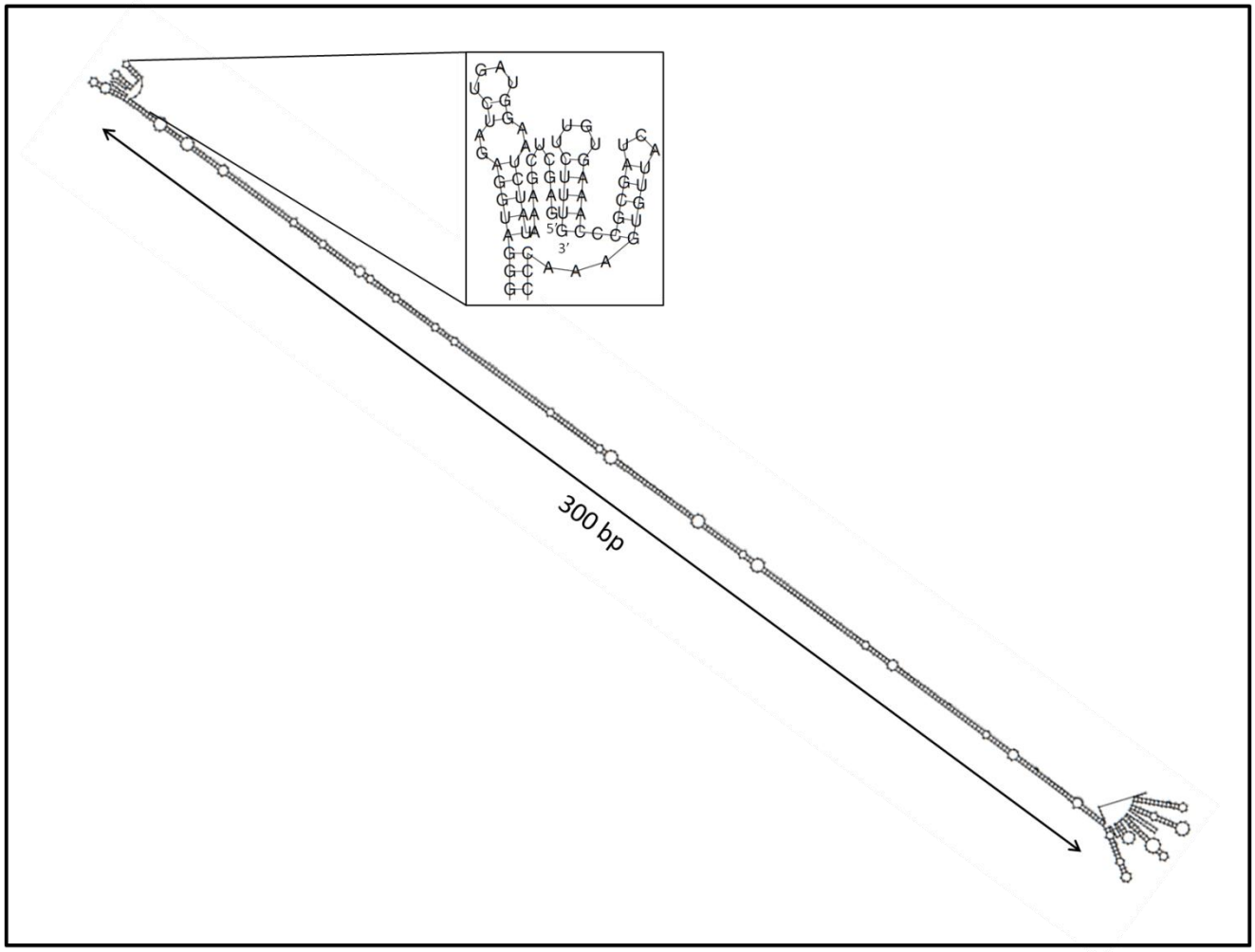

Figure 5. Secondary structure of 5'+3'S *Znf708 i*SINE containing 3'UTR. In order to create the 5'+3'S 3'UTR construct, the whole sequence between the stop codon and the first Alu was shortened to 50 bp. In addition, 150 bp were removed downstream of the SINEs. *i*SINEs could make a 300 bp double-stranded structure.

