Detection of a Subset of Posttranscriptional Transfer RNA Modifications in vivo With a Restriction Fragment Length Polymorphism-Based Method

Thomas F. Wulff[1,‡], Rafael J. Argüello[2,‡], Marc Molina Jordàn[1], Helena Roura Frigolé[1], Glenn Hauquier[1], Liudmila Filonava[1], Noelia Camacho[1], Evelina Gatti[2,3], Philippe Pierre[2,3], Lluís Ribas de Pouplana[1,4,] and Adrian G. Torres[1,*].*

1 Institute for Research in Biomedicine (IRB Barcelona), The Barcelona Institute of Science and Technology, Parc Científic de Barcelona, C/ Baldiri Reixac 10, 08028 Barcelona, Catalonia, Spain.

2 Centre d'Immunologie de Marseille-Luminy, Aix-Marseille Université U2M, Inserm, U1104, CNRS UMR7280, 13288 Marseille, France.

3 Institute for Research in Biomedicine (iBiMED) and Aveiro Health Sciences Program University of Aveiro, 3810-193 Aveiro, Portugal.

4 Catalan Institution for Research and Advanced Studies (ICREA), P/ Lluis Companys 23, 08010 Barcelona, Catalonia, Spain.

‡These authors contributed equally.

* Corresponding authors:

*Email: adriangabriel.torres@irbbarcelona.org (AGT). Tel: +34934034867. Fax: +34934034870

*Email: lluis.ribas@irbbarcelona.org (LRdP). Tel: +34934034868. Fax: +34934034870

ABSTRACT

Transfer RNAs (tRNAs) are among the most heavily modified RNA species. Posttranscriptional tRNA modifications (ptRMs) play fundamental roles in modulating tRNA structure and function, and are being increasingly linked to human physiology and disease. Detection of ptRMs is often challenging, expensive and laborious. Restriction Fragment Length Polymorphism (RFLP) analyses study the patterns of DNA cleavage after restriction enzyme treatment; and have been used for the qualitative detection of modified bases on messenger RNAs. It is known that some ptRMs induce specific and reproducible base 'mutations' when tRNAs are reverse transcribed. For example, inosine, which derives from the deamination of adenosine, is detected as a guanosine when an inosine-containing tRNA is reverse transcribed, PCR-amplified and sequenced. ptRM-dependent base changes on RT-PCR amplicons generated as a consequence of the reverse transcription reaction might create or abolish endonuclease restriction sites. The suitability of RFLP for the detection and/or quantification of ptRMs has not been studied thus far. Here we show that different ptRMs can be detected at specific sites of different tRNA types by RFLP. For the examples studied we show that this approach is able to reliably estimate the modification status of the sample, a feature that can be useful in the study of the regulatory role of tRNA modifications on gene expression.

INTRODUCTION

Transfer RNAs (tRNAs) are among the most heavily modified RNA species. Posttranscriptional modifications on tRNAs (ptRMs) may act as identity determinants for aminoacylation (e.g. posttranscriptional 5'-end guanosine addition on $tRNA^{His}$ is essential for tRNA charging); they can affect codon decoding by expanding or restricting codon recognition; or contribute to tRNA structure and stability [1]. In recent years ptRMs have become a strong focus of research, as mounting evidence points to roles for tRNA modifications in human pathologies such as cancer (e.g. 5-methylcytosine), diabetes (e.g. 2-methylthio-N6-threonylcarbamoyladenosine), neurological disorders (e.g. 5-methoxycarbonylmethyl-2-thiouridine), and mitochondrial-linked diseases (e.g. 5-taurinomethyluridine modifications on mitochondrial tRNAs), among others [2, 3]. Unfortunately, the complex set of potential tRNA modifications makes their identification in any given tRNA extremely challenging.

Available methods for the analysis of posttranscriptional RNA modification include: i) radiolabeling of RNA species followed by selective RNA digestion with specific RNases, and analysis of migration patterns by 2D-thin-layer chromatography [4]; ii) the use of liquid chromatography coupled to mass spectrometry (LC-MS/MS) [5, 6]; and iii) methods that involve the reverse transcription (RT) of RNAs [7]. Additional methods have also been developed to study specific modifications such as the PHA6 assay in which isopentyl-N^6-A37 can be detected based on the differential hybridization of an oligonucleotide against the anticodon loop of tRNAs carrying the modification [8]; or the acryloylaminophenylboronic (APB) acid-based boronic affinity electrophoresis to detect queuosine modifications on tRNAs [9].

Methods based on radiolabeling or mass spectrometry are usually laborious, low-throughput, and require specialized and expensive laboratory equipment and facilities. These methods may not

be easily applied to study ptRMs on specific tRNA residues, or to monitor physiological levels of endogenous ptRMs *in vivo*. In addition, RNA samples need to homogeneous, in some cases large amounts of starting material is required, and a specific expertise on experimental optimization and data interpretation is often necessary. Therefore, there is still a need for developing novel techniques for ptRM detection that would overcome some of these limitations.

During RT, modified RNA bases may be either read by the reverse transcriptase as if unmodified, as a different base (resulting in a change in the resulting DNA sequence), or may block the advance of the reverse transcriptase causing an arrest in the transcript. These possible outcomes of the RT reaction are dependent on the nature of the RNA modification [7]. Restriction Fragment Length Polymorphism (RFLP) is a technique that studies the patterns of DNA cleavage after restriction enzyme treatments [10]. It has been widely used in the past as a medium- to high-throughput technique to detect genetic polymorphisms, define DNA marker loci, perform genome mapping or gene tagging, and to construct genomic linkage relationships in human pedigrees [11]. More recently, RFLP has been used to detect editing events on messenger RNAs such as adenosine-to-inosine (A-to-G mutation on the PCR product) and cytidine-to-uridine (C-to-T mutation on the PCR product) modifications [12-14]; confirming that cDNA mutations caused by the reverse transcriptase can generate or abolish endonuclease restriction sites on PCR amplicons.

DEVELOPMENT OF THE METHOD

We envisaged that RFLP could be used to study ptRMs (**Figure 1**). Analyses of small-RNAseq data in search for mutated sites that are distinguishable from simple base-calling errors have proven useful to map several different ptRMs at genomic scale [15-18]; suggesting that several ptRMs induce the incorporation of alternative nucleotides in the generated cDNA when the reverse

transcriptase reads a modified tRNA. In the past, we validated these findings by using small-RNAseq to monitor the levels of inosine modifications on tRNAs in human cell lines [19].

