DISSERTATION

LOSS AND INTEGRATION: TRACING THE FUNCTIONAL REPLACEMENT OF MITOCHONDRIAL TRNA GENES

Submitted by

Jessica Marie Warren

Department of Biology

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2021

Doctoral Committee:

Advisor: Daniel Sloan

Rachel Mueller
Anireddy S.N. Reddy
Graham Peers
Timothy Stasevich

ABSTRACT

LOSS AND INTEGRATION: TRACING THE FUNCTIONAL REPLACEMENT OF MITOCHONDRIAL TRNA GENES

The endosymbiotic origins of mitochondria and plastids have resulted in a partitioned eukaryotic cell with multiple genomes. During the course of billions of years since the acquisition of endosymbiotic organelles, these bacterial-like genomes have experienced extensive gene loss. This genomic instability is apparent in today's mitochondria, with extant mitogenomes encoding only a few dozen genes. Despite this almost wholesale loss of coding content, all mitogenomes still retain at least some genes involved in protein synthesis – one of the only classes of mitochondrial genes universally retained. Transfer RNAs (tRNAs), the decoding molecules in translation, are present in the mitogenomes of most eukaryotes, with the same tRNA complement being conserved in some clades for hundreds of millions of years.

However, exceptions to this stability in mitochondrial tRNA (mt-tRNA) retention have been found, as mitogenome sequencing in flowering plants has revealed multiple lineages that appear to be rapidly losing these last vestiges of bacterial ancestry. In the flowering plant genus *Silene,* close relatives have vastly different mt-tRNA genes, with some species retaining as few as only two in their mitogenomes. Such extreme mitochondrial gene loss raises questions about the functional replacement process involved in maintaining mitochondrial translation.

It has been assumed that lost mt-tRNA genes are functionally replaced by nuclear-encoded counterparts; however, investigation into the *de novo* evolution of tRNA import has been, in part, hampered by methodological limitations in tRNA detection and quantification. Additionally, tRNAs are not standalone molecules in translation, and the replacement of bacterial-like mitochondrial tRNAs with eukaryotic, nuclear-encoded tRNAs raises numerous questions about the identity and localization of tRNA-interacting enzymes that must process tRNAs for functionality. In Chapter 1 (*Interchangeable*

parts: The evolutionarily dynamic tRNA population in plant mitochondria), I review some of the distinctive features and outstanding questions pertaining to the dynamic tRNA pools in plant mitochondria.

Using a combination of innovative sequencing technologies including recently developed tRNA sequencing (tRNA-seq) protocols as well as full-length, mRNA sequencing methods, I take advantage of the recent and ongoing loss of mitochondrial tRNA genes in the angiosperm tribe *Sileneae* to investigate fundamental questions in eukaryogenesis, including the integration of organelles, mitochondrial gene loss, and evolution of chimeric enzymatic networks encoded across separate genomes.

One of the first steps in elucidating the evolutionary mechanisms involved in the replacement of mt-tRNA genes is identification of the nuclear genes that gain mitochondrial function. Even though a tRNA was the first nucleic acid to be sequenced, tRNAs are now arguably the most difficult class of RNAs to analyze with modern sequencing technologies. The highly modified nature of tRNA bases prevents efficient and error-free reverse transcription (RT), and the tightly paired 3′ and 5′ termini of tRNAs inhibits adapter ligation necessary for many library construction protocols. These barriers have prevented the global sequencing of mitochondrial tRNA pools in most organisms, including plants.

The recent development of two tRNA-Seq technologies, including the utilization of demethylating enzymes to remove RT-inhibiting modifications and development of RNA-seq adapters that specifically target mature tRNAs, has shown promise in increasing the amenability of tRNAs to high throughput sequencing. These technologies, however, have never been applied in a combined fashion to increase both the detection and read length of tRNAs, nor have they been applied to any plant system. In Chapter 2 (*Combining tRNA sequencing methods to characterize plant tRNA expression and post-transcriptional modification*), I demonstrate that the use of a demethylating enzyme (AlkB) and specialized adapters (YAMAT-seq) increases the detection of unique tRNA genes in the model plant species *Arabidopsis thaliana* (with over 80% of all nuclear tRNA genes detected), providing full length tRNA sequences that can be definitively mapped to individual reference genes.

This successful application of tRNA-seq in plant systems opens the door for global characterization of mitochondrial tRNA pools. In Chapter 3 (*Rapid shifts in mitochondrial tRNA import in a plant lineage with extensive mitochondrial tRNA gene loss*), I present tRNA-seq data from isolated mitochondria from five *Sileneae* species exhibiting rapid and ongoing mt-tRNA loss, as well as in the distantly related species *Solanum tuberosum*. Through the sequencing of tRNAs from mitochondrial isolates in conjunction with differential abundance analysis, I characterize the nuclear-encoded tRNAs that are likely imported and excluded from the mitochondria of each species. I detected changes in these import patterns in response to mt-tRNA gene loss across the *Sileneae* clade. Shifts from exclusion to import of nuclear-encoded tRNAs that corresponded to the loss of a mt-tRNA counterpart support the hypothesis that mt-tRNA gene loss is being compensated for through an increase in cytosolic tRNAs to maintain mitochondrial translation. In addition to functional gain of cytosolic tRNA import, there was also evidence that some nuclear-encoded tRNAs are imported in species that have yet to lose the mitochondrial counterpart. This redundant tRNA import has not been previously detected in plants and may offer a mechanism by which mt-tRNA loss is tolerated through functional redundancy of equivalent mitochondrial- and nuclear-encoded tRNAs in mitochondrial translation.

The functional replacement of mt-tRNAs by nuclear-encoded counterparts raises multiple questions about the identity of tRNA-interacting enzymes that must recognize newly imported tRNAs. Mitochondrial- and nuclear-encoded tRNAs share little sequence homology, and enzymes such as aminoacyl-tRNA synthetases (aaRSs) that charge these tRNAs with amino acids are highly discriminating, using specific nucleotides for substrate recognition. In plants, this divergence largely necessitates separate enzymes for recognition of nuclear and organellar tRNAs, thereby maintaining phylogenetic congruence between bacterial-like tRNA/aaRS and eukaryotic tRNA/aaRSs interactions.

In Chapter 4 (*Extensive retargeting of plant aminoacyl tRNA synthetases associated with mitochondrial tRNA gene loss*), I investigate whether the evolution of cytosolic tRNA import into mitochondria precipitates changes in the subcellular localization of corresponding cytosolic aaRSs in the same five *Sileneae* species. By analyzing full-length mRNA transcripts with single-molecule sequencing

technology (PacBio Iso-Seq), I report cases of predicted retargeting of ancestrally cytosolic aaRSs to mitochondria as well as scenarios where enzyme localization does not appear to change despite functional mt-tRNA replacement. The retargeting of only certain cytosolic aaRSs may offer clues as to which organellar aaRSs have relaxed substrate specificity, as these aaRSs may now aminoacylate nuclear-encoded tRNAs. Conversely, the functional replacement of other tRNAs appears to necessitate the paired localization of a cytosolic aaRS and corresponding cytosolic tRNA to mitochondria.

Similar to the redundant import of tRNAs found in Chapter 3, there were also cases of mitochondrial localization of an aaRS prior to the loss of a corresponding mt-tRNA gene. Targeting of nuclear-encoded proteins to the mitochondria typically involves an N-terminal extension of the protein known as a transit peptide. In most cases of transit peptide acquisition, the cytosolic enzyme appears to have gained a coding extension prior to the loss of the cognate mitochondrial tRNA. The import of both cytosolic aaRSs and tRNAs into the mitochondria prior to mitochondrial tRNA gene loss suggests that *Sileneae* species may experience transitional states where organellar and cytosolic enzymes colocalize to mitochondria and tRNA metabolism involves both native mitochondrial tRNAs as well as imported nuclear-encoded counterparts.

Taken together, this work suggests that mt-tRNA replacement can occur through a process of redundancy and loss. However, functional redundancy can be achieved in various ways, and is dependent on the particular tRNA gene in question. Furthermore, the evolution of import is a common and necessary theme to any explanation of mitochondrial gene replacement. The work here provides a framework for understanding the process of mitochondrial tRNA gene loss by catching these molecular changes “in the act”.

ACKNOWLEDGEMENTS

I want to thank all the incredible people I've met in Fort Collins while completing this dissertation. The relationships gained here are the highlights of my life. Maggie, thank you for such a sincere and constant love – you are family. Lysa, you're a sister, and I want to thank you for the comfort and familiarity I experience around you. Ashley, you are beautiful inside and out, and I consider myself lucky to be your friend. Thank you for all the amazing times. Scott, your affection for me has made me a happier and better person, I am so grateful for your friendship. Tomas, thank you for so much love and care, I am so happy I've got to know you. I'll always be there for you two. Nate, you are a deeply grounded person, and the people in your life trust you. That's a rare quality in life. Xoco, thanks for all the trouble we got into. You made this adventure better! Rob, you made this friend group funnier, and there were times when your kind words made a big difference in my life. Theresa and Sean, you guys are home base for our group, thank you for being there and always offering the space in your heart and home. Cantrell, thank you for all the help planning some of the best adventures, you reinforced such a love for the night sky. Alyx, your silliness and dark sense of humor brought a lot of light into my life – keep following that light. Gretchen, thanks for all the talks (about everything), you're going to do great. Dave and Whitney, you guys are so admirable, I'm excited to see where you go in life.

I also want to thank the people who have helped shaped the person I am. Aunt Cindy – you're my rock in the middle of an ocean. Knowing that you're there has given me the courage to explore far. My uncle Eric, heck yes, the morse code machine makes it into the acknowledgments. Thank you for giving me an incredible childhood memory and my first engineering experience— it worked! You two have given me roots in this world, I am so grateful. Eric, you've hung around for a long time in my life, thanks for being someone I always knew I could call. Oliver, I will be forever grateful for you taking care of me in France.

And finally, I want to thank my advisor. Dan, I am so proud of what we've accomplished. I've wanted to be an evolutionary biologist since I was eleven, and you have given me so much time towards

that goal. Thank you for making me a better scientist and person. This was truly one of the best experiences of my life. Here's to many more years of cool science.

TABLE OF CONTENTS

CHAPTER 1: INTERCHANGEABLE PARTS: THE EVOLUTIONARILY DYNAMIC TRNA POPULATION IN PLANT MITOCHONDRIA[1]

Summary

Transfer RNAs (tRNAs) remain one of the very few classes of genes still encoded in the mitochondrial genome. These key components of the protein translation system must interact with a large enzymatic network of nuclear-encoded gene products to maintain mitochondrial function. Plants have an evolutionarily dynamic mitochondrial tRNA population, including ongoing tRNA gene loss and replacement by both horizontal gene transfer from diverse sources and import of nuclear-expressed tRNAs from the cytosol. Thus, plant mitochondria represent an excellent model for understanding how anciently divergent genes can act as "interchangeable parts" during the evolution of complex molecular systems. In particular, understanding the integration of the mitochondrial translation system with elements of the corresponding machinery used in cytosolic protein synthesis is a key area for eukaryotic cellular evolution. Here, we review the increasingly detailed phylogenetic data about the evolutionary history of mitochondrial tRNA gene loss, transfer, and functional replacement that has created extreme variation in mitochondrial tRNA populations across plant species. We describe emerging tRNA-seq methods with promise for refining our understanding of the expression and subcellular localization of tRNAs. Finally, we summarize current evidence and identify open questions related to coevolutionary changes in nuclear-encoded enzymes that have accompanied turnover in mitochondrial tRNA populations.

Introduction: decoding the mitochondrial genome

[1] Published in *Mitochondrion*, Volume 52, May 2020.

Authors: Jessica M. Warren, Daniel B. Sloan

The endosymbiotic origin of mitochondria has partitioned the eukaryotic cell into multiple genomic compartments, each with its own distinct information storage and processing pathways. Separate replication, transcription and translational machinery is required for the maintenance and expression of the mitochondrial genome (mitogenome) and is largely distinct from the machinery required for nuclear function. During the course of the approximately 2 billion years since the intracellular integration of the alphaproteobacterial progenitor of mitochondria and its host cell, the mitogenome has experienced extensive gene loss, with extant mitogenomes encoding only a small fraction of genes typically found in bacterial genomes (SLOAN *et al*. 2018). This reduction in coding content has resulted in almost all components required for DNA replication, repair, and transcription being lost outright or functionally replaced by gene products encoded in the nuclear genome and imported into the mitochondrial matrix (HUYNEN *et al*. 2013). Yet despite this massive loss, all mitogenomes still retain at least some genes involved in protein synthesis (ROGER *et al*. 2017). These stubbornly retained translation genes, in combination with those that have been functionally transferred to the nuclear genome or recruited from existing eukaryotic genes to function in the mitochondria, have created an enzymatic mosaic where mitochondrial protein synthesis depends on the orchestrated expression and assembly of gene products from both the nuclear and mitochondrial genomes.

Mitochondrial translation, like its counterpart in the cytosol, is the fundamental process of decoding messenger RNA (mRNA) molecules into polymers of amino acids. The site of protein synthesis is the ribosome, a massive enzyme complex composed of both RNAs (rRNAs) and proteins (PETROV *et al*. 2015). Amino acids are brought to the ribosome by transfer RNAs (tRNAs) where they are covalently bonded to form growing polypeptide chains. As the adapter molecules of translation, tRNAs are approximately 80 nt in length with a distinctive secondary "cloverleaf" structure composed of base-paired stems and unpaired loops, and a tertiary L-shaped configuration that must interact with the ribosome, mRNA, and amino acid (RODNINA *et al*. 2005; ZHANG AND FERRE-D'AMARE 2016). Translation is achieved based on a 3-nt anticodon sequence in each tRNA that is used to recognize appropriate codons with the mRNA. Amino acids are attached to their corresponding tRNAs by a group of specialized

enzymes called aminoacyl tRNA synthetases (aaRSs) (O'DONOGHUE AND LUTHEY-SCHULTEN 2003). In addition to their function in decoding, tRNAs can play a number of diverse biosynthetic and regulatory roles, including antibiotic biosynthesis, cell wall remodeling, translation control and even cellular signaling (BANERJEE *et al*. 2010; PHIZICKY AND HOPPER 2010; RAINA AND IBBA 2014; KIRCHNER AND IGNATOVA 2015). tRNA-derived RNA fragments (tRFs) generated from both nuclear and mitochondrial tRNAs have also come under increased scrutiny for potential functions (COGNAT *et al*. 2017). Finally, protein synthesis also requires a suite of additional enzymes involved in the maturation and processing of tRNAs and rRNAs as well as numerous initiation, elongation, and termination factors (MOTORIN AND HELM 2010; BETAT *et al*. 2014).

The retention of certain translation genes in mitogenomes is clearly nonrandom, with the large and small subunit mitochondrial rRNA genes being universally retained, whereas genes such as aaRSs are only encoded in the nuclear genome (DUCHÊNE *et al*. 2009; OTT *et al*. 2016). The presence and absence of other genes related to protein synthesis is much more lineage-specific. The number and identity of tRNAs encoded in the mitogenome is one of the greatest sources of variation. At one extreme, some eukaryotes have a sufficient set of mitochondrial tRNA (mt-tRNA) genes to decode all codons, but others lack any mitochondrial-encoded tRNAs whatsoever (SCHNEIDER 2011; SALINAS-GIEGÉ *et al*. 2015; GROSJEAN AND WESTHOF 2016). Yet, despite this variability in mt-tRNA gene content across the eukaryotic tree of life, tRNA complements have stabilized for long periods of time in some clades. For example, the vast majority of bilaterian animals have the same 22 mt-tRNA genes (LAVROV AND PETT 2016).

This stability is in stark contrast to vascular plants, which have exceptionally dynamic and varied mitogenome tRNA content. In some species, the frequent insertion of foreign DNA into plant mitogenomes has inflated the number of mt-tRNA genes. Some of these insertions have functionally replaced the native mt-tRNA genes inherited from the ancestor of mitochondria (JOYCE AND GRAY 1989; MARCHFELDER *et al*. 1990; DUCHÊNE AND MARECHAL-DROUARD 2001; KITAZAKI *et al*. 2011). As tRNAs are likely the only class of mitochondrial genes to be functionally acquired from multiple foreign

sources, they represent a fascinating system to study mitochondrial transcription and RNA processing pathways. Conversely, other plants have had their mt-tRNA gene content dwindle down to none (HECHT *et al*. 2011), with translation likely being maintained entirely by the import of nuclear-encoded tRNA counterparts (MURCHA *et al*. 2016). In other lineages, this loss processes has not gone to completion, but instead appears to be rapid and ongoing (SLOAN *et al*. 2012a; PETERSEN *et al*. 2015). As much of the enzymatic machinery required for mitochondrial tRNA maturation and function is distinct from cytosolic tRNA machinery, the evolution of cytosolic tRNA import raises numerous questions about the effect on the enzymatic network involved in mitochondrial translation.

This variation in tRNA metabolism makes plants one of the best systems to study the coevolutionary dynamics necessary for the replacement of one of the last remaining classes of mitochondrial genes. This review will focus on the origins of the heterogeneous nature of plant mt-tRNA genes and how the ongoing replacement of the bacterial-like mt-tRNAs with imported nuclear tRNAs can perturb the network of enzymatic interactions essential for translation. Additionally, we provide a short discussion on how recent advances in tRNA sequencing technologies make this an exciting time to test longstanding hypotheses about the expression and intermediate states necessary for functional mt-tRNA gene loss.

Plant mitochondrial tRNA populations: Rapid shifts and heterogeneous origins

The early sequencing of plant mitogenomes and the characterization of their mt-tRNA populations revealed a strikingly heterogeneous mixture of tRNAs (DIETRICH *et al*. 1992; MARÉCHAL-DROUARD *et al*. 1993; KNOOP 2004). It became evident that the ancestral tRNA gene content within plant mitogenomes had been reshaped by a complex history of gene transfer events from different sources including plastids, other plants, algae, bacteria and possibly other distantly related lineages (Fig. 1.1). As the availability of plant mitogenome sequences has steadily increased, a more detailed view of donors and recipients in this dynamic process has come into focus. Moreover, many tRNAs that function in mitochondrial translation are encoded by nuclear genes and imported from the cytosol. The extent of this cytosolic replacement of native mt-tRNA genes also varies dramatically among plant lineages. Recent

work involving instances of closely related plant species with radically different complements of mt-tRNA genes suggest that the reshaping of imported tRNA pools can happen remarkably fast. The following sections will outline how the propensity for gene transfer and the ongoing evolution of tRNA import have resulted in plant mitochondria having some of the most complex tRNA populations and tRNA metabolisms in the tree of life.

The ancestral set of mitochondrial tRNA genes

The variable number of tRNA genes encoded in extant mitogenomes raises questions about the ancestral set of tRNAs present in the last eukaryotic common ancestor (LECA) and those that were subsequently retained in the lineage leading to plants. Reconstructing the ancestral gene content of mitogenomes has been aided by sequencing efforts in diverse lineages of protists from across the eukaryotic tree of life that have unusually gene-rich mitogenomes, including jakobids, malawimonads, *Ancoracysta*, and *Diphylleia* (BURGER *et al*. 2013; VALACH *et al*. 2014; KAMIKAWA *et al*. 2016; JANOUSKOVEC *et al*. 2017). However, far more attention has been paid to proteins than tRNAs in inferring the ancestral mitogenome content (LANG *et al*. 1999; JOHNSTON AND WILLIAMS 2016; JANOUSKOVEC *et al*. 2017; ROGER *et al*. 2017). This bias may reflect some of the inherent challenges in studying tRNA evolution. The short length of tRNA genes (<100 bp) limits statistical power in applying standard sequence-alignment techniques to infer gene histories and relationships. Furthermore, because of substitutions in anticodons as well as the fact that mitogenomes are famous for evolving non-standard genetic codes, some tRNAs genes are now associated with different decoding functions than the genes from which they evolved (YONA *et al*. 2013; ROGERS AND GRIFFITHS-JONES 2014; NOUTAHI *et al*. 2019).

Despite the challenges and limitations described above, a relatively clear picture about tRNA gene content in the LECA mitogenome emerges by comparing presence/absence patterns across eukaryotes (Fig. 1.2). At least 27 of the 31 tRNA genes listed as being of mitochondrial origin in Figure 1.2 are likely to have been present in LECA (shown in bold), representing a sufficient set of tRNAs capable of decoding all amino acids. The tRNA-Arg(UCG) and tRNA-Leu(CAA) genes appear to

frequently arise via duplication and anticodon mutation of other genes within those isoacceptor families (i.e., tRNAs with the same amino acid but different anticodons) (TURMEL *et al*. 2013). As such, there is not necessarily compelling evidence for these two gene being in the LECA mitogenome despite their presence in the mitogenome of numerous extant eukaryotes. The tRNA-Ser(GGA) and tRNA-Val(GAC) genes are extremely rare in mitogenomes, making it difficult to confidently infer their origins, but it is possible that these two genes were also present in LECA based on the fact that there are found in the jakobid *Andalucia godoyi*, which harbors arguably the most ancestral-like mitogenome of any eukaryote studied to date (BURGER *et al*. 2013).

This review will focus on the dynamic nature of mt-tRNA evolution within land plants, with an emphasis on angiosperm mt-tRNA populations. The complement of ancestral mt-tRNA genes appears to have remained largely stable for many hundreds of millions of years after LECA in the lineage that eventually gave rise to land plants. Even though there exists a large amount of diversity of mt-tRNA gene content and lineages with extreme tRNA gene loss in green algae (e.g., the widely used model system *Chlamydomonas reinhardtii* has only three tRNA genes in its mitogenome; (VINOGRADOVA *et al*. 2009), many green algae, including the hypothesized sister lineage to land plants (Zygnematales), have relatively complete, ancestral-like gene sets (Fig. 1.2) (TURMEL *et al*. 2013; ONE THOUSAND PLANT TRANSCRIPTOMES 2019). Similarly, Non-vascular land plants (mosses, liverworts, and hornworts) also retained a majority of these ancestral tRNA genes, but with some clade-specific losses such as a suite of tRNA-Gln, tRNA-Arg, and tRNA-Ser genes being absent in hornwort mitochondria (XUE *et al*. 2010). Within vascular plants, there has been far more heterogeneity in mt-tRNA trajectories, including lineage-specific increases in the frequency of foreign DNA insertions and loss/replacement of native mt-tRNA genes.

Intracellular transfer of tRNA genes from plastids

While there are very few, if any, instances of bryophyte or algal mitogenomes having plastid-derived sequence, DNA of plastid origin is frequently inserted into the mitogenomes of vascular plants (WANG *et*

al. 2007; GREWE *et al*. 2009; LIU *et al*. 2011; SLOAN AND WU 2014; WANG *et al*. 2018). Known as mitochondrial plastid DNAs (MTPTs), these intracellular transferred sequences are sparsely distributed in lycophytes and ferns (GREWE *et al*. 2009; GUO *et al*. 2017), but are ubiquitous in gymnosperms and angiosperms (CHAW *et al*. 2008; GUO *et al*. 2016; WANG *et al*. 2018; JACKMAN *et al*. 2019; KAN *et al*. 2020), with sequenced mitogenomes containing anywhere from 0.1 to 11.5% plastid-derived DNA (ALVERSON *et al*. 2010; GANDINI AND SANCHEZ-PUERTA 2017). In extreme examples, almost 90% of the corresponding plastid genome can be represented in the cumulative MTPT content within a mitogenome (ALVERSON *et al*. 2010; RICE *et al*. 2013; SLOAN AND WU 2014). Although there have been some instances of MTPTs contributing to the sequence evolution or regulation of mitochondrial protein and rRNA genes (NAKAZONO *et al*. 1996; HAO AND PALMER 2009; SLOAN *et al*. 2010a; WANG *et al*. 2012), the vast majority of these insertions are assumed to be nonfunctional. The identity and amount of plastid-derived sequence can vary dramatically, with MTPTs being gained and lost rapidly sometimes even varying in presence within a species (CUMMINGS *et al*. 2003; ALVERSON *et al*. 2010; SLOAN *et al*. 2012c). However, plastid-derived tRNA genes have proven to be a fascinating exception to this general trend of MTPT "insert and decay," with cases of MTPT tRNAs being retained for over hundreds of millions of years (WANG *et al*. 2007; RICHARDSON *et al*. 2013). There is direct evidence that some of these plastid-derived sequences are transcribed into mature tRNAs, functionally replacing native mt-tRNA counterparts in mitochondrial translation (MARECHAL-DROUARD *et al*. 1990; FEY *et al*. 1997; MIYATA *et al*. 1998; DUCHÊNE AND MARECHAL-DROUARD 2001).

Some of the earliest functional replacements of native mt-tRNA genes appear to predate the divergence of gymnosperms and angiosperms, as all sequenced mitogenomes from both groups lack a native mitochondrial gene for tRNA-His and tRNA-Met (Fig. 1.2). Instead, a MTPT counterpart is encoded in the mitogenome of the majority of species (RICHARDSON *et al*. 2013). Additional functional replacements of native mt-tRNAs with MTPTs occurred early in angiosperm evolution, including the loss of the native tRNA-Asn and tRNA-Trp genes, which phylogenetically corresponds to the gain of plastid-derived counterparts (RICE *et al*. 2013; RICHARDSON *et al*. 2013). Interestingly, following the initial

MTPT transfer, the tRNA-Trp gene was secondarily transferred to a linear plasmid in maize (*Zea*) (LEON *et al*. 1989). Other plastid-derived tRNA insertions are much more recent, suggesting an ongoing and dynamic process of mt-tRNA replacement in some lineages (Fig. 1.2). For example, in grasses, the expression of MTPT tRNA-Cys and tRNA-Phe genes appears to compensate for native mt-tRNA gene losses (JOYCE AND GRAY 1989). This dynamic situation complicates cases of very recent MTPT insertions, as it becomes difficult to assess whether associated tRNA genes are functional in the mitochondria or simply destined for loss. The frequent insertion of plastid-derived sequence also means that the same plastid tRNA gene can be transferred multiple times independently, complicating analyses of transfer age and homology.

Additional foreign sources of mt-tRNA genes

The propensity for plant mitogenomes to take up foreign sequence is not limited to MTPTs and intracellular gene transfer (IGT). There are also multiple instances in which other organisms have served as donors, creating an even more heterogeneous complement of mt-tRNA genes. Functional horizontal gene transfer (HGT) involving tRNA genes from bacterial sources has been documented in vascular plant mitogenomes. The first identified example was a tRNA-Cys gene in the sugar beet (*Beta vulgaris*) mitogenome that had little sequence homology to any known native plant or plastid tRNA genes (KUBO *et al*. 2000). Found to be most similar in sequence to bacterial tRNAs, this gene has now been identified in numerous angiosperm mitogenomes (KITAZAKI *et al*. 2011). The bacterial-like tRNA-Cys gene was shown to be transcribed and aminoacylated in *Beta vulgaris*, suggesting that it is functional in mitochondrial translation (KITAZAKI *et al*. 2011). In other angiosperm species, this tRNA-Cys gene is absent, but a plastid-derived tRNA-Cys copy is present, demonstrating how tRNA counterparts with radically different origins can fill the same role. More recently, five additional bacterial-like tRNA genes (tRNA-Arg, tRNA-Asp, tRNA-Lys, and two tRNA-Ser genes) were found in many vascular plant mitogenomes, including those of some lycophytes, ferns and gymnosperms (Fig. 1.2) (KNIE *et al*. 2015; GUO *et al*. 2017). These genes exhibit apparent sequence homology to counterparts in Chlamydiae, a

clade of intracellular bacteria. The conservation of some of these tRNA genes across multiple plant lineages suggests that they were gained early in vascular plant history and are likely functional. Notably, after the early replacement of the native tRNA-Asp by a chlamydial-like tRNA in vascular plants, it was then replaced yet again in angiosperms, this time by a plastid-derived counterpart (KNIE *et al*. 2015).

The extent of tRNA HGT from bacteria has not been thoroughly explored across the larger green plant lineage, but there appear to be at least some instances of such acquisitions outside of land plants. For example, the tRNA-Thr(UGU) has been identified as non-native (TURMEL *et al*. 2013), and it exhibits perfect or near-perfect sequence identity with many bacterial homologs. In addition, the tRNA-Thr(GGU) gene is another likely example of bacterial HGT. The most similar sequences to this gene outside of green plants are found in bacteria, suggesting that it may have been acquired by HGT in a common ancestor of extant green plants.

One of the largest sources of foreign DNA in some plant mitogenomes is HGT from the mitogenomes of other plant species (BERGTHORSSON *et al*. 2003; RICE *et al*. 2013; PARK *et al*. 2015; SANCHEZ-PUERTA *et al*. 2019). Plant mitochondria are complex organelles that undergo frequent fusion and homologous recombination (ARIMURA *et al*. 2004), resulting in populations of mitochondria with varied mitogenome content and structure (ARRIETA-MONTIEL AND MACKENZIE 2011). Possibly due to this tendency to fuse, many plant mitogenomes have horizontally acquired mitochondrial sequence, including tRNA genes, from other plant species. An extreme example is the angiosperm *Amborella trichopoda,* which has an unusually large mitogenome, containing foreign mitochondrial sequence from other angiosperms, mosses and green algae. Remarkably, some of these insertions appear to represent entire mitogenomes from moss and algal donors (RICE *et al*. 2013; TAYLOR *et al*. 2015). In a few cases, HGT from other green plant mitogenomes was inferred to functionally replace the corresponding ancestral mt-tRNA gene in *Amborella trichopoda*, but expression of the foreign genes has not been directly tested. This form of horizontal transfer also serves as a mechanism by which plastid-derived sequences in mitogenomes can be moved from species to species (GANDINI AND SANCHEZ-PUERTA 2017).

More recently, the mitogenomes of orchids (one of the largest families of angiosperms) were shown to contain HGTs of fungal mitochondrial sequence, including three tRNA genes (SINN AND BARRETT 2019). There is no evidence that the fungal-derived tRNA genes function in mitochondrial translation within orchids. Nevertheless, it is intriguing that this cluster of tRNA genes was anciently transferred in the ancestor of orchids and is still retained in many extant lineages and even appears to have been replaced by gene conversion during a subsequent fungal transfer event in one major lineage.

In light of the numerous sources of foreign tRNA genes in plant mitogenomes, the lack of evidence for acquisition of tRNA genes from the nuclear genome is conspicuous. The nucleus is another common source of DNA insertions into plant mitogenomes (NOTSU *et al*. 2002; GOREMYKIN *et al*. 2012), but to our knowledge, there has never been a reported instance of a mitogenome with a functional nuclear-derived tRNA gene. This contrast may reflect simple probabilistic factors such as the low gene density of plant nuclear genomes, which makes it less likely that any random nuclear insert would contain a tRNA gene. In addition, unlike in plastids, which share some of the same transcriptional machinery with mitochondria (KÜHN *et al*. 2009), expression of nuclear genes is driven by entirely distinct regulatory systems (HUMMEL *et al*. 2019) such that nuclear tRNA genes may be more likely to be transcriptionally inactive ("dead on arrival") when inserted into the mitogenome.

The frequent insertion of foreign tRNA genes complicates the "head count" of functional mt-tRNA genes. For example, the large mitogenome of *Amborella trichopoda* contains in excess of 150 individual tRNA gene copies, but the majority of these sequences originate from the insertion of foreign DNA, including mitogenome-scale HGT from other plant and algal species, as well as MTPTs that cumulatively represent almost an entire plastid genome. It is likely that most of these foreign inserts are non-functional. Similarly, 11.5% of the mitogenome of *Curcurbita pepo* (zucchini) is derived from MTPTs, including 16 of its 26 unique tRNA genes (ALVERSON *et al*. 2010), demonstrating how gene insertion events can greatly inflate mt-tRNA gene counts. For comparison, a more typical angiosperm mitogenome from the model system *Arabidopsis thaliana* has 11 unique native tRNA genes and six plastid-derived genes. Previous northern analysis only detected the expression of four of the six plastid-

derived sequences, with no detection of tRNA-Met and tRNA-Trp (DUCHÊNE AND MARECHAL-DROUARD 2001). However, recent high-throughput tRNA sequencing (tRNA-seq) analysis in *Arabidopsis* did find the plastid-derived tRNA-Met gene to be expressed and post transcriptionally modified with a CCA tail (WARREN *et al*. 2019), implying that five (but probably not all six) of the MTPT tRNA genes in *Arabidopsis thaliana* are functional. Thus, commonly reported presence/absence matrices like the one shown in in Figure 1.2 are unlikely to fully reflect functionality or homology because of the fluid nature of mt-tRNA gene sets. And although some work has been done to test for the transcription and aminoacylation for tRNA genes derived from MTPTs and HGTs (JOYCE AND GRAY 1989; MARCHFELDER *et al*. 1990; DUCHÊNE AND MARECHAL-DROUARD 2001; KITAZAKI *et al*. 2011), the majority of functional inference is based on mitogenome sequence alone. Researchers are often limited to using MTPT length and sequence divergence as "circumstantial" evidence for tRNA functionality in more recent transfers (ALVERSON *et al*. 2010; RICHARDSON *et al*. 2013). Thus, despite tRNA genes being subject to extensive IGT and HGT, there is much work to do to explore the full scope of these phenomena and the mechanisms by which foreign tRNA genes can serve as interchangeable machinery in mitochondrial translation.

tRNA import from the cytosol complements missing mt-tRNA genes.

Despite their many native and foreign tRNA genes, angiosperm mitogenomes still lack a minimally sufficient set of tRNA genes needed to carry out translation (MURCHA *et al*. 2016). Instead, like other eukaryotes, plants depend on the import of expressed tRNAs from yet another source: the nuclear genome. Once thought to be uncommon, the import of nuclear-encoded tRNAs from the cytosol into the mitochondrial matrix has since been demonstrated to be the rule rather than the exception (SALINAS-GIEGÉ *et al*. 2015). All imported tRNAs are likely of eukaryotic origin, as we are not aware of a documented case of a mt-tRNA gene being functionally transferred to the nuclear genome and targeted back to the mitochondria (SCHNEIDER 2011). This is in sharp contrast to the numerous protein genes that were functionally transferred to the nuclear genome in the evolutionary history of mitochondria (ADAMS

AND PALMER 2003; TIMMIS *et al*. 2004), a process that is still highly active in plants (ADAMS *et al*. 2002). The degree of cytosolic tRNA import varies across eukaryotes but appears to be minimal in many animal and fungal lineages as their mitogenomes have a mt-tRNA gene set capable of decoding all amino acids (GROSJEAN AND WESTHOF 2016). However, even in taxa with a sufficient set of mt-tRNA genes, import of cytosolic tRNAs has been observed (RINEHART *et al*. 2005; RUBIO *et al*. 2008; MERCER *et al*. 2011). The function of imported tRNAs in such systems is not always clear, but in some cases these redundant tRNAs are important to maintain effective mitochondrial translation in stress conditions (MARTIN *et al*. 1979; KAMENSKI *et al*. 2007). Overall, however, the import of cytosolic tRNAs should generally account for a small percentage of the mitochondrial tRNA pool in these organisms (MERCER *et al*. 2011), and the identity of imported tRNAs has probably been stable for very long periods of time in some lineages (ALFONZO AND SÖLL 2009).

The apparent stasis observed in some eukaryotic lineages does not apply to plants, where the number of imported tRNAs has been shown to vary widely among species (KUMAR *et al*. 1996). Under the assumption that decoding deficiencies in mt-tRNAs are compensated for by cytosolic import of nuclear-encoded tRNAs, it is clear that multiple plant linages have evolved extensive tRNA import dependencies to maintain mitochondrial translation. The most extreme case in land plants is the spikemoss *Selaginella*, a lycophyte that lacks any mt-tRNA genes and presumably relies entirely on cytosolic import (HECHT *et al*. 2011). These losses appear to be lineage-specific because all other sequenced lycophytes, ferns, gymnosperms and angiosperms encode at least some tRNA genes in their mitogenome. The gymnosperm *Welwitschia* also likely depends more heavily on import of cytosolic tRNAs than most plants because its mitogenome retains only eight tRNA genes (Fig. 1.2) (GUO *et al*. 2016).

The large number of sequenced mitogenomes available for angiosperms have revealed incredibly diverse trajectories for history of loss amongst the inherited set of ancestral mt-tRNA genes (Fig. 1.3). Angiosperms are expected to import an average of ~30-50% of their mitochondrial tRNA pool (ESSER *et al*. 2006). However, based on observed patterns of tRNA gene loss from mitogenomes, this number may

be significantly higher in some species and exhibit dramatic variation even between close relatives (Fig. 1.3). The sequencing and assembly of multiple mitogenomes from the genus *Silene* (Caryophyllaceae) revealed a general reduction in mt-tRNA gene content within the genus and a clear history of recent and ongoing loss, with species having anywhere from nine down to only two tRNA genes remaining in the mitogenome (SLOAN *et al*. 2012a). Similarly extreme levels of reduction in mt-tRNA gene content have been more recently observed in species of the parasitic plant mistletoe (*Viscum*, Santalaceae), with *Viscum album* and *Viscum scurruloideum* retaining only five and three distinct mt-tRNA genes, respectively (PETERSEN *et al*. 2015; SKIPPINGTON *et al*. 2015).

Such cases of ongoing and rapid mt-tRNA gene loss and presumed replacement by cytosolic import are noteworthy because they point to import specificity evolving remarkably fast. There are many longstanding and still unresolved questions on tRNA import mechanisms (SALINAS *et al*. 2008), but what little is known about tRNA import in plants suggests that it is largely "complementary" (i.e., compensating for tRNA genes absent from the mitogenome rather than providing redundant tRNAs) and, at least in algae, "static" (i.e., tRNA pools were unaffected by manipulated mitogenome codon usage) (SALINAS *et al*. 2008; VINOGRADOVA *et al*. 2009; SALINAS *et al*. 2012). Both of these observations provide evidence that tRNA import is a coevolved process based on historical tRNA requirements and mitogenome content. As such, instances of rapid mt-tRNA loss and replacement should provide a valuable opportunity to investigate this coevolutionary process "in the act".

Characterizing the evolutionary mechanisms necessary for the *de novo* import of cytosolic tRNA has been hindered by incomplete investigations of mitochondrial tRNA pools, as there has never been an exhaustive study of which tRNAs are imported in plant mitochondria (SALINAS-GIEGÉ *et al*. 2015). One reason for the lack of data on plant tRNA populations at subcellular levels is related to technological limitations in applying conventional sequencing methods to tRNA transcripts. Thus, investigations have largely been limited to targeted, hybridization methods. However, recent advances in tRNA-seq are opening the door for fine-scale characterization of tRNA populations that may accelerate investigations of

tRNA import evolution (See Box 1: Challenges and progress in high-throughput sequencing of tRNA populations).

Box 1: Challenges and progress in high-throughput sequencing of tRNA populations

It is ironic that in the half-century since a *Saccharomyces cerevisiae* tRNA became the very first complete nucleic acid to be sequenced (HOLLEY *et al*. 1965), tRNAs have emerged as arguably the single most difficult class of molecules to analyze with modern sequencing technologies. One of the largest barriers to tRNA-seq is the extensive post-transcriptional modification of tRNA molecules (WILUSZ 2015b), which carry over 100 different identified modifications (MOTORIN AND HELM 2010). Reverse transcription (RT) into cDNA is required for almost all current RNA sequencing methods, but many base modifications can cause stalling or disassociation of the RT enzyme due to interference with base pairing or steric hindrance (MOTORIN *et al*. 2007). As a result, tRNAs often yield truncated sequences or cDNA molecules that are entirely unsequenceable due to a lack of RT readthrough to necessary 5′ adapters. The ligation of sequencing adapters to the 5′- and 3′-end of tRNAs has also proven to be problematic. The compact secondary and tertiary structure of mature tRNAs makes the base-paired termini less accessible for ligation reactions (LAMA *et al*. 2019). Because RT is normally primed off of the ligated 3′-adapter, molecules lacking these adapters are simply never sequenced.

Excitingly, the last five years has seen the introduction of multiple advances to circumvent these difficulties (see Figure 1.4). A breakthrough in tRNA-seq came from the utilization of the *Escherichia coli* dealkylating enzyme AlkB. Normally found in bacterial cells as a repair enzyme that removes aberrant DNA methylation (MISHINA AND HE 2006), wild type AlkB and engineered versions thereof were found to remove many RT-inhibiting modifications present on tRNAs (COZEN *et al*. 2015; ZHENG *et al*. 2015). Treatment with AlkB prior to RT has resulted in detecting higher abundance and diversity of tRNA reads in numerous recent studies (CLARK *et al*. 2016; TORRES *et al*. 2019; WARREN *et al*. 2019).

New methods have also been developed to increase the efficiency of adapter ligation and specifically target mature tRNA molecules within a heterogeneous RNA population. Y-shaped Adapter-ligated Mature tRNA sequencing, or YAMAT-seq, takes advantage of the unpaired discriminator base and CCA-tail found on all mature tRNAs by using adapters that complement this overhang (SHIGEMATSU *et al*. 2017). Other methods have tried to simultaneously circumvent inhibitory modifications and secondary-structure issues by partial hydrolysis of tRNAs (Hydro-tRNASeq; (GOGAKOS *et al*. 2017). The fragmentation of tRNAs appears to remove some modifications and free 3′- and 5′-ends for adapter ligation. This fragmentation, however, further complicates the mapping for what are already relatively short sequences, making it less suitable for certain biological questions that require distinguishing among genes from families of closely related sequences.

These protocols are very promising for tRNA biology; however, challenges remain for tRNA-seq. Namely, AlkB does not remove all modifications known to inhibit RT (COZEN *et al*. 2015; ZHENG *et al*. 2015), and biases in adapter ligation persist (LAMA *et al*. 2019), both of which lead to artefactual variation in tRNA-seq read abundance that may confound efforts to measure biological differences in tRNA expression levels. Nevertheless, the ability to generate full-length sequences from the tRNA populations of different cellular compartments represents a huge step forward in understanding the evolution of tRNA metabolism and key progress in investigating one of life's oldest and most fascinating molecules.

Coevolution and cytonuclear interactions in light of mt-tRNA gene loss and replacement

tRNAs are not standalone components in protein synthesis. Instead, they require large enzymatic networks for processing, editing, maturation and function (Table 1.1) (PHIZICKY AND HOPPER 2010). Like nuclear-encoded tRNAs, mt-tRNAs undergo 5′- and 3′-end processing from longer primary transcripts as well as a suite of editing, modification, and functional maturation steps before fulfilling their role of delivering amino acids to the ribosome (SALINAS-GIEGÉ *et al*. 2015). With the exception of

some components of the mitochondrial ribosome, all tRNA-interacting enzymes are encoded by the nuclear genome and imported into mitochondria. The incredibly dynamic nature of plant mt-tRNA gene loss and replacement raises questions about how the functional swapping of bacterial-like mt-tRNAs with anciently divergent counterparts perturbs the coevolved networks of interactions that are essential for mitochondrial translation.

The network of tRNA-interacting enzymes

The enzymes necessary for tRNA processing, maturation and function can be classified based on whether they function exclusively within one compartment (nucleus/cytosol, mitochondria, or plastids) or have activity in multiple compartments (CANINO *et al*. 2009; ROSSMANITH 2012; HOPPER AND NOSTRAMO 2019). In plants, the tRNA nucleotidyltransferase enzyme that adds the tail of CCA nucleotides found on the 3′ end of all mature tRNAs (CCAse) processes all cellular tRNAs regardless of where they are encoded or localized (VON BRAUN *et al*. 2007).

Other tRNA-interacting enzymes can be organelle-specific. A partial division between nuclear and mitochondrial tRNA metabolic networks is apparent among aaRSs. Each aaRS interacts with cognate tRNAs based on amino-acid decoding (i.e., there would typically be one aaRS for each of the 20 different amino acids) and uses nucleotide identities at key positions for substrate recognition (GIEGE *et al*. 1998). By and large, the bacterial-like tRNAs found in mitogenomes share little sequence similarity to eukaryotic tRNAs and are very poor substrates for eukaryotic aaRSs (SALINAS-GIEGÉ *et al*. 2015). Thus, eukaryotes must encode a largely separate set of aaRSs for mt-tRNA aminoacylation (DUCHÊNE *et al*. 2009). Naively, we might expect that cytosolic and mitochondrial aaRSs would reflect the archaeal and alphaproteobacterial legacy of the respective host and endosymbiont lineages that gave rise to extant eukaryotes. However, aaRS evolutionary history appears far more complex and likely involves substantial HGT from other sources (DOOLITTLE AND HANDY 1998; BRINDEFALK *et al*. 2007). Despite these complex origins, early eukaryotic evolution does appear to have largely established separate sets of aaRSs

that can be distinguished as either mitochondrial or cytosolic in function, although there are cases of aaRSs being shared between both compartments (TOLKUNOVA *et al*. 2000; DUCHÊNE *et al*. 2001; CHIEN *et al*. 2014).

The situation in plants is further complicated by the need to deliver aaRSs for a third translation system in the plastids. However, due to the widespread dual-functionality of aaRSs in both mitochondria and plastids, plants have fewer than the 60 aaRSs that would be expected if there were no overlap between the three translational systems. Possibly due to the bacterial-like nature of both plastid and mitochondrial tRNAs, many prokaryotic-like aaRSs are effective enzymatic partners for both classes of organellar tRNAs (DUCHÊNE *et al*. 2005a). In *Arabidopsis thaliana*, there are a total of 23 aaRSs thought to function in the mitochondria (Table 1.1), corresponding to 19 amino acids (there is no GlnRS in mitochondria because of the indirect pathway used to aminoacylate tRNA-Gln; see below). None of these 23 aaRSs seem to function exclusively in the mitochondria. A total of 17 *Arabidopsis* aaRSs have been identified as dual-targeted to the mitochondria and plastids but do not have cytosolic localization (Table 1.1). Only two of these enzymes (PheRS and SerRS) were assigned as the typical mitochondrial type inherited from the eukaryotic ancestor, implying an extensive history of horizonal gene transfer as most of the dual-targeted mitochondrial/plastid aaRSs were identified as having more recent bacterial origin. These include many aaRSs that can be traced to cyanobacteria and the presumed progenitor of plastids but also some that appear related to other bacterial lineages, implying additional sources of HGT and/or reflecting the limitations of phylogenetic resolution (BRINDEFALK *et al*. 2007; BRANDAO AND SILVA-FILHO 2011) concluded that all but two of the 17 dual-organellar aaRSs were of bacterial or mitochondrial origin, with HisRS and ProRS being the only exceptions. However, other phylogenetic analyses have placed this plant ProRS closest to bacterial enzymes, and the HisRS has exhibited ambiguous phylogenetic signal with association to a mix of both bacterial and archaeal enzymes (BRINDEFALK *et al*. 2007; BRANDAO AND SILVA-FILHO 2011). Therefore, it is possible that all 17 of the dual-organellar aaRSs in *Arabidopsis* have either bacterial or mitochondrial origin.

There are also six *Arabidopsis* aaRSs identified as having dual targeting to the mitochondria and cytosol, including two (AlaRS and ArgRS) that are also targeted to the plastids and thus appear to function in all three translation systems (Table 1.1). In the case of this ArgRS, there is only direct evidence of targeting to the plastid, but function in the mitochondria and cytosol is inferred based on the fact that disruption of the only other identified ArgRS gene in *Arabidopsis* does not affect viability (BERG *et al*. 2005; DUCHÊNE *et al*. 2009). In contrast to the overwhelming bacterial/mitochondrial origin of the dual-organellar aaRSs described above, all six of the aaRSs with localization in both the mitochondria and cytosol appear to be ancestral cytosolic enzymes that are shared with other eukaryotes and have gained mitochondrial targeting. In four of these six cases (AlaRS, GlyRS, ThrRS, and ValRS), the presence of a cytosolic-like aaRS in the mitochondria creates apparent redundancy with the presence of a bacterial-like dual-organellar aaRS for the same amino acid. However, it cannot be assumed that both aaRSs are involved in aminoacylation in these examples of redundancy, as investigations of the cytosolic-like GlyRS and ValRS enzymes suggest that they are inactive within mitochondria and that only their bacterial-like counterparts function in mitochondrial aminoacylation (DUCHÊNE *et al*. 2001; DUCHÊNE *et al*. 2009). It is not clear why these enzymes might be targeted to the mitochondria yet not function in aminoacylation, but it has been hypothesized that it might reflect a role in chaperoned tRNA import from the cytosol, and aaRSs are known to take on alternative functions within the cell (GUO AND SCHIMMEL 2013). Regardless, these examples of redundancy point to potential pathways for evolutionary transitions and loss/replacement of ancestral mitochondrial enzymes.

Other enzymes involved in organelle protein synthesis are specific to bacterial-like translational machinery and have no equivalent functioning in the nucleus or cytosol. In most bacteria, archaea, and plastids, as well as plant mitochondria, the aminoacylation of tRNA-Gln is achieved through an indirect pathway, meaning it is not aminoacylated by a dedicated GlnRS. Instead, the mt-tRNA-Gln is first "incorrectly" charged with the amino acid Glu by a nondiscriminating GluRS and subsequently converted to Gln by a tRNA-dependent amidotransferase complex called GatCAB. In plants, this complex is dual targeted to mitochondria and plastids, and the genes that encode its protein subunits are cyanobacterial in

origin, suggesting that they were acquired as part of plastid endosymbiosis (Pujol et al., 2008). Protein synthesis is also initiated differently in bacteria and organelles through the use of a formylmethionyl-tRNA, also known as tRNA-fMet (COFFIN AND COSSINS 1986; TAKEUCHI *et al*. 2001). The mt-tRNA-Met must first have a formyl group added to the amino group by an enzyme known as methionyl-tRNA formyltransferase (MTF) before it can be used to initiate translation (IBBA AND SÖLL 2004). Translation in plant mitochondria and plastids also relies on a bacterial-like modification of native tRNA-Ile, in which the cytidine in the CAU anticodon (which would typically decode AUG Met codons) is modified to lysidine by tRNA-Ile lysidine synthetase (TiLS), allowing it to decode AUA Ile codons (WEBER *et al*. 1990; SUZUKI AND MIYAUCHI 2010). The MTF and TiLS enzymes have not yet been described in plants, but obvious candidates exist based on sequence similarity (Table 1.1). Additional modifications are increasingly being recognized to play fundamental roles in mitochondrial translation, each with their own respective enzymatic machinery that must be imported into the mitochondrial matrix (PARIS AND ALFONZO 2018).

Still other plant mt-tRNA metabolism machinery is seemingly derived in eukaryotes and not inherited from either the archaeal-like host or the alphaproteobacterial progenitor of mitochondria, including the RNase P enzyme utilized for the cleavage of 5′-leader sequences of mitochondrial-encoded tRNA precursors. It was once thought that all domains of life utilized an RNase P enzyme with a ribozyme catalytic domain (ALTMAN 2007); however, a more complex picture emerged with the discovery that the enzyme responsible for RNase P activity in human and plant organelles is composed entirely of protein, now known as the proteinaceous RNase P or PRORP (HOLZMANN *et al*. 2008; GOBERT *et al*. 2010; GUTMANN *et al*. 2012). Based on its distribution across diverse eukaryotic lineages, it appears that PRORP was likely present alongside the conventional RNase P ribozyme in LECA (LECHNER *et al*. 2015). In addition to the originally described function in plant and animal mitochondria, some PRORPs have also been shown to have nuclear function, in some cases even serving as the only source of RNase P activity in the cell (GUTMANN *et al*. 2012; BONNARD *et al*. 2016). The 3′-end

processing of tRNAs relies on another type of enzyme known as tRNase Z. In *Arabidopsis thaliana*, there are four different tRNase Z genes with various targeting patterns (Canino et al., 2009), including two with mitochondrial localization (Table 1.1).

Overall, the many peculiarities of mt-tRNA metabolism highlight that, while translation is fundamentally similar across cellular compartments, organelle and cytosolic protein synthesis often require different, coevolved partners.

The coevolution of translational machinery across life and genomes

The genomic compartmentalization of mt-tRNAs and their nuclear-encoded enzymatic partners results in translational machinery being divided between two radically different genomes that differ in structure, mutation rates, inheritance and expression (SLOAN *et al*. 2018). There is now clear evidence that this division of translational machinery between multiple genomes in eukaryotes has resulted in both mitonuclear epistasis and coevolution in mitochondrial tRNA metabolism. Some of the best described examples of these interactions involve aaRS and their mt-tRNA substrates, likely because aaRSs and their tRNA partners are under strong selection to maintain faithful decoding of transcripts into proteins (SALAZAR *et al*. 2003; SALINAS-GIEGÉ *et al*. 2015). In the fruit fly *Drosophila melanogaster,* an amino acid substitution in a mitochondrial-targeted TyrRS was found to be incompatible with a single-nucleotide mt-tRNA-Tyr polymorphism in a closely related species, *Drosophila simulans* (MEIKLEJOHN *et al*. 2013). In separate genetic backgrounds, the nuclear-encoded aaRS amino acid substitution and the mitochondrial-encoded tRNA nucleotide substitution caused no fitness effects but, when present in the same organism, resulted in a significant decrease in development and fecundity. This epistatic effect demonstrates how single point mutations in tRNA networks can have large fitness effects on mt-tRNA metabolism. Similarly, mt-tRNA mutations disproportionally contribute to disease-causing phenotypes, with many of those pathologies resulting from perturbed aaRS interactions (PEARCE *et al*. 2013). However, other studies have found this "coevolutionary crosstalk" between tRNAs and enzymatic partners to be more nuanced, inferring that a large portion of mt-aaRS amino acid substitutions at deep

timescales are not driven by coevolutionary responses to mt-tRNA mutations (PETT AND LAVROV 2015; ADRION *et al*. 2016). Aside from aaRSs, other tRNA-interacting enzymes have adapted to mt-tRNA evolution, such as a modified EF-Tu enzyme which recognizes the unusual "armless" mt-tRNAs in nematodes (ARITA *et al*. 2006).

