

# Misfolded human tRNA isodecoder binds and neutralizes a 3' UTR-embedded Alu element

Joëlle Rudinger-Thirion, Alain Lescure, Caroline Paulus, and Magali Frugier<sup>1</sup>

Architecture et Réactivité de l'ARN (RNA Architecture and Reactivity), Université de Strasbourg, Centre National de la Recherche Scientifique, Institut de Biologie Moléculaire et Cellulaire, 15 rue René Descartes, 67084 Strasbourg Cedex, France

Edited\* by Paul Schimmel, The Skaggs Institute for Chemical Biology, La Jolla, CA, and approved August 9, 2011 (received for review March 11, 2011)

Several classes of small noncoding RNAs are key players in cellular metabolism including mRNA decoding, RNA processing, and mRNA stability. Here we show that a tRNA<sup>Asp</sup> isodecoder, corresponding to a human tRNA-derived sequence, binds to an embedded Alu RNA element contained in the 3' UTR of the human aspartyl-tRNA synthetase mRNA. This interaction between two well-known classes of RNA molecules, tRNA and Alu RNA, is driven by an unexpected structural motif and induces a global rearrangement of the 3' UTR. Besides, this 3' UTR contains two functional polyadenylation signals. We propose a model where the tRNA/Alu interaction would modulate the accessibility of the two alternative polyadenylation sites and regulate the stability of the mRNA. This unique regulation mechanism would link gene expression to RNA polymerase III transcription and may have implications in a primate-specific signal pathway.

aminoacyl-tRNA synthetase | posttranscriptional regulation

Alu elements are the most abundant repetitive element in primate genomes. They belong to the short interspersed element family and are classified into subfamilies according to their appearance during evolution. Typically, there are more than 1 million Alu elements in a primate genome, which represents 10% of the genome overall. Modern Alu elements are about 300 base-pairs (bp) long and adopt a conserved structure composed of two similar, but distinct, domains (left and right arms) joined by an A-rich linker and followed by a short polyadenylation [poly(A)] stretch. Moreover, Alu elements are transcribed by RNA polymerase III (pol III), and Alu RNA transcripts are present in the cytosol of primate cells (1).

Alu sequences are considered to be a huge reservoir of potential regulatory elements, which may have been involved in primate evolution and possibly facilitated their divergence from other mammals (2). Indeed, insertion of Alu elements could provide previously undescribed regulatory motifs to neighboring genes by creating previously undescribed transcription enhancers or promoters (3, 4). Alu elements can also modulate protein expression in a posttranscriptional way by at least two different mechanisms: (i) they are transcribed as free Alu RNA by the RNA PolIII, assembled into Alu ribonucleo-protein particles and act as transregulatory factors (2, 5, 6); or (ii) their transcription as a part of a mRNA will define them as cis-regulatory elements (7). If the Alu sequence is localized within the mRNA open reading frame, it can either affect gene expression by sequence disruption or, alternatively, simply add to the protein amino acid sequence (8). Nevertheless, most of the time they are present in 5' and 3' UTRs (2, 9, 10).

Several reports indicate that mRNAs containing Alu elements in their 3' UTRs are associated with cell growth and differentiation (11–15). Indeed, their presence could affect the processing of mRNAs at multiple levels (16). Specifically, they have been shown to modulate alternative splicing (e.g., ref. 17), RNA editing (e.g., ref. 18), nuclear retention (19), and STAU1-mediated mRNA decay (20) and to create AU-rich elements that regulate mRNA stability (21). In addition, Alu elements have been shown to introduce previously undescribed 3' UTR polyadenylation

sites (22), thereby affecting 3' end processing efficiency, as well as alternative 3' end selection. Remarkably, several examples stress the essential contribution of 3' end processing in physiological (e.g., immunity and inflammation) or pathological processes (e.g., cancer and viral infection) (23). Maturation of the 3' end of the mRNA is a nuclear cotranscriptional process that promotes transport of mRNAs from the nucleus to the cytoplasm and affects the stability and the translation efficiency of mRNAs. This process consists of the recognition of defined polyadenylation [poly(A)] signals in the pre-mRNA 3' UTR by a large cleavage/polyadenylation machinery. For transcripts containing more than one poly(A) signal, which is the majority of the pre-mRNAs (24–26), the basis of the mechanisms involved in poly(A) selection is still not understood and the role of regulatory factors remains undefined (27).

Interestingly, we made the observation that the mRNA encoding the human aspartyl-tRNA synthetase (AspRS) displays both a partial Alu insertion (3' end of the AluJ right arm) and two optimal poly(A) signals in its 3' UTR. Aminoacyl-tRNA synthetases (aaRSs) are ubiquitous enzymes that attach amino acids to their cognate tRNAs during protein synthesis. They ensure the high fidelity of the tRNA aminoacylation process, thereby maintaining the genetic code and contributing to cell viability (28). Our data show that (i) a human tRNA<sup>Asp</sup> sequence adopts a peculiar structure, (ii) it binds to the Alu insertion within the 3' UTR of AspRS mRNA, (iii) it stimulates specifically the expression of a luciferase reporter gene fused to this 3' UTR, (iv) it alters the 3' UTR overall folding, and (v) the regulatory function of the 3' UTR is related not only to the Alu insertion but also to two functional polyadenylation sites. We propose a model consistent with all these observations, in which the tRNA<sup>Asp</sup> isodecoder, by binding the Alu insertion, would direct alternative polyadenylation of AspRS mRNA and modulate its expression. We show that a tRNA isodecoder could act as a non-coding RNA (ncRNA), and we discuss its potential implication in a pathway important in cellular metabolism.

## Results

**AspRS mRNA 3' UTR Displays Two Poly(A) Signals and a Partial Alu Sequence.** Human AspRS is a protein of 500 amino acids, encoded on chromosome 2. The corresponding mature 2,215 nucleotide (nt) mRNA is characterized by a short 5' UTR (79 nt) and a predicted 633-nt 3' UTR containing two polyadenylation signals (Fig. 1A). This transcript is of low abundance and could not be detected by Northern blotting experiments, as verified by quan-

Author contributions: J.R.-T., A.L., and M.F. designed research; J.R.-T., A.L., C.P., and M.F. performed research; J.R.-T., A.L., and M.F. analyzed data; and J.R.-T., A.L., and M.F. wrote the paper.