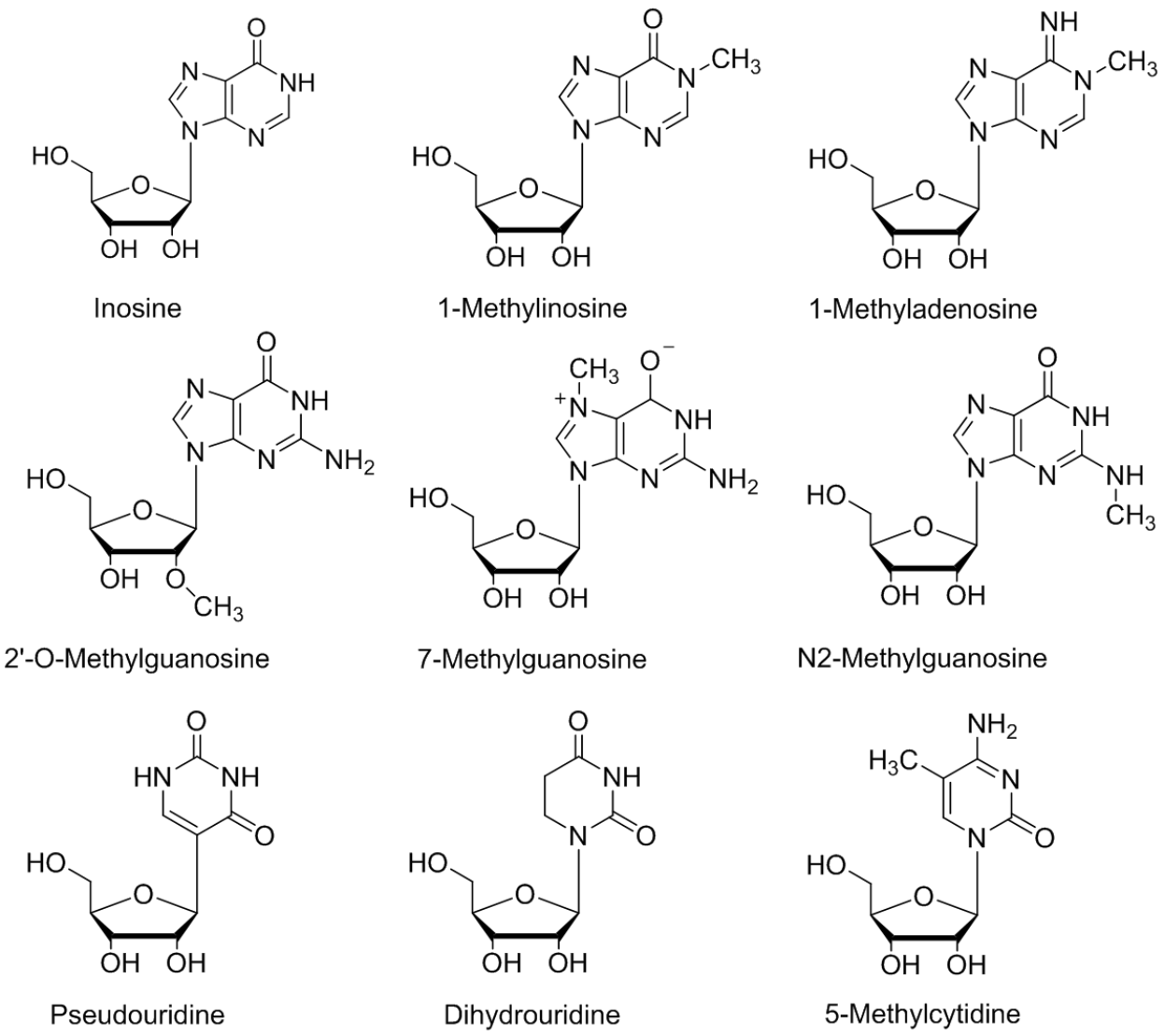

Figure 1. Posttranscriptional tRNA modifications discussed in this manuscript. Inosine, 1-Methylinosine and 1-Methyladenosine nucleosides are derived from Adenosine. 2'-O-Methylguanosine, 7-Methylguanosine and N2-Methylguanosine are derived from Guanosine.

Pseudouridine and Dihydrouridine are derived from Uridine. 5-Methylcytidine is derived from Cytidine.

In this context, RFLP could be beneficial because, by avoiding the need of performing RNA sequencing, it would allow detection of inosine and other ptRMs in standard laboratory facilities, rapidly, and in a cost-effective manner. The RFLP method to detect ptRMs that we present here has four simple steps: RT, PCR amplification with specific primers, incubation of the PCR product with a specific restriction enzyme, and detection of the cleavage patterns by standard gel electrophoresis (**Figure 2**). Here, we have applied this method to study different ptRMs in different tRNA types, using *in vitro* and *in vivo* tRNA samples. We additionally show direct applications of the method to study, for example, the activity of a tRNA modification enzyme *in vitro*, or to monitor the levels of ptRMs *in vivo*. We believe that this method offers a new and useful approach to the detection and study of tRNA modifications.

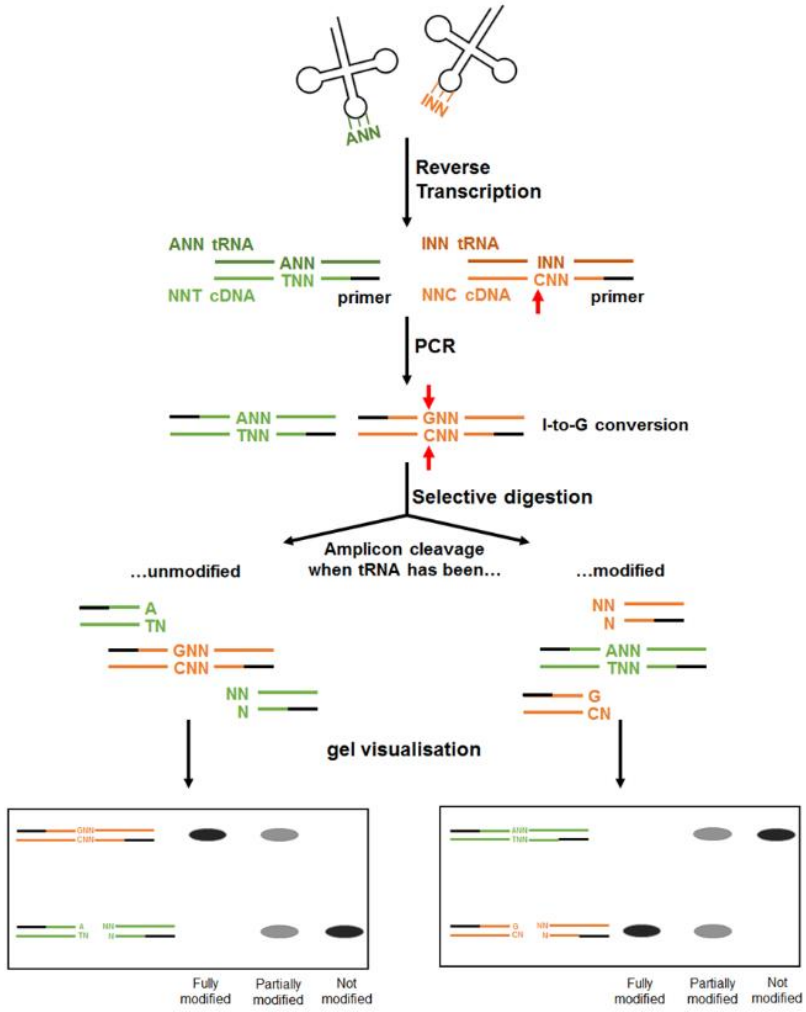

Figure 2. Schematic representation of the RFLP method for ptRM detection. Detection of inosine at the first anticodon position (position 34; I34) on tRNAs is shown as an example. During RT, the reverse transcriptase reads I34 as a G and incorporates a C on the cDNA strand. Therefore, after PCR amplification, the original I34 position is read as a G34. The obtained amplicon is then digested with a specific restriction enzyme that would recognize and cleave the target in a modified/unmodified residue-dependent manner. The products of digestion are then resolved by gel electrophoresis and the proportion of cleaved:uncleaved amplicon can be quantified to estimate the levels of modified tRNA.

VALIDATION OF THE METHOD

Detection of inosine from *in vitro* tRNA samples. As a proof of concept, we first decided to apply the RFLP method to *in vitro* tRNA samples, and to focus on inosine detection (**Figure 1**). Inosine is present at the first position of the anticodon (position 34; 'wobble' position) of bacterial $tRNA^{Arg}_{ACG}$ and 7-8 eukaryotic tRNAs. It is the result of a deamination reaction of adenosine, catalyzed in eukaryotes by the heterodimeric adenosine deaminase acting on tRNAs (hetADAT). Inosine is one of the few essential ptRMs described, as it is required to expand the number of codons that genetically encoded A34-containing tRNAs can recognise [20].

In vitro transcribed human $tRNA^{Ala}_{AGC}$, a natural substrate of hetADAT, was incubated with purified human hetADAT. The reaction product was then reverse transcribed and amplified by PCR. Deamination was confirmed by sequencing the PCR amplicon, a commonly used strategy [19, 21-25] (**Figure 3A**). As expected, A34 was detected on the *in vitro* transcribed tRNA (**Figure 3A** upper panel), but was fully converted to I34 (detected as G34) upon incubation with hetADAT (**Figure 3A** bottom panel). An *in silico* search for restriction enzymes revealed that *Bpu1102I* (*BlpI*) would recognize and cleave a PCR amplicon derived from unmodified $tRNA^{Ala}_{AGC}$ (A34) but would not cleave the $tRNA^{Ala}_{IGC}$-derived counterpart (G34) (**Figure 3B**). Indeed, upon incubation with *Bpu1102I*, only the PCR amplicon derived from the unmodified tRNA was cleaved (**Figure 3C**). Importantly, no cleavage was observed for the modified tRNA, consistent with the 100% modification status of the *in vitro* transcribed tRNA upon incubation with hetADAT observed by sequencing (**Figure 3A**). Similar results were obtained for *in vitro* transcribed $tRNA^{Thr}_{AGU}$ (**Supporting Information Figure S1 A-C**). This set of experiments shows that results obtained with the RFLP method are in agreement with those obtained by sequencing.