While the above examples demonstrate possible coevolutionary responses to changes in mt-tRNA sequences, other work has focused on the effects of outright mt-tRNA gene loss and functional replacement. Certain lineages of non-bilaterian animals (sponges, placozoans, cnidarians, and ctenophores) have experienced extreme mt-tRNA loss. This loss of mt-tRNA genes appears to trigger the loss of nuclear-encoded enzymes formerly associated with the mt-tRNAs (KOHN *et al*. 2012; PETT AND LAVROV 2015). The absence of mt-tRNA genes generally corresponds to the parallel loss of the associated aaRS in the respective nuclear genomes of these animals (HAEN *et al*. 2010b). In species completely lacking mt-tRNAs genes, other nuclear tRNA-processing enzymes have been lost as well, including the genes encoding the GatCAB complex and the mitochondrial PRORP (PETT AND LAVROV 2015). The kinetoplastids present another case of extreme aaRS reduction precipitated by mt-tRNA gene loss in eukaryotes (TAN *et al*. 2002). The mitogenome of the kinetoplastid *Trypanosoma brucei* does not encode any mt-tRNAs (SCHNEIDER 2001), which largely removes the requirement for two separate sets of aaRS for mitochondrial and cytosolic tRNA aminoacylation. In accordance with this, *Trypanosoma brucei* only encodes 23 aaRS genes (CHARRIERE *et al*. 2006). However, surprisingly, even in this extreme example of merging cytosolic and mitochondrial tRNA metabolisms, separate eukaryotic aaRSs exist to aminoacylate the mitochondrial-imported tRNAs for Asp, Lys, and Trp, even though all tRNAs are encoded in the nuclear genome (CHARRIERE *et al*. 2006; CHARRIERE *et al*. 2009). The use of different aaRSs may be due to differences in post-transcriptional modification between tRNAs functioning in the cytosol vs. the mitochondria, which may make the two classes of tRNAs too dissimilar to be substrates for a single aaRS (ALFONZO *et al*. 1999; CHARRIERE *et al*. 2009).

These studies point to broad effects of mt-tRNA gene loss, but the deep evolutionary timescales at which many of the gene loss/replacement events occurred can make it difficult to identify the

evolutionary mechanisms and intermediate states that are required for the functional replacement of mt-tRNAs. These are significant questions as the functional replacement of native mt-tRNA genes with imported cytosolic counterparts represents swapping molecular parts after billions of years of evolutionary divergence.

Using plants to elucidate the evolutionary dynamics of mt-tRNA functional replacement

The ongoing replacement of mt-tRNA genes in plant lineages via the import of nuclear-encoded cytosolic counterparts poses questions about which enzymatic partners are processing, modifying, and charging these tRNAs. Multiple scenarios could be envisioned, one being that that aaRSs and other nuclear-encoded tRNA-processing enzymes that were historically present only in the cytosol or nucleus have also gained import into the mitochondrial matrix and maintain their same interactions with newly imported cytosolic tRNA substrates. Mt-tRNA gene loss appears to precipitate this outcome at deep evolutionary timescales in groups with very few or no mt-tRNA genes remaining, such as the aforementioned Cnidaria and Ctenophora (PETT AND LAVROV 2015). These species largely lack mitochondrial-like aaRSs, which are assumed to be replaced by cytosolic versions of the enzymes. However, the import of nuclear tRNAs and cytosolic enzymes would seem to have required more or less simultaneous evolution because the import of either one without the other would have no functional effect. This "chicken and egg" problem of functional mt-tRNA replacement has been raised numerous times in mt-tRNA biology (SMALL *et al*. 1999; SCHNEIDER 2011). Characterizing the subcellular localization of tRNAs and interacting enzymes in plant lineages with ongoing mt-tRNA loss represents a valuable opportunity to test hypotheses on the intermediate states and order of events necessary for functional mt-tRNA replacements, such as possible periods of redundant tRNA/enzyme import and the evolution of mitochondrial targeting peptides.

Conversely, a second evolutionary scenario could have arisen where mitochondrial aaRSs and other tRNA-interacting enzymes have been retained, adapting to recognize and process nuclear-encoded tRNAs. There is some evidence that this process has played a role in plant mt-tRNA metabolism as the organellar, bacterial-like GlyRS has gained substrate recognition for a cytosolically imported tRNA-Gly.

Both a mitochondrial-encoded tRNA-Gly and an imported nuclear-encoded tRNA-Gly are necessary for mitochondrial translation in eudicot angiosperms because the mitochondrial-encoded tRNA-Gly(GCC) is unable to decode all four GGN Gly codons (BRUBACHER-KAUFFMANN *et al*. 1999). It was demonstrated that organellar GlyRSs from multiple angiosperms are able to aminoacylate both tRNAs (DUCHÊNE *et al*. 2001; SALINAS *et al*. 2005). This was surprising because prokaryotic and eukaryotic GlyRSs use a specific nucleotide identity at position 73 for activity, which is different between the two (nuclear and mitochondrial) tRNA-Gly genes (GIEGE *et al*. 1998). This ability for the organellar, bacterial-like GlyRS to recognize cytosolic tRNA-Gly suggests the evolution of expanded substrate recognition based on comparisons with the GlyRS from *Escherichia coli*, which has limited activity on cytosolic tRNA substrates (NAMEKI *et al*. 1997; DUCHÊNE *et al*. 2001).

More generally, data from *Arabidopsis thaliana* on tRNA gene content in the mitogenome and subcellular localization of aaRSs provide multiple examples supporting both of the evolutionary scenarios described above (DUCHÊNE *et al*. 2009). The absence of tRNA-Arg and tRNA-Leu genes from the *Arabidopsis* mitogenome suggests cytosolic import of these tRNAs (which has been confirmed in other angiosperms; (SALINAS-GIEGÉ *et al*. 2015), and the only mitochondrial-localized aaRSs for these tRNAs are also cytosolic in origin (Table 1.1). Therefore, these appear to be cases of matched retargeting of cytosolic tRNAs and interacting enzymes, as described in the first scenario. In contrast, for other examples of import of cytosolic tRNAs into *Arabidopsis* mitochondria that have been inferred (tRNA-Ile) or demonstrated (tRNA-Phe and tRNA-Trp; (DUCHÊNE AND MARECHAL-DROUARD 2001), the only corresponding aaRS localized to the mitochondria is bacterial-like, supporting the second evolutionary scenario, where existing mitochondrial machinery has adapted to function on cytosolic tRNAs. The cytosolic import of tRNA-Val and the aforementioned case of tRNA-Gly also fall in this category because only the bacterial-like aaRSs appear to be active in aminoacylation within the mitochondria even though a cytosolic-like aaRS is also localized to the mitochondria (DUCHÊNE *et al*. 2009). The import of cytosolic tRNA-Ala and tRNA-Thr is less clear because bacterial-like and cytosolic-like aaRSs are both targeted to the mitochondria, and their relative contributions to aminoacylation are not known.

Although the adaptation of bacterial-like aaRSs to charge cytosolic tRNAs appears to be a viable evolutionary route in some cases, the dual-targeted nature of mitochondrial and plastid aaRSs could constrain enzymatic evolution, as very often the same enzyme must recognize both plastid and mitochondrial-encoded tRNAs (Table 1.1). It has yet to be determined if amino acid substitutions in organellar aaRSs and other tRNA-interacting proteins would have pleiotropic effects on plastid tRNA metabolism, possibility necessitating gene duplication and neofunctionalization events.

The two scenarios proposed above are not mutually exclusive, and it will be interesting to determine if certain evolutionary routes are taken repeatedly for lineages with independent mt-tRNA gene losses, or if tRNA-interacting networks are perturbed differently depending on the specific mt-tRNA gene. In addition, these two proposed coevolutionary responses to mt-tRNA loss do not represent the only possibilities. One fascinating example of an alternative mechanism to achieve functional mt-tRNA gene replacement comes from the apicomplexan *Toxoplasma gondii*, which depends entirely on cytosolic tRNA import for mitochondrial translation (ESSEIVA *et al*. 2004). What makes its mt-tRNA metabolism so exceptional is that no aaRSs have been found to be localized to the mitochondria in this species, meaning that tRNAs are imported already aminoacylated. This mechanism removes the evolutionary constraint of *de novo* aaRS import but raises questions about the recycling of tRNAs in these organisms. Because of the incredibly long half-life of tRNA molecules (up to multiple days; (HOPPER 2013), a single tRNA can be assumed to be aminoacylated numerous times throughout its lifetime. Are the imported tRNAs in *Toxoplasma gondii* repeatedly transported back and forth across the mitochondrial membranes to be reloaded with amino acids? Alternatively, are imported tRNAs effectively "single-use" molecules that are degraded inside the mitochondrial matrix after unloading their amino acid? Both of these models seem energetically inefficient, but they represent the only obvious possibilities if all imported tRNAs are indeed aminoacylated in the cytosol. This apicomplexan aminoacylation mechanism is less likely to be ocurring in plants, as data on subcellular localization of aaRSs shows that plant mitochondria still have extensive aaRS import (DUCHÊNE *et al*. 2005b). But it is worth noting that other types of enzymatic interactions that need only occur once in the lifetime of a tRNA (e.g., end-processing, CCA-tailing, base

modifications, etc.) may not suffer any inefficiency if the activity remains in the nucleus/cytosol followed by mitochondrial import of mature tRNAs. Regardless, the aminoacylation system in apicomplexans illustrates the diversity of mechanisms by which eukaryotic life maintains translation in the face of the mosaic origins of mitochondrial molecular biology.

Outlook: Catching functional mt-tRNA replacement in the act

It is a striking contrast that interactions between tRNAs and their enzyme partners are so coevolved that single nucleotide/amino-acid polymorphisms found in *Drosophila* species are sufficient to create highly deleterious incompatibilities (MEIKLEJOHN *et al*. 2013) but, at the same time, there is such evolutionary interchangeability among anciently divergent tRNAs in plants and other lineages (Fig. 1.2). Reconciling this apparent paradox represents a key challenge for tRNA biology. The sequencing of mitogenomes from angiosperms such as *Silene* and *Viscum* (SLOAN *et al*. 2012a; PETERSEN *et al*. 2015; SKIPPINGTON *et al*. 2015) has revealed closely related species with very different complements of mt-tRNA genes, reflecting histories of gene loss that are both recent and extreme. These systems appear to have an even more active state of mt-tRNA gene replacement than what is already an incredibly dynamic plant mt-RNA metabolism. As such, they may offer new opportunities to describe the molecular coevolution and intermediate steps necessary for functional mt-tRNA replacement while it is actively occurring. Investigations into protein-RNA interactions, cellular trafficking, and genome evolution involved in mt-tRNA gene loss have the potential shed light on some of the most fundamental questions in eukaryotic cell evolution, with the aid of advances in multiple fields of cell and molecular biology, including breakthroughs in tRNA-seq. And tracing the steps needed to replace such anciently divergence translational systems could have ripple effects in numerous fields including human health, synthetic biology, speciation, and the origin of eukaryotic life.

Funding and support

This work was supported by an NSF GAUSSI graduate research fellowship (DGE-1450032), NSF award IOS-1829176, and Colorado State University.

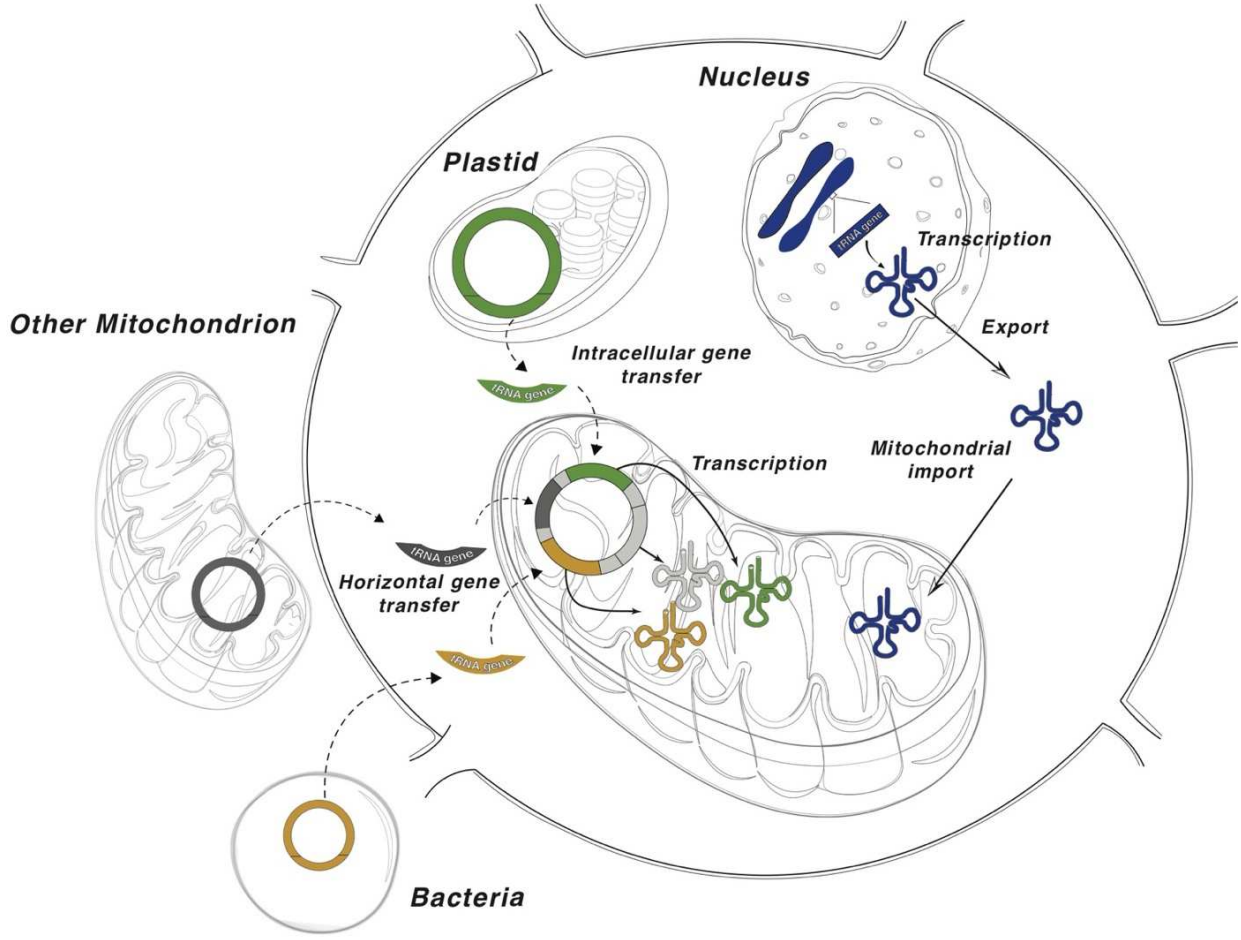


Figure 1.1| The heterogeneous nature of the plant mitochondrial tRNA populations. The pool of tRNAs in the mitochondrial matrix of plants can contain a mix of tRNAs transcribed from native mitochondrial genes (light gray), intracellularly transferred genes from plastids (green), and horizontally transferred genes from other species (gold and dark gray), as well tRNAs expressed from nuclear genes and imported from the cytosol (blue). The light gray mt-tRNAs represent the native mitochondrial tRNAs.

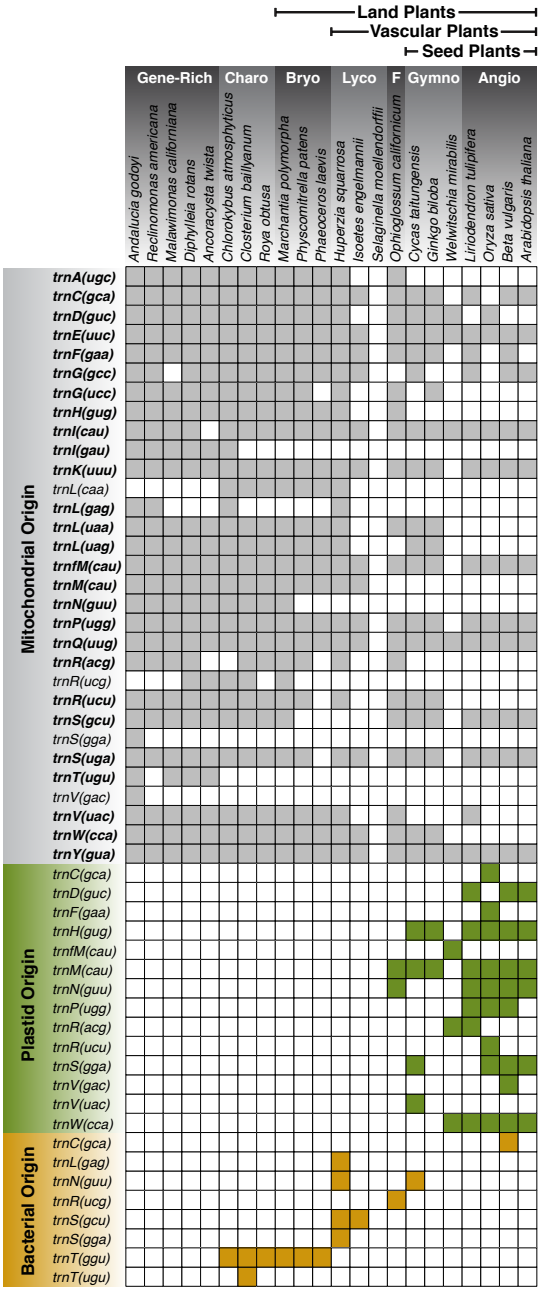


Figure 1.2| Summary of tRNA gene content in mitogenomes from a sample of diverse streptophytic green algae and land plants. Filled squares indicate the presence of an intact gene sequence, with genes named based on a single-letter amino-acid abbreviation and with anticodon indicated parenthetically. For example, *trnA(ugc)* corresponds to tRNA-Ala with a UGC anticodon. Formyl-methionine is abbreviated fM. Genes shown in bold text are inferred to have been present in LECA (see main text). Disrupted or incomplete sequences inferred to be pseudogenes are not included. In some cases, gene classification is based on observed or inferred C-to-U RNA editing in the anticodon sequence (Grewe et al. 2009, Guo et al. 2017). Taxon abbreviations are defined as follows. Angio: angiosperms; Gymno: gymnosperms; F: Ferns; Lyco: lycophytes; Bryo: bryophytes; Charo: charophytes (i.e., streptophytic algae); Gene-Rich: other eukaryotes from lineages with unusually high retention of ancestral gene content in their mitogenomes.

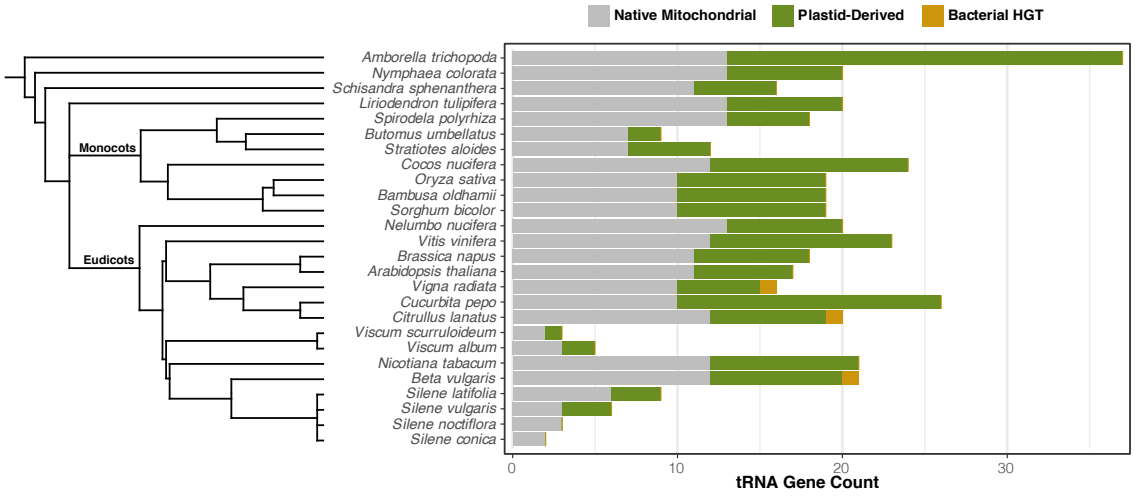


Figure 1.3| Extensive variation in tRNA gene counts within the mitogenomes of a diverse sampling of angiosperms. Counts are based on intact gene sequences, excluding duplicate copies of the same gene. Genes are categorized based on their evolutionary origin, as indicated by color code. The tree on the left indicates evolutionary relationships among species, with the topology roughly following the Angiosperm Phylogeny Website v14 (http://www.mobot.org/MOBOT/research/APweb/). For *Amborella trichopoda*, which has been the recipient of extensive HGT from other plant species, the count for native mitochondrial tRNAs is limited to those inferred to be functional by Rice et al. (2013). These include cases of apparent vertical inheritance of the ancestral mitochondrial tRNA genes, as well as inferred cases of functional replacement by HGT of an orthologous gene copy from the mitogenome of another plant species. In contrast, HGTs that would have augmented the set of mitochondrial-derived tRNAs were assumed to be non-functional by Rice et al. (2013).

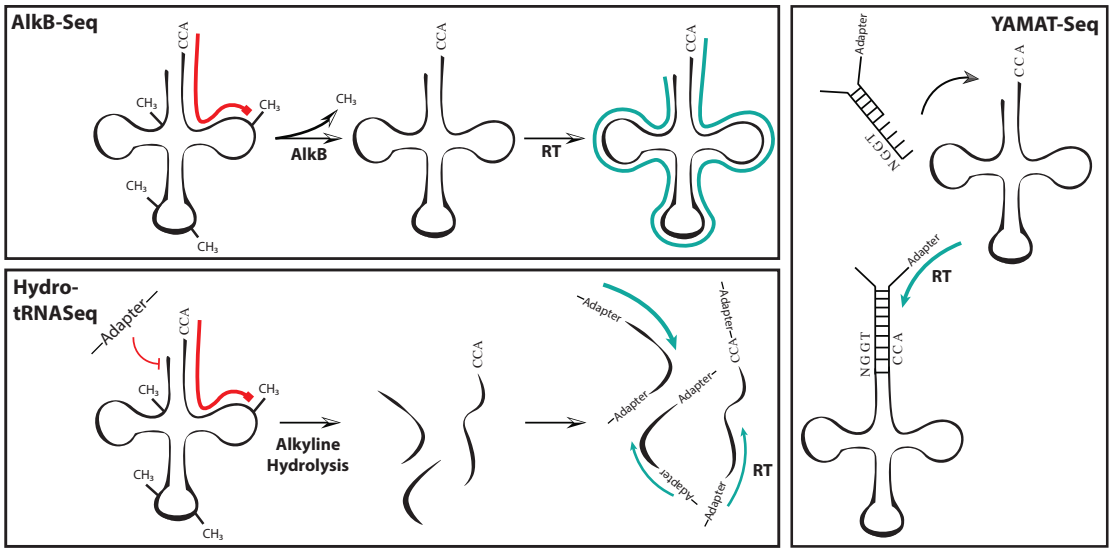


Figure 1.4| Methods to increase the efficiency of reverse transcription (RT) and adapter ligation of tRNA molecules. Top left panel: RT-inhibiting modifications are removed with the enzyme AlkB prior to RT, allowing for production of full-sized cDNA molecules. Lower left panel: The base-paired 3′- and 5′-ends of tRNA molecules inhibit ligation reactions. The partial alkaline hydrolysis of tRNAs generates fragments which are more amenable for sequencing. Right panel: In order to specifically ligate adapters to mature tRNAs, Y-shaped adapters with an overhang that is complementary to the discriminator base and the CCA tail on mature tRNAs are utilized.

Table 1.1| mt-tRNA-interacting enzymes in *Arabidopsis thaliana*

Enzyme	Arabidopsis ID	Function	Compartment[1]	Enzyme Origin[2]	Target Mitochondrial tRNAs[3]
AlaRS	At1g50200	aminoacylation	M|P|N	Eukaryotic: Cytosolic	trnA-cyto
AlaRS	At5g22800	aminoacylation	M|P	Cyanobacterial	trnA-cyto
ArgRS	At4g26300	aminoacylation	M|P|N[6]	Eukaryotic: Cytosolic	trnR-cyto
AsnRS	At4g17300	aminoacylation	M|P	Cyanobacterial[7]	trnN(guu)-cp
AspRS	At4g33760	aminoacylation	M|P	Cyanobacterial	trnD(guc)-cp
CysRS	At2g31170	aminoacylation	M|P	Other Bacterial	trnC(gca)-mt
GluRS	At5g64050	aminoacylation	M|P	Cyanobacterial[7]	trnE(uuc)-mt, trnQ(uug)-mt
GlyRS	At1g29880[5]	aminoacylation	M|N	Eukaryotic: Cytosolic	trnG(gcc)-mt, trnG-cyto
GlyRS	At3g48110	aminoacylation	M|P	Other Bacterial	trnG(gcc)-mt, trnG-cyto
HisRS	At3g46100	aminoacylation	M|P	Unclear[8]	trnH(gug)-cp
IleRS	At5g49030	aminoacylation	M|P	Cyanobacterial	trnI(cau)-mt, trnI-cyto
LeuRS	At1g09620	aminoacylation	M|N	Eukaryotic: Cytosolic	trnL-cyto
LysRS	At3g13490	aminoacylation	M|P	Cyanobacterial	trnK(uuu)-mt
MetRS	At3g55400	aminoacylation	M|P	Cyanobacterial	trnfM(cau)-mt, trnM(cau)-cp[10]
PheRS	At3g58140	aminoacylation	M|P	Eukaryotic: Mitochondrial	trnF-cyto
ProRS	At5g52520	aminoacylation	M|P	Other Bacterial[9]	trnP(ugg)-mt
SerRS	At1g11870	aminoacylation	M|P	Eukaryotic: Mitochondrial	trnS(gcu)-mt, trnS(gga)-cp, trnS(uga)-mt
ThrRS	At5g26830	aminoacylation	M|N	Eukaryotic: Cytosolic	trnT-cyto
ThrRS	At2g04842	aminoacylation	M|P	Cyanobacterial	trnT-cyto
TrpRS	At2g25840	aminoacylation	M|P	Cyanobacterial	trnW-cyto[11]
TyrRS	At3g02660	aminoacylation	M|P	Other Bacterial	trnY(gua)-mt
ValRS	At1g14610[5]	aminoacylation	M|N	Eukaryotic: Cytosolic	trnV-cyto
ValRS	At5g16715	aminoacylation	M|P	Other Bacterial	trnV-cyto
PRORP	At2g32230	5'-end processing	M|P	Eukaryotic	All
tRNase Z	At3g16260	3'-end processing	M	Eukaryotic	All
tRNase Z	At1g52160	3'-end processing	M|N	Eukaryotic	All
CCAse	At1g22660	CCA-tail addition	M|P|N	Eukaryotic	All
GatA	At3g25660	Glu/Gln amidation	M|P	Cyanobacterial	trnQ(uug)-mt
GatB	At1g48520	Glu/Gln amidation	M|P	Cyanobacterial	trnQ(uug)-mt
GatC	At4g32915	Glu/Gln amidation	M|P	Cyanobacterial	trnQ(uug)-mt
MTF[4]	At1g66520	met. formylation	M|P		trnfM(cau)-mt
TiLS[4]	At3g24560	trnI lysidine mod.	M|P		trnI(cau)-mt

[1]Compartment classifications (M: mitochondrial; P: plastid; N: nuclear/cytosolic) are based on targeting information from Duchêne et al. (2005), von Braun et al. (2007), Pujol et al. (2008), Canino et al. (2009), and Gobert et al. (2010).

[2]Enzyme origins are curated from Duchêne et al. (2005), Brindefalk et al. (2007), Pujol et al. (2008), Brandão et al. (2011), and Lechner et al. (2015). Enzymes identified as eukaryotic are inferred to have been present in LECA. For aaRSs, "mitochondrial" and "cytosolic" indicate the inferred functional role in LECA but do not necessarily imply deeper alphaproteobacterial or archaeal origins.

[3]In *Arabidopsis*, import of cytosolic tRNAs into the mitochondria has only been demonstrated for trnF and trnW. Other examples of cytosolic tRNAs are inferred to be imported based on the absence of corresponding genes in the mitogenome and documented import in other angiosperms (Salinas-Giegé et al. 2015).

[4]MTF and TiLS have not been characterized in plants, but the listed *Arabidopsis* genes represent candidates based on sequence similarity.

[5]Despite evidence of mitochondrial targeting, cytosolic-like GlyRS (At1g29880) and ValRS (At1g14610) may not be active in mitochondrial aminoacylation (Duchêne et al. 2009).

[6]Empirical evidence for plastid-targeting was obtained for ArgRS (At4g26300), but it is inferred to also function in the mitochondria and cytosol. See main text.

[7]Duchêne et al. (2005) identified AsnRS (At4g17300) and GluRS (At5g64050) as cyanobacterial in origin, but other studies have produced conflicting or uncertain results potentially pointing to other bacterial sources (Brindefalk et al. 2007, Brandão et al. 2011), so these classifications are tentative.

[8]Duchêne et al. (2005) identified HisRS (At3g46100) as archaeal in origin, but more recent studies have indicated an unresolved placement relative to bacteria and archaea (Brindefalk et al. 2007, Brandão et al. 2011).

[9]Duchêne et al. (2005) identified ProRS (At5g52520) as cytosolic in origin, but more recent studies have indicated a bacterial origin (Brindefalk et al. 2007, Brandão et al. 2011).

[10]Previous hybridization analysis of the plastid-derived trnM(cau) gene did not detect expression in *Arabidopsis* mitochondria (Duchêne and Maréchal-Drouard 2001), but we recently found evidence of mature tRNA transcripts from this gene using high-throughput sequencing (Warren et al. 2019).

[11]Previous hybridization analysis of the plastid-derived trnW(cca) gene did not detect expression in *Arabidopsis* mitochondria and instead detected import of a cytosolic counterpart (Duchêne and Maréchal-Drouard 2001).

LITERATURE CITED

Adams, K.L., Palmer, J.D., 2003. Evolution of mitochondrial gene content: gene loss and transfer to the nucleus. Mol Phylogenet Evol 29, 380-395.

Adams, K.L., Qiu, Y.L., Stoutemyer, M., Palmer, J.D., 2002. Punctuated evolution of mitochondrial gene content: high and variable rates of mitochondrial gene loss and transfer to the nucleus during angiosperm evolution. Proc Natl Acad Sci U S A 99, 9905-9912.

Adrion, J.R., White, P.S., Montooth, K.L., 2016. The Roles of Compensatory Evolution and Constraint in Aminoacyl tRNA Synthetase Evolution. Mol Biol Evol 33, 152-161.

Alfonzo, J.D., Blanc, V., Estévez, A.M., Rubio, M.A.T., Simpson, L., 1999. C to U editing of the anticodon of imported mitochondrial tRNATrp allows decoding of the UGA stop codon in Leishmania tarentolae. The EMBO journal 18, 7056-7062.

Alfonzo, J.D., Söll, D., 2009. Mitochondrial tRNA import--the challenge to understand has just begun. Biol Chem 390, 717-722.

Altman, S., 2007. A view of RNase P. Mol Biosyst 3, 604-607.

Alverson, A.J., Wei, X.X., Rice, D.W., Stern, D.B., Barry, K., Palmer, J.D., 2010. Insights into the Evolution of Mitochondrial Genome Size from Complete Sequences of Citrullus lanatus and Cucurbita pepo (Cucurbitaceae). Mol Biol Evol 27, 1436-1448.

Arimura, S., Yamamoto, J., Aida, G.P., Nakazono, M., Tsutsumi, N., 2004. Frequent fusion and fission of plant mitochondria with unequal nucleoid distribution. Proc Natl Acad Sci U S A 101, 7805-7808.

Arita, M., Suematsu, T., Osanai, A., Inaba, T., Kamiya, H., Kita, K., Sisido, M., Watanabe, Y., Ohtsuki, T., 2006. An evolutionary 'intermediate state' of mitochondrial translation systems found in Trichinella species of parasitic nematodes: co-evolution of tRNA and EF-Tu. Nucleic Acids Res 34, 5291-5299.

Arrieta-Montiel, M.P., Mackenzie, S.A., 2011. Plant Mitochondrial Genomes and Recombination, in: Kempken, F. (Ed.), Plant Mitochondria. Springer New York, New York, NY, pp. 65-82.

Banerjee, R., Chen, S., Dare, K., Gilreath, M., Praetorius-Ibba, M., Raina, M., Reynolds, N.M., Rogers, T., Roy, H., Yadavalli, S.S., Ibba, M., 2010. tRNAs: cellular barcodes for amino acids. FEBS Lett 584, 387-395.

Berg, M., Rogers, R., Muralla, R., Meinke, D., 2005. Requirement of aminoacyl-tRNA synthetases for gametogenesis and embryo development in Arabidopsis. Plant J 44, 866-878.

Bergthorsson, U., Adams, K.L., Thomason, B., Palmer, J.D., 2003. Widespread horizontal transfer of mitochondrial genes in flowering plants. Nature 424, 197-201.

Betat, H., Long, Y., Jackman, J.E., Morl, M., 2014. From end to end: tRNA editing at 5'- and 3'-terminal positions. Int J Mol Sci 15, 23975-23998.

Bonnard, G., Gobert, A., Arrivé, M., Pinker, F., Salinas-Giegé, T., Giegé, P., 2016. Transfer RNA maturation in Chlamydomonas mitochondria, chloroplast and the nucleus by a single RNase P protein. Plant J 87, 270-280.

Brandao, M.M., Silva-Filho, M.C., 2011. Evolutionary history of Arabidopsis thaliana aminoacyl-tRNA synthetase dual-targeted proteins. Mol Biol Evol 28, 79-85.

Brindefalk, B., Viklund, J., Larsson, D., Thollesson, M., Andersson, S.G., 2007. Origin and evolution of the mitochondrial aminoacyl-tRNA synthetases. Mol Biol Evol 24, 743-756.

Brubacher-Kauffmann, S., Maréchal-Drouard, L., Cosset, A., Dietrich, A., Duchêne, A.M., 1999. Differential import of nuclear-encoded tRNAGly isoacceptors into solanum Tuberosum mitochondria. Nucleic Acids Res 27, 2037-2042.

Burger, G., Gray, M.W., Forget, L., Lang, B.F., 2013. Strikingly bacteria-like and gene-rich mitochondrial genomes throughout jakobid protists. Genome Biol Evol 5, 418-438.

Canino, G., Bocian, E., Barbezier, N., Echeverria, M., Forner, J., Binder, S., Marchfelder, A., 2009. Arabidopsis encodes four tRNase Z enzymes. Plant Physiol 150, 1494-1502.

Charriere, F., Helgadottir, S., Horn, E.K., Söll, D., Schneider, A., 2006. Dual targeting of a single tRNA(Trp) requires two different tryptophanyl-tRNA synthetases in Trypanosoma brucei. Proc Natl Acad Sci U S A 103, 6847-6852.

Charriere, F., O'Donoghue, P., Helgadottir, S., Maréchal-Drouard, L., Cristodero, M., Horn, E.K., Söll, D., Schneider, A., 2009. Dual targeting of a tRNAAsp requires two different aspartyl-tRNA synthetases in Trypanosoma brucei. J Biol Chem 284, 16210-16217.

Chaw, S.M., Shih, A.C., Wang, D., Wu, Y.W., Liu, S.M., Chou, T.Y., 2008. The mitochondrial genome of the gymnosperm Cycas taitungensis contains a novel family of short interspersed elements, Bpu sequences, and abundant RNA editing sites. Mol Biol Evol 25, 603-615.

Chien, C.I., Chen, Y.W., Wu, Y.H., Chang, C.Y., Wang, T.L., Wang, C.C., 2014. Functional substitution of a eukaryotic glycyl-tRNA synthetase with an evolutionarily unrelated bacterial cognate enzyme. PLoS One 9, e94659.

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, 1771-1784.

Coffin, J.W., Cossins, E.A., 1986. Mitochondrial folates and methionyl-tRNA transformylase activity during germination and early growth of seeds. Phytochemistry 25, 2481-2487.

Cognat, V., Morelle, G., Megel, C., Lalande, S., Molinier, J., Vincent, T., Small, I., Duchêne, A.M., Maréchal-Drouard, L., 2017. The nuclear and organellar tRNA-derived RNA fragment population in Arabidopsis thaliana is highly dynamic. Nucleic Acids Res 45, 3460-3472.

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, 879-+.

Cummings, M.P., Nugent, J.M., Olmstead, R.G., Palmer, J.D., 2003. Phylogenetic analysis reveals five independent transfers of the chloroplast gene rbcL to the mitochondrial genome in angiosperms. Curr Genet 43, 131-138.

Dietrich, A., Weil, J.H., Maréchal-Drouard, L., 1992. Nuclear-encoded transfer RNAs in plant mitochondria. Annu Rev Cell Biol 8, 115-131.

Doolittle, R.F., Handy, J., 1998. Evolutionary anomalies among the aminoacyl-tRNA synthetases. Curr Opin Genet Dev 8, 630-636.

Duchêne, A.M., Giritch, A., Hoffmann, B., Cognat, V., Lancelin, D., Peeters, N.M., Zaepfel, M., Maréchal-Drouard, L., Small, I.D., 2005. Dual targeting is the rule for organellar aminoacyl-tRNA synthetases in Arabidopsis thaliana. Proc Natl Acad Sci U S A 102, 16484-16489.

Duchêne, A.M., Maréchal-Drouard, L., 2001. The chloroplast-derived trnW and trnM-e genes are not expressed in Arabidopsis mitochondria. Biochem Biophys Res Commun 285, 1213-1216.

Duchêne, A.M., Peeters, N., Dietrich, A., Cosset, A., Small, I.D., Wintz, H., 2001. Overlapping destinations for two dual targeted glycyl-tRNA synthetases in Arabidopsis thaliana and Phaseolus vulgaris. Journal of Biological Chemistry 276, 15275-15283.

Duchêne, A.M., Pujol, C., Maréchal-Drouard, L., 2009. Import of tRNAs and aminoacyl-tRNA synthetases into mitochondria. Curr Genet 55, 1-18.

Esseiva, A.C., Naguleswaran, A., Hemphill, A., Schneider, A., 2004. Mitochondrial tRNA import in Toxoplasma gondii. Journal of Biological Chemistry 279, 42363-42368.

Esser, K., Lüttge, U., Beyschlag, W., Murata, J., 2006. Progress in Botany 68. Springer Berlin Heidelberg.

Fey, J., Dietrich, A., Cosset, A., Desprez, T., Maréchal-Drouard, L., 1997. Evolutionary aspects of "chloroplast-like" trnN and trnH expression in higher-plant mitochondria. Curr Genet 32, 358-360.

Gandini, C.L., Sanchez-Puerta, M.V., 2017. Foreign Plastid Sequences in Plant Mitochondria are Frequently Acquired Via Mitochondrion-to-Mitochondrion Horizontal Transfer. Sci Rep 7, 43402.

Giegé, R., Sissler, M., Florentz, C., 1998. Universal rules and idiosyncratic features in tRNA identity. Nucleic Acids Research 26, 5017-5035.

Gobert, A., Gutmann, B., Taschner, A., Gößringer, M., Holzmann, J., Hartmann, R.K., Rossmanith, W., Giegé, P., 2010. A single Arabidopsis organellar protein has RNase P activity. Nat Struct Mol Biol 17, 740-744.

Gogakos, T., Brown, M., Garzia, A., Meyer, C., Hafner, M., Tuschl, T., 2017. Characterizing Expression and Processing of Precursor and Mature Human tRNAs by Hydro-tRNAseq and PAR-CLIP. Cell Rep 20, 1463-1475.

Goremykin, V.V., Lockhart, P.J., Viola, R., Velasco, R., 2012. The mitochondrial genome of Malus domestica and the import-driven hypothesis of mitochondrial genome expansion in seed plants. Plant J 71, 615-626.

Grewe, F., Viehoever, P., Weisshaar, B., Knoop, V., 2009. A trans-splicing group I intron and tRNA-hyperediting in the mitochondrial genome of the lycophyte Isoetes engelmannii. Nucleic Acids Res 37, 5093-5104.

Grosjean, H., Westhof, E., 2016. An integrated, structure- and energy-based view of the genetic code. Nucleic Acids Res 44, 8020-8040.

Guo, M., Schimmel, P., 2013. Essential nontranslational functions of tRNA synthetases. Nat Chem Biol 9, 145-153.

Guo, W., Grewe, F., Fan, W., Young, G.J., Knoop, V., Palmer, J.D., Mower, J.P., 2016. Ginkgo and Welwitschia Mitogenomes Reveal Extreme Contrasts in Gymnosperm Mitochondrial Evolution. Mol Biol Evol 33, 1448-1460.

Guo, W., Zhu, A., Fan, W., Mower, J.P., 2017. Complete mitochondrial genomes from the ferns Ophioglossum californicum and Psilotum nudum are highly repetitive with the largest organellar introns. New Phytol 213, 391-403.

Gutmann, B., Gobert, A., Giegé, P., 2012. PRORP proteins support RNase P activity in both organelles and the nucleus in Arabidopsis. Genes Dev 26, 1022-1027.

Haen, K.M., Pett, W., Lavrov, D.V., 2010. Parallel loss of nuclear-encoded mitochondrial aminoacyl-tRNA synthetases and mtDNA-encoded tRNAs in Cnidaria. Mol Biol Evol 27, 2216-2219.

Hao, W., Palmer, J.D., 2009. Fine-scale mergers of chloroplast and mitochondrial genes create functional, transcompartmentally chimeric mitochondrial genes. Proc Natl Acad Sci U S A 106, 16728-16733.

Hecht, J., Grewe, F., Knoop, V., 2011. Extreme RNA editing in coding islands and abundant microsatellites in repeat sequences of Selaginella moellendorffii mitochondria: the root of frequent plant mtDNA recombination in early tracheophytes. Genome Biol Evol 3, 344-358.

Holley, R.W., Apgar, J., Everett, G.A., Madison, J.T., Marquisee, M., Merrill, S.H., Penswick, J.R., Zamir, A., 1965. Structure of a Ribonucleic Acid. Science 147, 1462-1465.

Holzmann, J., Frank, P., Löffler, E., Bennett, K.L., Gerner, C., Rossmanith, W., 2008. RNase P without RNA: identification and functional reconstitution of the human mitochondrial tRNA processing enzyme. Cell 135, 462-474.

Hopper, A.K., 2013. Transfer RNA post-transcriptional processing, turnover, and subcellular dynamics in the yeast Saccharomyces cerevisiae. Genetics 194, 43-67.

Hopper, A.K., Nostramo, R.T., 2019. tRNA Processing and Subcellular Trafficking Proteins Multitask in Pathways for Other RNAs. Front Genet 10, 96.

Hummel, G., Warren, J., Drouard, L., 2019. The multi-faceted regulation of nuclear tRNA gene transcription. IUBMB Life 71, 1099-1108.

Huynen, M.A., Duarte, I., Szklarczyk, R., 2013. Loss, replacement and gain of proteins at the origin of the mitochondria. Bba-Bioenergetics 1827, 224-231.

Ibba, M., Söll, D., 2004. Aminoacyl-tRNAs: setting the limits of the genetic code. Genes Dev 18, 731-738.

Jackman, S.D., Coombe, L., Warren, R.L., Kirk, H., Trinh, E., McLeod, T., Pleasance, S., Pandoh, P., Zhao, Y., Coope, R.J., Bousquet, J., Bohlmann, J., Jones, S.J.M., Birol, I., 2019. Largest Complete Mitochondrial Genome of a Gymnosperm, Sitka Spruce (Picea sitchensis), Indicates Complex Physical Structure. bioRxiv, 601104.

Janouškovec, J., Tikhonenkov, D.V., Burki, F., Howe, A.T., Rohwer, F.L., Mylnikov, A.P., Keeling, P.J., 2017. A New Lineage of Eukaryotes Illuminates Early Mitochondrial Genome Reduction. Curr Biol 27, 3717-3724 e3715.

Johnston, I.G., Williams, B.P., 2016. Evolutionary Inference across Eukaryotes Identifies Specific Pressures Favoring Mitochondrial Gene Retention. Cell Syst 2, 101-111.

Joyce, P.B., Gray, M.W., 1989. Chloroplast-like transfer RNA genes expressed in wheat mitochondria. Nucleic Acids Res 17, 5461-5476.

Kamenski, P., Kolesnikova, O., Jubenot, V., Entelis, N., Krasheninnikov, I.A., Martin, R.P., Tarassov, I., 2007. Evidence for an adaptation mechanism of mitochondrial translation via tRNA import from the cytosol. Mol Cell 26, 625-637.

Kamikawa, R., Shiratori, T., Ishida, K., Miyashita, H., Roger, A.J., 2016. Group II Intron-Mediated Trans-Splicing in the Gene-Rich Mitochondrial Genome of an Enigmatic Eukaryote, Diphylleia rotans. Genome Biol Evol 8, 458-466.

Kan, S.L., Shen, T.T., Gong, P., Ran, J.H., Wang, X.Q., 2020. The complete mitochondrial genome of Taxus cuspidata (Taxaceae): eight protein-coding genes have transferred to the nuclear genome. Bmc Evol Biol 20, 10.

Kirchner, S., Ignatova, Z., 2015. Emerging roles of tRNA in adaptive translation, signalling dynamics and disease. Nat Rev Genet 16, 98-112.

Kitazaki, K., Kubo, T., Kagami, H., Matsumoto, T., Fujita, A., Matsuhira, H., Matsunaga, M., Mikami, T., 2011. A horizontally transferred tRNA(Cys) gene in the sugar beet mitochondrial genome: evidence that the gene is present in diverse angiosperms and its transcript is aminoacylated. Plant J 68, 262-272.

Knie, N., Polsakiewicz, M., Knoop, V., 2015. Horizontal Gene Transfer of Chlamydial-Like tRNA Genes into Early Vascular Plant Mitochondria. Mol Biol Evol 32, 629-634.

Knoop, V., 2004. The mitochondrial DNA of land plants: peculiarities in phylogenetic perspective. Curr Genet 46, 123-139.

Kohn, A.B., Citarella, M.R., Kocot, K.M., Bobkova, Y.V., Halanych, K.M., Moroz, L.L., 2012. Rapid evolution of the compact and unusual mitochondrial genome in the ctenophore, Pleurobrachia bachei. Mol Phylogenet Evol 63, 203-207.

Kubo, T., Nishizawa, S., Sugawara, A., Itchoda, N., Estiati, A., Mikami, T., 2000. The complete nucleotide sequence of the mitochondrial genome of sugar beet (Beta vulgaris L.) reveals a novel gene for tRNA(Cys)(GCA). Nucleic Acids Res 28, 2571-2576.

Kühn, K., Richter, U., Meyer, E.H., Delannoy, E., de Longevialle, A.F., O'Toole, N., Börner, T., Millar, A.H., Small, I.D., Whelan, J., 2009. Phage-type RNA polymerase RPOTmp performs gene-specific transcription in mitochondria of Arabidopsis thaliana. Plant Cell 21, 2762-2779.

Kumar, R., Maréchal-Drouard, L., Akama, K., Small, I., 1996. Striking differences in mitochondrial tRNA import between different plant species. Mol Gen Genet 252, 404-411.

Lama, L., Cobo, J., Buenaventura, D., Ryan, K., 2019. Small RNA-seq: The RNA 5'-end adapter ligation problem and how to circumvent it. J Biol Methods 6.

Lang, B.F., Gray, M.W., Burger, G., 1999. Mitochondrial genome evolution and the origin of eukaryotes. Annu Rev Genet 33, 351-397.

Lavrov, D.V., Pett, W., 2016. Animal Mitochondrial DNA as We Do Not Know It: mt-Genome Organization and Evolution in Nonbilaterian Lineages. Genome Biol Evol 8, 2896-2913.

Lechner, M., Rossmanith, W., Hartmann, R.K.,Thölken, C., Gutmann, B., Giegé, P., Gobert, A., 2015. Distribution of Ribonucleoprotein and Protein-Only RNase P in Eukarya. Mol Biol Evol 32, 3186-3193.

Leon, P., Walbot, V., Bedinger, P., 1989. Molecular analysis of the linear 2.3 kb plasmid of maize mitochondria: apparent capture of tRNA genes. Nucleic Acids Res 17, 4089-4099.

Liu, Y., Xue, J.Y., Wang, B., Li, L., Qiu, Y.L., 2011. The mitochondrial genomes of the early land plants Treubia lacunosa and Anomodon rugelii: dynamic and conservative evolution. PLoS One 6, e25836.

Marchfelder, A., Schuster, W., Brennicke, A., 1990. In vitro processing of mitochondrial and plastid derived tRNA precursors in a plant mitochondrial extract. Nucleic Acids Res 18, 1401-1406.

Maréchal-Drouard, L., Guillemaut, P., Cosset, A., Arbogast, M., Weber, F., Weil, J.H., Dietrich, A., 1990. Transfer RNAs of potato (Solanum tuberosum) mitochondria have different genetic origins. Nucleic Acids Res 18, 3689-3696.

Maréchal-Drouard, L., Weil, J., Dietrich, A., 1993. Transfer RNAs and transfer RNA genes in plants. Annual review of plant biology 44, 13-32.

Martin, R.P., Schneller, J.M., Stahl, A.J., Dirheimer, G., 1979. Import of nuclear deoxyribonucleic acid coded lysine-accepting transfer ribonucleic acid (anticodon C-U-U) into yeast mitochondria. Biochemistry 18, 4600-4605.

Meiklejohn, C.D., Holmbeck, M.A., Siddiq, M.A., Abt, D.N., Rand, D.M., Montooth, K.L., 2013. An Incompatibility between a Mitochondrial tRNA and Its Nuclear-Encoded tRNA Synthetase Compromises Development and Fitness in Drosophila. Plos Genet 9.

Mercer, T.R., Neph, S., Dinger, M.E., Crawford, J., Smith, M.A., Shearwood, A.M., Haugen, E., Bracken, C.P., Rackham, O., Stamatoyannopoulos, J.A., Filipovska, A., Mattick, J.S., 2011. The human mitochondrial transcriptome. Cell 146, 645-658.

Mishina, Y., He, C., 2006. Oxidative dealkylation DNA repair mediated by the mononuclear non-heme iron AlkB proteins. J Inorg Biochem 100, 670-678.

Miyata, S., Nakazono, M., Hirai, A., 1998. Transcription of plastid-derived tRNA genes in rice mitochondria. Curr Genet 34, 216-220.

Motorin, Y., Helm, M., 2010. tRNA stabilization by modified nucleotides. Biochemistry 49, 4934-4944.

Motorin, Y., Muller, S., Behm-Ansmant, I., Branlant, C., 2007. Identification of modified residues in RNAs by reverse transcription-based methods. Methods Enzymol 425, 21-53.

Murcha, M.W., Kubiszewski-Jakubiak, S., Teixeira, P.F., Gügel, I.L., Kmiec, B., Narsai, R., Ivanova, A., Megel, C., Schock, A., Kraus, S., Berkowitz, O., Glaser, E., Philippar, K., Maréchal-Drouard, L.,

Soll, J., Whelan, J., 2016. Plant-Specific Preprotein and Amino Acid Transporter Proteins Are Required for tRNA Import into Mitochondria. Plant Physiol 172, 2471-2490.

Nakazono, M., Nishiwaki, S., Tsutsumi, N., Hirai, A., 1996. A chloroplast-derived sequence is utilized as a source of promoter sequences for the gene for subunit 9 of NADH dehydrogenase (nad9) in rice mitochondria. Mol Gen Genet 252, 371-378.

Nameki, N., Tamura, K., Asahara, H., Hasegawa, T., 1997. Recognition of tRNA(Gly) by three widely diverged glycyl-tRNA synthetases. J Mol Biol 268, 640-647.

Notsu, Y., Masood, S., Nishikawa, T., Kubo, N., Akiduki, G., Nakazono, M., Hirai, A., Kadowaki, K., 2002. The complete sequence of the rice (Oryza sativa L.) mitochondrial genome: frequent DNA sequence acquisition and loss during the evolution of flowering plants. Mol Genet Genomics 268, 434-445.

Noutahi, E., Calderon, V., Blanchette, M., El-Mabrouk, N., Lang, B.F., 2019. Rapid Genetic Code Evolution in Green Algal Mitochondrial Genomes. Mol Biol Evol 36, 766-783.

O'Donoghue, P., Luthey-Schulten, Z., 2003. On the evolution of structure in aminoacyl-tRNA synthetases. Microbiol Mol Biol Rev 67, 550-573.

One Thousand Plant Transcriptomes, I., 2019. One thousand plant transcriptomes and the phylogenomics of green plants. Nature 574, 679-685.

Ott, M., Amunts, A., Brown, A., 2016. Organization and Regulation of Mitochondrial Protein Synthesis. Annu Rev Biochem 85, 77-101.

Paris, Z., Alfonzo, J.D., 2018. How the intracellular partitioning of tRNA and tRNA modification enzymes affects mitochondrial function. IUBMB Life 70, 1207-1213.

Park, S., Grewe, F., Zhu, A., Ruhlman, T.A., Sabir, J., Mower, J.P., Jansen, R.K., 2015. Dynamic evolution of Geranium mitochondrial genomes through multiple horizontal and intracellular gene transfers. New Phytol 208, 570-583.

Pearce, S., Nezich, C.L., Spinazzola, A., 2013. Mitochondrial diseases: translation matters. Mol Cell Neurosci 55, 1-12.

Petersen, G., Cuenca, A., Møller, I.M., Seberg, O., 2015. Massive gene loss in mistletoe (Viscum, Viscaceae) mitochondria. Sci Rep 5, 17588.

Petrov, A.S., Gulen, B., Norris, A.M., Kovacs, N.A., Bernier, C.R., Lanier, K.A., Fox, G.E., Harvey, S.C., Wartell, R.M., Hud, N.V., Williams, L.D., 2015. History of the ribosome and the origin of translation. Proc Natl Acad Sci U S A 112, 15396-15401.

Pett, W., Lavrov, D.V., 2015. Cytonuclear Interactions in the Evolution of Animal Mitochondrial tRNA Metabolism. Genome Biol Evol 7, 2089-2101.

Pujol, C., Bailly, M., Kern, D., Marechal-Drouard, L., Becker, H., Duchêne, A.M., 2008. Dual-targeted tRNA-dependent amidotransferase ensures both mitochondrial and chloroplastic Gln-tRNAGln synthesis in plants. Proc Natl Acad Sci U S A 105, 6481-6485.

Phizicky, E.M., Hopper, A.K., 2010. tRNA biology charges to the front. Genes Dev 24, 1832-1860.

Raina, M., Ibba, M., 2014. tRNAs as regulators of biological processes. Front Genet 5, 171.

Rice, D.W., Alverson, A.J., Richardson, A.O., Young, G.J., Sanchez-Puerta, M.V., Munzinger, J., Barry, K., Boore, J.L., Zhang, Y., dePamphilis, C.W., Knox, E.B., Palmer, J.D., 2013. Horizontal transfer of entire genomes via mitochondrial fusion in the angiosperm Amborella. Science 342, 1468-1473.

Richardson, A.O., Rice, D.W., Young, G.J., Alverson, A.J., Palmer, J.D., 2013. The "fossilized" mitochondrial genome of Liriodendron tulipifera: ancestral gene content and order, ancestral editing sites, and extraordinarily low mutation rate. BMC Biol 11, 29.

Rinehart, J., Krett, B., Rubio, M.A., Alfonzo, J.D., Söll, D., 2005. Saccharomyces cerevisiae imports the cytosolic pathway for Gln-tRNA synthesis into the mitochondrion. Genes Dev 19, 583-592.