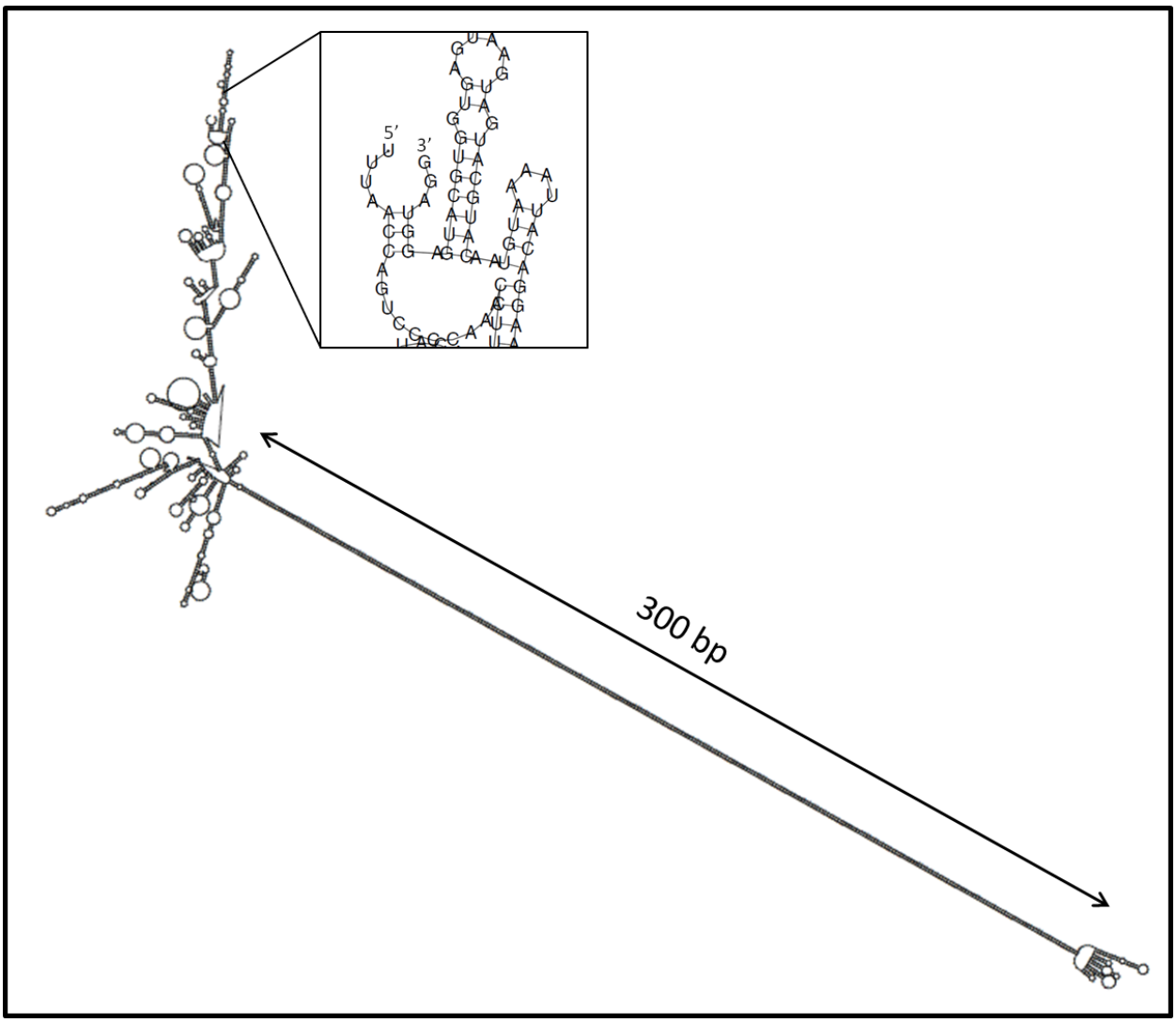

Figure 6. Secondary structure of perfect *(p)i*SINE containing 3'UTR. In the *Znf708* containing construct, Alu Sc was replaced with Alu Sg to make a perfect *i*SINE (*pi*SINE). *pi*SINEs are able to make a perfect double-stranded structure which is 300 bp long.

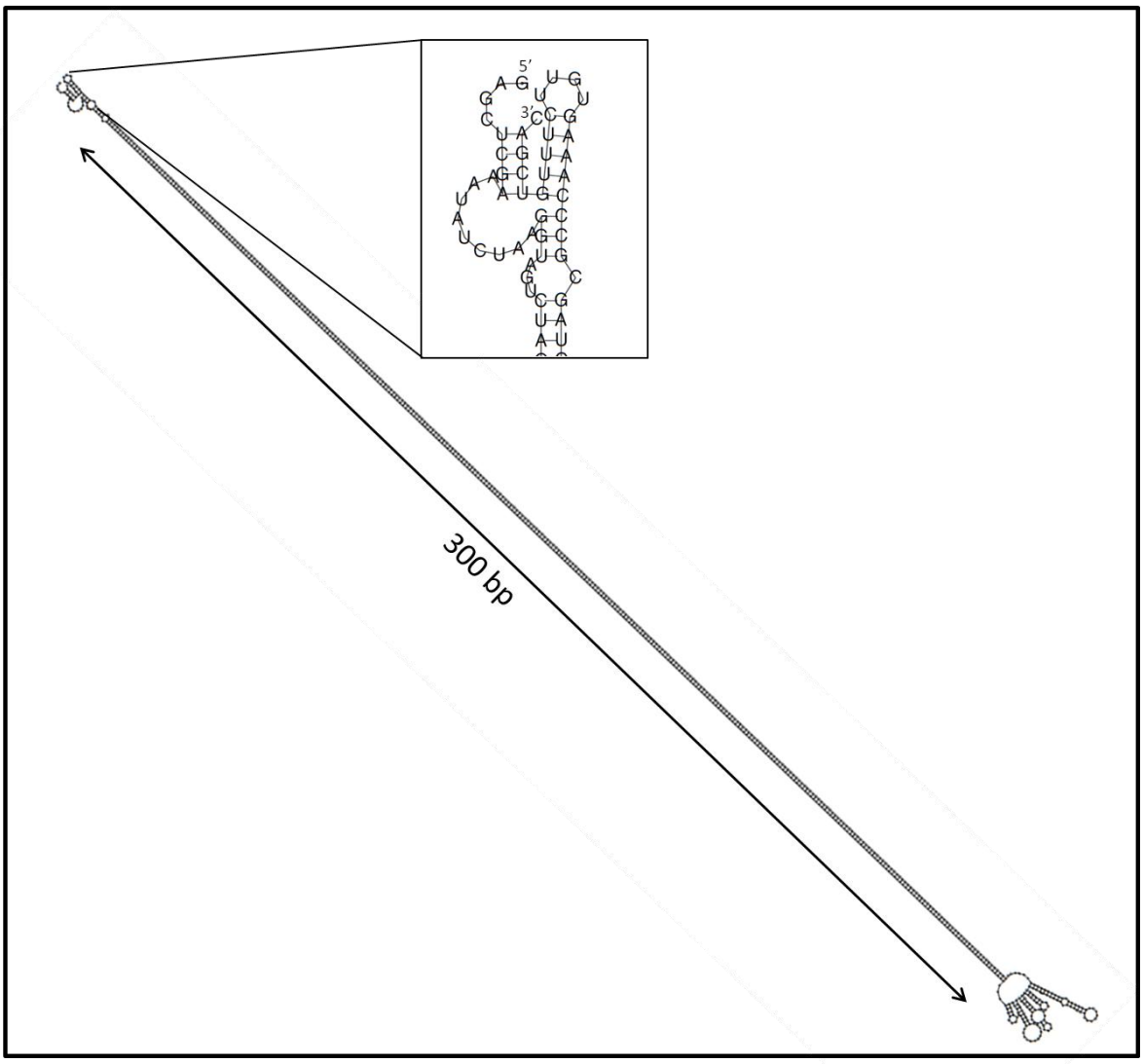

Figure 7. Secondary structure of perfect *(p)i*SINE containing 3'UTR, in the context of 5'+3'S *Znf708*. In the 5'+3'S *Znf708* containing construct, Alu Sc was replaced with Alu Sg to make a perfect *i*SINE (*pi*SINE). *pi*SINEs are able to make a perfect double-stranded structure which is 300 bp long.