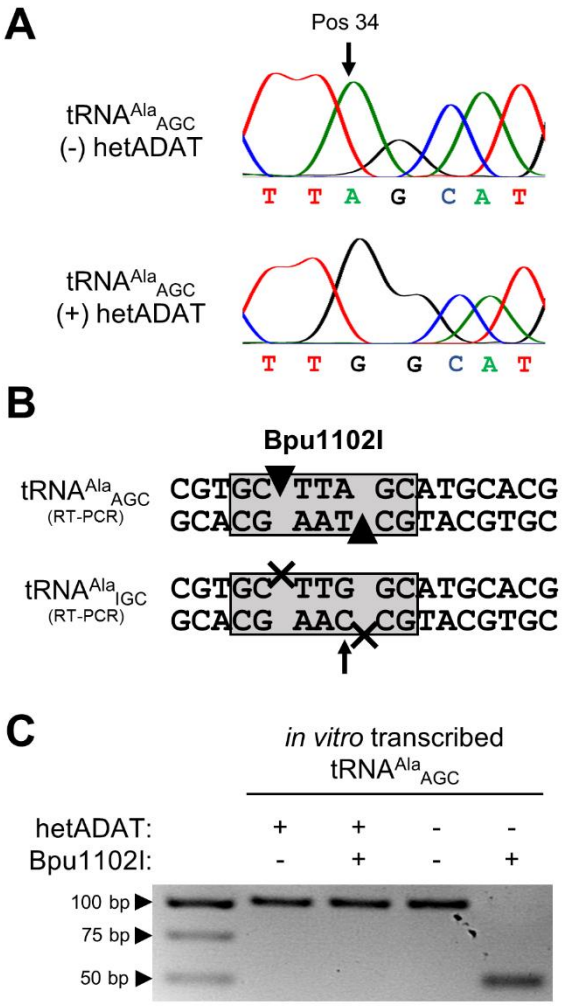

Figure 3. Evaluation of *in vitro* deamination by human hetADAT on *in vitro* transcribed tRNA$^{Ala}_{AGC}$. (A) Sequencing spectrum for *in vitro* transcribed tRNA$^{Ala}_{AGC}$ incubated or not with human hetADAT. I34 is detected as G34. (B) Schematic representation of the recognition site for *Bpu1102l* for amplicons derived from unmodified (upper panel) or modified (lower panel) tRNA. The arrow indicates the modification site (Position 34 on the tRNA). Bpu1102l will cleave the amplicon only when it is derived from unmodified tRNA$^{Ala}_{AGC}$. (C) Evaluation of *in vitro* deamination by human hetADAT on *in vitro* transcribed tRNA$^{Ala}_{AGC}$ using the RFLP method. After cDNA synthesis, PCR amplification and digestion with *Bpu1102l*, samples were resolved in

a 2% agarose gel. As expected, the full length amplicon is fully cleaved only when A34 is not modified.

To further characterize this method, we tested its sensitivity towards I34 detection and its potential to be used in a semi-quantitative manner. Fully modified- and fully unmodified-*in vitro* transcribed $tRNA^{Ala}_{AGC}$ were mixed at known proportions. The levels of I34 modification were next evaluated by RFLP and the observed gel bands were quantified. **Figure 4A** shows a representative gel obtained for this type of experiment. Gel quantification of three independent replicates revealed a linear correlation ($R^2=0.99$) between the expected and the observed levels of I34 (**Figure 4B**) suggesting that the method can be semi-quantitative. Importantly, the method proved to be sensitive enough to reliably detect low levels of modification (expected: 1% I34; observed: 2.80 ± 1.05 % I34) (**Figure 4B**).

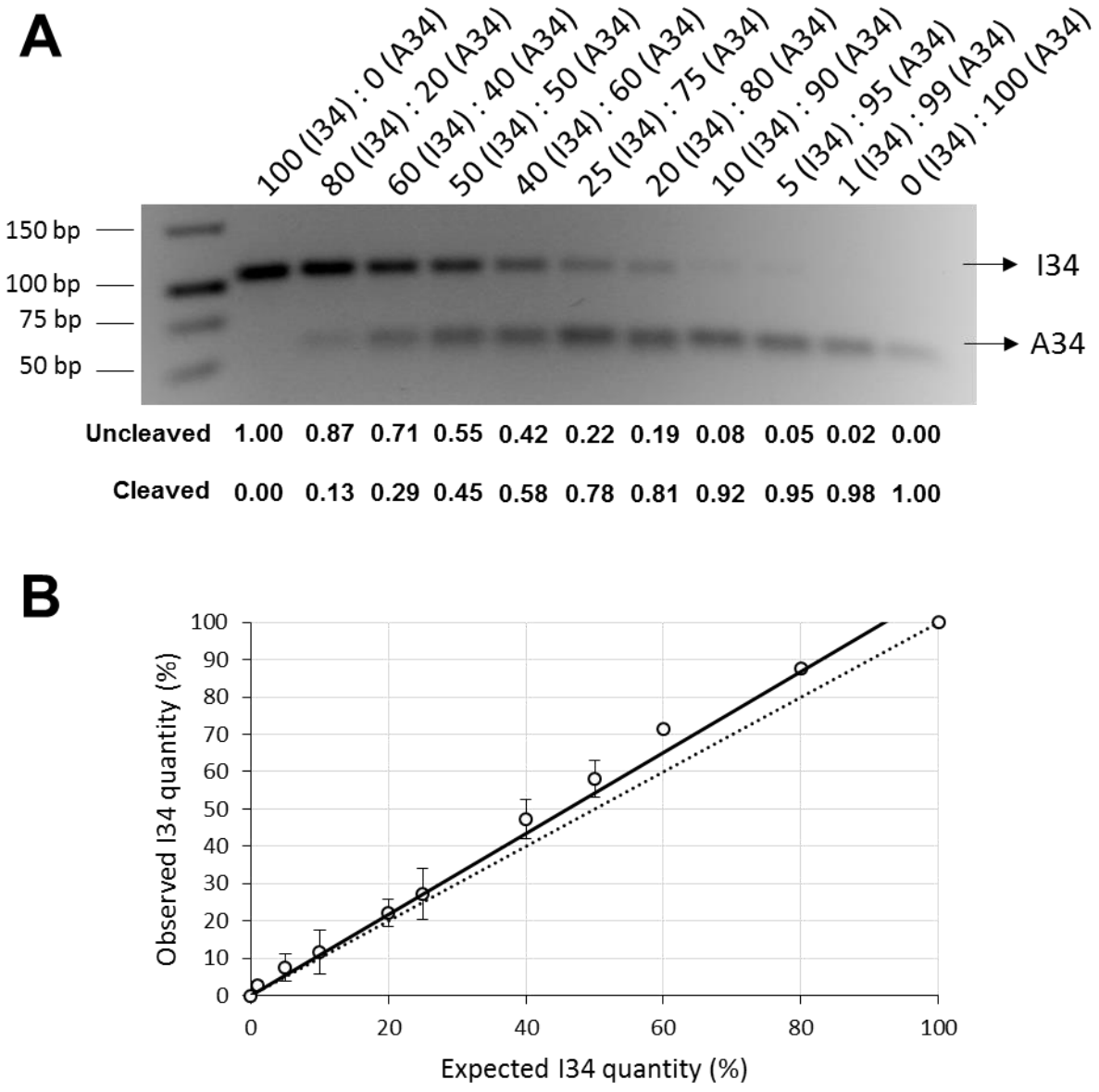

Figure 4. Evaluation of the sensitivity and accuracy of the RFLP method for ptRM detection. Fully modified- (I34) and fully unmodified- (A34) *in vitro* transcribed $tRNA^{Ala}_{AGC}$ were mixed at the indicated proportions. The RNA mix was then reverse transcribed, PCR amplified, digested with *Bpu1102I*, and ran in a 2% agarose gel. The cleavage pattern was quantified using ImageJ. (A) A representative gel showing the obtained quantification results. Band quantification was normalized to total lane intensity (uncleaved + cleaved band). (B) Plot of the estimated levels of I34 on *in vitro*

transcribed $tRNA^{Ala}$ by RFLP (continuous regression line: $y=1.084x$; $R^2=0.99$) against the theoretically expected levels of modification (dotted regression line: $y=x$; $R^2=1$). Mean and standard deviations are shown for three independent replicates. Curve fitting was performed following a linear regression model.

We next applied the RFLP method to monitor the effect of different reaction conditions on hetADAT activity by varying the pH and magnesium (Mg^{2+}) concentrations of the reaction buffer. As shown in **Figure 5**, optimal deamination of $tRNA^{Ala}_{AGC}$ was achieved at pH 8.0 and 1 mM Mg^{2+}. These results suggest that hetADAT activity *in vitro* is sensitive to pH and Mg^{2+} concentrations; and illustrate the potential use of the method in reaction optimization.

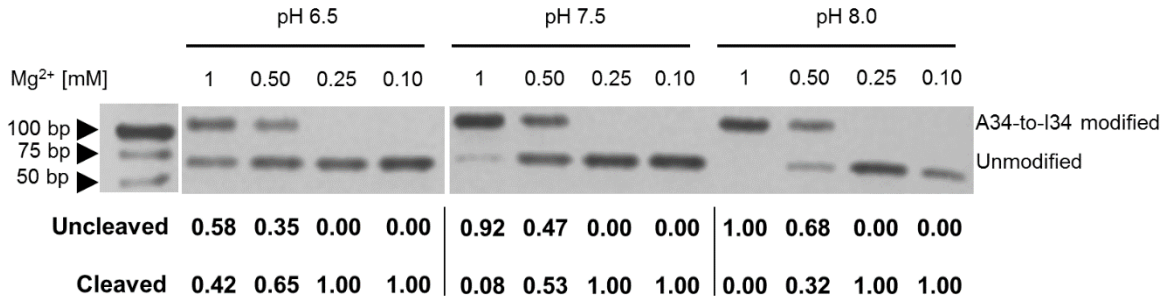

Figure 5. Evaluation of *in vitro* deamination efficiency by human hetADAT on *in vitro* transcribed $tRNA^{Ala}_{AGC}$. Different reaction conditions (variable pH and Mg^{2+}) were evaluated using the RFLP method. Samples were resolved in a 2% agarose gel. Band quantification was normalized to total lane intensity (uncleaved + cleaved band). Optimal reaction conditions were observed at pH 8.0 and 1 mM Mg^{2+}.