Rodnina, M.V., Gromadski, K.B., Kothe, U., Wieden, H.J., 2005. Recognition and selection of tRNA in translation. FEBS Lett 579, 938-942.

Roger, A.J., Munoz-Gomez, S.A., Kamikawa, R., 2017. The Origin and Diversification of Mitochondria. Curr Biol 27, R1177-R1192.

Rogers, H.H., Griffiths-Jones, S., 2014. tRNA anticodon shifts in eukaryotic genomes. RNA 20, 269-281.

Rossmanith, W., 2012. Of P and Z: mitochondrial tRNA processing enzymes. Biochim Biophys Acta 1819, 1017-1026.

Rubio, M.A., Rinehart, J.J., Krett, B., Duvezin-Caubet, S., Reichert, A.S., Söll, D., Alfonzo, J.D., 2008. Mammalian mitochondria have the innate ability to import tRNAs by a mechanism distinct from protein import. Proc Natl Acad Sci U S A 105, 9186-9191.

Salazar, J.C., Ahel, I., Orellana, O., Tumbula-Hansen, D., Krieger, R., Daniels, L., Söll, D., 2003. Coevolution of an aminoacyl-tRNA synthetase with its tRNA substrates. Proc Natl Acad Sci U S A 100, 13863-13868.

Salinas, T., Duby, F., Larosa, V., Coosemans, N., Bonnefoy, N., Motte, P., Maréchal-Drouard, L., Remacle, C., 2012. Co-evolution of mitochondrial tRNA import and codon usage determines translational efficiency in the green alga Chlamydomonas. Plos Genet 8, e1002946.

Salinas, T., Duchêne, A.M., Maréchal-Drouard, L., 2008. Recent advances in tRNA mitochondrial import. Trends Biochem Sci 33, 320-329.

Salinas, T., Schaeffer, C., Maréchal-Drouard, L., Duchêne, A.M., 2005. Sequence dependence of tRNA(Gly) import into tobacco mitochondria. Biochimie 87, 863-872.

Salinas-Giegé, T., Giegé, R., Giegé, P., 2015. tRNA biology in mitochondria. Int J Mol Sci 16, 4518-4559.

Sanchez-Puerta, M.V., Edera, A., Gandini, C.L., Williams, A.V., Howell, K.A., Nevill, P.G., Small, I., 2019. Genome-scale transfer of mitochondrial DNA from legume hosts to the holoparasite Lophophytum mirabile (Balanophoraceae). Mol Phylogenet Evol 132, 243-250.

Schneider, A., 2001. Unique aspects of mitochondrial biogenesis in trypanosomatids. International journal for parasitology 31, 1403-1415.

Schneider, A., 2011. Mitochondrial tRNA Import and Its Consequences for Mitochondrial Translation. Annu Rev Biochem 80, 1033-1053.

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.

Sinn, B.T., Barrett, C.F., 2019. Ancient mitochondrial gene transfer between fungi and the orchids. Mol Biol Evol.

Skippington, E., Barkman, T.J., Rice, D.W., Palmer, J.D., 2015. Miniaturized mitogenome of the parasitic plant Viscum scurruloideum is extremely divergent and dynamic and has lost all nad genes. Proc Natl Acad Sci U S A 112, E3515-3524.

Sloan, D.B., Alverson, A.J., Chuckalovcak, J.P., Wu, M., McCauley, D.E., Palmer, J.D., Taylor, D.R., 2012a. Rapid Evolution of Enormous, Multichromosomal Genomes in Flowering Plant Mitochondria with Exceptionally High Mutation Rates. Plos Biol 10.

Sloan, D.B., Alverson, A.J.,Storchová, H., Palmer, J.D., Taylor, D.R., 2010. Extensive loss of translational genes in the structurally dynamic mitochondrial genome of the angiosperm Silene latifolia. Bmc Evol Biol 10.

Sloan, D.B., Müller, K., McCauley, D.E., Taylor, D.R.,Storchová, H., 2012b. Intraspecific variation in mitochondrial genome sequence, structure, and gene content in Silene vulgaris, an angiosperm with pervasive cytoplasmic male sterility. New Phytol 196, 1228-1239.

Sloan, D.B., Warren, J.M., Williams, A.M., Wu, Z., Abdel-Ghany, S.E., Chicco, A.J., Havird, J.C., 2018. Cytonuclear integration and co-evolution. Nat Rev Genet 19, 635-648.

Sloan, D.B., Wu, Z., 2014. History of plastid DNA insertions reveals weak deletion and at mutation biases in angiosperm mitochondrial genomes. Genome Biol Evol 6, 3210-3221.

Small, I., Akashi, K., Chapron, A., Dietrich, A., Duchêne, A.M., Lancelin, D., Maréchal-Drouard, L., Menand, B., Mireau, H., Moudden, Y., Ovesna, J., Peeters, N., Sakamoto, W., Souciet, G., Wintz, H., 1999. The strange evolutionary history of plant mitochondrial tRNAs and their aminoacyl-tRNA synthetases. J Hered 90, 333-337.

Suzuki, T., Miyauchi, K., 2010. Discovery and characterization of tRNAIle lysidine synthetase (TilS). FEBS Lett 584, 272-277.

Takeuchi, N., Vial, L., Panvert, M., Schmitt, E., Watanabe, K., Mechulam, Y., Blanquet, S., 2001. Recognition of tRNAs by Methionyl-tRNA transformylase from mammalian mitochondria. J Biol Chem 276, 20064-20068.

Tan, T.H., Pach, R., Crausaz, A., Ivens, A., Schneider, A., 2002. tRNAs in Trypanosoma brucei: genomic organization, expression, and mitochondrial import. Mol Cell Biol 22, 3707-3717.

Taylor, Z.N., Rice, D.W., Palmer, J.D., 2015. The Complete Moss Mitochondrial Genome in the Angiosperm Amborella Is a Chimera Derived from Two Moss Whole-Genome Transfers. PLoS One 10, e0137532.

Timmis, J.N., Ayliffe, M.A., Huang, C.Y., Martin, W., 2004. Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. Nat Rev Genet 5, 123-135.

Tolkunova, E., Park, H., Xia, J., King, M.P., Davidson, E., 2000. The human lysyl-tRNA synthetase gene encodes both the cytoplasmic and mitochondrial enzymes by means of an unusual alternative splicing of the primary transcript. J Biol Chem 275, 35063-35069.

Torres, A.G., Reina, O., Stephan-Otto Attolini, C., Ribas de Pouplana, L., 2019. Differential expression of human tRNA genes drives the abundance of tRNA-derived fragments. Proc Natl Acad Sci U S A 116, 8451-8456.

Turmel, M., Otis, C., Lemieux, C., 2013. Tracing the evolution of streptophyte algae and their mitochondrial genome. Genome Biol Evol 5, 1817-1835.

Valach, M., Burger, G., Gray, M.W., Lang, B.F., 2014. Widespread occurrence of organelle genome-encoded 5S rRNAs including permuted molecules. Nucleic Acids Res 42, 13764-13777.

Vinogradova, E., Salinas, T., Cognat, V., Remacle, C., Maréchal-Drouard, L., 2009. Steady-state levels of imported tRNAs in Chlamydomonas mitochondria are correlated with both cytosolic and mitochondrial codon usages. Nucleic Acids Res 37, 1521-1528.

von Braun, S.S., Sabetti, A., Hanic-Joyce, P.J., Gu, J., Schleiff, E., Joyce, P.B., 2007. Dual targeting of the tRNA nucleotidyltransferase in plants: not just the signal. J Exp Bot 58, 4083-4093.

Wang, D., Rousseau-Gueutin, M., Timmis, J.N., 2012. Plastid sequences contribute to some plant mitochondrial genes. Mol Biol Evol 29, 1707-1711.

Wang, D., Wu, Y.W., Shih, A.C., Wu, C.S., Wang, Y.N., Chaw, S.M., 2007. Transfer of chloroplast genomic DNA to mitochondrial genome occurred at least 300 MYA. Mol Biol Evol 24, 2040-2048.

Wang, X.C., Chen, H., Yang, D., Liu, C., 2018. Diversity of mitochondrial plastid DNAs (MTPTs) in seed plants. Mitochondrial DNA A DNA Mapp Seq Anal 29, 635-642.

Warren, J.M., Salinas-Giegé, T., Hummel, G., Coots, N.L., Svendsen, J.M., Brown, K.C., Maréchal-Drouard, L., Sloan, D.B., 2019. Combining tRNA sequencing methods to characterize plant tRNA expression and post-transcriptional modification. bioRxiv, 790451.

Weber, F., Dietrich, A., Weil, J.H., Maréchal-Drouard, L., 1990. A potato mitochondrial isoleucine tRNA is coded for by a mitochondrial gene possessing a methionine anticodon. Nucleic Acids Res 18, 5027-5030.

Wilusz, J.E., 2015. Removing roadblocks to deep sequencing of modified RNAs. Nat Methods 12, 821-822.

Xue, J.Y., Liu, Y., Li, L., Wang, B., Qiu, Y.L., 2010. The complete mitochondrial genome sequence of the hornwort Phaeoceros laevis: retention of many ancient pseudogenes and conservative evolution of mitochondrial genomes in hornworts. Curr Genet 56, 53-61.

Yona, A.H., Bloom-Ackermann, Z., Frumkin, I., Hanson-Smith, V., Charpak-Amikam, Y., Feng, Q., Boeke, J.D., Dahan, O., Pilpel, Y., 2013. tRNA genes rapidly change in evolution to meet novel translational demands. Elife 2, e01339.

Zhang, J., Ferre-D'Amare, A.R., 2016. The tRNA Elbow in Structure, Recognition and Evolution. Life (Basel) 6.

Zheng, G., Qin, Y., Clark, W.C., Dai, Q., Yi, C., He, C., Lambowitz, A.M., Pan, T., 2015. Efficient and quantitative high-throughput tRNA sequencing. Nat Methods 12, 835-837.

CHAPTER 2: COMBINING TRNA SEQUENCING METHODS TO CHARACTERIZE PLANT TRNA EXPRESSION AND POST-TRANSCRIPTIONAL MODIFICATION[2]

Summary

Differences in tRNA expression have been implicated in a remarkable number of biological processes. There is growing evidence that tRNA genes can play dramatically different roles depending on both expression and post-transcriptional modification, yet sequencing tRNAs to measure abundance and detect modifications remains challenging. Their secondary structure and extensive post-transcriptional modifications interfere with RNA-seq library preparation methods and have limited the utility of high-throughput sequencing technologies. Here, we combine two modifications to standard RNA-seq methods by treating with the demethylating enzyme AlkB and ligating with tRNA-specific adapters in order to sequence tRNAs from four species of flowering plants, a group that has been shown to have some of the most extensive rates of post-transcriptional tRNA modifications. This protocol has the advantage of detecting full-length tRNAs and sequence variants that can be used to infer many post-transcriptional modifications. We used the resulting data to produce a modification index of almost all unique reference tRNAs in *Arabidopsis thaliana*, which exhibited many anciently conserved similarities with humans but also positions that appear to be "hot spots" for modifications in angiosperm tRNAs. We also found evidence based on northern blot analysis and droplet digital PCR that, even after demethylation treatment, tRNA-seq can produce highly biased estimates of absolute expression levels most likely due to biased reverse transcription. Nevertheless, the generation of full-length tRNA sequences with modification data is still promising for assessing differences in relative tRNA expression across treatments, tissues or subcellular fractions and help elucidate the functional roles of tRNA modifications.

[2] Published in *RNA Biology*, Volume 18, January 2021.
Authors: Jessica M. Warren, Thalia Salinas-Giegé, Guillaume Hummel, Nicole L. Coots, Joshua M. Svendsen, Kristen C. Brown, Laurence Drouard, Daniel B. Sloan

Introduction

Despite their central function in cell physiology and the increasing interest in quantifying their expression, transfer RNAs (tRNAs) remain difficult to sequence using conventional RNA sequencing (RNA-seq) methods. tRNAs are poorly suited to high-throughput sequencing library preparation protocols for two reasons. First, tRNAs are extensively post-transcriptionally modified, and certain base modifications occurring at the Watson-Crick face stall or terminate reverse transcription (RT) through interference with base pairing or steric hindrance, effectively blocking cDNA synthesis (MOTORIN *et al*. 2007). The majority of RNA-seq protocols require that RNA is reverse transcribed into cDNA before sequencing, but RT of tRNAs often results in truncated cDNA products that lack 5′ adapters and are never sequenced. Second, RT for RNA-seq is normally primed off of a ligated 3′ adapter, but mature tRNAs have compact secondary and tertiary structure with tightly base-paired 5′ and 3′ termini, which can prevent adapter ligation (WILUSZ 2015b). Hybridization methods such as microarrays and northern blot analysis have been used to quantify multiple tRNA species, but these approaches require prior knowledge of an organism's tRNA repertoire, and they may not able to evaluate expression levels of near-identical tRNAs because of cross-hybridization (SHIGEMATSU *et al*. 2017). Additionally, RNA-seq has higher sensitivity for genes that are very highly or lowly expressed and can detect a much larger range of expression (WANG *et al*. 2009).

One significant technological advance in developing effective tRNA-seq protocols was the discovery that certain tRNA base modifications could be removed prior to RT using the demethylating enzyme AlkB from *Escherichia coli*. AlkB was shown to remove 1-methyladenine and 3-methylcytosine damage in single- and double-stranded DNA (TREWICK *et al*. 2002) and later utilized by Zheng et al. (ZHENG *et al*. 2015) and Cozen et al. (COZEN *et al*. 2015) to remove some RT-inhibiting modifications present on tRNAs. In these studies, tRNA-seq libraries treated with AlkB had substantially higher abundance and diversity of tRNA reads. This was followed by the development of mutant forms of AlkB engineered to target specific modifications that appeared to be recalcitrant to demethylation by wild type

AlkB treatment (ZHENG *et al*. 2015; DAI *et al*. 2017). Although it is still not possible to remove all RT-inhibiting modifications with AlkB, the use of this enzyme represented a large step forward in tRNA sequencing.

In order to overcome inefficient adapter ligation, multiple methods have been developed involving modified adapter ligations (PANG *et al*. 2014a; SMITH *et al*. 2015; ZHONG *et al*. 2015; SHIGEMATSU *et al*. 2017) or a template-switching thermostable group II intron reverse transcriptase that eliminates the adapter ligation step entirely (DM-TGIRT) (ZHENG *et al*. 2015). In the case of YAMAT-seq (SHIGEMATSU *et al*. 2017), Y-shaped DNA/RNA hybrid adapters are utilized to specifically bind to the unpaired discriminator base and the -CCA sequence motif added to the 3′ end of all mature tRNAs. The utility of adapter protocols that utilize CCA-complementarity and ligation of both the 5′ and 3′ adapters to intact RNAs is that full-length, mature tRNAs are preferentially sequenced even from total RNA samples. In contrast, methods that ligate adapters in RNA/DNA step-wise fashion (PANG *et al*. 2014a) or eliminate adapter ligation in the case of DM-TGIRT have been effective at broadly detecting tRNA gene transcription, but largely capture tRNA fragments. These truncated sequence reads are then difficult to confidently map to a single gene, requiring additional predictive models and bioinformatic solutions (GOGAKOS *et al*. 2017; HOFFMANN *et al*. 2018; TORRES *et al*. 2019). Despite the progress being made with AlkB to remove inhibitory modifications and methods such as YAMAT-seq to specifically target full-length tRNAs, no study has yet applied both approaches to combine the benefits of each.

Although post-transcriptional tRNA modifications have hindered efforts to sequence tRNAs, they are a universal and fundamental aspect of tRNA biology (ROJAS-BENITEZ *et al*. 2015; LYONS *et al*. 2018). Base modifications are involved in tRNA folding and stability (SAMPSON AND UHLENBECK 1988; VERMEULEN *et al*. 2005; PHIZICKY AND ALFONZO 2010), translational accuracy and reading-frame maintenance (URBONAVICIUS *et al*. 2001; HOU *et al*. 2015), and tRNA fragment generation (LYONS *et al*. 2018; HUBER *et al*. 2019). Not surprisingly given these fundamental roles, they have been implicated in numerous diseases (PHIZICKY AND HOPPER 2015), and there has been a long historical interest to identify,

map, and quantify these modifications (KUCHINO *et al*. 1987). Traditional methods to identify base modifications involved tRNA purification and either direct sequencing and fingerprinting or complete digestion of RNAs into nucleosides followed by liquid chromatography-mass spectrometry (LC-MS) (KOWALAK *et al*. 1993; ROSS *et al*. 2016). Soon after the discovery of reverse transcriptase enzymes, signatures of RT inhibition were also recognized as a means to map base modification positions (YOUVAN AND HEARST 1979). Typical protocols involved primer extension assays and ^{32}P-labeled DNA primers to detect blocked or paused primer extension signals (SCHWARTZ AND MOTORIN 2017). RT-based RNA modification detection assays were later extended through the use of chemical reagents that reacted with specific modified nucleotides to further confirm the identity of the modifications (MOTORIN *et al*. 2007). The advent of high-throughput sequencing has now facilitated the generation of entire maps or "indexes" of the modifications that cause RT-misincorporations across all tRNAs found in a species (CLARK *et al*. 2016). These signatures include fragment generation because of RT termination but also read-through misincorporations, i.e. base misincorporation and indels produced from the RT enzyme reading through a modified base (KIETRYS *et al*. 2017; MOTORIN AND HELM 2019). Despite these advances, modification indexes have been produced for few species, and there has been limited investigation of tRNA modification profiles of isodecoders (tRNAs that share the same anticodon but have sequence variation at other parts along the tRNA). In addition, the function of some modifications is still unknown, and new modifications are still being discovered using a combination of RT-based modification mapping and mass spectrometry (KIMURA *et al*. 2019).

tRNA gene organization, expression, and modification patterns affect a wide diversity of biological processes, and the role of individual tRNAs has seen increasing interest in the past twenty years (HUMMEL *et al*. 2019). Species commonly have conserved multigene isoacceptor families (a group of tRNAs that are acylated with the same amino acid but can have different anticodons) (BERMUDEZ-SANTANA *et al*. 2010), and there is growing evidence that certain tRNA genes have functionally unique roles (KONDO *et al*. 1990; GOODARZI *et al*. 2016). One evolutionary lineage that has particularly complex

tRNA metabolism is angiosperms, or flowering plants. Plant nuclear tRNAs can be present in numerous copies, some of which are organized into large, tandemly repeated arrays (THEOLOGIS *et al*. 2000; COGNAT *et al*. 2013). Furthermore, the generation and function of tRNA-derived fragments (tRFs) has come under increasing scrutiny because their presence has been directly linked to plant cell growth and response to stresses (COGNAT *et al*. 2017; PARK AND KIM 2018; SOPRANO *et al*. 2018). Plants also have a high degree of tRNA compartmentalization and trafficking because of the presence of two endosymbiotically derived organelles, the plastid and the mitochondria. Each organelle has its own set of tRNAs, but plant mitochondria are exceptional compared to bilaterian animals in that they require extensive import of nuclear-encoded tRNAs into the mitochondrial matrix to maintain translation (SALINAS-GIEGE *et al*. 2015). Additionally, plants have been shown to have some of the most post-transcriptionally modified tRNAs identified to date (IIDA *et al*. 2009; MACHNICKA *et al*. 2014). These modifications are increasingly being assigned new biological roles, making plants an ideal group to study the function of tRNA modifications.

Unfortunately, the reasons that make plant tRNA biology so fascinating (i.e. complex metabolism and extreme modification rates) have historically hindered investigation because of difficulties in applying high throughput sequencing and mapping methods, and recent advances in tRNA-seq such as the use of AlkB and CCA-complementary adapters have not yet been applied to plant systems. Here, we have combined these two tRNA-seq methods to detect full-length tRNAs, map them to unique reference gene sequences and identify RT-induced sequence variants in four different angiosperm species, focusing on the model system *Arabidopsis thaliana*.

Results

YAMAT-seq captures mature, full-length tRNAs from the majority of A. thaliana tRNA reference genes

tRNAs were sequenced with a combination of two complementary methods, YAMAT-seq (SHIGEMATSU *et al*. 2017) and AlkB treatment (COZEN *et al*. 2015; ZHENG *et al*. 2015) to generate full-length, mature tRNA reads from *A. thaliana* leaf total-cellular RNA. The demethylating enzyme AlkB was used in conjunction with YAMAT-seq's Y-shaped adapters that are complementary to the CCA motif found at the 3′ end of mature tRNAs to generate Illumina libraries composed almost entirely of tRNA sequences, with representatives from most genes in the reference *A. thaliana* database (all sequences in the database can be found in supp. Table 2.1 with the corresponding tRNA gene identifiers). On average, 95% of all reads passed quality filters after trimming, and of those, greater than 99% of the reads mapped to a tRNA sequence when BLASTed to an *A. thaliana* tRNA database (e-value cutoff of 1e-6) (supp. Table 2.2). The remaining reads were largely derived from cytosolic/plastid rRNAs and adapter artifacts (supp. Table 2.3). Over 98% of the reads had a 3′-terminal CCA sequence. Although the proportion of reads that matched the reference sequence length varied among isoacceptor families (Figure 2.1), the vast majority of the reads were 74-95 nt after adapter trimming (representing the expected range of mature tRNA sequence lengths in *A. thaliana*), indicating that our library size selection successfully targeted full-length tRNAs. A small fraction of the reads, even those with a CCA tail, were truncated. These fragments (defined as having less than 70% hit coverage to a reference tRNA) appear to be largely generated from a subset of tRNAs, particularly certain nuclear tRNA-Thr and tRNA-Ala genes. Many of these fragments appeared to involve sequences that inverted and re-primed off of short regions of sequence similarity on the complementary strand, producing a "U-turn" molecule.

AlkB treatment improves detection of A. thaliana reference tRNAs

Base modifications known to stall or terminate RT are prevalent in tRNAs (MOTORIN *et al*. 2007). The demethylating enzyme AlkB has been shown to effectively remove some of these modifications (N^1-methyladenosine [m^1A], N^3-methylcytosine [m^3C]), thereby making certain tRNAs more amenable to

sequencing (COZEN *et al*. 2015; ZHENG *et al*. 2015). Our preliminary observations from testing AlkB treatments found a considerable effect on *A. thaliana* RNA integrity, with the majority of degradation occurring with exposure to the AlkB reaction buffer, regardless of whether the AlkB enzyme was included (supp. Figure 2.1). Even though this degradation predominantly affected larger transcripts and had little noticeable effect on the tRNA size fraction, we decided to use two types of negative controls in our experimental design to isolate the effects of AlkB from the buffer alone. Total RNA was either treated with the AlkB enzyme in the reaction buffer (AlkB+), treated with the reaction buffer alone (AlkB-), or left entirely untreated prior to RT. Libraries sequenced after the reaction buffer treatment alone and those that were entirely untreated did not substantially differ from each other in reference sequence detection and frequency (supp. Figure 2.2). However, treatment with AlkB resulted in better representation of the majority of cytosolic tRNAs and a moderate increase in the detection of some organellar tRNAs (Figure 2.2, supp. Table 2.4). Only 95-115 of the 183 nuclear tRNA reference genes were detected in untreated and AlkB- libraries (supp. Table 2.5), whereas AlkB+ treatment increased this range to 138-149. These reference coverage counts do not include tRNA reference genes that were exclusively detected with reads that were an equally good match to another reference sequence, which was the case for 4-13 genes per library (supp. Table 2.5). All 30 plastid and 17 of the 19 mitochondrial reference tRNAs were detected in at least one library with one mitochondrial tRNA-SerTGA gene only detected in a single AlkB+ library. No reads were detected for the mitochondrial genes tRNA-SerGCT-3360 and tRNA-SerGGA-3359. Ser and Tyr isoacceptors were the most likely to be undetected or have very low abundance in all libraries. Interestingly, three mitochondrial tRNA genes that are homologous to tRNA-PheGAA but have a mutated GTA (Tyr) anticodon and were previously annotated to be pseudogenes in the *A. thaliana* mitochondrial genome (GenBank: NC_037304.1) were detected as mature, expressed tRNAs in this analysis, which is consistent with earlier detection of expressed copies of one of these genes (CHEN *et al*. 1997). In addition, the plastid-derived tRNA-Met in the mitochondrial genome was detected in this study whereas previous studies using hybridization methods alone failed to detect expression of this gene (DUCHÊNE AND MARECHAL-DROUARD 2001).

In addition to the wild type AlkB, an AlkB mutant (D135S) has been specifically engineered to remove the modification N^1-methylguanosine (m^1G), which is known to inhibit RT activity (ZHENG *et al*. 2015). We performed tRNA-seq on three additional *A. thaliana* total-cellular RNA samples treated with either wild type AlkB, or a 2:1 ratio of wild type AlkB and D135S AlkB to test for further improvements in tRNA detection. We found only a moderate increase in the detection of a few genes when performing differential expression analysis on libraries treated with D135S (supp. Table 2.6). One D135S library did have a single read for a tRNA-TyrGTA gene that was undetected in the wild type libraries, but otherwise there was no increase in the number of genes detected in D135S libraries (supp. Figure 2.3).

tRNA-seq profiles are dominated by nuclear tRNA-Pro and plastid tRNA-GlyGCC genes in four angiosperm species

The number of reads mapped to each reference tRNA sequence varied drastically and was heavily skewed towards multiple nuclear tRNA-Pro genes and a plastid tRNA-GlyGCC (Figure 2.3). Together, nuclear tRNAs-Pro and the plastid tRNA-GlyGCC sequences comprised 86-90% of all reads in AlkB+ libraries and 91-93% in untreated and AlkB- libraries. To test whether this dominance of tRNA-Pro and tRNA-GlyGCC reads was unique to *A. thaliana* or a more widespread pattern in flowering plants, tRNAs were sequenced from leaf total-cellular RNA from another rosid (*Medicago truncatula*), an asterid (*Solanum tuberosum*) and a monocot (*Oryza sativa*), using the same tRNA-seq method described above with wild type AlkB. The resulting reads from all three species showed a similarly extreme skew towards nuclear tRNA-Pro and plastid tRNA-GlyGCC (Figure 2.4).

Persistent reverse transcription bias contributes to tRNA-seq coverage variation

The dominance of multiple nuclear tRNA-Pro genes and a plastid tRNA-GlyGCC gene was surprising because tRNA expression is generally expected to reflect codon usage after accounting for base-pairing modifications (NOVOA *et al*. 2012). One possible artifactual source of variation in tRNA-seq read

abundance is biased adapter ligation during library construction (FUCHS *et al*. 2015). In order to determine the abundance of tRNA-derived cDNA molecules independent of adapter ligation, droplet digital PCR (ddPCR) was performed on reverse transcribed, unligated subsamples of the three original *A. thaliana* RNA replicates using internal primers for four tRNA genes (supp. Table 2.7). There was a strong correlation between counts per million tRNA-seq reads (CPM) and ddPCR copies per nanogram ($p = 0.03$ and adjusted $R^2 = 0.91$; Figure 2.5). These data suggest that adapter ligation bias is not the primary determinant of tRNA-seq coverage skew as cDNA copy number in ddPCR is reflective of the number of final tRNA-seq Illumina reads.

An additional source of sequencing bias could result from preferential RT of tRNAs with fewer modifications or less inhibitory secondary structures. In order to quantify tRNA abundance in total-cellular RNA without the intermediate step of RT, northern blot analysis was performed by probing for four *A. thaliana* tRNAs representing a range of CPM values from the tRNA-seq data. All four labeled probes were hybridized to three RNA replicates as well as a dilution series of a complementary oligonucleotide in order to quantify the tRNA signal (supp. Table 2.8). The tRNA-Pro gene with the highest CPM showed the weakest hybridization signal, and the highly expressed plastid tRNA-GlyGCC did not have the strongest intensity of the two plastid tRNAs probed (Figure 2.6). The concentration estimates from the northern blot analysis present a truly striking contrast with the tRNA-seq data because the read abundance for the sampled tRNA-ProTGG was approximately 70,000-fold higher than one of the other sampled nuclear tRNAs (tRNA-SerCGA). This massive incongruence strongly points to biased tRNA RT (even after AlkB treatment) as contributing to the extreme coverage variation found in our tRNA-seq data.

Because tRNAs are multicopy genes with high sequence similarity, a "mismatch oligo" with one or two noncomplementary nucleotides (relative to the probe) was included on each membrane to test for probe specificity and cross-hybridization. There was clearly a signal of cross-hybridization to these mismatch oligos (which was expected given the permissive hybridization temperature of 48°C). Thus, the

signal for the RNA samples likely reflects some additional hybridization for other isodecoders with very similar sequence at the probe region. There are multiple isodecoders that are similar in sequence to tRNA-ProTGG-112 and tRNA-SerCGA-3245 at the probe region, and the mismatch oligos were designed to be identical to these similar isodecoders. There are no similar tRNAs to GlyGCC-3370 and IleGAT-3373 (the mismatch oligos for these genes were therefore "synthetic" with no biological match to another gene). Therefore, the already low hybridization signal for tRNA-ProTGG-112 and tRNA-SerCGA-3245 is likely an overestimation of expression, further exacerbating the incongruence between the northern blot analysis and the massive tRNA-seq CPM values for tRNA-ProTGG-112.

RT-induced misincorporations identify positions of base modifications in plant tRNAs

tRNA base modifications at the Watson-Crick face can interfere with the base pairing that is necessary for RT. Such modifications may not only stall or terminate reverse transcriptase activity but can also result in the misincorporation or deletion of nucleotides in the resulting cDNA. These misincorporations can be used to infer both the position and modification type, producing a modification map or "index" of a tRNA (CLARK *et al*. 2016; POTAPOV *et al*. 2018; VANDIVIER *et al*. 2019). Reads were globally aligned to reference sequences to identify all nucleotide positions that differed from the reference gene (supp. Table 2.9). Even after treatment with AlkB, a signal of RT misincorporation was still present in almost one-third of the tRNAs in at least one position. We identified a position as confidently modified if ≥30% of the mapped reads varied from the reference sequence at that position. There was evidence of multiple tRNAs being modified at the same site, with the same position being modified in up to 17% of all reference tRNAs (Figure 2.7). Given that AlkB may act on only a subset of modifications on certain bases (CLARK *et al*. 2016), it was unsurprising that positions with a modified T were largely insensitive to AlkB treatment (Table 2.1). Similar to work with RT-based modification detection in human cell lines and yeast (COZEN *et al*. 2015; ZHENG *et al*. 2015; CLARK *et al*. 2016), we found that demethylation treatment with

AlkB had a strong effect at only certain tRNA positions (e.g., 9, 26, and 58). The most frequent type of misincorporation differed by both position and reference base, but Gs were most likely to be deleted whereas the other three bases were most likely to be misread as a substitution (Table 2.2).

In order to ensure that the majority of reads were being effectively mapped to a reference sequence, all mapped reads in the AlkB treated library 1 were BLASTed to the entire nuclear genome of *A. thaliana* to check if reads had a better hit to a nuclear location than one of the reference sequences in the database. Only 1106 reads in the AlkB1 treated library (0.097%) mapped to a genomic location that was not within 5bp window of a reference sequence. Of those, only 182 reads (0.016%) would have passed the 90% reference coverage threshold to be used for modification analysis (supp. Table 2.10). Given that we applied a 30% threshold to identify modified sites, mismapping of reads derived from unannotated nuclear tRNA genes or tRNA-like sequences appears to have a negligible effect on these predictions.

Discussion

The promise and pitfalls of quantifying tRNA expression using tRNA-seq methods

In addition to the fundamental role of tRNAs as decoders of the genetic code, they are increasingly being recognized as integral players in a wide range of developmental, stress, tumorigenesis, biosynthetic, and amino acid delivery pathways (HOPPER AND PHIZICKY 2003; BANERJEE *et al*. 2010; PHIZICKY AND HOPPER 2010; RAINA AND IBBA 2014; KIRCHNER AND IGNATOVA 2015; WILUSZ 2015a; HUANG *et al*. 2018). As such, accurately detecting and quantifying tRNAs is key to gaining a more complete understanding of expression and regulation of numerous biological processes.

There have been substantial efforts in the past decade to make high-throughput sequencing methods more applicable to tRNAs. Nevertheless, we found that, even after treatment with AlkB and use of YAMAT adapters to specifically capture mature tRNAs, there still remained persistent RT-related

sequencing bias, as evidenced by the extreme skew in tRNA-seq reads towards certain sequences (Figures 2.3 and 2.4). The observed dominance of multiple nuclear tRNA-Pro genes and a plastid tRNA-GlyGCC likely does not reflect biological reality, as there was no agreement between the tRNA-seq read abundance estimates and the intensity of hybridization in northern blot analysis, where tRNAs are probed directly without an intervening RT step. Moreover, the massive inferred expression level of plastid tRNA-GlyGCC is at odds with previous 2D-gel analysis of purified chloroplast tRNAs, which did not identify tRNA-GlyGCC as unusually abundant (SELDEN *et al*. 1983; BERGMANN *et al*. 1984; MUBUMBILA *et al*. 1984). "Jackpot" tRNAs with extremely and presumably artefactually high read abundance have been reported in prior tRNA-seq results (PANG *et al*. 2014a; JACOB *et al*. 2019), and the original YAMAT-seq paper reported a single tRNA-LysCTT gene having approximately 10-fold higher expression than the next most highly expressed gene (SHIGEMATSU *et al*. 2017).

We found that four divergent angiosperm species showed the same general tRNA-seq profile dominated by nuclear tRNA-Pro and plastid tRNA-GlyGCC, suggesting an underlying cause of sequencing bias that is broadly shared across angiosperms. One possibility is that the overrepresented genes are less modified than other tRNAs. The plastid tRNA-GlyGCC and almost all of the nuclear tRNA-Pro isoacceptors had no RT misincorporations (based on our 30% threshold) after AlkB treatment. However, many other tRNAs lacked a strong signal for modification at any position (supp. Table 2.9) but did not show the same high abundance as nuclear tRNA-Pro and plastid tRNA-GlyGCC. The fact that the original YAMAT-seq data generated from human cell lines shows a skewed distribution of entirely different isoacceptors suggests that tRNAs from different eukaryotic lineages vary in their properties that make them more (or less) amenable to sequencing.

Modifications that predominantly result in the termination of RT ("RT falloff") without other signatures of RT inhibition (i.e. base misincorporations and indels) are not detectable by the YAMAT-seq method that we employed because molecules are not sequenced unless RT proceeds all the way through the tRNA and captures the 5′ adapter. Thus, variation across different tRNAs in the presence of

modifications that induce RT falloff could be responsible for observed RT bias. Additionally, the effect of certain modifications on RT behavior has been shown to be dependent on the nucleotide 3′ of modifications in the template RNA (HAUENSCHILD *et al*. 2015), suggesting that even if a modification is detected and identified, it may have a different effect on expression analysis depending on the tRNA sequence. However, a study on cancer type tRFs found no association with the 5′ ends of tRFs and known modified positions (TELONIS *et al*. 2019), suggesting that the presence of base modifications was not leading to truncated reads. Thus, the RT of tRNAs may then be affected by secondary structure (MOTORIN *et al*. 2007). In the cases of the dominance of certain isoacceptor reads in this study, there may be otherwise ubiquitous secondary structure characteristics that are not present in plastid tRNA-GlyGCC or multiple nuclear tRNA-Pro isoacceptors, making these tRNAs more amenable to RT. These represent important areas of investigation to further understand and alleviate sources of bias in tRNA-seq methods.

The biases that we have identified make it clear that more work must be done in developing a tRNA sequencing method that can accurately quantify *absolute* levels of expression. Nevertheless, the combination of YAMAT-seq and AlkB treatment has a number of advantages and great promise for analyzing changes in *relative* tRNA expression across treatments, tissues or subcellular fractions. In particular, we were able to generate a high proportion of full-length tRNA reads which can be confidently mapped to a single reference gene. Mapping of tRNA-seq reads to loci can be problematic because of the large number of similar but non-identical tRNA gene sequences. Other tRNA sequencing methods that hydrolyze tRNA (Hydro-tRNA-seq) (GOGAKOS *et al*. 2017), or utilize a template-switching reverse transcriptase [DM-TGIRT-seq] (ZHENG *et al*. 2015) produce a large proportion of tRNA fragments that can ambiguously map to multiple genes. In general, it is common for tRNA-seq methods to report overrepresentation of certain isoacceptors (PANG *et al*. 2014a; ZHENG *et al*. 2015; GOGAKOS *et al*. 2017; SHIGEMATSU *et al*. 2017; JACOB *et al*. 2019), suggesting sequencing bias is a frequent problem. However, the skew we observed with YAMAT-seq was especially extreme (Figure 2.3). Therefore, alternatives to YAMAT-seq may benefit from a reduced level of sequencing bias, albeit often at the expense of only

generating partial tRNA sequences. As there is increasing attention on the regulation and expression of specific tRNA genes (HUMMEL *et al*. 2019; TORRES *et al*. 2019), the combination of tRNA-seq methods that we employed could be effectively used for differential expression analysis when trying to tease apart the transcriptional activity and turnover of individual genes within large gene families.

It is important to note that plants, like many eukaryotes, have multiple copies of identical tRNA genes, and tRNA-seq cannot distinguish which copy, or copies, are expressed. This facet of tRNA expression is of growing interest in the regulation tRNA gene clusters such as the high-copy, tandemly repeated clusters of Ser, Tyr and Pro tRNA genes in *A. thaliana* (THEOLOGIS *et al*. 2000; HUMMEL *et al*. 2019). The database of tRNA references used in this study represent collapsed identical sequences of tRNA genes (see supp. Table 2.1 for all gene copy numbers by reference sequence). These identical copies can be found even across multiple genomes due to extensive organelle DNA insertions in the nucleus (STUPAR *et al*. 2001; HUANG *et al*. 2003). Although most intracellular gene transfers from organelles to the nuclear genome are thought to be nonfunctional, there is evidence in some eukaryotes that such organelle-derived sequences can be transcribed in the nucleus (TELONIS *et al*. 2014; TELONIS *et al*. 2015). In the *A. thaliana* Col-0 accession, all but one identical, organellar-derived tRNA gene found in the nuclear genome are located in a large 620-kb insertion of mitochondrial DNA on chromosome two, comprising almost the entire mitochondrial genome (STUPAR *et al*. 2001). This insertion is in a strongly heterochromatic, transcriptionally inactive region (SEQUEIRA-MENDES *et al*. 2014). There are other, nonidentical of organelle-derived located in other nuclear locations, but no tRNA-seq reads from this analysis were found to be better matches to any of those nuclear genes than to the organelle copy (supp. Table 2.10).

Looking forward, technologies that directly analyze RNA molecules without the intermediate step of RT, such as nanopore sequencing (JAIN *et al*. 2016), offer promise to accurately quantify tRNA expression without the confounding effects of RT bias. However, nanopore technologies still require substantial development to achieve sufficient sequencing accuracy and differentiation of tRNA species,

especially in the context of the extensive base modifications present in tRNAs (SMITH *et al*. 2015). In the interim, our combined approach may represent one of the more effective means to quantify relative tRNA expression changes at a single-gene level.

The landscape of base modifications in plant tRNAs

Although base modifications likely contribute to biased quantification of tRNA expression, there remains a large benefit of utilizing a RT-based tRNA-seq methods because RT misincorporation behavior provides insight into the location and identity of base modifications. In addition to modifications that terminate cDNA synthesis, RTs can read through methylations at the Watson-Crick face with varying efficiencies, resulting in misincorporations and indels in the cDNA (TSEROVSKI *et al*. 2016; PAN 2018; POTAPOV *et al*. 2018). Here, we took advantage of RT-based expression analysis to present one of the most extensive modification landscapes of angiosperm tRNAs to date. The modification peaks produced from sequencing complete *A. thaliana* tRNAs (Figure 2.7) provide an informative comparison to the annotated modification indexes previously generated for human tRNAs (CLARK *et al*. 2016) as well as in the recently published PRMdb database (MA *et al*. 2020) of predicted plant tRNA base modifications. Given the widely conserved presence of some tRNA modifications known to inhibit RT (JACKMAN AND ALFONZO 2013), it was unsurprising that the modification indexes from plants and humans shared many similarities, including a high rate of modification at nucleotide positions 9, 20, 26, 34, 37, and 58 (Figure 2.8). We reported modifications based on their actual position in tRNA genes. Thus, in some cases, the same modification at functionally analogous sites in different tRNAs can be represented by peaks spanning two or three nt positions due to differences in tRNA length.

Multiple software programs have been developed to predict RNA base modifications using misincorporation signatures in transcripts. For example, HAMR ((VANDIVIER *et al*. 2019)) has been used by the database PRMdb(MA *et al*. 2020) to predict modifications in plant RNAs. Many positions are

predicted to be modified by both HAMR and our analysis, including positions 9, 20, 26, 37, and 58 (supp. Figure 2.4). There do, however, exist some differences between the two methods. The modification index generated in this analysis has strong misincorporation signals in the variable loop (positions 46-48), which are missing from the PRMdb analysis. This may be partly due to the PRMdb tRNA database containing predominantly nuclear tRNAs, whereas our analysis used all nuclear, plastid and mitochondrial tRNAs to construct a modification index. In particular, plastid and mitochondrial tRNAs appear to be very commonly modified at positions 46-48 (supp. Figure 2.5). Even so, our analysis found multiple nuclear tRNA-Thr, tRNA-Tyr and tRNA-Asn isodecoders with misincorporation signals at positions 46 and 48 that were not reported by HAMR. Similarly, our dataset lacks modification predictions at positions 67-69, likely because of low coverage of some tRNA-Ser and tRNA-Leu isodecoders.

Previous work (ZHENG *et al*. 2015) has confirmed position 9 as frequently having an m^1G modification and position 58 having an m^1A modification. These modifications are known targets of AlkB demethylation, and we found that AlkB treatment resulted in almost complete reduction of the modification index peaks at positions 9 and 58 (some of this signal appears at position 57 in our modification index because of differences in tRNA length) (Figure 2.7), adding support for the existence of these same modifications in plants. Two consecutive dihydrouridines are commonly found in the D loop around position 20 in eukaryotes (HENDRICKSON 2001), a modification motif that may also be common in plant tRNAs, as we found a modification peak around positions 19-21. Position 26 is often modified to N2,N2-dimethylguanosine (m^2_2G), which also appears to be mildly sensitive to AlkB demethylation treatment (CLARK *et al*. 2016). We found a modification peak at position 26 and a small reduction in RT misincorporation rate after AlkB treatment in *A. thaliana*. Adenosine-to-inosine modifications are frequently found at position 34 in eukaryotic tRNAs (TORRES *et al*. 2015). RT of inosine is expected to cause A-to-G changes because inosine base pairs with C instead of T (MOTORIN *et al*. 2007). Indeed, we found almost 100% misincorporation of G when the reference was A at position 34.

It is important to note that some tRNA modifications may not be present in all copies (ARIMBASSERI *et al*. 2015) or may not always cause RT misincorporation (POTAPOV *et al*. 2018). Thus, our modification index with a cutoff of ≥30% misincorporation is focusing on highly modified positions in the corresponding tRNAs (CLARK *et al*. 2016).

Our analysis also identified clear differences in the modification index between plants vs. humans. In particular, we found a strong misincorporation peak at positions 46-48 for T reference nucleotides that was not observed in humans (Figure 2.8) (CLARK *et al*. 2016). The misread Ts at position 46-48, could correspond to dihydrouridines, which are commonly found in the variable loop (position 47 in standard tRNA nomenclature) of eukaryotic and plastid tRNAs (XING *et al*. 2004; LORENZ *et al*. 2017). Taken together, over 14% of all *A. thaliana* reference tRNAs appear to have a modified T at the position, including all cytosolic and plastid tRNA-Ile isoacceptors, as well as all cytosolic tRNA-ThrAGT, tRNA-AsnGTT, and tRNA-TyrGTA isodecoders. This trend of frequent modification at position 47 was conserved in all angiosperm species tested (supp. Figure 2.6, supp. Table 2.11). It is interesting to note that the misincorporation profile (deletions and T-to-C misincorporations) for dihydrouridines at positions 19-21 is similar to the misincorporation profiles that we observe for positions 46-48. Many *A. thaliana* tRNAs have a G in their reference sequence at position 46 and frequently exhibited a deletion at this site. This pattern likely reflects a modification to m^7G, which is predicted to be common in cytosolic Viridiplantae tRNAs at that position (MACHNICKA *et al*. 2014). Adding to this prediction is the complete lack of an AlkB effect on the modification peak at position 46, as AlkB is not expected to remove m^7G modifications. Our analysis demonstrates how tRNA-seq methods can be used to elucidate similarities and differences in global tRNA modification patterns between divergent taxonomic groups. These approaches could be extended to assess whether different tissues or even subcellular compartments depend more heavily on certain tRNA modifications for functionality.

In addition to generating a general modification index across the entire tRNA population, this tRNA-seq method can be used to identify the modification profiles of individual tRNAs or specific

groups of tRNAs. In other eukaryotes, organellar tRNAs have been found to be less modified than their cytosolic counterparts (MACHNICKA *et al*. 2014; SALINAS-GIEGE *et al*. 2015). In agreement with this, we detected fewer modifications on average in *A. thaliana* organellar tRNAs. Cytosolic tRNAs with a misincorporation signal had an average of 3.0 modifications in AlkB- libraries, whereas organelle tRNAs with misincorporations had an average of 2.1 modifications. Likely owing to greater rates of modification in cytosolic tRNAs, AlkB treatment had a substantially larger effect on improving the detection of cytosolic tRNAs than on organellar tRNAs (Figure 2.2).

It is sometimes assumed that tRNAs with the same anticodon but different body sequences will have similar modification patterns (TORRES *et al*. 2019). Nevertheless, we found examples of large differences in modification patterns among isodecoders. For example, the large nuclear tRNA-GlyGCC isodecoder family has a range of one to six inferred modifications depending on the reference sequence and show a range of sensitivity to AlkB treatment. That some tRNAs, even those with the same anticodon, have such radically different modification patterns represents an interesting and largely unexplored facet of tRNA biology, and offers support to the finding that isodecoders may differ in function (GESLAIN AND PAN 2010).

We found that AlkB treatment sometimes created novel RT misincorporations instead of restoring the correct nucleotide. For example, the C in the anticodon of the plastid tRNA-IleCAT is known to be modified to lysidine, which has base-paring behavior like that of U (KASHDAN AND DUDOCK 1982; ALKATIB *et al*. 2012), and we found this modification to be consistently misread as a deletion in AlkB+ libraries, whereas it was detected as a T, or even rarely as a C, in untreated libraries. Additionally, an increase in abundance of some tRNA fragments was found in AlkB+ libraries (supp. Table 2.12) and spatially associated with modified bases. For example, the "U-turn" fragments generated from tRNA-ThrCGT-3316 invert at what appears to be a modified G25, and the abundance of these fragments increased with AlkB treatment. In addition, treatment with AlkB as well as the AlkB reaction buffer had a

significant effect on RNA integrity (supp. Figure 2.1), meriting further investigation into the cause of RNA integrity loss and whether this degradation is system specific.

Historically, comprehensive mapping of base modifications has been laborious and depended on tRNA purification followed by digestion and mass spectrometry (SU *et al*. 2014). Quickly and accurately identifying tRNA modifications through RT-based methods makes it possible to test hypotheses regarding the presence and function of modifications at phylogenetic, developmental, and subcellular levels. tRNAs are by far the most post-transcriptionally modified gene class (BOCCALETTO *et al*. 2018). Research is just now starting to tease apart the many roles these modifications play (BJORK *et al*. 2001; PHIZICKY AND ALFONZO 2010; PEREIRA *et al*. 2018; DE CRECY-LAGARD *et al*. 2019). Not all modifications are detectable through sequencing, but it has been estimated that RT-based methods can identify approximately 25% of tRNA modifications in humans (CLARK *et al*. 2016), and the modifications that can be detected by sequencing are known to have critical roles in tRNA folding, rigidity and stability (PAN 2018). Given the diversity of tRNA modifications, the full scope of RT behavior and the effect of AlkB treatment on tRNA base modifications has yet to be fully explored, but work is currently being done to further characterize RT behavior when encountering specific modifications (EBHARDT *et al*. 2009; POTAPOV *et al*. 2018; WERNER *et al*. 2020). Moving beyond RT-based methods to direct sequencing of tRNAs and their modifications is another exciting development on the horizon. However, major advances must still be made in techniques such as nanopore sequencing because current technologies cannot be used to reliably sequence tRNA molecules in biological total RNA samples because they are likely unable to discriminate between more than two tRNAs at a time (HENLEY *et al*. 2016; POODARI 2019). Additionally, efforts to characterize the resistance properties of alternative tRNA base modifications in nanopores are underway (ONANUGA *et al*. 2017) but still in their infancy. Thus, along with the promise of additional technologies on the horizon, current RT-based methods offer easily accessible and exciting tools to study base modification profiles of full-length tRNA sequences and how they vary across taxa, treatments, tissues, and subcellular locations.

Methods

Plant material and growth conditions

For tRNA-seq material, Arabidopsis thaliana Columbia ecotype (Col-0) was grown in Pro-Mix BX General Purpose soil supplemented with vermiculite and perlite. Plants were germinated and kept in growth chambers at Colorado State University 23°C with a 16-hr/8-hr light/dark cycle (light intensity of 100 $\mu E \cdot m^{-2} \cdot s^{-1}$). *Solanum tuberosum* var. Gladstone (PI: 182477), *Oryza sativa* var. Ai-Chiao-Hong (PI:584576) and *Medicago truncatula* (PI:660383) were acquired from the U.S. National Plant Germplasm System (https://www.ars-grin.gov/npgs/). Tissue culture plantlets of *S. tuberosum* were transferred to Pro-Mix BX General Purpose soil supplemented with vermiculite and perlite and kept in growth chambers with the same settings as above. Seeds of *O. sativa,* and *M. truncatula* were germinated in SC7 Cone-tainers (Stuewe and Sons) on a mist bench under supplemental lighting (16-hr/8-hr light/dark cycle) in the Colorado State University greenhouse, then moved to a growth chamber with the same settings as above after 4 weeks. For northern blot analysis, *A. thaliana* Col-0 was grown in a 12-hr/12-hr light/dark cycle in growth chambers at the Institute of Molecular Biology of Plants.

RNA extraction

For our primary tRNA-seq experiment and ddPCR analysis, RNA from *A. thaliana* and *O. sativa* was extracted from 7-week-old leaf tissue. RNA from *M. truncatula* was extracted from 3-week-old leaf tissue, and RNA from *S. tuberosum* plantlets was extracted from leaf tissue 3 weeks after transfer into soil. Extractions were performed using a modified version of the Jordon-Thaden et al. (JORDON-THADEN *et al*. 2015) protocol. In brief, tissue was frozen in liquid nitrogen and pulverized with a mortar and pestle and then vortexed and centrifuged with 900 µl of hexadecyltrimethylammonium bromide (CTAB) lysis buffer supplemented with 1% polyvinylpyrrolidone (PVP) and 0.2% β-mercaptoethanol (BME). Samples were then centrifuged, and the aqueous solution was removed. 800 µL of a chloroform:isoamyl alcohol

(24:1) solution was added, mixed by inverting and centrifuged. The aqueous phase was removed, and 900 µl of TRIzol was added. The solution was then mixed by inversion and centrifuged. The aqueous phase was removed, 200 µl of chloroform was added, and the solution was then mixed and centrifuged again. The aqueous phase was removed followed by isopropanol RNA precipitation, and cleaned pellets were resuspended in dH_2O. RNA was checked for integrity with a TapeStation 2200 and purity on a Nanodrop 2000. Later RNA extractions from *A. thaliana* for the AlkB D135S tRNA-seq experiment (8-week-old tissue) and northern blot analysis (4-week-old tissue) were performed with a simplified protocol. Leaf tissue was frozen in liquid nitrogen and pulverized with a mortar and pestle prior to the addition of TRIzol but otherwise followed the TRIzol manufacturer's RNA extraction protocol. Three biological replicates (different plants) were used for the *A. thaliana* experiments, and a single sample was used for each of the three other angiosperms.

AlkB purification

Plasmids containing cloned wild type AlkB protein (pET24a-AlkB deltaN11 [plasmid #73622]) and D135S mutant protein (pET30a-AlkB-D135S [plasmid #79051]) were obtained from Addgene (http://www.addgene.org/). Protein was expressed and purified at CSU Biochemistry and Molecular Biology Protein Expression and Purification Facility. Cells were grown at 37°C to an OD_{600} of 0.6, at which time, 1 mM of isopropyl-β-D-thiogalactoside was added, and temperature was lowered to 30°C. Cells were harvested after 3 hr by centrifugation and resuspended in 10 mM Tris (pH 7.3), 2 mM $CaCl_2$, 300 mM NaCl, 10 mM $MgCl_2$, 5% glycerol, and 1 mM BME. Resuspension was homogenized by sonication, and lysate was recovered by centrifugation. The supernatant was loaded onto HisTrap HP 5 ml columns (GE Healthcare) and was washed and eluted by a linear gradient of 0-500 mM imidazole. The fractions containing AlkB were pooled and concentrated with ultrafiltration using Amicon Ultra-15 MWCO 10kDa (Millipore). The concentrated sample was then loaded onto HiLoad 16/60 Superdex 200

prep grade size exclusion column (GE Healthcare) in 20 mM Tris pH 8.0, 200 mM NaCl, 2 mM DTT, and 10% glycerol. The fractions containing AlkB were pooled and concentrated, aliquoted, flash-frozen in the presence of 20% glycerol and stored at -70°C.

Demethylation reaction

AlkB reactions were performed using a modified version of existing protocols (COZEN *et al*. 2015; ZHENG *et al*. 2015; CHEN *et al*. 2016). Demethylation was performed by treating 10 µg of total cellular RNA with 400 pmols of AlkB in a reaction volume of 80 µl containing: 70 µM ammonium iron(II) sulfate hexahydrate (Sigma-Aldrich, 203505), 0.93 mM α-ketoglutaric acid disodium salt dihydrate (Sigma-Aldrich, 75892), 1.86 mM ascorbic acid (EMD, AX1775-3), and 46.5 mM HEPES (pH 8.0-HCl, Corning, 61-034-RO), incubated at 37°C for 60 min in 0.2 ml PCR strip tubes (Fisherbrand) in a preheated thermal cycler (Bio-Rad, C1000 Touch). The reaction was quenched by adding 4 µl of 100 mM EDTA (Invitrogen, AM9262) followed by a phenol-chloroform RNA extraction, ethanol precipitation with the addition of 0.08 µg of RNase-free glycogen (Thermo Scientific, R0551), and resuspension in RNase-free water (Invitrogen, 10-977-023). Prior to the reaction, the AlkB enzyme was diluted to 100 pmol/µl with a dilution buffer containing 10 mM Tris-HCl pH 8.0 (National Diagnostics, EC-406), 100 mM NaCl (Fisher, S271-3), and 1 mM BME (Acros Organics, 125472500). The reaction buffer was prepared fresh with reagents added in the order listed above. The diluted AlkB enzyme was added to the reaction buffer followed by the RNA, and the reaction was brought to the final volume using RNase-free dH_2O. RNA integrity was checked on a TapeStation 2200. The same procedure was followed in parallel for AlkB- control libraries except that the AlkB enzyme in the reaction volume was replaced with dH_2O.