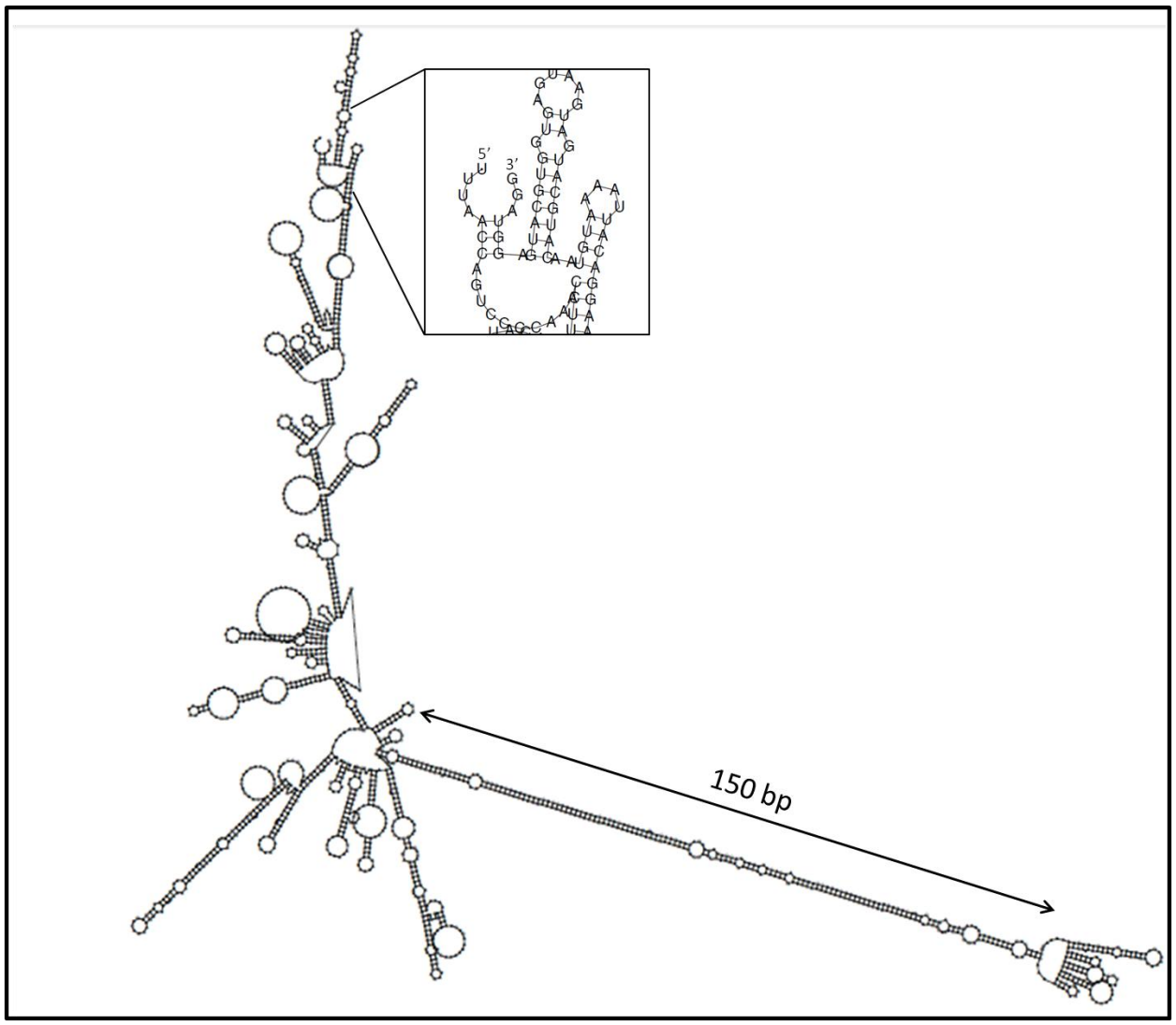

Figure 8. Secondary structure of mouse *i*SINEs containing 3'UTR. To make *i*SINE containing construct derived from B1 elements, we used the human *Znf708* 3'UTR as a starting construct. The *Znf708* Alu elements were replaced by B1 elements of the mouse *car5b* gene. Inverted B1 elemnts are able to make 150 bp double stranded structure.

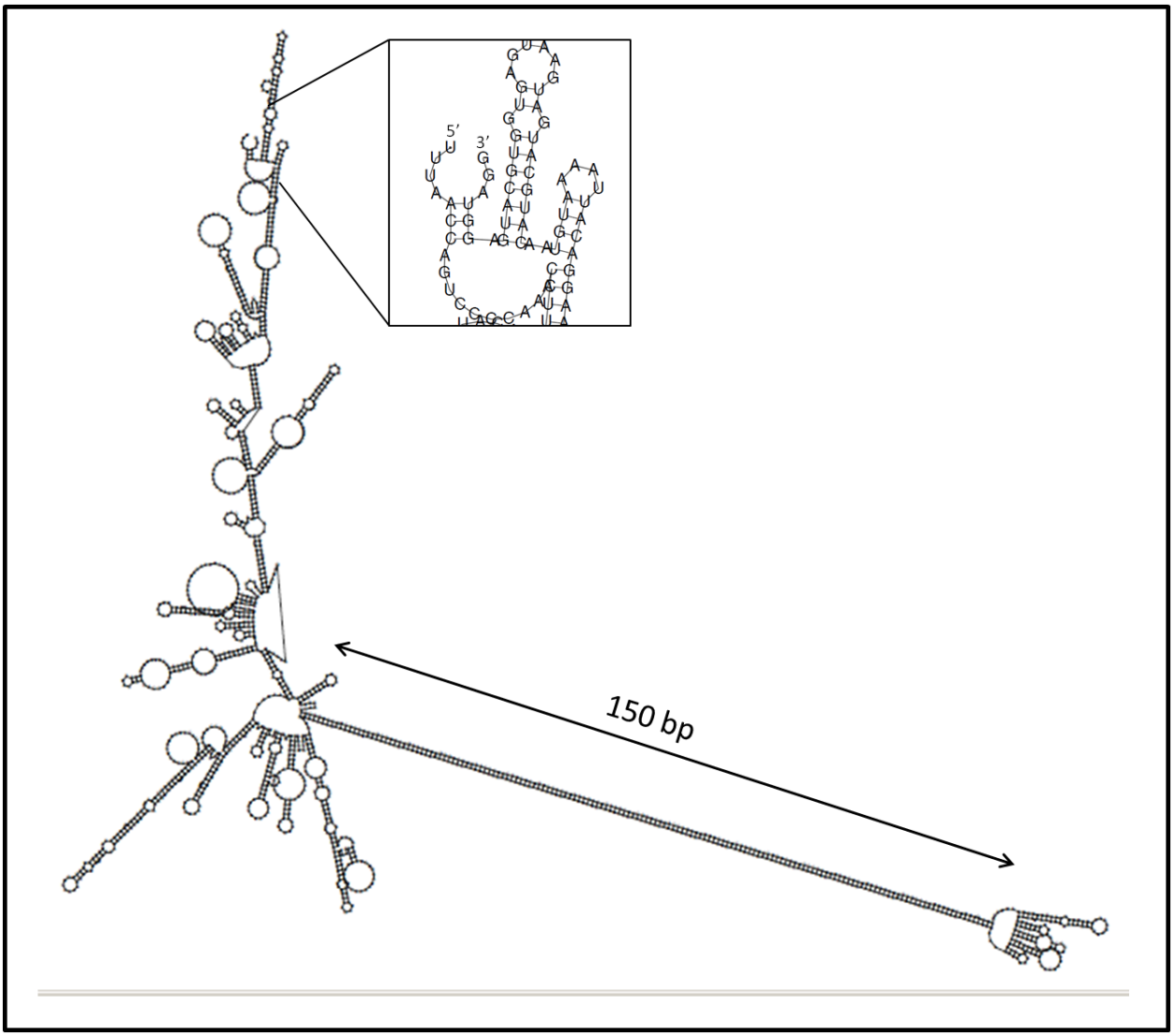

Figure 9. Secondary structure of a mouse perfect *i*SINEs containing 3'UTR. To make a perfect *i*SINE containing construct, the second B1 element in ZNF708 containing construct (supplementary figure 7) was replaced by the first B1 element in inverted orientations. Inverted B1 elements are able to make a perfect double stranded structure, which is 150 bp long.