Detection of inosine from *in vivo* tRNA samples. We then applied the RFLP method to the detection of ptRMs in tRNAs from *in vivo* samples. HeLa cells expressing a short hairpin RNA (shRNA) against human ADAT2, the catalytic subunit of hetADAT, were generated. We have previously shown that knockdown (KD) of ADAT2 in a human cell line causes partial downregulation of I34 levels on all 8 tRNA substrates of hetADAT [19]. Using quantitative real time PCR, ADAT2 was found downregulated in shADAT2 cell lines as compared to cells expressing a non-target control shRNA (shCV) (**Figure 6A**). The RFLP method for ptRM detection was then used to evaluate the levels of I34 on endogenous $tRNA^{Ala}_{AGC}$ in these cell lines. As previously shown (**Figure 3**), the amplicon derived from $tRNA^{Ala}_{IGC}$ should not be cleaved by *Bpu1102I*. As expected, a higher degree of amplicon cleavage was observed in shADAT2 cells, consistent with lower levels of I34 for this tRNA (**Figure 6B**). Similar results were obtained for endogenous $tRNA^{Thr}_{AGU}$, another natural substrate of hetADAT (**Supporting Information Figure S1D**).

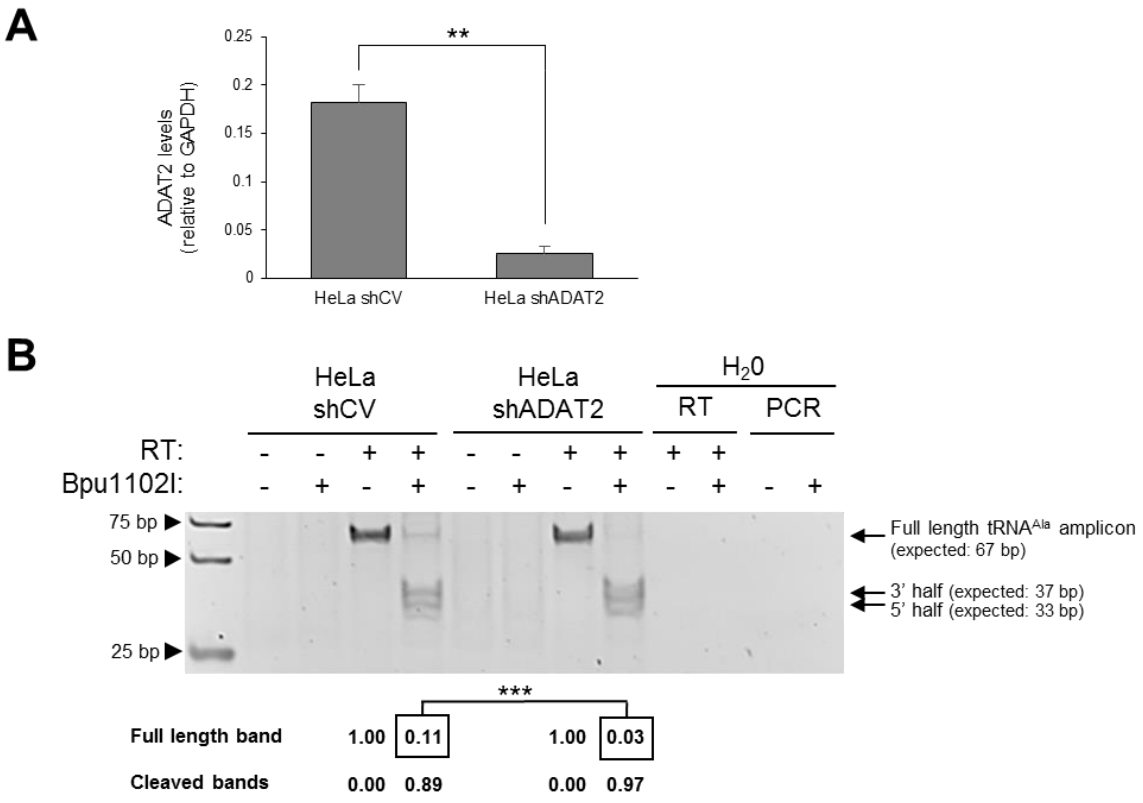

Figure 6. Evaluation of *in vivo* I34 levels on $tRNA^{Ala}_{AGC}$ in HeLa cells. (A) Quantitative Real-Time PCR for ADAT2 in HeLa cells expressing a control shRNA (shCV) or an ADAT2 shRNA (shADAT2). ADAT2 levels are expressed relative to GAPDH. Statistical significance was obtained with a *t*-test (n = 3; **: p-val < 0.01). (B) Evaluation of *in vivo* I34 levels on $tRNA^{Ala}_{AGC}$ in HeLa shCV and HeLa shADAT2 cells using the RFLP method. A representative experiment is shown. Samples were resolved in a 15% polyacrylamide gel. RT (-): no reverse transcriptase in RT reaction; H_2O RT (+): reverse transcription using water as template (no RNA template); H_2O PCR: PCR reaction using water as template (no cDNA template). Band quantification normalized to total lane intensity (uncleaved + cleaved bands) is shown. Quantification of 'Cleaved bands' corresponds to the sum of both 3' and 5' amplicon halves. Statistical significance was obtained with a *t*-test (***: p-val < 0.001) and corresponds to the comparison between the remaining uncleaved band after *Bpu1102I* digestion for HeLa shCV and HeLa shADAT2 cells on three independent replicates (n = 3).

tRNAs genetically encoded with A34 are known to be almost fully modified to I34 when fully matured [20, 26]. It was therefore surprising to observe a significant amount of amplicon cleavage for $tRNA^{Ala}_{AGC}$ and $tRNA^{Thr}_{AGU}$ in HeLa shCV cells (**Figure 6B** and **Supporting Information Figure S1D**). We sequenced the obtained amplicons for these samples and confirmed that A34 was present on both tRNAs (**Supporting Information Figure S2**). After ruling out possible contamination of the PCR reaction with genomic DNA (**Figure 6B**) we concluded that part of the obtained amplicons were derived from unmodified precursor tRNA species. It is well known that RT of fully matured tRNAs is challenging and that in complex RNA samples, there is a bias towards reverse transcribing precursor tRNAs over mature tRNAs [19, 27-29]. Despite amplifying

precursor tRNAs, our data shows a trend for lower levels of I34 in RNA samples derived from ADAT2 KD cells. This demonstrates that the RFLP method can be used for the detection and relative semi-quantification of ptRMs *in vivo* (see Discussion section). To facilitate the use of the RFLP method to detect I34 levels on tRNAs, we have compiled the tRNA substrates that could potentially be studied with this technique and the required restriction enzymes that would recognize either the A34 or I34 version of the PCR amplicon, for all model organisms where I34 has been studied (**Table 1**) [19, 21-25].

Table 1. $tRNA_{ANN}$ substrates of A34-to-I34 editing in model organisms and restriction enzymes that could be used to detect such modification using the RFLP method.

tRNA undergoing A34-to-I34 editing		Restriction enzyme for amplicon cleavage[#]	
		Unmodified (ANN RNA / ANN cDNA)	Modified (INN RNA / GNN cDNA)
Escherichia coli K12			
Arginine	ACG	-	*TseI, Fnu4HI*
Homo sapiens			
Alanine	AGC	*Bpu1102I*	-
Arginine	ACG	-	*AciI*
Isoleucine	AAU	-	-
Leucine	AAG	*MseI*	-
Proline	AGG	*DdeI, Bpu10I*	-
Serine	AGA	*XspI*	-
Threonine	AGU	*XspI*	-
Valine	AAC	-	-

Arabidopsis thaliana			
Alanine	AGC	*Bpu1102I*	-
Arginine	ACG	*AccI, Hpy166II*	-
Isoleucine	AAU	-	-
Leucine	AAG	*MseI*	-
Proline	AGG	*DdeI*	-
Serine	AGA	*XspI*	*BsrI*
Threonine	AGU	-	-
Valine	AAC	-	*Hpy188I*
Trypanosoma brucei			
Alanine	AGC	*Bpu1102I*	*BglI*
Arginine	ACG	*TspGWI*	*AciI*
Isoleucine	AAU	-	*Hpy188I*
Leucine	AAG	*MseI*	*MnlI* (2 cutter)
Proline	AGG	*Bsu36I, DdeI*	*StyI*
Serine	AGA	*XspI*	*BsrI*
Threonine	AGU	*XspI*	*StyD4I, BstNI*
Valine	AAC	*MseI*	-
Saccharomyces cerevisiae			
Alanine	AGC	*Bpu10I, DdeI*	*BsaJI, StyI*
Arginine	ACG	-	*TseI, Fnu4HI, BbvI*
Isoleucine	AAU	-	-
Proline	AGG	*Bpu10I, DdeI*	-
Serine	AGA	-	-
Threonine	AGU	*SpeI, XspI*	*BsrI, TspRI*
Valine	AAC	*MseI*	-

Schizosaccharomyces pombe			
Alanine	AGC	*Bpu10I, DdeI*	*BsaJI, StyI*
Arginine	ACG	-	*AciI* (2 cutter)
Isoleucine	AAU	-	*Hpy188I*
Leucine	AAG	*MseI*	-
Proline	AGG	*DdeI*	-
Serine	AGA	*BfaI*	*BsrI*
Threonine	AGU	*BfaI*	*StyD4I, BstNI*
Valine	AAC	*MseI*	-

[#] Only one isoschizomer is shown.