The authors declare no conflict of interest.

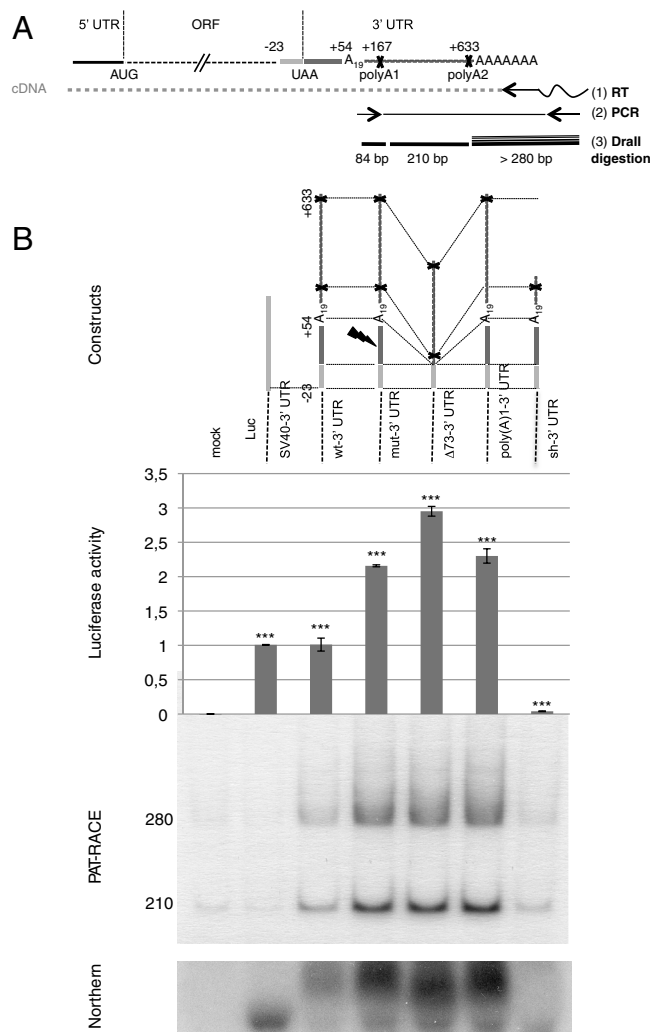
\*This Direct Submission article had a prearranged editor.

See Commentary on page 16489.

<sup>1</sup>To whom correspondence should be addressed. E-mail: M.Frugier@ibmc-cnrs.unistra.fr.

See Author Summary on page 16497.

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1103698108/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1103698108/-DCSupplemental).



**Fig. 1.** AspRS 3' UTR variants and luciferase reporter gene expression. (A) Global organization of the complete AspRS mRNA and PAT-RACE principle: UTRs and ORF are indicated. Positions numbers are relative to the UAA stop codon. The 3' UTR contains an Alu sequence located downstream the stop codon (up to position +54) and two predicted polyadenylation signals (polyA1 and polyA2 at positions 167 and 633, respectively). In vitro probing of the junction between the ORF and the 3' UTR showed that the Alu sequence has the capacity to form a stem loop with the last 23 nts of the ORF; therefore, we chose to fuse this entire domain in the reporter system in order to present the Alu sequence in its native structural context (cotranscriptional folding). In PAT-RACE experiments, this mRNA is first reverse transcribed using a "PAT adapter" primer. The resulting cDNA is then amplified using a primer complementary to the adapter and a primer designed intentionally to hybridize after the internal polyA sequence ( $A_{19}$ ) to avoid background. DraII digestion of radioactive PCR products lead to three fragments of 84, 210, and about 280 nts. This longest fragment is specified by the presence of different polyA tails, initially used to reverse transcribe the mRNA and the shortest fragment was not visible on gels. (B) Schematic representation of reporter constructs: WT-3' UTR, wild-type sequence including the 23 last nts of AspRS ORF and the 633 nt of the 3' UTR, Mut-3' UTR is the Alu mutated version (sequence  $_{31}\text{GATGACAG}_{38}$  was replaced by  $_{31}\text{TAGTCACT}_{38}$ ),  $\Delta 73$ -3' UTR corresponds to the 3' UTR missing the 54-nt-long Alu sequence and the 19-A stretch, poly(A)1-3' UTR has its first poly(A) signal AATAAA mutated into AATGCA and Sh-3' UTR ends at nt 213 and displays only the first poly(A) signal. Finally, SV40-3' UTR was used as a control. Gene expression is monitored at the protein level with the luciferase activity (Middle, standard deviation values of three independent transfections) and at the mRNA level (Bottom); relative mRNA concentrations are estimated by RACE-PAT (30) and by Northern blotting. Error bars indicating the standard deviation are given. \*\*\* shows statistical significance (AspRS 3' UTR fusions versus SV40 3' UTR) at  $P < 0.001$  (Welch's  $t$  test).

