

The landscape of Arabidopsis tRNA aminoacylation

Luis F. Ceriotti^{1,2}, Jessica M. Warren^{3,4}, M. Virginia Sanchez-Puerta^{1,2}  and Daniel B. Sloan^{5,*} 

¹IBAM, Universidad Nacional de Cuyo, CONICET, Facultad de Ciencias Agrarias, Mendoza, Argentina,

²Facultad de Ciencias Exactas y Naturales, Universidad Nacional de Cuyo, Mendoza, Argentina,

³Biodesign Institute and School of Life Sciences, Arizona State University, Tempe, Arizona, USA,

⁴Howard Hughes Medical Institute, Chevy Chase, Maryland, USA, and

⁵Department of Biology, Colorado State University, Fort Collins, Colorado, USA

Received 12 September 2024; revised 2 November 2024; accepted 6 November 2024.

*For correspondence (e-mail dan.sloan@colostate.edu).

SUMMARY

The function of transfer RNAs (tRNAs) depends on enzymes that cleave primary transcript ends, add a 3' CCA tail, introduce post-transcriptional base modifications, and charge (aminoacylate) mature tRNAs with the correct amino acid. Maintaining an available pool of the resulting aminoacylated tRNAs is essential for protein synthesis. High-throughput sequencing techniques have recently been developed to provide a comprehensive view of aminoacylation state in a tRNA-specific fashion. However, these methods have never been applied to plants. Here, we treated *Arabidopsis thaliana* RNA samples with periodate and then performed tRNA-seq to distinguish between aminoacylated and uncharged tRNAs. This approach successfully captured every tRNA isodecoder family and detected expression of additional tRNA-like transcripts. We found that estimated aminoacylation rates and CCA tail integrity were significantly higher on average for organellar (mitochondrial and plastid) tRNAs than for nuclear/cytosolic tRNAs. Reanalysis of previously published human cell line data showed a similar pattern. Base modifications result in nucleotide misincorporations and truncations during reverse transcription, which we quantified and used to test for relationships with aminoacylation levels. We also determined that the *Arabidopsis* tRNA-like sequences (t-elements) that are cleaved from the ends of some mitochondrial messenger RNAs have post-transcriptionally modified bases and CCA-tail addition. However, these t-elements are not aminoacylated, indicating that they are only recognized by a subset of tRNA-interacting enzymes and do not play a role in translation. Overall, this work provides a characterization of the baseline landscape of plant tRNA aminoacylation rates and demonstrates an approach for investigating environmental and genetic perturbations to plant translation machinery.

Keywords: aminoacylation, *Arabidopsis thaliana*, mitochondrial t-elements, post-transcriptional modifications, tRNAs.

INTRODUCTION

Transfer RNAs (tRNAs) play a pivotal role in the central dogma of molecular biology by recognizing codons in messenger RNAs (mRNAs) and delivering the corresponding amino acid during protein synthesis. The expression and maturation of tRNAs depends on interactions with numerous specialized enzymes (Phizicky & Hopper, 2010). Precursor tRNA transcripts are processed by the endonucleases RNase P and RNase Z to remove their 5' leader sequence and 3' extension, respectively (Hartmann et al., 2009). The mature 5' and 3' ends then base pair to form the acceptor stem with a 3' overhanging nucleotide known as the discriminator base, which is followed by a 3' trinucleotide CCA sequence. This CCA tail can be genomically encoded but is more often incorporated

post-transcriptionally by a CCA-adding tRNA nucleotidyl-transferase (CCAse), which plays an important role in tRNA quality control (Betat & Mörl, 2015). The maturation of tRNAs also involves extensive post-transcriptional base modifications, which can be essential for proper tRNA folding, stability, codon-anticodon recognition, and charging by aminoacyl-tRNA synthetases (aaRSs) (Giegé & Eriani, 2023; Motorin & Helm, 2010; Suzuki, 2021). Aminoacylated tRNAs deliver their amino acid to the ribosome for addition to the growing polypeptide chain during protein synthesis, after which they can cyclically repeat the process through multiple rounds of aminoacylation.

Availability of aminoacylated tRNAs is, thus, essential for sustaining protein synthesis. The levels of aminoacylation in the cellular tRNA pool are also environmentally

responsive and play important signaling roles (Avicilar-Kucukgoze et al., 2016; Dittmar et al., 2005; Raina & Ibba, 2014; Subramaniam et al., 2013). Being able to measure the relative abundance of aminoacylated and uncharged tRNAs is vital for understanding the associated regulatory mechanisms.

In addition to canonical tRNAs, there are many transcripts with tRNA-like sequences that can be recognized by the same suite of enzymatic machinery. For example, the maturation of human mascRNA (MALAT1-associated small cytoplasmic RNA) depends on a tRNA-like structure that recruits RNase P and RNase Z for cleavage of a long non-coding RNA (lncRNA) precursor transcript. However, mascRNA is sufficiently divergent in the structure of its D stem and anticodon region that it is not aminoacylated by aaRSs and does not interfere with translation (Skeparnias et al., 2024). Another example is the tRNA-like structure involved in the replication, translation, and genome encapsidation of the brome mosaic virus in plants. In this case, however, recognition and aminoacylation by a host aaRS is necessary for the full functionality of the viral element (Bonilla et al., 2021).

Plant mitochondrial genomes contain another class of tRNA-like sequences (t-elements) that are homologous to tRNA genes and encode transcripts predicted to retain typical tRNA-like secondary structures (Forner et al., 2007; Hanic-Joyce et al., 1990). Because t-elements often occur at the 5' or 3' ends of plant mitochondrial genes, recognition and cleavage of these structures by RNase P and RNase Z can define the ends of mature ribosomal RNA (rRNA) and mRNA transcripts (Forner et al., 2007; Warren, Salinas-Giegé, Triant, et al., 2021). Initial characterization of t-elements also indicated that they could act as CCAse substrates *in vitro*, but they were not detectable *in vivo*, suggesting they were unstable and rapidly removed (Hanic-Joyce et al., 1990). However, subsequent investigation found that t-element transcripts are at least transiently present *in vivo* and can carry post-transcriptionally added 3' nucleotides that are not genomically encoded (Forner et al., 2007). Although this combination of findings suggests the action of a CCAse, a typical CCA motif was rarely observed in the analysis of *in vivo* transcripts. Indeed, in the most extensively studied example (the t-element that is transcribed at the 5' end of *cox1*), sequenced transcripts had various strings of Cs and/or As but none of them had a standard CCA tail (Forner et al., 2007). These findings raise multiple alternative hypotheses: (i) the observed post-transcriptional addition of 3' nucleotides to t-elements *in vivo* is not mediated by a conventional CCAse, (ii) CCAse activity is less specific when not acting on canonical tRNA substrates, or (iii) t-elements with true CCA tails are aminoacylated by aaRSs, inhibiting their detection with RNA sequencing because ligation is blocked by the resulting amino acid at the 3' end. A further uncertainty about

t-element metabolism is whether they are recognized by the enzymes responsible for post-transcriptional base modifications and/or aminoacylation of tRNAs. Therefore, a more comprehensive understanding of t-element processing and function necessitates a detailed investigation of their relationships with the full suite of tRNA-interacting enzymes.

Recent advances have been made in sequencing-based methods that can quantify expression and determine aminoacylation state for tRNAs in a gene-specific fashion (Behrens et al., 2021; Davidsen & Sullivan, 2024; Evans et al., 2017; Pavlova et al., 2020; Watkins et al., 2022). These approaches incorporate a classic method to distinguish between aminoacylated and uncharged tRNAs that has been termed the "Whitfield reaction" (Davidsen & Sullivan, 2024). This method involves treatment with sodium periodate to elicit ring opening and destabilization of the 3' ribose followed by removal of the 3' nucleotide by incubation under basic conditions. Because of the presence of an amino acid on the 3' nucleotide of a tRNA protects the transcript from this reaction, there is a resulting length difference between aminoacylated tRNAs (full-length) and uncharged tRNAs (lacking their 3' nucleotide).

In principle, the length difference generated by periodate treatment can be assessed on a global level with tRNA-seq. However, the extensive post-transcriptional base modifications in tRNAs (Cappannini et al., 2024) create substantial technical challenges because they can block reverse transcriptase (RT) progression during generation of complementary DNA (cDNA) (Padhiar et al., 2024; Wilusz, 2015). Some tRNA-seq methods such as YAMAT-seq specifically capture full-length tRNAs by targeting transcripts with a 3' NCCA overhang for double-stranded adapter ligation (Shigematsu et al., 2017). However, this approach produces a highly biased pool of sequences because it excludes tRNAs with base modifications that block RTs. The limitations of sequencing full-length tRNAs can be partially overcome through treatment with the dealkylating enzyme AlkB to remove some inhibitory base modification or use of specialized RTs with high processivity (Cozen et al., 2015; Scheepbouwer et al., 2023; Warren, Salinas-Giegé, Hummel, et al., 2021; Zheng et al., 2015). However, these approaches have their own drawbacks because of the extensive RNA degradation introduced by AlkB buffer conditions and increased potential for RT artifacts. An alternative or complementary approach is to circularize cDNA or ligate a second adapter to its 3' end after cDNA synthesis so that tRNA-derived sequences are captured regardless of whether they are full-length or truncated (Behrens et al., 2021; Pinkard et al., 2020; Scheepbouwer et al., 2023; Watkins et al., 2022).

In addition to the types of base modifications that inhibit RT processivity, tRNAs contain many modifications

that lead to nucleotide misincorporations or indels by RTs during cDNA synthesis. The resulting sequence variants provide the opportunity to infer the location and identity of modified tRNA bases (Arrivé et al., 2023; Behrens et al., 2021; Clark et al., 2016; Enroth et al., 2019; Gołębiewska et al., 2024; Potapov et al., 2018; Vandivier et al., 2019; Warren, Salinas-Giegé, Hummel, et al., 2021), and these modification patterns have been compared with aminoacylation levels to infer their role in tRNA charging (Hernandez-Alias et al., 2023).

Plants are fascinating models for studying tRNA metabolism because they maintain three parallel translation systems (nuclear/cytosolic, mitochondrial, and plastid). However, despite the advances in using tRNA-seq to measure aminoacylation levels, none of these approaches have yet been applied to plants. Here, we use periodate treatment in combination with a specialized tRNA-seq method (Watkins et al., 2022) to distinguish between aminoacylated and uncharged tRNAs and tRNA-like transcripts in the model angiosperm *Arabidopsis thaliana*. We also complement this analysis with previously published YAMAT-seq data (Warren, Salinas-Giegé, Hummel, et al., 2021) to investigate the enigmatic plant mitochondrial t-elements.

RESULTS

MSR-seq captures all *Arabidopsis* tRNA isodecoder families

We used total RNA samples from *A. thaliana* Col-0 rosettes (4 weeks old, three biological replicates) to generate and sequence multiplex small RNA-seq (MSR-seq) libraries (Watkins et al., 2022) (Table S1). On average, 90.64% of the processed sequences mapped to a reference database consisting of annotated tRNAs in *A. thaliana* (Cognat et al., 2022), with 70.36, 19.62, and 0.66% of the processed sequences mapping to nuclear, plastid, and mitochondrial tRNA genes, respectively. All 220 unique tRNA gene sequences had mapped reads in one of the libraries, and all but two had mapped reads in all libraries (mapped read counts available via https://github.com/dbsloan/Arabidopsis_aminoacylation). Some of the reference sequences are similar enough that reads can produce equal mapping scores for two or more reference genes. When these ambiguously mapping reads are excluded, nine of the 220 reference genes lacked any coverage (eight tRNA-TyrGTA genes and one tRNA-SerAGA gene), but all isodecoder families (i.e., sets of tRNAs that share the same anticodon) found in *Arabidopsis* were still represented. Abundance estimates were highly correlated across biological replicates and different library treatment methods (Figure S1). Raw read abundances should be interpreted with caution because the tRNA sequencing process can result in heavily biased representation of

specific genes, likely related to secondary structure, post-transcriptional base modifications, and other challenges associated with tRNA-seq (Ma et al., 2021; Warren, Salinas-Giegé, Hummel, et al., 2021). In this dataset, tRNA-Trp is the most abundant isoacceptor family (i.e., a set of tRNAs charged with the same amino acid), accounting for an average of 19.6% of all nuclear tRNA reads. At the other extreme, tRNA-Cys (0.3%) and tRNA-His (0.8%) isoacceptors were very rare, accounting for less than 1% of all nuclear tRNA reads. Given the rarity of Trp in the proteome, it is unlikely that tRNA-Trp isoacceptors are as abundant as suggested by read counts or that the large differences among isoacceptors faithfully reflect biological abundances. Accordingly, relative comparisons across samples or treatment types may be the most informative interpretation of MSR-seq (and other tRNA-seq) datasets.