Illumina tRNA-seq library construction

All adapter and primer sequences used in library construction can be found in supp. Table 2.13. In order to remove amino acids from the mature tRNAs (deacylation), demethylated or control RNA was incubated in 20 mM Tris HCl (pH 9.0) at 37°C for 40 min. Following deacylation, adapter ligation was performed using a modified protocol from (SHIGEMATSU *et al*. 2017). A 9 µl reaction volume containing 1 µg of deacylated RNA and 1 pmol of each Y-5′ adapter (4 pmols total) and 4 pmols of the Y-3′ adapter was incubated at 90°C for 2 min. 1 µl of an annealing buffer containing 500 mM Tris-HCl (pH 8.0) and 100 mM $MgCl_2$ was added to the reaction mixture and incubated for 15 min at 37°C. Ligation was performed by adding 1 unit of T4 RNA Ligase 2 enzyme (New England Biolabs) in 10 µl of 1X reaction buffer and incubating the reaction at 37°C for 60 min, followed by overnight incubation at 4°C. We found that adapter ligation with the input of 80 pmols of adapters called for in the original protocol resulted in an excess of adapter dimers and other adapter-related products after PCR and that using $1/10^{th}$ of that adapter concentration maximized yield in the expected size range for ligated tRNA products, while minimizing adapter-related products.

RT of ligated RNA was performed using SuperScript IV (Invitrogen) according to the manufacturer's protocol. Briefly, 1 µl of 2 µM RT primer, and 1 µl of 10 mM dNTP mix was added to 11 µl of the deacylated RNA from each sample. The mixture was briefly vortexed, centrifuged and incubated at 65°C for 5 min. Then, 4 µl of 5X SSIV buffer, 1 µl 100 mM DTT, 1 µl RNaseOUT, and 1 µl of SuperScriptIV were added to each reaction. The mixture was then incubated for 10 min at 55°C for RT and inactivated by incubating at 80°C for 10 min.

The resulting cDNA was then amplified by polymerase chain reaction (PCR) in a 50 µl reaction containing 7 µl of the RT reaction, 25 µl of the NEBNext 2X PCR Master Mix, 1 µl of the PCR forward primer, 1 µl of the PCR reverse primer, and 15.5 µl dH_2O. Ten cycles of PCR were performed on a Bio-Rad C1000 Touch thermal cycler with an initial 1 min incubation at 98°C and 10 cycles of 30 s at 98°C, 30 s at 60°C and 30 s at 72°C, followed by 5 min at 72°C.

Size selection of the resulting PCR products was done on a BluePippin (Sage Science) with Q3 3% agarose gel cassettes following the manufacturer's protocol. The size selection parameters were set to a range of 180-231 bp, with a target of 206 bp. Size-selected products were then cleaned using solid phase reversable immobilization beads and resuspended in 10 mM Tris (pH 8.0).

Sequencing and read processing

The nine original *A. thaliana* tRNA-seq libraries (three AlkB+, three AlkB-, and three entirely untreated) were single-indexed and sequenced on an Illumina MiSeq with single-end, 150 bp reads. Libraries from *M. truncatula*, *O. sativa, S. tuberosum*, and the *A. thaliana* D135S AlkB mutant and wild type AlkB libraries were dual-indexed and sequenced on an Illumina NovaSeq 6000 with paired-end, 150-base pair reads. Sequencing reads are available via the NCBI Sequence Read Archive under BioProject PRJNA562543. Adapters were trimmed using Cutadapt version 1.16 (MARTIN 2011) with a quality-cutoff parameter of 10 for the 3′ end of each read. A minimum length filter of 5 bp was applied to reads from the MiSeq sequencing platform. Read length filters of a minimum of 50 bp and a maximum of 95 bp were applied to reads produced from the NovaSeq 6000, as a much larger percentage of the reads from those libraries were <20 bp after adapter trimming. For paired-end data, BBMerge from the BBTools software package was used to merge R1 and R2 read pairs into a consensus sequence (BUSHNELL *et al*. 2017). Identical reads were summed and collapsed into read families using the FASTQ/A Collapser tool from the FASTX-Toolkit version 0.0.13 (http://hannonlab.cshl.edu/fastx_toolkit/index.html).

Sequence alignments and mapping

Mapping for all processed and collapsed reads was performed with a custom Perl script using reference tRNA databases from the PlantRNA website (http://seve.ibmp.unistra.fr/plantrna/, downloaded 9/8/2018). The database for *A. thaliana* comprises 579 nuclear tRNA genes, 23 mitochondrial tRNAs, and 37 plastid

tRNAs that collapse down to 232 unique read families. This database is manually curated and updated to reflect current tRNA biology in model plant species. Each collapsed read family was BLASTed (blastn, 1e-6) against a complete set of (non-identical) nuclear, mitochondrial and plastid reference tRNA sequences from the corresponding species, and the best match was assigned to each read family. The read count of all assigned read families was then summed for each reference sequence. When a read family was an equally good match to multiple references, the read count was divided by the total number of tied references. Multiple BLAST statistics were recorded for each read family/reference hit pair including e-value, hit score, hit length, and percent identity. The read family sequences, tRNA identifier and the gene copy number can be found for every database sequence in supp. Table 2.1.

Because we were mainly interested in mature tRNA sequences, only reads that had 80% or greater hit coverage to a reference sequence were used in downstream analysis. Complete datasets including reads that fell below the 80% hit coverage threshold were used to determine read length distributions and the proportion of reads that did not BLAST to a tRNA reference sequence.

Efficiency of mapping was calculated by BLASTing the first AlkB treated library (AlkB1) to all *A. thaliana* chromosomes and custom Perl scripts were used to check for reads that had a better hit (raw bit score) to any genomic location than the mapped reference sequence. The coordinates of the genomic positions of the hit of any reads with a better hit were then compared to the genomic positions of the reference sequences using a custom Perl script.

Generation of modification index

In order to capture all possible misincorporations and indels including those that could have occurred at the end of the reads, a custom Perl script and the alignment software MAFFT version 7.407 (KATOH *et al*. 2017) was used to produce full-length alignments of reads to their top BLAST reference tRNA hit and identify all base substitutions and indels in the resulting alignment. Only reads that BLASTed to their

reference sequence with an e-value of 1e-10 or less and had 90% hit coverage to the reference were retained for modification index calling. Only positions that were covered by greater than five reads were used for modification index calling, and a position was considered confidently modified if ≥30% of the mapped reads at that position differed from the reference nucleotide by either a substation or deletion. Scripts used for read processing, mapping, and sequence analysis are available via GitHub (https://github.com/warrenjessica/YAMAT-scripts).

The modification index of predicted tRNA modifications from PRMdb (MA *et al*. 2020) was generated by downloading all tRNA modification data for *Arabidopsis thaliana* found at http://www.biosequencing.cn/PRMdb/index.html (downloaded on 5/5/2020). The total number of tRNAs with a modification at each position was summed and divided by the total number of tRNAs in that database.

ddPCR

To generate cDNA for ddPCR quantification, 1 µg of RNA from each of the three original AlkB+ *A. thaliana* samples was treated with DNase I (Invitrogen) according to the manufacturer's protocol. For each sample, 146 ng of DNase-treated RNA was then subjected to RT using iScript Reverse Transcription Supermix (BioRad) according to manufacturer's protocol in a 20-µl reaction volume. cDNA was then diluted in a four step concentration series to 0.001 ng/µl, 0.01 ng/µl, 0.1 ng/µl, and 1 ng/µl. ddPCR was performed with each of the four primer pairs (supp. Table 2.6) using this concentration series of template cDNA to determine ideal concentration. All ddPCR amplifications were set up in 20-µL volumes with Bio-Rad QX200 ddPCR EvaGreen Supermix and a 2 µM concentration of each primer before mixing into an oil emulsion with a Bio-Rad QX200 Droplet Generator. The final template input amounts were 0.1 ng for amplification of tRNA-GlyGCC-3370 and tRNA-CysGCA-3354, 0.001 ng for amplification of tRNA-ProTGG-112, and 1 ng for amplification of tRNA-GlyCCC-93. Amplification was performed on a Bio-

Rad C1000 Touch Thermal Cycler with an initial 5 min incubation at 95°C and 40 cycles of 30 s at 95°C and 1 min at 60°C, followed by signal stabilization via 5 min at 4°C and 5 min at 95°C. The resulting droplets were read on a Bio-Rad QX200 Droplet Reader. Copy numbers for each PCR target were calculated based on a Poisson distribution using the Bio-Rad QuantaSoft package.

Northern blots

RNAs and synthesized oligonucleotides (supp. Table 2.7) were separated on 15% (w/v) polyacrylamide gels. Gels were then electrotransfered onto Hybond-N^+ nylon membranes (Amersham) for 90 min at 300mA/250V and UV-crosslinked. All membranes were hybridized with ^{32}P-radiolabeled oligonucleotide probes (supp. Table 2.7) at 48°C in a 6X saline-sodium citrate (SSC) buffer with 0.5% sodium dodecyl sulphate (SDS) solution for 14 hr. Hybridization was followed by two washes (10 min) with 2X SSC buffer at 48°C buffer followed by one wash (30 min) at 48°C in 2X SSC with 0.1% SDS. Membranes were imaged on a Typhoon FLA 9500 biomolecular imager (GE Healthcare Life Sciences) after 16 hr of exposure on film.

Northern blots were analyzed using ImageJ, version Java 1.8.0_172 (64-bit) (https://imagej.nih.gov/ij/) to quantify signal intensity for each band. Signal peaks were quantified by first drawing a straight (0.00 degree) line beginning at the farthest left point of contact of the signal to the opposite side of the window. To define only the major signal peak of interest, two lines were drawn perpendicular to the horizontal. The position of these lines was determined by eye based on where the peak's curves began to smooth. The coordinates of these vertical boundaries were used for all subsequent plots of the same membrane. The area was then calculated between the horizontal line and the peak signal bounded by the two vertical lines. Using the known concentration of the loaded oligos and the corresponding signal area, the estimated pmols of each tRNA target was calculated by fitting a standard curve to the $\log_{10}$ values of signal intensity and oligo concentration (supp. Table 2.7).

Statistics and figure generation

Differential expression analysis between the AlkB treated and untreated libraries in the original analysis as well as the wild type and D135S mutant AlkB libraries was done with the R package EdgeR version 3.24.3 (ROBINSON *et al*. 2010), using a dataset with only reads that had 100% hit coverage to a reference sequence. A linear regression was performed in R with the lm function to test the relationship between tRNA-seq CPM values and ddPCR copy-number estimates per ng of input RNA.

Funding and support

We thank Deyu Li for providing valuable guidance in optimizing the AlkB reaction and all of the research personnel at the Institute of Molecular Biology of Plants for graciously hosting J.M.W. We would also like to thank Dr. Carol Wilusz and Adam Heck for test RNA samples and valuable discussions, and Dr. John Hunt and Dr. Victor Wong at Columbia University for sending samples of AlkB for initial testing. Two anonymous reviewers provided helpful comments on an earlier version of this manuscript. This work was supported by a Chateaubriand Fellowship from the Embassy of France in the United States, an NSF GAUSSI graduate research fellowship (DGE-1450032), NSF award IOS-1829176, and start-up funds from Colorado State University.

Data availability statement

Supplementary tables are available at the publisher's website for this article (same title) and can be accessed at https://www.tandfonline.com/doi/abs/10.1080/15476286.2020.1792089?journalCode=krnb20.

Sequencing reads are available via the NCBI Sequence Read Archive under BioProject PRJNA562543.

Table 2.1| Total proportion of reference nucleotides misread ≥30% of the time in *Arabidopsis thaliana* libraries.

		Reference			
		A	C	G	T
Library	AlkB+ (rep 1)	0.0187	0.0028	0.0147	0.0149
	AlkB+ (rep 2)	0.0168	0.0028	0.0113	0.0140
	AlkB+ (rep 3)	0.0136	0.0011	0.0100	0.0112
	AlkB- (rep 1)	0.0350	0.0038	0.0241	0.0100
	AlkB- (rep 2)	0.0361	0.0032	0.0258	0.0089
	AlkB- (rep 3)	0.0299	0.0036	0.0224	0.0098
	Untreated (rep 1)	0.0347	0.0038	0.0237	0.0112
	Untreated (rep 2)	0.0294	0.0032	0.0213	0.0086
	Untreated (rep 3)	0.0339	0.0041	0.0213	0.0100

Table 2.2| Total misincorporation frequency of tRNA reference bases across all nine *Arabidopsis thaliana* tRNA-seq libraries (see Table 2.1).

		Reference			
		A	**C**	**G**	**T**
Misincorporation	**A**	**-**	**0.141**	**0.077**	**0.037**
	C	**0.041**	**-**	**0.075**	**0.275**
	G	**0.294**	**0.15**	**-**	**0.315**
	T	**0.535**	**0.508**	**0.209**	**-**
	Deletion	**0.129**	**0.202**	**0.639**	**0.373**

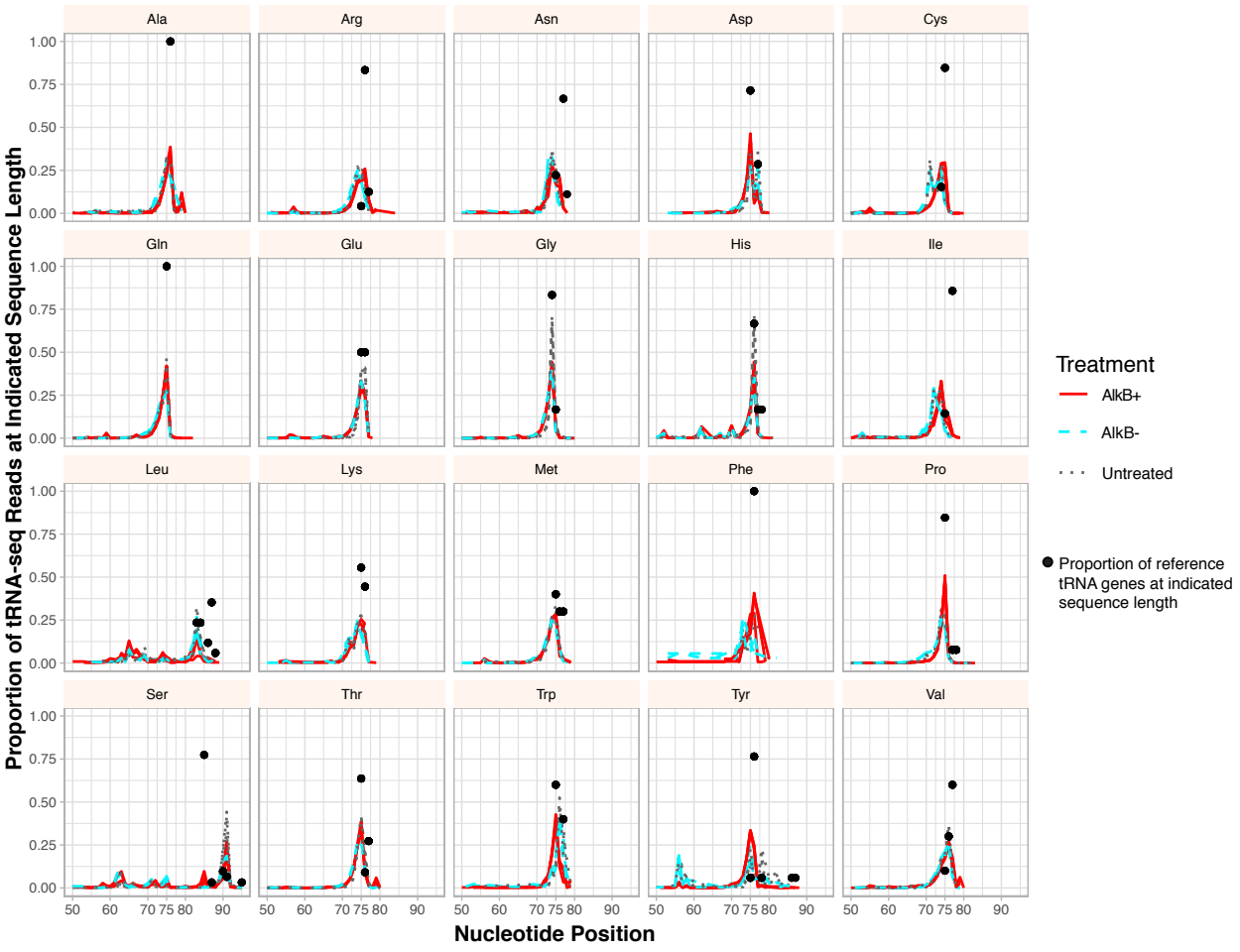


Figure 2.1| The proportion of tRNA-seq reads at different lengths based on isoacceptor family. Black dots represent the proportion of isoacceptor reference sequences at indicated length (e.g. 100% of the Ala-tRNA reference genes are 76 nt in length). Three biological replicates are plotted for each treatment (AlkB+, AlkB- and untreated). Note that the dominant peak for read length often corresponds to the length of the reference sequence(s).

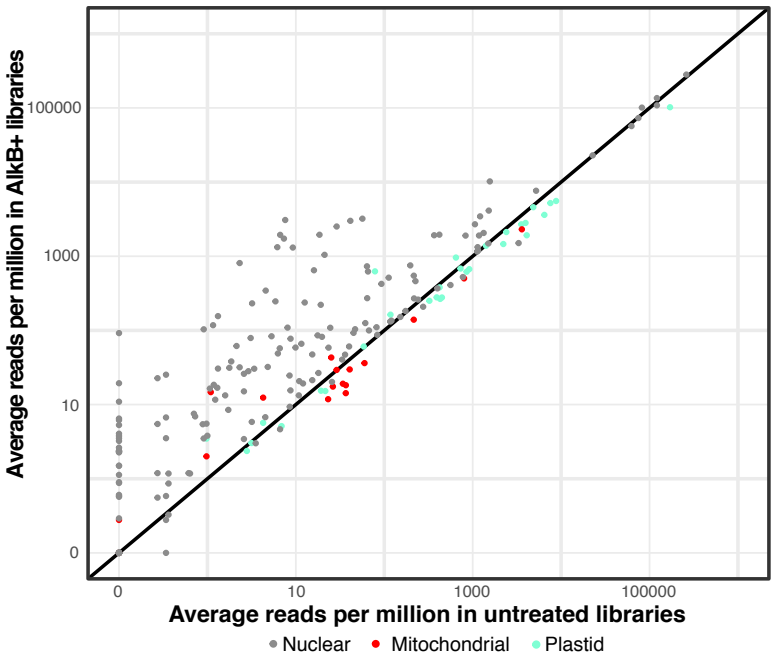


Figure 2.2| Demethylation treatment with AlkB increases the proportion of tRNA-seq reads for many tRNAs (points falling above the 1:1 line). Average read counts per million across three biological replicates are shown for each unique *A. thaliana* tRNA reference sequence in AlkB+ vs. untreated libraries.

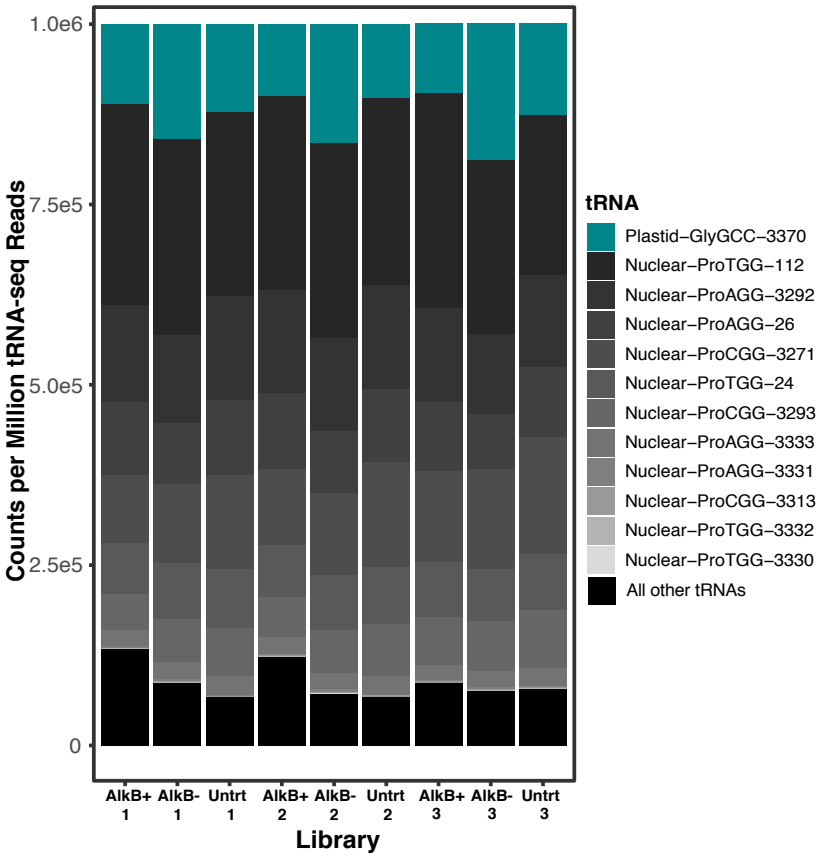


Figure 2.3| tRNA-seq read abundance in *A. thaliana* tRNA is dominated by nuclear tRNA-Pro and plastid tRNA-GlyGCC genes. All nuclear tRNA-Pro genes are shown and indicated by gray colors, with the darkest grays indicating the highest abundance, and the plastid tRNA-GlyGCC is indicated by teal. All other tRNAs have been grouped and shown in black.

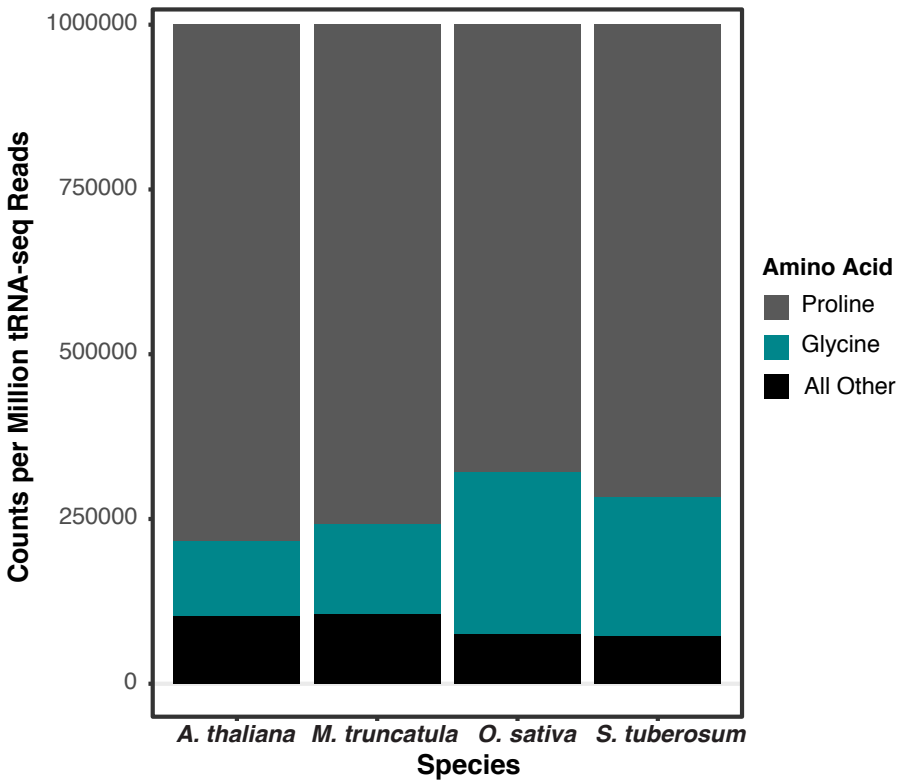


Figure 2.4| tRNA-seq read abundance is similarly dominated by only two isoacceptor families (tRNA-Pro and tRNA-Gly) in four angiosperm species.

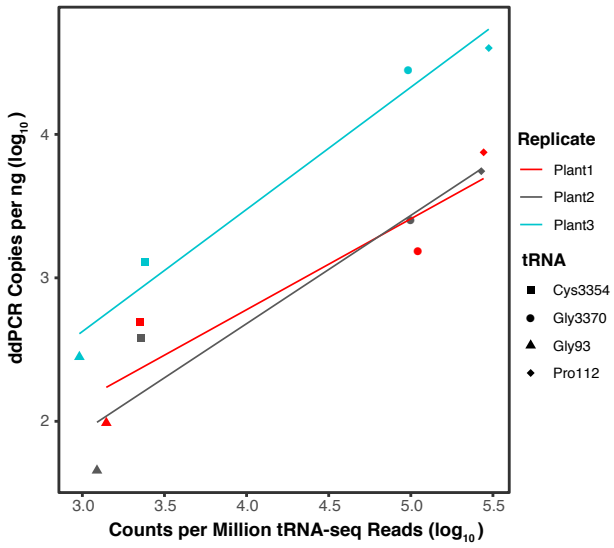


Figure 2.5| Droplet digital PCR (ddPCR) copies correlate with number of tRNA-seq reads. ddPCR copies per ng of cDNA plotted against counts per million tRNA-seq reads for four *A. thaliana* tRNA genes. Adjusted R^2 values for separate linear regressions on biological replicates 1, 2, and 3 were 0.79, 0.85, and 0.96, respectively. When data points were averaged across biological replicates, linear regression yielded an adjusted R^2 of 0.91 and a *p*-value of 0.03.

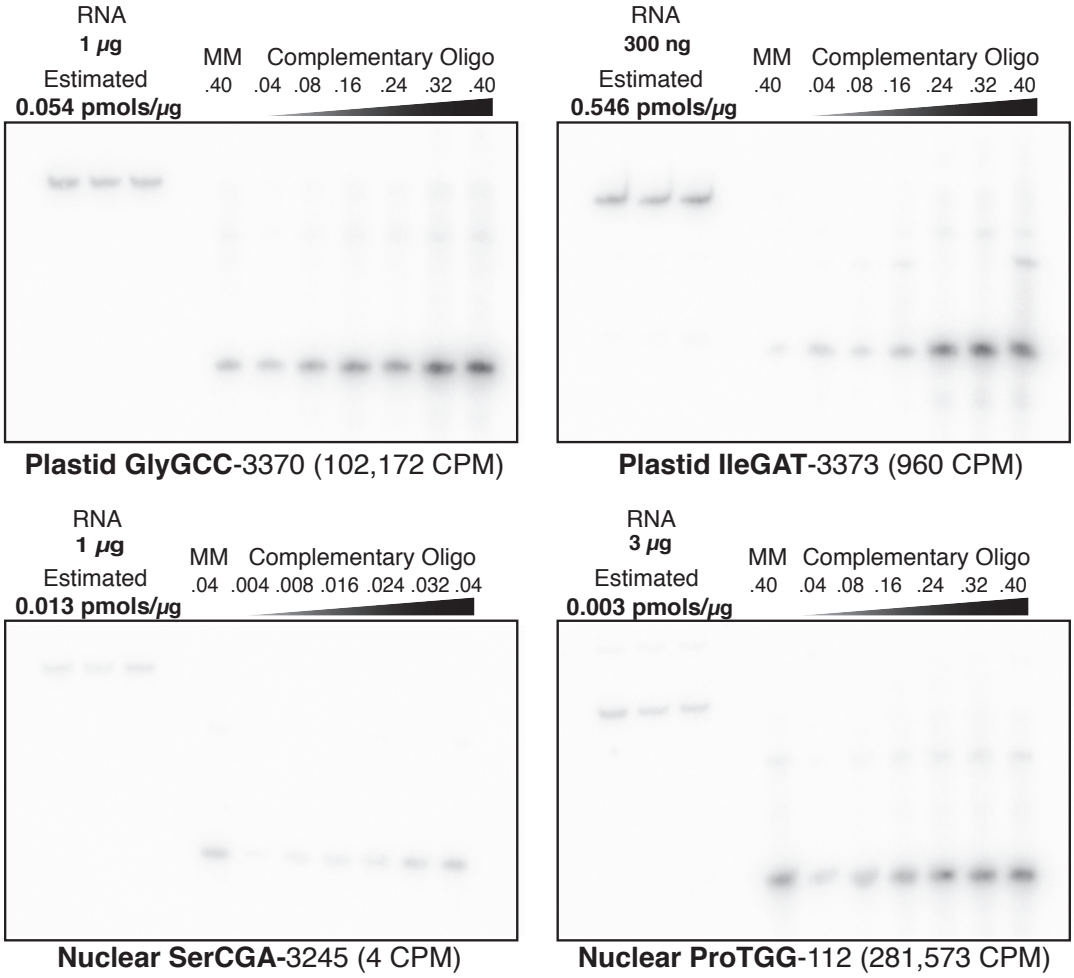


Figure 2.6| Northern blot analysis does not show the same expression dominance of nuclear tRNA-Pro and plastid tRNA-GlyGCC that was observed with tRNA-seq. Four different *A. thaliana* tRNA genes were probed from total cellular RNA and quantified. Each tRNA target membrane had three replicates of total *A. thaliana* RNA, which was quantified using a dilution series of a synthesized 38-nt oligonucleotide complementary to the corresponding probe. A single lane of a 38-nt mismatched oligonucleotide (MM), which had one or two non-complementary nucleotides relative to the probe, was also included on each membrane to test for cross hybridization of probes. The amounts of total RNA and oligonucleotides loaded on each blot were varied according to expected hybridization signal based on preliminary analyses. The mass of RNA and pmols of oligonucleotides are indicated above each corresponding lane. The estimated concentration of the target tRNA based on analysis of signal intensity in ImageJ is reported as pmols per μg of input RNA. The average Illumina tRNA-seq read counts per million (CPM) is indicated parenthetically for each tRNA.

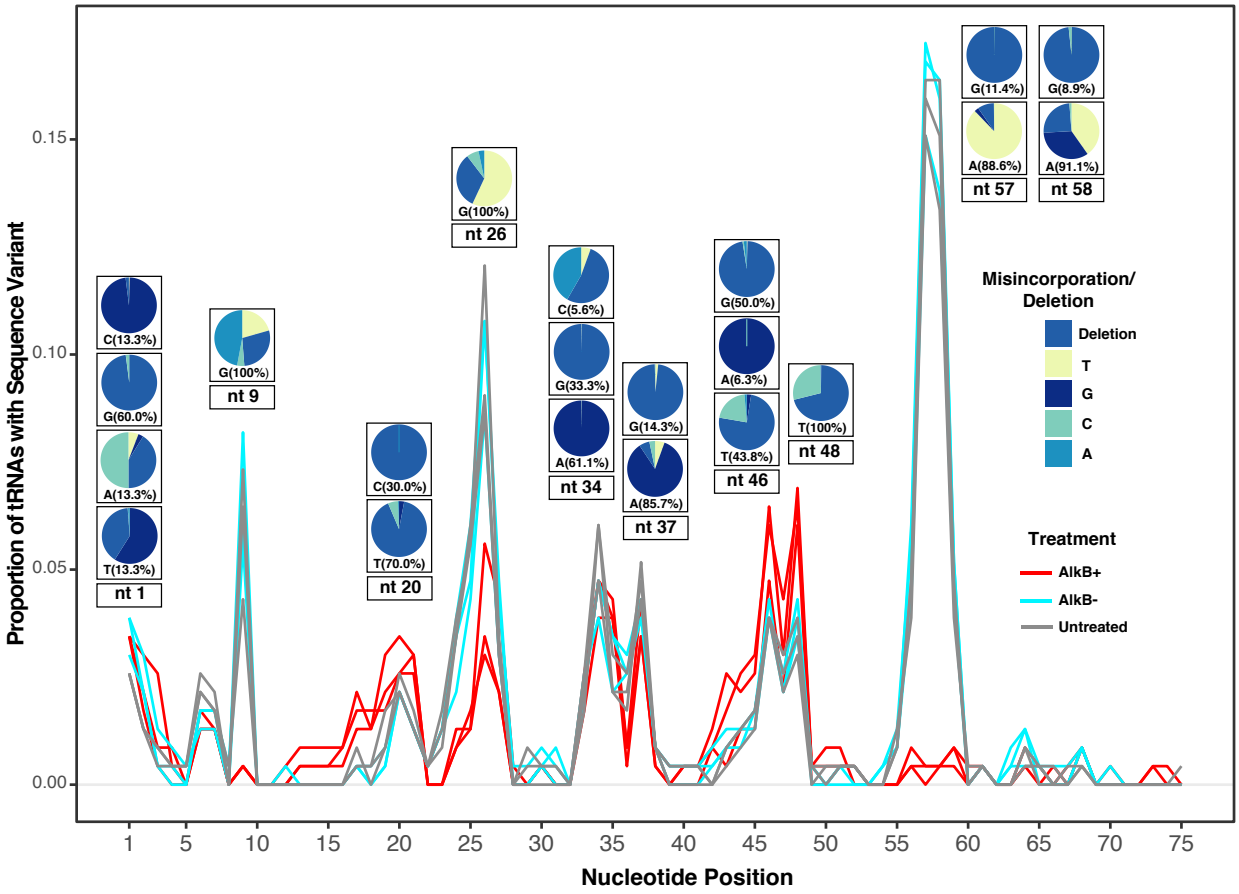


Figure 2.7| A tRNA modification index showing the proportion of all *A. thaliana* tRNA reference sequences with a misincorporation/deletion at each nucleotide position. A position was considered modified if ≥30% of the mapped reads differed from the reference sequence and the sequence was detected by more than five reads. Pie charts show the identity of the misincorporated base at some of the most frequently modified positions. A separate pie chart is provided for each observed reference base at that position, and the percentage indicates what proportion of modified tRNA sequences have that reference nucleotide. All three replicates of each treatment are shown and indicated by line color. Because the y-axis is expressed as a proportion of tRNA reference sequences (and not total reads), genes such as nuclear tRNA-Pro and plastid tRNA-GlyGCC are not preferentially weighted in this analysis even though they represent the majority of reads.

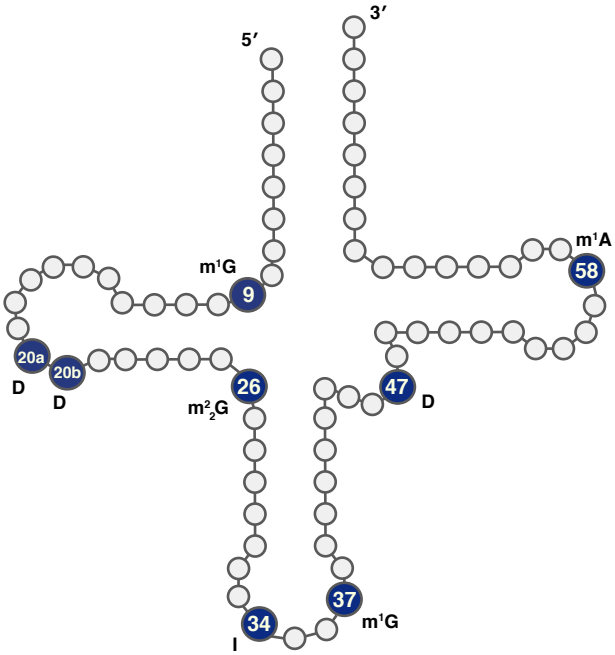


Figure 2.8| tRNA positions labeled with possible modifications. Positions with strong misincorporation signals are indicated with a known modification at that position in other species. D: dihydrouridine; I: inosine; m^1A: N1-methyladenosine; m^1G: N1-methylguanosine; m^2_2G: N2,N2-dimethylguanosine.

LITERATURE CITED

Adams, K. L., and J. D. Palmer, 2003 Evolution of mitochondrial gene content: gene loss and transfer to the nucleus. Mol Phylogenet Evol 29**:** 380-395.

Adams, K. L., Y. L. Qiu, M. Stoutemyer and J. D. Palmer, 2002 Punctuated evolution of mitochondrial gene content: high and variable rates of mitochondrial gene loss and transfer to the nucleus during angiosperm evolution. Proc Natl Acad Sci U S A 99**:** 9905-9912.

Adrion, J. R., P. S. White and K. L. Montooth, 2016 The Roles of Compensatory Evolution and Constraint in Aminoacyl tRNA Synthetase Evolution. Mol Biol Evol 33**:** 152-161.

Alfonzo, J. D., V. Blanc, A. M. Estévez, M. A. T. Rubio and L. Simpson, 1999 C to U editing of the anticodon of imported mitochondrial tRNATrp allows decoding of the UGA stop codon in Leishmania tarentolae. The EMBO journal 18**:** 7056-7062.

Alfonzo, J. D., and D. Söll, 2009 Mitochondrial tRNA import--the challenge to understand has just begun. Biol Chem 390**:** 717-722.

Alkatib, S., T. T. Fleischmann, L. B. Scharff and R. Bock, 2012 Evolutionary constraints on the plastid tRNA set decoding methionine and isoleucine. Nucleic Acids Res 40**:** 6713-6724.

Almagro Armenteros, J. J., M. Salvatore, O. Emanuelsson, O. Winther, G. von Heijne *et al.*, 2019 Detecting sequence signals in targeting peptides using deep learning. Life Sci Alliance 2.

Altman, S., 2007 A view of RNase P. Mol Biosyst 3**:** 604-607.

Alverson, A. J., X. X. Wei, D. W. Rice, D. B. Stern, K. Barry *et al.*, 2010 Insights into the Evolution of Mitochondrial Genome Size from Complete Sequences of Citrullus lanatus and Cucurbita pepo (Cucurbitaceae). Molecular Biology and Evolution 27**:** 1436-1448.

Anderson, S., A. T. Bankier, B. G. Barrell, M. H. de Bruijn, A. R. Coulson *et al.*, 1981 Sequence and organization of the human mitochondrial genome. Nature 290**:** 457-465.

Arimbasseri, A. G., N. H. Blewett, J. R. Iben, T. N. Lamichhane, V. Cherkasova *et al.*, 2015 RNA Polymerase III Output Is Functionally Linked to tRNA Dimethyl-G26 Modification. PLoS Genet 11**:** e1005671.

Arimura, S., J. Yamamoto, G. P. Aida, M. Nakazono and N. Tsutsumi, 2004 Frequent fusion and fission of plant mitochondria with unequal nucleoid distribution. Proc Natl Acad Sci U S A 101**:** 7805-7808.

Arita, M., T. Suematsu, A. Osanai, T. Inaba, H. Kamiya *et al.*, 2006 An evolutionary 'intermediate state' of mitochondrial translation systems found in Trichinella species of parasitic nematodes: co-evolution of tRNA and EF-Tu. Nucleic Acids Res 34**:** 5291-5299.

Arrieta-Montiel, M. P., and S. A. Mackenzie, 2011 Plant Mitochondrial Genomes and Recombination, pp. 65-82 in *Plant Mitochondria*, edited by F. Kempken. Springer New York, New York, NY.

Artimo, P., M. Jonnalagedda, K. Arnold, D. Baratin, G. Csardi *et al.*, 2012 ExPASy: SIB bioinformatics resource portal. Nucleic Acids Res 40**:** W597-603.

Ayan, G. B., H. J. Park and J. Gallie, 2020 The birth of a bacterial tRNA gene by large-scale, tandem duplication events. Elife 9.

Baker, A., and G. Schatz, 1987 Sequences from a prokaryotic genome or the mouse dihydrofolate reductase gene can restore the import of a truncated precursor protein into yeast mitochondria. Proc Natl Acad Sci U S A 84**:** 3117-3121.

Banerjee, R., S. Chen, K. Dare, M. Gilreath, M. Praetorius-Ibba *et al.*, 2010 tRNAs: cellular barcodes for amino acids. FEBS Lett 584**:** 387-395.

Berg, M., R. Rogers, R. Muralla and D. Meinke, 2005 Requirement of aminoacyl-tRNA synthetases for gametogenesis and embryo development in Arabidopsis. Plant J 44**:** 866-878.

Berglund, A. K., E. Spanning, H. Biverstahl, G. Maddalo, C. Tellgren-Roth *et al.*, 2009 Dual targeting to mitochondria and chloroplasts: characterization of Thr-tRNA synthetase targeting peptide. Mol Plant 2**:** 1298-1309.

Bergmann, P., P. Seyer, G. Burkard and J. H. Weil, 1984 Mapping of transfer RNA genes on tobacco chloroplast DNA. Plant Mol Biol 3**:** 29-36.

Bergthorsson, U., K. L. Adams, B. Thomason and J. D. Palmer, 2003 Widespread horizontal transfer of mitochondrial genes in flowering plants. Nature 424**:** 197-201.

Bermudez-Santana, C., C. S. Attolini, T. Kirsten, J. Engelhardt, S. J. Prohaska *et al.*, 2010 Genomic organization of eukaryotic tRNAs. BMC Genomics 11**:** 270.

Betat, H., Y. Long, J. E. Jackman and M. Morl, 2014 From end to end: tRNA editing at 5'- and 3'-terminal positions. Int J Mol Sci 15**:** 23975-23998.

Binder, S., A. Marchfelder and A. Brennicke, 1994 RNA editing of tRNA(Phe) and tRNA(Cys) in mitochondria of Oenothera berteriana is initiated in precursor molecules. Mol Gen Genet 244**:** 67-74.

Binder, S., A. Marchfelder, A. Brennicke and B. Wissinger, 1992 RNA editing in trans-splicing intron sequences of nad2 mRNAs in Oenothera mitochondria. J Biol Chem 267**:** 7615-7623.

Bjork, G. R., K. Jacobsson, K. Nilsson, M. J. Johansson, A. S. Bystrom *et al.*, 2001 A primordial tRNA modification required for the evolution of life? EMBO J 20**:** 231-239.

Boccaletto, P., M. A. Machnicka, E. Purta, P. Piatkowski, B. Baginski *et al.*, 2018 MODOMICS: a database of RNA modification pathways. 2017 update. Nucleic Acids Res 46**:** D303-D307.

Bonnard, G., A. Gobert, M. Arrivé, F. Pinker, T. Salinas-Giegé *et al.*, 2016 Transfer RNA maturation in Chlamydomonas mitochondria, chloroplast and the nucleus by a single RNase P protein. Plant J 87**:** 270-280.

Boore, J. L., J. R. Macey and M. Medina, 2005 Sequencing and comparing whole mitochondrial genomes of animals. Methods Enzymol 395**:** 311-348.

Brandao, M. M., and M. C. Silva-Filho, 2011 Evolutionary history of Arabidopsis thaliana aminoacyl-tRNA synthetase dual-targeted proteins. Mol Biol Evol 28**:** 79-85.

Brandvain, Y., and M. J. Wade, 2009 The functional transfer of genes from the mitochondria to the nucleus: the effects of selection, mutation, population size and rate of self-fertilization. Genetics 182**:** 1129-1139.

Brindefalk, B., J. Viklund, D. Larsson, M. Thollesson and S. G. Andersson, 2007 Origin and evolution of the mitochondrial aminoacyl-tRNA synthetases. Mol Biol Evol 24**:** 743-756.

Brown, W. M., M. George, Jr. and A. C. Wilson, 1979 Rapid evolution of animal mitochondrial DNA. Proc Natl Acad Sci U S A 76**:** 1967-1971.

Broz, A. K., G. Waneka, Z. Wu, M. Fernandes Gyorfy and D. B. Sloan, 2021 Detecting de novo mitochondrial mutations in angiosperms with highly divergent evolutionary rates. Genetics 218.

Brubacher-Kauffmann, S., L. Marechal-Drouard, A. Cosset, A. Dietrich and A. M. Duchêne, 1999 Differential import of nuclear-encoded tRNAGly isoacceptors into solanum Tuberosum mitochondria. Nucleic Acids Res 27**:** 2037-2042.

Bruce, B. D., 2001 The paradox of plastid transit peptides: conservation of function despite divergence in primary structure. Biochim Biophys Acta 1541**:** 2-21.

Burger, G., M. W. Gray, L. Forget and B. F. Lang, 2013 Strikingly bacteria-like and gene-rich mitochondrial genomes throughout jakobid protists. Genome Biol Evol 5**:** 418-438.

Bushnell, B., J. Rood and E. Singer, 2017 BBMerge - Accurate paired shotgun read merging via overlap. Plos One 12.

Calvo, S. E., and V. K. Mootha, 2010 The mitochondrial proteome and human disease. Annu Rev Genomics Hum Genet 11**:** 25-44.

Canino, G., E. Bocian, N. Barbezier, M. Echeverria, J. Forner *et al.*, 2009 Arabidopsis encodes four tRNase Z enzymes. Plant Physiol 150**:** 1494-1502.

Carrie, C., and I. Small, 2013 A reevaluation of dual-targeting of proteins to mitochondria and chloroplasts. Biochim Biophys Acta 1833**:** 253-259.

Ceci, L. R., P. Veronico, L. Siculella and R. Gallerani, 1995 Identification and mapping of trnI, trnE and trnfM genes in the sunflower mitochondrial genome. DNA Seq 5**:** 315-318.

Chan, P. P., and T. M. Lowe, 2019 tRNAscan-SE: Searching for tRNA Genes in Genomic Sequences. Methods Mol Biol 1962**:** 1-14.

Chapdelaine, Y., and L. Bonen, 1991 The wheat mitochondrial gene for subunit I of the NADH dehydrogenase complex: a trans-splicing model for this gene-in-pieces. Cell 65**:** 465-472.

Charriere, F., S. Helgadottir, E. K. Horn, D. Söll and A. Schneider, 2006 Dual targeting of a single tRNA(Trp) requires two different tryptophanyl-tRNA synthetases in Trypanosoma brucei. Proc Natl Acad Sci U S A 103**:** 6847-6852.

Charriere, F., P. O'Donoghue, S. Helgadottir, L. Maréchal-Drouard, M. Cristodero *et al.*, 2009 Dual targeting of a tRNAAsp requires two different aspartyl-tRNA synthetases in Trypanosoma brucei. J Biol Chem 284**:** 16210-16217.

Chaw, S. M., A. C. Shih, D. Wang, Y. W. Wu, S. M. Liu *et al.*, 2008 The mitochondrial genome of the gymnosperm Cycas taitungensis contains a novel family of short interspersed elements, Bpu sequences, and abundant RNA editing sites. Mol Biol Evol 25**:** 603-615.

Chen, F., Q. Tang, K. Bian, Z. T. Humulock, X. Yang *et al.*, 2016 Adaptive Response Enzyme AlkB Preferentially Repairs 1-Methylguanine and 3-Methylthymine Adducts in Double-Stranded DNA. Chem Res Toxicol 29**:** 687-693.

Chen, H. C., M. Viry-Moussaid, A. Dietrich and H. Wintz, 1997 Evolution of a mitochondrial tRNA PHE gene in A. thaliana: import of cytosolic tRNA PHE into mitochondria. Biochem Biophys Res Commun 237**:** 432-437.

Chien, C. I., Y. W. Chen, Y. H. Wu, C. Y. Chang, T. L. Wang *et al.*, 2014 Functional substitution of a eukaryotic glycyl-tRNA synthetase with an evolutionarily unrelated bacterial cognate enzyme. PLoS One 9**:** e94659.

Chimnaronk, S., M. Gravers Jeppesen, T. Suzuki, J. Nyborg and K. Watanabe, 2005 Dual-mode recognition of noncanonical tRNAs(Ser) by seryl-tRNA synthetase in mammalian mitochondria. EMBO J 24**:** 3369-3379.

Christensen, A. C., A. Lyznik, S. Mohammed, C. G. Elowsky, A. Elo *et al.*, 2005 Dual-domain, dual-targeting organellar protein presequences in Arabidopsis can use non-AUG start codons. Plant Cell 17**:** 2805-2816.

Christian, R. W., S. L. Hewitt, G. Nelson, E. H. Roalson and A. Dhingra, 2020 Plastid transit peptides-where do they come from and where do they all belong? Multi-genome and pan-genomic assessment of chloroplast transit peptide evolution. PeerJ 8**:** e9772.

Clark, W. C., M. E. Evans, D. Dominissini, G. Zheng and T. Pan, 2016 tRNA base methylation identification and quantification via high-throughput sequencing. RNA 22**:** 1771-1784.

Clary, D. O., and D. R. Wolstenholme, 1984 The Drosophila mitochondrial genome. Oxf Surv Eukaryot Genes 1**:** 1-35.

Coffin, J. W., and E. A. Cossins, 1986 Mitochondrial folates and methionyl-tRNA transformylase activity during germination and early growth of seeds. Phytochemistry 25**:** 2481-2487.

Cognat, V., G. Morelle, C. Megel, S. Lalande, J. Molinier *et al.*, 2017 The nuclear and organellar tRNA-derived RNA fragment population in Arabidopsis thaliana is highly dynamic. Nucleic Acids Res 45**:** 3460-3472.

Cognat, V., G. Pawlak, A. M. Duchene, M. Daujat, A. Gigant *et al.*, 2013 PlantRNA, a database for tRNAs of photosynthetic eukaryotes. Nucleic Acids Res 41**:** D273-279.

Costello, R., D. M. Emms and S. Kelly, 2020 Gene Duplication Accelerates the Pace of Protein Gain and Loss from Plant Organelles. Mol Biol Evol 37**:** 969-981.

Cozen, A. E., E. Quartley, A. D. Holmes, E. Hrabeta-Robinson, E. M. Phizicky *et al.*, 2015 ARM-seq: AlkB-facilitated RNA methylation sequencing reveals a complex landscape of modified tRNA fragments. Nature Methods 12**:** 879-+.

Cummings, M. P., J. M. Nugent, R. G. Olmstead and J. D. Palmer, 2003 Phylogenetic analysis reveals five independent transfers of the chloroplast gene rbcL to the mitochondrial genome in angiosperms. Curr Genet 43**:** 131-138.

Dai, Q., G. Zheng, M. H. Schwartz, W. C. Clark and T. Pan, 2017 Selective Enzymatic Demethylation of N(2) ,N(2) -Dimethylguanosine in RNA and Its Application in High-Throughput tRNA Sequencing. Angew Chem Int Ed Engl 56**:** 5017-5020.

Darty, K., A. Denise and Y. Ponty, 2009 VARNA: Interactive drawing and editing of the RNA secondary structure. Bioinformatics 25**:** 1974-1975.

de Crecy-Lagard, V., P. Boccaletto, C. G. Mangleburg, P. Sharma, T. M. Lowe *et al.*, 2019 Matching tRNA modifications in humans to their known and predicted enzymes. Nucleic Acids Res 47**:** 2143-2159.

Delage, L., A. Dietrich, A. Cosset and L. Marechal-Drouard, 2003 In vitro import of a nuclearly encoded tRNA into mitochondria of Solanum tuberosum. Mol Cell Biol 23**:** 4000-4012.

Dewe, J. M., J. M. Whipple, I. Chernyakov, L. N. Jaramillo and E. M. Phizicky, 2012 The yeast rapid tRNA decay pathway competes with elongation factor 1A for substrate tRNAs and acts on tRNAs lacking one or more of several modifications. RNA 18**:** 1886-1896.

Dietrich, A., J. H. Weil and L. Maréchal-Drouard, 1992 Nuclear-encoded transfer RNAs in plant mitochondria. Annu Rev Cell Biol 8**:** 115-131.

Doolittle, R. F., and J. Handy, 1998 Evolutionary anomalies among the aminoacyl-tRNA synthetases. Curr Opin Genet Dev 8**:** 630-636.

Doyle, J. J., and J. L. Doyle, 1987 A rapid DNA isolation procedure for small quantities of fresh leaf tissue. PHYTOCHEMICAL BULLETIN v.19(1):11-15.

Duchêne, A. M., A. Giritch, B. Hoffmann, V. Cognat, D. Lancelin *et al.*, 2005a Dual targeting is the rule for organellar aminoacyl-tRNA synthetases in Arabidopsis thaliana. Proc Natl Acad Sci U S A 102**:** 16484-16489.

Duchêne, A. M., A. Giritch, B. Hoffmann, V. Cognat, D. Lancelin *et al.*, 2005b Dual targeting is the rule for organellar aminoacyl-tRNA synthetases in Arabidopsis thaliana. Proceedings of the National Academy of Sciences of the United States of America 102**:** 16484-16489.

Duchêne, A. M., and L. Marechal-Drouard, 2001 The chloroplast-derived trnW and trnM-e genes are not expressed in Arabidopsis mitochondria. Biochem Biophys Res Commun 285**:** 1213-1216.

Duchêne, A. M., N. Peeters, A. Dietrich, A. Cosset, I. D. Small *et al.*, 2001 Overlapping destinations for two dual targeted glycyl-tRNA synthetases in Arabidopsis thaliana and Phaseolus vulgaris. Journal of Biological Chemistry 276**:** 15275-15283.

Duchêne, A. M., C. Pujol and L. Marechal-Drouard, 2009 Import of tRNAs and aminoacyl-tRNA synthetases into mitochondria. Curr Genet 55**:** 1-18.

Ebhardt, H. A., H. H. Tsang, D. C. Dai, Y. Liu, B. Bostan *et al.*, 2009 Meta-analysis of small RNA-sequencing errors reveals ubiquitous post-transcriptional RNA modifications. Nucleic Acids Res 37**:** 2461-2470.

Esseiva, A. C., A. Naguleswaran, A. Hemphill and A. Schneider, 2004 Mitochondrial tRNA import in Toxoplasma gondii. Journal of Biological Chemistry 279**:** 42363-42368.

Esser, K., U. Lüttge, W. Beyschlag and J. Murata, 2006 *Progress in Botany 68*. Springer Berlin Heidelberg.

Fey, J., A. Dietrich, A. Cosset, T. Desprez and L. Marechal-Drouard, 1997 Evolutionary aspects of "chloroplast-like" trnN and trnH expression in higher-plant mitochondria. Curr Genet 32**:** 358-360.

Forner, J., B. Weber, S. Thuss, S. Wildum and S. Binder, 2007 Mapping of mitochondrial mRNA termini in Arabidopsis thaliana: t-elements contribute to 5' and 3' end formation. Nucleic Acids Res 35**:** 3676-3692.

Fu, L., B. Niu, Z. Zhu, S. Wu and W. Li, 2012 CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics 28**:** 3150-3152.

Fuchs, R. T., Z. Sun, F. Zhuang and G. B. Robb, 2015 Bias in ligation-based small RNA sequencing library construction is determined by adaptor and RNA structure. PLoS One 10**:** e0126049.