Detection of other ptRMs from *in vivo* tRNA samples. We performed an *in silico* search for additional restriction sites in unmodified human $tRNA^{Ala}_{AGC}$ sequences that may be useful for the detection of additional ptRMs. A restriction site for the enzyme *SphI* (GCATGC) was found in positions 35-40, a region that can incorporate 1-methylinosine at position 37 (m^1I37), pseudouridine at position 38 (Ψ38) and 2'-O-methylguanosine at position 39 (Gm39) [30] (**Figure 1**). We found that under conditions where the PCR amplicon derived from *in vitro* transcribed $tRNA^{Ala}_{AGC}$ (unmodified tRNA) is fully cleaved, the amplicon obtained from *in vivo* $tRNA^{Ala}_{AGC}$ is only partially cleaved (**Figure 7A**). This is consistent with at least one of the aforementioned modifications inducing a detectable 'mutation' on the PCR amplicon. Of note, m^1I37 has been reported to induce such a mutation as detected by RNAseq [15, 19]. To show the robustness of the method using *in vivo* tRNA samples, we performed three independent replicates for this experiment (**Figure 7B**). The results suggest that the RFLP method can detect other modifications

besides inosine, although in this particular case it is not possible to determine the exact nature of the modification(s) (see Discussion section).

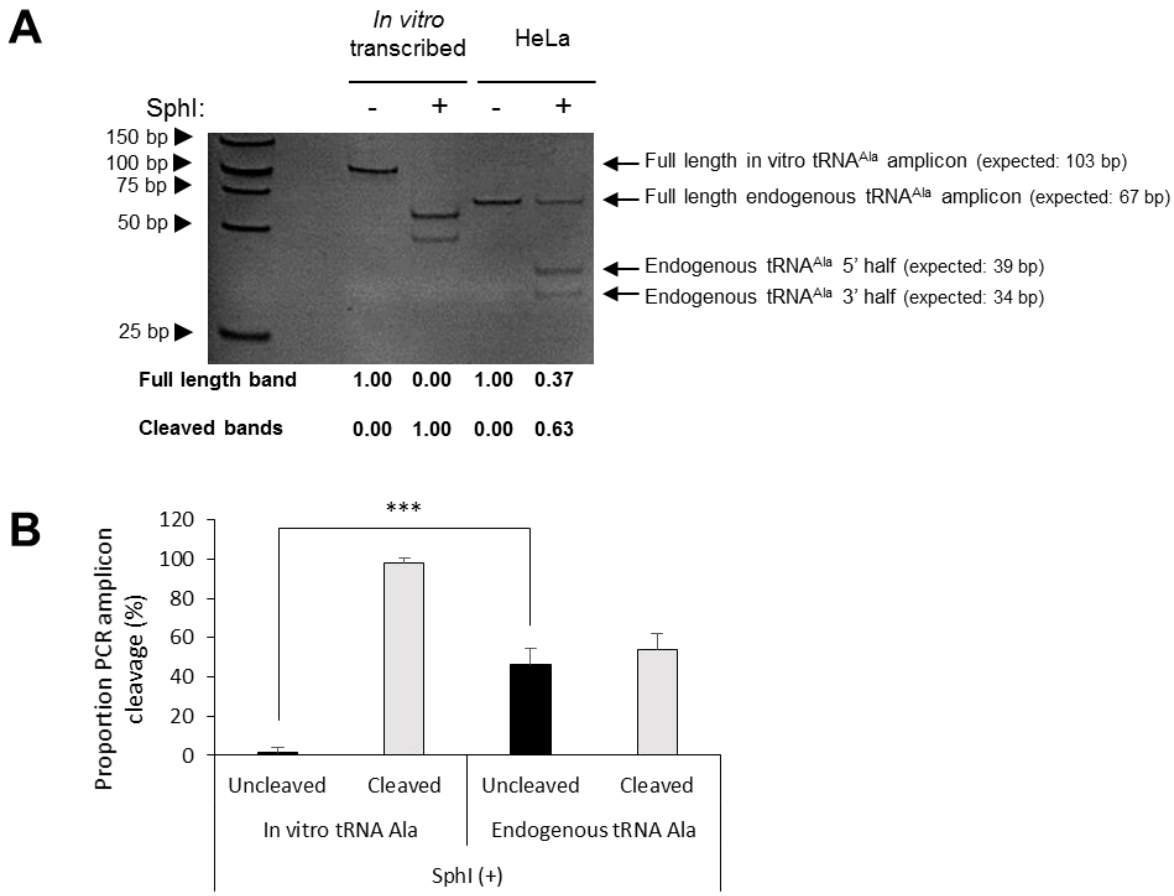

Figure 7. Evaluation of *in vivo* ptRMs on $tRNA^{Ala}_{AGC}$ in HeLa cells using the RFLP method. (A) *SphI* recognizes a region of unmodified $tRNA^{Ala}_{AGC}$ comprising positions 35-40 of the tRNA expected to be modified to m^1I37, Ψ38, and Gm39. Samples were resolved in a 15% polyacrylamide gel. *In vitro* $tRNA^{Ala}_{AGC}$ and the endogenous $tRNA^{Ala}_{AGC}$ were amplified using different primers (see **Supporting Information Table S1**); hence the difference in their full length size. Expected sizes of the endogenous $tRNA^{Ala}_{AGC}$ amplicon 5'- and 3'-halves are shown. Bands were quantified and normalized to total lane intensity (uncleaved + cleaved bands). Quantification of 'Cleaved bands' corresponds to the sum of both 3' and 5' amplicon halves. (B) Quantification

of three independent experiments as that shown in (A). Statistical significance was obtained with a *t*-test (n = 3; ***: p-val < 0.001).

Lastly, we wanted to study RNA modifications on a different tRNA substrate. We chose $tRNA^{Val}$, and used *in vitro* transcribed $tRNA^{Val}_{AAC}$ as unmodified tRNA control. We reverse transcribed and PCR amplified *in vitro* transcribed $tRNA^{Val}$, and endogenous $tRNA^{Val}$ from HeLa cells, using the same primers. We sequenced the PCR amplicon obtained from the endogenous tRNA and found that two isoacceptors were being amplified: $tRNA^{Val}_{AAC}$ and $tRNA^{Val}_{CAC}$ (**Supporting Information Figure S3**); this is expected as the sequences of these two isoacceptors are very similar to each other [31]. Notably, at position 34 both C (from $tRNA^{Val}_{CAC}$) and G were found. Given that there are no predicted $tRNA^{Val}_{GAC}$ genes in the human genome [31], this result is consistent with A34 from $tRNA^{Val}_{AAC}$ being deaminated to I34 (detected as G34).

We next searched for restriction enzymes that would recognize other regions of $tRNA^{Val}$ expected to be modified. We chose two enzymes: *HpyCH4IV* (ACGT) and *AvaII* (GGWCC), that target sites likely to contain N2-methylguanosine (m^2G) and Ψ; and 7-methylguanosine (m^7G), dihydrouridine (D) and 5-methylcytidine (m^5C), respectively (**Fig 8A**). m^2G has been reported to induce RT-dependent mutations as observed by RNAseq [17]. As a negative control, we chose *Taq$^{\alpha}$I* (TCGA). This enzyme targets a site for 1-methyladenosine (m^1A) and Ψ. Since m^1A is known to cause a hard-stop on reverse transcription [28, 29, 32, 33] and Ψ is considered an RT-silent modification [34] (i.e. it does not induce a hard-stop on the RT and the reverse transcriptase recognizes it as an unmodified U), the RFLP method should not be able to detect these modifications and therefore, a PCR amplicon derived from tRNAs modified in this region should be fully cleaved by the restriction enzyme. All three enzymes tested fully cleaved the PCR amplicon derived from the *in*

vitro (unmodified) tRNA^{Val}. As expected, the PCR amplicon derived from endogenous tRNA^{Val} was also fully cleaved by Taq$^{\alpha}$I. However, this amplicon was only partially cleaved by *HpyCH4IV* and *AvalI*, consistent with partial tRNA modification on these regions (**Figure 8B** and **Figure 8C**).