CCA tail integrity and inferred aminoacylation levels vary among tRNAs and among genomic compartments

For MSR-seq library construction, we split the RNA samples from each biological replicate (see above) into three subsamples. One subsample was treated with periodate to distinguish between aminoacylated and uncharged tRNAs (see “Introduction” section). The other two subsamples were used as a negative control (no periodate) or positive control (deacylation prior to periodate treatment) for the removal of the 3' terminal nucleotide. Results from these three different treatments followed the expected pattern (Figure 1), as the overall percentage of mapped reads with an intact CCA tail in the test samples was intermediate (mean of 55.1%) between the negative control (75.7%) and the positive control (15.1%). Here and throughout, these “CCA percentages” were calculated after excluding reads lacking more than just a single 3' nucleotide (see below for analysis of other read types that were missing additional 3' sequence). The fact that the 3' terminal nucleotide was not removed from 100% of tRNA reads in positive control samples indicates that the periodate treatment was not fully efficient in eliminating exposed 3' nucleotides and/or that the preceding deacylation step was not fully efficient in removing the amino acids to expose these ends. Likewise, the fact that tRNA reads from negative control libraries did not exhibit 100% retention of an intact CCA tail indicates that a substantial percentage of transcripts in the *in vivo* tRNA pool lack the 3' terminal nucleotide and/or that tRNAs were subject to some degradation during the RNA extraction and library construction process. These alternative interpretations are investigated below in the next two sections.

In the periodate-treated test sample, inferred aminoacylation levels differed substantially among tRNA genes. The percentage of reads with intact CCA tails ranged over two-fold across isoacceptor families (35.97–77.75%;

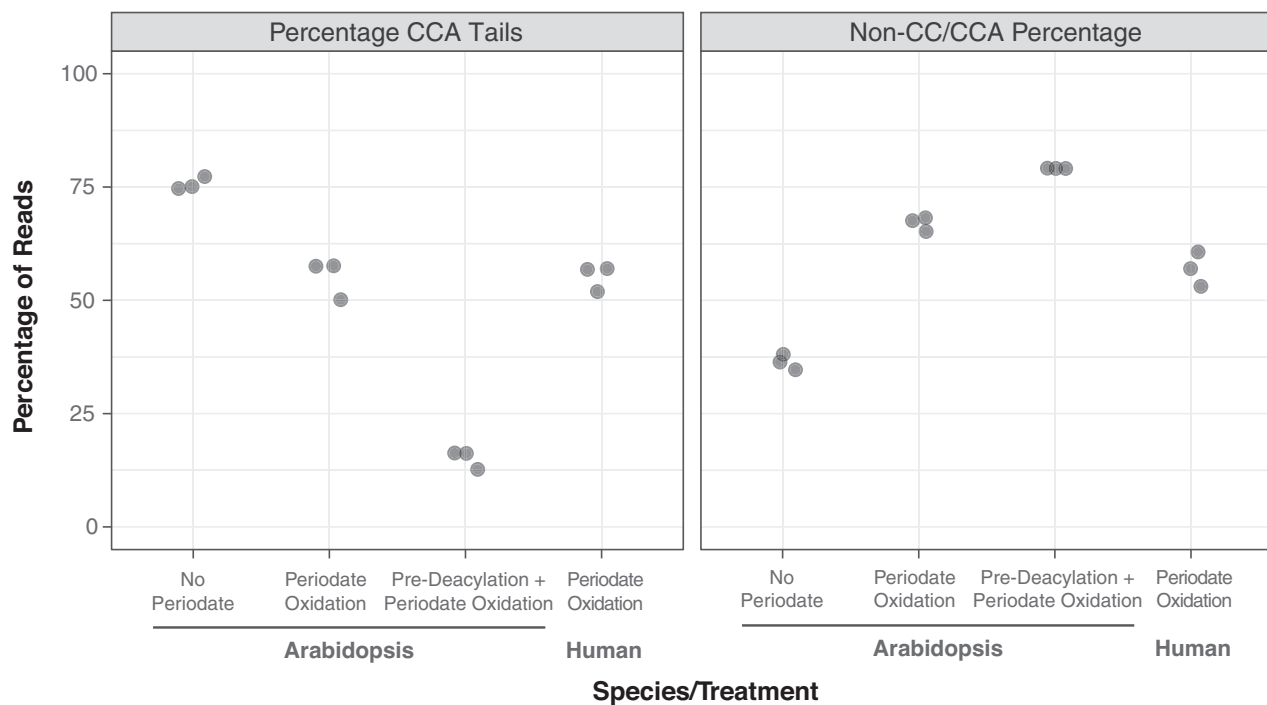


Figure 1. MSR-seq data from Arabidopsis (this study) and human cell line tRNAs (Watkins et al., 2022). Periodate treatment prior to MSR-seq library construction was used to estimate aminoacylation levels based on protection of the terminal nucleotide of the CCA by the amino acid. Reported values for CCA tails (left) represent the percentages of reads with intact CCA tails after excluding reads that lacked more than just a single 3' nucleotide. In contrast, non-CC/CCA percentages (right) represent the fraction of *all* tRNA reads that lack two or more nucleotides at their 3' ends, so values between the two panels do not sum to 100%. Such reads represent a large percentage of the dataset, especially in periodate-treated libraries. Three biological replicates were sequenced per treatment, with highly significant differences for both CCA tail ($P = 6.3e-7$, one-way ANOVA) and non-CC/CCA percentages ($P = 5.1e-8$, one-way ANOVA). The (periodate-treated) human samples are very similar to their Arabidopsis counterparts in terms of both CCA percentage and non-CC/CCA percentage.

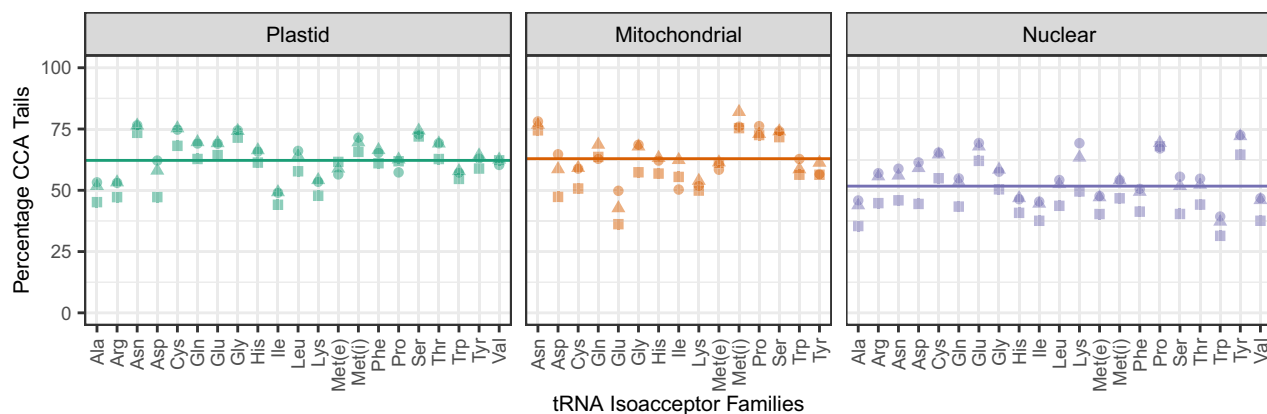


Figure 2. Variation among Arabidopsis tRNA isoacceptor families and genomic compartments in CCA tail integrity in response to periodate treatment. Reported values represent the percentages of reads with intact CCA tails after excluding reads that lacked more than just a single 3' nucleotide. Biological replicates are indicated by different shapes. The average frequency of intact CCA tails differs significantly across genomic compartments ($P = 0.0005$; one-way ANOVA), with a lower rate for nuclear-encoded tRNAs than organellar tRNAs, as indicated by horizontal lines (means) in each panel. Met(e) and Met(i) refer to elongator and initiator tRNA-Met genes, respectively. We also found a similar difference between organellar and nuclear reads when retention of CCA tails was expressed as a percentage of total reads and not relative to just reads ending in CCA or CC ($P = 1.5e-7$; one-way ANOVA). A similar analysis subdivided into isodecoder families is presented in Figure S2.

Figure 2). Interestingly, the CCA percentage was significantly higher for tRNAs encoded in organellar genomes than in nuclear genomes when calculated as averages

across isoacceptor families (Figure 2) or isodecoder families (Figure S2). Based on this observation, we reanalyzed a previously published MSR-seq dataset from human cell

lines (Watkins et al., 2022) and found a similar pattern, with significantly higher percentages of intact CCA tails in mitochondrial tRNAs than in nuclear-encoded tRNAs (Figure S3).

The premise of using periodate treatment to infer aminoacylation levels is that uncharged tRNAs are susceptible to removal of their terminal 3' nucleotide. However, we found that tRNAs in negative control (no periodate) libraries had frequencies of intact CCA tails that were strongly correlated with the periodate-treated test samples ($r = 0.77$; $P = 8.9\text{e-}40$), albeit with the control libraries at higher levels (Figure 3a). Accordingly, organellar tRNAs had a higher percentage of intact CCA tails than nuclear-encoded tRNAs even without periodate treatment (Figure 3a) ($P = 6.9\text{e-}9$; t -test). This observation implies that nuclear-encoded tRNAs have lower aminoacylation levels, in part because a smaller fraction of them within the cellular pool have an intact CCA tail. It is also possible that these tRNAs are more vulnerable to degradation during RNA extraction and library construction for some reason (including that they might be less likely to be protected by an amino acid at their 3' end). Therefore, to further investigate variation in aminoacylation levels, we repeated the comparison of CCA tail percentage between nuclear and organellar tRNAs, after dividing the value for each gene under periodate treatment by the corresponding value from no-periodate controls. This analysis also found higher values for CCA tail frequencies for organellar tRNAs than for nuclear tRNAs ($P = 1.3\text{e-}5$; t -test), meaning that organellar tRNAs are more likely to keep their CCA tail even when adjusting for their higher starting percentage of CCA tails.

In contrast to the strong correlation in the frequency of intact CCA tails between test samples and negative controls, there was little relationship between test samples and positive controls (pre-deacylated with high pH treatment prior to periodate treatment; Figure 3a; $r = 0.04$; $P = 0.55$). In principle, pre-deacylation should make tRNAs universally susceptible to the periodate. Many tRNA genes showed the expected pattern, with intact CCA tail percentages dropping to ~0% in the pre-deacylated libraries. However, many other tRNAs exhibited little or no difference when a pre-deacylation treatment was applied, suggesting that effectiveness of deacylation may vary across tRNAs and that longer incubations may be necessary for complete deacylation (Davidsen & Sullivan, 2024). Certain isoacceptor families such as Ile, Thr, and Val showed especially high retention of intact CCA tails despite pre-deacylation treatment (Figure 3b). Overall, the retention of CCA tails following pre-deacylation and periodate treatment exhibited a striking correlation (Figure 3c) with previous estimates of aminoacyl-tRNA half-life for *Escherichia coli* tRNAs under similar deacylation conditions (Hentzen et al., 1972).

RNA extraction method has minimal effect on CCA tail integrity, but periodate treatment causes extensive RNA degradation

Many Arabidopsis tRNA reads were missing more than just the terminal 3' nucleotide. The frequency of such reads was higher in periodate-treated samples than in negative controls (no periodate) and increased further when we performed a deacylation step prior to treating with periodate (Figure 1). We also reanalyzed previously published MSR-seq data (Watkins et al., 2022) from human cell lines and found similar frequencies (Figure 1). The presence of shortened reads with incomplete CCA tails even in control libraries could reflect the true presence of these molecules *in vivo*, as CCA tail integrity can be responsive to environmental and subcellular conditions (Czech, 2020; Lucas et al., 2023; Scheepbouwer et al., 2024). However, artifactual RNA damage may also contribute to this pool of shortened transcripts. Therefore, we performed a follow-up experiment to investigate the possible role of damage during RNA extraction and specifically the low pH in the acid-phenol extraction method that is typically used to preserve tRNA aminoacylation state (Zaborske et al., 2009). In this follow-up experiment, we generated libraries for Arabidopsis RNA samples that were extracted in parallel with either an acid-phenol protocol or a conventional Trizol protocol. We found that the results from acid-phenol and Trizol RNA samples were highly correlated (Figure S4) with significantly higher percentages of reads with intact CCA tails in the Trizol samples than in acid-phenol samples ($P = 3.5\text{e-}32$ and $P = 2.3\text{e-}34$ in control and periodate treatments, respectively). However, the difference between the extraction methods was very small (mean difference of 2.3% and 3.1% in control and periodate treatments, respectively). The two extraction methods did not differ in the percentage of shortened reads that lacked more than just the terminal 3' nucleotide ($P = 0.06$ and $P = 0.72$ in control and periodate treatments, respectively). Therefore, the acid-phenol extraction approach does not appear to have caused a substantial increase in damage over conventional Trizol methods. However, somewhat surprisingly, the similar readouts from the different extraction methods and the fact that Trizol produced *higher* CCA tail percentages also suggest that our acid-phenol extractions may not be better than Trizol protocols at preserving the tRNA aminoacylation state.