Gandini, C. L., and M. V. Sanchez-Puerta, 2017 Foreign Plastid Sequences in Plant Mitochondria are Frequently Acquired Via Mitochondrion-to-Mitochondrion Horizontal Transfer. Sci Rep 7**:** 43402.

Ge, C., E. Spanning, E. Glaser and A. Wieslander, 2014 Import determinants of organelle-specific and dual targeting peptides of mitochondria and chloroplasts in Arabidopsis thaliana. Mol Plant 7**:** 121-136.

Geslain, R., and T. Pan, 2010 Functional analysis of human tRNA isodecoders. J Mol Biol 396**:** 821-831.

Ghifari, A. S., M. Gill-Hille and M. W. Murcha, 2018 Plant mitochondrial protein import: the ins and outs. Biochem J 475**:** 2191-2208.

Giege, R., M. Sissler and C. Florentz, 1998 Universal rules and idiosyncratic features in tRNA identity. Nucleic Acids Research 26**:** 5017-5035.

Glover, K. E., D. F. Spencer and M. W. Gray, 2001 Identification and structural characterization of nucleus-encoded transfer RNAs imported into wheat mitochondria. J Biol Chem 276**:** 639-648.

Gobert, A., B. Gutmann, A. Taschner, M. Gossringer, J. Holzmann *et al.*, 2010 A single Arabidopsis organellar protein has RNase P activity. Nat Struct Mol Biol 17**:** 740-744.

Gogakos, T., M. Brown, A. Garzia, C. Meyer, M. Hafner *et al.*, 2017 Characterizing Expression and Processing of Precursor and Mature Human tRNAs by Hydro-tRNAseq and PAR-CLIP. Cell Rep 20**:** 1463-1475.

Goodarzi, H., H. C. B. Nguyen, S. Zhang, B. D. Dill, H. Molina *et al.*, 2016 Modulated Expression of Specific tRNAs Drives Gene Expression and Cancer Progression. Cell 165**:** 1416-1427.

Goremykin, V. V., P. J. Lockhart, R. Viola and R. Velasco, 2012 The mitochondrial genome of Malus domestica and the import-driven hypothesis of mitochondrial genome expansion in seed plants. Plant J 71**:** 615-626.

Grabherr, M. G., B. J. Haas, M. Yassour, J. Z. Levin, D. A. Thompson *et al.*, 2011 Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol 29**:** 644-652.

Gray, M. W., 2012 Mitochondrial evolution. Cold Spring Harb Perspect Biol 4**:** a011403.

Greiner, S., P. Lehwark and R. Bock, 2019 OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. Nucleic Acids Res 47**:** W59-W64.

Grewe, F., P. Viehoever, B. Weisshaar and V. Knoop, 2009 A trans-splicing group I intron and tRNA-hyperediting in the mitochondrial genome of the lycophyte Isoetes engelmannii. Nucleic Acids Res 37**:** 5093-5104.

Grosjean, H., and E. Westhof, 2016 An integrated, structure- and energy-based view of the genetic code. Nucleic Acids Res 44**:** 8020-8040.

Guo, M., and P. Schimmel, 2013 Essential nontranslational functions of tRNA synthetases. Nat Chem Biol 9**:** 145-153.

Guo, W., F. Grewe, W. Fan, G. J. Young, V. Knoop *et al.*, 2016 Ginkgo and Welwitschia Mitogenomes Reveal Extreme Contrasts in Gymnosperm Mitochondrial Evolution. Mol Biol Evol 33**:** 1448-1460.

Guo, W., A. Zhu, W. Fan, R. P. Adams and J. P. Mower, 2020 Extensive Shifts from Cis- to Trans-splicing of Gymnosperm Mitochondrial Introns. Mol Biol Evol 37**:** 1615-1620.

Guo, W., A. Zhu, W. Fan and J. P. Mower, 2017 Complete mitochondrial genomes from the ferns Ophioglossum californicum and Psilotum nudum are highly repetitive with the largest organellar introns. New Phytol 213**:** 391-403.

Gutmann, B., A. Gobert and P. Giegé, 2012 PRORP proteins support RNase P activity in both organelles and the nucleus in Arabidopsis. Genes Dev 26**:** 1022-1027.

Haen, K. M., W. Pett and D. V. Lavrov, 2010a Parallel Loss of Nuclear-Encoded Mitochondrial Aminoacyl-tRNA Synthetases and mtDNA-Encoded tRNAs in Cnidaria. Molecular Biology and Evolution 27**:** 2216-2219.

Haen, K. M., W. Pett and D. V. Lavrov, 2010b Parallel loss of nuclear-encoded mitochondrial aminoacyl-tRNA synthetases and mtDNA-encoded tRNAs in Cnidaria. Mol Biol Evol 27**:** 2216-2219.

Hancock, K., and S. L. Hajduk, 1990 The mitochondrial tRNAs of Trypanosoma brucei are nuclear encoded. J Biol Chem 265**:** 19208-19215.

Hao, W., and J. D. Palmer, 2009 Fine-scale mergers of chloroplast and mitochondrial genes create functional, transcompartmentally chimeric mitochondrial genes. Proc Natl Acad Sci U S A 106**:** 16728-16733.

Hauenschild, R., L. Tserovski, K. Schmid, K. Thuring, M. L. Winz *et al.*, 2015 The reverse transcription signature of N-1-methyladenosine in RNA-Seq is sequence dependent. Nucleic Acids Res 43**:** 9950-9964.

Hecht, J., F. Grewe and V. Knoop, 2011 Extreme RNA editing in coding islands and abundant microsatellites in repeat sequences of Selaginella moellendorffii mitochondria: the root of frequent plant mtDNA recombination in early tracheophytes. Genome Biol Evol 3**:** 344-358.

Hendrickson, T. L., 2001 Recognizing the D-loop of transfer RNAs. Proc Natl Acad Sci U S A 98**:** 13473-13475.

Henley, R. Y., S. Carson and M. Wanunu, 2016 Studies of RNA Sequence and Structure Using Nanopores. Prog Mol Biol Transl Sci 139**:** 73-99.

Hoffmann, A., J. Fallmann, E. Vilardo, M. Morl, P. F. Stadler *et al.*, 2018 Accurate mapping of tRNA reads. Bioinformatics 34**:** 2339.

Holley, R. W., J. Apgar, G. A. Everett, J. T. Madison, M. Marquisee *et al.*, 1965 Structure of a Ribonucleic Acid. Science 147**:** 1462-1465.

Holzmann, J., P. Frank, E. Loffler, K. L. Bennett, C. Gerner *et al.*, 2008 RNase P without RNA: identification and functional reconstitution of the human mitochondrial tRNA processing enzyme. Cell 135**:** 462-474.

Hopper, A. K., 2013 Transfer RNA post-transcriptional processing, turnover, and subcellular dynamics in the yeast Saccharomyces cerevisiae. Genetics 194**:** 43-67.

Hopper, A. K., and R. T. Nostramo, 2019 tRNA Processing and Subcellular Trafficking Proteins Multitask in Pathways for Other RNAs. Front Genet 10**:** 96.

Hopper, A. K., and E. M. Phizicky, 2003 tRNA transfers to the limelight. Genes Dev 17**:** 162-180.

Hou, Y. M., H. Gamper and W. Yang, 2015 Post-transcriptional modifications to tRNA--a response to the genetic code degeneracy. RNA 21**:** 642-644.

Huang, C. Y., M. A. Ayliffe and J. N. Timmis, 2003 Direct measurement of the transfer rate of chloroplast DNA into the nucleus. Nature 422**:** 72-76.

Huang, S., N. L. Taylor, J. Whelan and A. H. Millar, 2009 Refining the definition of plant mitochondrial presequences through analysis of sorting signals, N-terminal modifications, and cleavage motifs. Plant Physiol 150**:** 1272-1285.

Huang, S. Q., B. Sun, Z. P. Xiong, Y. Shu, H. H. Zhou *et al.*, 2018 The dysregulation of tRNAs and tRNA derivatives in cancer. J Exp Clin Cancer Res 37**:** 101.

Huber, S. M., A. Leonardi, P. C. Dedon and T. J. Begley, 2019 The Versatile Roles of the tRNA Epitranscriptome during Cellular Responses to Toxic Exposures and Environmental Stress. Toxics 7.

Hummel, G., J. Warren and L. Drouard, 2019 The multi-faceted regulation of nuclear tRNA gene transcription. IUBMB Life 71**:** 1099-1108.

Huynen, M. A., I. Duarte and R. Szklarczyk, 2013 Loss, replacement and gain of proteins at the origin of the mitochondria. Biochimica Et Biophysica Acta-Bioenergetics 1827**:** 224-231.

Iams, K. P., J. E. Heckman and J. H. Sinclair, 1985 Sequence of histidyl tRNA, present as a chloroplast insert in mtDNA ofZea mays. Plant Mol Biol 4**:** 225-232.

Ibba, M., and C. Francklyn, 2004 Turning tRNA upside down: When aminoacylation is not a prerequisite to protein synthesis. Proc Natl Acad Sci U S A 101**:** 7493-7494.

Ibba, M., and D. Söll, 2004 Aminoacyl-tRNAs: setting the limits of the genetic code. Genes Dev 18**:** 731-738.

Iida, K., H. Jin and J. K. Zhu, 2009 Bioinformatics analysis suggests base modifications of tRNAs and miRNAs in Arabidopsis thaliana. BMC Genomics 10**:** 155.

Jackman, J. E., and J. D. Alfonzo, 2013 Transfer RNA modifications: nature's combinatorial chemistry playground. Wiley Interdiscip Rev RNA 4**:** 35-48.

Jackman, S. D., L. Coombe, R. L. Warren, H. Kirk, E. Trinh *et al.*, 2019 Largest Complete Mitochondrial Genome of a Gymnosperm, Sitka Spruce (Picea sitchensis), Indicates Complex Physical Structure. bioRxiv**:** 601104.

Jacob, D., K. Thuring, A. Galliot, V. Marchand, A. Galvanin *et al.*, 2019 Absolute Quantification of Noncoding RNA by Microscale Thermophoresis. Angew Chem Int Ed Engl 58**:** 9565-9569.

Jafari, F., S. Zarre, A. Gholipour, F. Eggens, R. K. Rabeler *et al.*, 2020 A new taxonomic backbone for the infrageneric classification of the species-rich genus Silene (Caryophyllaceae). TAXON 69**:** 337-368.

Jain, M., H. E. Olsen, B. Paten and M. Akeson, 2016 The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. Genome Biol 17**:** 239.

Janouskovec, J., D. V. Tikhonenkov, F. Burki, A. T. Howe, F. L. Rohwer *et al.*, 2017 A New Lineage of Eukaryotes Illuminates Early Mitochondrial Genome Reduction. Curr Biol 27**:** 3717-3724 e3715.

Johnston, I. G., and B. P. Williams, 2016 Evolutionary Inference across Eukaryotes Identifies Specific Pressures Favoring Mitochondrial Gene Retention. Cell Syst 2**:** 101-111.

Jordon-Thaden, I. E., A. S. Chanderbali, M. A. Gitzendanner and D. E. Soltis, 2015 Modified CTAB and TRIzol protocols improve RNA extraction from chemically complex Embryophyta. Appl Plant Sci 3.

Joyce, P. B., and M. W. Gray, 1989 Chloroplast-like transfer RNA genes expressed in wheat mitochondria. Nucleic Acids Res 17**:** 5461-5476.

Juhling, F., M. Morl, R. K. Hartmann, M. Sprinzl, P. F. Stadler *et al.*, 2009 tRNAdb 2009: compilation of tRNA sequences and tRNA genes. Nucleic Acids Res 37**:** D159-162.

Juhling, T., E. Duchardt-Ferner, S. Bonin, J. Wohnert, J. Putz *et al.*, 2018 Small but large enough: structural properties of armless mitochondrial tRNAs from the nematode Romanomermis culicivorax. Nucleic Acids Res 46**:** 9170-9180.

Kamenski, P., O. Kolesnikova, V. Jubenot, N. Entelis, I. A. Krasheninnikov *et al.*, 2007 Evidence for an adaptation mechanism of mitochondrial translation via tRNA import from the cytosol. Mol Cell 26**:** 625-637.

Kamikawa, R., T. Shiratori, K. Ishida, H. Miyashita and A. J. Roger, 2016 Group II Intron-Mediated Trans-Splicing in the Gene-Rich Mitochondrial Genome of an Enigmatic Eukaryote, Diphylleia rotans. Genome Biol Evol 8**:** 458-466.

Kan, S. L., T. T. Shen, P. Gong, J. H. Ran and X. Q. Wang, 2020 The complete mitochondrial genome of Taxus cuspidata (Taxaceae): eight protein-coding genes have transferred to the nuclear genome. BMC Evol Biol 20**:** 10.

Kashdan, M. A., and B. S. Dudock, 1982 The gene for a spinach chloroplast isoleucine tRNA has a methionine anticodon. J Biol Chem 257**:** 11191-11194.

Katoh, K., J. Rozewicki and K. D. Yamada, 2017 MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Brief Bioinform.

Katoh, K., and D. M. Standley, 2013 MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Molecular Biology and Evolution 30**:** 772-780.

Kietrys, A. M., W. A. Velema and E. T. Kool, 2017 Fingerprints of Modified RNA Bases from Deep Sequencing Profiles. J Am Chem Soc 139**:** 17074-17081.

Kimura, S., P. C. Dedon and M. K. Waldor, 2019 Surveying the landscape of tRNA modifications by combining tRNA sequencing and RNA mass spectrometry. bioRxiv**:** 723049.

Kirchner, S., and Z. Ignatova, 2015 Emerging roles of tRNA in adaptive translation, signalling dynamics and disease. Nat Rev Genet 16**:** 98-112.

Kitazaki, K., T. Kubo, H. Kagami, T. Matsumoto, A. Fujita *et al.*, 2011 A horizontally transferred tRNA(Cys) gene in the sugar beet mitochondrial genome: evidence that the gene is present in diverse angiosperms and its transcript is aminoacylated. Plant J 68**:** 262-272.

Knie, N., M. Polsakiewicz and V. Knoop, 2015 Horizontal Gene Transfer of Chlamydial-Like tRNA Genes into Early Vascular Plant Mitochondria. Molecular Biology and Evolution 32**:** 629-634.

Knoop, V., 2004 The mitochondrial DNA of land plants: peculiarities in phylogenetic perspective. Curr Genet 46**:** 123-139.

Knoop, V., W. Schuster, B. Wissinger and A. Brennicke, 1991 Trans splicing integrates an exon of 22 nucleotides into the nad5 mRNA in higher plant mitochondria. EMBO J 10**:** 3483-3493.

Kohn, A. B., M. R. Citarella, K. M. Kocot, Y. V. Bobkova, K. M. Halanych *et al.*, 2012 Rapid evolution of the compact and unusual mitochondrial genome in the ctenophore, Pleurobrachia bachei. Mol Phylogenet Evol 63**:** 203-207.

Kondo, K., B. Makovec, R. H. Waterston and J. Hodgkin, 1990 Genetic and molecular analysis of eight tRNA(Trp) amber suppressors in Caenorhabditis elegans. J Mol Biol 215**:** 7-19.

Kowalak, J. A., S. C. Pomerantz, P. F. Crain and J. A. McCloskey, 1993 A novel method for the determination of post-transcriptional modification in RNA by mass spectrometry. Nucleic Acids Res 21**:** 4577-4585.

Krasovec, M., M. Chester, K. Ridout and D. A. Filatov, 2018 The Mutation Rate and the Age of the Sex Chromosomes in Silene latifolia. Curr Biol 28**:** 1832-1838 e1834.

Kubo, T., S. Nishizawa, A. Sugawara, N. Itchoda, A. Estiati *et al.*, 2000 The complete nucleotide sequence of the mitochondrial genome of sugar beet (Beta vulgaris L.) reveals a novel gene for tRNA(Cys)(GCA). Nucleic Acids Res 28**:** 2571-2576.

Kuchino, Y., N. Hanyu and S. Nishimura, 1987 Analysis of modified nucleosides and nucleotide sequence of tRNA. Methods Enzymol 155**:** 379-396.

Kühn, K., U. Richter, E. H. Meyer, E. Delannoy, A. F. de Longevialle *et al.*, 2009 Phage-type RNA polymerase RPOTmp performs gene-specific transcription in mitochondria of Arabidopsis thaliana. Plant Cell 21**:** 2762-2779.

Kumar, R., L. Marechal-Drouard, K. Akama and I. Small, 1996 Striking differences in mitochondrial tRNA import between different plant species. Mol Gen Genet 252**:** 404-411.

Kunze, M., and J. Berger, 2015 The similarity between N-terminal targeting signals for protein import into different organelles and its evolutionary relevance. Front Physiol 6**:** 259.

Lama, L., J. Cobo, D. Buenaventura and K. Ryan, 2019 Small RNA-seq: The RNA 5'-end adapter ligation problem and how to circumvent it. J Biol Methods 6.

Lang, B. F., M. W. Gray and G. Burger, 1999 Mitochondrial genome evolution and the origin of eukaryotes. Annu Rev Genet 33**:** 351-397.

Lavrov, D. V., and W. Pett, 2016 Animal Mitochondrial DNA as We Do Not Know It: mt-Genome Organization and Evolution in Nonbilaterian Lineages. Genome Biol Evol 8**:** 2896-2913.

Lechner, M., W. Rossmanith, R. K. Hartmann, C. Tholken, B. Gutmann *et al.*, 2015 Distribution of Ribonucleoprotein and Protein-Only RNase P in Eukarya. Mol Biol Evol 32**:** 3186-3193.

Lee, D. W., J. K. Kim, S. Lee, S. Choi, S. Kim *et al.*, 2008 Arabidopsis nuclear-encoded plastid transit peptides contain multiple sequence subgroups with distinctive chloroplast-targeting sequence motifs. Plant Cell 20**:** 1603-1622.

Leon, P., V. Walbot and P. Bedinger, 1989 Molecular analysis of the linear 2.3 kb plasmid of maize mitochondria: apparent capture of tRNA genes. Nucleic Acids Res 17**:** 4089-4099.

Liu, Y., and Y. Chen, 2020 Mitochondrial tRNA Mutations Associated With Essential Hypertension: From Molecular Genetics to Function. Front Cell Dev Biol 8**:** 634137.

Liu, Y., J. Y. Xue, B. Wang, L. Li and Y. L. Qiu, 2011 The mitochondrial genomes of the early land plants Treubia lacunosa and Anomodon rugelii: dynamic and conservative evolution. PLoS One 6**:** e25836.

Lorenz, C., C. E. Lunse and M. Morl, 2017 tRNA Modifications: Impact on Structure and Thermal Adaptation. Biomolecules 7.

Lorenz, R., S. H. Bernhart, C. Honer Zu Siederdissen, H. Tafer, C. Flamm *et al.*, 2011 ViennaRNA Package 2.0. Algorithms Mol Biol 6**:** 26.

Lyons, S. M., M. M. Fay and P. Ivanov, 2018 The role of RNA modifications in the regulation of tRNA cleavage. FEBS Lett 592**:** 2828-2844.

Ma, X., F. Si, X. Liu and W. Luan, 2020 PRMdb: a repository of predicted RNA modifications in plants. Plant and Cell Physiology In press.

Machnicka, M. A., A. Olchowik, H. Grosjean and J. M. Bujnicki, 2014 Distribution and frequencies of post-transcriptional modifications in tRNAs. RNA Biol 11**:** 1619-1629.

Mackenzie, S. A., 2005 Plant organellar protein targeting: a traffic plan still under construction. Trends Cell Biol 15**:** 548-554.

Magallon, S., S. Gomez-Acevedo, L. L. Sanchez-Reyes and T. Hernandez-Hernandez, 2015 A metacalibrated time-tree documents the early rise of flowering plant phylogenetic diversity. New Phytol 207**:** 437-453.

Marchfelder, A., W. Schuster and A. Brennicke, 1990 In vitro processing of mitochondrial and plastid derived tRNA precursors in a plant mitochondrial extract. Nucleic Acids Res 18**:** 1401-1406.

Marechal-Drouard, L., P. Guillemaut, A. Cosset, M. Arbogast, F. Weber *et al.*, 1990 Transfer RNAs of potato (Solanum tuberosum) mitochondria have different genetic origins. Nucleic Acids Res 18**:** 3689-3696.

Marechal-Drouard, L., D. Ramamonjisoa, A. Cosset, J. H. Weil and A. Dietrich, 1993 Editing corrects mispairing in the acceptor stem of bean and potato mitochondrial phenylalanine transfer RNAs. Nucleic Acids Res 21**:** 4909-4914.

Maréchal-Drouard, L., J. Weil and A. Dietrich, 1993 Transfer RNAs and transfer RNA genes in plants. Annual review of plant biology 44**:** 13-32.

Martin, M., 2011 Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet. journal 17**:** 10-12.

Martin, R. P., J. M. Schneller, A. J. Stahl and G. Dirheimer, 1979 Import of nuclear deoxyribonucleic acid coded lysine-accepting transfer ribonucleic acid (anticodon C-U-U) into yeast mitochondria. Biochemistry 18**:** 4600-4605.

Meiklejohn, C. D., M. A. Holmbeck, M. A. Siddiq, D. N. Abt, D. M. Rand *et al.*, 2013 An Incompatibility between a Mitochondrial tRNA and Its Nuclear-Encoded tRNA Synthetase Compromises Development and Fitness in Drosophila. Plos Genetics 9.

Mercer, T. R., S. Neph, M. E. Dinger, J. Crawford, M. A. Smith *et al.*, 2011 The human mitochondrial transcriptome. Cell 146**:** 645-658.

Michaud, M., V. Cognat, A. M. Duchêne and L. Marechal-Drouard, 2011a A global picture of tRNA genes in plant genomes. Plant Journal 66**:** 80-93.

Michaud, M., V. Cognat, A. M. Duchêne and L. Marechal-Drouard, 2011b A global picture of tRNA genes in plant genomes. Plant J 66**:** 80-93.

Mireau, H., D. Lancelin and I. D. Small, 1996 The same Arabidopsis gene encodes both cytosolic and mitochondrial alanyl-tRNA synthetases. Plant Cell 8**:** 1027-1039.

Mishina, Y., and C. He, 2006 Oxidative dealkylation DNA repair mediated by the mononuclear non-heme iron AlkB proteins. J Inorg Biochem 100**:** 670-678.

Miyata, S., M. Nakazono and A. Hirai, 1998 Transcription of plastid-derived tRNA genes in rice mitochondria. Curr Genet 34**:** 216-220.

Motorin, Y., and M. Helm, 2010 tRNA stabilization by modified nucleotides. Biochemistry 49**:** 4934-4944.

Motorin, Y., and M. Helm, 2019 Methods for RNA Modification Mapping Using Deep Sequencing: Established and New Emerging Technologies. Genes 10.

Motorin, Y., S. Muller, I. Behm-Ansmant and C. Branlant, 2007 Identification of modified residues in RNAs by reverse transcription-based methods. Methods Enzymol 425**:** 21-53.

Mower, J. P., P. Touzet, J. S. Gummow, L. F. Delph and J. D. Palmer, 2007 Extensive variation in synonymous substitution rates in mitochondrial genes of seed plants. BMC Evol Biol 7**:** 135.

Mubumbila, M., E. J. Crouse and J. H. Weil, 1984 Transfer RNAs and tRNA genes of Vicia faba chloroplasts. Curr Genet 8**:** 379-385.

Murcha, M. W., B. Kmiec, S. Kubiszewski-Jakubiak, P. F. Teixeira, E. Glaser *et al.*, 2014 Protein import into plant mitochondria: signals, machinery, processing, and regulation. J Exp Bot 65**:** 6301-6335.

Murcha, M. W., S. Kubiszewski-Jakubiak, P. F. Teixeira, I. L. Gugel, B. Kmiec *et al.*, 2016 Plant-Specific Preprotein and Amino Acid Transporter Proteins Are Required for tRNA Import into Mitochondria. Plant Physiol 172**:** 2471-2490.

Nakazono, M., S. Nishiwaki, N. Tsutsumi and A. Hirai, 1996 A chloroplast-derived sequence is utilized as a source of promoter sequences for the gene for subunit 9 of NADH dehydrogenase (nad9) in rice mitochondria. Mol Gen Genet 252**:** 371-378.

Nameki, N., K. Tamura, H. Asahara and T. Hasegawa, 1997 Recognition of tRNA(Gly) by three widely diverged glycyl-tRNA synthetases. J Mol Biol 268**:** 640-647.

Notsu, Y., S. Masood, T. Nishikawa, N. Kubo, G. Akiduki *et al.*, 2002 The complete sequence of the rice (Oryza sativa L.) mitochondrial genome: frequent DNA sequence acquisition and loss during the evolution of flowering plants. Mol Genet Genomics 268**:** 434-445.

Noutahi, E., V. Calderon, M. Blanchette, N. El-Mabrouk and B. F. Lang, 2019 Rapid Genetic Code Evolution in Green Algal Mitochondrial Genomes. Mol Biol Evol 36**:** 766-783.

Novoa, E. M., M. Pavon-Eternod, T. Pan and L. Ribas de Pouplana, 2012 A role for tRNA modifications in genome structure and codon usage. Cell 149**:** 202-213.

Nurk, S., A. Bankevich, D. Antipov, A. A. Gurevich, A. Korobeynikov *et al.*, 2013 Assembling single-cell genomes and mini-metagenomes from chimeric MDA products. J Comput Biol 20**:** 714-737.

O'Donoghue, P., and Z. Luthey-Schulten, 2003 On the evolution of structure in aminoacyl-tRNA synthetases. Microbiol Mol Biol Rev 67**:** 550-573.

Onanuga, K., T. J. Begley, A. A. Chen and S. V. Ranganathan, 2017 tRNA Modification Detection Using Graphene Nanopores: A Simulation Study. Biomolecules 7.

One Thousand Plant Transcriptomes, I., 2019 One thousand plant transcriptomes and the phylogenomics of green plants. Nature 574**:** 679-685.

Ott, M., A. Amunts and A. Brown, 2016 Organization and Regulation of Mitochondrial Protein Synthesis. Annu Rev Biochem 85**:** 77-101.

Pan, T., 2018 Modifications and functional genomics of human transfer RNA. Cell Res 28**:** 395-404.

Pang, Y. L., R. Abo, S. S. Levine and P. C. Dedon, 2014a Diverse cell stresses induce unique patterns of tRNA up- and down-regulation: tRNA-seq for quantifying changes in tRNA copy number. Nucleic Acids Res 42**:** e170.

Pang, Y. L., K. Poruri and S. A. Martinis, 2014b tRNA synthetase: tRNA aminoacylation and beyond. Wiley Interdiscip Rev RNA 5**:** 461-480.

Paris, Z., and J. D. Alfonzo, 2018 How the intracellular partitioning of tRNA and tRNA modification enzymes affects mitochondrial function. IUBMB Life 70**:** 1207-1213.

Park, E. J., and T. H. Kim, 2018 Fine-Tuning of Gene Expression by tRNA-Derived Fragments during Abiotic Stress Signal Transduction. Int J Mol Sci 19.

Park, S., F. Grewe, A. Zhu, T. A. Ruhlman, J. Sabir *et al.*, 2015 Dynamic evolution of Geranium mitochondrial genomes through multiple horizontal and intracellular gene transfers. New Phytol 208**:** 570-583.

Pawar, K., M. Shigematsu, P. Loher, S. Honda, I. Rigoutsos *et al.*, 2019 Exploration of CCA-added RNAs revealed the expression of mitochondrial non-coding RNAs regulated by CCA-adding enzyme. RNA Biol 16**:** 1817-1825.

Pearce, S., C. L. Nezich and A. Spinazzola, 2013 Mitochondrial diseases: translation matters. Mol Cell Neurosci 55**:** 1-12.

Peeters, N., and I. Small, 2001 Dual targeting to mitochondria and chloroplasts. Biochim Biophys Acta 1541**:** 54-63.

Pereira, M., S. Francisco, A. S. Varanda, M. Santos, M. A. S. Santos *et al.*, 2018 Impact of tRNA Modifications and tRNA-Modifying Enzymes on Proteostasis and Human Disease. Int J Mol Sci 19.

Petersen, G., A. Cuenca, I. M. Moller and O. Seberg, 2015 Massive gene loss in mistletoe (Viscum, Viscaceae) mitochondria. Sci Rep 5**:** 17588.

Peterson, K. J., J. B. Lyons, K. S. Nowak, C. M. Takacs, M. J. Wargo *et al.*, 2004 Estimating metazoan divergence times with a molecular clock. Proc Natl Acad Sci U S A 101**:** 6536-6541.

Petrov, A. S., B. Gulen, A. M. Norris, N. A. Kovacs, C. R. Bernier *et al.*, 2015 History of the ribosome and the origin of translation. Proc Natl Acad Sci U S A 112**:** 15396-15401.

Pett, W., and D. V. Lavrov, 2015 Cytonuclear Interactions in the Evolution of Animal Mitochondrial tRNA Metabolism. Genome Biol Evol 7**:** 2089-2101.

Phizicky, E. M., and J. D. Alfonzo, 2010 Do all modifications benefit all tRNAs? FEBS Lett 584**:** 265-271.

Phizicky, E. M., and A. K. Hopper, 2010 tRNA biology charges to the front. Genes Dev 24**:** 1832-1860.

Phizicky, E. M., and A. K. Hopper, 2015 tRNA processing, modification, and subcellular dynamics: past, present, and future. RNA 21**:** 483-485.

Pons, J., P. Bover, L. Bidegaray-Batista and M. A. Arnedo, 2019 Arm-less mitochondrial tRNAs conserved for over 30 millions of years in spiders. BMC Genomics 20**:** 665.

Poodari, V. C., 2019 Direct RNA Sequencing of E. coli initiator tRNA Using the MinION Sequencing Platform, pp. UC Santa Cruz.

Potapov, V., X. Fu, N. Dai, I. R. Correa, Jr., N. A. Tanner *et al.*, 2018 Base modifications affecting RNA polymerase and reverse transcriptase fidelity. Nucleic Acids Res 46**:** 5753-5763.

Pujol, C., M. Bailly, D. Kern, L. Marechal-Drouard, H. Becker *et al.*, 2008 Dual-targeted tRNA-dependent amidotransferase ensures both mitochondrial and chloroplastic Gln-tRNAGln synthesis in plants. Proc Natl Acad Sci U S A 105**:** 6481-6485.

Pujol, C., L. Marechal-Drouard and A. M. Duchêne, 2007 How can organellar protein N-terminal sequences be dual targeting signals? In silico analysis and mutagenesis approach. J Mol Biol 369**:** 356-367.

Qiu, Y. L., and J. D. Palmer, 2004 Many independent origins of trans splicing of a plant mitochondrial group II intron. J Mol Evol 59**:** 80-89.

Quinlan, A. R., and I. M. Hall, 2010 BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26**:** 841-842.

Raina, M., and M. Ibba, 2014 tRNAs as regulators of biological processes. Front Genet 5**:** 171.

Reinbothe, S., C. Rossig, J. Gray, S. Rustgi, D. von Wettstein *et al.*, 2021 tRNA-Dependent Import of a Transit Sequence-Less Aminoacyl-tRNA Synthetase (LeuRS2) into the Mitochondria of Arabidopsis. Int J Mol Sci 22.

Rice, D. W., A. J. Alverson, A. O. Richardson, G. J. Young, M. V. Sanchez-Puerta *et al.*, 2013 Horizontal transfer of entire genomes via mitochondrial fusion in the angiosperm Amborella. Science 342**:** 1468-1473.

Rice, P., I. Longden and A. Bleasby, 2000 EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet 16**:** 276-277.

Richardson, A. O., D. W. Rice, G. J. Young, A. J. Alverson and J. D. Palmer, 2013 The "fossilized" mitochondrial genome of Liriodendron tulipifera: ancestral gene content and order, ancestral editing sites, and extraordinarily low mutation rate. BMC Biol 11**:** 29.

Rinehart, J., B. Krett, M. A. Rubio, J. D. Alfonzo and D. Söll, 2005 Saccharomyces cerevisiae imports the cytosolic pathway for Gln-tRNA synthesis into the mitochondrion. Genes Dev 19**:** 583-592.

Robinson, M. D., D. J. McCarthy and G. K. Smyth, 2010 edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26**:** 139-140.

Rodnina, M. V., K. B. Gromadski, U. Kothe and H. J. Wieden, 2005 Recognition and selection of tRNA in translation. FEBS Lett 579**:** 938-942.

Roger, A. J., S. A. Munoz-Gomez and R. Kamikawa, 2017 The Origin and Diversification of Mitochondria. Curr Biol 27**:** R1177-R1192.

Rogers, H. H., and S. Griffiths-Jones, 2014 tRNA anticodon shifts in eukaryotic genomes. RNA 20**:** 269-281.

Rojas-Benitez, D., P. C. Thiaville, V. de Crecy-Lagard and A. Glavic, 2015 The Levels of a Universally Conserved tRNA Modification Regulate Cell Growth. J Biol Chem 290**:** 18699-18707.

Ross, R., X. Cao, N. Yu and P. A. Limbach, 2016 Sequence mapping of transfer RNA chemical modifications by liquid chromatography tandem mass spectrometry. Methods 107**:** 73-78.

Rossmanith, W., 2012 Of P and Z: mitochondrial tRNA processing enzymes. Biochim Biophys Acta 1819**:** 1017-1026.

Rubio Gomez, M. A., and M. Ibba, 2020 Aminoacyl-tRNA synthetases. RNA 26**:** 910-936.

Rubio, M. A., J. J. Rinehart, B. Krett, S. Duvezin-Caubet, A. S. Reichert *et al.*, 2008 Mammalian mitochondria have the innate ability to import tRNAs by a mechanism distinct from protein import. Proc Natl Acad Sci U S A 105**:** 9186-9191.

Salazar, J. C., I. Ahel, O. Orellana, D. Tumbula-Hansen, R. Krieger *et al.*, 2003 Coevolution of an aminoacyl-tRNA synthetase with its tRNA substrates. Proc Natl Acad Sci U S A 100**:** 13863-13868.

Salinas, T., F. Duby, V. Larosa, N. Coosemans, N. Bonnefoy *et al.*, 2012 Co-evolution of mitochondrial tRNA import and codon usage determines translational efficiency in the green alga Chlamydomonas. PLoS Genet 8**:** e1002946.

Salinas, T., A. M. Duchêne and L. Marechal-Drouard, 2008 Recent advances in tRNA mitochondrial import. Trends Biochem Sci 33**:** 320-329.

Salinas, T., C. Schaeffer, L. Marechal-Drouard and A. M. Duchêne, 2005 Sequence dependence of tRNA(Gly) import into tobacco mitochondria. Biochimie 87**:** 863-872.

Salinas-Giege, T., R. Giege and P. Giege, 2015 tRNA Biology in Mitochondria. International Journal of Molecular Sciences 16**:** 4518-4559.

Salinas-Giegé, T., R. Giegé and P. Giegé, 2015 tRNA biology in mitochondria. Int J Mol Sci 16**:** 4518-4559.

Sampson, J. R., and O. C. Uhlenbeck, 1988 Biochemical and physical characterization of an unmodified yeast phenylalanine transfer RNA transcribed in vitro. Proc Natl Acad Sci U S A 85**:** 1033-1037.

Sanchez-Puerta, M. V., A. Edera, C. L. Gandini, A. V. Williams, K. A. Howell *et al.*, 2019 Genome-scale transfer of mitochondrial DNA from legume hosts to the holoparasite Lophophytum mirabile (Balanophoraceae). Mol Phylogenet Evol 132**:** 243-250.

Schmidt, O., N. Pfanner and C. Meisinger, 2010 Mitochondrial protein import: from proteomics to functional mechanisms. Nat Rev Mol Cell Biol 11**:** 655-667.

Schneider, A., 2001 Unique aspects of mitochondrial biogenesis in trypanosomatids. International journal for parasitology 31**:** 1403-1415.

Schneider, A., 2011 Mitochondrial tRNA Import and Its Consequences for Mitochondrial Translation. Annual Review of Biochemistry, Vol 80 80**:** 1033-1053.

Schwartz, S., and Y. Motorin, 2017 Next-generation sequencing technologies for detection of modified nucleotides in RNAs. RNA Biol 14**:** 1124-1137.

Selden, R. F., A. Steinmetz, L. McIntosh, L. Bogorad, G. Burkard *et al.*, 1983 Transfer RNA genes ofZea mays chloroplast DNA. Plant Mol Biol 2**:** 141-153.

Sequeira-Mendes, J., I. Araguez, R. Peiro, R. Mendez-Giraldez, X. Zhang *et al.*, 2014 The Functional Topography of the Arabidopsis Genome Is Organized in a Reduced Number of Linear Motifs of Chromatin States. Plant Cell 26**:** 2351-2366.

Shigematsu, M., S. Honda, P. Loher, A. G. Telonis, I. Rigoutsos *et al.*, 2017 YAMAT-seq: an efficient method for high-throughput sequencing of mature transfer RNAs. Nucleic Acids Research 45.

Simpson, A. M., Y. Suyama, H. Dewes, D. A. Campbell and L. Simpson, 1989 Kinetoplastid Mitochondria Contain Functional Transfer-Rnas Which Are Encoded in Nuclear-DNA and Also Contain Small Minicircle and Maxicircle Transcripts of Unknown Function. Nucleic Acids Research 17**:** 5427-5445.

Sinn, B. T., and C. F. Barrett, 2019 Ancient mitochondrial gene transfer between fungi and the orchids. Mol Biol Evol.

Sissler, M., J. Putz, F. Fasiolo and C. Florentz, 2005 Mitochondrial Aminoacyl-tRNA Synthetases, pp. in *Madame Curie Bioscience Database [Internet]*. Landes Bioscience, Austin (TX).

Skippington, E., T. J. Barkman, D. W. Rice and J. D. Palmer, 2015 Miniaturized mitogenome of the parasitic plant Viscum scurruloideum is extremely divergent and dynamic and has lost all nad genes. Proc Natl Acad Sci U S A 112**:** E3515-3524.

Sloan, D. B., A. J. Alverson, J. P. Chuckalovcak, M. Wu, D. E. McCauley *et al.*, 2012a Rapid Evolution of Enormous, Multichromosomal Genomes in Flowering Plant Mitochondria with Exceptionally High Mutation Rates. Plos Biology 10.

Sloan, D. B., A. J. Alverson, H. Storchova, J. D. Palmer and D. R. Taylor, 2010a Extensive loss of translational genes in the structurally dynamic mitochondrial genome of the angiosperm Silene latifolia. Bmc Evolutionary Biology 10.

Sloan, D. B., A. J. Alverson, H. Storchova, J. D. Palmer and D. R. Taylor, 2010b Extensive loss of translational genes in the structurally dynamic mitochondrial genome of the angiosperm Silene latifolia. BMC Evol Biol 10**:** 274.

Sloan, D. B., A. J. Alverson, M. Wu, J. D. Palmer and D. R. Taylor, 2012b Recent acceleration of plastid sequence and structural evolution coincides with extreme mitochondrial divergence in the angiosperm genus Silene. Genome Biol Evol 4**:** 294-306.

Sloan, D. B., K. Muller, D. E. McCauley, D. R. Taylor and H. Storchova, 2012c Intraspecific variation in mitochondrial genome sequence, structure, and gene content in Silene vulgaris, an angiosperm with pervasive cytoplasmic male sterility. New Phytol 196**:** 1228-1239.

Sloan, D. B., B. Oxelman, A. Rautenberg and D. R. Taylor, 2009 Phylogenetic analysis of mitochondrial substitution rate variation in the angiosperm tribe Sileneae. BMC Evol Biol 9**:** 260.

Sloan, D. B., D. A. Triant, N. J. Forrester, L. M. Bergner, M. Wu *et al.*, 2014a A recurring syndrome of accelerated plastid genome evolution in the angiosperm tribe Sileneae (Caryophyllaceae). Mol Phylogenet Evol 72**:** 82-89.

Sloan, D. B., D. A. Triant, M. Wu and D. R. Taylor, 2014b Cytonuclear interactions and relaxed selection accelerate sequence evolution in organelle ribosomes. Mol Biol Evol 31**:** 673-682.

Sloan, D. B., J. M. Warren, A. M. Williams, Z. Wu, S. E. Abdel-Ghany *et al.*, 2018 Cytonuclear integration and co-evolution. Nat Rev Genet 19**:** 635-648.

Sloan, D. B., and Z. Wu, 2014 History of plastid DNA insertions reveals weak deletion and at mutation biases in angiosperm mitochondrial genomes. Genome Biol Evol 6**:** 3210-3221.

Small, I., K. Akashi, A. Chapron, A. Dietrich, A. M. Duchêne *et al.*, 1999 The strange evolutionary history of plant mitochondrial tRNAs and their aminoacyl-tRNA synthetases. Journal of Heredity 90**:** 333-337.

Small, I., N. Peeters, F. Legeai and C. Lurin, 2004 Predotar: A tool for rapidly screening proteomes for N-terminal targeting sequences. Proteomics 4**:** 1581-1590.

Smith, A. M., R. Abu-Shumays, M. Akeson and D. L. Bernick, 2015 Capture, Unfolding, and Detection of Individual tRNA Molecules Using a Nanopore Device. Front Bioeng Biotechnol 3**:** 91.

Soprano, A. S., J. H. C. Smetana and C. E. Benedetti, 2018 Regulation of tRNA biogenesis in plants and its link to plant growth and response to pathogens. Biochim Biophys Acta Gene Regul Mech 1861**:** 344-353.

Sperschneider, J., A. M. Catanzariti, K. DeBoer, B. Petre, D. M. Gardiner *et al.*, 2017 LOCALIZER: subcellular localization prediction of both plant and effector proteins in the plant cell. Sci Rep 7**:** 44598.

Stamatakis, A., 2014 RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30**:** 1312-1313.

Stupar, R. M., J. W. Lilly, C. D. Town, Z. Cheng, S. Kaul *et al.*, 2001 Complex mtDNA constitutes an approximate 620-kb insertion on Arabidopsis thaliana chromosome 2: implication of potential sequencing errors caused by large-unit repeats. Proc Natl Acad Sci U S A 98**:** 5099-5103.

Su, D., C. T. Chan, C. Gu, K. S. Lim, Y. H. Chionh *et al.*, 2014 Quantitative analysis of ribonucleoside modifications in tRNA by HPLC-coupled mass spectrometry. Nat Protoc 9**:** 828-841.

Suzuki, T., and K. Miyauchi, 2010 Discovery and characterization of tRNAIle lysidine synthetase (TilS). FEBS Lett 584**:** 272-277.

Takeuchi, N., L. Vial, M. Panvert, E. Schmitt, K. Watanabe *et al.*, 2001 Recognition of tRNAs by Methionyl-tRNA transformylase from mammalian mitochondria. J Biol Chem 276**:** 20064-20068.

Tan, T. H., R. Pach, A. Crausaz, A. Ivens and A. Schneider, 2002 tRNAs in Trypanosoma brucei: genomic organization, expression, and mitochondrial import. Mol Cell Biol 22**:** 3707-3717.

Taylor, Z. N., D. W. Rice and J. D. Palmer, 2015 The Complete Moss Mitochondrial Genome in the Angiosperm Amborella Is a Chimera Derived from Two Moss Whole-Genome Transfers. PLoS One 10**:** e0137532.

Telonis, A. G., Y. Kirino and I. Rigoutsos, 2015 Mitochondrial tRNA-lookalikes in nuclear chromosomes: could they be functional? RNA Biol 12**:** 375-380.

Telonis, A. G., P. Loher, Y. Kirino and I. Rigoutsos, 2014 Nuclear and mitochondrial tRNA-lookalikes in the human genome. Front Genet 5**:** 344.

Telonis, A. G., P. Loher, R. Magee, V. Pliatsika, E. Londin *et al.*, 2019 tRNA Fragments Show Intertwining with mRNAs of Specific Repeat Content and Have Links to Disparities. Cancer Res 79**:** 3034-3049.

Theologis, A., J. R. Ecker, C. J. Palm, N. A. Federspiel, S. Kaul *et al.*, 2000 Sequence and analysis of chromosome 1 of the plant Arabidopsis thaliana. Nature 408**:** 816-820.

Timmis, J. N., M. A. Ayliffe, C. Y. Huang and W. Martin, 2004 Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. Nat Rev Genet 5**:** 123-135.

Tolkunova, E., H. Park, J. Xia, M. P. King and E. Davidson, 2000 The human lysyl-tRNA synthetase gene encodes both the cytoplasmic and mitochondrial enzymes by means of an unusual alternative splicing of the primary transcript. J Biol Chem 275**:** 35063-35069.

Torres, A. G., D. Pineyro, M. Rodriguez-Escriba, N. Camacho, O. Reina *et al.*, 2015 Inosine modifications in human tRNAs are incorporated at the precursor tRNA level. Nucleic Acids Res 43**:** 5145-5157.

Torres, A. G., O. Reina, C. Stephan-Otto Attolini and L. Ribas de Pouplana, 2019 Differential expression of human tRNA genes drives the abundance of tRNA-derived fragments. Proc Natl Acad Sci U S A 116**:** 8451-8456.

Trewick, S. C., T. F. Henshaw, R. P. Hausinger, T. Lindahl and B. Sedgwick, 2002 Oxidative demethylation by Escherichia coli AlkB directly reverts DNA base damage. Nature 419**:** 174-178.

Tserovski, L., V. Marchand, R. Hauenschild, F. Blanloeil-Oillo, M. Helm *et al.*, 2016 High-throughput sequencing for 1-methyladenosine (m(1)A) mapping in RNA. Methods 107**:** 110-121.

Turmel, M., C. Otis and C. Lemieux, 2013 Tracing the evolution of streptophyte algae and their mitochondrial genome. Genome Biol Evol 5**:** 1817-1835.

Urbonavicius, J., Q. Qian, J. M. Durand, T. G. Hagervall and G. R. Bjork, 2001 Improvement of reading frame maintenance is a common function for several tRNA modifications. EMBO J 20**:** 4863-4873.

Valach, M., G. Burger, M. W. Gray and B. F. Lang, 2014 Widespread occurrence of organelle genome-encoded 5S rRNAs including permuted molecules. Nucleic Acids Res 42**:** 13764-13777.

Vandivier, L. E., Z. D. Anderson and B. D. Gregory, 2019 HAMR: High-Throughput Annotation of Modified Ribonucleotides. Methods Mol Biol 1870**:** 51-67.

Varre, J. S., N. D'Agostino, P. Touzet, S. Gallina, R. Tamburino *et al.*, 2019 Complete Sequence, Multichromosomal Architecture and Transcriptome Analysis of the Solanum tuberosum Mitochondrial Genome. Int J Mol Sci 20.

Vermeulen, A., S. A. McCallum and A. Pardi, 2005 Comparison of the global structure and dynamics of native and unmodified tRNAval. Biochemistry 44**:** 6024-6033.

Vinogradova, E., T. Salinas, V. Cognat, C. Remacle and L. Marechal-Drouard, 2009 Steady-state levels of imported tRNAs in Chlamydomonas mitochondria are correlated with both cytosolic and mitochondrial codon usages. Nucleic Acids Res 37**:** 1521-1528.

von Braun, S. S., A. Sabetti, P. J. Hanic-Joyce, J. Gu, E. Schleiff *et al.*, 2007 Dual targeting of the tRNA nucleotidyltransferase in plants: not just the signal. J Exp Bot 58**:** 4083-4093.

Wang, D., M. Rousseau-Gueutin and J. N. Timmis, 2012 Plastid sequences contribute to some plant mitochondrial genes. Mol Biol Evol 29**:** 1707-1711.

Wang, D., Y. W. Wu, A. C. Shih, C. S. Wu, Y. N. Wang *et al.*, 2007 Transfer of chloroplast genomic DNA to mitochondrial genome occurred at least 300 MYA. Mol Biol Evol 24**:** 2040-2048.

Wang, X. C., H. Chen, D. Yang and C. Liu, 2018 Diversity of mitochondrial plastid DNAs (MTPTs) in seed plants. Mitochondrial DNA A DNA Mapp Seq Anal 29**:** 635-642.

Wang, Z., M. Gerstein and M. Snyder, 2009 RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 10**:** 57-63.

Warren, J. M., T. Salinas-Giege, G. Hummel, N. L. Coots, J. M. Svendsen *et al.*, 2020 Combining tRNA sequencing methods to characterize plant tRNA expression and post-transcriptional modification. RNA Biol**:** 1-15.

Warren, J. M., T. Salinas-Giegé, G. Hummel, N. L. Coots, J. M. Svendsen *et al.*, 2021 Combining tRNA sequencing methods to characterize plant tRNA expression and post-transcriptional modification. RNA Biology 18**:** 64-78.

Warren, J. M., T. Salinas-Giegé, G. Hummel, N. L. Coots, J. M. Svendsen *et al.*, 2019 Combining tRNA sequencing methods to characterize plant tRNA expression and post-transcriptional modification. bioRxiv**:** 790451.

Warren, J. M., and D. B. Sloan, 2020a Hopeful monsters: Unintended sequencing of famously malformed mite mitochondrial tRNAs reveals widespread expression and processing of sense-antisense pairs. bioRxiv**:** 2020.2009.2008.286963.

Warren, J. M., and D. B. Sloan, 2020b Interchangeable parts: The evolutionarily dynamic tRNA population in plant mitochondria. Mitochondrion 52**:** 144-156.

Watanabe, Y., T. Suematsu and T. Ohtsuki, 2014 Losing the stem-loop structure from metazoan mitochondrial tRNAs and co-evolution of interacting factors. Front Genet 5**:** 109.

Weber, F., A. Dietrich, J. H. Weil and L. Maréchal-Drouard, 1990 A potato mitochondrial isoleucine tRNA is coded for by a mitochondrial gene possessing a methionine anticodon. Nucleic Acids Res 18**:** 5027-5030.

Werner, S., L. Schmidt, V. Marchand, T. Kemmer, C. Falschlunger *et al.*, 2020 Machine learning of reverse transcription signatures of variegated polymerases allows mapping and discrimination of methylated purines in limited transcriptomes. Nucleic Acids Res 48**:** 3734-3746.

Williams, A. M., M. W. Itgen, A. K. Broz, O. G. Carter and D. B. Sloan, 2020 Long-read transcriptome and other genomic resources for the angiosperm Silene noctiflora. bioRxiv**:** 2020.2008.2009.243378.

Williams, E. J., C. Pal and L. D. Hurst, 2000a The molecular evolution of signal peptides. Gene 253**:** 313-322.

Williams, M. A., Y. Johzuka and R. M. Mulligan, 2000b Addition of non-genomically encoded nucleotides to the 3'-terminus of maize mitochondrial mRNAs: truncated rps12 mRNAs frequently terminate with CCA. Nucleic Acids Res 28**:** 4444-4451.

Wilusz, J. E., 2015a Controlling translation via modulation of tRNA levels. Wiley Interdiscip Rev RNA 6**:** 453-470.

Wilusz, J. E., 2015b Removing roadblocks to deep sequencing of modified RNAs. Nat Methods 12**:** 821-822.

Wissinger, B., W. Schuster and A. Brennicke, 1991 Trans splicing in Oenothera mitochondria: nad1 mRNAs are edited in exon and trans-splicing group II intron sequences. Cell 65**:** 473-482.

Wu, Z., J. M. Cuthbert, D. R. Taylor and D. B. Sloan, 2015 The massive mitochondrial genome of the angiosperm Silene noctiflora is evolving by gain or loss of entire chromosomes. Proc Natl Acad Sci U S A 112**:** 10185-10191.

Xing, F., S. L. Hiley, T. R. Hughes and E. M. Phizicky, 2004 The specificities of four yeast dihydrouridine synthases for cytoplasmic tRNAs. J Biol Chem 279**:** 17850-17860.

Xu, L., C. Carrie, S. R. Law, M. W. Murcha and J. Whelan, 2013 Acquisition, conservation, and loss of dual-targeted proteins in land plants. Plant Physiol 161**:** 644-662.

Xue, J. Y., Y. Liu, L. Li, B. Wang and Y. L. Qiu, 2010 The complete mitochondrial genome sequence of the hornwort Phaeoceros laevis: retention of many ancient pseudogenes and conservative evolution of mitochondrial genomes in hornworts. Curr Genet 56**:** 53-61.

Yarham, J. W., J. L. Elson, E. L. Blakely, R. McFarland and R. W. Taylor, 2010 Mitochondrial tRNA mutations and disease. Wiley Interdiscip Rev RNA 1**:** 304-324.

Yogev, O., and O. Pines, 2011 Dual targeting of mitochondrial proteins: mechanism, regulation and function. Biochim Biophys Acta 1808**:** 1012-1020.

Yona, A. H., Z. Bloom-Ackermann, I. Frumkin, V. Hanson-Smith, Y. Charpak-Amikam *et al.*, 2013 tRNA genes rapidly change in evolution to meet novel translational demands. Elife 2**:** e01339.

Youvan, D. C., and J. E. Hearst, 1979 Reverse transcriptase pauses at N2-methylguanine during in vitro transcription of Escherichia coli 16S ribosomal RNA. Proc Natl Acad Sci U S A 76**:** 3751-3754.

Yu, C. H., J. Y. Liao, H. Zhou and L. H. Qu, 2008 The rat mitochondrial Ori L encodes a novel small RNA resembling an ancestral tRNA. Biochem Biophys Res Commun 372**:** 634-638.

Zhang, J., and A. R. Ferre-D'Amare, 2016 The tRNA Elbow in Structure, Recognition and Evolution. Life (Basel) 6.

Zhao, L., H. Zhang, M. V. Kohnen, K. Prasad, L. Gu *et al.*, 2019 Analysis of Transcriptome and Epitranscriptome in Plants Using PacBio Iso-Seq and Nanopore-Based Direct RNA Sequencing. Front Genet 10**:** 253.

Zheng, G., Y. Qin, W. C. Clark, Q. Dai, C. Yi *et al.*, 2015 Efficient and quantitative high-throughput tRNA sequencing. Nat Methods 12**:** 835-837.

Zhong, J., C. Xiao, W. Gu, G. Du, X. Sun *et al.*, 2015 Transfer RNAs Mediate the Rapid Adaptation of Escherichia coli to Oxidative Stress. PLoS Genet 11**:** e1005302.