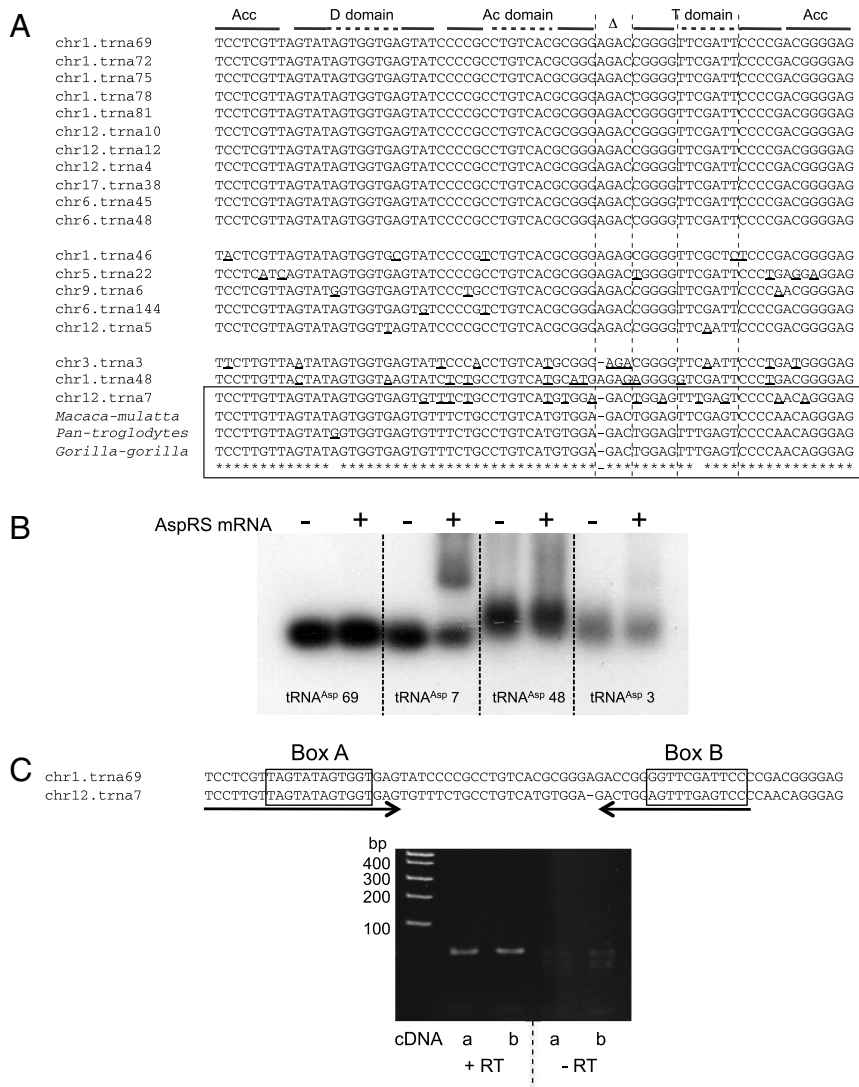
titative PCR. Therefore, we used RNA ligase mediated RACE by PCR (RLM-RT-RACE PCR) to establish the nucleotide sequence of the different 3' ends of AspRS mRNA transcripts produced in humans (Fig. S1A). This approach was performed on three different RNA samples purified from human brain, muscle, and breast tissues. Cloning and sequencing of DNA products revealed the presence of two transcripts that matched the genomic sequence (Fig. S1B, arrows). The shortest 3' UTR sequence corresponds to a proximal cleavage site positioned at nt 167 in the 3' UTR, whereas the longest corresponds to a distal cleavage site located at position 633, both of them ending with a cytidine thymidine dinucleotide. Each transcript contains a canonical poly(A) signal, defined by two primary sequence elements: a AAUAAA hexamer found 13 and 11 nt upstream and a U/GU-rich region located 15 and 19 nt downstream of both cleavage sites, respectively. Furthermore, UGUA elements, which serve as additional anchors for the 3' end processing machinery, are also present in the two transcripts (29).

Another characteristic of this 3' UTR sequence is the presence of a short insertion 11 nt downstream of the stop codon (boxed in Fig. S1B). This insertion is conserved in all primate mRNA sequences identified (*Pan troglodytes*, *Macaca mulatta*, *Callithrix jacchus*, *Pongo abelii*, and *Gorilla gorilla*), but absent in other mammal sequences. It matches the sequence of the Alu right arm (in the sense orientation) followed by a poly(A) stretch (19 to 29 A). A second insertion specific to primate sequences, not investigated in this study, is found close to the mRNA 3' end.

**Functional Motifs Within the 3' UTR.** To assess the importance of both the Alu insertion and the proximal and distal poly(A) signals for gene control, the AspRS mRNA 3' UTR or different variants were fused directly downstream of the stop codon of the luciferase (*Luc*) reporter gene (Fig. 1B). All constructs include the last 23 nt from the AspRS ORF because this sequence is partially complementary to the Alu insertion, an observation that was confirmed by in vitro probing experiments. Constructs were transfected into mouse Hepa 1–6 cells and the luciferase activity was measured (Fig. 1B). In comparison with the WT AspRS 3' UTR (WT-3' UTR), substantial stimulation of luciferase activity was observed when (i) the Alu insertion is mutated (mut-3' UTR) or deleted ( $\Delta 73$ -3' UTR) and (ii) the proximal poly(A) signal is removed [poly(A)1-3' UTR]. On the contrary, shortening the 3' UTR at position 213, downstream of the proximal poly(A) signal (sh-3' UTR), leads to a strong reduction in luciferase expression.

In order to measure specifically the amount of the corresponding mRNAs, we used Northern blotting and the RACE Poly(A) Test (RACE-PAT) (30). RACE-PAT products derived from *Luc* fusions were further digested with DraII and two specific fragments, corresponding to the longest transcript, were detected: a 210-bp fragment and a 280-bp fragment ending with the poly(A) tail (Fig. 1A and B). The shortest transcript was not detected by this technique, indicating that it is highly unstable. Variations in mRNA expression level correlate with luciferase activity; this suggests that *Luc* expression is controlled by mRNA concentration, which in turn is governed by distinct motifs within its 3' UTR: The region 167–633 stabilizes the mRNA, whereas the Alu insertion, as well as the proximal poly(A) signal, destabilizes the mRNA.

**Three Striking tRNA<sup>Asp</sup> Isodecoders Are Encoded by the Human Genome.** Remarkably, we observed that the Alu insertion present in the AspRS mRNA 3' UTR is complementary to the tRNA<sup>Asp</sup> sequence. A total of 19 human tRNA<sup>Asp</sup> sequences are referenced in the genomic tRNA database (<http://lowelab.ucsc.edu/GtRNAdb/>), all corresponding to a unique isoacceptor (GUC anticodon), but nine tRNA<sup>Asp</sup> isodecoders: tRNAs with the same anticodon, but different cores (31) (Fig. 2A). Three families can