To further investigate potential artifactual sources of RNA damage, we compared the integrity of total Arabidopsis RNA in a sample treated with a series of periodate, ribose, sodium tetraborate, and T4 polynucleotide kinase incubations according to the MSR-seq protocol (see "Methods" section) to an untreated control sample. This comparison revealed extensive degradation in the treated sample (Figure S5), confirming that effects are not limited

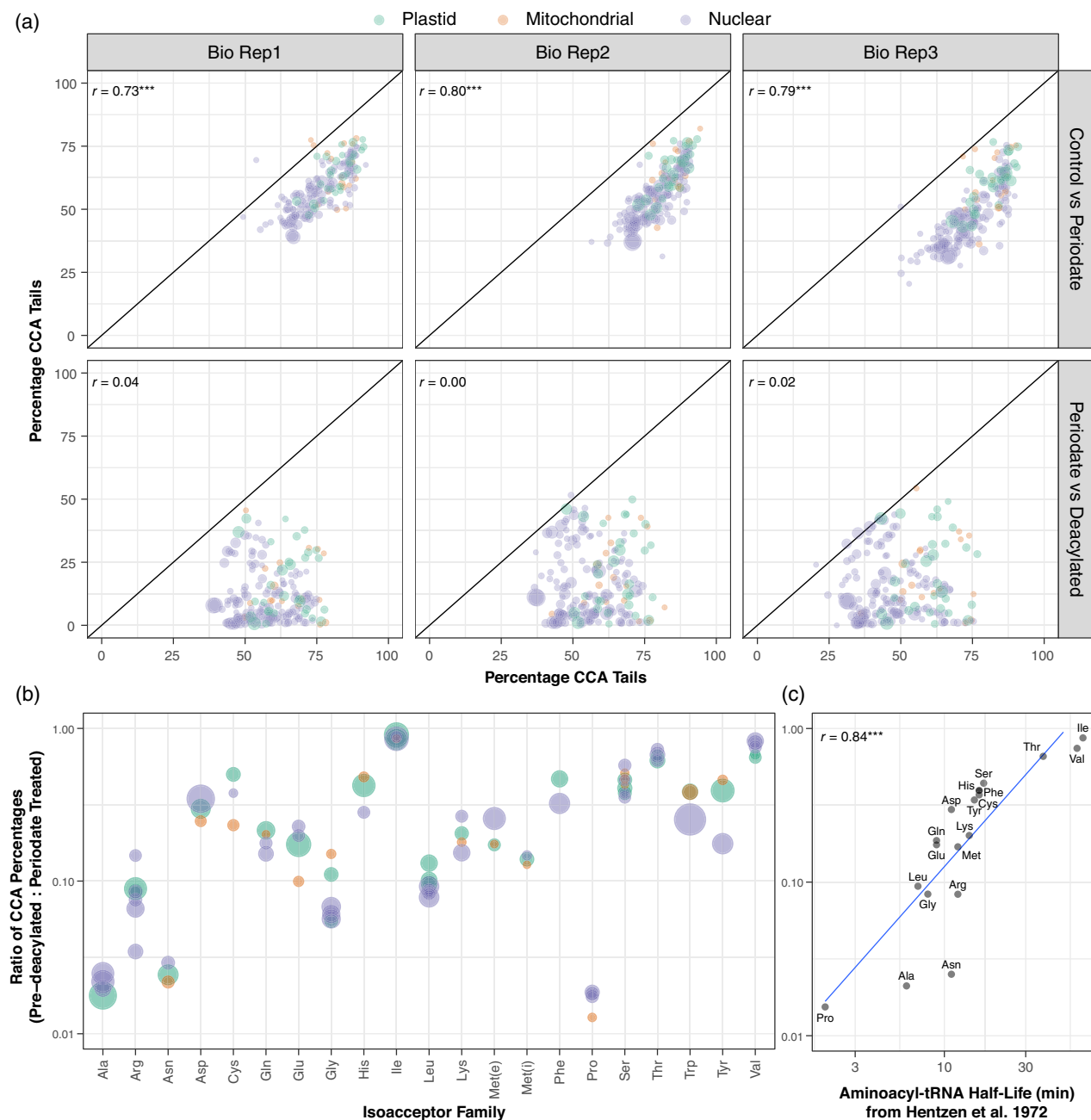


Figure 3. Comparison of CCA tail retention across treatments.

(a) Correlation in percentage of Arabidopsis reads with intact CCA tails are shown for each biological replicate. Reported values represent the percentages of reads with intact CCA tails after excluding reads that lacked more than just a single 3' nucleotide. The top row shows the comparison between the negative control libraries (no periodate) on the x-axis and the periodate-treated test samples on the y-axis. The bottom row shows the comparison between the periodate-treated test samples on the x-axis and the positive control samples (pre-deacylated prior to periodate treatment) on the y-axis. Each point represents an individual tRNA gene (minimum of 100 reads per gene), with color indicating genome of origin and size proportional to the average number of reads mapping to that gene. A one-to-one line is plotted in each panel.

(b) Values on y-axis represent the ratio of CCA percentages for the pre-deacylated treatment versus the standard periodate treatment by isoeceptor family. A value of 1 indicates that pre-deacylation did not result in any reduction of CCA tail percentage, whereas a value of 0 indicates that pre-deacylation completely eliminated CCA tails. Each point represents an isodecoder family averaged across genes with a minimum of 100 reads in that family and three biological replicates. Point size reflects average read abundance for that isodecoder family. Met(e) and Met(i) refer to elongator and initiator tRNA-Met genes, respectively.

(c) Comparison of CCA retention rates [unweighted average of values in panel (b)] with aminoacyl-tRNA estimated for *Escherichia coli* tRNAs at 37°C, pH 8.6 (Hentzen et al., 1972). These half-life estimates are strongly correlated with the observed retention of CCA tails in Arabidopsis after a 30-min pre-deacylation at 37°C, pH 9.0 followed by periodate treatment ($P = 7.9 \times 10^{-6}$, log-transformed). Reported r values in panels (a) and (c) are Pearson correlation coefficients. ***Statistical significance at a $P < 0.001$ threshold.

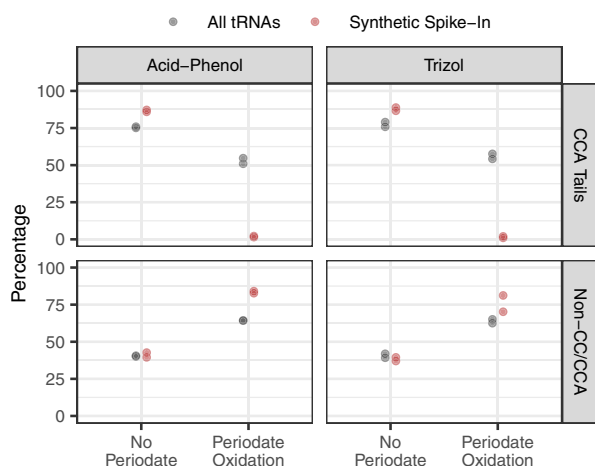


Figure 4. Inclusion of a synthetic tRNA spike-in control in MSR-seq libraries.

Reported values for CCA tails (top row) represent the percentages of reads with intact CCA tails after excluding reads that lacked more than just a single 3' nucleotide. In contrast, non-CC/CCA percentages (bottom row) represent the fraction of *all* tRNA reads that lack two or more nucleotides at their 3' ends, so values between the panels do not sum to 100%. Values are shown for all *Arabidopsis* tRNAs (gray) and the synthetic spike-in tRNAs (red). Two biological replicates were performed for each of two RNA extraction methods: acid-phenol (left panels) and Trizol (right panels). For each biological replicate, control and periodate-treated libraries were generated.

to removal of just a single 3' nucleotide. Recent optimization efforts have shown that the desired outcome of periodate treatment can be obtained with much shorter periodate incubation times at lower temperatures and replacement of sodium tetraborate incubation with a less basic alternative to limit these additional forms of damage (Davidsen & Sullivan, 2024).

Periodate treatment results in near-complete removal of 3' terminal nucleotide from synthetic spike-in control tRNAs

Given that deacylating RNA samples prior to periodate treatment did not lead to complete elimination of 3' nucleotides (Figures 1 and 3), a more robust positive control for periodate-mediated removal would be desirable. Therefore, in the follow-up experiment that compared acid-phenol and Trizol extraction methods, we also included a synthetic "spike-in" tRNA that was added to each *Arabidopsis* RNA sample prior to library construction. Because these synthetic tRNAs are not aminoacylated, periodate treatment would be expected to fully remove terminal 3' nucleotides. Indeed, reads derived from the spike-in exhibited near-complete elimination of the 3' nucleotide, with an intact CCA tail percentage of only 1.7% in periodate-treated samples (Figure 4). This finding implies that the failure to completely remove 3' nucleotides from some *Arabidopsis* tRNAs (Figures 1 and 3) was likely the result of incomplete deacylation rather than low efficiency of the periodate reaction itself.

The addition of a synthetic spike-in tRNA also provides insights into the origins of shortened reads that lack two or more nucleotides from their 3' ends. In principle, the synthetic tRNAs should all retain intact CCA tails in control (no periodate) libraries, and they should only have a single 3' nucleotide removed in periodate-treated libraries. We found that the CCA tail percentage in control libraries was higher for the synthetic tRNAs than for *Arabidopsis* tRNAs, but it still fell short of 100% (Figure 4). Furthermore, the majority of reads derived from the synthetic tRNAs in periodate-treated libraries lacked two or more 3' nucleotides (Figure 4). This observation supports the conclusion that periodate treatment and downstream reactions can result in more extensive damage than the expected removal of a single 3' nucleotide (Davidsen & Sullivan, 2024). In addition, even in control libraries, there was a substantial frequency of these shortened reads derived from synthetic tRNAs (Figure 4). Thus, it appears that RNA handling and the process of library construction itself (e.g., overnight adapter ligation reaction) may result in some 3' end degradation.

Truncated tRNA sequences reveal "hard-stop" base modification and tRNA fragmentation points

Some tRNA post-transcriptional base modifications inhibit RT progression ("hard stops") and result in truncation of cDNA products at internal positions within the tRNA template (Padhiar et al., 2024; Wilusz, 2015). In addition, there is an abundance of tRNA-derived fragments (tRFs) within cells (Keam & Hutvagner, 2015; Lalande et al., 2020). Mapping the positions where sequencing reads start relative to the 5' end of mature tRNAs can be used to infer positions of hard-stop base modifications and/or tRF breakpoints (Figure 5; Figure S6). We found that some *Arabidopsis* isodecoder families produced almost exclusively full-length tRNA sequences (e.g., nuclear tRNA-AspGTC and tRNA-GluTTC), whereas a few were almost always internally truncated (e.g., nuclear tRNA-LeuCAG and tRNA-LysTTT). Many others exhibited a mix of full-length and internally truncated reads (e.g., nuclear tRNA-AlaAGC and tRNA-SerGCA). Internal truncations were concentrated at the positions immediately 3' of known base modifications (Figure 6), including m¹A, m¹G, m²₂G, and the diversity of modifications at the anticodon wobble position (base 34) and immediately after the anticodon (base 37) (Cozen et al., 2015; Gołębiewska et al., 2024; Masuda & Hou, 2024; Warren, Salinas-Giegé, Hummel, et al., 2021; Zheng et al., 2015). Because size selection in our library construction restricted the analysis to reads of at least ~35 nt in length, truncations near the 3' end of tRNAs were not investigated.

The distribution of 5' mapping positions for a given isodecoder family was generally similar between periodate-treated and control libraries. However, nuclear