CHAPTER 3: RAPID SHIFTS IN MITOCHONDRIAL TRNA IMPORT IN A PLANT LINEAGE WITH EXTENSIVE MITOCHONDRIAL TRNA GENE LOSS

Summary

In most eukaryotes, transfer RNAs (tRNAs) are one of the very few classes of genes remaining in the mitochondrial genome, but some mitochondria have lost these vestiges of their prokaryotic ancestry. Sequencing of mitogenomes from the flowering plant genus *Silene* previously revealed a large range in tRNA gene content, suggesting rapid and ongoing gene loss/replacement. Here, we use this system to test longstanding hypotheses about how mitochondrial tRNA genes are replaced by importing nuclear-encoded tRNAs. We traced the evolutionary history of these gene loss events by sequencing mitochondrial genomes from key outgroups (*Agrostemma githago* and *Silene* [=*Lychnis*] *chalcedonica*). We then performed the first global sequencing of purified plant mitochondrial tRNA populations to characterize the expression of mitochondrial-encoded tRNAs and the identity of imported nuclear-encoded tRNAs. We also confirmed the utility of high-throughput sequencing methods for the detection of tRNA import by sequencing mitochondrial tRNA populations in a species (*Solanum tuberosum*) with known tRNA trafficking patterns. Mitochondrial tRNA sequencing in *Silene* revealed substantial shifts in the abundance of some nuclear-encoded tRNAs in conjunction with their recent history of mt-tRNA gene loss and surprising cases where tRNAs with anticodons still encoded in the mitochondrial genome also appeared to be imported. These data suggest that nuclear-encoded counterparts are likely replacing mitochondrial tRNAs even in systems with recent mitochondrial tRNA gene loss, and the redundant import of a nuclear-encoded tRNA may provide a mechanism for functional replacement between translation systems separated by billions of years of evolutionary divergence.

Introduction

The existence of multiple genomes within eukaryotic cells necessitates multiple gene expression systems. Protein synthesis occurs separately in endosymbiotically-derived organelles (mitochondria and plastids) and the cytosol. The coding capacity of mitochondrial genomes (mitogenomes) has dwindled from an

estimated thousands of genes in the mitochondrial progenitor to only a few dozen in most eukaryotes (GRAY 2012). Import of gene products encoded in the nuclear genome compensates for many of these organellar gene losses (HUYNEN *et al*. 2013). Yet despite their reduced gene content, almost all mitogenomes still encode at least some components of the translational machinery (SLOAN *et al*. 2018).

Clear patterns have emerged in the retention and loss of certain genes involved in organellar translation systems. The genes encoding the large and small subunit mitochondrial ribosomal RNAs (rRNAs) are almost universally retained in the mitogenome, whereas other gene products, such as aminoacyl tRNA synthetases (aaRSs), are exclusively encoded in the nuclear genome and must be imported into the mitochondrial matrix (SISSLER *et al*. 2005). The genes for transfer RNAs (tRNAs) exhibit much more heterogeneity with respect to their retention in mitogenomes.

A set of 22 mt-tRNA genes that is sufficient to decode all codons has been retained for over 500 million years in some animal mitogenomes (ANDERSON *et al*. 1981; CLARY AND WOLSTENHOLME 1984; PETERSON *et al*. 2004), but the sampling of mitogenomes from diverse taxonomic groups has revealed extensive variation in complements of mt-tRNAs with examples of extreme mt-tRNA loss. Trypanosomatids such as *Leishmania tarentolae* and *Trypanosoma brucei* (SIMPSON *et al*. 1989; HANCOCK AND HAJDUK 1990) entirely lack tRNA genes in their respective mitogenomes, and some bilaterian animals only have a single mt-tRNA gene (LAVROV AND PETT 2016). In contrast, plants have an intermediate and heterogenous set of mt-tRNA genes. The mitogenomes of flowering plants typically have 11-13 native (mitochondrial in origin) tRNA genes in addition to a variable number of intracellularly and horizontally transferred tRNAs from diverse origins including plastids, bacteria, fungi, and the mitochondria of other plants (IAMS *et al*. 1985; SMALL *et al*. 1999; KUBO *et al*. 2000; RICE *et al*. 2013; KNIE *et al*. 2015; SANCHEZ-PUERTA *et al*. 2019; WARREN AND SLOAN 2020b). Despite this mosaic of mt-tRNA genes, no plant species appears to have a sufficient set in its mitogenome to decode all amino acids (MICHAUD *et al*. 2011b).

The import of nuclear-encoded tRNAs in mitochondria has been demonstrated in a handful of plant species and is assumed to compensate for the insufficient coding capacity of mt-tRNAs in the

mitogenome (CECI *et al*. 1995; KUMAR *et al*. 1996; BRUBACHER-KAUFFMANN *et al*. 1999; DUCHÊNE AND MARECHAL-DROUARD 2001; GLOVER *et al*. 2001). Because most plants still retain many mt-tRNA genes, they require only a specific subset of imported tRNAs for mitochondrial function. The import of nuclear-encoded tRNAs into plant mitochondria from the cytosol has been shown to be complementary rather than redundant, meaning that tRNAs corresponding to the genes functionally lost from the mitogenome are selectively imported (KUMAR *et al*. 1996; BRUBACHER-KAUFFMANN *et al*. 1999; SALINAS-GIEGÉ *et al*. 2015). This specificity of cytosolic tRNA import even extends beyond the anticodon level within isoacceptor families, as tRNA-Gly transcripts have been shown to be differentially imported or excluded in *Solanum tuberosum* mitochondria depending on the anticodon (BRUBACHER-KAUFFMANN *et al*. 1999). In addition, tRNA import has been shown to be associated with translation optimization and codon usage in the green alga *Chlamydomonas reinhardtii* (VINOGRADOVA *et al*. 2009).

Heterogeneity in mt-tRNA content appears to arise over short evolutionary timeframes in plants, raising questions about the evolutionary dynamics of functional replacement of tRNAs in light of the observed import specificity. The sequencing of mitogenomes from the angiosperm genus *Silene* revealed multiple species with greatly reduced numbers of tRNA genes (Fig. 3.1) (SLOAN *et al*. 2012a). Many of these mitochondrial gene losses appear to have happened recently as the closely related species *S. latifolia*, *S. vulgaris*, *S. noctiflora*, and *S. conica* had nine, four, three, and two mt-tRNA genes respectively (Fig. 3.1). These four species represent different sections within the same subgenus and have been difficult to resolve phylogenetically (JAFARI *et al*. 2020), making their precise history of shared and independent tRNA-gene losses unclear.

This rapid and ongoing loss of mt-tRNAs in *Silene* presents a unique opportunity to study the evolutionary mechanisms involved in the functional replacement of mt-tRNAs with nuclear-encoded counterparts. However, characterizing the mt-tRNA pools (including both imported and mitochondrially encoded tRNAs) presents numerous challenges—the first being the intrinsic difficulty in sequencing tRNAs. The vast majority of RNA-seq methods require the reverse transcription of the molecules before sequencing, but the extensive base modifications in tRNAs can stall or terminate reverse transcriptase

enzymes (WILUSZ 2015b). Additionally, the tightly base paired 5′ and 3′ termini of tRNAs can prevent the ligation of sequencing adapters (SHIGEMATSU *et al*. 2017). To date, the sequencing of purified mt-tRNA pools has only been done in humans (MERCER *et al*. 2011). A second challenge in mt-tRNA genetics arises from the degenerate nature of mt-tRNAs. Although mt-tRNAs in plants generally have canonical shape (i.e. a cloverleaf-shape with a D-arm, a T-arm, an acceptor stem, an anticodon stem, and associated loops (RAINA AND IBBA 2014)), mt-tRNAs from other lineages (particularly metazoans) can have such aberrant structure that determining the number and location of mt-tRNAs is unreliable using only genomic data and predictive software (BOORE *et al*. 2005; JUHLING *et al*. 2009; WARREN AND SLOAN 2020a). The existence of degenerate mt-tRNAs in some lineages raises the possibility that putatively lost tRNA genes in *Silene* are still present in the mitogenome but simply undetected. A third challenge is that the majority of tRNA import research in plants has relied on northern blot analysis (KUMAR *et al*. 1996; DELAGE *et al*. 2003). This approach requires previously predicted tRNA sequences for hybridization and can lack clear resolution of probed sequences because of cross-hybridization to undesired targets. Therefore, multiple questions remain about the composition of tRNA populations in the plant mitochondrial matrix.

Here, we trace the evolutionary history of a plant lineage with extreme mt-tRNA loss using recently developed tRNA-seq methods and mitogenome sequencing. We provide the first global sequencing of tRNA pools from isolated plant mitochondria, characterizing the expression of mt-tRNA genes as well the population of imported cytosolic sequences. Using these analyses, we address the following questions: First, have *Silene* mitogenomes really experienced an extreme reduction in mt-tRNA content or do they contain "hidden" genes that are unrecognizable with standard annotation methods? If mt-tRNAs have been lost, are these genes functionally replaced by the import of cytosolic tRNAs? And finally, do *Silene* species preserve the import specificity that has been demonstrated in other plant species, maintaining the avoidance of functionally redundant tRNA import?

This study supports the conclusion from previous mitogenome annotations that *Silene* species have experienced extensive mt-tRNA gene loss and reveals that there are widespread shifts in the abundance of certain nuclear-encoded tRNAs in isolated mitochondria across *Sileneae* species – suggesting that import of tRNAs from the cytosol compensates for mt-tRNA gene losses. Furthermore, we show instances where tRNAs with anticodons still encoded in the mitogenome appear to be imported, indicating a loss of the strict complementary in tRNA import previously seen in plant mitochondria.

Results

Tracing the evolution of mt-tRNA gene loss in Sileneae by sequencing the mitochondrial genomes of Agrostemma githago and Silene chalcedonica

All previously sequenced *Silene* mitogenomes have been found to have extensive mt-tRNA gene loss (SLOAN *et al*. 2012a); however, all of the species that have been sequenced are from a single subgenus, *Behenantha*. In order to better reconstruct the history and timing of mt-tRNA gene losses, we expanded mitogenome sampling by sequencing a representative of the subgenus *Lychnis* (*Silene chalcedonica*) (JAFARI *et al*. 2020), as well as *Agrostemma githago*, which is also a member of the tribe *Sileneae* (Fig. 3.1). Prior to this work, the closest sequenced relative to the genus *Silene* was *Beta vulgaris* (sugar beet), representing an estimated 70 Myr of divergence from *Silene* (MAGALLON *et al*. 2015). The mitogenome of *B. vulgaris* has a more typical set of 19 mt-tRNA genes (KUBO *et al*. 2000) and lacks the history of gene loss seen in *Silene*.

The assembly of the *A. githago* mitogenome produced a 262,903 bp master circle with a GC content of 44.7%. The assembly included four copies of a 2,427 bp "core" repeat region, with additional repeated sequence flanking this core region but only present in a subset of the four copies. This structure of large, identical repeat regions is similar to the mitogenome of *S. latifolia*, which has six copies of a (different) large repeat (SLOAN *et al*. 2010b). These repeats likely undergo intra- and intermolecular recombination, resulting in multiple genome configurations not depicted by the master circle represented here (Supp. Fig. 3.1).

The protein-coding gene inventory of *A. githago* is similar to that of members of the genus *Silene* (Supp. Fig. 3.1). Interestingly, the intron-containing gene *nad7* appears to be *trans*-spliced in *A. githago*. While the first four exons are found in a standard *cis* configuration, the fifth exon is located elsewhere in the genome (Supp. Fig. 3.1). Other NADH-ubiquinone oxidoreductase genes (specifically *nad1*, *nad2*, and *nad5*) typically contain *trans*-splicing introns (CHAPDELAINE AND BONEN 1991; KNOOP *et al*. 1991; WISSINGER *et al*. 1991; BINDER *et al*. 1992), but this is not the case for *nad7* in most angiosperm mitogenomes. Such transitions from *cis*- to *trans*-splicing are relatively common in plant mitochondria (QIU AND PALMER 2004), and it was recently shown that the Pinaceae has independently evolved *trans*-splicing of this same *nad7* intron (GUO *et al*. 2020).

The complement of 14 mt-tRNA genes in *A. githago* also falls below the typical number encoded in angiosperm mitogenomes, but it is greater than the previously documented mt-tRNA gene content in *Silene* (Fig. 3.1), putting *A. githago* at an intermediate state of mt-tRNA gene loss. There is, however, evidence of independent mt-tRNA gene loss in *A. githago* as both tRNA-Phe and tRNA-Pro have been lost in *A. githago* but they are present in at least one *Silene* species (Fig. 3.1). Two distinct copies of tRNA-Cys are present in *A. githago*: the native mitochondrial copy and the bacterial, horizontally transferred gene that was first reported in the mitogenome of *B. vulgaris* (KUBO *et al*. 2000; KITAZAKI *et al*. 2011) (Fig. 3.1, Supp. Table 3.1). The native tRNA-Gly and tRNA-SerTGA, as well as the plastid-derived tRNA-Met and tRNA-SerGGA, appear to have been lost before the divergence of *Agrostemma* and *Silene* as they are absent from *A. githago* and all currently sequenced *Silene* species (Fig. 3.1). As *A. githago* does have reduced mt-tRNA gene content compared to most angiosperms, the process of tRNA gene loss likely initiated prior to divergence of *Sileneae,* but it is has proceeded to much greater extents in the sampled species from *Silene* subgenus *Behenantha*.

The mitogenome of *S. chalcedonica* was found to be considerably larger than that of *A. githago* with an estimated size of approximately 880 kb and a far more complex repeat structure. Because large, highly repetitive genomes are difficult to “close”, we did not try to produce a master circle assembly for

the species. However, sequencing resulted in high coverage (>50× on average) of assembled mitochondrial contigs, suggesting that we likely captured all sequence content for identification of mt-tRNA genes. *Silene chalcedonica* was found to have 12 predicted mt-tRNA genes (two fewer than in *A. githago*). Multiple tRNA gene losses are shared by other *Silene* species (Fig. 3.1). For example, tRNA-Gln was likely lost before the divergence of *S. chalcedonica* and the other *Silene* species, as none of the sequenced members of the genus have intact copy of the gene. However, mt-tRNA-SerGCU and mt-tRNA-Phe are still present in *S. chalcedonica*, whereas they have been lost in the other sequenced *Silene* species (Fig. 3.1).

Confirming the expression of mitochondrially encoded tRNA genes in Sileneae species

To validate the *in silico* annotations of mitochondrially encoded tRNA genes and characterize global mt-tRNA populations, mitochondria were isolated from leaf tissue of five *Sileneae* species (*A. githago, S. conica, S. latifolia, S. noctiflora, and S. vulgaris*) using differential centrifugation and discontinuous Percoll gradients. Paired, total-cellular RNA (whole leaf tissue) samples were processed in parallel with each mitochondrial sample. tRNAs were then sequenced with recently developed tRNA-seq protocols including treatment with the dealkylating enzyme AlkB to remove base modifications that inhibit reverse transcription (COZEN *et al*. 2015; ZHENG *et al*. 2015) and the use of adapters with complementarity to the CCA-tail found on all mature tRNAs (SHIGEMATSU *et al*. 2017). Sequences were mapped to annotated mt-tRNA genes, as well as the entire mitogenome of each respective species to search for previously undetected tRNAs. Reads were also mapped to tRNA genes predicted from nuclear assemblies of the five species by tRNAscan-SE (CHAN AND LOWE 2019).

As described in detail below, the purified mitochondrial samples showed strong enrichment for mitochondrially encoded tRNAs (Fig. 3.2), confirming the efficacy of the mitochondrial isolations. All mt-tRNAs previously predicted to be functional in *S. conica, S. latifolia, S. noctiflora and S. vulgaris* were found to be expressed (Supp. Table 3.2). Two *A. githago* mt-tRNAs (tRNA-Asn, and tRNA-Arg)

with identical sequence identity to plastid copies were depleted in mitochondrial isolates (Supp. Table 3.3), suggesting that these tRNAs may not be functional in the mitochondrial matrix. Both tRNAs have 100% sequence identity to plastid tRNAs and are present within a larger plastid-derived sequence block (i.e., a "mtpt"), indicating that these sequences are recent insertions of plastid DNA. The tRNA-Asn gene is likely different from the ancestral plastid-derived tRNA-Asn copy found in most angiosperm mitogenomes (Fig. 3.1), which can be differentiated from plastid copies by mitochondrial-specific sequence variants (RICHARDSON *et al*. 2013). The lack of detection of both genes suggests that that reduction in mt-tRNA gene content in *A. githago* may be more extreme than apparent from genomic data alone. Differential expression (enrichment) analysis for all mitochondrial genes can be found for each species in Supp. Tables 3.3-3.8.

The post-transcriptional addition of a CCA tail is widespread in Sileneae *mitochondria and occurs on multiple non-tRNA transcripts*

Although the vast majority of reads from the mitogenome mapped to an annotated tRNA gene, a small percentage mapped to other positions (Table 3.1, Supp. Table 3.9). These reads frequently had CCA nucleotides at a terminus and occurred at the boundaries of rRNA or protein coding genes (Supp. Table 3.9). For some of these sequences, the CCA nucleotides appear to be post-transcriptionally added as the sequence is not genomically encoded. Many of these reads formed stem-loop structures that act as t-element processing signals previously described in other angiosperms (FORNER *et al*. 2007; VARRE *et al*. 2019). T-elements are tRNA-like sequences bordering other genes that act as signals for endonucleolytic cleavage by RNase Z and/or RNase P for mRNA or rRNA transcript maturation (FORNER *et al*. 2007). A conserved t-element structure internal to the annotated *nad6* open reading frame was expressed and post transcriptionally modified with a CCA-tail in *A. githago*, *S. latifolia*, and *S. vulgaris* (Supp. Table 3.9, Supp. Fig. 3.2).

Although most of these stem-loops lacked canonical tRNA structure, they did have stem structure at the 5′- and 3′-termini (Supp. Fig. 3.2). The CCA-tailing of stem-loop structures has been reported in numerous other mitochondrial systems including rat (YU *et al*. 2008), human (PAWAR *et al*. 2019), mites (WARREN AND SLOAN 2020a), and other plants (WILLIAMS *et al*. 2000b). Interestingly, the species with the fewest mt-tRNAs (*S. conica* and *S. noctiflora*) also had the fewest of these detected stem-loops (Table 3.1, Supp. Table 3.9). Only 0.0005-0.0314% of the CCA-containing reads in these two species mapped to a non-tRNA position in the mitogenome, and these reads almost exclusively originated from the boundaries of rRNA genes (Supp. Table 3.9). In comparison, the percentage of reads mapping to non-tRNA position in the mitogenomes in the other three species ranged from 0.0438% up to 4.4819%. The relatively abundant reads mapping to a stem-loop structure in *S. vulgaris* derived mainly from a region of the mitogenome with two identical step-loop sequences that were not immediately up- or downstream of an annotated gene (over 6 kb away from *cob* and over 4.5 kb away from the fourth exon of *nad5*). *Silene vulgaris* also had a relatively highly abundant stem-loop structure originating from a tRNA-fMet gene (over 51,000 reads combined in all six libraries). This sequence is annotated as a pseudogene as it lacks multiple features of a tRNA, and it is predicted to form an arm-less, degenerate stem-loop structure with multiple bulges (Supp. Fig. 3.2). It is only 58 bp away from *atp6*, so it is likely that it now serves as a t-element maturation signal.

One stem-loop structure mapped to the mitogenome of *A. githago* did have a cloverleaf-like structure similar to a canonical tRNA shape but did not correspond to an annotated tRNA gene (Fig. 3.3, Supp. Table 3.2a). Predictive folding under a maximum free energy model puts the anticodon as tRNA-LysUUU (Fig. 3.3). Unlike t-elements, this stem-loop is not closely associated with any annotated gene, as it is 2,414 bp upstream of the nearest gene (the fourth exon of *nad1*, Supp. Table 3.9). Homologous sequences are present in the mitogenome of *Beta vulgaris* (GenBank: BA000024) and *Chenopodium quinoa* (GenBank: NC_041093), as well as copies located within a region of mitochondrial-like sequence in the nuclear genome assembly of *Linum usitatissimum* (GenBank: CP027627). However, this sequence

is not present in the mitogenomes of any of the sequenced *Silene* species. Although the evolution of novel tRNAs through the duplication of existing tRNA genes has been shown phylogenetically and experimentally (AYAN *et al*. 2020), this sequence in *Agrostemma* does not have identifiable homology to any known tRNA and does not appear to form a tRNA-like structure in *B. vulgaris* or *C. quinoa*. Whether this is just a non-functional stem-loop that happened to converge on a predicted cloverleaf structure or a functional decoding molecule representing the *de novo* birth of a tRNA gene is not clear.

For multiple reasons, it is unlikely that the vast majority of the "stem-loops" detected in this dataset represent functional tRNAs. First, with the exception of the tRNA-like Lys structure in *A. githago*, these stem-loops have structural deformities that would be generally inconsistent with tRNA function, including multiple mismatches, a lack of a cloverleaf or L-shape, and incorrectly sized anticodon loops. Even in extremely aberrant metazoan mt-tRNAs that lack arms, an L-shape tertiary structure is still expected (WATANABE *et al*. 2014; JUHLING *et al*. 2018). Mismatched bases in stems of plant mt-tRNAs have been predicted from genomic data but found to be corrected post-transcriptionally (MARECHAL-DROUARD *et al*. 1993; BINDER *et al*. 1994). The sequenced stem-loops found in this study maintained these base mismatches. Secondly, many of these structures originate from known t-element regions or are directly up or downstream of other genes. T-elements are broadly present in angiosperms (including those with much larger mt-tRNA gene sets) and are known substrates for other tRNA-processing machinery (i.e., RNase P/Z) (FORNER *et al*. 2007). Thus, it is likely that these sequences are just processed with a CCA-tail as a byproduct. Lastly, these stem-loop sequences were of low abundance and almost absent in the species with the fewest mt-tRNA genes. Therefore, it does not appear that the apparent mt-tRNA gene loss seen in *Sileneae* species can be explained by a large class of previously undetected tRNA genes in the mitogenome.

Characterization of imported tRNA pools in Sileneae *mitochondria reveals extensive sharing of the same nuclear-encoded tRNAs between the cytosol and the mitochondria*

The average enrichment of mitochondrial-encoded tRNAs in purified *Sileneae* mitochondrial fractions was anywhere from 37- to 146-fold relative to total-cellular samples (Supp. Tables 3.3-3.8) depending on the species. Almost no nuclear-encoded tRNAs in any of the five *Sileneae* species reached the enrichment level of the mitochondrially-encoded tRNAs (Fig. 3.2), suggesting that very few, if any, nuclear-encoded tRNAs exclusively function in the mitochondria. Therefore, nuclear-encoded tRNA that are imported into the mitochondria must also be present in substantial abundances in the cytosol or elsewhere in the cell. The most notable exception was a low-abundance nuclear-encoded tRNA in *S. vulgaris* with homology to tRNA-AlaCGC. This tRNA was part of a family of related genes (Supp. Table 3.1, Reference ID#s 964, 1005-1014) with multiple sequence variants (Fig. 3.4) and reached comparable enrichment levels to those observed for mitochondrial genes (Fig. 3.2). The anticodon of these tRNAs is uncertain, as an insertion in the anticodon loop may change the anticodon sequence from CGC to ACG (tRNA-Arg) or possibly GCA (tRNA-Cys), depending on how the tRNA folds *in vivo* (Fig. 3.4, Supp. Fig. 3.4). Multiple tRNAs originating from these genes were found to be enriched in mitochondrial isolates, but detection was low and only one (nuclear ID: 1009, Supp. Table 3.1) survived the minimum read count threshold for differential expression analysis (Fig. 3.2, Supp. Table 3.8). Both tRNA-Arg and tRNA-Ala are expected to be imported into all angiosperm mitochondria (MICHAUD *et al*. 2011b), and tRNA-Cys is expected to be imported into *S. vulgaris* because of recent loss of a mt-tRNA-Cys gene, but why enrichment of these noncanonical tRNAs would be so high in mitochondrial isolates is unclear. It is possible that this is not a case of exclusive import but instead a failure of quality control, as import may have occurred prior to degradation, and the cytosolic surveillance mechanisms that degrade tRNAs without proper structures (DEWE *et al*. 2012) may not be active in the mitochondrial matrix. Other explanations for this enrichment are possible, but regardless these genes were the exception, as there was a clear distinction in the enrichment of mt-tRNAs versus the nuclear-encoded tRNA populations (Fig. 3.2).

Confirming the utility of tRNA-seq methods to detect the import of nuclear-encoded tRNAs by sequencing mitochondrial tRNA pools in Solanum tuberosum

Because *S. tuberosum* (potato) has had the most extensive prior characterization of mitochondrial tRNA import using northern blot analysis (MARECHAL-DROUARD *et al*. 1990; KUMAR *et al*. 1996; BRUBACHER-KAUFFMANN *et al*. 1999; DELAGE *et al*. 2003; PUJOL *et al*. 2008), we sequenced the mitochondrial tRNA pools in *S. tuberosum* in order to compare the effectiveness of detecting tRNA import using the tRNA-seq methods employed in this study. Mitochondria were isolated from young *S. tuberosum* leaves with the same protocol and Percoll gradients used for *Sileneae* species.

Enrichment of *S. tuberosum* mitochondria in the isolated fractions was apparent from high (21-fold-change) relative abundance of mt-tRNA and mitochondrial stem-loop sequences compared to the total-cellular samples (Supp. Fig. 3.3, Supp. Table 3.6). Likewise, plastid tRNAs were greatly depleted relative to mitochondrial tRNAs, indicating that the preparation of the mitochondrial isolates was also effective (although not perfect) at removing plastid tRNAs (Table 3.2). As there is no known import of plastid tRNAs into mitochondria, any detection of plastid tRNAs in mitochondrial fractions likely represents contaminating plastids in the mitochondrial gradient purification.

Previous experiments testing the exclusion of cytosolic tRNAs in *S. tuberosum* (BRUBACHER-KAUFFMANN *et al*. 1999) reported a strong exclusion of tRNAs with the same anticodon as those that are still encoded in the mitogenome. In order to compare the sequence data generated here to previously performed hybridization methods (which detects the abundance of multiple tRNAs with the same anticodon), we calculated an average mitochondrial enrichment ($\log_2$ fold-change) weighted by expression of all tRNA sequences sharing the same anticodon (Supp. Table 3.10). The resulting mitochondrial enrichment values for tRNAs grouped by anticodon were highly consistent with previous northern blot results on enrichment and depletion of nuclear-encoded tRNAs in the mitochondrial fraction (Supp. Table 3.11). The seven tRNAs previously shown to be imported almost universally had higher enrichment values than the 17 tRNAs previously shown to be excluded – the only exception being that that tRNA-Gln(TTG) (excluded) had a slightly higher enrichment value than tRNA-Thr(TGT) (imported) (Supp. Table 3.11). The support for expected import and exclusion patterns for cytosolic tRNAs in *S.*

tuberosum included clear differential import of tRNA-Gly(CCC) and tRNA-Gly(UCC) transcripts but exclusion of tRNA-Gly(GCC) (Supp. Table 3.11) (BRUBACHER-KAUFFMANN *et al*. 1999).

Because only a subset of *S. tuberosum* tRNAs have been previously assayed for mitochondrial import by northern blot, we extended our analysis of the tRNA-seq data by predicting import/exclusion status for the remaining anticodons. An isodecoder family was predicted to be excluded if a tRNA for that anticodon was still encoded in the *S. tuberosum* mitogenome or imported if there was no corresponding tRNA encoded in the mitogenome. Correspondence between tRNA-seq enrichment/depletion and expected import status was highly significant (i.e., tRNAs that were predicted to be imported were significantly more enriched in the mitochondrial fraction than those expected to be excluded; Welch's t-test, *p*-value = 3.24e-07 Fig. 3.5).

Extensive shifts in mitochondrial import and exclusion of tRNAs in Sileneae with possible cases of redundant tRNA import

We wanted to test whether rapid mt-tRNA gene loss in *Sileneae* precipitated shifts in cytosolic tRNA trafficking through changes in the import and exclusion of certain tRNA families. Like in *Solanum*, a signal of tRNA exclusion and import was also recovered for *Sileneae* species, with tRNA exclusion corresponding to some anticodons retained in the mitogenomes (i.e., a signal of non-redundant import of nuclear-encoded tRNAs). For all five *Sileneae* species, cytosolic tRNAs that were predicted to be excluded from mitochondria because they share an anticodon with a mitochondrially encoded tRNA showed lower relative abundance on average than those predicted to be imported (Fig. 3.5). This difference was not statistically significant for *S. conica* and *S. vulgaris* (Fig. 3.5), but that may simply reflect the limited statistical power in species where so few tRNAs are predicted to be excluded because of extensive tRNA gene loss from the mitogenome. Even in these species, the cytosolic tRNAs that corresponded to the few remaining anticodons remaining in the mitogenomes were often the most depleted tRNAs (Supp. Tables 3.3-3.8, and Supp. Table 3.10). For example, cytosolic tRNA-Ile(TAT) was always one of the most strongly excluded tRNA families in all species, and the functionally

equivalent mt-tRNA-Ile(CAT) is the only mt-tRNA gene retained in all species tested in this analysis (Fig. 3.1).

We predicted that *Sileneae* species that had lost a mt-tRNA gene would exhibit evidence of increased import of the corresponding nuclear-encoded tRNA from the cytosol. Overall, of the 12 mt-tRNA genes lost in at least one *Silene* species since divergence from *Agrostemma*, seven corresponding cytosolic tRNAs (tRNA-Asp(GTC), tRNA-Cys(GCA), tRNA-Gln(TTG), tRNA-Glu(TTC), tRNA-His(GTG), tRNA-Trp(CCA) and tRNA-Tyr(GTA)) showed increased representation in the mitochondrial fraction in all species that had lost the mt-tRNA counterpart compared to both *Solanum* and *Agrostemma* (Figs. 3.6-3.7, Supp. Fig 3.4). Conversely, there were no cases showing the opposite pattern, where all *Silene* species with a mt-tRNA gene loss had lower enrichment than both *Solanum* and *Agrostemma* for the corresponding cytosolic tRNA. In addition, there are two mt-tRNA genes that were lost in the *Sileneae* lineage prior to the divergence of *Agrostemma* and *Silene*: tRNA-Gly(GCC) and tRNA-Met(CAT) (Fig. 3.1). In both of these cases, the corresponding cytosolic isodecoder families showed higher mitochondrial enrichment in all five *Sileneae* species than *Solanum* (Supp. Fig. 3.5). In contrast to these tRNAs with a recent history of loss of the corresponding mt-tRNA genes in *Silene*, the nuclear-encoded tRNAs that are expected to be universally imported in angiosperms because of more ancient losses from the mitogenome did not show consistent trends towards mitochondrial enrichment in *Silene* relative to *Agrostemma* and *Solanum* (Supp. Fig. 3.5)

In some cases of mt-tRNA gene loss, we did not see expected changes in import. For example, the cytosolic initiator tRNA-Met did not exhibit positive enrichment in any species (Supp. Fig. 3.4), including *S. vulgaris* which lacks and intact copy of the functionally equivalent tRNA-fMet from its mitogenome – possibly suggesting that initiator tRNA-Met may not be used for mitochondrial translation initiation even in systems that lose tRNA-fMet. Multiple tRNA-Ser anticodons failed to show a consistent pattern of increased enrichment despite multiple losses of tRNA-Ser genes in *Sileneae* mitogenomes but did exhibit substantial heterogeneity in enrichment across the different species (Supp. Fig. 3.5).

Interestingly, there were multiple cases of increased mitochondrial enrichment of a cytosolic tRNA in a *Sileneae* species that still retained the corresponding mt-tRNA gene. Such cases suggest that gain of mitochondrial import may have occurred ancestrally, followed by a period of redundant presence of mitochondrially encoded tRNAs and imported cytosolic tRNAs with the same anticodon. For tRNA-His(GTG) in *S. latifolia*, tRNA-Glu(TTC) in *S. vulgaris*, and tRNA-Asn(GTT) in *A. githago* (Figs. 3.6-3.7, Supp. Fig. 3.4 and 3.5), there was a positive enrichment value despite the species retaining a mitochondrial counterpart in the mitogenome. These would represent some of the first reported cases of redundant tRNA import as previously tested plant species have shown strictly complementary import of tRNAs in plant mitochondria (KUMAR *et al*. 1996; BRUBACHER-KAUFFMANN *et al*. 1999; DUCHÊNE AND MARECHAL-DROUARD 2001; DELAGE *et al*. 2003).

Discussion

The import of tRNAs across mitochondrial membranes occurs in various eukaryotic lineages (SALINAS-GIEGÉ *et al*. 2015), but questions remain about the evolutionary mechanisms that facilitate the functional replacement and integration of nuclear-encoded tRNAs in mitochondrial translation.

Except for a handful of tRNAs related to tRNA-AlaCGC in *S. vulgaris*, we did not find nuclear-encoded tRNAs enriched to the same level as those encoded in the mitogenome (Fig. 3.2). This result suggests an absence of dedicated tRNAs for mitochondrial import in these systems with recent mitochondrial tRNA gene replacement. The widespread sharing of nuclear-encoded tRNAs between both compartments is notably different than the imported mitochondrial proteome, which has been estimated to only be 15% dual-localized between the mitochondria and another cellular location (CALVO AND MOOTHA 2010). This lack of tRNA specialization contrasts with mitochondrial tRNA import studies performed in the green alga *Chlamydomonas* in which multiple nuclear-encoded tRNAs had varying degrees of mitochondrial specialization, and a nuclear-encoded tRNA-Lys(UUU) was found to be exclusively localized to the mitochondria (VINOGRADOVA *et al*. 2009).

One caveat to this interpretation is that, even if tRNAs are destined for exclusive function in the mitochondria, there may be a substantial fraction "en route" at any given point in time. Therefore, they would still be present in substantial quantities elsewhere in the cell, producing a muted mitochondrial enrichment signal. This caveat aside, the dual localization of tRNAs between the cytosol and mitochondria represents a large-scale homogenization of host (nuclear) and endosymbiont (mitochondrial) translational systems.

Sequencing tRNA pools from isolated mitochondria of multiple *Sileneae* species revealed substantial changes in the relative abundance of nuclear-encoded tRNAs corresponding to mt-tRNA gene loss (Figs. 3.6-3.7 and Supp. Fig. 3.4). These fluctuations occurred on phylogenetically recent timeframes, suggesting that changes to tRNA import specificity can occur rapidly. However, the *de novo* evolution of tRNA import did not appear to be a perfectly binary switch from excluded to imported. There were cases where a nuclear-encoded tRNA family showed partially elevated abundance in the mitochondrial fraction of a species that still retained a mitochondrially encoded counterpart (see tRNA-Glu and tRNA-His in Fig. 3.6). Such potential losses of exclusion signals prior to mt-tRNA gene loss may point to a model where the evolution of tRNA import is gradual process whereby some cytosolic "leakiness" is tolerated. If these imported tRNAs are functional, we hypothesize that leaky tRNA import may provide the necessary first step in tolerating mt-tRNA gene losses by providing functional redundancy. And finally, it is interesting to note that some of the mt-tRNAs most recalcitrant to loss in *Silene* (tRNA-Ile, tRNA-Trp, and tRNA-Tyr) corresponded to the most excluded cytosolic tRNAs in *Solanum*. It may be that certain tRNAs interact with additional enzymatic partners in the cell during cellular trafficking resulting in strong exclusion, and successful delivery to the mitochondria would require additional coevolutionary hurdles to overcome before functional replacement.

tRNAs are not standalone molecules in the process of translation, as entire networks of enzymes must interact with tRNAs for maturation and function. Because some tRNA-interacting enzymes are highly specific (CHIMNARONK *et al*. 2005; SALINAS-GIEGÉ *et al*. 2015), separate sets of enzymes are normally required for the maturation and aminoacylation of cytosolic vs. organellar tRNAs (DUCHÊNE *et*

al. 2009). How nuclear-encoded tRNAs are effectively "swapped" into the mitochondrial translational machinery has remained a long-standing question in plant mitochondrial genetics (WARREN AND SLOAN 2020b). Notably, the charging of tRNAs with the correct amino acid is necessary for protein synthesis and relies on aaRSs, which use specific nucleotide identities along the tRNA for recognition and interaction with cognate tRNAs (GIEGE *et al*. 1998; MEIKLEJOHN *et al*. 2013). Without an effective aaRS partner, a newly imported tRNA would be nonfunctional in translation. One possible evolutionary mechanism would involve a cytosolic aaRS also gaining mitochondrial import and charging the cognate, imported tRNA present in the mitochondrial matrix. Massive tRNA loss in the mitogenomes of some non-bilaterian animals (e.g., sponges) has been associated with the loss of mitochondrial aaRSs, suggesting that imported tRNAs are being charged by partnered cytosolic aaRSs (i.e., replacement of both tRNA and aaRS) (HAEN *et al*. 2010a; PETT AND LAVROV 2015). However, this presents what has been previously termed the "chicken and egg" problem in plant mt-tRNA replacement (SMALL *et al*. 1999). Does the aaRS or the tRNA gain import first if nether would be functional without the other? Our results from sequencing of tRNAs from the mitochondrial fractions of *Sileneae* species as well as *S. tuberosum* may provide some insight into this conundrum—cytosolic tRNAs not essential for mitochondrial translation (because of the presence and expression of a native mt-tRNA counterpart) may frequently be found in the mitochondrial matrix before the loss of mt-tRNA genes. It is yet to be determined if these imported tRNAs are functional, and if so, whether a mitochondrial aaRSs has evolved increased substrate recognition to charge both mitochondrial- and nuclear-encoded tRNAs, or if cytosolic aaRSs has also gained mitochondrial import alongside the cognate tRNA.

The homogenization of tRNA metabolism in plant mitochondria represents a much different evolutionary route than that taken by most bilaterian animals, which have retained an independent and conserved set mt-tRNAs with very limited tRNA import into mitochondria (GROSJEAN AND WESTHOF 2016). Fundamental differences in the historical mutation rates of mitogenomes in animals versus plants may have created a situation whereby plant mt-tRNAs are more predisposed to the functional replacement with that of cytosolic counterparts. The rate of mitochondrial sequence evolution in land plants are among

some of the slowest ever estimated (MOWER *et al*. 2007) and has likely been a factor in the maintenance of the largely canonical shape of plant mt-tRNAs. This conserved shape may contribute to interchangeable nature of plant mt-tRNA genes. It is possible that slow rate of evolution in plant mitogenomes has resulted in tRNA enzymatic networks that are more amenable to nuclear-encoded substrates because of shared sequence identity and structure between plant mt-tRNAs and nuclear counterparts. In contrast, the much faster evolving animal mitogenomes (BROWN *et al*. 1979; LAVROV AND PETT 2016) may have experienced early tRNA mutations and compensatory enzymatic substitutions that make nuclear-encoded tRNAs poor substrates for the tRNA-interacting enzymes that function in the mitochondria (WATANABE *et al*. 2014; PONS *et al*. 2019) — thereby locking in independence of the mitochondrial protein synthesis apparatus.

Despite this generally low mutation rate, some plant lineages have experienced a recent and dramatic acceleration in mutation rates in their mitogenomes, including multiple species in *Silene* (MOWER *et al*. 2007; SLOAN *et al*. 2009; BROZ *et al*. 2021). This accelerated rate may have then facilitated a rapid replacement of mt-tRNA genes in some plant lineages. *Silene conica* and *S. noctiflora* both have highly accelerated rates of mitochondrial sequence evolution, as well as the fewest reported mt-tRNA genes in angiosperms, with only two and three mt-tRNA genes, respectively (Fig. 3.1). It has been proposed that high mutation rates increase the likelihood of disruptive mutations in tRNAs and provide a selective pressure to functionally replace mt-tRNA genes with imported nuclear-encoded counterparts – thereby ratcheting a gene replacement process (BRANDVAIN AND WADE 2009). Another plant lineage with a high rate of sequence evolution, *Viscum* (mistletoe), also has a severely reduced mt-tRNA gene content (PETERSEN *et al*. 2015; SKIPPINGTON *et al*. 2015). The heterogeneity in mitogenomes across angiosperm diversity creates exciting opportunities to test these hypothesized relationships between import specificity, mutation rates, and the functional loss/replacement of mt-tRNAs.

Methods

Agrostemma githago and Silene chalcedonica mitochondrial genome sequencing

Mitochondrial DNA was extracted from *A. githago* and *S. chalcedonica* using methods described previously (SLOAN *et al*. 2010b; SLOAN *et al*. 2012b). Tissue was collected from the same maternal families previously used for plastid genome sequencing in these species (SLOAN *et al*. 2014a). Mitochondrial DNA from each species was used to construct 3kb paired-end libraries which were sequenced on a Roche 454 GS-FLX platform with Titanium reagents. Library construction and sequencing were performed at the University of Virginia's Genomics Core Facility. Reads were assembled with Roche's GS de novo Assembler v2.3 ("Newbler") using default settings, resulting in average read depths of >50× for the mitogenomes of each species. The mitogenome of A. githago was then manually assembled into a master circle conformation based on the Newbler assembly graph as described previously (SLOAN et al. 2012b). We did not attempt to manually close the assembly of the larger and more complex S. chalcedonica mitogenome. In order to correct the known high insertion and deletion error rates associated with long homopolymer regions from 454 sequencing, we performed Illumina sequencing on total-cellular DNA from an individual from the same *A. githago* maternal family (Giles County) used for mitogenome sequencing. DNA was extracted from a 3-day old seedling using a CTAB protocol (DOYLE AND DOYLE 1987). An Illumina library was constructed with the NEBNext Ultra II FS DNA Library Prep Kit, using approximately 30 ng of input DNA, a 20 min fragmentation time, and 8 cycles of PCR amplification. The library was dual-indexed and multiplexed with other libraries on a NovaSeq 6000 S4 sequencing lane (paired-end, 150-bp reads) at the University of Colorado Cancer Center, generating 49.6M read pairs. These reads were then mapped to the mitogenome to correct homopolymer length errors in an iterative fashion as described previously (SLOAN *et al*. 2012b).

The *A. githago* mitogenome was annotated with Mitofy (ALVERSON *et al*. 2010), BLAST homology searches, and tRNAscan-SE (CHAN AND LOWE 2019). tRNAscan-SE was also used to identify tRNA genes in the mitochondrial contigs of the *S. chalcedonica* assembly. The mitogenome plot (Supp. Fig. 3.1) was produced with OGDRAW version 1.3.1 (GREINER *et al*. 2019). The annotated *A. githago* mitogenome sequence and the contigs from the *S. chalcedonica* assembly were deposited to GenBank

(accession numbers MW553037 and MW580967-MW581000, respectively). The raw 454 and Illumina reads are available via NCBI SRA (accessions SRX9983705, SRX9983703, and SRX9983702).

Tissue growth for tRNA analysis

Agrostemma githago seeds were obtained from the Kew Gardens Millennium Seed Bank (Kew 0053084), germinated in small plastic pots with Pro-Mix BX General Purpose soil supplemented with vermiculite and perlite, and grown in growth chambers held at 23°C with a 16-hr/8-hr light/dark cycle (light intensity of 100 $\mu E \cdot m^{-2} \cdot s^{-1}$). *Silene latifolia* seeds were from a female derived from the line originally used for mitogenome sequencing (SLOAN *et al*. 2010b) fertilized by a male obtained from the Kew Gardens Millennium Seed Bank (Kew 32982). Leaf tissue for *S. tuberosum* was generated by planting seed potato (tubers) cultivar White Kennebec. *Silene latifolia*, *S. noctiflora* BRP (WU *et al*. 2015), and *S. vulgaris* S9L (SLOAN *et al*. 2012c) seeds were germinated in SC7 Cone-tainers (Stuewe and Sons) with the same soil as above on a mist bench under supplemental lighting (16-hr/8-hr light/dark cycle) in the Colorado State University greenhouse. After germination, *S. latifolia*, *S. noctiflora*, and *S. vulgaris* seedlings were moved to a bench with the same 16-hr/8-hr light/dark cycle supplemental lighting until harvest. *S. conica* ABR (SLOAN *et al*. 2012a) seeds were germinated in the same soil and containers but grown in a growth room under short-day conditions (10-hr/14-hr light/dark cycle) in order to increase rosette growth. And finally, *S. tuberosum* seed potato was planted in the same soil but in one-gallon pots in the same chamber and short-day conditions (10-hr/14-hr light/dark cycle) to promote leaf tissue growth.

Mitochondrial isolations

The age of the plants at harvest time ranged depending on species, as certain species required more growth time to produce a sufficient amount of tissue for mitochondrial isolations. *Agrostemma githago* plants were harvested at 4 weeks, *S. conica* at 14 weeks, *S. latifolia* at 6 weeks, *S. noctiflora* at 14 weeks, *S. vulgaris* at 8 months, and *S. tuberosum* at 4 weeks. 75 g of leaf and stem tissue (entire above ground

tissue) was collected for each replicate for each respective species. This represented 5 potato plants and anywhere from 24 to 65 *Sileneae* plants per replicate. Leaf tissue was disrupted in a Nutri Ninja Blender for 2 x 2-sec short bursts, and 1 x 4-sec blending in 350 mL of a homogenization buffer containing: 0.3 M sucrose, 5 mM tetrasodium pyrophosphate, 2 mM EDTA, 10 mM KH_2PO_4, 1% PVP-40, 1% BSA, 20 mM ascorbic acid, and 5 mM cysteine (pH 7.5-KOH) in a cold room. Homogenate was then poured over four layers of cheese cloth and 1 layer of miracloth and squeezed to filter out all solids. The liquid homogenate was then subjected to differential centrifugation to remove nuclei, plastids, and cellular debris in a Beckman-Coulter Avanti JXN-30 centrifuge with a JA14.50 fixed-angle rotor at 4°C. Homogenate was spun at 500 rcf, 1500 rcf, and 3000 rcf for 10 min at each speed. After each centrifugation step, the supernatant was transferred into a clean centrifuge bottle. Mitochondria were then pelleted by centrifugation for 10 min at 20,000 rcf with the brake off. Supernatant was discarded and the pellet was resuspended using a goat-hair paintbrush and 2 mL wash buffer containing 0.3 M sucrose, 10 mM MOPS, 1 mM EGTA (pH 7.2-KOH). Supernatant was then transferred to 32 mL tubes for centrifugation on a swinging bucket rotor (JS-24.38). The homogenate was centrifuged for 5 min at 3000 rcf to pellet residual plastid and nuclear contaminants, and the supernatant was transferred to a clean 32 mL tube. Mitochondria were once again pelleted by centrifugation for 10 min at 20,000 rcf. The pellet was resuspended with 500 μL wash buffer and a paint brush. The mitochondrial suspension was then added to a glass Dounce homogenizer and homogenized with three strokes.

Homogenized mitochondria were then suspended on top of a Percoll gradient with the following Percoll density layers, 18%, 25%, 50%. The gradient was then centrifuged at 40,000 rcf for 45 min with the brake off. The mitochondrial band at the 25%:50% interface was then aspirated off of the gradient and diluted with 30 mL of wash buffer. The diluted mitochondria were then centrifuged at 20,000 rcf for 10 min. The supernatant was vacuum aspirated, and the mitochondrial pellet was resuspended in a fresh 30 mL of wash buffer and centrifuged at 10,000 rcf for 10 min. The supernatant was vacuum aspirated, and the mitochondrial pellet was resuspended in 1000 μL of fresh wash buffer. Resuspended mitochondria

were centrifuged at 10,000 rcf for 10 min. Supernatant was removed with a pipette and the mitochondrial pellet immediately went into RNA extraction procedures using 1000 μL TRIzol following the manufacturer's RNA extraction protocol.

The total-cellular samples for *A. githago, S. latifolia, and S. noctiflora* were produced by freezing an entire, single plant excluding any root tissue in liquid nitrogen and grinding with a mortar and pestle into powder and performing an RNA extraction using the Trizol manufacturer's protocol. In the case of *S. tuberosum*, *S. vulgaris, and S. conica*, one shoot from one plant with all leaves was used.

Only two replicates of *S. latifolia* were performed because of tissue limitations, and only two mitochondrial isolations from *S. vulgaris* were used in downstream analysis because of an error quantifying the AlkB/RNA ratio in one of the isolations. Two replicates of mitochondrial isolations and total cellular samples were produced for *S. tuberosum*. Three replicates (mitochondrial isolations and total-cellular samples) were produced for all other species.

AlkB treatment and YAMAT-seq library construction

See WARREN *et al*. (2021) for detailed protocols for AlkB expression/purification, RNA treatment and YAMAT-seq (tRNA-seq) library construction. Briefly, a plasmid containing cloned wild type AlkB protein (pET24a-AlkB deltaN11 [plasmid #73622]) was obtained from Addgene (http://www.addgene.org/), and the AlkB protein was expressed and purified at the CSU Biochemistry and Molecular Biology Protein Expression and Purification Facility.

For the demethylation (AlkB) reaction, 6 μg of total RNA was treated with 250 pmols of AlkB for 1 hr at 37°C in a reaction buffer containing 70 μM ammonium iron(II) sulfate hexahydrate, 0.93 mM α-ketoglutaric acid disodium salt dihydrate, 1.86 mM ascorbic acid, and 46.5 mM HEPES (pH 8.0-HCl). The reaction was quenched by adding 4 µl of 100 mM EDTA. RNA was then extracted with a phenol-chloroform RNA extraction, followed by an ethanol precipitation with the addition of 0.08 µg of RNase-free glycogen and resuspension in RNase-free water. RNA integrity was checked on a TapeStation 2200.

To remove amino acids attached to the 3′-end of tRNAs prior to adapter ligation, the demethylated RNA was deacylated in 20 mM Tris HCl (pH 9.0) at 37°C for 40 min. Following deacylation, adapter ligation was performed using a modified protocol from SHIGEMATSU *et al*. (2017). A 9 μl reaction volume containing 1 μg of deacylated RNA and 1 pmol of each Y-5′ adapter (4 pmols total) and 4 pmols of the Y-3′ adapter was incubated at 90°C for 2 min. 1 μl of an annealing buffer containing 500 mM Tris-HCl (pH 8.0) and 100 mM $MgCl_2$ was added to the reaction mixture and incubated for 15 min at 37 °C. Ligation was performed by adding 1 unit of T4 RNA Ligase 2 enzyme (New England Biolabs) in 10 μl of 1X reaction buffer and incubating the reaction at 37°C for 60 min, followed by overnight incubation at 4°C. Reverse transcription of ligated RNA was performed using SuperScript IV (Invitrogen) according to the manufacturer's protocol.

The resulting cDNA was then amplified by PCR using NEBNext PCR Master mix with 7 μl of the reverse transcription reaction. Ten cycles of PCR were performed for *A. githago*, *S. noctiflora,* and *S. vulgaris*. To generate greater library yields, we increased the number of PCR cycles to 12 cycles for *S. conica*, *S. latifolia*, and *S. tuberosum*. PCR was performed on a on a Bio-Rad C1000 Touch thermal cycler with an initial 1 min incubation at 98°C and 10/12 cycles of 30 s at 98°C, 30 s at 60°C and 30 s at 72°C, followed by 5 min at 72°C. Size selection of the resulting PCR products was done on a BluePippin (Sage Science) with Q3 3% agarose gel cassettes following the manufacturer's protocol. The size selection parameters were set to a range of 180-231 bp. Size-selected products were then cleaned using solid phase reversable immobilization (SPRI) beads and resuspended in 10 mM Tris (pH 8.0). Libraries were dual-indexed and sequenced on an Illumina NovaSeq 6000 S4 lane with paired-end, 150-bp reads at the University of Colorado Cancer Center. YAMAT-seq reads for *A. githago*, *S. conica*, *S. latifolia*, *S. noctiflora* and *S. tuberosum* are available via the NCBI Sequence Read Archive under BioProject PRJNA698234 and sequencing reads for *S. vulgaris* are available under BioProject PRJNA662108.

YAMAT-seq read processing

Adapters were trimmed using Cutadapt version 1.16 (MARTIN 2011) with a nextseq trim quality-cutoff parameter of 20 (option: --nextseq-trim=20). A minimum length filter of 50 bp and a maximum of 95 bp was applied to only retain full-length tRNA sequences and remove a substantial number of reads containing only adapter sequence without an insert (SHIGEMATSU *et al*. 2017). BBMerge from the BBTools software package was used to merge R1 and R2 read pairs into a consensus sequence (BUSHNELL *et al*. 2017). Identical reads were summed and collapsed into read families using the FASTQ/A Collapser tool from the FASTX-Toolkit version 0.0.13 (http://hannonlab.cshl.edu/fastx_toolkit/index.html).

Nuclear genome assemblies and reference tRNA genes

In order to generate reference tRNA gene sequences for each species, nuclear genome assemblies were either generated from Illumina short-read datasets or obtained from published sources. Illumina reads for an A. githago nuclear assembly were generated from total-cellular DNA extracted from a 5-month-old plant germinated from a line started from Kew 0053084 seeds. DNA was extracted using a Qiagen DNeasy kit, and the NEBNext Ultra II FS DNA Library Prep Kit was used as described above. Sequencing was performed on an Illumina NovaSeq 6000 S4 Lane with paired-end 150-bp reads at the University of Colorado Cancer Center. Shotgun sequencing reads for *A. githago* are available via the NCBI Sequence Read Archive under BioProject PRJNA698248. Illumina reads for an S. vulgaris were generated from plants were grown from seeds collected from natural populations in Seefield, Austria. Leaf tissue was collected from a single hermaphrodite plant that had been crossed to the F2 generation and grown under standard greenhouse conditions at the University of Virginia. Genomic DNA was extracted with a modified CTAB protocol (DOYLE AND DOYLE 1987) and sent to Global Biologics (Columbia, MO) for the construction of Illumina libraries with 180 bp inserts. Paired-end (2 x100) reads were sequenced at the Yale Center for Genome Analysis on a single Illumina HiSeq 2000 lane. Raw reads have been deposited into the NCBI Short Read Archive (SRA) with accession SRX10073976. *Silene conica* ABR Illumina reads were obtained from a previously published dataset (BROZ et al. 2021).

Genomic reads for *A. githago*, *S. conica*, and *S. vulgaris* were trimmed with cutadapt ver. 1.16 (MARTIN 2011) with a quality-cutoff parameter of 10 for the 3′ end of each read and a minimum length filter of 5 bp. Reads were then assembled using SPAdes v3.11.0, (NURK *et al*. 2013) using the following command options: -k 21,33,55,77,99 -t 40 -m 900. Assemblies for S. noctiflora OPL (GenBank assembly accession: VHZZ00000000.1) and *S. latifolia* (GenBank assembly accession: QBIE00000000.1) were taken from previous studies (KRASOVEC et al. 2018; WILLIAMS et al. 2020).

A database of nuclear tRNA reference genes was produced by searching each of the nuclear assemblies from *A. githago*, *S. conica, S. latifolia*, *S. noctiflora*, and *S. vulgaris* with tRNAscan-SE. ver. 2.0.3 (CHAN AND LOWE 2019) using the General search option (-G). Introns predicted from tRNAscan-SE were removed, and identical tRNA sequences were collapsed into a single reference using custom Perl scripts. It is important to note that plant genomes have multiple copies of identical tRNA genes, but for mapping purposes all of these identical sequences are collapsed into a single reference.

The nuclear references for *S. tuberosum* were obtained from the curated PlantRNA website (http://seve.ibmp.unistra.fr/plantrna/, downloaded Oct. 26, 2020).