**Fig. 2.** Sequences of *Homo sapiens* tRNA<sup>Asp</sup> genes and characterization of the most divergent isodecoders. (A) Sequences of tRNA<sup>Asp</sup> genes registered in the genomic tRNA database (<http://lowelab.ucsc.edu/GtRNAdb/>) are listed with their numbering and chromosome localization. The general 2D organization of tRNAs is specified on the top: "Acc" stands for acceptor arm, "D domain" for D arm and loop, "Ac domain" for anticodon arm and loop, "Δ" symbolizes the variable region, and "T domain" stands for T arm and loop. Nucleotides in tRNA<sup>Asp</sup> isodecoders that differ from the most frequent tRNA<sup>Asp</sup> sequence are underlined. A unique tRNA<sup>Asp7</sup> homologue was identified in three primate genomes, namely *M. mulatta*, *P. troglodytes*, and *G. gorilla* (boxed). (B) Detection of putative complexes between tRNA isodecoders and AspRS mRNA. Gel shift assays were performed with radio-labeled tRNA<sup>Asp</sup> 69, 7, 48, or 3 in the presence (+) and absence (-) of 500 nM AspRS mRNA. (C) Detection of tRNA<sup>Asp7</sup> in human total RNA. The RNA product corresponding to the trna7 gene, encoded on chromosome 12, was detected only by RT-PCR. A couple of primers have been specifically designed to hybridize (arrows) on the sequences surrounding the putative A and B boxes of tRNA<sup>Asp7</sup>. For two independent experiments (a and b), PCR controls were performed on the same RNA samples (DNase treated), not subjected previously to reverse transcription. Resultant PCR products were then cloned and sequenced; they all matched tRNA<sup>Asp7</sup> sequence.

be distinguished: (i) Eleven tRNA sequences are identical and considered as the reference molecules (tRNA<sup>Asp69</sup>, 72, 75, 78, 81, 10, 12, 4, 38, 45, and 48), (ii) five contain at most 6 mutations that do not affect conserved residues (tRNA<sup>Asp46</sup>, 22, 6, 144, and 5), and (iii) three sequences differ significantly from the reference by 11 to 13 mutations (tRNA<sup>Asp3</sup>, 48, and 7). This last family displays major variations in the variable region and/or in the T-loop sequence. Indeed, tRNA<sup>Asp48</sup> has a G54 instead of the strictly conserved U54 in its T loop, whereas tRNA<sup>Asp3</sup> and 7 have both a shorter variable region made up of only 2 nt instead of 3. In addition, tRNA<sup>Asp7</sup> has a T56 instead of a C56 in the conserved motif TYC defining the T loop. Despite these major sequence variations, all three unconventional tRNA<sup>Asp</sup> sequences could still adopt, in theory, the classical cloverleaf tRNA structure; however, tertiary interactions between the D and T loops are unlikely (Fig. S2).

The reference tRNA<sup>Asp</sup> sequence (tRNA<sup>Asp69</sup>), as well as tRNA<sup>Asp3</sup>, tRNA<sup>Asp48</sup>, and tRNA<sup>Asp7</sup>, were produced by in vitro transcription and assayed for their capacity to be aminoacylated by human AspRS. Surprisingly, whereas tRNA<sup>Asp69</sup> is efficiently aspartylated, none of the other isodecoders tested are substrates for AspRS in vitro (Fig. S2), despite the conservation of the major aspartate identity nucleotides (32). All four tRNAs show large sequence complementarity with the Alu insertion of AspRS mRNA 3' UTR. Transcripts were tested for their ability to bind AspRS mRNA in vitro (Fig. 2B). Among them, only tRNA<sup>Asp7</sup> formed a stable duplex with AspRS mRNA. Indeed, its migration is delayed when the full-length AspRS mRNA is added to the binding assay. This indicates that specific sequence or structural motifs in the tRNA<sup>Asp7</sup> molecule induce a stable complex with the mRNA.

Expression of endogenous tRNA<sup>Asp7</sup> was verified in HeLa cells. It was detectable only by RT-PCR and verified by sequencing, indicating that it is expressed at a very low level (Fig. 2C).

**tRNA<sup>Asp7</sup> Binds Specifically the Alu Insertion.** Various transcripts of AspRS mRNA were tested for their capacity to bind tRNA<sup>Asp7</sup> in vitro (Fig. 3). <sup>32</sup>P-labeled tRNA<sup>Asp7</sup> transcript was incubated with increasing concentrations of (i) the full-length mRNA, (ii) the 3' UTR sequence (−23 + 633), (iii) a shorter version of the 3' UTR (−23 + 279), (iv) the wild-type Alu sequence (−23 + 54), or (v) a mutated Alu sequence. Whereas none of the mRNA variants interact with the canonical tRNA<sup>Asp69</sup> (data shown for the full-length mRNA only), all constructs, which contain the intact Alu insertion, strongly bind tRNA<sup>Asp7</sup>. Indeed, apparent  $K_D$  is about 250 nM, and only mutations within the Alu sequence hinder completely the interaction.

**tRNA<sup>Asp7</sup> Does Not Fold as a Cloverleaf.** Structure mapping experiments were performed on the tRNA<sup>Asp7</sup> transcript using enzymatic probes. Nucleases T1 and S1 are known to specifically cleave unpaired nucleotides, whereas nuclease V1 is specific for