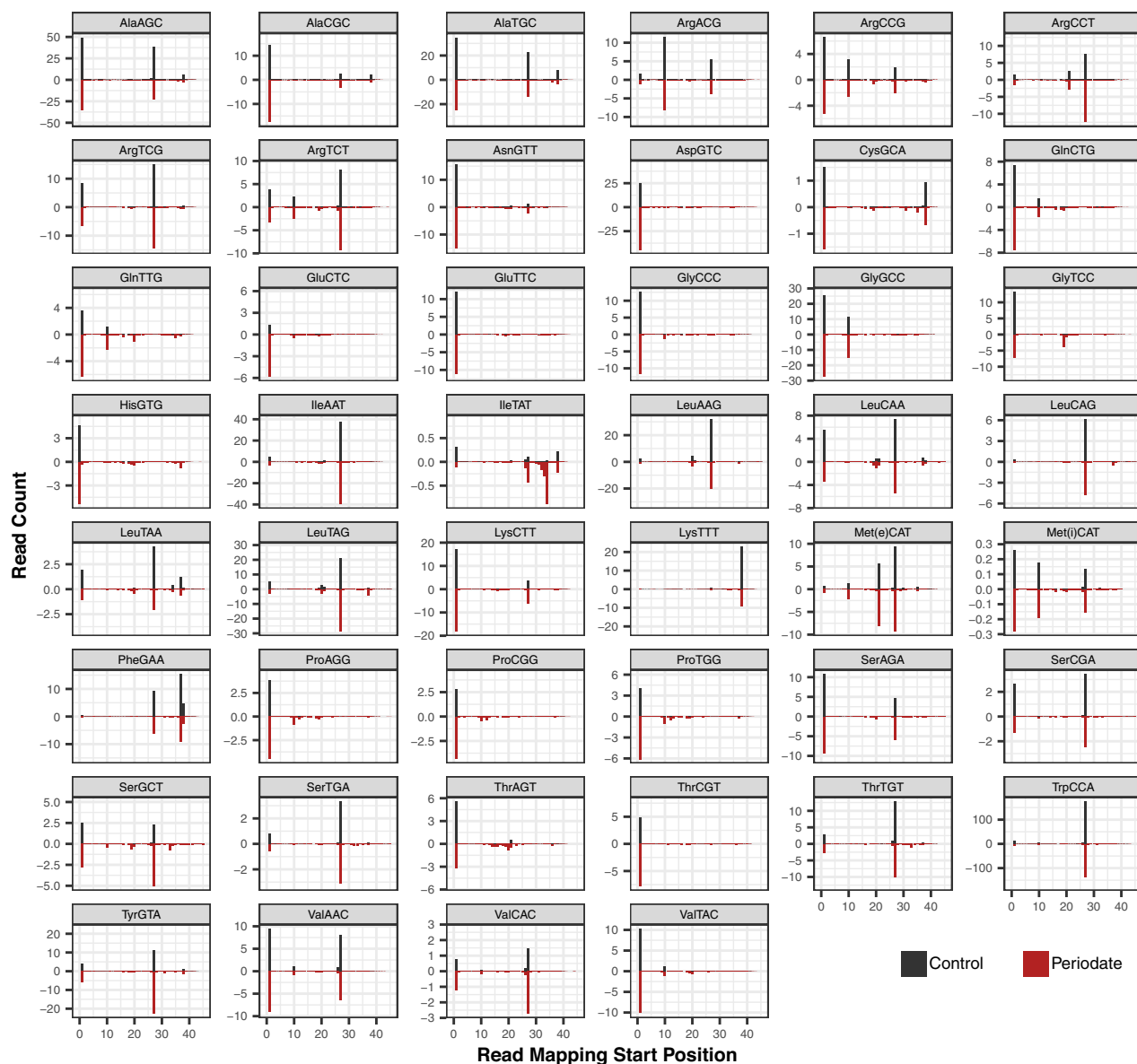


Figure 5. 5' Mapping position of MSR-seq reads for all Arabidopsis tRNA isodecoder families found in the nuclear genome. Counts are represented per thousand reads that mapped to the nuclear tRNA gene set (averaged across the three biological replicates). Control libraries (black bars) are shown as positive values above the x-axis, while periodate-treated libraries (red bars) are shown as negative values below the x-axis. Mapping positions are standardized based on the Sprinzl coordinate system (Sprinzl et al., 1998).

tRNA-IleTAT (Figure 5) and plastid tRNA-LysTTT (Figure S6) exhibited intriguing exceptions. For these two isodecoder families, there was a major truncation point indicated by reads starting at nt position 34 (the wobble position of the anticodon) in periodate-treated samples, which was at very low frequency in no-periodate controls. We also found a large increase in nucleotide misincorporations or deletions at the neighboring position 33 in the subset of reads that are not truncated for these two isodecoder families in periodate-treated libraries (Figure S7). Similar observations in other species have been attributed

to reactivity of periodate with 2-thio modifications at the wobble position (Davidsen & Sullivan, 2024; Katanski et al., 2022). However, given that 2-thio modifications are present in many tRNAs, it is not clear why this dramatic effect of periodate treatment was only observed in nuclear tRNA-IleTAT and plastid tRNA-LysTTT. More generally, many isodecoder families showed increases in the frequency of partial reads distributed across numerous low-frequency internal truncation points in periodate-treated libraries (Figure 5; Figure S6), suggesting more widespread effects of periodate treatment.

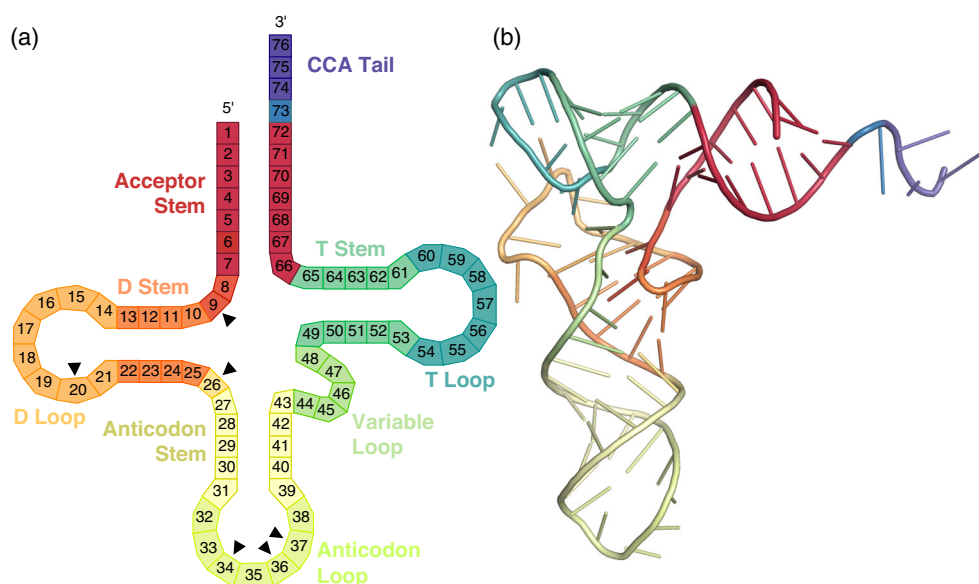


Figure 6. Canonical tRNA structures.

(a) Secondary structure with bases numbered according to the Sprinzl coordinate system (Sprinzl et al., 1998). Black triangles indicate positions that were commonly associated with 5' truncations. Note that because of size selection, short fragments resulting from truncations towards the end of molecule would not have been detectable.

(b) Representative tRNA 3D structure (*Thermus thermophilus* tRNA-Glu; PDB 1G59), using the same color scheme as in panel (a).

Mapping 3' read end positions (Figure 7; Figure S8) can identify the breakpoints for 5' tRFs (note that, unlike truncations at the 5' end, shortening on the 3' end of a read cannot be caused by hard-stop modifications and RT inhibition). Although most isodecoder families lacked major internal 3' truncations points, there were some notable exceptions, such as nuclear tRNA-ValAAC (Figure 7) and plastid tRNA-IleCAT (Figure S8). When mapped at the level of individual genes, 5' tRFs were most associated with breakpoints in the anticodon loops of nuclear tRNAs (Figure S9). However, it should be noted size selection during library construction limited detection of tRFs <~35 nt in length. Comparative mapping of end positions between periodate-treated and control libraries also illustrates the effect of periodate on 3' nucleotide removal, as well as the sizeable population of reads that lack more than a single 3' nucleotide (Figure 7; Figure S8).

Post-transcriptional base modifications inferred from misincorporation patterns

Mapping the positions of nucleotide substitutions and indels onto reference sequences can identify bases that are post-transcriptionally modified (Arrivé et al., 2023; Behrens et al., 2021; Clark et al., 2016; Enroth et al., 2019; Gołębiewska et al., 2024; Potapov et al., 2018; Vandivier et al., 2019; Warren, Salinas-Giegé, Hummel, et al., 2021). Using data from control (no periodate) libraries, we generated a global summary of position-specific misincorporation frequencies across all *Arabidopsis* tRNAs (Figure 8), revealing striking

patterns such as the ubiquitous signature of the m¹A58 modification in nuclear tRNAs (but not organellar tRNAs) and widespread misincorporations corresponding to the m²G26 modification (Motorin & Helm, 2010). Many other positions known to harbor modifications show high rates of misincorporations in multiple tRNAs (Cappannini et al., 2024; Suzuki, 2021). It should be noted, however, that some modifications can be "silent" with regard to generation of RT misincorporations. Therefore, the absence of a misincorporation signature should not be taken as evidence that a base is completely free of modifications.

Inferring relationships between base modification and aminoacylation rates

Base modifications can alter tRNA stability and molecular interactions. Therefore, they have the potential to affect aminoacylation rates (Hernandez-Alias et al., 2023). To investigate this possibility, we first tested for a difference in misincorporation rates in periodate-treated libraries between reads with intact CCA tails (aminoacylated tRNAs) and reads with CC tails (uncharged tRNAs). This analysis showed that misincorporation rates were highly correlated between these two read classes with no significant difference in rates between CCA and CC reads ($P = 0.10$; Figure S10). Therefore, misincorporation data did not provide evidence for an effect of single base modifications on aminoacylation rates.

We next considered 5' truncations as another proxy for base modifications. In this case, we did find that

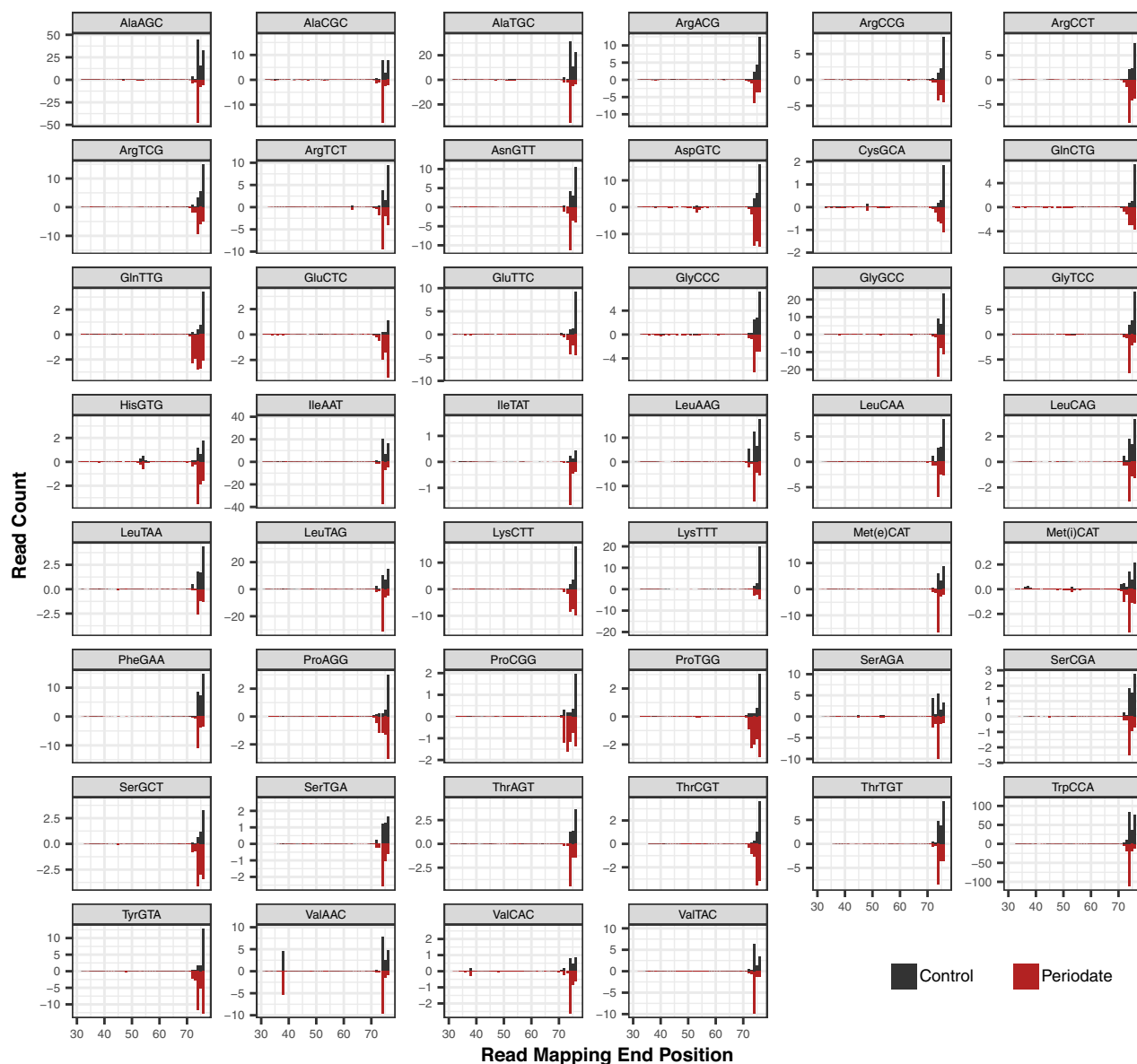


Figure 7. 3' Mapping position of MSR-seq reads for all Arabidopsis tRNA isodecoder families found in the nuclear genome. Counts are represented per thousand reads that mapped to the nuclear tRNA gene set (averaged across the three biological replicates). Control libraries (black bars) are shown as positive values above the x-axis, while periodate-treated libraries (red bars) are shown as negative values below the x-axis. Mapping positions are standardized based on the Sprinzl coordinate system (Sprinzl et al., 1998).