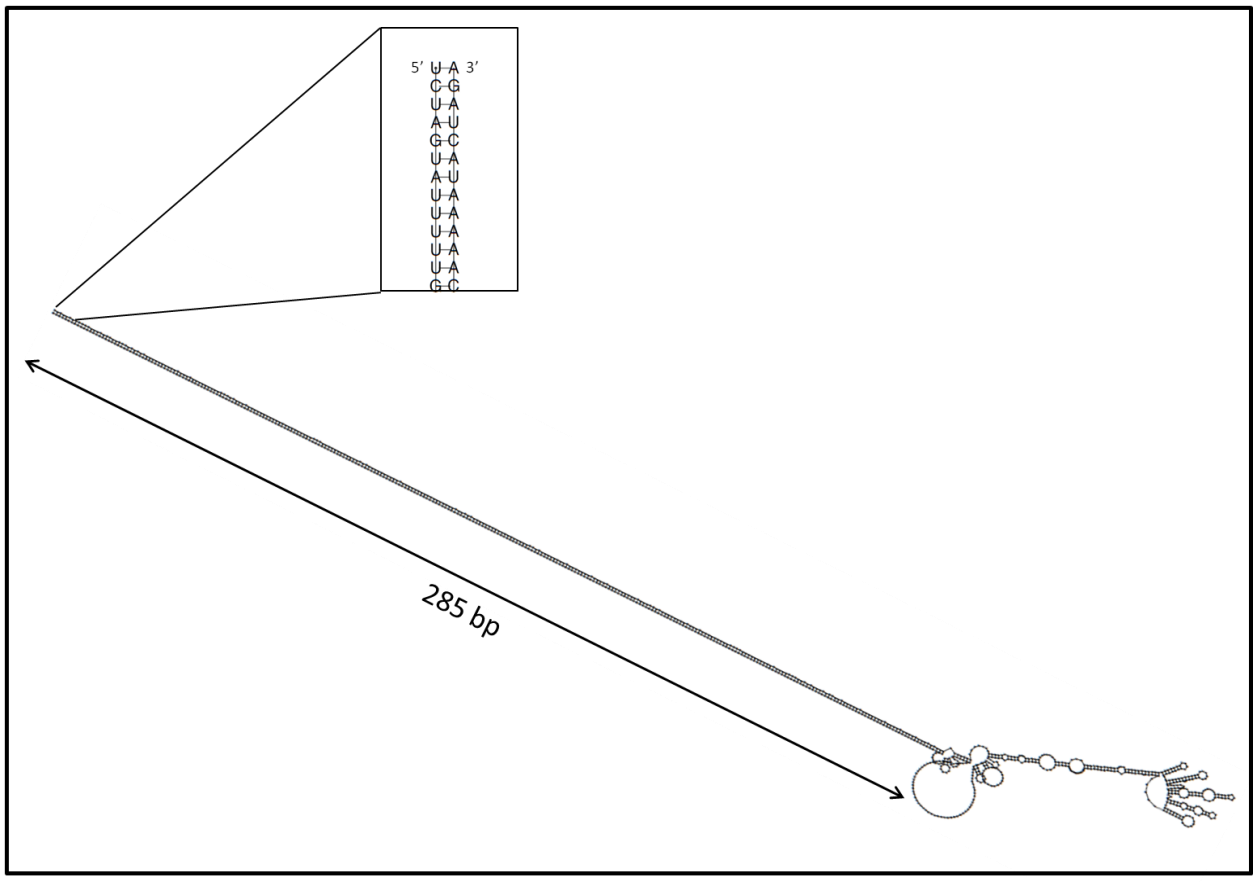

Figure 10. Secondary sructure of inverted artificial repeat containing (*i*AR) 3'UTR. in order to make *i*AR construct**,** 285 nucleotides of the normal Pyruvate Kinase 3'UTR were duplicated and inserted next to the 3' end of the UTR in inverted orientation. Artificial repeats could make a 285 bp long double-stranded structure.

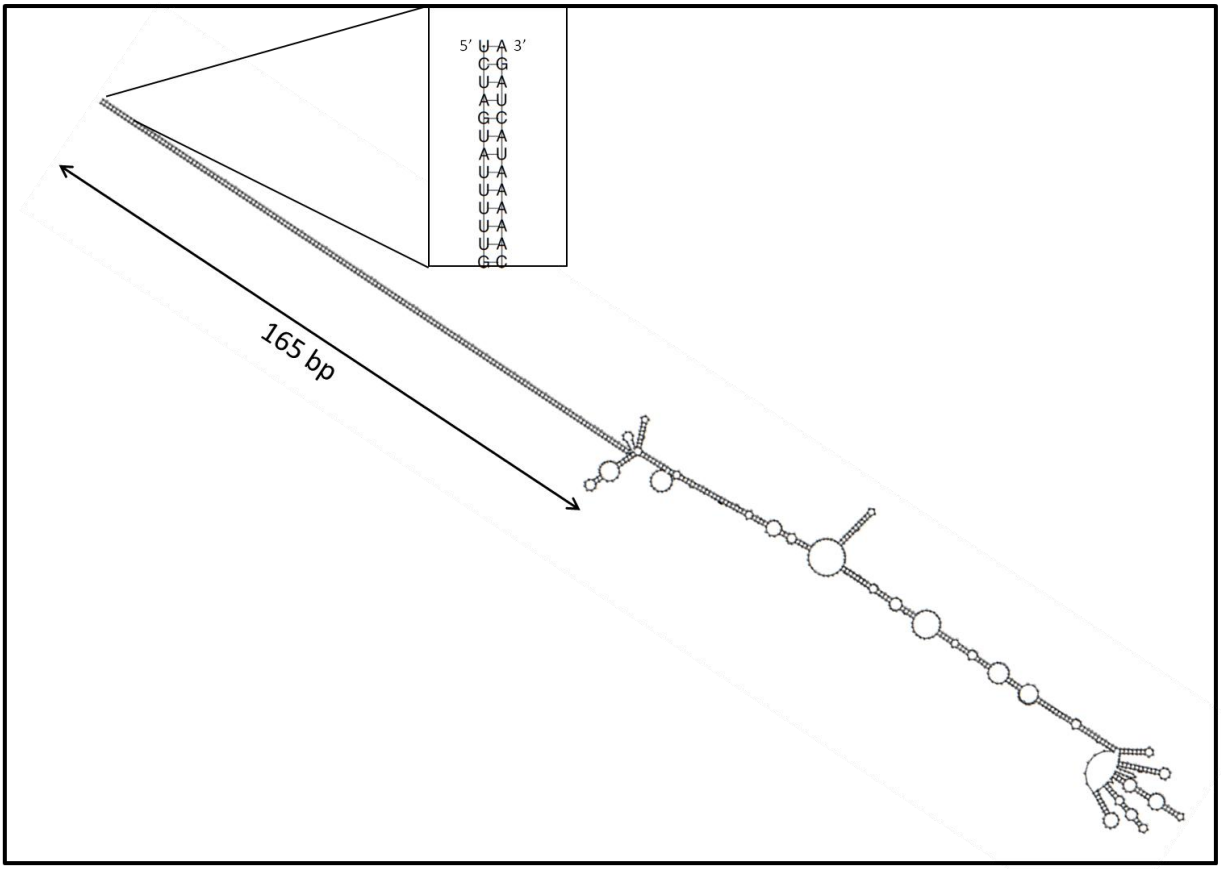

Figure 11. Secondary structure of a 165 bp long inverted artificial repeat containing (*i*AR 150bp) 3'UTR. 165 nucleotides of the normal Pyruvate Kinase 3'UTR were duplicated and inserted next to the 3' end of the UTR in inverted orientation. Artificial repeats could make a 165 bp long double-stranded structure.