Mitochondrial genome mapping and stem-loop references

To search for previously undetected mitochondrial tRNA sequences captured by YAMAT-seq, as well as to identify any expressed stem-loop regions modified with a CCA-tail, all YAMAT reads for each species were mapped to the entire mitogenome using BLAST (blastn, e-value of < 1e-6, low complexity regions not filtered [-dust no]). To ensure only high confidence mitogenome hits, reads mapping to the mitogenome were then filtered so that only reads with greater than or equal to 90% identical matches and hit coverage to a mitogenome location were retained. The “closest” function in bedtools ver. 2.27.1 (QUINLAN AND HALL 2010) was used to assign the closest annotated gene for each of these retained reads, including any reads that mapped/overlapped with a previously annotated tRNA. All reads that did not map to a tRNA gene were then further analyzed to determine if they occurred at the boundary of protein-coding or RNA gene (indicating that these transcripts may be t-elements). Those that did not

occur at a gene boundary were considered to represent a previously undocumented stem-loop or tRNA-like structure. Highly expressed t-elements and stem-loops were added to the mapping database as reference sequences.

Folding predictions of mitochondrial t-elements and stem-loops

Folding prediction analysis for Supp. Fig. 3.2 was done with RNAFold (ver. 2.4.11, (LORENZ *et al*. 2011)), using the using maximum free energy (MFE) model. Because stem-loops also had RT misincorporations at some sites, the YAMAT sequence with 100% nucleotide identity to the mitogenome was used for the folding prediction. Folding diagrams were created with VARNA (ver. 3.9, (DARTY *et al*. 2009)).

Contamination removal

Although the vast majority of the YAMAT reads originated from the sampled plant species, there were numerous reads of contaminating origins such as soil bacteria and plant pests. In order to remove reads that did not originate from the plant itself, all unique sequences with three or more reads were BLASTED to the NCBI nucleotide collection (nt) database (posting date of Oct 27, 2019, downloaded Feb 02, 2020) (blastn, e-value of < 1e-3), and the taxonomy information for the top two hits for each read was pulled using the NCBI taxonomy database (downloaded Feb 02, 2020). All queries that did not have the taxon name Viridiplantae in the first two hits were retained to make a contamination database. GenBank files were downloaded for all contaminating accessions that had more than 50 reads assigned, and tRNA annotations were extracted from those GenBank files. Some accessions did not have annotated tRNA genes, so the contaminating hit was added to the database manually. Many of the contaminating references were identical or nearly identical to each other (e.g., many bacterial tRNA sequences were identical or only differed by slight length variation due to annotations). Thus, the contaminant database was collapsed to retain only unique sequences and only the longest reference of otherwise identical sequences. Contaminating sequences were added to the mapping database as reference sequences. We

also filtered YAMAT-seq reads that did not map to any reference with at least 90% sequence identity and 80% coverage (see below) to ensure that that any low-abundance, contaminating reads were not retained in the analysis.

Read mapping

Mapping of processed reads was performed with a custom Perl script previously published in WARREN *et al*. (2021) that can be found at (https://github.com/warrenjessica/YAMAT-scripts). Briefly, each read was BLASTed (blastn, e-value threshold of < 1e-6) to all references. A mapped read was retained if it had 90% or more sequence identity and at least 80% coverage to a reference sequence. This ensured that full-length tRNA sequences were used in analysis but also allowed for some mismatches due to reverse transcription-induced misincorporations (WARREN *et al*. 2021). Furthermore, only sequences that uniquely mapped to a single reference were retained for downstream analysis. All mapped and filtered reads were then counted for differential expression analysis.

EdgeR differential expression analysis

Differential expression analysis was done on mapped counts with the R package EdgeR version 3.24.3 (ROBINSON *et al*. 2010) using a gene-wise negative binomial generalized linear models with quasi-likelihood tests (glmQLFit). Library sizes were normalized with the function calcNormFactors(y) and low-abundance genes were filtered by expression using the filterByExpr(y) function. Because misincorporations in cDNA are common when reverse transcribing tRNAs, lowly detected (< 2 counts per million or CPM) sequences in this analysis likely represent rare misincorporation profiles of tRNA sequences (and not true unique genes) which occasionally mapped to unique references. To summarize the level of mitochondrial enrichment by anticodon (isodecoder) family, a weighted average enrichment level for all tRNAs with the same anticodon was calculated and presented in Figures 3.6 and 3.7. The average enrichment was calculating by weighting each sequence by its percentage of total expression (CPM) within the isodecoder family. Count data as well as differential expression results can be found in

Supp. Tables 3.2-3.8, and the weighted enrichment for each anticodon group for each species can be found in Supp. Table 3.11.

Funding and support

We would like to thank Laura Bergner for mitochondrial DNA extractions of *Agrostemma githago* and *Silene chalcedonica* and Zhiqiang Wu for generation of preliminary tRNA-seq data. This work was supported by an NSF award (MCB-2048407), a Chateaubriand Fellowship from the Embassy of France in the United States, an NSF GAUSSI graduate research fellowship (DGE-1450032), and by the Centre National de la Recherche Scientifique (CNRS) and the "Initiative d'Excellence" (MITOCROSS ANR-11-LABX-0057).

Data availability statement

Supplementary tables are available at GitHub:

https://gitub.com/warrenjessica/Chap.3_dissertation_supplementary_files

Table 3.1| Percentage of reads with a CCA-tail that mapped to a mitochondrial location other than a tRNA gene. Only sequences that had more than two reads (in all six libraries combined) were used for global mitogenome mapping and this analysis. Only two replicates of *S. latifolia* mitochondrial isolations, *S. latifolia* total-cellular samples, and *S. vulgaris* mitochondrial isolations were performed; see Methods for details.

Species	Mitochondrial isolate 1	Mitochondrial isolate 2	Mitochondrial isolate 3	Total cellular 1	Total cellular 2	Total cellular 3
Agrostemma	0.2316%	0.3290%	0.2057%	0.0085%	0.0072%	0.0086%
S. latifolia	0.0700%	0.0438%	NA	0.0009%	0.0003%	NA
S. vulgaris	4.4819%	NA	3.6261%	0.1084%	0.1553%	0.3070%
S. noctiflora	0.0243%	0.0314%	0.0234%	0.0012%	0.0018%	0.0059%
S. conica	0.0005%	0.0008%	0.0014%	0.0000%	0.0000%	0.0000%

Table 3.2| Depletion of tRNAs from the nuclear and plastid genomes in mitochondrial isolates from six angiosperm species relative to mitochondrially encoded tRNAs. Nuclear tRNA depletion was calculated as the difference between the average $\log_2$ fold change of all mt-tRNA genes (including mitochondrial stem-loops) and the average $\log_2$ fold change of nuclear-encoded tRNAs. Average plastid tRNA depletion was calculated similarly as the difference between the average mt-tRNA $\log_2$ fold change and the average $\log_2$ fold change of plastid tRNAs.

*Note that the depletion values for nuclear-encoded tRNAs do not account for the fact that some of these tRNAs are presumably imported into the mitochondria. Thus, the values reported here likely underestimate the true depletion levels for non-imported, nuclear-encoded tRNAs.

	Average plastid tRNA depletion	**Average nuclear tRNA depletion***
Species	Relative $\log_2$ fold difference	Relative $\log_2$ fold difference
Solanum	-4.43	-4.75
Agrostemma	-6.76	-6.55
S. latifolia	-7.56	-8.57
S. vulgaris	-6.37	-5.78
S. noctiflora	-8.68	-6.99
S. conica	-7.45	-4.88

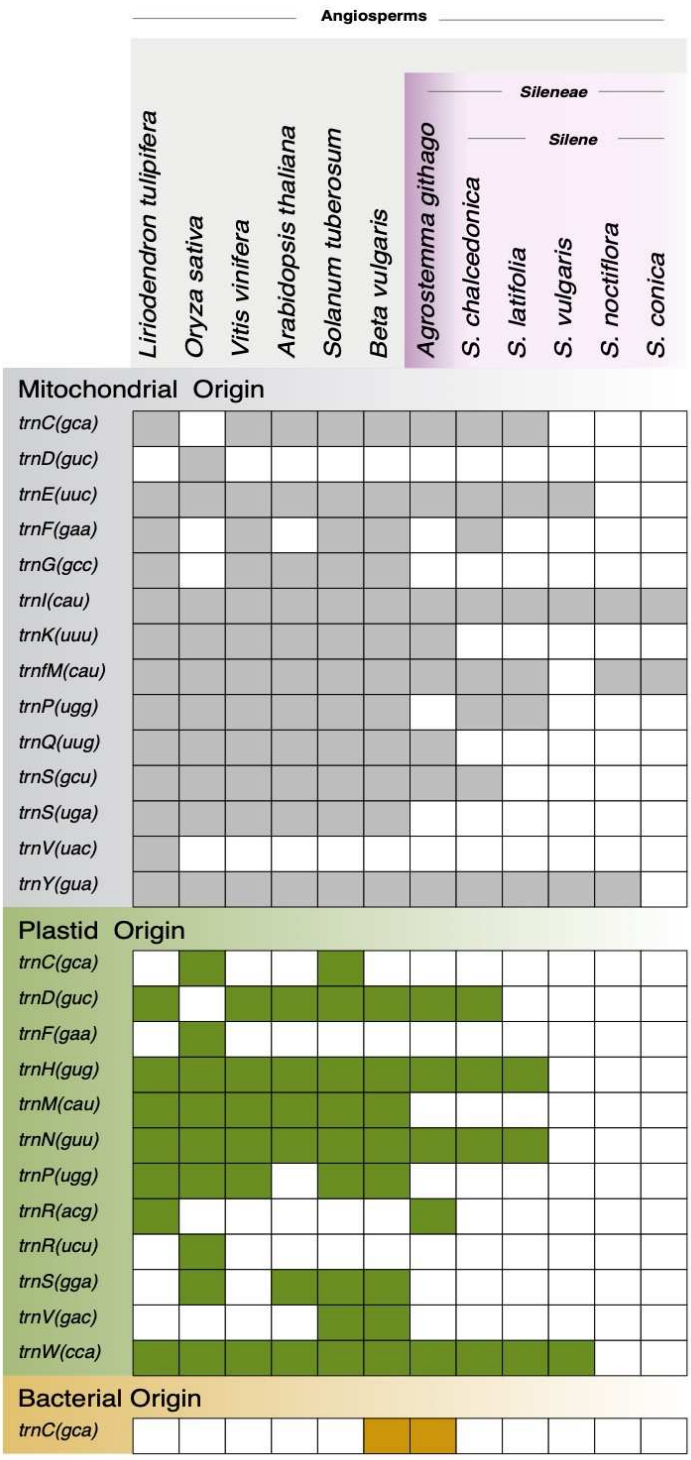


Figure 3.1| Summary of tRNA gene content in a sample of angiosperm mitogenomes with an emphasis on the family Caryophyllaceae. Filled squares indicate the presence of an intact gene sequence, with genes named based on a single-letter amino-acid abbreviation and with anticodon indicated parenthetically. Formyl-methionine is abbreviated fM. Different colors represent the ancestral origin of the tRNA gene. Disrupted or incomplete sequences inferred to be pseudogenes are not included.

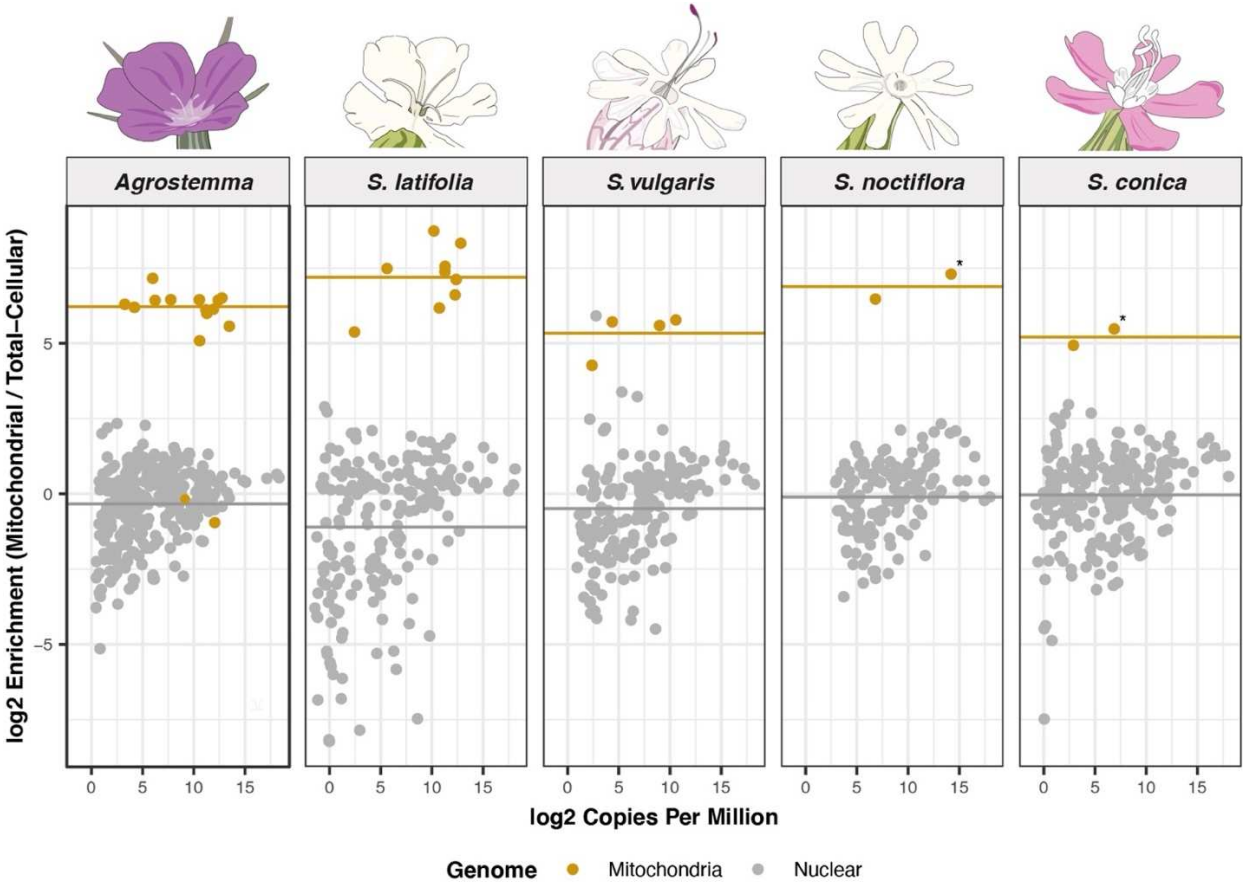


Figure 3.2| Enrichment of mitochondrially encoded and nuclear-encoded tRNAs in mitochondrial isolates relative to total-cellular samples. Each dot represents a unique tRNA sequence. The heavy gray line represents the average enrichment of nuclear-encoded genes. The genomic origin of the tRNA is indicated by color, with gray being nuclear-encoded and gold being mitochondrial. The two mt-tRNAs around the 0-line for *Agrostemma* are tRNA-Asp and tRNA-Arg, which are identical in sequence to the plastid-encoded tRNA counterparts and were not included in the calculation of average enrichment of mitochondrially-encoded tRNAs for the species. Expression of the multiple (but non-identical) copies of formyl-methionine (fMet) tRNAs present in in the mitogenomes of *S. noctiflora* and *S. conica* are marked with an * and were summed for expression analysis. Stem-loops are not shown in this figure.

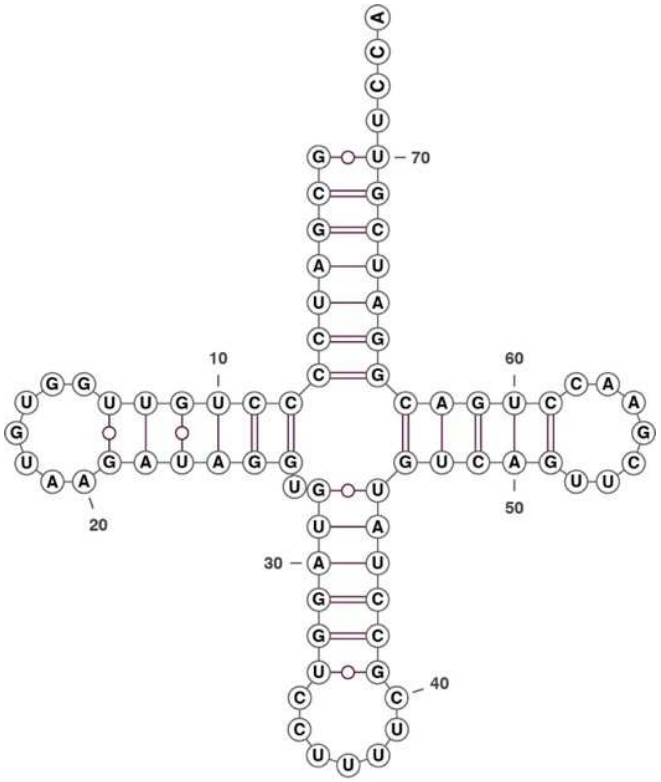


Figure 3.3| Predicted cloverleaf folding transcript expressed and post-transcriptionally modified with a CCA-tail in *A. githago*. Despite the predicted tRNA-like secondary structure, this sequence is not homologous to any previously annotated mt-tRNA genes. Folding represents the maximum free energy folding model using the program the program RNAfold (LORENZ *et al*. 2011), and structure diagrams were created using the program VARNA ver. 3.9 (DARTY *et al*. 2009).

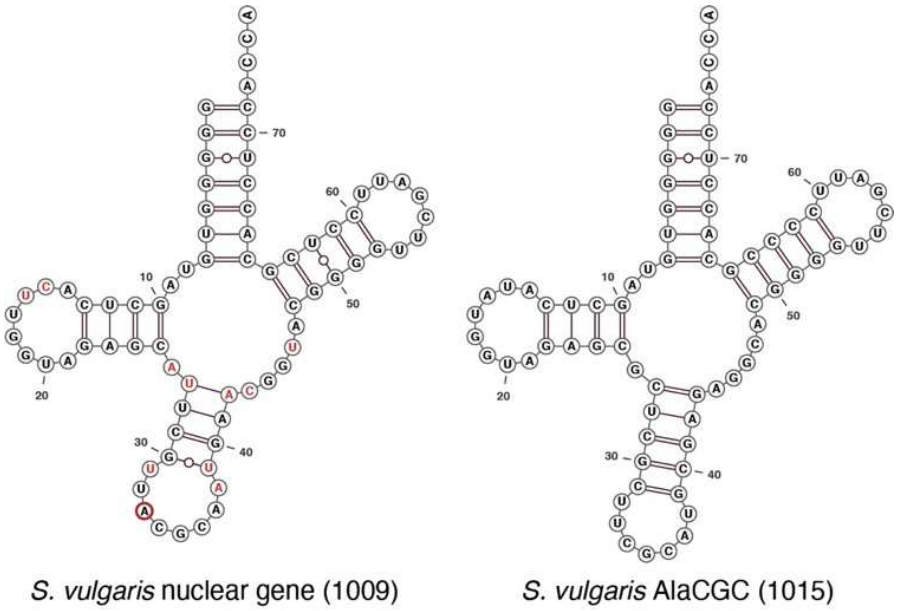


Figure 3.4| Sequence and predictive folding of a highly enriched nuclear-encoded gene (#1009) in *S. vulgaris* mitochondrial isolates and the closest *S. vulgaris* reference with a tRNAscan-determined anticodon. The unique reference ID is in parentheses. Red letters indicate base pair differences between the two tRNAs, and the thick, red circle indicates the insertion in the anticodon mentioned in the main text. Nuclear gene 1009 had multiple more closely related sequences but without definitive anticodons see Supp. Table 3.1 nuclear reference ID#s 964, 1005-1014.

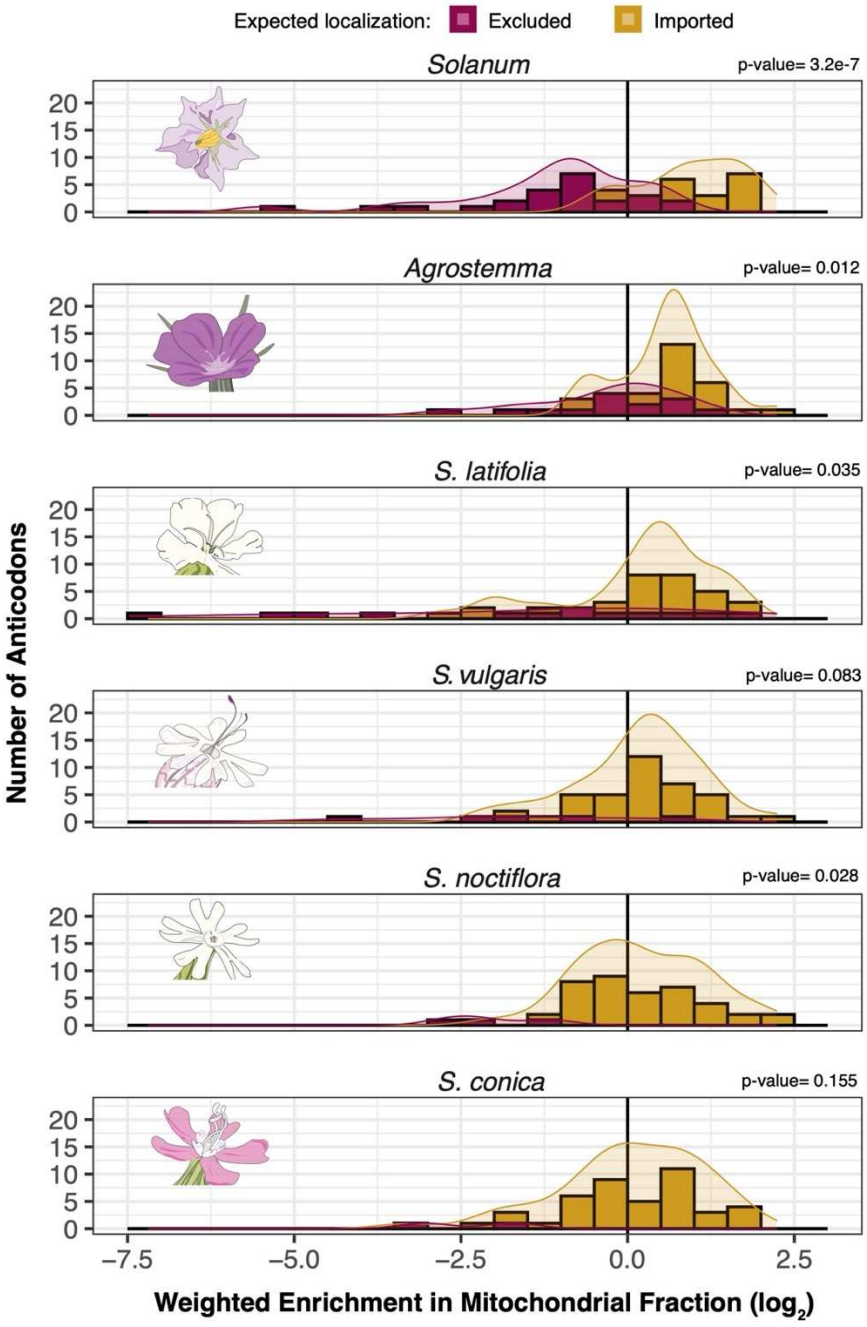


Figure 3.5| Histograms and density curves showing the distribution of average mitochondrial enrichments of nuclear-encoded tRNA anticodon families in *S. tuberosum, A. githago* and four *Silene* species. tRNAs with anticodons that are expected to be imported due to the lack of a corresponding mt-tRNA gene are shown in gold. Those that are expected to be excluded are shown in magenta. Reported *p*-values are based on *t*-tests between these two groups. Enrichment values are calculated based on expression-weighted average of all tRNAs with the same anticodon (see Methods). Positive and negative values indicate enrichment and depletion in the mitochondrial fraction, respectively.

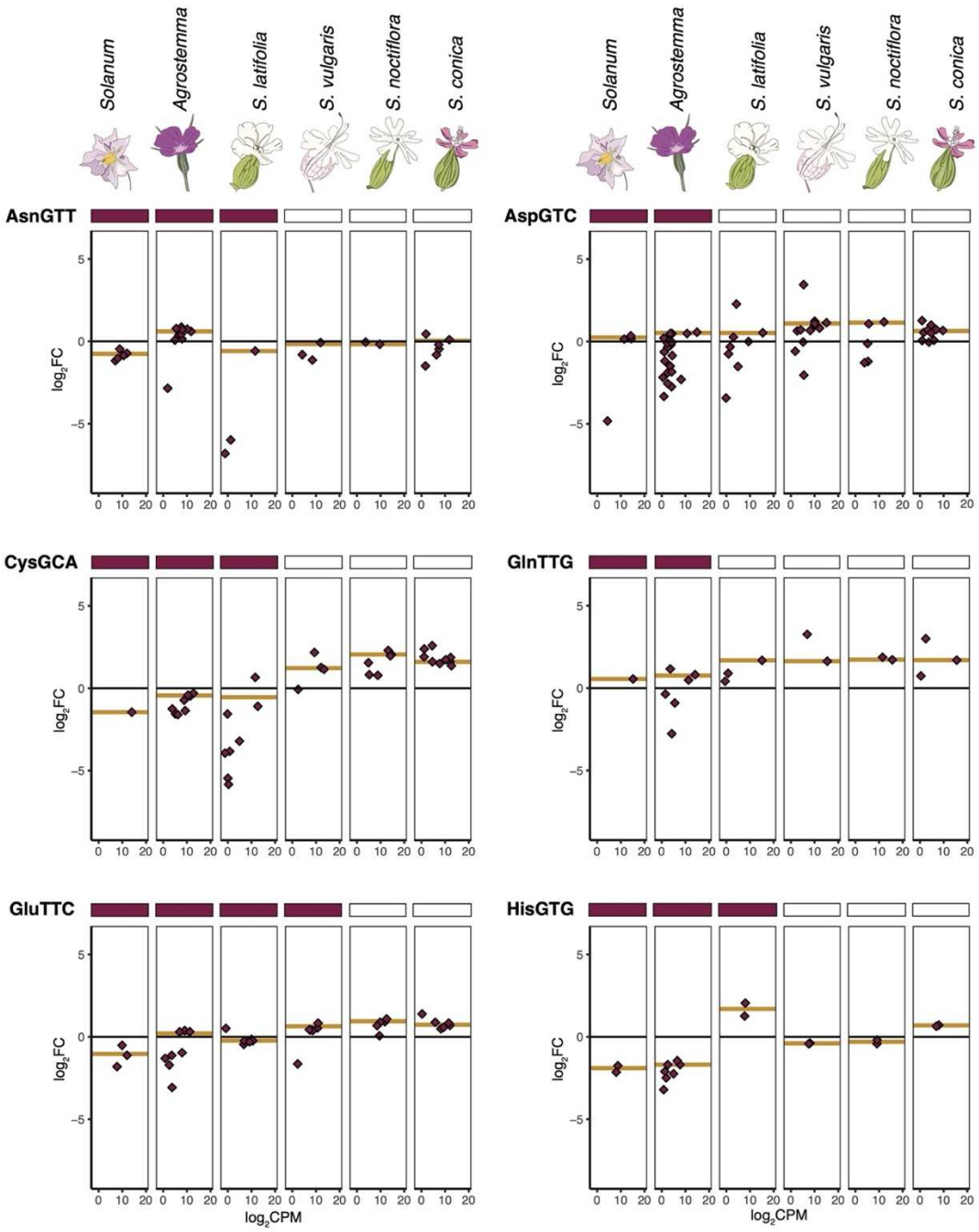


Figure 3.6| Enrichment and expression of individual cytosolic tRNA genes corresponding to mitochondrial genes with a history of loss in *Silene* since divergence from *Agrostemma* (first half, see Fig. 3.7 for second half). The y-axis is enrichment in $\log_2$ fold change in mitochondrial isolates versus total-cellular samples. The x-axis is expression level in counts per million on a $\log_2$ scale. Points represent individual reference sequences, and gold lines represent the average enrichment of all nuclear-encoded tRNAs with that anticodon weighted by expression (see Methods; enrichment data for all anticodons can be found in Supp. Table 3.10). A filled rectangle above a panel indicates that a tRNA gene with that anticodon is encoded in the mitogenome of that species.

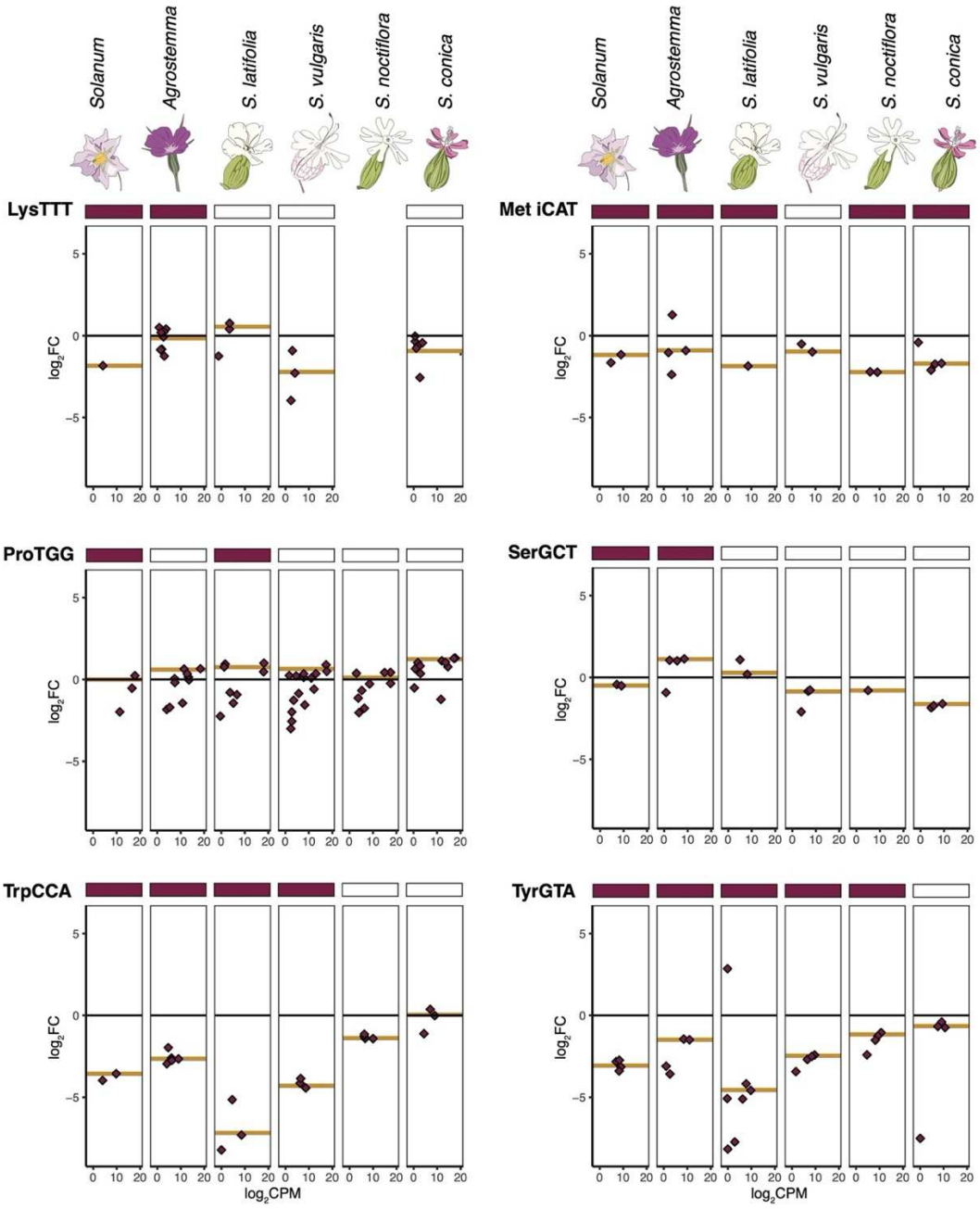


Figure 3.7| Enrichment and expression of individual cytosolic tRNA genes corresponding to mitochondrial genes with a history of loss in *Silene* since divergence from *Agrostemma* (second half, see Fig. 3.6 for first half). The y-axis is enrichment in $\log_2$ fold change in mitochondrial isolates versus total-cellular samples. The x-axis is expression level in counts per million on a $\log_2$ scale. Points represent individual reference sequences, and gold lines represent the average enrichment of all nuclear-encoded tRNAs with that anticodon weighted by expression (see Methods; enrichment data for all anticodons can be found in Supp. Table 3.10). A filled rectangle above a panel indicates that a tRNA gene with that anticodon is encoded in the mitogenome of that species. No panel is reported for *S. noctiflora* tRNA-Lys because no tRNA-Lys gene was detected as sufficient expression levels of analysis of differential expression in that species.

LITERATURE CITED

Alverson, A. J., X. X. Wei, D. W. Rice, D. B. Stern, K. Barry *et al.*, 2010 Insights into the Evolution of Mitochondrial Genome Size from Complete Sequences of Citrullus lanatus and Cucurbita pepo (Cucurbitaceae). Molecular Biology and Evolution 27**:** 1436-1448.

Anderson, S., A. T. Bankier, B. G. Barrell, M. H. de Bruijn, A. R. Coulson *et al.*, 1981 Sequence and organization of the human mitochondrial genome. Nature 290**:** 457-465.

Ayan, G. B., H. J. Park and J. Gallie, 2020 The birth of a bacterial tRNA gene by large-scale, tandem duplication events. Elife 9.

Binder, S., A. Marchfelder and A. Brennicke, 1994 RNA editing of tRNA(Phe) and tRNA(Cys) in mitochondria of Oenothera berteriana is initiated in precursor molecules. Mol Gen Genet 244**:** 67-74.

Binder, S., A. Marchfelder, A. Brennicke and B. Wissinger, 1992 RNA editing in trans-splicing intron sequences of nad2 mRNAs in Oenothera mitochondria. J Biol Chem 267**:** 7615-7623.

Boore, J. L., J. R. Macey and M. Medina, 2005 Sequencing and comparing whole mitochondrial genomes of animals. Methods Enzymol 395**:** 311-348.

Brandvain, Y., and M. J. Wade, 2009 The functional transfer of genes from the mitochondria to the nucleus: the effects of selection, mutation, population size and rate of self-fertilization. Genetics 182**:** 1129-1139.

Brown, W. M., M. George, Jr. and A. C. Wilson, 1979 Rapid evolution of animal mitochondrial DNA. Proc Natl Acad Sci U S A 76**:** 1967-1971.

Broz, A. K., G. Waneka, Z. Wu, M. Fernandes Gyorfy and D. B. Sloan, 2021 Detecting de novo mitochondrial mutations in angiosperms with highly divergent evolutionary rates. Genetics 218.

Brubacher-Kauffmann, S., L. Marechal-Drouard, A. Cosset, A. Dietrich and A. M. Duchêne, 1999 Differential import of nuclear-encoded tRNAGly isoacceptors into solanum Tuberosum mitochondria. Nucleic Acids Res 27**:** 2037-2042.

Bushnell, B., J. Rood and E. Singer, 2017 BBMerge - Accurate paired shotgun read merging via overlap. Plos One 12.

Calvo, S. E., and V. K. Mootha, 2010 The mitochondrial proteome and human disease. Annu Rev Genomics Hum Genet 11**:** 25-44.

Ceci, L. R., P. Veronico, L. Siculella and R. Gallerani, 1995 Identification and mapping of trnI, trnE and trnfM genes in the sunflower mitochondrial genome. DNA Seq 5**:** 315-318.

Chan, P. P., and T. M. Lowe, 2019 tRNAscan-SE: Searching for tRNA Genes in Genomic Sequences. Methods Mol Biol 1962**:** 1-14.

Chapdelaine, Y., and L. Bonen, 1991 The wheat mitochondrial gene for subunit I of the NADH dehydrogenase complex: a trans-splicing model for this gene-in-pieces. Cell 65**:** 465-472.

Chimnaronk, S., M. Gravers Jeppesen, T. Suzuki, J. Nyborg and K. Watanabe, 2005 Dual-mode recognition of noncanonical tRNAs(Ser) by seryl-tRNA synthetase in mammalian mitochondria. EMBO J 24**:** 3369-3379.

Clary, D. O., and D. R. Wolstenholme, 1984 The Drosophila mitochondrial genome. Oxf Surv Eukaryot Genes 1**:** 1-35.

Cozen, A. E., E. Quartley, A. D. Holmes, E. Hrabeta-Robinson, E. M. Phizicky *et al.*, 2015 ARM-seq: AlkB-facilitated RNA methylation sequencing reveals a complex landscape of modified tRNA fragments. Nature Methods 12**:** 879-+.

Darty, K., A. Denise and Y. Ponty, 2009 VARNA: Interactive drawing and editing of the RNA secondary structure. Bioinformatics 25**:** 1974-1975.

Delage, L., A. Dietrich, A. Cosset and L. Marechal-Drouard, 2003 In vitro import of a nuclearly encoded tRNA into mitochondria of Solanum tuberosum. Mol Cell Biol 23**:** 4000-4012.

Dewe, J. M., J. M. Whipple, I. Chernyakov, L. N. Jaramillo and E. M. Phizicky, 2012 The yeast rapid tRNA decay pathway competes with elongation factor 1A for substrate tRNAs and acts on tRNAs lacking one or more of several modifications. RNA 18**:** 1886-1896.

Doyle, J. J., and J. L. Doyle, 1987 A rapid DNA isolation procedure for small quantities of fresh leaf tissue. PHYTOCHEMICAL BULLETIN v.19(1):11-15.

Duchêne, A. M., and L. Marechal-Drouard, 2001 The chloroplast-derived trnW and trnM-e genes are not expressed in Arabidopsis mitochondria. Biochem Biophys Res Commun 285**:** 1213-1216.

Duchêne, A. M., C. Pujol and L. Marechal-Drouard, 2009 Import of tRNAs and aminoacyl-tRNA synthetases into mitochondria. Curr Genet 55**:** 1-18.

Forner, J., B. Weber, S. Thuss, S. Wildum and S. Binder, 2007 Mapping of mitochondrial mRNA termini in Arabidopsis thaliana: t-elements contribute to 5' and 3' end formation. Nucleic Acids Res 35**:** 3676-3692.

Giege, R., M. Sissler and C. Florentz, 1998 Universal rules and idiosyncratic features in tRNA identity. Nucleic Acids Research 26**:** 5017-5035.

Glover, K. E., D. F. Spencer and M. W. Gray, 2001 Identification and structural characterization of nucleus-encoded transfer RNAs imported into wheat mitochondria. J Biol Chem 276**:** 639-648.

Gray, M. W., 2012 Mitochondrial evolution. Cold Spring Harb Perspect Biol 4**:** a011403.

Greiner, S., P. Lehwark and R. Bock, 2019 OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. Nucleic Acids Res 47**:** W59-W64.

Grosjean, H., and E. Westhof, 2016 An integrated, structure- and energy-based view of the genetic code. Nucleic Acids Res 44**:** 8020-8040.

Guo, W., A. Zhu, W. Fan, R. P. Adams and J. P. Mower, 2020 Extensive Shifts from Cis- to Trans-splicing of Gymnosperm Mitochondrial Introns. Mol Biol Evol 37**:** 1615-1620.

Haen, K. M., W. Pett and D. V. Lavrov, 2010 Parallel Loss of Nuclear-Encoded Mitochondrial Aminoacyl-tRNA Synthetases and mtDNA-Encoded tRNAs in Cnidaria. Molecular Biology and Evolution 27**:** 2216-2219.

Hancock, K., and S. L. Hajduk, 1990 The mitochondrial tRNAs of Trypanosoma brucei are nuclear encoded. J Biol Chem 265**:** 19208-19215.

Huynen, M. A., I. Duarte and R. Szklarczyk, 2013 Loss, replacement and gain of proteins at the origin of the mitochondria. Biochimica Et Biophysica Acta-Bioenergetics 1827**:** 224-231.

Iams, K. P., J. E. Heckman and J. H. Sinclair, 1985 Sequence of histidyl tRNA, present as a chloroplast insert in mtDNA ofZea mays. Plant Mol Biol 4**:** 225-232.

Jafari, F., S. Zarre, A. Gholipour, F. Eggens, R. K. Rabeler *et al.*, 2020 A new taxonomic backbone for the infrageneric classification of the species-rich genus Silene (Caryophyllaceae). TAXON 69**:** 337-368.

Juhling, F., M. Morl, R. K. Hartmann, M. Sprinzl, P. F. Stadler *et al.*, 2009 tRNAdb 2009: compilation of tRNA sequences and tRNA genes. Nucleic Acids Res 37**:** D159-162.

Juhling, T., E. Duchardt-Ferner, S. Bonin, J. Wohnert, J. Putz *et al.*, 2018 Small but large enough: structural properties of armless mitochondrial tRNAs from the nematode Romanomermis culicivorax. Nucleic Acids Res 46**:** 9170-9180.

Kitazaki, K., T. Kubo, H. Kagami, T. Matsumoto, A. Fujita *et al.*, 2011 A horizontally transferred tRNA(Cys) gene in the sugar beet mitochondrial genome: evidence that the gene is present in diverse angiosperms and its transcript is aminoacylated. Plant J 68**:** 262-272.

Knie, N., M. Polsakiewicz and V. Knoop, 2015 Horizontal Gene Transfer of Chlamydial-Like tRNA Genes into Early Vascular Plant Mitochondria. Molecular Biology and Evolution 32**:** 629-634.

Knoop, V., W. Schuster, B. Wissinger and A. Brennicke, 1991 Trans splicing integrates an exon of 22 nucleotides into the nad5 mRNA in higher plant mitochondria. EMBO J 10**:** 3483-3493.

Krasovec, M., M. Chester, K. Ridout and D. A. Filatov, 2018 The Mutation Rate and the Age of the Sex Chromosomes in Silene latifolia. Curr Biol 28**:** 1832-1838 e1834.

Kubo, T., S. Nishizawa, A. Sugawara, N. Itchoda, A. Estiati *et al.*, 2000 The complete nucleotide sequence of the mitochondrial genome of sugar beet (Beta vulgaris L.) reveals a novel gene for tRNA(Cys)(GCA). Nucleic Acids Res 28**:** 2571-2576.

Kumar, R., L. Marechal-Drouard, K. Akama and I. Small, 1996 Striking differences in mitochondrial tRNA import between different plant species. Mol Gen Genet 252**:** 404-411.

Lavrov, D. V., and W. Pett, 2016 Animal Mitochondrial DNA as We Do Not Know It: mt-Genome Organization and Evolution in Nonbilaterian Lineages. Genome Biol Evol 8**:** 2896-2913.

Lorenz, R., S. H. Bernhart, C. Honer Zu Siederdissen, H. Tafer, C. Flamm *et al.*, 2011 ViennaRNA Package 2.0. Algorithms Mol Biol 6**:** 26.

Magallon, S., S. Gomez-Acevedo, L. L. Sanchez-Reyes and T. Hernandez-Hernandez, 2015 A metacalibrated time-tree documents the early rise of flowering plant phylogenetic diversity. New Phytol 207**:** 437-453.

Marechal-Drouard, L., P. Guillemaut, A. Cosset, M. Arbogast, F. Weber *et al.*, 1990 Transfer RNAs of potato (Solanum tuberosum) mitochondria have different genetic origins. Nucleic Acids Res 18**:** 3689-3696.

Marechal-Drouard, L., D. Ramamonjisoa, A. Cosset, J. H. Weil and A. Dietrich, 1993 Editing corrects mispairing in the acceptor stem of bean and potato mitochondrial phenylalanine transfer RNAs. Nucleic Acids Res 21**:** 4909-4914.

Martin, M., 2011 Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet. journal 17**:** 10-12.

Meiklejohn, C. D., M. A. Holmbeck, M. A. Siddiq, D. N. Abt, D. M. Rand *et al.*, 2013 An Incompatibility between a Mitochondrial tRNA and Its Nuclear-Encoded tRNA Synthetase Compromises Development and Fitness in Drosophila. Plos Genetics 9.

Mercer, T. R., S. Neph, M. E. Dinger, J. Crawford, M. A. Smith *et al.*, 2011 The human mitochondrial transcriptome. Cell 146**:** 645-658.

Michaud, M., V. Cognat, A. M. Duchêne and L. Marechal-Drouard, 2011 A global picture of tRNA genes in plant genomes. Plant J 66**:** 80-93.

Mower, J. P., P. Touzet, J. S. Gummow, L. F. Delph and J. D. Palmer, 2007 Extensive variation in synonymous substitution rates in mitochondrial genes of seed plants. BMC Evol Biol 7**:** 135.

Nurk, S., A. Bankevich, D. Antipov, A. A. Gurevich, A. Korobeynikov *et al.*, 2013 Assembling single-cell genomes and mini-metagenomes from chimeric MDA products. J Comput Biol 20**:** 714-737.

Pawar, K., M. Shigematsu, P. Loher, S. Honda, I. Rigoutsos *et al.*, 2019 Exploration of CCA-added RNAs revealed the expression of mitochondrial non-coding RNAs regulated by CCA-adding enzyme. RNA Biol 16**:** 1817-1825.

Petersen, G., A. Cuenca, I. M. Moller and O. Seberg, 2015 Massive gene loss in mistletoe (Viscum, Viscaceae) mitochondria. Sci Rep 5**:** 17588.

Peterson, K. J., J. B. Lyons, K. S. Nowak, C. M. Takacs, M. J. Wargo *et al.*, 2004 Estimating metazoan divergence times with a molecular clock. Proc Natl Acad Sci U S A 101**:** 6536-6541.

Pett, W., and D. V. Lavrov, 2015 Cytonuclear Interactions in the Evolution of Animal Mitochondrial tRNA Metabolism. Genome Biol Evol 7**:** 2089-2101.

Pons, J., P. Bover, L. Bidegaray-Batista and M. A. Arnedo, 2019 Arm-less mitochondrial tRNAs conserved for over 30 millions of years in spiders. BMC Genomics 20**:** 665.

Pujol, C., M. Bailly, D. Kern, L. Marechal-Drouard, H. Becker *et al.*, 2008 Dual-targeted tRNA-dependent amidotransferase ensures both mitochondrial and chloroplastic Gln-tRNAGln synthesis in plants. Proc Natl Acad Sci U S A 105**:** 6481-6485.

Qiu, Y. L., and J. D. Palmer, 2004 Many independent origins of trans splicing of a plant mitochondrial group II intron. J Mol Evol 59**:** 80-89.

Quinlan, A. R., and I. M. Hall, 2010 BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26**:** 841-842.

Raina, M., and M. Ibba, 2014 tRNAs as regulators of biological processes. Front Genet 5**:** 171.

Rice, D. W., A. J. Alverson, A. O. Richardson, G. J. Young, M. V. Sanchez-Puerta *et al.*, 2013 Horizontal transfer of entire genomes via mitochondrial fusion in the angiosperm Amborella. Science 342**:** 1468-1473.

Richardson, A. O., D. W. Rice, G. J. Young, A. J. Alverson and J. D. Palmer, 2013 The "fossilized" mitochondrial genome of Liriodendron tulipifera: ancestral gene content and order, ancestral editing sites, and extraordinarily low mutation rate. BMC Biol 11**:** 29.

Robinson, M. D., D. J. McCarthy and G. K. Smyth, 2010 edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26**:** 139-140.

Salinas-Giegé, T., R. Giegé and P. Giegé, 2015 tRNA biology in mitochondria. Int J Mol Sci 16**:** 4518-4559.

Sanchez-Puerta, M. V., A. Edera, C. L. Gandini, A. V. Williams, K. A. Howell *et al.*, 2019 Genome-scale transfer of mitochondrial DNA from legume hosts to the holoparasite Lophophytum mirabile (Balanophoraceae). Mol Phylogenet Evol 132**:** 243-250.

Shigematsu, M., S. Honda, P. Loher, A. G. Telonis, I. Rigoutsos *et al.*, 2017 YAMAT-seq: an efficient method for high-throughput sequencing of mature transfer RNAs. Nucleic Acids Research 45.

Simpson, A. M., Y. Suyama, H. Dewes, D. A. Campbell and L. Simpson, 1989 Kinetoplastid Mitochondria Contain Functional Transfer-Rnas Which Are Encoded in Nuclear-DNA and Also Contain Small Minicircle and Maxicircle Transcripts of Unknown Function. Nucleic Acids Research 17**:** 5427-5445.

Sissler, M., J. Putz, F. Fasiolo and C. Florentz, 2005 Mitochondrial Aminoacyl-tRNA Synthetases, pp. in *Madame Curie Bioscience Database [Internet]*. Landes Bioscience, Austin (TX).

Skippington, E., T. J. Barkman, D. W. Rice and J. D. Palmer, 2015 Miniaturized mitogenome of the parasitic plant Viscum scurruloideum is extremely divergent and dynamic and has lost all nad genes. Proc Natl Acad Sci U S A 112**:** E3515-3524.

Sloan, D. B., A. J. Alverson, J. P. Chuckalovcak, M. Wu, D. E. McCauley *et al.*, 2012a Rapid Evolution of Enormous, Multichromosomal Genomes in Flowering Plant Mitochondria with Exceptionally High Mutation Rates. Plos Biology 10.

Sloan, D. B., A. J. Alverson, H. Storchova, J. D. Palmer and D. R. Taylor, 2010 Extensive loss of translational genes in the structurally dynamic mitochondrial genome of the angiosperm Silene latifolia. BMC Evol Biol 10**:** 274.

Sloan, D. B., A. J. Alverson, M. Wu, J. D. Palmer and D. R. Taylor, 2012b Recent acceleration of plastid sequence and structural evolution coincides with extreme mitochondrial divergence in the angiosperm genus Silene. Genome Biol Evol 4**:** 294-306.

Sloan, D. B., K. Muller, D. E. McCauley, D. R. Taylor and H. Storchova, 2012c Intraspecific variation in mitochondrial genome sequence, structure, and gene content in Silene vulgaris, an angiosperm with pervasive cytoplasmic male sterility. New Phytol 196**:** 1228-1239.

Sloan, D. B., B. Oxelman, A. Rautenberg and D. R. Taylor, 2009 Phylogenetic analysis of mitochondrial substitution rate variation in the angiosperm tribe Sileneae. BMC Evol Biol 9**:** 260.

Sloan, D. B., D. A. Triant, N. J. Forrester, L. M. Bergner, M. Wu *et al.*, 2014 A recurring syndrome of accelerated plastid genome evolution in the angiosperm tribe Sileneae (Caryophyllaceae). Mol Phylogenet Evol 72**:** 82-89.

Sloan, D. B., J. M. Warren, A. M. Williams, Z. Wu, S. E. Abdel-Ghany *et al.*, 2018 Cytonuclear integration and co-evolution. Nat Rev Genet 19**:** 635-648.

Small, I., K. Akashi, A. Chapron, A. Dietrich, A. M. Duchêne *et al.*, 1999 The strange evolutionary history of plant mitochondrial tRNAs and their aminoacyl-tRNA synthetases. Journal of Heredity 90**:** 333-337.

Varre, J. S., N. D'Agostino, P. Touzet, S. Gallina, R. Tamburino *et al.*, 2019 Complete Sequence, Multichromosomal Architecture and Transcriptome Analysis of the Solanum tuberosum Mitochondrial Genome. Int J Mol Sci 20.

Vinogradova, E., T. Salinas, V. Cognat, C. Remacle and L. Marechal-Drouard, 2009 Steady-state levels of imported tRNAs in Chlamydomonas mitochondria are correlated with both cytosolic and mitochondrial codon usages. Nucleic Acids Res 37**:** 1521-1528.

Warren, J. M., T. Salinas-Giegé, G. Hummel, N. L. Coots, J. M. Svendsen *et al.*, 2021 Combining tRNA sequencing methods to characterize plant tRNA expression and post-transcriptional modification. RNA Biology 18**:** 64-78.

Warren, J. M., and D. B. Sloan, 2020a Hopeful monsters: Unintended sequencing of famously malformed mite mitochondrial tRNAs reveals widespread expression and processing of sense-antisense pairs. bioRxiv**:** 2020.2009.2008.286963.

Warren, J. M., and D. B. Sloan, 2020b Interchangeable parts: The evolutionarily dynamic tRNA population in plant mitochondria. Mitochondrion 52**:** 144-156.

Watanabe, Y., T. Suematsu and T. Ohtsuki, 2014 Losing the stem-loop structure from metazoan mitochondrial tRNAs and co-evolution of interacting factors. Front Genet 5**:** 109.

Williams, A. M., M. W. Itgen, A. K. Broz, O. G. Carter and D. B. Sloan, 2020 Long-read transcriptome and other genomic resources for the angiosperm Silene noctiflora. bioRxiv**:** 2020.2008.2009.243378.

Williams, M. A., Y. Johzuka and R. M. Mulligan, 2000 Addition of non-genomically encoded nucleotides to the 3'-terminus of maize mitochondrial mRNAs: truncated rps12 mRNAs frequently terminate with CCA. Nucleic Acids Res 28**:** 4444-4451.

Wilusz, J. E., 2015 Removing roadblocks to deep sequencing of modified RNAs. Nat Methods 12**:** 821-822.

Wissinger, B., W. Schuster and A. Brennicke, 1991 Trans splicing in Oenothera mitochondria: nad1 mRNAs are edited in exon and trans-splicing group II intron sequences. Cell 65**:** 473-482.

Wu, Z., J. M. Cuthbert, D. R. Taylor and D. B. Sloan, 2015 The massive mitochondrial genome of the angiosperm Silene noctiflora is evolving by gain or loss of entire chromosomes. Proc Natl Acad Sci U S A 112**:** 10185-10191.

Yu, C. H., J. Y. Liao, H. Zhou and L. H. Qu, 2008 The rat mitochondrial Ori L encodes a novel small RNA resembling an ancestral tRNA. Biochem Biophys Res Commun 372**:** 634-638.

Zheng, G., Y. Qin, W. C. Clark, Q. Dai, C. Yi *et al.*, 2015 Efficient and quantitative high-throughput tRNA sequencing. Nat Methods 12**:** 835-837.