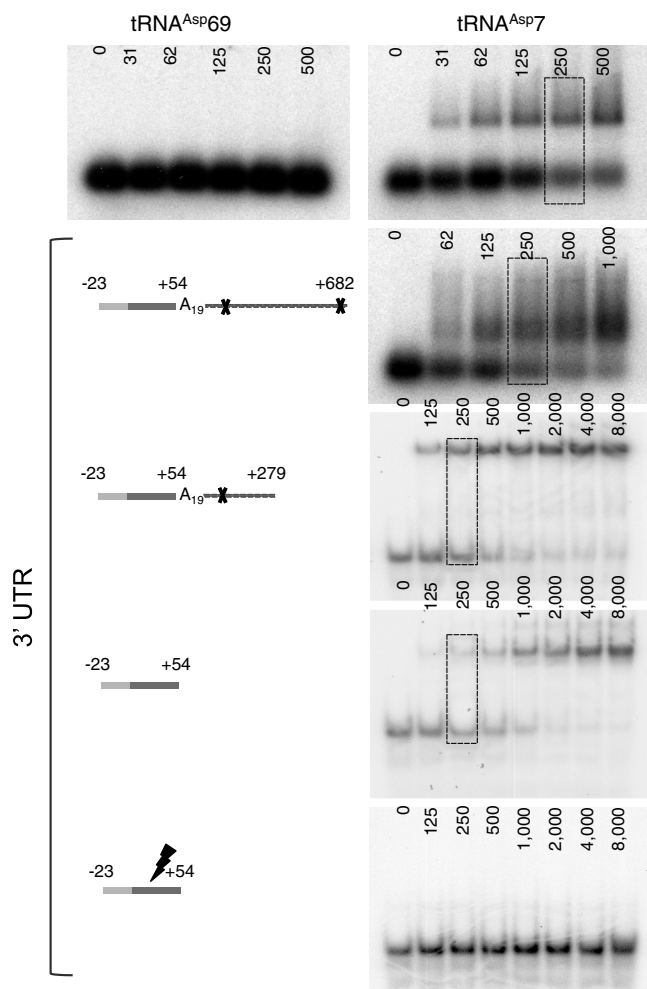
double-stranded and structured regions. Analysis of the cleavage pattern reveals that tRNA<sup>Asp7</sup> does not fold into the expected canonical cloverleaf structure [see lanes minus Alu (−23 + 54) on the autoradiogram in Fig. S3A]. Indeed, accessibilities to single-stranded probes T1 and S1 (circled nt in Fig. 4A) are clearly present in the T, anticodon (Ac), and D arms, whereas in the double-stranded probe, V1 (arrows) cuts in the D loop of the putative cloverleaf structure.

Lead (Pb<sup>2+</sup>) was further used to probe and compare the global accessibility of both tRNA<sup>Asp69</sup> (reference tRNA<sup>Asp</sup>) and tRNA<sup>Asp7</sup> (Fig. S3B). Profiles are drastically different between the two molecules. The classical profile expected for the canonical cloverleaf structure was obtained for tRNA<sup>Asp69</sup>; however, the profile corresponding to tRNA<sup>Asp7</sup> reveals clear discrepancies, confirming the previous observations obtained with enzymatic probes (Fig. 3B). In particular, the frequency of lead cleavages (lane 10 mM Pb<sup>2+</sup>) is much lower for tRNA<sup>Asp7</sup> compared to the reference tRNA, suggesting a more compact structure. Only positions 24–30 and 50–59 show stronger hydrolysis. Based on all our in vitro probing data, we propose an alternative structure (Fig. 3C), where tRNA<sup>Asp7</sup> folds into an elongated structure containing a large internal loop made up by 16 nt. The only tRNA feature conserved in this structure corresponds to the acceptor stem.

However, the presence of two unexpected V1 cleavages (nt 44 and 46) as well as the specific reactivity of residue G10, in both enzymatic (T1, Fig. S3A) and lead (Fig. S3B) probing experiments, suggest that this loop may adopt a unique structure. Interestingly, the sequence context (nt 9–12 and 44–46) allows the folding of the internal loop into an E-loop module, originally found in 5SrRNAs (33), with the residue G10 protruding from the structure. The E-loop structure is dependent on Mg<sup>2+</sup>. Thus, lead probing was repeated in the presence of decreasing concentrations of Mg<sup>2+</sup> ions (Fig. S3B). Results indicate that the absence of Mg<sup>2+</sup> ions alters significantly the cleavage profile of tRNA<sup>Asp7</sup>. Above all, the lead cleavage at positions 10 disappears and confirms the presence of an E-loop structure. Besides, lead cleavages decrease significantly at positions 13 and 14 and increase at positions 53 to 57 in the large internal loop.

Structure mapping experiments were performed on the tRNA<sup>Asp7</sup> transcript alone or bound to the Alu (−23 + 54) sequence, using enzymatic probes (Fig. S3A). The data of these footprinting experiments support not only the intermolecular interaction but also the alternative fold. Indeed, 16 potential bp can form between tRNA<sup>Asp7</sup> and the Alu sequence (Fig. 4C). In particular, disappearance of single stranded specific cleavages as well as appearance of double-stranded V1 cuts support the existence of at least two out of three blocks of interactions (1 and 3). In addition, most of the nucleotides protected by Alu (−23 + 54) binding are distributed around the E-loop platform. This suggests that this motif has a functional role in properly presenting the tRNA<sup>Asp7</sup> residues that bind to the mRNA 3' UTR. This was further confirmed by the progressive disappearance of the tRNA<sup>Asp7</sup>/Alu sequence duplex in the presence of decreasing concentrations of Mg<sup>2+</sup> ions in the test (Fig. S3C).

**Binding of tRNA<sup>Asp7</sup> Has an Allosteric Effect on 3' UTR Folding.** Long-range structural effects triggered by the binding of tRNA<sup>Asp7</sup> on the mRNA were investigated by probing nt 120 to 250 of the 3' UTR. Profiles, in the presence or absence of tRNA<sup>Asp7</sup>, were compared, revealing numerous changes in reactivity (Fig. 5A). First, the occurrence of numerous strong stops in the control experiment (without RNase treatment) indicates that the 3' UTR is structured and interrupts reverse transcription. This observation is supported by the fact that, in general, accessibilities for RNases are rare in this region. Thus, any precise interpretation would be difficult. However, we clearly see that, when tRNA<sup>Asp7</sup> is bound, the global pattern changes drastically, even 200 nt away



**Fig. 3.** Complex formation between tRNA<sup>Asp7</sup> and AspRS 3' UTR. The global organization of the different 3' UTR variants is displayed. The resolution of the gels was adapted to the length of the tested mRNA: 1% agarose gels for the longest mRNA transcripts (more than 700 nts) and 6% polyacrylamide gels for shorter ones (less than 300 nts). Increasing mRNA concentrations are indicated at the tops of the autoradiograms, and dashed boxes designate the mRNA concentration for which about 50% of the radio-labeled tRNA<sup>Asp7</sup> is shifted ( $K_D$ ). For each variant tested, tRNA<sup>Asp69</sup> was used as a control; here only the experiment performed with the full-length mRNA is presented.





However, because our results were derived from a reporter system, we cannot exclude that additional mechanisms might be involved in the endogenous mRNA regulation, such as nuclear export, specific localization, degradation, or translational control.