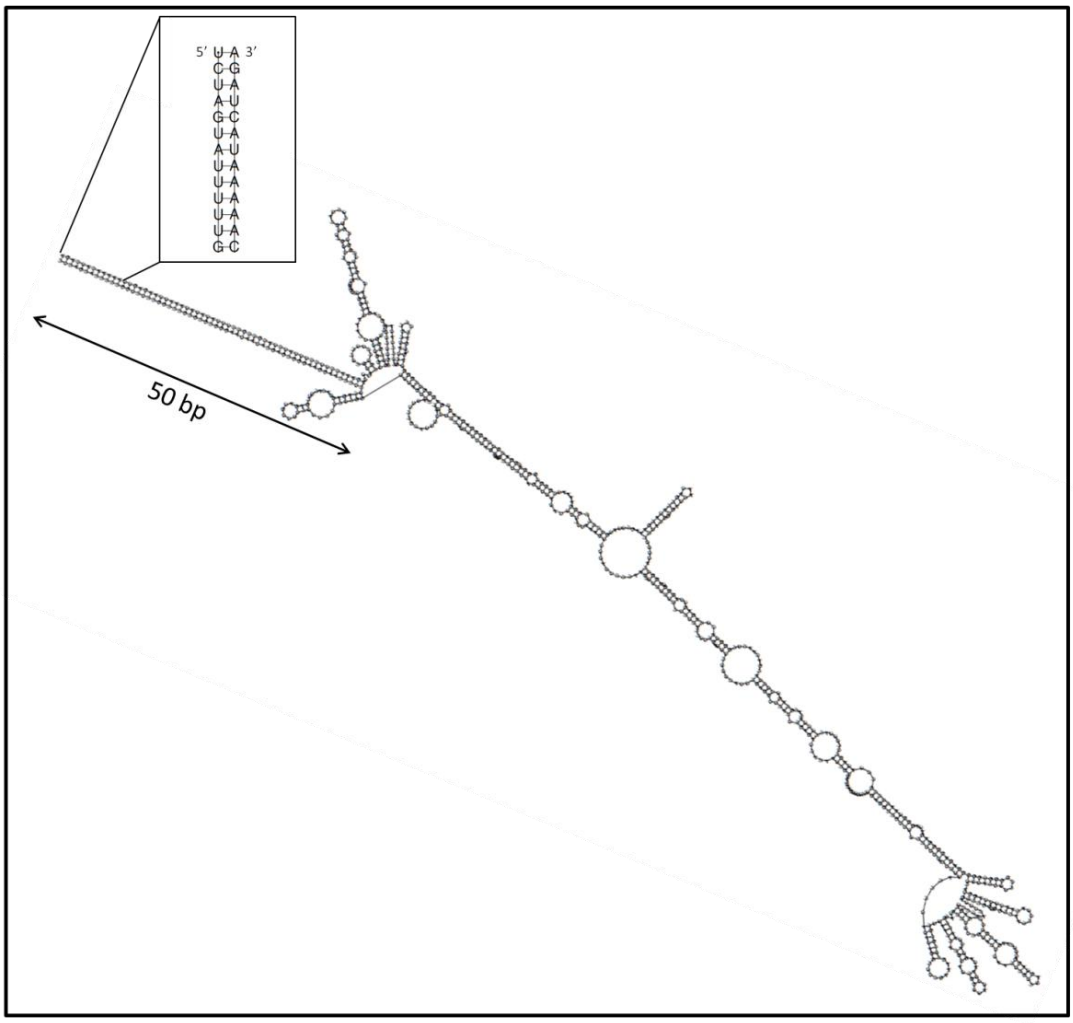

Figure 12. Secondary sructure of a 50 bp long inverted artificial repeat containing (*i*AR 50bp) 3'UTR. 50 nucleotides of the normal Pyruvate Kinase 3'UTR were duplicated and inserted next to the 3' end of the UTR in inverted orientation. Artificial repeats could make a double-stranded structure.

References:

Abraham, N., D. F. Stojdl, et al. (1999). "Characterization of transgenic mice with targeted disruption of the catalytic domain of the double-stranded RNA-dependent protein kinase, PKR." Journal of Biological Chemistry **274**(9): 5953-5962.

Allen, T. A., S. Von Kaenel, et al. (2004). "The SINE-encoded mouse B2 RNA represses mRNA transcription in response to heat shock." Nature Structural & Molecular Biology **11**(9): 816-821.

Athanasiadis, A., A. Rich, et al. (2004). "Widespread A-to-I RNA editing of alu-containing mRNAs in the human transcriptome." PLoS biology **2**(12): 2144-2158.

Ausubel F.M., B. R., Kingston R.E., Moore D.D., Seidmann J.G., Smith J.A., Struhl K. and K. Janssen, Eds. (1987). Current protocols in molecular biology. Current protocols in molecular biology. New York, John Wiley & Sons, Inc.

Backofen, B., R. Jacob, et al. (2002). "Cloning and characterization of the mammalian-specific nicolin 1 gene (NICN1) encoding a nuclear 24 kDa protein." European Journal of Biochemistry **269**(21): 5240-5245.

Batzer, M. A. and P. L. Deininger (2002). "Alu repeats and human genomic diversity." Nature Reviews Genetics **3**(5): 370-379.

Belancio, V. P., D. J. Hedges, et al. (2008). "Mammalian non-LTR retrotransposons: For better or worse, in sickness and in health." Genome research **18**(3): 343-358.

Bensaude, O. (2011). "Inhibiting eukaryotic transcription: Which compound to choose? How to evaluate its activity?" Transcription **2**(3): 103-108.

Berger, A. and K. Strub (2011). "Multiple Roles of Alu-Related Noncoding RNAs." Long Non-Coding Rnas, Vol 51 **51**: 119-146.

Britten, R. J. and D. E. Kohne (1968). "Repeated Sequences in DNA." Science **161**(3841): 529-&.