reads with 5' truncations had higher percentages of intact CCA tails than full-length reads in periodate-treated libraries (Figure S11). This pattern was observed for plastid ($P = 2.5e-11$), mitochondrial ($P = 1.1e-5$), and nuclear ($P = 2.4e-6$) tRNAs, but the average difference in CCA percentages was approximately twice as large for organellar tRNAs than for nuclear tRNAs (Figure S11). The control (no periodate) libraries did not show this same pattern ($P > 0.05$ for each of the three genomes), which is consistent with the hypothesis that some of the

base modifications that cause 5' truncations during cDNA synthesis can facilitate aminoacylation. However, the magnitude of these effects is small (Figures S11–S13), and they run counter to the results from misincorporation data, which did not show an effect. Therefore, alternative interpretations should also be considered. For example, it is possible that the secondary structure of purified tRNAs is correlated with aminoacylation state and affects RT progression. Some of the reads with truncations at their 5' ends might also result from 3'

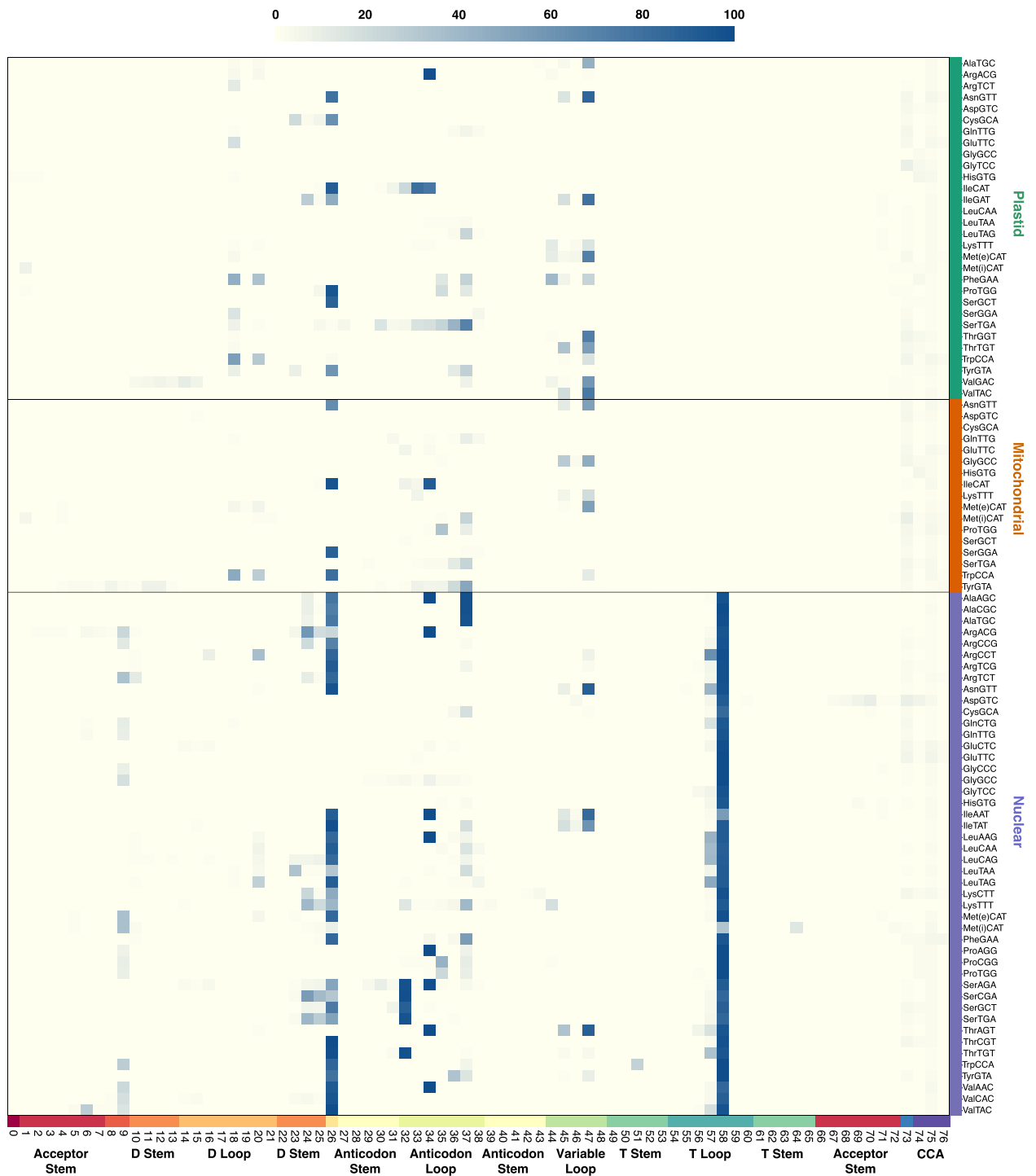


Figure 8. Heatmap representation of RT misincorporations across different tRNA isodecoder families and nucleotide positions in MSR-seq control (no periodate) libraries.

The color scale reflects the average percentage of nucleotide substitutions and deletions for three biological replicates. Data are only reported for replicates with a read count >50. Mapping positions are standardized based on the Sprinzl coordinate system (Sprinzl et al., 1998), and corresponding structural positions are shown on the x-axis (see Figure 6 for definition of structural elements).

tRFs (including those that are part of nicked but structurally intact tRNAs; Costa et al., 2023) rather than hard-stop modifications that block RT progression. If tRFs

exhibit different rates of aminoacylation than full-length tRNAs, it could produce detectable differences in this analysis of 5' truncations.

Mitochondrial t-elements are transcribed, cleaved, base-modified, and CCA tailed – but not aminoacylated

The *A. thaliana Col-0* mitochondrial reference genome (GenBank NC_037304.1) has eight annotated t-elements. To detect additional CCA-tailed transcripts that may be expressed and processed by tRNA-interacting enzymes, we reanalyzed a previously published Arabidopsis YAMAT-seq dataset (Warren, Salinas-Giegé, Hummel, et al., 2021). This analysis detected expression and CCA-tailing of a family of tRNA-PheGAA-like sequences. One of these sequences is included in the plantRNA database (Cognat et al., 2022) as tRNA-TyrGTA-171 843 because of a nucleotide substitution converting the GAA anticodon (Phe) to GTA (Tyr). However, this and other tRNA-Phe-like sequences in the Arabidopsis mitochondrial genome have been previously annotated as potential pseudogenes due to their divergence in sequence (Sloan et al., 2018). In total, there are seven unique tRNA-Phe-like sequences in the Arabidopsis mitochondrial genome (corresponding to 11 different loci due to additional identical duplicates). Some of these retain the GAA anticodon, whereas others share the derived GTA anticodon. We excluded one of these unique sequences from subsequent analysis because it represented only a small fragment and was not detected by YAMAT-seq. Analysis of YAMAT-seq data also found expression of a sequence that is 3' of the mitochondrial *orf315*. This 63-nt transcript does not exhibit obvious homology to any tRNA gene, but it has sequence complementarity between its 5' and 3' ends, suggesting a "stem-loop" structure with a single-nucleotide 3' overhang and added CCA tail. The same sequence is also present in the mitochondrial genomes of many other angiosperms.

Having cataloged the set of CCA-tailed mitochondrial transcripts, we then determined whether these were captured in our MSR-seq dataset. Overall, we found MSR-seq reads for five of eight previously annotated t-elements, all six of the analyzed tRNA-Phe-like sequences, and the *orf315*-adjacent sequence. In some cases, read abundance was extremely low (Figure S14), but we detected dozens to hundreds of reads per library for the 5' *cox1* t-element, 3' *ccmC* t-element, tRNA-Phe-like sequences, and the *orf315*-adjacent stem-loop (Dataset S1), enabling further investigation.

MSR-seq control (no periodate) libraries identified a diversity of post-transcriptionally added 3' tails on the *ccmC* and *cox1* t-elements (Table 1) (Forner et al., 2007). The most common addition for the *ccmC* t-element was a canonical CCA tail, but many non-canonical variants were also detected. For the *cox1* t-element, a CCA tail was also abundant but slightly less common than a non-canonical CA tail. For reasons that are not clear, the distribution of post-transcriptionally added tails for the *cox1* t-element was notably different than observed in a previous analysis

Table 1 Various post-transcriptional additions to the 3' end of sequenced mitochondrial t-elements

3' Addition	Raw read counts in control (no periodate) libraries		
	<i>ccmC</i> t-element	<i>cox1</i> t-element (this study)	<i>cox1</i> t-element (Forner et al., 2007)
CCA	86	151	0
CC	35	47	3
C	12	19	4
None	0	12	6
CA	0	180	1
CCAAA	0	1	1
CAAAA	0	0	2
CCAA	4	4	0
CCAC	0	3	0
CAAA	0	1	0
CAA	0	3	0
CACA	0	1	0
AC	0	1	0
A	0	0	2

(Table 1) (Forner et al., 2007). Specifically, past sequencing of 19 cDNA clones did not identify any with a canonical CCA tail in contrast to the abundance of CCA tails in our dataset.

We also identified clear signatures of RT misincorporations in the *ccmC* t-element at the 26G nucleotide, a major target for post-transcriptional base modifications in tRNAs (Figure 9b). The mitochondrial *ccmC* t-element is homologous to plastid tRNA-IleCAT. The misincorporation profile at G26 is very similar between these two transcripts (dominated by deletions and G → C substitutions; Figure 9b). Therefore, it is possible that the *ccmC* t-element has the typical m²G26 modification that is common at this position. However, plastid tRNA-IleCAT also shows substantial hard-stop activity at this site, with a truncation rate of >50% at the 3' nucleotide (position 27), whereas the *ccmC* t-element does not show any evidence of RT inhibition at the homologous position (Figure 9). This contrast raises uncertainty as to whether the *ccmC* t-element and plastid tRNA-IleCAT share the same modification at this position. Although RT misincorporation and truncation patterns for the same base modification can vary depending on sequence context (Behrens et al., 2021; Hauenschild et al., 2015), the *ccmC* t-element and plastid tRNA-IleCAT share identical flanking sequences, so any differential effects on RT would have to be mediated by the extensive sequence divergence in the middle of the transcripts (Figure 9b). Determination of the precise modification at the G26 position in the *ccmC* t-element will require a more direct approach such as liquid chromatography–mass spectrometry. In addition to the misincorporation signature at G26, the two homologs share a high frequency of 3' truncations after the 64U nucleotide (Figure 9). However, the t-element lacks the strong modification signatures that

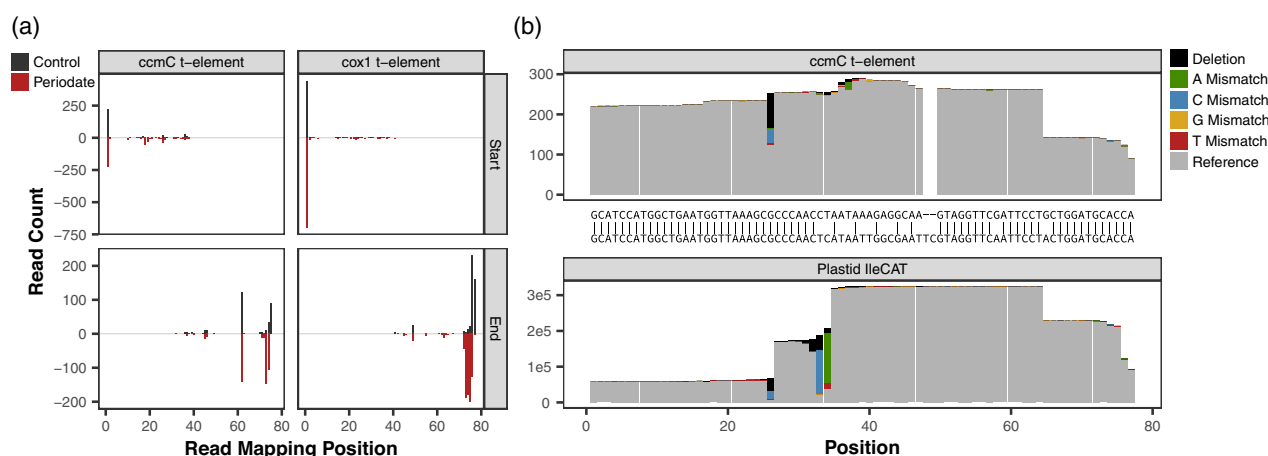


Figure 9. Arabidopsis mitochondrial t-element end-positions and misincorporations. The 5' *cox1* and 3' *ccmC* t-elements were both detected at substantial abundance in MSR-seq dataset.

(a) The 5' end (top) and 3' end (bottom) mapping positions of reads are depicted in the same fashion as in Figures 5 and 7, respectively. Both *ccmC* (left) and *cox1* (right) t-elements are predominantly full-length on the 5' end, whereas *ccmC* shows a high-frequency internal breakpoint among the 3' ends. Periodate treatment leads to almost complete removal of the terminal 3' nucleotide, indicating that the t-elements are not aminoacylated. (b) The mitochondrial *ccmC* t-element (top) is homologous to plastid tRNA-IleCAT (bottom). Nucleotide misincorporation patterns reveal an abundance deleted bases and nucleotide substitutions at G26 in both transcripts, indicating that the t-element likely shares the typical base modification at this position, although it appears to act as a "hard-stop" in tRNA-IleCAT but not in the t-element. We did not detect modified bases in the *cox1* t-element, which is not shown in panel (b).

are present 33U and 34C in plastid tRNA-IleCAT, which is not surprising because of divergence in the anticodon loop that distinguishes the two sequences (Figure 9b).