**An Alternative Fold of a tRNA-Derived Sequence Is Functional.** tRNA isodecoders are defined as tRNA molecules sharing the same anticodon but diverging elsewhere in their sequence (up to 274 different tRNA species are produced from 446 genes in human) (31). Like all tRNA isodecoders, the tRNA<sup>Asp7</sup> sequence contains all the information necessary to maintain the classical cloverleaf structure (Fig. S2). It also displays the functional residues identified as aminoacylation determinants (41). However, tRNA<sup>Asp7</sup> is not a substrate for aminoacylation; the main reason is that tRNA<sup>Asp7</sup> adopts an alternative fold, deprived of the structural tRNA characteristics. This is consistent with its low tRNAscan score (32.3) compared to the reference tRNA<sup>Asp69</sup> (72.9) (<http://gtrnadb.ucsc.edu/>). However, the alternative architecture is able to drive duplex formation with the AspRS mRNA 3' UTR. The E-loop module present in the alternative fold plays a major role in stabilizing and in organizing the surrounding nucleotides to properly present the residues involved in the interaction. This study indicates that among eukaryotic tRNA isodecoders, some of them may direct posttranscriptional regulation. This possibility has already been anticipated by colleagues, who proposed that such functional relevance might explain the higher diversity in human tRNA isodecoders (42).

On the other hand, Alu insertions are widely spread over the human genome, and some of them could also be the target of tRNA<sup>Asp7</sup>. However, only AspRS mRNA carries this exact Alu insertion, and a specific scaffold certainly plays also a major role to permit optimal binding between both RNA molecules.

**Toward a Function for a Primate-Specific Control of AspRS Expression: A Bundle of Evidence. Alternative Polyadenylation Regulates AspRS.** The involvement of polyadenylation in eukaryotic gene control is becoming increasingly apparent, and our findings reveal that primate cells can use a tRNA-derived ncRNA to select alternative polyadenylation sites. Thus, it controls the stability of AspRS mRNA, which encodes a protein involved in translation, a key biochemical process. The major housekeeping function of aaRSs is indeed the attachment of amino acids to their cognate tRNAs during translation (28). As aminoacylation of tRNAs is an essential step in protein synthesis, this process proceeds with high specificity (41) and necessitates the tuned expression of tRNAs and aaRSs. Several aaRS regulation pathways have been studied, especially in bacteria (43). They are not conserved in evolution and differ from one aaRS to another, the only shared occurrence being the involvement of tRNA molecules or tRNA-like structures in the regulatory mechanism (39). In the present study, because tRNA<sup>Asp7</sup> is a RNA polymerase III transcript, it could act as a sensor for tRNA transcription activity. This would link two essential processes, transcription and translation for overall coordination and homeostasis of the cell.

**AaRSs Are Housekeeping, yet Moonlighting Proteins.** The relevance of the aminoacylation activities of aaRSs is obvious. However, it has recently become clear that some human aaRSs also possess tissue-specific functions, distinct from their role in translation. Indeed, different aaRSs define pathways such as angiogenesis [TyrRS, TrpRS, and AIMP1 (44–46)], the immune response [LysRS (47)], inflammation [GluProRS (48)], apoptosis [GlnRS (49)], but also neural development [AIMP1, GlyRS, AIMP2 (50–52)], and tumorigenesis [LeuRS (53)]. The switch between the canonical aminoacylation function of an aaRS and other alternative functions might rely on a different process: (i) the oligomerization state, (ii) the subcellular localization, or (iii) the differential expression of the protein. In our case, the existence

of such a sophisticated regulatory mechanism controlling AspRS expression suggests that AspRS could also have additional functions involved in physiological processes.

**Toward Tissue-Specific Expression of AspRS.** We propose that the present regulation mechanism would control AspRS tissue-specific accumulation in a primate-specific regulation pathway such as immunity, development, or tumorigenesis. In an attempt to identify such a putative pathway, we explored the large amount of microarray data available (<http://www.ncbi.nlm.nih.gov/sites/GDSbrowser>), searching for reliable results where only human AspRS mRNA would be overexpressed, compared to other aaRS mRNAs. We found four conditions where AspRS mRNA indeed accumulates (2-fold increase for AspRS mRNA compared to an average 2-fold reduction for the other aaRS mRNAs). Interestingly, the relevant microarray data concern different cell types, such as macrophages (GDS2036), B lymphocytes (GDS1807), breast cancer cells (GDS2758), and aortic smooth muscle cells (GDS3112). Yet these different tissues commonly overexpress AspRS mRNA when subjected to hypoxia. Hypoxia is a common condition present in many pathological situations, including inflamed tissues, malignant tumors, atherosclerotic plaques, chronic venous insufficiency, healing wounds, etc. For example, tumor-associated macrophages respond to hypoxia by up-regulating a broad array of genes, encoding proteins that promote the proliferation, invasion, and metastasis of cells as well as angiogenesis (reviewed in refs. 54 and 55). These data support the necessity for the control of AspRS accumulation in hypoxic tissues.

## Conclusion

The fraction of tRNA isodecoder genes increases all along the phylogenetic spectrum, and the expansion of Alu elements is characteristic of primate genomes. For primates, this opens large perspectives regarding function and structure for the most divergent sequences and reveals the reservoir of potential regulatory functions associated with tRNA- and Alu-derived sequences. Finally, the observation of this original posttranscriptional regulation mechanism points toward the discovery of a unique physiological pathway in which the contribution of AspRS is implicated besides its usual roles in translation.

## Methods

**RT-PCR.** The cDNA of tRNA<sup>Asp7</sup> was amplified from HeLa cells total RNA using the OneStep RT-PCR (Qiagen) following the supplier's protocol. Two gene specific primers were designed: 5'-GGACTCAAACCTCCAGTC-3' was first used for the reverse transcription and then together with 5'-TCCTTGTTAGTATAGTGGTGAG-3' for the PCR reaction. The PCR products were cloned into pCR2.1 (PCR cloning kit from Qiagen) and sequenced.