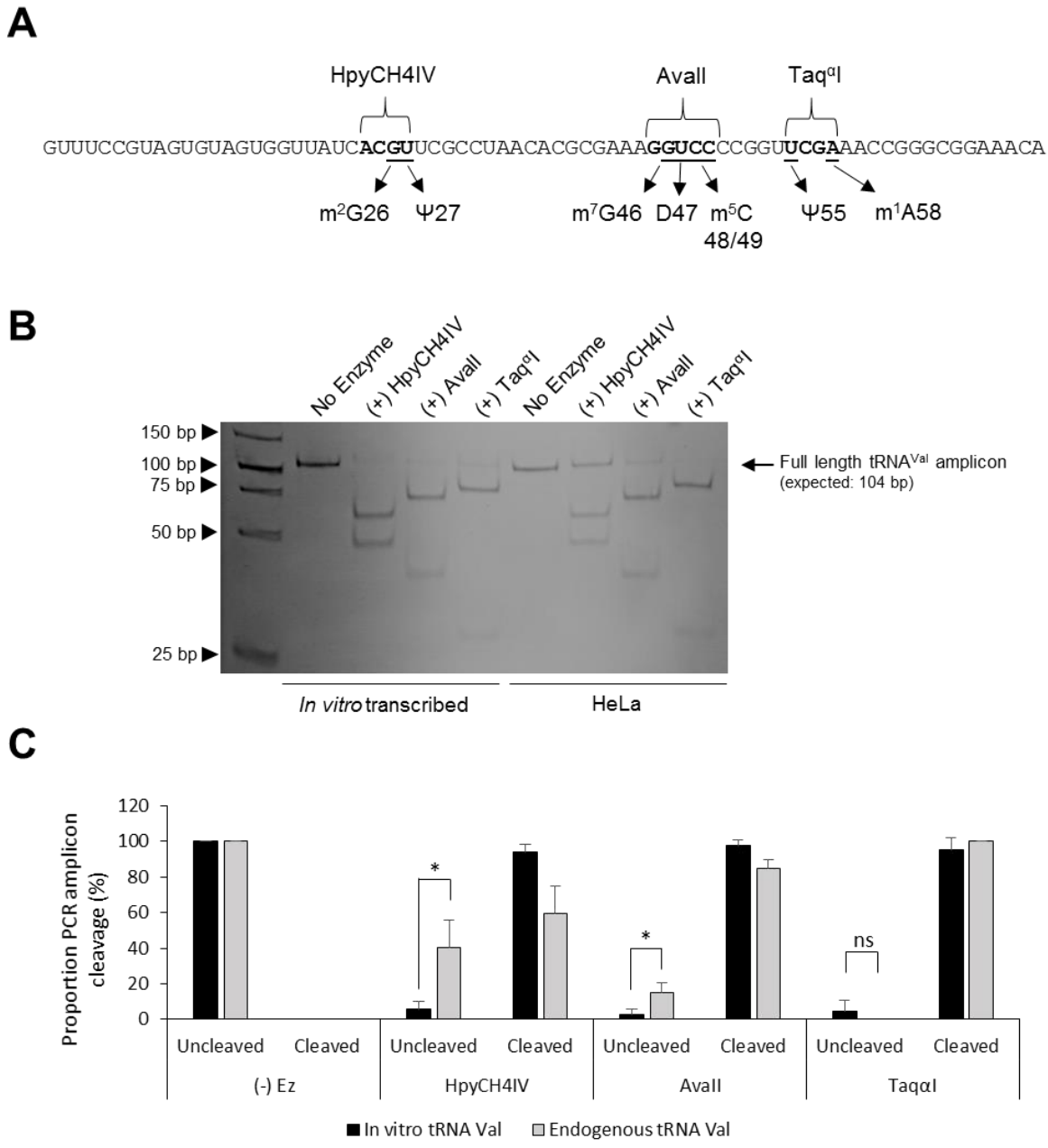

Figure 8. Evaluation of *in vivo* ptRMs on tRNA^{Val} in HeLa cells using the RFLP method. (A) tRNA^{Val} sequence. Bold residues show the recognition site for the indicated restriction enzyme.

Underlined residues are those expected to be modified within the restriction site. Expected modifications are shown: m^2G26 (N2-Methylguanosine at position 26), Ψ27 (Pseudouridine at position 27), m^7G46 (7-Methylguanosine at position 46), D47 (Dihydrouridine at position 47), $m^5C48/49$ (5-Methylcytidine at positions 48 and 49), Ψ55 (Pseudouridine at position 55), m^1A58 (1-Methyladenosine at position 58); see also **Figure 1**. (B) Restriction enzyme cleavage patterns observed for PCR amplicons derived from *in vitro* transcribed $tRNA^{Val}$ (*in vitro* transcribed) or endogenous $tRNA^{Val}$ (HeLa). Samples were resolved in a 15% polyacrylamide gel. Expected band sizes: full length 104 bp (arrow), *HpyCH4IV* 5'-half 45 bp and 3'-half 59 bp, *AvaII* 5'-half 66 bp and 3'-half 38 bp, $Taq^{\alpha}I$ 5'-half 76 bp and 3'-half 28 bp. (C) Quantification of three independent experiments as that shown in (B). Bands were quantified and normalized to total lane intensity (uncleaved + cleaved bands). Quantification of 'Cleaved bands' corresponds to the sum of both 3' and 5' amplicon halves. Statistical significance was obtained with a *t*-test (n = 3; *: p-val < 0.05; ns: not statistically significant).

DISCUSSION

In recent years RNA modifications in tRNAs have emerged as a fundamental aspect of tRNA function and regulation, and highly relevant to human physiology [2, 3]. Moreover, while in the past ptRMs were considered rather static, mounting evidence now suggests that they can be dynamic [5, 6]. This prompts for the development of new techniques to efficiently detect, quantify, and monitor ptRMs. The RFLP method for ptRM detection can semi-quantitatively detect ptRMs on specific tRNA bases, and in different tRNA species. The method involves four steps: cDNA synthesis, PCR amplification with specific primers, digestion of the obtained amplicon with restriction

endonucleases, and detection of cleavage patterns by gel electrophoresis (**Figure 2**). Altogether, results can be obtained in less than 10 hours (see Step-by-step method description section).

We applied the RFLP method to evaluate the levels of ptRMs on *in vitro* transcribed tRNAs and on endogenous *in vivo* tRNA species. We show that, at least for I34 detection, the method can be used in a semi-quantitative manner, and has enough sensitivity to detect the modification even when it is present in only 1 % of the tRNA population (**Figure 4**). We monitored A34-to-I34 editing by human hetADAT *in vitro* and under different reaction conditions using this approach, and showed that this method can be used to determine the optimal parameters for the deamination reaction (**Figure 5**). It is likely that the RFLP method could also be useful to perform kinetic (time course) assays of hetADAT activity *in vitro*. We are currently performing a full *in vitro* characterization of hetADAT activity and we are hoping to shed light into this matter in the near future.

Two types of techniques are currently being used to detect ptRMs: LC-MS/MS and RNAseq. The principal advantages of RNAseq over LC-MS/MS methods, are the fact that small amounts of RNA are required, and that detected ptRMs can be directly assigned to the tRNA sequence [15-19, 27-29]. The major drawback of RNAseq is that it often results in the detection of partially modified precursor tRNAs, rather than fully mature tRNAs [19, 27].

In our analysis of I34 modifications on endogenous tRNAs, sequences from unmodified precursor tRNAs were also detected (**Figure 6B** and **Supporting Information Figure S2**). However, we were able to show that RFLP can be used to detect a relative decrease of I34 nucleotides in two tRNAs from HeLa cells when ADAT2 expression is silenced (**Figure 6** and **Supporting Information Figure S1D**). The partial downregulation of I34 levels in ADAT2 KD

HeLa cells observed with the RFLP method was consistent with that reported in ADAT2 KD HEK293T cells using an RNAseq strategy [19].

We also analyzed the potential of RFLP to detect other ptRMs *in vivo*. We intentionally focused on regions of tRNAs where several ptRMs might be present and have selected sites on $tRNA^{Ala}_{AGC}$ and $tRNA^{Val}_{AAC/CAC}$ in order to study m^1I, Gm, D, m^2G, m^7G, and m^5C. RFLP on all the studied sites of potential ptRMs demonstrated partial cleavage of the PCR amplicon (indicating the presence of a ptRM) (**Figure 7** and **Figure 8**). Treatment of $tRNA^{Val}$ amplicons with *HpyCH4IV* revealed the presence of m^2G26 and/or Ψ27 on *in vivo* tRNAs. Given that Ψ cannot be readily detected with the RFLP method (see above), our results indicate that RFLP is suitable to detect m^2G26 modifications (**Figure 8**). This is in agreement with the report by Iida and colleagues of m^2G sites located on the D-loop of tRNAs that are detected as C, U or A 'mutations' on PCR amplicons following deep sequencing [17].