Overall, these observations provide evidence that mitochondrial t-elements and other tRNA-like sequences are recognized and processed by RNase P, RNase Z, CCase, and enzymes that install post-transcriptional base modifications. However, our MSR-seq data indicate that the captured t-elements, tRNA-Phe-like sequences, and *orf315*-adjacent stem-loop were not aminoacylated because periodate treatment resulted in essentially complete elimination of the terminal 3' nucleotide (Figure 9a; Dataset S1).

DISCUSSION

The effectiveness of MSR-seq in measuring plant tRNA expression and aminoacylation

The history of tRNA sequencing has long been a glass half-full, half-empty affair. On one hand, tRNAs played a pioneering role in the sequencing era, as a yeast tRNA was the first nucleic acid ever sequenced (Holley et al., 1965). On the other hand, they have arguably become the most challenging class of RNAs to accurately and quantitatively sequence. Recent innovations have begun to overcome this challenge, and our findings show that the MSR-seq workflow (Watkins et al., 2022) has promise for a more complete representation of plant tRNA pools. Because it can capture truncated RT products, this sequencing approach does not discriminate against transcripts

carrying hard-stop base modifications. Accordingly, many tRNAs were found almost exclusively as truncated products in our dataset (Figure 5; Figure S6). These include isodecoder families that were essentially undetectable in a previous YAMAT-seq analysis of Arabidopsis (e.g., mitochondrial TyrGTA, SerTGA in all three genomes, and all nuclear Leu isodecoder families) and many others that were only detectable at appreciable levels with YAMAT-seq if samples were pre-treated with dealkylating enzyme AlkB to remove inhibitory base modifications (e.g., nuclear IleAAT, LysTTT, eMetCAT, PheGAA, ThrTGT, and TrpCCA) (Warren, Salinas-Giegé, Hummel, et al., 2021).

Therefore, capturing truncated products resulting from RT inhibition greatly expands accessibility of the tRNA pool. However, there are obvious drawbacks to sequencing these truncated fragments. For example, unsequenced regions cannot be assessed for base modification profiles (Figure 8), and shorter sequences can be difficult to map unambiguously when multiple genes share similarity at their 3' ends. In this study, our library size selection restricted the analysis to sequences of >35 nt in length, so we did not investigate truncation patterns for extremely short fragments. Although treating RNA samples with the AlkB has been partially successful in circumventing inhibitory base modifications (Cozen et al., 2015; Scheepbouwer et al., 2023; Warren, Salinas-Giegé, Hummel, et al., 2021; Zheng et al., 2015), the RNA degradation caused by its iron-containing buffer is likely to be especially problematic in the context of assays designed to use CCA-tail integrity as a proxy for aminoacylation state. Continued

optimization of RT activity may instead be key to improving representation of full-length tRNAs. The MSR-seq protocol includes an unusually long (overnight) RT incubation step to partially overcome the inhibitory effects of base modifications. Template-switching RTs derived from group II introns have also been used in to improve processivity in some tRNA-seq studies (Gołębiewska et al., 2024; Scheepbouwer et al., 2023; Zheng et al., 2015), although they have not always yielded superior results and there is potential for template switching to yield artifactual chimeric products. Recent strides have also been made to circumvent RT-associated challenges entirely by direct sequencing of tRNAs with nanopore or mass spectrometry technology (Lucas et al., 2023; Sun et al., 2023; Thomas et al., 2021; White et al., 2024; Yuan et al., 2024).

Our main objective in this study was to infer the plant tRNA aminoacylation landscape, and our results suggest that MSR-seq coupled with periodate treatment is a promising approach. In particular, this method appears to be highly effective at distinguishing a synthetic spike-in tRNA with no amino acid from functional biological tRNAs (Figure 4), which lends confidence to biological conclusions about tRNA-like sequences that do not appear to be aminoacylated and are, thus, unlikely to participate in translation (see below). However, our study also identified limitations associated with using periodate-induced read-length variation as a proxy for aminoacylation. The fact that current treatment conditions do not consistently remove one (and only one) nucleotide from uncharged tRNAs or consistently leave aminoacylated tRNAs intact makes the raw estimates of aminoacylation rates unreliable. This inconsistency also confounds the protective effect from amino acids with any true biological differences in tRNA end length. Fortunately, an extensive effort was recently made to optimize periodate treatments for tRNA-seq (Davidsen & Sullivan, 2024), which identified multiple factors that would likely reduce damage-related artifacts observed in our study. Potential modifications to our protocol include performing periodate treatment in the dark, reducing the periodate incubation temperature and duration, and using a lysine-containing buffer (pH 8.0) as a source of primary amine instead of the more alkaline sodium tetraborate treatment. Our data also suggest that incomplete deacylation can be a significant source of bias in tRNA-seq more generally (Figure 3), so optimization to ensure complete deacylation may also be important.

Variation in aminoacylation rates between nuclear and organellar tRNAs

One unexpected result from our analysis was the difference in average CCA tail integrity between nuclear and organellar tRNAs (Figure 2; Figure S2). This observation raises the question as to whether these two classes have true biological differences in the standing ratio of

aminoacylated vs. uncharged tRNAs or differences in susceptibility to degradation that affects CCA tail integrity. The notion that aminoacylation rates and/or CCA tail integrity could differ substantially across subcellular environments is plausible. For example, it was recently shown that mammals sequester tRNAs in extracellular vesicle that differ in CCA tail integrity from the rest of the cellular tRNA pool despite similar base-modification profiles (Scheepbouwer et al., 2024). However, the potential for methodological artifacts should also be considered. We found similar differences in CCA tail integrity between nuclear and mitochondrial tRNAs from human cell culture (Figure S3) in a previously published MSR-seq dataset generated with the same periodate treatment protocol (Watkins et al., 2022). However, other tRNA-seq analyses of aminoacylation levels in human cell lines either did not see this same pattern (Evans et al., 2017) or even showed the opposite trend with lower inferred aminoacylation levels for mitochondrial tRNAs (Davidsen & Sullivan, 2024). Therefore, future investigations will be needed to pinpoint the cause of this variation across studies. Regardless, focusing on gene-level differences will also be important because it appears that variation among genes is larger than the average difference between organellar and nuclear tRNAs.

Functionality of mitochondrial t-elements and other tRNA-like sequences

To illustrate the utility of tRNA-seq approaches, we used our data to interpret the expression and functional interactions of mitochondrial t-elements. These tRNA-like sequences are already known to be transcribed and cleaved as part of expression with adjoining mRNA sequences, in addition to evidence for post-transcriptional addition of nucleotides to their 3' end (Forner et al., 2007; Hanic-Joyce et al., 1990). We found that a full CCA tail was the most or one of the most common post-transcriptional additions (none of the t-elements have a genomically encoded CCA tail) (Table 1). This observation suggests that the 3' additions are mediated by a conventional CCase (von Braun et al., 2007) despite evidence for substantial non-canonical tailing (Forner et al., 2007). Notably, some of the annotated t-elements in the Arabidopsis mitochondrial genome were rarely or never detected by either YAMAT-seq or MSR-seq (Dataset S1), raising the possibilities that they are either expressed at very low levels (or not at all) or rapidly degraded.

In addition to CCA-tailing, our data show that mitochondrial t-elements can have post-transcriptional base modifications (Figure 9b), implying an even larger set of tRNA-interacting enzymes are capable of recognizing t-elements. Importantly, however, we see almost complete elimination of t-element terminal 3' nucleotides in response to periodate treatment (Figure 9a; Dataset S1). Despite caveats about inferring aminoacylation levels (see

above), the extreme difference in CCA tail integrity in periodate-treated libraries relative to actual tRNAs that are known to function in translation appears to be strong evidence that these transcripts are rarely, if ever, recognized and charged by aaRS. We also found that the expressed family of tRNA-Phe-like sequences and the *orf315*-adjacent stem-loop in the mitochondrial genome show little or no evidence of aminoacylation. It is, thus, unlikely that any of these elements play a canonical tRNA role as decoder molecules in translation, but whether they play some other mitochondrial or cellular function remains to be determined.

Although our study has only investigated canonical tRNAs in plastids, it is known that RNA secondary structure plays an important role in defining plastid mRNA end maturation and that the combined action of RNases and RNA-binding proteins can create small non-coding RNAs in plastids (Castandet et al., 2019; Ruwe et al., 2016; Small et al., 2023). These observations raise the question as to whether any plastid sequences play tRNA-like roles or function in a fashion similar to mitochondrial t-elements. To our knowledge, it has not been determined whether any of these transcripts interact with enzymatic machinery typically involved in tRNA metabolism. However, it has been hypothesized that antisense tRNAs expressed adjacent to protein-coding genes might serve as a suitable substrate because the “mirrored” nature of these transcripts leads them to adopt a tRNA-like secondary structure (Gawroński et al., 2020). Even though many transcripts that contain sequence antisense to plastid tRNA genes are produced (Hotto et al., 2011), an earlier analysis of the angiosperm *Silene vulgaris* did not find convincing evidence that these “mirror tRNAs” were post-transcriptionally CCA-tailed (Warren & Sloan, 2021). Nevertheless, future studies should comprehensively investigate the expression, processing, post-transcriptional modification, and even possible aminoacylation of the plastid non-coding transcript pool.

Overall, our findings highlight the level of caution and specificity that is needed in interpreting functionality of tRNA(-like) genes. Although it is tempting to use transcription, cleavage, CCA-tailing, and post-transcriptional modifications as evidence for a functional role in translation, it is important to test for aminoacylation and, even better, loading of tRNAs onto ribosomes (Chen & Tanaka, 2018) to make such determinations. Fortunately, the rapid and ongoing advances in specialized tRNA-seq methods is creating opportunities for such comprehensive assessments of tRNA metabolism.

METHODS

Plant growth and RNA extractions

Arabidopsis thaliana Col-0 seeds were surface sterilized for 10 min in a 50% bleach solution with Tween 20 (VWR; 0777A)

followed by three washes with sterile dH₂O. They were then sown onto MS-agar plates, consisting of 2.2 g of Murashige and Skoog basal salts, 0.5 g β-(N-morpholino)ethanesulfonic acid, and 7.5 g of Phytoblend plant agar (PTP01; Caisson Labs) per liter. Plates were then sealed in parafilm and cold-stratified at 4°C for 2 days before being transferred to an incubator for growth under a 16 h–8 h light/dark cycle. Plants were grown for 4 weeks for the first MSR-seq experiment. Another batch of seeds was sown and grown for 3 weeks for the second experiment. For each biological replicate (three in the first experiment and two for each of the two RNA extraction methods in the second experiment), one rosette was pinched off from its root system with forceps and flash frozen in liquid N₂ (~15 to 20 mg of tissue per rosette).

All samples in the first experiment and half of the samples in the second experiment were extracted with a modified version of an acid-phenol protocol (Zaborske et al., 2009). Flash-frozen tissue was immediately ground with a micropestle and resuspended in 500 μl of lysis buffer consisting of 300 mM NaOAc pH 4.5 (Thermo Fisher; AAJ63669AE), 10 mM EDTA, pH 8.0 (Invitrogen; AM9262), and 1% sodium dodecyl sulfate in RNase-free dH₂O. Five hundred microlitres of an acid-phenol, chloroform, isoamyl-alcohol mixture (125:24:1) pH 4.5 (Invitrogen; AM9720) was then added and followed by three bouts of ~15 to 30 sec of vortexing, keeping samples on ice between bouts. Samples were then centrifuged for 15 min at 20 000 g and 4°C. The resulting aqueous was then transferred to an additional 500 μl of the acid-phenol, chloroform, isoamyl-alcohol mixture and vortexed, followed by centrifugation under the same settings as above. After centrifugation, the aqueous phase was collected and ethanol precipitated, and the final pellet was dissolved in 20 μl of 10 mM NaOAc pH 4.5. The other half of the samples from the second experiment were extracted with Trizol (Ambion; 15596018), following manufacturer's instruction. The final pellet from each Trizol extraction was dissolved in 20 μl RNase-free dH₂O.

MSR-seq library construction and sequencing

Sequencing libraries were generated following the published MSR-seq protocol (Watkins et al., 2022) with minor modifications. Specific differences relative to the published version included the following: (i) we added 0.1% Tween 20 (VWR 0777A) in the high-salt buffer, (ii) we used a single capture hairpin oligo (CHO; Table S2), so each library was processed in a separate tube with no multiplexing until the conventional Illumina indexes were incorporated at the PCR stage, (iii) prior to PCR amplification, beads with ligated adapter molecules were stored overnight at –20°C because we found that doing so resulted in more consistent amplification across samples, and (4) after amplification, libraries were pooled for subsequent size selection on a BluePipin (Sage Science) with a 3% cassette and Q3 Marker, using “Range” settings and size limits of 170–275 bp. This size selection window effectively eliminated small RNAs (e.g., miRNAs) and longer mRNAs/rRNAs, yielding libraries that were dominated by tRNAs (Table S1) without any other forms of targeted enrichment. A detailed version of the protocol used in these experiments is available as Supporting Information and via GitHub (https://github.com/dbsloan/Arabidopsis_aminoacylation).