Burns, K. H. and J. D. Boeke (2012). "Human Transposon Tectonics." Cell **149**(4): 740-752.

Capshew, C. R., K. L. Dusenbury, et al. (2012). "Inverted Alu dsRNA structures do not affect localization but can alter translation efficiency of human mRNAs independent of RNA editing." Nucleic acids research **40**(17): 8637-8645.

Chen, L. L. and G. G. Carmichael (2008). "Gene regulation by SINES and inosines: biological consequences of A-to-I editing of Alu element inverted repeats." Cell cycle **7**(21): 3294-3301.

Chen, L. L., J. N. DeCerbo, et al. (2008). "Alu element-mediated gene silencing." The EMBO journal **27**(12): 1694-1705.

Chu, W. M., R. Ballard, et al. (1998). "Potential Alu function: Regulation of the activity of double-stranded RNA-activated kinase PKR." Molecular and cellular biology **18**(1): 58-68.

Cordaux, R. and M. A. Batzer (2009). "The impact of retrotransposons on human genome evolution." Nature reviews. Genetics **10**(10): 691-703.

Dangkulwanich, M., T. Ishibashi, et al. (2014). "Molecular Mechanisms of Transcription through Single-Molecule Experiments." Chemical Reviews **114**(6): 3203-3223.

Daniel, C., G. Silberberg, et al. (2014). "Alu elements shape the primate transcriptome by cis-regulation of RNA editing." Genome Biology **15**(2).

Dewannieux, M., C. Esnault, et al. (2003). "LINE-mediated retrotransposition of marked Alu sequences." Nature genetics **35**(1): 41-48.

Dugre-Brisson, S., G. Elvira, et al. (2005). "Interaction of Staufen1 with the 5 ' end of mRNA facilitates translation of these RNAs." Nucleic acids research **33**(15): 4797-4812.

Elbarbary, R. A., W. C. Li, et al. (2013). "STAU1 binding 3 ' UTR IRAlus complements nuclear retention to protect cells from PKR-mediated translational shutdown." Genes & Development **27**(13): 1495-1510.

Fedoroff, N. V. (2012). "PRESIDENTIAL ADDRESS Transposable Elements, Epigenetics, and Genome Evolution." Science **338**(6108): 758-767.
Feng, Q. H., J. V. Moran, et al. (1996). "Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition." Cell **87**(5): 905-916.
Garcia, M. A., E. F. Meurs, et al. (2007). "The dsRNA protein kinase PKR: Virus and cell control." Biochimie **89**(6-7): 799-811.
Gong, C. G. and L. E. Maquat (2011). "lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3 ' UTRs via Alu elements." Nature **470**(7333): 284-+.
Gong, C. G., Y. L. Tang, et al. (2013). "mRNA-mRNA duplexes that autoelicit Staufen1-mediated mRNA decay." Nature Structural & Molecular Biology **20**(10): 1214-+.
Han, J. S., S. T. Szak, et al. (2004). "Transcriptional disruption by the L1 retrotransposon and implications for mammalian transcriptomes." Nature **429**(6989): 268-274.
Hasler, J., Samuelsson, T. and Strub, K. (2007). "Useful 'junk': Alu RNAs in the human transcriptome." Cellular and molecular life sciences;CMLS **64**: 1793-1800.
Hasler, J. and K. Strub (2006). "Alu RNP and Alu RNA regulate translation initiation in vitro." Nucleic acids research **34**(8): 2374-2385.
Hasler, J. and K. Strub (2007). "Survey and summary - Alu elements as regulators of gene expression (vol 34, pg 5491, 2006)." Nucleic acids research **35**(4): 1389-1389.
Hellemans, J., G. Mortier, et al. (2007). "qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data." Genome Biology **8**(2): R19.
Hellmannblumberg, U., M. F. M. Hintz, et al. (1993). "Developmental Differences in Methylation of Human Alu Repeats." Molecular and cellular biology **13**(8): 4523-4530.
Heras, S. R., S. Macias, et al. (2013). "The Microprocessor controls the activity of mammalian retrotransposons." Nature Structural & Molecular Biology **20**(10): 1173-+.
Hundley, H. A. and B. L. Bass (2010). "ADAR editing in double-stranded UTRs and other noncoding RNA sequences." Trends in Biochemical Sciences **35**(7): 377-383.
Hundley, H. A., A. A. Krauchuk, et al. (2008). "C. elegans and H. sapiens mRNAs with edited 3' UTRs are present on polysomes." RNA **14**(10): 2050-2060.
Ichiyanagi, K. (2013). "Epigenetic regulation of transcription and possible functions of mammalian short interspersed elements, SINEs." Genes & Genetic Systems **88**(1): 19-29.
Ichiyanagi, K., Y. F. Li, et al. (2011). "Locus- and domain-dependent control of DNA methylation at mouse B1 retrotransposons during male germ cell development." Genome research **21**(12): 2058-2066.
Jantsch, M. F. and J. G. Gall (1992). "Assembly and localization of the U1-specific snRNP C protein in the amphibian oocyte." J Cell Biol **119**(5): 1037-1046.
Jonkers, I., H. Kwak, et al. (2014). "Genome-wide dynamics of Pol II elongation and its interplay with promoter proximal pausing, chromatin, and exons." Elife **3**.
Kameoka, S., P. Duque, et al. (2004). "P54(nrb) associates with the 5 ' splice site within large transcription/splicing complexes." Embo Journal **23**(8): 1782-1791.
Kaneko, S., O. Rozenblatt-Rosen, et al. (2007). "The multifunctional protein p54nrb/PSF recruits the exonuclease XRN2 to facilitate pre-mRNA 3 ' processing and transcription termination." Genes & Development **21**(14): 1779-1789.
Keren, H., G. Lev-Maor, et al. (2010). "Alternative splicing and evolution: diversification, exon definition and function." Nature Reviews Genetics **11**(5): 345-355.
Kim, Y., J. H. Lee, et al. (2014). "PKR is activated by cellular dsRNAs during mitosis and acts as a mitotic regulator." Genes & Development **28**(12): 1310-1322.
Kim, Y. K., L. Furic, et al. (2005). "Mammalian Staufen1 recruits Upf1 to specific mRNA 3'UTRs so as to elicit mRNA decay." Cell **120**(2): 195-208.