Similar to Ψ, both D and m^5C modifications are considered RT silent ptRMs [35]. D and m^5C ptRMs are present in a region of $tRNA^{Val}$ that also contains m^7G (m^7G46, D47, m^5C48, and m^5C49 respectively) [30], and is a potential restriction site for *AvaII* (**Figure 8A**). Given the silent nature of D and m^5C, the observed partial cleavage of $tRNA^{Val}$ amplicons with *AvaII* (**Figure 8B** and **Figure 8C**) suggests that RFLP can also detect the m^7G46 modification on this tRNA.

Finally, we also assessed the detection of m^1I37, Ψ38 and Gm39 on $tRNA^{Ala}_{AGC}$ by means of amplicon cleavage by *SphI* (**Figure 7**). We detected partial amplicon cleavage of *in vivo* $tRNA^{Ala}_{AGC}$ consistent with ptRM detection on at least one of these residues. Gm39 is known to be an RT-silent ptRM [7, 35], therefore m^1I37 is likely responsible for the observed partial cleavage of the PCR amplicon. We and others have previously shown that sequencing of m^1I results mainly in A-to-T or A-to-G 'mutations' on PCR amplicons, and that such 'mutations' can be detected in

amplicons derived from precursor tRNAs [15, 19]. Altogether, these results support the use of the RFLP method to detect, *in vitro* or *in vivo*, different ptRMs such as I, m^2G, m^7G and m^1I in different tRNA species.

As mentioned above, the RFLP method for ptRM detection has some limitations. First, it requires the reverse transcription and amplification of the tRNA of interest. This is a significant problem for amplifying mature tRNAs, that may contain other ptRMs incompatible with RT reactions [7]. As discussed above, this may introduce a bias towards reverse transcribing partially modified precursor tRNAs over fully modified mature tRNAs. However, we believe that once the RT step is completed the PCR step of the RFLP method will not introduce additional biases (i.e. the obtained cDNA template contains standard DNA bases). Despite this limitation, we show that it is possible to assess the modification status of tRNAs when the ptRM is incorporated at the precursor tRNA level, as it is the case for I34 and m^1I37 (**Figure 6B**, **Supporting Information Figure S1D**, and **Figure 7**) [19].

However, we note that the detected levels of ptRM under these conditions do not quantitatively represent the physiological levels of ptRMs on fully mature tRNAs. For example, in **Figure 6B** and **Supporting Information Figure S1D,** control cells showed ~10% of I34 modification on substrates expected to be ~100% modified when fully matured. Therefore, the RFLP method can accurately quantify the levels of ptRMs on *in vitro* tRNA substrates (**Figure 4**); and it allows for a relative quantification with respect to control cells in *in vivo* samples. In the aforementioned example, shADAT2 cells show ~4 fold less I34 on $tRNA^{Ala}$ and $tRNA^{Val}$ as compared to shCV cells.

Obviously the RFLP method requires that the studied ptRM causes the incorporation of an alternative nucleotide on the reverse transcribed cDNA (e.g. Ψ could not be detected; **Figure 8**).

Here we have focused on four modifications that have this effect: m^2G26, I34, m^1I37 and m^7G46. Other ptRMs that share this property, and might be detected by RFLP, are cytidine deaminations (C-to-T), 1-methylguanosine (m^1G) (G-to-T or C), 3-methylcytosine (m^3C) (C-to-T or A), N2-N2-dimethylguanosine ($m^2{}_2G$) (G-to-A or T), N4-acetylcytidine (ac^4C) (C-to-A), and N6-threonylcarbamoyladenosine (t^6A) (A-to-C) [15-17, 25, 26, 36].

Some of these ptRMs have been reported to cause hard-stops on RT, so they may not be easily detected by RFLP; as in the case for m^1A58 modification (**Figure 8**). However, it is important to note that the ptRM-mediated misincorporation of nucleotides during cDNA synthesis has been reported to be sequence context dependent [33, 36]; therefore a comprehensive and systematic analysis of reverse transcriptase behavior when encountering ptRMs in a sequence context dependent manner is needed.

Finally, RFLP is limited by the need of a restriction enzyme that will differentially recognize the region of interest where ptRMs induce a change of sequence in their transcripts. In general, we have found useful restriction enzymes for most of our targets of interest. As an example, **Table 1** gives an estimation of how frequently a restriction enzyme exists to identify A34-to-I34 conversion on tRNAs in different organisms. This limitation may be overcome with the advent of CRISPR/Cas9-based methods developed to facilitate the *in vitro* digestion of DNA at sites for which no endonuclease is available [37].

Although RFLP is designed to evaluate the modification status of tRNA residues known (or suspected) to be modified (e.g. **Figure 7** and **Figure 8**), it can also be used as an initial screening step to detect new modification sites. In this case, we would recommend using as many restriction enzymes as possible to ensure a significant tRNA sequence coverage. Additionally, the selected restriction enzyme should recognize the unmodified residue and, if full cleavage is not observed,

the presence of a modification can be suspected, and other independent means (e.g. LC-MS/MS using the purified specific tRNA type of interest) may be used to characterize it.

STEP-BY-STEP DESCRIPTION OF THE RFLP METHOD FOR ptRM DETECTION

RNA was reverse transcribed with the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems; Part number: 4368814). RT of tRNA species was carried out using 100 ng *in vitro* transcribed tRNA or 2 µg total RNA as RT template in a final reaction volume of 20 µL. These amounts of starting material are significantly lower than what used for other methods. For example, an optimized LC-MS/MS method for ptRM detection has been recently reported in which 12 µg of total tRNA pool or 2 µg of a single tRNA was required as starting material [38]. For experiments depicted in **Figure 4** fully modified- and fully unmodified-*in vitro* transcribed $tRNA^{Ala}_{AGC}$ were mixed at the indicated proportions for a total amount of 100 ng RNA (e.g. 80 (I34): 20 (A34) proportions corresponds to 80 ng of *in vitro* transcribed $tRNA^{Ala}_{IGC}$ mixed with 20 ng of *in vitro* transcribed $tRNA^{Ala}_{AGC}$, prior to RT reaction). 10 pmol of specific RT-primers (**Supporting Information Table S1**) were used for each tRNA species. RT-cycle: 10 min at 25 ºC; 60 min at 37 ºC; 5 min at 85 ºC. Estimated time required for this protocol step: 1 hour 30 minutes.

3 µL total RNA-derived cDNA were used for a standard 50 µL PCR reaction. 10 µL *in vitro* transcribed tRNA-derived cDNA were used for a standard 100 µL PCR reaction. Forward and Reverse primers used for PCR amplification of each target are detailed in **Supporting Information Table S1**. PCR cycle: 5 min at 94 ºC (1x cycle); 30 sec at 94 ºC, 15 sec at target-specific annealing temperature (**Supporting Information Table S2**), 15 sec at 72 ºC (35x cycles); 5 min at 72 ºC (1x cycle). Estimated time required for this protocol step: 2 hours.

When required, PCR amplicons were purified using the NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel; Ref number: 740609.250) and were sequenced using their respective forward and reverse PCR primers as sequencing primers (GATC Biotech). PCR amplicons were digested with 5 U of the appropriate restriction enzyme in 1x recommended reaction buffer; following the manufacturer's protocol, except that we extended the recommended incubation time for an additional 2 hours to ensure full cleavage of the target PCR amplicon. Enzymes used in this study were as follows: *Bpu1102I* (Thermo Scientific; Cat Number: ER0091); *XspI* and *SphI* (Takara/Clontech; Cat Number: 1095A and 1246A, respectively); *HpyCH4IV*, *AvaII* and *Taq$^{\alpha}$1* (New England Biolabs; Cat Number: R0619S, R0153S and R0149S, respectively). PCR amplicon digestions using *HpyCH4IV*, *AvaII* and *Taq$^{\alpha}$1* were carried out with a total of 30 U of enzyme for 4 hours, by adding 15 U of enzyme every 2 hours to the reaction. Estimated time required for this protocol step: between 3 hours and 6 hours depending on the restriction enzyme digestion conditions.

Digested samples were then resolved in 2 % agarose gels or 15 % polyacrylamide gels (ran for 1 hour at 120 V). Similar results were obtained when using either gel type (**Supporting Information Figure S1C**), but in some cases higher resolution may be achieved by PAGE. When required, gel bands were quantified using ImageJ. In our experience the full RFLP method for ptRM detection can be usually performed in less than 10 hours.