In our first experiment, RNA samples were subjected to one of three different treatments: (i) standard periodate treatment to infer aminoacylation state, (ii) periodate treatment preceded by a deacylation treatment, and (iii) a no-periodate control treatment. Deacylation reactions (which were required for treatments #2 and #3) involved a 30 min incubation at 37°C in 50 mM Tris-HCl pH 9.0, followed by ethanol precipitation. Periodate treatments were

performed as described in the original MSR-seq publication (Watkins et al., 2022), using 500 ng of input RNA (either untreated or pre-deacylated). For the no-periodate controls, 1 µg of deacylated RNA was used as input for the library construction protocol.

In our second experiment, only the standard periodate and no-periodate control treatments were performed (i.e., there was no pre-deacylation treatment group), using RNA samples from the two different extraction protocols (acid-phenol and Trizol). We also included an internal control in these libraries by spiking in 1 ng of a synthesized tRNA (with the sequence of *Bacillus subtilis* tRNA-Ile [GenBank NC_000964.3: 31932–32008]; Table S2). This spike-in was added immediately prior to the periodate or deacylation step in the respective treatment.

After amplification and size selection, the final library pools were sent to Novogene for 2 × 150 bp paired-end sequencing on an Illumina NovaSeq platform. Raw reads for each library have been deposited into the NCBI Sequence Read Archive (SRA) (Table S1).

MSR-seq read processing and mapping

All scripts used in read processing and subsequent data analysis are available via our GitHub repository (https://github.com/dbsloan/Arabidopsis_aminoacylation). BBMerge (Bushnell et al., 2017) was used to combine each pair of R1 and R2 reads into a single sequence with the following parameters: ordered=t qtrim=r minoverlap=30 mismatches=4. Merged reads were then further trimmed with a custom script to remove the adapter sequence from the MSR-seq CHO. Read pairs that could not be combined by BBMerge or did not have the expected CHO adapter sequence (Table S1) were excluded from downstream steps.

All successfully processed reads were mapped against a reference that consisted of the unique *Arabidopsis thaliana* Col-0 tRNA sequences in the plantRNA database (Cognat et al., 2022) as of December 20, 2023. Later versions of our reference set also included the synthetic spike-in *B. subtilis* tRNA-Ile sequence and mitochondrial t-element and stem-loop sequences (see below). Mapping was conducted with Bowtie v2.2.5 (Langmead & Salzberg, 2012), using the following sensitive parameters to facilitate detection of reads with mismatches and indels that were introduced due to post-transcriptional base modifications: -p 12 -L 10 -i C,1 --mp 5,2 --score-min L,-0.7,-0.7. However, even with these sensitive settings, mapping MSR-seq reads to tRNA references can be problematic because RT terminal transferase activity results in extra nucleotides being added past the 5' end of the original RNA template and because reads that extend beyond the ends of reference sequences are heavily penalized under the Bowtie global alignment settings. Therefore, we used a custom local alignment pipeline based on NCBI BLASTN 2.14.1+ (Camacho et al., 2009) to pre-process reads and trim 5' extensions prior to Bowtie mapping. A custom script was then used to parse the Sequence Alignment Map (SAM) file generated by Bowtie, quantifying the number of reads that mapped to each reference sequence and determining whether they contained a full CCA tail or were lacking one or more 3' nucleotides. This script was run both with and without an option to exclude reads that map equally well to two or more reference sequences.

Detecting RT misincorporations and truncations

SAM files were also parsed to identify reads with mismatches or small indels relative to the reference sequence, which can be caused by RT misincorporations in response to post-transcriptionally modified bases in the RNA template (Arrivé et al., 2023; Behrens et al., 2021; Clark et al., 2016; Enroth

et al., 2019; Gołębiewska et al., 2024; Potapov et al., 2018; Vandivier et al., 2019; Warren, Salinas-Giegé, Hummel, et al., 2021). For a global analysis of variants across all reference tRNAs (Figure 8), we used a previously published variant calling workflow (Edera & Sanchez-Puerta, 2021). We also used Perbase v0.9.0 (<https://github.com/sstadick/perbase>) for analysis of sequence variants in some specific tRNAs of interest (Figure 9; Figure S7). In addition, the mapping positions of 5' read ends were also parsed from SAM files to infer internal truncation points, which may result from base-modifications that block RT progression or from strand breaks (i.e., 3' tRFs). Scripts used to perform SAM file parsing and data visualizations are all available via our GitHub repository (https://github.com/dbsloan/Arabidopsis_aminoacylation).

Sprinzi coordinate system

Although tRNAs share a typical structural organization comprised of multiple stems and loops (Figure 7), the exact nucleotide positioning is not identical across tRNAs because of differences in the lengths of some of these structural motifs and certain “variable” elements. Therefore, we applied a standardized numbering scheme (Sprinzi et al., 1998) that anchored gene-specific nucleotide positions to their position in the canonical tRNA structure, using the classification of structural elements from the plantRNA database (Cognat et al., 2022) and a custom script (https://github.com/dbsloan/Arabidopsis_aminoacylation).

Analysis of mitochondrial t-elements, pseudogenes, and stem-loops

To investigate whether mitochondrial t-elements and other tRNA-like sequences are expressed, post-transcriptionally processed, and aminoacylated, we first used existing YAMAT-seq data (Warren, Salinas-Giegé, Hummel, et al., 2021) with a workflow to identify and quantify mitochondrially mapping reads that do not overlap with annotated tRNAs or other genes as previously described (Warren, Salinas-Giegé, Triant, et al., 2021). Based on the results of this analysis, we included the eight annotated t-elements in the Arabidopsis reference mitochondrial genome (NC_037304.1), as well as a set of tRNA-Phe-like sequences, and an expressed stem-loop associated with *orf315* in our tRNA reference set used for mapping (Dataset S1). This reference set is available via GitHub (https://github.com/dbsloan/Arabidopsis_aminoacylation). The MSR-seq mapping workflow described above was then re-run with this updated reference to quantify read abundance for these mitochondrial sequences and determine whether they retained intact CCA tails.

Analysis of 5' tRFs

To identify tRFs that represent the 5' portion of tRNA genes, we parsed SAM files and screened for reads that terminated >7 nt from the annotated 3' end of the reference tRNA sequence. We identified tRFs as high frequency if the reads ending at that position represented >5% of all reads mapping to that tRNA. If other positions separated by three or fewer nt also cleared this threshold, we considered them to be a single tRF and assigned the tRF position to 3'-most site above this threshold. Custom scripts were used to apply these criteria (https://github.com/dbsloan/Arabidopsis_aminoacylation).

Reanalysis of published MSR-seq data from human cell lines

To determine whether previously published MSR-seq data from human cell lines (Watkins et al., 2022) showed similar frequencies

of non-CC/CCA reads and differences between nuclear and mitochondrial tRNAs in CCA percentages as we observed in Arabidopsis, we re-analyzed sequences obtained from NCBI SRA (SRR18302277, SRR18302278, and SRR18302279) with our pipeline. These three datasets were from control samples (i.e., no stress treatment), with a standard periodate protocol used in library construction (Watkins et al., 2022).

ACKNOWLEDGEMENTS

We thank Lio Lee and Amanda Broz for assistance with plant growth and Chris Katanski for helpful methodological advice. We also thank two anonymous reviewers for helpful comments on an earlier version of the manuscript. This work was supported by funding from the National Science Foundation (MCB-2322154), an IUBMB Wood-Whelan Research Fellowship, and an HHMI Hanna H. Gray Fellowship.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

MSR-seq reads have been deposited to NCBI SRA (BioProject PRJNA1154980). Custom scripts and processed data are available via GitHub (https://github.com/dbsloan/Arabidopsis_aminoacylation).

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Figure S1. Pairwise matrix showing correlation in read abundance measurements for each library.

Figure S2. Variation among Arabidopsis tRNA isodecoder families and genomic compartments in CCA tail integrity in response to periodate treatment.

Figure S3. Variation among human cell line (HEK293T) tRNA isoacceptor families and genomic compartments in CCA tail integrity in response to periodate treatment.

Figure S4. Comparison of acid-phenol and Trizol RNA extraction methods as input into Arabidopsis MSR-seq libraries.

Figure S5. Comparison of RNA samples before (gray) and after (red) treatment with sodium periodate, ribose, sodium tetraborate, and T4 polynucleotide kinase according to the MSR-seq protocol.

Figure S6. 5' Mapping position of MSR-seq reads for Arabidopsis organellar tRNAs.

Figure S7. Read depth and misincorporation profiles for two tRNA isodecoder families that both show a major 5' truncation point induced by periodate treatment.

Figure S8. 3' Mapping position of MSR-seq reads for Arabidopsis organellar tRNAs.

Figure S9. Summary of 5' tRFs.

Figure S10. Relationship between nucleotide misincorporation rates in reads with intact CCA tails (aminoacylated tRNAs) vs. reads with CC tails (uncharged tRNAs).

Figure S11. Relationship between percentage of reads with intact CCA tails and 5' truncations as a proxy for "hard-stop" base modifications.

Figure S12. 5' Mapping position of MSR-seq reads from periodate-treated libraries for Arabidopsis nuclear tRNAs.

Figure S13. 5' Mapping position of MSR-seq reads from periodate-treated libraries for Arabidopsis organellar tRNAs.

Figure S14. Comparison between read abundance for mitochondrial tRNAs (left) and mitochondrial tRNA-like sequences (right), including t-elements, tRNA-Phe-like sequences, and the orf315 stem-loop.

Table S1. Summary of MSR-seq libraries generated for this study, including number of reads produced, successfully processed.

Table S2. Oligonucleotides used in library construction.

REFERENCES

- Arrivé, M., Bruggeman, M., Skaltsogiannis, V., Coudray, L., Quan, Y.-F., Schelcher, C. et al. (2023) A tRNA-modifying enzyme facilitates RNase P activity in Arabidopsis nuclei. *Nature Plants*, **9**(12), 2031–2041.
- Avçilar-Kucukgoze, I., Bartholomäus, A., Cordero Varela, J.A., Kaml, R.F.-X., Neubauer, P., Budisa, N. et al. (2016) Discharging tRNAs: a tug of war between translation and detoxification in *Escherichia coli*. *Nucleic Acids Research*, **44**(17), 8324–8334.
- Behrens, A., Rodschinka, G. & Nedialkova, D.D. (2021) High-resolution quantitative profiling of tRNA abundance and modification status in eukaryotes by mim-tRNAseq. *Molecular Cell*, **81**(8), 1802–1815.e7.
- Betat, H. & Mörl, M. (2015) The CCA-adding enzyme: a central scrutinizer in tRNA quality control. *BioEssays*, **37**(9), 975–982.
- Bonilla, S.L., Sherlock, M.E., MacFadden, A. & Kieft, J.S. (2021) A viral RNA hijacks host machinery using dynamic conformational changes of a tRNA-like structure. *Science*, **374**(6570), 955–960.
- Bushnell, B., Rood, J. & Singer, E. (2017) BBMerge—accurate paired shotgun read merging via overlap. *PLoS One*, **12**(10), e0185056.
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K. et al. (2009) BLAST+: architecture and applications. *BMC Bioinformatics*, **10**, 421.
- Cappannini, A., Ray, A., Purta, E., Mukherjee, S., Boccaletto, P., Moafinejad, S.N. et al. (2024) MODOMICS: a database of RNA modifications and related information. 2023 update. *Nucleic Acids Research*, **52**(D1), D239–D244.
- Castandet, B., Germain, A., Hotto, A.M. & Stern, D.B. (2019) Systematic sequencing of chloroplast transcript termini from *Arabidopsis thaliana* reveals >200 transcription initiation sites and the extensive imprints of RNA-binding proteins and secondary structures. *Nucleic Acids Research*, **47**(22), 11889–11905.
- Chen, C.-W. & Tanaka, M. (2018) Genome-wide translation profiling by ribosome-bound tRNA capture. *Cell Reports*, **23**(2), 608–621.
- Clark, W.C., Evans, M.E., Dominissini, D., Zheng, G. & Pan, T. (2016) tRNA base methylation identification and quantification via high-throughput sequencing. *RNA*, **22**(11), 1771–1784.
- Cognat, V., Pawlak, G., Pflieger, D. & Drouard, L. (2022) PlantRNA 2.0: an updated database dedicated to tRNAs of photosynthetic eukaryotes. *The Plant Journal*, **112**(4), 1112–1119.
- Costa, B., Li Calzi, M., Castellano, M., Blanco, V., Cuevasanta, E., Litvan, I. et al. (2023) Nicked tRNAs are stable reservoirs of tRNA halves in cells and biofluids. *Proceedings of the National Academy of Sciences of the United States of America*, **120**(4), e2216330120.
- Cozen, A.E., Quartley, E., Holmes, A.D., Hrabeta-Robinson, E., Phizicky, E.M. & Lowe, T.M. (2015) ARM-seq: AlkB-facilitated RNA methylation sequencing reveals a complex landscape of modified tRNA fragments. *Nature Methods*, **12**(9), 879–884.
- Czech, A. (2020) Deep sequencing of tRNA's 3'-termini sheds light on CCA-tail integrity and maturation. *RNA*, **26**(2), 199–208.
- Davidson, K. & Sullivan, L.B. (2024) A robust method for measuring aminoacylation through tRNA-seq. *eLife*, **12**, 91554.
- Dittmar, K.A., Sørensen, M.A., Elf, J., Ehrenberg, M. & Pan, T. (2005) Selective charging of tRNA isoacceptors induced by amino-acid starvation. *EMBO Reports*, **6**(2), 151–157.
- Edera, A.A. & Sanchez-Puerta, M.V. (2021) Computational detection of plant RNA editing events. *Methods in Molecular Biology*, **2181**, 13–34.
- Enroth, C., Poulsen, L.D., Iversen, S., Kirpekar, F., Albrechtsen, A. & Vinther, J. (2019) Detection of internal N7-methylguanosine (m7G) RNA modifications by mutational profiling sequencing. *Nucleic Acids Research*, **47**(20), e126.