Kim, Y. K., L. Furic, et al. (2007). "Staufen1 regulates diverse classes of mammalian transcripts." Embo Journal **26**(11): 2670-2681.
Kohrmann, M., M. Luo, et al. (1999). "Microtubule-dependent recruitment of Staufen-Green fluorescent protein into large RNA-containing granules and subsequent dendritic transport in living hippocampal neurons." Molecular Biology of the Cell **10**(9): 2945-2953.
Kriegs, J. O., G. Churakov, et al. (2007). "Evolutionary history of 7SL RNA-derived SINEs in supraprimates." Trends in Genetics **23**(4): 158-161.
Lander, E. S., I. H. G. S. Consortium, et al. (2001). "Initial sequencing and analysis of the human genome." Nature **409**(6822): 860-921.
Lev-Maor, G., R. Sorek, et al. (2007). "RNA-editing-mediated exon evolution." Genome Biology **8**(2).
Levanon, E. Y., E. Eisenberg, et al. (2004). "Systematic identification of abundant A-to-I editing sites in the human transcriptome." Nature biotechnology **22**(8): 1001-1005.
Liu, W. M., W. M. Chu, et al. (1995). "Cell Stress and Translational Inhibitors Transiently Increase the Abundance of Mammalian Sine Transcripts." Nucleic acids research **23**(10): 1758-1765.
Luan, D. D., M. H. Korman, et al. (1993). "Reverse Transcription of R2bm Rna Is Primed by a Nick at the Chromosomal Target Site - a Mechanism for Non-Ltr Retrotransposition." Cell **72**(4): 595-605.
Magdinier, F., S. Ribieras, et al. (1998). "Down-regulation of BRCA1 in human sporadic breast cancer; analysis of DNA methylation patterns of the putative promoter region." Oncogene **17**(24): 3169-3176.
Mariner, P. D., R. D. Walters, et al. (2008). "Human Alu RNA is a modular transacting repressor of mRNA transcription during heat shock." Molecular cell **29**(4): 499-509.
Matthew S. Comeaux, A. M. R.-E., Dale J. Hedges and Prescott L. Deininger (2009). "Diverse cis factors controlling Alu retrotransposition: What causes Alu elements to die?" Genome research **19**(4): 545-555.
Minasaki, R., D. Rudel, et al. (2014). "Increased sensitivity and accuracy of a single-stranded DNA splint-mediated ligation assay (sPAT) reveals poly(A) tail length dynamics of developmentally regulated mRNAs." RNA biology **11**(2): 111-123.
Modrek, B. and C. Lee (2002). "A genomic view of alternative splicing." Nature genetics **30**(1): 13-19.
Moran, J. V., S. E. Holmes, et al. (1996). "High frequency retrotransposition in cultured mammalian cells." Cell **87**(5): 917-927.
Neeman, Y., E. Y. Levanon, et al. (2006). "RNA editing level in the mouse is determined by the genomic repeat repertoire." RNA **12**(10): 1802-1809.
Nejepinska, J., R. Malik, et al. (2012). "dsRNA expression in the mouse elicits RNAi in oocytes and low adenosine deamination in somatic cells." Nucleic Acids Research **40**(1): 399-413.
Pace, J. K. and C. Feschotte (2007). "The evolutionary history of human DNA transposons: Evidence for intense activity in the primate lineage." Genome research **17**(4): 422-432.
Pan, Q., O. Shai, et al. (2008). "Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing." Nature genetics **40**(12): 1413-1415.
Park, E. and L. E. Maquat (2013). "Staufen-mediated mRNA decay." Wiley Interdisciplinary Reviews-Rna **4**(4): 423-435.
Peculis, B. A. and J. G. Gall (1992). "Localization of the Nucleolar Protein-No38 in Amphibian Oocytes." Journal of Cell Biology **116**(1): 1-14.
Philipp, S. and V. Flockerzi (1997). "Molecular characterization of a novel human PDZ domain protein with homology to INAD from Drosophila melanogaster." Febs Letters **413**(2): 243-248.
Ponicsan, S. L., S. Houel, et al. (2013). "The Non-Coding B2 RNA Binds to the DNA Cleft and Active-Site Region of RNA Polymerase II." Journal of molecular biology **425**(19): 3625-3638.
Ponicsan, S. L., J. F. Kugel, et al. (2010). "Genomic gems: SINE RNAs regulate mRNA production." Current Opinion in Genetics & Development **20**(2): 149-155.

Prasanth, K. V., S. G. Prasanth, et al. (2005). "Regulating gene expression through RNA nuclear retention." Cell **123**(2): 249-263.
Ricci, E. P., A. Kucukural, et al. (2014). "Staufen1 senses overall transcript secondary structure to regulate translation." Nature Structural & Molecular Biology **21**(1): 26-+.
Rosbash, M., P. J. Ford, et al. (1974). "Analysis of C-Value Paradox by Molecular Hybridization." Proceedings of the National Academy of Sciences of the United States of America **71**(9): 3746-3750.
Rueter, S. M., T. R. Dawson, et al. (1999). "Regulation of alternative splicing by RNA editing." Nature **399**(6731): 75-80.
Sela, N., B. Mersch, et al. (2007). "Comparative analysis of transposed element insertion within human and mouse genomes reveals Alu's unique role in shaping the human transcriptome." Genome Biology **8**(6).
Shattuckeidens, D., M. Mcclure, et al. (1995). "A Collaborative Survey of 80 Mutations in the Brca1 Breast-Cancer and Ovarian-Cancer Susceptibility Gene - Implications for Presymptomatic Testing and Screening." Jama-Journal of the American Medical Association **273**(7): 535-541.
Sobczak, K. and W. J. Krzyzosiak (2002). "Structural determinants of BRCA1 translational regulation." Journal of Biological Chemistry **277**(19): 17349-17358.
Sommer, B., M. Kohler, et al. (1991). "RNA editing in brain controls a determinant of ion flow in glutamate-gated channels." Cell **67**(1): 11-19.
Sorek, R., G. Ast, et al. (2002). "Alu-containing exons are alternatively spliced." Genome research **12**(7): 1060-1067.
Tarallo, V., Y. Hirano, et al. (2012). "DICER1 Loss and Alu RNA Induce Age-Related Macular Degeneration via the NLRP3 Inflammasome and MyD88." Cell **149**(4): 847-859.
Toulokhonov, I., I. Artsimovitch, et al. (2001). "Allosteric control of RNA polymerase by a site that contacts nascent RNA hairpins." Science **292**(5517): 730-733.
Versteeg, R., B. D. C. van Schaik, et al. (2003). "The human transcriptome map reveals extremes in gene density, intron length, GC content, and repeat pattern for domains of highly and weakly expressed genes." Genome research **13**(9): 1998-2004.
Vessey, J. P., P. Macchi, et al. (2008). "A loss of function allele for murine Staufen1 leads to impairment of dendritic Staufen1-RNP delivery and dendritic spine morphogenesis." Proceedings of the National Academy of Sciences of the United States of America **105**(42): 16374-16379.
Wagschal, A., E. Rousset, et al. (2012). "Microprocessor, Setx, Xrn2, and Rrp6 Co-operate to Induce Premature Termination of Transcription by RNAPII." Cell **150**(6): 1147-1157.
Wicker, T., W. Zimmermann, et al. (2005). "A detailed look at 7 million years of genome evolution in a 439 kb contiguous sequence at the barley Hv-eIF4E locus: recombination, rearrangements and repeats." Plant Journal **41**(2): 184-194.
Williams, B. R. G. (1999). "PKR; a sentinel kinase for cellular stress." Oncogene **18**(45): 6112-6120.
Xie, H. H., M. Wang, et al. (2009). "High-throughput sequence-based epigenomic analysis of Alu repeats in human cerebellum." Nucleic acids research **37**(13): 4331-4340.
Xu, C. F., M. A. Brown, et al. (1995). "Distinct Transcription Start Sites Generate 2 Forms of Brca1 Messenger-Rna." Human molecular genetics **4**(12): 2259-2264.
Yakovchuk, P., J. A. Goodrich, et al. (2009). "B2 RNA and Alu RNA repress transcription by disrupting contacts between RNA polymerase II and promoter DNA within assembled complexes." Proceedings of the National Academy of Sciences of the United States of America **106**(14): 5569-5574.
Zhang, Z. and G. G. Carmichael (2001). "The fate of dsRNA in the nucleus: A p54(nrb)-containing complex retention of promiscuously mediates the nuclear A-to-I edited RNAs." Cell **106**(4): 465-475.