**3' RLM-RACE PCR.** RLM-RACE by RLM-RACE PCR was run on three different human RNA samples (brain, breast, and muscle, Ambion). Reverse transcription and nested PCR reactions were performed using the first choice RLM-RACE kit (Ambion). PCRs were run with two primers complementary to the 3' adaptor sequence and two nested primers hybridizing (i) at the end of the AspRS mRNA ORF sequence (5'-CCATGTTCCCTCGTGATCCCAAACG-3') and (ii) directly downstream of the 19 A ending the Alu insertion in the mRNA 3' UTR (5'-GTAACCTGCTAGTGACAGCTGTACTTTAGG-3'). PCR-amplified fragments were cloned into pDrive cloning vector (PCR cloning kit, Qiagen) and sequenced.

**Preparation of Templates for in Vitro Transcription.** Human tRNA<sup>Asp</sup> transcripts were produced in vitro using the "transzyme" system (56). The transcripts were purified on native 12% PAGE and electroeluted (Schleicher and Schuell). DNA sequences encoding the full-length AspRS mRNA (2,215 nt including the ORF and both the 5' and 3' noncoding regions) as well as the 3' UTR (-23 + 682) were PCR-amplified from a cDNA library (human muscle cDNA library, Invitrogen) and cloned between EcoRI and XhoI in pXJ41 downstream the T7 RNA polymerase promoter. Plasmids were linearized with XhoI before in vitro transcription (57). The reaction was stopped by acidic phenol/chloroforme extraction and transcripts were purified on a Sephadex G-25 column (NAP-5, GE Healthcare). Shorter 3' UTR variants, 3' UTR (-23 + 279)

and Alu (−23 + 54), were cloned into pUC119 between BamHI and EcoRI restriction sites, downstream of the T7 RNA polymerase promoter. In the mutated version of the Alu insertion sequence <sub>31</sub>GATGACAG<sub>38</sub> was replaced by <sub>31</sub>TAGTCACT<sub>38</sub>. Plasmids were linearized with EcoRI before in vitro transcription and transcripts were purified on a native 12% gel.

**Aspartylation Assays.** In vitro aspartylation of human tRNA<sup>Asp</sup> transcripts was performed as described previously (58).

**Band-Shift Assays.** <sup>32</sup>P 3′-labeled tRNA<sup>Asp7</sup> (1,000 cpm/μL) (59) was incubated for 20 min at 37 °C in 25 mM Tris-HCl pH 7.5, 10 mM MgCl<sub>2</sub>, 50 mM NaCl, 10% glycerol, 40 ng/μL 5SRNA with increasing concentrations of mRNA variants (30 to 8,000 nM). Bound and unbound RNA molecules were separated either on a 1% agarose gel, with Tris borate EDTA buffer containing 1 mM MgCl<sub>2</sub> for 2 h at 50 V [full-length mRNA and 3′ UTR (−23 + 682)] or on a 6% acryl/bisacrylamide [(37.5/1) gel 1 × TBE, for 90 min at 140 V [3′ UTR (−23 + 279) and Alu (−23 + 54)].

**Structural Mapping Procedures.** Structural mappings of tRNA<sup>Asp</sup> transcripts and AspRS mRNA were done using chemical and enzymatic probes (Pb<sup>2+</sup> and T1, S1 and V1 nucleases). RNA digestions were performed in 25 mM Tris-HCl pH 7.5, 10 mM MgCl<sub>2</sub>, and 50 mM NaCl, whereas lead-induced cleavages were performed in 50 mM Tris acetate pH 7.5, 5 mM Mg acetate, 50 mM KOH acetate, and 1 mM lead acetate (60).

tRNA transcripts were <sup>32</sup>P 5′-labeled (61) and purified on NAP5. This molecule was used in (i) direct probing experiments and (ii) in RNA/RNA footprinting experiments where tRNA<sup>Asp7</sup> was preincubated in the presence of 20 pmoles of Alu (−23 + 54) transcript prior modifications.

Longer transcripts such as 3′ UTR (−23 + 682) or full-length mRNA were first incubated with structural probes, and cleavage products were then revealed by reverse transcription as described elsewhere (62). Radio-labeled fragments were separated on a denaturing (8 M urea) 12% PAGE.

**Constructs for Mammalian Cell Transfection.** Wild-type 3′ UTR (−23 + 682) of AspRS mRNA and its variants were PCR-amplified from genomic DNA and inserted into pRL-CMV vector (Promega) between the XbaI and BamHI, downstream of the *Rluc* reporter gene, instead of the SV40 late polyadenylation signal. The cloned PCR fragment displays 50 nts below the distal polyadenylation signal, in order to conserve all the downstream elements necessary for an efficient maturation. Luciferase activity assays were achieved as indicated in the dual-Luciferase Reporter Assay System manual (Promega). Statistical significances were assessed using the Welch's *t* test. Probability *P* values were calculated using the GraphPad Prism software.

tRNA<sup>Asp</sup> genes were PCR-amplified from genomic DNA (previously digested with KpnI and SphI). The resulting 1,450-bp amplicon spans positions −700 to +700 relative to the tRNA sequence and incorporates all the

upstream promoter and downstream regions required for efficient transcription and maturation. Regarding tRNA<sup>Asp7</sup>, another construct was designed with only 40 residues upstream from the tRNA sequence to remove potential upstream external transcriptional promoters. PCR products were then cloned into pSK(−), between the NotI and XhoI.

**Cell Culture and Transfection.** HeLa (human epithelial) and Hepa 1–6 (mouse hepatocytes) cells were cultured in DMEM, supplemented with 10% fetal calf serum, and 20 U/mL streptomycin and penicillin at 37 °C with 5% CO<sub>2</sub>. At 70–80% confluence, cells were transfected using DreamFect Gold reagent (OZBioscience). Hepa 1–6 were transfected using 100 ng pRL-CMV constructs, 50 ng pGL3 vector (transfection efficiency control), and 10 ng pSK-tRNA per well (24 well plate format). After 6 h incubation, the medium was changed, and the cells were further grown for 20 h. Mouse hepatocytes were chosen as a heterologous system because neither tRNA<sup>Asp7</sup> nor the Alu insertion is present; hence the studied regulatory mechanism is absent in this cellular context.

**Nucleus Purification.** HeLa cells were transfected with 6 μg DNA (tRNA<sup>Asp7</sup>) per 10-cm dish for 24 h. Nuclei were purified from the cytosolic fraction as described at <http://www.lamondlab.com/f7nucleolarprotocol.htm>. RNA was extracted from both nuclear and cytosolic fractions following the Tri-Reagent protocol (Sigma-Aldrich).