- Evans, M.E., Clark, W.C., Zheng, G. & Pan, T. (2017) Determination of tRNA aminoacylation levels by high-throughput sequencing. *Nucleic Acids Research*, **45**(14), e133.
- Forner, J., Weber, B., Thuss, S., Wildum, S. & Binder, S. (2007) Mapping of mitochondrial mRNA termini in *Arabidopsis thaliana*: t-elements contribute to 5' and 3' end formation. *Nucleic Acids Research*, **35**(11), 3676–3692.
- Gawronski, P., Palac, A. & Scharff, L.B. (2020) Secondary structure of chloroplast mRNAs in vivo and in vitro. *Plants*, **9**(3), 323.
- Giegé, R. & Eriani, G. (2023) The tRNA identity landscape for aminoacylation and beyond. *Nucleic Acids Research*, **51**, 1528–1570.
- Golebiewska, K., Gregorová, P., Peter Sarin, L. & Gawronski, P. (2024) Deciphering the RNA modification landscape in Arabidopsis chloroplast tRNAs and rRNAs reveals a blend of ancestral and acquired characteristics. *bioRxiv*. 2024.06.14.598963. Available from: <https://doi.org/10.1101/2024.06.14.598963>
- Hanic-Joyce, P.J., Spencer, D.F. & Gray, M.W. (1990) In vitro processing of transcripts containing novel tRNA-like sequences ('t-elements') encoded by wheat mitochondrial DNA. *Plant Molecular Biology*, **15**(4), 551–559.
- Hartmann, R., Gössringer, M., Späth, B., Fischer, S. & Marchfelder, A. (2009) The making of tRNAs and more – RNase P and tRNase Z. *Progress in Molecular Biology and Translational Science*, **85**, 319–368.
- Hauenschild, R., Tserovski, L., Schmid, K., Thüning, K., Winz, M.-L., Sharma, S. et al. (2015) The reverse transcription signature of N1-methyladenosine in RNA-seq is sequence dependent. *Nucleic Acids Research*, **43**(20), 9950–9964.
- Hentzen, D., Mandel, P. & Garel, J.P. (1972) Relation between aminoacyl-tRNA stability and the fixed amino acid. *Biochimica et Biophysica Acta*, **281**(2), 228–232.
- Hernandez-Alias, X., Katanski, C.D., Zhang, W., Assari, M., Watkins, C.P., Schaefer, M.H. et al. (2023) Single-read tRNA-seq analysis reveals coordination of tRNA modification and aminoacylation and fragmentation. *Nucleic Acids Research*, **51**(3), e17.
- Holley, R.W., Appgar, J., Everett, G.A., Madison, J.T., Marquisee, M., Merrill, S.H. et al. (1965) Structure of a ribonucleic acid. *Science*, **147**(3664), 1462–1465.
- Hotto, A.M., Schmitz, R.J., Fei, Z., Ecker, J.R. & Stern, D.B. (2011) Unexpected diversity of chloroplast noncoding RNAs as revealed by deep sequencing of the Arabidopsis transcriptome. *G3 (Bethesda)*, **1**(7), 559–570.
- Katanski, C.D., Watkins, C.P., Zhang, W., Reyer, M., Miller, S. & Pan, T. (2022) Analysis of queuosine and 2-thio tRNA modifications by high throughput sequencing. *Nucleic Acids Research*, **50**(17), e99.
- Keam, S.P. & Hutvagner, G. (2015) tRNA-derived fragments (tRFs): emerging new roles for an ancient RNA in the regulation of gene expression. *Life (Basel)*, **5**(4), 1638–1651.
- Lalonde, S., Merret, R., Salinas-Giegé, T. & Drouard, L. (2020) Arabidopsis tRNA-derived fragments as potential modulators of translation. *RNA Biology*, **17**(8), 1137–1148.
- Langmead, B. & Salzberg, S.L. (2012) Fast gapped-read alignment with bowtie 2. *Nature Methods*, **9**(4), 357–359.
- Lucas, M.C., Prysacz, L.P., Medina, R., Milenkovic, I., Camacho, N., Marchand, V. et al. (2023) Quantitative analysis of tRNA abundance and modifications by nanopore RNA sequencing. *Nature Biotechnology*, **42**(1), 72–86.
- Ma, X., Liu, C., Kong, X., Liu, J., Zhang, S., Liang, S. et al. (2021) Extensive profiling of the expressions of tRNAs and tRNA-derived fragments (tRFs) reveals the complexities of tRNA and tRF populations in plants. *Science China. Life Sciences*, **64**(4), 495–511.
- Masuda, I. & Hou, Y.-M. (2024) A tRNA modification pattern that facilitates interpretation of the genetic code. *Frontiers in Microbiology*, **15**, 1415100.
- Motorin, Y. & Helm, M. (2010) tRNA stabilization by modified nucleotides. *Biochemistry*, **49**(24), 4934–4944.
- Padhiar, N.H., Katneni, U., Komar, A.A., Motorin, Y. & Kimchi-Sarfaty, C. (2024) Advances in methods for tRNA sequencing and quantification. *Trends in Genetics*, **40**(3), 276–290.
- Pavlova, N.N., King, B., Josselson, R.H., Violante, S., Macera, V.L., Vardhana, S.A. et al. (2020) Translation in amino-acid-poor environments is limited by tRNAGln charging. *eLife*, **9**, e62307.
- Phizicky, E.M. & Hopper, A.K. (2010) tRNA biology charges to the front. *Genes & Development*, **24**(17), 1832–1860.
- Pinkard, O., McFarland, S., Sweet, T. & Collier, J. (2020) Quantitative tRNA-sequencing uncovers metazoan tissue-specific tRNA regulation. *Nature Communications*, **11**(1), 4104.
- Potapov, V., Fu, X., Dai, N., Correa, I.R., Jr., Tanner, N.A. & Ong, J.L. (2018) Base modifications affecting RNA polymerase and reverse transcriptase fidelity. *Nucleic Acids Research*, **46**(11), 5753–5763.
- Raina, M. & Ibba, M. (2014) tRNAs as regulators of biological processes. *Frontiers in Genetics*, **5**, 171.
- Ruwe, H., Wang, G., Gusewski, S. & Schmitz-Linneweber, C. (2016) Systematic analysis of plant mitochondrial and chloroplast small RNAs suggests organelle-specific mRNA stabilization mechanisms. *Nucleic Acids Research*, **44**(15), 7406–7417.
- Scheepbouwer, C., Aparicio-Puerta, E., Gómez-Martin, C., van Eijndhoven, M.A.J., Drees, E.E.E., Bosch, L. et al. (2024) Full-length tRNAs lacking a functional CCA tail are selectively sorted into the lumen of extracellular vesicles. *bioRxiv*. Available from: <https://doi.org/10.1101/2024.05.12.593148>
- Scheepbouwer, C., Aparicio-Puerta, E., Gomez-Martin, C., Verschuere, H., van Eijndhoven, M., Wedekind, L.E. et al. (2023) ALL-tRNAseq enables robust tRNA profiling in tissue samples. *Genes & Development*, **37**(5–6), 243–257.
- Shigematsu, M., Honda, S., Loher, P., Telonis, A.G., Rigoutsos, I. & Kirino, Y. (2017) YAMAT-seq: an efficient method for high-throughput sequencing of mature transfer RNAs. *Nucleic Acids Research*, **45**(9), e70.
- Skeparnis, I., Bou-Nader, C., Anastasakis, D.G., Fan, L., Wang, Y.-X., Hafner, M. et al. (2024) Structural basis of MALAT1 RNA maturation and mscRNA biogenesis. *Nature Structural & Molecular Biology*, **31**, 1655–1668. Available from: <https://doi.org/10.1038/s41594-024-01340-4>
- Sloan, D.B., Wu, Z. & Sharbrough, J. (2018) Correction of persistent errors in Arabidopsis reference mitochondrial genomes. *The Plant Cell*, **30**(3), 525–527.
- Small, I., Melonek, J., Böhne, A.-V., Nickelsen, J. & Schmitz-Linneweber, C. (2023) Plant organellar RNA maturation. *The Plant Cell*, **35**(6), 1727–1751.
- Sprinzel, M., Horn, C., Brown, M., Ioudovitch, A. & Steinberg, S. (1998) Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Research*, **26**(1), 148–153.
- Subramaniam, A.R., Pan, T. & Cluzel, P. (2013) Environmental perturbations lift the degeneracy of the genetic code to regulate protein levels in bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, **110**(6), 2419–2424.
- Sun, Y., Piechotta, M., Naarmann-de Vries, I., Dieterich, C. & Ehrenhofer-Murray, A.E. (2023) Detection of queuosine and queuosine precursors in tRNAs by direct RNA sequencing. *Nucleic Acids Research*, **51**(20), 11197–11212.
- Suzuki, T. (2021) The expanding world of tRNA modifications and their disease relevance. *Nature Reviews. Molecular Cell Biology*, **22**(6), 375–392.
- Thomas, N.K., Poodari, V.C., Jain, M., Olsen, H.E., Akeson, M. & Abushumays, R.L. (2021) Direct nanopore sequencing of individual full length tRNA strands. *ACS Nano*, **15**(10), 16642–16653.
- Vandivier, L.E., Anderson, Z.D. & Gregory, B.D. (2019) HAMR: high-throughput annotation of modified ribonucleotides. *Methods in Molecular Biology*, **1870**, 51–67.
- von Braun, S.S., Sabetti, A., Hanic-Joyce, P.J., Gu, J., Schleiff, E. & Joyce, P.B.M. (2007) Dual targeting of the tRNA nucleotidyltransferase in plants: not just the signal. *Journal of Experimental Botany*, **58**(15–16), 4083–4093.
- Warren, J.M., Salinas-Giegé, T., Hummel, G., Coots, N.L., Svendsen, J.M., Brown, K.C. et al. (2021) Combining tRNA sequencing methods to characterize plant tRNA expression and post-transcriptional modification. *RNA Biology*, **18**, 64–78.
- Warren, J.M., Salinas-Giegé, T., Triant, D.A., Taylor, D.R., Drouard, L. & Sloan, D.B. (2021) Rapid shifts in mitochondrial tRNA import in a plant lineage with extensive mitochondrial tRNA gene loss. *Molecular Biology and Evolution*, **38**(12), 5735–5751.
- Warren, J.M. & Sloan, D.B. (2021) Hopeful monsters: unintended sequencing of famously malformed mite mitochondrial tRNAs reveals widespread expression and processing of sense-antisense pairs. *NAR Genomics and Bioinformatics*, **3**(1), lqaa111.
- Watkins, C.P., Zhang, W., Wylder, A.C., Katanski, C.D. & Pan, T. (2022) A multiplex platform for small RNA sequencing elucidates multifaceted tRNA stress response and translational regulation. *Nature Communications*, **13**(1), 2491.

- White, L.K., Dobson, K., Del Pozo, S., Bilodeaux, J.M., Andersen, S.E., Baldwin, A. *et al.* (2024) Comparative analysis of 43 distinct RNA modifications by nanopore tRNA sequencing. *bioRxiv*. Available from: <https://doi.org/10.1101/2024.07.23.604651>
- Wilusz, J.E. (2015) Removing roadblocks to deep sequencing of modified RNAs. *Nature Methods*, **12**(9), 821–822.
- Yuan, X., Su, Y., Johnson, B., Kirchner, M., Zhang, X., Xu, S. *et al.* (2024) Mass spectrometry-based direct sequencing of tRNAs de novo and quantitative mapping of multiple RNA modifications. *Journal of the American Chemical Society*, **146**, 25600–25613. Available from: <https://doi.org/10.1021/jacs.4c07280>
- Zaborske, J.M., Narasimhan, J., Jiang, L., Wek, S.A., Dittmar, K.A., Freimoser, F. *et al.* (2009) Genome-wide analysis of tRNA charging and activation of the eIF2 kinase Gcn2p. *The Journal of Biological Chemistry*, **284** (37), 25254–25267.
- Zheng, G., Qin, Y., Clark, W.C., Dai, Q., Yi, C., He, C. *et al.* (2015) Efficient and quantitative high-throughput tRNA sequencing. *Nature Methods*, **12** (9), 835–837.