Zheng, W. X., F. Luo, et al. (2000). "Reduction of BRCA1 expression in sporadic ovarian cancer." Gynecologic oncology **76**(3): 294-300.

Acknowledgment

First, I would like to express my gratitude to my supervisor, Michael Jantsch, for offering me this fully challenging project. It was a privilege to work with him towards my PhD. By being a member of his research group, I have learnt how to deal with scientific challenges and enjoy doing that. I am really grateful to have the opportunity to work in his lab.

One of the most memorable and greatest experiences during PhD is pursuing your goals with great colleagues, so, I especially like to thank all our lab-members who made a great working atmosphere through these years; Aamira, Silpi, Maja and Mamta who have been good friends and support for me in different steps of my PhD and also our former lab members, Conny and Florian, and current ones Laura, Emmilio, Prajakta and Utkarsh. Moreover, I appreciate the fruitful discussions with Konstantin and his collaboration in this project.

I would like to thank my wonderful friends whom I met in Vienna and I could not accomplish my PhD without their supports. Peyman and Mona, Who are my family in Vienna. There is no word to show my gratefulness for having you guys. Yammile and Zahra, who have been my colleagues and turned to be my amazing friends during this journey. I am also thankful to all my Persian wonderful friends Tara, Serwe, Sara, Ladan and Parin.

My special thanks go to my Family; my parents, Mohsen and Majid. You have been 4000 km away, but your endless support and love gave me incredible energy and ambition to struggle with obstacles and go for more and more achievements. You are the most valuable gifts in my life.

Curriculum Vitae

General information

First name:	Mansoureh
Last name:	Tajaddod
Date and place of birth:	17 July 1983, Isfahan, Iran
Current address:	Max F. Perutz Laboratories, Dept. of Chromosome Biology. Dr. Bohr Gasse 9/5, 1030, Vienna, Austria.
E-mail:	mansoureh.tajaddod@univie.ac.at
Current activity:	PhD student in Dr. Michael Jantsch Laboratory

Education

2010 – present	PhD in Molecular Biology, Max F. Perutz labrotaries (MFPL), Vienna, Austria. Project: 'Molecular mechanism of gene silencing through inverted Alu elements'; laboratory of Dr Michael Jantsch.
2005 – 2008	Diploma thesis, department of Biology, University of Isfahan, Isfahan, Iran. Project: 'Identification and analysis of STR markers around HLA-DRB1 region'; laboratory of Dr Sadegh Valian.
2000 – 2005	Studies of Celular and Molecular Biology–Genetic, University of Isfahan, Iran

Conferences and Presentations

- **Tajaddod, M**., Vallian, S. and Hojati, Z. 2007. Identification and characterization of short tandem repeat (STR) markers in HLA-DRB1 region. Accepted as a poster in the ninth Iranian congress of biochemistry & second international congress of biochemistry and molecular biology (29 Oct-1 Nov. 2007) Shiraz-Iran. Published in Archive of Iranian Medicine. Vol 10(4):207.

- Vallian, S., **Tajaddod, M**. and Hojati, Z. 2007. Computational and molecular analysis of short tandem repeat (STR) markers in HLA-DRB1 region. Accepted as a poster in the first international congress on health genomics & biotechnology (24-26 Nov. 2007) Tehran-Iran. Published in Iranian Biomedical Journal. Vol 11(1):175.

- **Tajaddod, M**., Shopof, S. and Jantsch, M.F. The effect of inverted iSINE in gene expression. Accepted as poster in the RNA Editing Gordon Research Conference, January 9-14, 2011.

- **Tajaddod, M**., Shopof, S. and Jantsch, M.F. The effect of (inverted) SINE in gene expression. Accepted as poster in The EMBO Meeting, 10 - 13 Sept 2011 in Vienna.

- **Tajaddod, M**., Licht, K., Huber, F., Schopoff, S. and Jantsch, M.F. The effect of (inverted) SINE in gene expression. Accepted as poster in 18th annual meeting of the RNA society, 11-16 June 2013 in Davos, Switzerland.

- **Tajaddod, M**., Licht, K., Huber, F., Schopoff, S. and Jantsch, M.F. The effect of (inverted) SINE in gene expression. Accepted as poster in “RNA 3' ends: Mechanism and biological function in eukaryotic genomes” workshop, 25-29 Sep 2013 in Oxford.

Publications

- Vallian, S., **Tajadod, M**. and Hojati, Z. *Characterization of D6S2806 and D6S2879 short tandem repeat loci in HLA-DRB1 region in Iranian population*. Tissue Antigens. 2010 Jul;76(1):60-3.

- Vesely C, Tauber S, Sedlazeck FJ, **Tajaddod M**, Ivon Haeseler A, Jantsch MF ADAR2 induces reproducible changes in sequence and abundance of mature microRNAs in the mouse brain. Nucleic Acids Res. 2014 Oct 29;42(19): 12155–12168.

Awards

- 2001, Accepted in public university. (Only 10% of students who completed their high school accepted into the bachelor's program in public university).
- 2004, First ranked genetic students in the Genetics Department in Isfahan University.
- 2005, Graduated as a second ranked genetic student from Isfahan University.
- 2005, Introduced as an Exceptional Talent in Isfahan University.
- 2005, Invited for studding Master in Isfahan University.
- 2008, Submitting several STR allele sequences from Iranian population in Gene Bank.
- 2009, Accepted in international PhD program of RNA Doctoral School (Vienna, Austria)

Courses and Training

- DK RNA Biology & SFB RNA-REG workshop, *'Coaching - how to make the audience listen to your presentation'*, Vienna/Austria, 2012.
- CSF course: *'Introduction to Applied Statistics'*, Vienna/Austria 2013.
- DK RNA Biology & SFB RNA-REG workshop, *'Translational research workshop'*, Vienna/Austria 2014.
- Molecular mechanism of cell signaling workshop, *'ENCODE Workshop'*, Vienna/Austria 2014.

- DK RNA Biology & SFB RNA-REG workshop, *'Basic statistics and graphics in R'*, Vienna/Austria 2014.