**RACE-PAT.** cDNAs were synthesized from total RNA using a poly-T/adaptor primer (gcgagacagaataatcagactcactatagg(t)<sub>12</sub>). The amount of mRNAs within the sample was determined by a standard PCR reaction in the presence of [ $\alpha$ -<sup>32</sup>P] dCTP.

**Northern Blotting.** Total cellular RNAs (15 μg) were run on 1% agarose gel and analyzed by Northern blotting using the NorthernMax kit (Ambion) and <sup>32</sup>P labeled probes prepared with the NonaPrimer kit (Quantum Appligene).

**Premature Chain Termination Primer Extension.** Different tRNA<sup>Asp</sup> entities were specifically detected by chain termination primer extension (63), using dideoxynucleotides complementary to tRNA<sup>Asp69</sup> (5′-GGGAATC-GAACCCCGTCT-3′), tRNA<sup>Asp7</sup> (5′-AGTCTCCACATGACAG-3′), and SnRNA<sub>6</sub> (5′-GCAGGGGCCATGCTAATC-TTCTCTGTAT-3′).

**ACKNOWLEDGMENTS.** We thank Pascale Romby, Catherine Schuster, and Eric Westhof for advice and stimulating discussions; Oliver Miller and Marie Messmer for critical reading of the manuscript. This work was supported by grants from the Centre National de la Recherche Scientifique, Université de Strasbourg, and "La Ligue Contre Le Cancer."

- Batzer MA, Deininger PL (2002) Alu repeats and human genomic diversity. *Nat Rev Genet* 3:370–379.
- Hasler J, Strub K (2006) Alu elements as regulators of gene expression. *Nucleic Acids Res* 34:5491–5497.
- Tomilin NV (1999) Control of genes by mammalian retroposons. *Int Rev Cytol* 186:1–48.
- Brosius J (1999) Genomes were forged by massive bombardments with retroelements and retrosequences. *Genetica* 107:209–238.
- Yakovchuk P, Goodrich JA, Kugel JF (2009) B2 RNA and AluRNA repress transcription by disrupting contacts between RNA polymerase II and promoter DNA within assembled complexes. *Proc Natl Acad Sci USA* 106:5569–5574.
- Mariner PD, et al. (2008) Human Alu RNA is a modular transacting repressor of mRNA transcription during heat shock. *Mol Cell* 29:499–509.
- Landry JR, Medstrand P, Mager DL (2001) Repetitive elements in the 5′ untranslated region of a human zinc-finger gene modulate transcription and translation efficiency. *Genomics* 76:110–116.
- Britten RJ (2006) Almost all human genes resulted from ancient duplication. *Proc Natl Acad Sci USA* 103:19027–19032.
- Hasler J, Samuelsson T, Strub K (2007) Useful 'junk': Alu RNAs in the human transcriptome. *Cell Mol Life Sci* 64:1793–1800.
- Sobczak K, Krzyzosiak WJ (2002) Structural determinants of BRCA1 translational regulation. *J Biol Chem* 277:17349–17358.
- Spence J, Duggan BM, Eckhardt C, McClelland M, Mercola D (2006) Messenger RNAs under differential translational control in Ki-ras-transformed cells. *Mol Cancer Res* 4(1):47–60.
- Vila MR, et al. (2003) Higher processing rates of Alu-containing sequences in kidney tumors and cell lines with overexpressed Alu-mRNAs. *Oncol Rep* 10:1903–1909.
- Stuart JJ, Egly LA, Wong GH, Kaspar RL (2000) The 3′ UTR of human MnSOD mRNA hybridizes to a small cytoplasmic RNA and inhibits gene expression. *Biochem Biophys Res Commun* 274:641–648.
- Krichevsky AM, Metzger E, Rosen H (1999) Translational control of specific genes during differentiation of HL-60 cells. *J Biol Chem* 274:14295–14305.
- Vidal F, et al. (1993) Coordinated posttranscriptional control of gene expression by modular elements including Alu-like repetitive sequences. *Proc Natl Acad Sci USA* 90:208–212.
- Kramerov DA, Vassetzky NS (2005) Short retroposons in eukaryotic genomes. *Int Rev Cytol* 247:165–221.
- Sorek R, Ast G, Graur D (2002) Alu-containing exons are alternatively spliced. *Genome Res* 12:1060–1067.
- Levanon D, Groner Y (2004) Structure and regulated expression of mammalian RUNX genes. *Oncogene* 23:4211–4219.
- Lin-Ling C, DeCervo JN, Carmichael GG (2008) Alu element-mediated gene silencing. *EMBO J* 27:1694–1705.
- Gong C, Maquat LE (2011) lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3′ UTRs via Alu elements. *Nature* 470:284–290.
- An HJ, Lee D, Lee KH, Bhak J (2004) The association of Alu repeats with the generation of potential AU-rich elements (ARE) at 3′ untranslated regions. *BMC Genomics* 5:97.
- Roy-Engel AM, et al. (2005) Human retroelements may introduce intragenic polyadenylation signals. *Cytogenet Genome Res* 110:365–371.
- Wang ET, et al. (2008) Alternative isoform regulation in human tissue transcriptomes. *Nature* 456:470–476.
- Koscielny G, et al. (2009) ASTD: The alternative splicing and transcript diversity database. *Genomics* 93(3):213–220.
- Tian B, Hu J, Zhang H, Lutz CS (2005) A large-scale analysis of mRNA polyadenylation of human and mouse genes. *Nucleic Acids Res* 33:201–212.
- Yan J, Marr TG (2005) Computational analysis of 3′-ends of ESTs shows four classes of alternative polyadenylation in human, mouse, and rat. *Genome Res* 15:369–375.
- Lutz CS (2008) Alternative polyadenylation: A twist on mRNA 3′ end formation. *ACS Chem Biol* 3:609–617.
- Ibba M, Francklyn C, Cusack S (2005) *The aminoacyl-tRNA synthetases* (Landes Biosciences, Georgetown, TX).
- Millevoi S, Vagner S (2010) Molecular mechanisms of eukaryotic pre-mRNA 3′ end processing regulation. *Nucleic Acids Res* 38:2757–2774.