ASSOCIATED CONTENT

Supporting Information Available

Experimental procedures. Evaluation of A34-to I34 editing of $tRNA^{Thr}_{AGU}$ *in vitro* and *in vivo*. Sequencing spectrum of amplicons derived from endogenous $tRNA^{Ala}_{AGC}$ and $tRNA^{Val}$.

Oligonucleotides (DNA oligonucleotides) used in this study. Combination of PCR primers used for amplifying the targets described in this work and their respective annealing temperatures used for PCR reactions. This material is available free of charge (file type, PDF).

FUNDING SOURCES

This work was supported by the Spanish Ministry of Economy and Competitiveness [FPDI-2013-17742 to AGT, BIO2015-64572-R to LRdP]; by the Bayer Science and Education Foundation [Otto Bayer Fellowship to TFW]; by the Spanish Ministry of Education, Culture and Sports [FPU2014-01315 to HRF]; by the European Union's Seventh Framework Programme for research, technological development and demonstration [IRBPostPro2.0 600404 to LF]; and by the "Investissements d'Avenir" French government program [A*MIDEX project ANR-11-IDEX-0001-02, and CSI A*MAAP-INME-14.01 to PP and EG].

ACKNOWLEDGMENT

We would like to thank the Ribas de Pouplana lab for helpful discussions and, in particular Enric Ros for assistance in generating Figure 1.

REFERENCES

1. Piñeyro, D.; Torres, A. G.; Ribas de Pouplana, L. (2014) in *Fungal RNA Biology* (Sesma, A.; Von der Haar, T., Eds.), pp 233-267, Springer International Publishing, Switzerland.
2. Torres, A. G.; Batlle, E.; Ribas de Pouplana, L. (2014) *Trends Mol Med.* 20, 306-314.
3. Torres, A. G.; Ribas de Pouplana, L. (2016) in *Modified Nucleic Acids in Biology and Medicine* (Jurga, S.; Erdmann, V. A.; Barciszewski, J., Eds.), pp 1-26, Springer International Publishing, Switzerland.
4. Grosjean, H.; Droogmans, L.; Roovers, M.; Keith, G. (2007) *Methods Enzymol.* 425, 55-101.

5. Chan, C. T.; Dyavaiah, M.; DeMott, M. S.; Taghizadeh, K.; Dedon, P. C.; Begley, T. J. (2010) *PLoS Genet.* 6, e1001247.
6. Chan, C. T.; Pang, Y. L.; Deng, W.; Babu, I. R.; Dyavaiah, M.; Begley, T. J.; Dedon, P. C. (2012) *Nat Commun.* 3, 937.
7. Motorin, Y.; Muller, S.; Behm-Ansmant, I.; Branlant, C. (2007) *Methods Enzymol.* 425, 21-53.
8. Lamichhane, T. N.; Arimbasseri, A. G.; Rijal, K.; Iben, J. R.; Wei, F. Y.; Tomizawa, K.; Maraia, R. J. (2016) *RNA.* 22, 583-596.
9. Igloi, G. L.; Kossel, H. (1985) *Nucleic Acids Res.* 13, 6881-6898.
10. Botstein, D.; White, R. L.; Skolnick, M.; Davis, R. W. (1980) *Am J Hum Genet.* 32, 314-331.
11. Grover, A.; Sharma, P. C. (2016) *Crit Rev Biotechnol.* 36, 290-302.
12. Gu, T.; Buaas, F. W.; Simons, A. K.; Ackert-Bicknell, C. L.; Braun, R. E.; Hibbs, M. A. (2012) *PLoS One.* 7, e33720.
13. Kawahara, Y.; Ito, K.; Sun, H.; Kanazawa, I.; Kwak, S. (2003) *Eur J Neurosci.* 18, 23-33.
14. Palladino, M. J.; Keegan, L. P.; O'Connell, M. A.; Reenan, R. A. (2000) *Cell.* 102, 437-449.
15. Ryvkin, P.; Leung, Y. Y.; Silverman, I. M.; Childress, M.; Valladares, O.; Dragomir, I.; Gregory, B. D.; Wang, L. S. (2013) *RNA.* 19, 1684-1692.
16. Findeiss, S.; Langenberger, D.; Stadler, P. F.; Hoffmann, S. (2011) *Biol Chem.* 392, 305-313.
17. Iida, K.; Jin, H.; Zhu, J. K. (2009) *BMC Genomics.* 10, 155.
18. Ebhardt, H. A.; Tsang, H. H.; Dai, D. C.; Liu, Y.; Bostan, B.; Fahlman, R. P. (2009) *Nucleic Acids Res.* 37, 2461-2470.
19. Torres, A. G.; Pineyro, D.; Rodriguez-Escriba, M.; Camacho, N.; Reina, O.; Saint-Leger, A.; Filonava, L.; Batlle, E.; Ribas de Pouplana, L. (2015) *Nucleic Acids Res.* 43, 5145-5157.
20. Torres, A. G.; Pineyro, D.; Filonava, L.; Stracker, T. H.; Batlle, E.; Ribas de Pouplana, L. (2014) *FEBS Lett.* 588, 4279-4286.
21. Gerber, A. P.; Keller, W. (1999) *Science.* 286, 1146-1149.
22. Wolf, J.; Gerber, A. P.; Keller, W. (2002) *EMBO J.* 21, 3841-3851.
23. Tsutsumi, S.; Sugiura, R.; Ma, Y.; Tokuoka, H.; Ohta, K.; Ohte, R.; Noma, A.; Suzuki, T.; Kuno, T. (2007) *J Biol Chem.* 282, 33459-33465.
24. Zhou, W.; Karcher, D.; Bock, R. (2014) *Plant Physiol.* 166, 1985-1997.
25. Rubio, M. A.; Ragone, F. L.; Gaston, K. W.; Ibba, M.; Alfonzo, J. D. (2006) *J Biol Chem.* 281, 115-120.
26. Arimbasseri, A. G.; Blewett, N. H.; Iben, J. R.; Lamichhane, T. N.; Cherkasova, V.; Hafner, M.; Maraia, R. J. (2015) *PLoS Genet.* 11, e1005671.
27. Pang, Y. L.; Abo, R.; Levine, S. S.; Dedon, P. C. (2014) *Nucleic Acids Res.* 42, e170.
28. Zheng, G.; Qin, Y.; Clark, W. C.; Dai, Q.; Yi, C.; He, C.; Lambowitz, A. M.; Pan, T. (2015) *Nat Methods.* 12, 835-837.
29. Cozen, A. E.; Quartley, E.; Holmes, A. D.; Hrabeta-Robinson, E.; Phizicky, E. M.; Lowe, T. M. (2015) *Nat Methods.* 12, 879-884.
30. Machnicka, M. A.; Milanowska, K.; Osman Oglou, O.; Purta, E.; Kurkowska, M.; Olchowik, A.; Januszewski, W.; Kalinowski, S.; Dunin-Horkawicz, S.; Rother, K. M.; Helm, M.; Bujnicki, J. M.; Grosjean, H. (2013) *Nucleic Acids Res.* 41, D262-267.
31. Chan, P. P.; Lowe, T. M. (2009) *Nucleic Acids Res.* 37, D93-97.

32. Clark, W. C.; Evans, M. E.; Dominissini, D.; Zheng, G.; Pan, T. (2016) *RNA.* 22, 1771-1784.
33. Hauenschild, R.; Tserovski, L.; Schmid, K.; Thuring, K.; Winz, M. L.; Sharma, S.; Entian, K. D.; Wacheul, L.; Lafontaine, D. L.; Anderson, J.; Alfonzo, J.; Hildebrandt, A.; Jaschke, A.; Motorin, Y.; Helm, M. (2015) *Nucleic Acids Res.* 43, 9950-9964.
34. Zaringhalam, M.; Papavasiliou, F. N. (2016) *Methods.* 107, 63-72.
35. Behm-Ansmant, I.; Helm, M.; Motorin, Y. (2011) *J Nucleic Acids.* 2011, 408053.
36. Clark, W. C.; Evans, M. E.; Dominissini, D.; Zheng, G.; Pan, T. (2016) *RNA.*
37. Wang, J. W.; Wang, A.; Li, K.; Wang, B.; Jin, S.; Reiser, M.; Lockey, R. F. (2015) *Biotechniques.* 58, 161-170.
38. Ross, R.; Cao, X.; Yu, N.; Limbach, P. A. (2016) *Methods.* 107, 73-78.

For Table of Contents Use Only

Detection of a Subset of Posttranscriptional Transfer RNA Modifications in vivo With a Restriction Fragment Length Polymorphism-Based Method

Thomas F. Wulff1, ‡, Rafael J. Argüello2, ‡, Marc Molina Jordán1, Helena Roura Frigolé1, Glenn Hauquier1, Liudmila Filonava1, Noelia Camacho1, Evelina Gatti2,3, Philippe Pierre2,3, Lluís Ribas de Pouplana1,4, and Adrian G. Torres1,*.*