CHAPTER 4: EXTENSIVE RETARGETING OF PLANT AMINOACYL TRNA SYNTHETASES ASSOCIATED WITH MITOCHONDRIAL TRNA GENE LOSS

Summary

The number of tRNAs encoded in the mitogenomes of plants varies considerably, with some lineages experiencing rapid reductions in tRNA gene content. Loss of these tRNA genes necessitates the import of nuclear-encoded counterparts that share little sequence similarity to the bacterial-like mitochondrial tRNAs, which raises multiple questions about the identity, evolution, and trafficking of the enzymes necessary for the maturation and function of these newly imported tRNAs. In particular, the aminoacyl tRNA synthetases (aaRSs) that charge tRNAs are usually found in distinct enzyme classes that specialize on either organellar or cytosolic tRNAs. Here, we investigated the evolution of subcellular localization of all aaRSs in five different species from a plant lineage (*Sileneae*) that has experienced some of the most extreme and rapid mitochondrial tRNA loss documented in angiosperms. By analyzing full-length mRNA transcripts with single-molecule sequencing technology (PacBio Iso-Seq) and searching genome assemblies, we found instances of predicted retargeting of an ancestrally cytosolic synthetase to the mitochondrion as well as scenarios where enzyme localization does not appear to change despite functional tRNA replacement. In most cases of transit peptide acquisition, the cytosolic enzyme appears to have gained a coding extension prior to the loss of the cognate mitochondrial tRNAs – suggesting a transitional state where organellar and cytosolic enzymes colocalize to mitochondria. In addition, multiple organellar aaRSs were predicted to lose mitochondrial localization (while retaining plastid localization) after the functional replacement of a mitochondrial tRNA and the gain of a transit peptide on a cytosolic synthetase. The ongoing functional replacement of mitochondrial tRNAs in *Sileneae* and subsequent rapid evolution of tRNA-interacting enzymes presents a unique opportunity to study the coevolutionary dynamics of plant cytonuclear genetics.

Introduction

Mitochondrial protein synthesis requires the expression and interaction of gene products encoded in two separate genomes. Despite the vast majority of genes involved in translation being encoded in the nuclear genome, all mitochondrial genomes (mitogenomes) encode some translational machinery (ROGER *et al*. 2017). For example, mitogenomes typically encode at least some transfer RNA (tRNA) genes but lack all the enzymes necessary for tRNA maturation, modification, and aminoacylation. This partitioning of translational proteins and RNAs across distinct genomes has resulted in chimeric enzymatic networks, with gene products being produced in separate subcellular locations by disparate genetic systems (mitochondrial and nuclear) (SLOAN *et al*. 2018). Vascular plants have even more complex translational systems because of the presence of a second endosymbiotic partner, the plastid. Plastids are additional sites for protein synthesis and exploit some of the same machinery utilized in mitochondrial translation (DUCHÊNE *et al*. 2005b; PUJOL *et al*. 2008). Therefore, the molecular evolution of the plant mitochondrial translation apparatus represents a unique opportunity to study mitonuclear and organellar coevolution.

One of the most fundamental enzymatic processes required by all translation systems is the interaction between tRNAs and a cognate aminoacyl-tRNA synthetase (aaRS or RS). Before carrying amino acids to the ribosome for polypeptide synthesis, the correct amino acid must first be ligated to the 3′-end of the tRNA by an aaRS. There are a total of a total of 23 described aaRSs, one for each of the twenty proteinogenic amino acids (with the exception of lysine, which has two), and two additional aaRSs for the rare amino acids pyrolysine and phosphoserine (RUBIO GOMEZ AND IBBA 2020). In vascular plants, all aaRS genes are encoded by the nuclear genome (DUCHÊNE *et al*. 2009). Therefore, aaRSs that function in organellar protein synthesis must be translated in the cytosol on cytosolic ribosomes, targeted to the correct organelle, and translocated across multiple membranes for functionality (DUCHÊNE *et al*. 2009; GHIFARI *et al*. 2018). These subcellularly trafficked aaRSs are generally bacterial-like and distinct from their cytosolic counterparts—largely the result of intracellular gene transfers (plastid and mitochondrial transfers to the nuclear genome) or horizontal gene transfers from other bacterial sources

(DOOLITTLE AND HANDY 1998; DUCHÊNE *et al*. 2005b; BRANDAO AND SILVA-FILHO 2011; RUBIO GOMEZ AND IBBA 2020).

Somewhat surprisingly, an analysis of aaRS genes in *Arabidopsis thaliana* did not find the expected 20 aaRS genes for each subcellular compartment (cytosol, mitochondria, and plastid) (SMALL *et al*. 1999; DUCHÊNE *et al*. 2005b). Instead, most aaRSs that are imported into the mitochondria are also trafficked to plastids in a dual-targeted fashion, resulting in only 45 aaRS genes in *A. thaliana* (DUCHÊNE *et al*. 2005b). These dual-targeted aaRSs must then interact with both mitochondrial tRNAs (mt-tRNAs) and plastid tRNAs to enable translation in both bacterial-like systems.

This trafficking of nuclear-encoded aaRSs to plant organelles is accomplished through an import pathway involving a series of amino acids located on the N-terminus of the protein that are recognized by translocase proteins on outer organelle membranes (BERGLUND *et al*. 2009; GE *et al*. 2014; GHIFARI *et al*. 2018). These N-terminal presequences, or transit peptides, are cleaved after translocation and can vary considerably in length from less than 20 amino acids to over 100 (averaging around 42-50 residues) (HUANG *et al*. 2009; GE *et al*. 2014; MURCHA *et al*. 2014). Similar to mitochondrial transit peptides in other domains of life, peptides recognized by plant mitochondria often have the capacity to form amphipathic alpha helices with many hydrophobic and positively charged amino acids (HUANG *et al*. 2009; SCHMIDT *et al*. 2010). Plant mitochondrial transit peptides are also particularly rich in Ser residues, and many have a loosely conserved motif near the peptide cleavage site containing an Arg residue (HUANG *et al*. 2009; GE *et al*. 2014). Despite these general structural features, studies have shown that there is almost no primary amino acid sequence conservation in transit peptides (LEE *et al*. 2008; KUNZE AND BERGER 2015), and these domains are often considered some of the fastest evolving (non-neutral) sites (WILLIAMS *et al*. 2000a; CHRISTIAN *et al*. 2020).

Dual-targeted aaRSs that function in both mitochondria and plastids contain an ambiguous N-terminal transit peptide that is recognized by each of the organelle outer membranes (PEETERS AND SMALL 2001; DUCHÊNE *et al*. 2005b). A considerable amount of work has gone into characterizing dual-targeted (ambiguous) versus organelle-specific peptides in plant biology, but how the nonhomologous

protein translocation machinery on plastid and mitochondrial outer membranes can interact with the same transit peptide is still unclear (YOGEV AND PINES 2011; CARRIE AND SMALL 2013). While plastid-specific transit peptide sequences generally lack the helical structure found on mitochondrial transit peptides, both organelle transit peptides have very similar amino acid compositions with many hydrophobic and positively charged residues (BRUCE 2001; GE *et al*. 2014; CHRISTIAN *et al*. 2020). Not surprisingly, ambiguous transit peptides often exhibit intermediate properties between plastid and mitochondrial-specific transit peptides (PUJOL *et al*. 2007; BERGLUND *et al*. 2009).

Despite the predominance of dual-targeted, bacterial-like aaRSs imported into plant organelles, alternative scenarios also exist. In *A. thaliana*, there are five eukaryotic-like (cytosolic) aaRSs that are dual-localized to mitochondria and the cytosol (MIREAU *et al*. 1996; DUCHÊNE *et al*. 2005b). The import of these eukaryotic-like aaRSs contributes to the complex nature of mitochondrial tRNA metabolism in plants, where the import of some nuclear-encoded tRNAs is also necessary due to an insufficient number of tRNAs being encoded in the mitochondrial genome (MICHAUD *et al*. 2011a). These five aaRS enzymes shared between the cytosol and mitochondrial compartments in *A. thaliana* correspond to cognate tRNAs that are also imported from the cytosol – thereby maintaining phylogenetic congruence between the imported tRNA and interacting enzyme (DUCHÊNE *et al*. 2005b). This coevolutionary pairing of tRNA and aaRS is likely necessary due to the highly discriminating nature of aaRSs (RUBIO GOMEZ AND IBBA 2020). Because the correct attachment of the corresponding amino acid to RNA is essential for the faithful decoding of the genome, tRNA aminoacylation is a highly accurate process whereby aaRS enzymes use certain nucleotide positions (identity elements) on the tRNA for substrate recognition (GIEGE *et al*. 1998). As eukaryotic tRNAs share little sequence homology to the bacterial-like mitochondrial and plastid tRNAs, they would be expected to make poor substrates for bacterial-like aaRSs (SALINAS-GIEGÉ *et al*. 2015). However, there are a few cases of an aaRS and tRNA not sharing a kingdom-level relationship (DUCHÊNE *et al*. 2005b; WARREN AND SLOAN 2020b). For example, a eukaryotic ProRS appears to have replaced all bacterial counterparts in the genome of *A. thaliana*, despite the species retaining a bacterial-like mitochondrial tRNA-Pro. Therefore, mitochondrial tRNA-Pro must

then be aminoacylated by a eukaryotic enzyme. Notably, two eukaryotic ProRSs exist in the *A. thaliana* genome, and only one of those genes contains an organelle transit peptide – suggesting that some enzymatic differentiation may be necessary for bacterial-like tRNA recognition (DUCHÊNE *et al*. 2005b).

Despite a few aaRS/tRNA phylogenic incongruencies, there exists a general rule of tRNAs encoded in mitochondrial genomes being charged by enzymes that are bacterial in nature. Questions then arise as to the trafficking of aaRSs in plants that have undergone recent and extensive mt-tRNA loss. For example, the sequencing of multiple mitogenomes from the angiosperm genus *Silene* revealed some of the most extensive losses of mt-tRNA genes reported in flowering plants (Fig. 4.1) (SLOAN *et al*. 2012a).

Recent analysis suggests that these mt-tRNAs are being functionally replaced by nuclear-encoded counterparts which have gained mitochondrial import (Chapter 3). The almost complete replacement of native mt-tRNAs raises multiple alternative scenarios as to the identity of the aaRSs that aminoacylate the newly imported tRNAs. It is possible that the bacterial-like enzymes retain mitochondrial localization and have adapted to recognize a novel substrate (a eukaryotic tRNA) or that these organelle-targeted aaRSs are just less discriminating and can effectively charge substantially divergent tRNA molecules. Alternatively, cytosolic aaRSs may also gain targeting to the mitochondria, possibly through the *de novo* acquisition of a mitochondrial transit peptide, and act on the newly imported tRNA – effectively replacing both partners in the tRNA/aaRS system with eukaryotic counterparts (HAEN *et al*. 2010a; PETT AND LAVROV 2015).

Here, using data from a combination of single-molecule RNA-seq methods that capture full-length transcripts, conventional Illumina mRNA sequencing, and genome assemblies, we test for the evolution of mitochondrial transit peptides, gene duplication events, and changes in transit peptide targeting in four different *Silene* species as well as *Agrostemma githago*, a close outgroup in the tribe *Sileneae*. We find that functional replacement of mt-tRNAs in these species precipitates rapid aaRS evolution and predicted changes in organelle trafficking but that the evolutionary responses differ depending on the species and the identity of the aaRS.

Results

Identification and characterization of Sileneae aaRS gene content

To test for the functional gain of mitochondrial import of cytosolic aaRSs in response to the import of nuclear-encoded tRNAs, we used PacBio Iso-Seq technology (ZHAO *et al*. 2019) to sequence full-length mRNA transcripts from five *Sileneae* species: *A. githago, S. conica*, *S. latifolia, S. noctiflora*, and *S. vulgaris*. In addition, previously generated genome assemblies from the same species (Chapter 3;(KRASOVEC *et al*. 2018; WILLIAMS *et al*. 2020)were searched for genes or putative transit peptides that may not have been detected by Iso-Seq because of lower expression levels.

Transcripts identified as homologous to annotated *A. thaliana* aaRSs were aligned, and maximum likelihood trees were generated to infer the history of each gene family. To investigate the possible acquisition of mitochondrial transit peptides, the subcellular targeting of each aaRS was predicted using multiple software packages: TargetP v.2.0 (ALMAGRO ARMENTEROS *et al*. 2019), LOCALIZER v.1.0.4 (SPERSCHNEIDER *et al*. 2017), and Predotar v.1.04 (SMALL *et al*. 2004).

This analysis successfully identified transcripts from five *Sileneae* species corresponding to known *A. thaliana* organellar and cytosolic aaRSs for each amino acid (Table 4.1; Figs. 4.2-4.19). As expected, *Sileneae* aaRSs that were homologous to organelle-targeted aaRSs in *A. thaliana* had very high predicted probabilities of being localized to mitochondria, plastids, or both (Figs. 4.4-4.6, 4.8-4.11, 4.13, 4.15-4.19). Gene trees and targeting predictions for aaRS families corresponding to 14 mt-tRNA genes that have been lost and replaced within *Sileneae* are presented in Figs. 4.3-4.19. The aaRS families with cognate mt-tRNAs that are absent in all angiosperm mitogenomes (i.e., were lost in the ancestor of angiosperms) have targeting predictions shown in Fig. 4.2 but are not presented as individual gene trees.

Frequent and rapid acquisition of N-terminal extensions in cytosolic aaRSs genes associated with mt-tRNA loss in Sileneae

In *Sileneae*, mt-tRNA genes decoding 13 amino acids have been lost in one or more species compared to *A. thaliana* (Fig. 4.1), raising the question as to which aaRSs are charging the newly imported tRNA replacements. In six cases (GlnRS, LysRS, MetRS, ProRS, TrpRS, and TyrRS), we identified an N-terminal extension of the cytosolic enzyme in multiple *Sileneae* species that was absent in the corresponding *A. thaliana* enzyme (Figs. 4.2, 4.7, 4.11-4.12, 4.16, 4.18-4.19). Many aaRS genes existed as multicopy gene families, and there were multiple cases where an N-terminal extension was only present on a specific clade of genes within an aaRS family. GlnRS (Fig. 4.7), MetRS (Fig. 4.12), ProRS (Fig. 4.16), and TyrRS (Fig. 4.19) all had multiple well-supported clades for each enzyme, but only one group of enzymes had acquired a 5´-coding extension, suggesting that mitochondrial localization happened following a gene duplication event. The age of these duplications varied considerably, as the two groups of cytosolic MetRS enzymes predate the divergence of *A. thaliana* and *Sileneae* (Fig. 4.12), whereas the duplication of GlnRS, ProRS, and TyrRS were specific to the lineage leading to *Sileneae*. There were also multiple LysRS genes in *A. githago*, *S. latifolia* and *S. vulgaris* with and without coding extensions; however, these genes did not form clearly distinct clades (Fig. 4.11). TrpRS was the only cytosolic aaRS enzyme predicted to gain a mitochondrial transit peptide that was clearly present as a single copy in the *Silene* species (Fig. 4.18).

For aaRS genes predicted to be mitochondrially-localized, alternative transcription start sites commonly resulted in the expression of two isoforms – producing transcripts with and without coding extensions. For MetRS, GlnRS, LysRS, and TrpRS, an isoform that lacked an N-terminal extension (but was otherwise identical or nearly identical to the extension-containing transcripts) generally exhibited much higher expression levels (inferred Iso-Seq read counts) than transcripts with coding extensions. One exception was the expression of LysRS in *S. noctiflora* where transcripts with extensions outnumbered those without. The alternative transcription of an aaRS gene with dual localization to the cytosol (shorter protein without transit peptide) and mitochondria (longer protein) has been previously described in *A. thaliana* (MIREAU *et al*. 1996). In contrast to the above cases, only transcripts containing an extension

were detected for the TyrRS gene in *S. conica* that was predicted to have gained mitochondrial localization.

The identified N-terminal extensions were generally found on orthologs from multiple species within the *Sileneae* clade (Fig. 4.2), suggesting a single origin preceding the divergence of these species. A handful of exceptions to this pattern did exist where the presence of an N-terminal extension appeared to have occurred uniquely within a species. For example, a duplicate cytosolic AspRS gene in *S. vulgaris* was discovered through a genome assembly search and was strongly predicted to be mitochondrially targeted (Fig. 4.5). Additionally, N-terminal extensions on a cytosolic ProRS in *A. githago* (Fig. 4.16) and a cytosolic TyrRS in *S. vulgaris* (Fig. 4.19) appear to be nonhomologous to the extensions found in other species because no significant similarity was found with a blastn comparison at an e-value threshold of 0.1. Therefore, these cases are likely independently derived.

In the majority of cases where a cytosolic enzyme was predicted to have gained a mitochondrial targeting, an N-terminal extension was obvious in alignments and was contiguous with the downstream coding sequence in the respective genome assembly (i.e., not seperated by an intron). However, alternative splicing for GluRS in *A. githago* (low expression; two Iso-Seq reads) and *S. conica* (moderate expression; 50 Iso-Seq reads) resulted in isoforms that were predicted to be mitochondrially-targeted. These alternative isoforms were not detected in other species (Fig. 4.2) and were shorter than the more typical isoforms predicted for these GluRS genes. In each case, a 5′ portion of the open reading frame (ORF) was replaced by a shorter, novel exon (the alternative *S. conica* and *A. githago* isoforms were predicted to be 65 and 21 amino acids shorter, respectively). Therefore, despite the predicted gain of mitochondrial targeting, the truncated length of the core ORFs raises questions about whether these alternatively spliced transcripts produce aaRSs that are functional in aminoacylation.

Functional replacement of mt-tRNAs is not always associated with retargeting of cytosolic aaRSs

Although evolution of N-terminal transit peptides appears to have occurred repeatedly in *Sileneae* aaRSs, there were also numerous examples where a cytosolic enzyme did not experience an obvious change in N-

terminal sequence despite a change in mitochondrial gene content and tRNA import. For AsnRS, CysRS, HisRS, PheRS, and SerRS, organelle localization was not predicted by any of the software programs (Figs. 4.3, 4.6, 4.10, 4.14, 4.17), and the length of the enzymes did not differ substantially from the corresponding *A. thaliana* ortholog(s) in alignments. The organellar aaRSs for Asn (Fig. 4.4), Cys (Fig. 4.6), His (Fig. 4.10), and Phe (Fig. 4.15) did retain putative transit peptides for mitochondrial localization, suggesting that these organellar aaRSs might now be charging the newly imported cytosolic tRNAs. SerRS presented a unique situation because neither the cytosolic nor organellar genes in *Sileneae* are predicted to be mitochondrially targeted (Fig. 4.17). Instead, the organellar SerRS is predicted to be plastid-specific (Fig. 4.17), making the identity of the aaRS that charges mitochondrially imported tRNA-Ser unclear.

Although *A. thaliana* lacks mt-tRNA-Phe, the loss of this mt-tRNA gene in *Sileneae* appears to be an independent event, as multiple relatives (including *S. chalcedonica*) still retain a mt-tRNA-Phe in the mitogenome (Fig. 4.1). In *A. thaliana,* only a single organellar PheRS was found to localize to both plastids and mitochondria and was therefore expected to charge both native plastid tRNAs as well as the imported tRNA-Phe in mitochondria. However, the enzymatic coevolutionary response to losing this mt-tRNA may be sustainably different in *Sileneae* as there are two phylogenetically distinct PheRS gene families in the clade (Fig. 4.15). One clade of PheRS genes is strongly predicted to be mitochondrially targeted and only moderately plastid-targeted, whereas the inverse is true for the other clade (Fig. 4.15). This history of gene duplication and N-terminal sequence evolution may point to specialization of these organellar-targeted genes and the adaptation of one clade of orthologs for cytosolic tRNA recognition.

Some substantially truncated transcripts originating from HisRS and AspRS resulted in ORFs that began in the middle or the C-terminal end of the gene body and had a high probability of being mitochondrially targeted. These lowly expressed transcripts could represent transcript degradation, and we did not consider them to be strong evidence of aaRS function in the mitochondria due to the lack of an intact full-length protein typically required for aminoacylation. Some aaRS protein domains are known to have functions outside of aminoacylation (IBBA AND FRANCKLYN 2004; PANG *et al*. 2014b), so we cannot

rule out the possibility that a truncated aaRS gene product is organelle localized and has novel functions, but we did not include these truncated sequences in this aaRS analysis.

The evolution of putative aaRS transit peptides can occur prior to the loss of mt-tRNAs

Phylogenetically, the acquisition of putative transit peptides by cytosolic aaRSs in *Sileneae* rarely matched the precise timing of the loss of the cognate mt-tRNA (Fig. 4.2). A perfect correlation between mt-tRNA loss and predicted cytosolic aaRS retargeting only occurred in GlnRS and potentially in TrpRS. In the latter, some extensions were detected on a cytosolic TrpRS in species that still retain a mt-tRNA-Trp gene, but they failed to make the targeting prediction cutoff for strong likelihood of mitochondrial localization. For the cytosolic enzymes GluRS, LysRS, TyrRS, and ProRS, an N-terminal extension was present in species that still retained a mt-tRNA (Fig. 4.2). For example, a predicted targeting peptide was present on the cytosolic LysRS transcripts from *A. githago* despite *A. githago* retaining both a mt-tRNA-Lys gene and an organellar LysRS with predicted mitochondrial targeting. Similarly, mitochondrial localization was also predicted for cytosolic TrpRS and TyrRS in *S. vulgaris*, yet both cognate bacterial-like tRNAs are still encoded in the *S. vulgaris* mitogenome. If these targeting predictions are correct, then the colocalization to the mitochondria of two aaRSs (both an organellar and cytosolic type) would be a widespread phenomenon in *Sileneae* (Fig. 4.2), raising numerous questions about the enzymatic function and substrates of both aaRSs.

Retargeting of cytosolic aaRSs to mitochondria may precipitate loss of organellar aaRS dual-targeting

Predicting organelle-specific versus dual-targeted enzymes with *in silico* methods is difficult due to the shared characteristics of mitochondrial, plastid, and dual transit peptides. Nevertheless, the gain of mitochondrial localization of a cytosolic aaRS enzyme leads to multiple hypothesis about the fate of the organelle-targeted counterpart. The wholesale replacement of both tRNA and aaRS in mitochondrial translation is unlikely to lead to the loss of the organellar aaRS because plant cells contain an additional

endosymbiotic partner (the plastid) that retains a full complement of tRNA genes that require an organellar aaRS for aminoacylation. This dual functionality of many plant aaRSs likely constrains outright aaRS gene loss because of the pleiotropic effects on plastid translation. However, one testable prediction is that ancestrally dual-targeted aaRSs that localize to both plastids and mitochondria lose mitochondrial targeting when a cytosolic counterpart gains mitochondrial localization. In agreement with this, we observed a decreased probability of mitochondrial targeting for some organellar aaRS genes when a cytosolic enzyme was predicted to be retargeted to the mitochondria.

The targeting of GlyRS enzymes presents an interesting situation in *A. thaliana* where both a cytosolic enzyme and a dual-targeted organellar enzyme are localized to the mitochondria (Figs. 4.2 and 4.9). In *Sileneae*, the presence of a putative transit peptide on the cytosolic GlyRS is conserved, but unlike in *A. thaliana*, the *Sileneae* species have lost the mt-tRNA-Gly gene, suggesting a replacement of the bacterial Gly decoding system in *Sileneae* (Fig. 4.2). This functional replacement of tRNA/aaRS appears to correspond to a decrease in the predicted probably of mitochondrial localization and is replaced by a largely plastid-specific signal (Fig. 4.9).

Retargeting of cytosolic MetRS is also associated with changes in dual-targeting predictions. Although multiple organellar MetRS genes experienced only a marginal decrease in mitochondrial targeting prediction compared to *A. thaliana*, *S. vulgaris* had virtually no signal of mitochondrial localization (Fig. 4.13) and is the only species in the lineage that has lost both tRNA-Met genes (elongator Met and initiator fMet, Fig. 4.1). This observation raises the possibly that the loss of both genes has obviated the need for an organellar MetRS in *S. vulgaris* mitochondria.

Finally, there was a correspondence between the strength of the prediction for mitochondrial targeting of a cytosolic TrpRS and the reduction of a dual-targeting signal in the organellar TrpRS counterpart in *Silene* (Fig. 4.18). This association was strongest in the species that have lost the cognate mt-tRNA gene, as *S. noctiflora* and *S. conica* had the greatest reduction in mitochondrial targeting probability for the organellar enzymes (Fig. 4.2, Fig. 4.18). *Silene vulgaris* shows some indication of mitochondrial targeting of the cytosolic TrpRS but still retains the mt-tRNA-Trp gene. Interestingly, this

species has two different copies of the organellar TrpRS, one of which is predicted to be dual targeted, while the other is predicted to be exclusively plastid localized (Fig. 4.18).

Although there were multiple instances where bacterial-like aaRSs appeared to lose mitochondrial localization after the gain of a putative transit peptide on a cytosolic enzyme, this trend was not universal. For ProRS, LysRS, and TryRS, the gain of N-terminal extensions for cytosolic aaRSs was not associated with substantial changes in the predicted localization of the existing organellar aaRSs

Discussion

The rapid functional replacement of *Sileneae* mt-tRNA genes with nuclear-encoded counterparts raises questions about the identity of the aaRSs that aminoacylate the newly imported tRNAs. Here, we used predictive subcellular localization software in conjunction with full-length cDNA sequencing and nuclear genome assembly data to test for hypothesized changes in aaRS targeting and import in association with tRNA replacement.

If the aaRS targeting predictions are accurate (see below for discussion of some of the limitations of *in silico* predictions), in roughly half of the cases where an mt-tRNA gene is lost, a cognate cytosolic aaRS has also gained mitochondrial import – thereby maintaining the functional interaction between a eukaryotic-like tRNA/aaRS pair. The acquisition of mitochondrial transit peptides appears to evolve quickly, as there were multiple instances where the evolution of a transit peptide happened in a single species, and even instances where a class of aaRS genes gained transit peptides independently multiple times (e.g., ProRS [Fig. 4.16], and TyrRS [Fig. 4.19]). The likelihood and rate of proteins gaining organelle import have been experimentally explored by testing the localization of bacterial-derived sequences (BAKER AND SCHATZ 1987) as well as with *in silico* methods tracing the gain and loss of putative transit sequences in gene histories and the predictive localization of randomized sequences (CHRISTENSEN *et al*. 2005; XU *et al*. 2013; CHRISTIAN *et al*. 2020; COSTELLO *et al*. 2020). Both experimental analyses with foreign sequences and genomic efforts have suggested that the process of

protein retargeting is relatively common. This high probability of *de novo* evolution of organellar transit sequences may be related to the typical amino acid composition of these presequences, as the most frequently occurring residues in transit peptides are represented by the most abundant codons in the genetic code (Ser, Arg, Leu with 6-codon families, and Ala with a 4-codon family), whereas the least abundant amino acids are Asp and Glu with only 2-codon families (MACKENZIE 2005). Thus, the dynamic evolution of protein targeting may be the result, at least in part, of mutational biases towards amino acid compositions that are recognized as mitochondrial transit peptides.

The frequent appearance of N-terminal extensions on cytosolic aaRSs before the loss of the cognate mt-tRNA gene presents a possible intermediate state of mitochondrial translation whereby both an organellar system (a bacterial-like tRNA/aaRS) and a cytosolic system (eukaryotic-like tRNA/aaRS) are cofunctional. For LysRS, TyrRS, ProRS and possibly TrpRS and GluRS, a *Sileneae* species is predicted to have both a cytosolic and organellar aaRS localized to the mitochondria for the same amino acid. A similar situation exists in *A. thaliana* where both imported tRNA-Gly and mt-tRNA-Gly are necessary for translation. The dual existence of both imported and native tRNAs that decode the same amino acid (but different codons) is mirrored by the import of both an organellar and cytosolic GlyRS (Fig. 4.2) (DUCHÊNE *et al*. 2005b). The bacterial-like GlyRS was found to effectively aminoacylate both tRNA counterparts whereas the cytosolic GlyRS had poor actively with a mt-tRNA-Gly substrate (DUCHÊNE *et al*. 2001). The cytosolic GlyRS was suggested to be inactive in the mitochondrial matrix in *A. thaliana* because the enzymatic extracts from mitochondria were hydrophobic and appeared to be inactive when fractionated by chromatography (DUCHÊNE *et al*. 2001). Nevertheless, it would be interesting to determine whether similar scenarios exist in *Sileneae* where an organellar or cytosolic aaRS has cross-functionality and is able to effectively charge both tRNA counterparts. Alternatively, there may exist a functional partitioning of matched aaRSs and tRNAs based on phylogenetic origins. Once both aaRSs are functional within the mitochondria, it becomes easy to imagine a scenario where an inactivating mutation in the mt-tRNA gene makes the system wholly dependent on the cytosolic tRNA/aaRS pair.

For the remaining roughly half of the cases of mt-tRNA gene loss and functional replacement in *Sileneae*, no N-terminal extension could be found on a corresponding cytosolic aaRS. The retention of putative mitochondrial transit peptides on at least some of the organellar aaRSs points to imported cytosolic tRNAs being charged by a bacterial-like aaRS. This type of mismatch between a bacterial-like aaRS and a eukaryotic tRNA is not unprecedented. The evolutionary history of aaRSs across the tree of life is complex, and likely involves substantial horizontal gene transfer, making aaRSs somewhat modular genes within interaction networks on long timescales (DOOLITTLE AND HANDY 1998; BRINDEFALK *et al*. 2007).

In *A. thaliana*, the cytosolic TrpRS, PheRS, IleRS, and the aforementioned GlyRS enzymes all lack obvious transit peptides, yet nuclear-encoded tRNAs for those amino acids are imported from the cytosol, suggesting that the eukaryotic tRNAs are effective substrates for some organellar aaRSs (DUCHÊNE AND MARECHAL-DROUARD 2001; DUCHÊNE *et al*. 2005b; DUCHÊNE *et al*. 2009). In agreement with this model, aminoacylation studies on aaRSs in plants and other taxa have found organellar aaRSs to be less discriminating than bacterial or eukaryotic counterparts (SALINAS-GIEGÉ *et al*. 2015).

Despite what appears to be a relaxation of tRNA specificity in some organellar aaRSs, there are multiple examples of even single nucleotide substitutions in tRNAs resulting in severe reductions in aminoacylation (YARHAM *et al*. 2010; LIU AND CHEN 2020). In one described case of aaRS/tRNA incompatibility in *Drosophila*, a single amino acid polymorphism in the mitochondrial TyrRS negatively interacted with a base-pair polymorphism in mt-tRNA-Tyr to produce a diseased phenotype of delayed development and reduced fecundity (MEIKLEJOHN *et al*. 2013).

Given that mitochondrial-encoded and nuclear-encoded tRNA counterparts often lack detectable similarity in primary nucleotide sequence, the enzymatic interaction of an organellar aaRS and a nuclear-encoded tRNA represents a radical change in substrate recognition and may necessitate compensatory changes in the bacterial-like enzyme. In the case of the organellar PheRS enzymes in *Sileneae*, our data suggest that there was a duplication, resulting in mitochondrial- and plastid-specific forms of this enzyme

(Fig. 4.15). Such duplications may enable enzymatic specificity where mutations in the enzyme responsible for charging the imported cytosolic tRNAs in mitochondria are incompatible with plastid tRNA recognition. More broadly, the retargeting of certain cytosolic aaRSs may offer clues as to which organellar aaRS have maintained substrate specificity thereby requiring the import of a phylogenetically congruent (cytosolic) aaRS for functional mt-tRNA gene replacement.

Although the localization of aaRSs to plant mitochondria has been demonstrated to be largely achieved through the presence and interaction of N-terminal presequences, there are two cases in *A. thaliana* (LeuRS and ArgRS) where a cytosolic enzyme is localized to mitochondria but lacks an identifiable N-terminal transit peptide (DUCHÊNE *et al*. 2005b). For cytosolic LeuRS, western blots and *in vivo* uptake studies have confirmed its presence in *A. thaliana* mitochondria, and recent work has shown that the import is dependent on the cognate tRNA-Leu (REINBOTHE *et al*. 2021). The cytosolic ArgRS in *A. thaliana* also lacks a transit peptide, but the ArgRS enzyme with an N-terminal extension was shown to exclusively localize to plastids (DUCHÊNE *et al*. 2005b) – suggesting that the cytosolic ArgRS is targeted to the mitochondria by an unknown mechanism. These examples highlight the need for experimental investigation into the localization of aaRS, particularly in systems like *Sileneae* undergoing rapid evolution in mt-tRNA gene content. The curious example of SerRS localization in *Sileneae*, where neither the cytosolic nor organellar enzymes were highly predicted to be mitochondrially localized, may point to either a failure to recognize a mitochondrial transit peptide or an import mechanism that is not dependent on an N-terminal sequence. Interestingly, cytosolic SerRS enzymes have undergone a duplication, something seen in other cytosolic aaRS classes that have gained mitochondrial localization.

Depending on the substrate specificity of organellar aaRSs, the loss of mt-tRNAs may result in a loss of function for the cognate aaRS. An enzyme that is no longer functional in the mitochondrial matrix would then be expected to eventually be lost due to a lack of selective constraint. This trend of organellar aaRS gene loss correlating with mt-tRNA gene loss has been described in multiple non-bilaterian animal systems, including cnidarians, comb jellies, and sponges (HAEN *et al*. 2010a; PETT AND LAVROV 2015). Instead of outright gene loss, we did see changes in the predicted targeting of multiple aaRSs (GlyRS,

MetRS, TrpRS) to the mitochondria following mt-tRNA loss, resulting in aaRS targeting predictions that were largely plastid-specific. The loss of dual targeting in plants has been described as infrequent, but it does occur and is associated with gene duplication and neofunctionalization, resulting in two organelle-specific proteins (XU *et al*. 2013). If *Sileneae* is experiencing large scale loss of mitochondrial targeting in response to cytosolic aaRS retargeting, this would represent a complex chain of coevolutionary events involving three genomic compartments. The rapid acquisition of N-terminal extensions on cytosolic aaRSs and shifts in predicted localization of organellar aaRSs suggest that aaRS retargeting can happen rapidly in systems with ongoing mt-tRNA gene loss and represent a key arena for cytonuclear coevolutionary studies.

Methods

Tissue generation and growth conditions

Tissue generation, RNA extraction, and library construction for *S. noctiflora* was done in a previously described study (WILLIAMS *et al*. 2020), while data for the other four *Sileneae* species were newly generated for this study. The following seed collections or accessions were used: *A. githago* Kew Gardens Millennium Seed Bank (0053084), *S. vulgaris* S9L (SLOAN *et al*. 2012c), *S. latifolia* UK2600 (from the line originally used for mitochondrial genome sequencing in SLOAN *et al*. (2010b)), and *S. conica* ABR (SLOAN *et al*. 2012a). Seeds were germinated in small plastic pots with Plantorium Greenhouse brand potting soil in a growth chamber at 23 °C with a light setting of 8-hour light/16-hour dark at 100 μE m^{-1} s^{-1}. One month after germination, chamber settings were modified to promote flowering ("long-day" conditions) with 16-hour light/8-hour dark, 100 μE $m^{-1}s^{-1}$.

RNA extraction and Iso-Seq library construction

RNA was extracted from *A. githago* (hermaphrodite), *S. conica* (hermaphrodite), *S. latifolia* (male), and *S. vulgaris* (male-fertile hermaphrodite) with a Qiagen RNeasy Plant Mini Kit, using RLT

buffer with 10 μl beta-mercaptoethanol. RNA was DNase treated with a Qiagen RNase-Free DNase Set. Separate RNA extractions were performed on leaf tissue and an immature flower sample (~5 days post flower development) for *A. githago*, *S. vulgaris*, and *S. latifolia*. Two different tissues were used to increase detection of diverse transcripts, but the two RNA samples were pooled equally by mass for each species prior to library construction, so individual reads cannot be assigned to leaf or floral tissues. Only leaf tissue was used for *S. conica* as the individual had not yet begun flowering at the time of RNA extraction. Both tissue types were harvested at 4 weeks post-germination, and RNA integrity and purity were checked on a TapeStation 2200 and a Nanodrop 2000.

Iso-Seq library construction and sequencing was performed at the Arizona Genomics Institute. Library construction was done using PacBio's SMRTbell Express Template Prep Kit 2.0. The four libraries were barcoded and pooled. The multiplexed pool was sequenced with a PacBio Sequel II platform on two SMRT Cells using a Sequencing Primer V4, Sequel II Bind Kit 2.0, Internal Control 1.0, and Sequel II Sequencing Kit 2.0. Raw movie files were processed to generate circular consensus sequences (CCSs) using PacBio's SMRT Link v9.0.0.92188 software (Pacific Biosciences 2020). Demultiplexing was performed with lima v2.0.0 and the --isoseq option. Full-length non-chimeric (FLNC) sequences were generated with the refine command and the --require_polya option in the IsoSeq3 (v3.4.0) pipeline. Clustering of FLNCs into isoforms was then performed with the cluster command in IsoSeq3 with the --use-qvs option. The two SMRT Cells produced similar outputs with 5.8M and 5.9M raw reads, which resulted in 3.9M CCSs for each cell (3.5M and 3.4M retained after demultiplexing). The results of demultiplexing, FLNC filtering, and clustering are shown in Table 4.2.

Extraction of *aaRS transcript sequences*

Arabidopsis aaRS genes were identified from published sources (DUCHÊNE *et al*. 2005b; WARREN AND SLOAN 2020b) and the corresponding protein sequences were obtained from the Araport11 genome annotation (201606 release). Homologs from the high-quality (HQ) clustered isoforms from each species

(including the previously published *S. noctiflora* dataset; (WILLIAMS *et al*. 2020)) were identified with a custom Perl script (iso-seq_blast_pipeline.pl available at GitHub: https://github.com/warrenjessica/Iso-Seq_scripts) that performed a tBLASTn search with each *Arabidopsis* aaRS sequence, requiring a minimum sequence identity of 50% and a minimum query length coverage of 50%. All HQ clusters that satisfied these criteria were retained by setting the --min_read parameter to 2 (the IsoSeq3 clustering step already excludes singleton transcripts).

Transcript processing and targeting prediction

The longest ORF was extracted from each aaRS transcript using the EMBOSS v. 6.6.0 (RICE *et al*. 2000) getorf program with the options: -minsize 75 -find 1. Many Iso-Seq transcripts differed in length by only by a few nucleotides in the UTR region but resulted in identical ORFs. Therefore, all identical ORFs were collapsed for downstream targeting and phylogenetic analysis. Collapsed ORFs were translated into protein coding sequences for localization analysis. TargetP v.2.0 (ALMAGRO ARMENTEROS *et al*. 2019), LOCALIZER v.1.0.4 (SPERSCHNEIDER *et al*. 2017), and Predotar v.1.04 (SMALL *et al*. 2004) were each used to predict targeting probabilities of each coding sequence. All programs were run with the plant option.

Determination of gene copy number and genome assembly scanning for undetected genes

Very similar transcripts can be the product of different genes, alleles, or sequencing errors. In order to infer the number of unique genes for each related set of transcripts in a species, CD-HIT-EST v. 4.8.1 (FU *et al*. 2012) was used to further cluster transcripts into groups. For this clustering step, sequences were first aligned with MAFFT v. 7.245 (KATOH AND STANDLEY 2013) with default settings and trimmed by eye to remove terminal sequence ends with gaps and N-terminal extensions that were not present on all sequences. Any two sequences in which the coding region shared greater than 98% sequence similarity were collapsed into a single gene cluster (CD-HIT-EST options -c 0.98 -n 5 -d 0).

Each cluster of transcripts was considered a single gene, and the transcript with the highest expression and longest length was retained as the representative sequence for the gene.

To check for the possibility that a cytosolic gene had gained a transit peptide but was undetected in Iso-Seq data (due to low expression or representation in the sequencing library), all cytosolic genes that appeared to lack transit peptides were checked for immediately upstream start codons in the corresponding nuclear genome assembly (see Chapter 3). Representative transcripts from each gene cluster were translated and BLASTed (tblastn) against the nuclear assembly, and scaffolds with a hit to the first exon of the protein were extracted and analyzed with the ExPASy Translate tool (ARTIMO *et al*. 2012). The ORF found in the genome assembly was then compared to the ORF generated from the transcript and inspected for length differences. If an upstream Met was present, the upstream sequence was appended to the rest of the gene and re-run through the targeting prediction software described above.

Occasionally, when BLASTing cytosolic aaRS proteins to nuclear assemblies, additional genes were discovered that were entirely absent from the Iso-Seq data (genes marked with ** in Figs. 4.3-4.19). In these cases, the region that aligned to the first exon of the expressed paralog was used for phylogenetic and targeting analysis.

Sequence alignment and maximum likelihood phylogenetic analysis

After clustering transcripts by sequence similarity (see above), the coding region of the longest transcript for each gene was retained for phylogenetic analysis. If two or more transcripts were tied for the longest length, the one with higher expression level was used. Retained sequences for each aaRS gene family were aligned using MAFFT v. 7.245 (KATOH AND STANDLEY 2013) with default settings. Sequences were trimmed by eye to remove poorly aligned regions, and maximum likelihood trees were produced using RAxML v.8.2.12 (STAMATAKIS 2014) with a GTRGAMMA model and rapid bootstrap analysis with a 100 replicates.

Table 4.1| Iso-Seq transcript coverage for identified aaRS genes. Only transcripts with coverage of two or more Iso-Seq reads were retained for analysis (i.e., singleton transcripts were discarded). Single *Sileneae* transcripts often mapped to multiple *A. thaliana* aaRS genes because of complex gene histories, including lineage-specific gene duplication events. Organelle localizations marked with * have been experimentally shown (DUCHÊNE *et al*. 2005b). Cytosolic localization were inferred based on the lack of a transit peptide, and a ? indicates that a transit peptide with potential mitochondrial targeting is present, but experimental localization studies were inconclusive (DUCHÊNE *et al*. 2005b) .

aaRS	**Mapped TAIR gene**	**Localization in *A. thaliana***	**Iso-Seq transcript coverage**				
			Agrostemma	*S. conica*	*S. latifolia*	*S. noctiflora*	*S. vulgaris*
AlaRS	At1g50200	Cytosol/Mitochondria*/Plastid*	1084	786	2630	60	1617
AlaRS	At5g22800	Mitochondria*/Plastid*	189	228	143	9	365
ArgRS	At1g66530:At4g26300	Cytosol:Plastid*	190	139	265	54	511
AsnRS	At1g70980:At5g56680	Cytosol:Cytosol	135	85	206	42	208
AsnRS	At4g17300	Mitochondria*/Plastid*	68	48	67	14	40
AsnRS	At3g07420	Cytosol	23	0	0	4	2
AspRS	At4g26870:At4g31180	Cytosol:Cytosol	165	62	114	82	140
AspRS	At4g33760	Mitochondria*/Plastid*	44	37	37	6	61
CysRS	At3g56300:At5g38830:At2g31170	Cytosol:Cytosol:Mitochondria*/Plastid*	135	102	229	93	116
GlnRS	At1g25350	Cytosol	640	194	301	65	554
GluRS	At5g26710	Cytosol	139	510	342	97	268
GluRS	At5g64050	Mitochondria*/Plastid*	92	48	130	35	86
GlyRS	At1g29870:At1g29880:At3g44740	Cytosol:Mitochondria*:Cytosol	896	445	1225	90	552
GlyRS	At3g48110	Mitochondria*/Plastid*	114	73	165	6	225
HisRS	At3g02760	Cytosol	525	221	357	11	468
HisRS	At3g46100	Mitochondria*/Plastid*	62	23	36	19	29
IleRS	At4g10320	Cytosol	713	687	812	31	684
IleRS	At5g49030	?/Plastid*	472	258	498	6	331
LeuRS	At1g09620	Cytosol/Mitochondria*	1441	732	829	27	1672
LeuRS	At4g04350	Plastid*	90	27	77	2	94
LysRS	At3g13490	Mitochondria*/Plastid*	107	148	199	21	376
LysRS	At3g11710	Cytosol	172	102	501	61	314
MetRS	At3g55400	Mitochondria*/Plastid*	197	11	77	10	51
MetRS	At2g40660	Cytosol	53	13	19	33	26
MetRS	At4g13780	Cytosol	202	258	162	62	363
PheRS	At3g58140	Mitochondria*/Plastid*	56	22	33	23	35
PheRS	At1g72550	Cytosol	191	63	140	37	293

PheRS	At4g39280	Cytosol	81	118	161	69	165
ProRS	At3g62120	Cytosol	133	233	384	116	356
ProRS	At5g52520	Mitochondria*/Plastid*	48	51	73	45	64
SerRS	At5g27470	Cytosol	167	67	68	98	187
SerRS	At1g11870	Mitochondria*/Plastid*	35	3	20	0	17
ThrRS	At2g04842	Mitochondria*/Plastid*	120	155	65	53	135
ThrRS	At1g17960:At5g26830	Cytosol:Mitochondria*	390	238	384	66	184
TrpRS	At3g04600	Cytosol	29	21	24	32	47
TrpRS	At2g25840	Mitochondria*/Plastid*	19	8	5	6	18
TyrRS	At1g28350:At2g33840	Cytosol:Cytosol	326	12	97	14	189
TyrRS	At3g02660	Mitochondria*/Plastid*	12	0	8	13	12
ValRS	At1g14610:At1g27160	Mitochondria*:Cytosol	923	483	600	56	858
ValRS	At5g16715	?/Plastid*	66	25	108	7	128

Table 4.2| Yields from Iso-Seq libraries.

Species	**CCSs**	**FLNCs**	**Clusters (HQ)**
A. githago	2034058	2029458	153293
S. conica	1535114	1525890	94736
S. latifolia	1584689	1580565	122274
S. vulgaris	1765441	1762081	126377

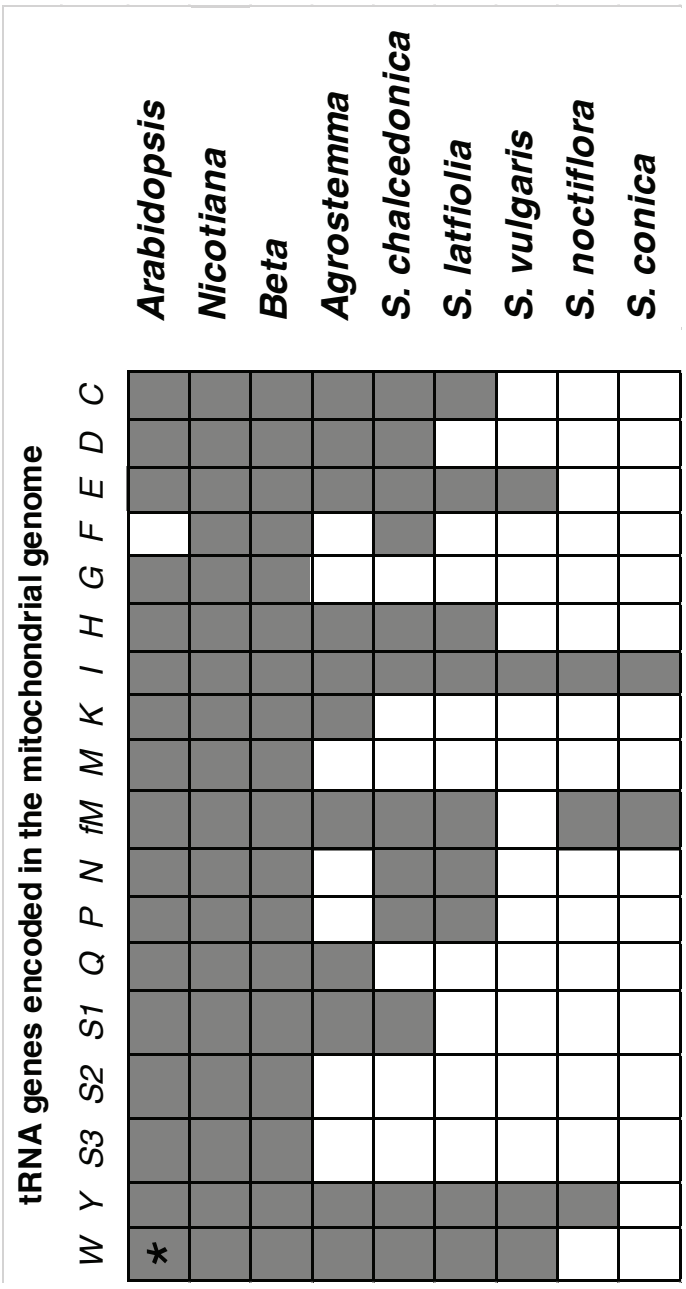


Figure 4.1| The loss of mt-tRNA genes in the mitogenomes of multiple *Sileneae* species. Singe-letter codes are used for amino acid abbreviations. Gray filled squares indicate the presence of the corresponding tRNA gene in the mitogenome. *Arabidopsis* tRNA-Trp (W) is marked with an * because multiple studies have failed to detect its expression and the import of a nuclear-encoded tRNA-Trp has been demonstrated (DUCHÊNE AND MARECHAL-DROUARD 2001; WARREN *et al*. 2020)

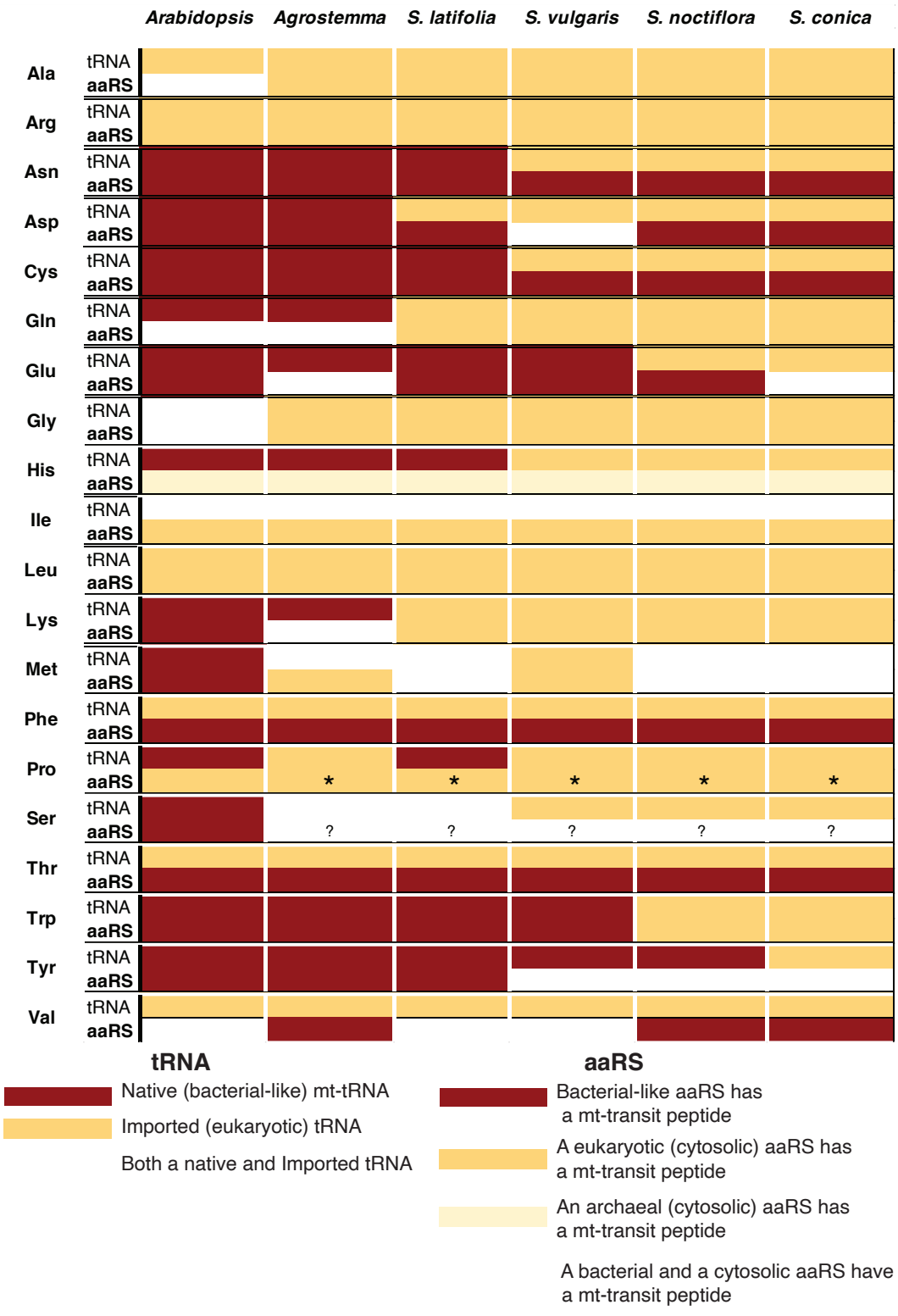

Arabidopsis
Agrostemma
S. latifolia
S. vulgaris
S. noctiflora
S. conica
Ala
Arg
Asn
Asp
Cys
Gln
Glu
Gly
His
Ile
Leu
Lys
Met
Phe
Pro
Ser
Thr
Trp
Tyr
Val
tRNA
aaRS
*
?
tRNA
Native (bacterial-like) mt-tRNA
Imported (eukaryotic) tRNA
Both a native and Imported tRNA
aaRS
Bacterial-like aaRS has a mt-transit peptide
A eukaryotic (cytosolic) aaRS has a mt-transit peptide
An archaeal (cytosolic) aaRS has a mt-transit peptide
A bacterial and a cytosolic aaRS have a mt-transit peptide

Figure 4.2| The origin of mitochondrial tRNAs and the cognate mitochondrially-localized aaRSs in *A. thaliana* and multiple *Sileneae* species. The mitochondrial localization of aaRSs has been experimentally demonstrated in *A. thaliana* (see Table 4.1, (DUCHÊNE *et al*. 2005b)) and has been assigned for *Silenae* species based on localization prediction software. Mt-tRNAs and aaRSs have complex gene histories. Thus, tRNAs and aaRSs classified as "bacterial-like" include those derived from the mitochondria (alphaproteobacterial-like), the plastids (cyanobacterial-like), or other bacterial origins. All bacterial-like tRNAs are still encoded in the mitogenomes, whereas all eukaryotic mitochondrial tRNAs are imported from the cytosol. Solid colors represent the presence of only a single phylogenetic class of a tRNA or aaRS, while striped boxes indicate the presence of multiple aaRSs/tRNAs with differing phylogenic origins. The cytosolic TyrRS in angiosperms is also archaeal in origin, and it appears to have been retargeted to the mitochondria in some *Silene* species. An aaRS was classified as being predicted to be mitochondrially-targeted if the enzyme had 50 or more percentage points of targeting likelihood to the organelle (cumulatively between the three targeting prediction software programs). See Figs. 3-19 for more detailed targeting data for each gene family. An organellar GlnRS does not exist in most plant mitochondria (including *Arabidopsis*) as tRNA-Gln is aminoacylated by an indirect transamidation pathway with a nondiscriminating GluRS (SALINAS-GIEGÉ *et al*. 2015). This ancestral state appears to be retained in *Agrostemma*. All ProRS enzymes have eukaryotic origins, but organelle and cytosolic genes are distinct. ProRS enzymes marked with an * indicate that both organellar and cytosolic copies have a coding extension predicted to function as a transit peptide. No SerRS enzymes in *Sileneae* were strongly predicted to be targeted to the mitochondria.

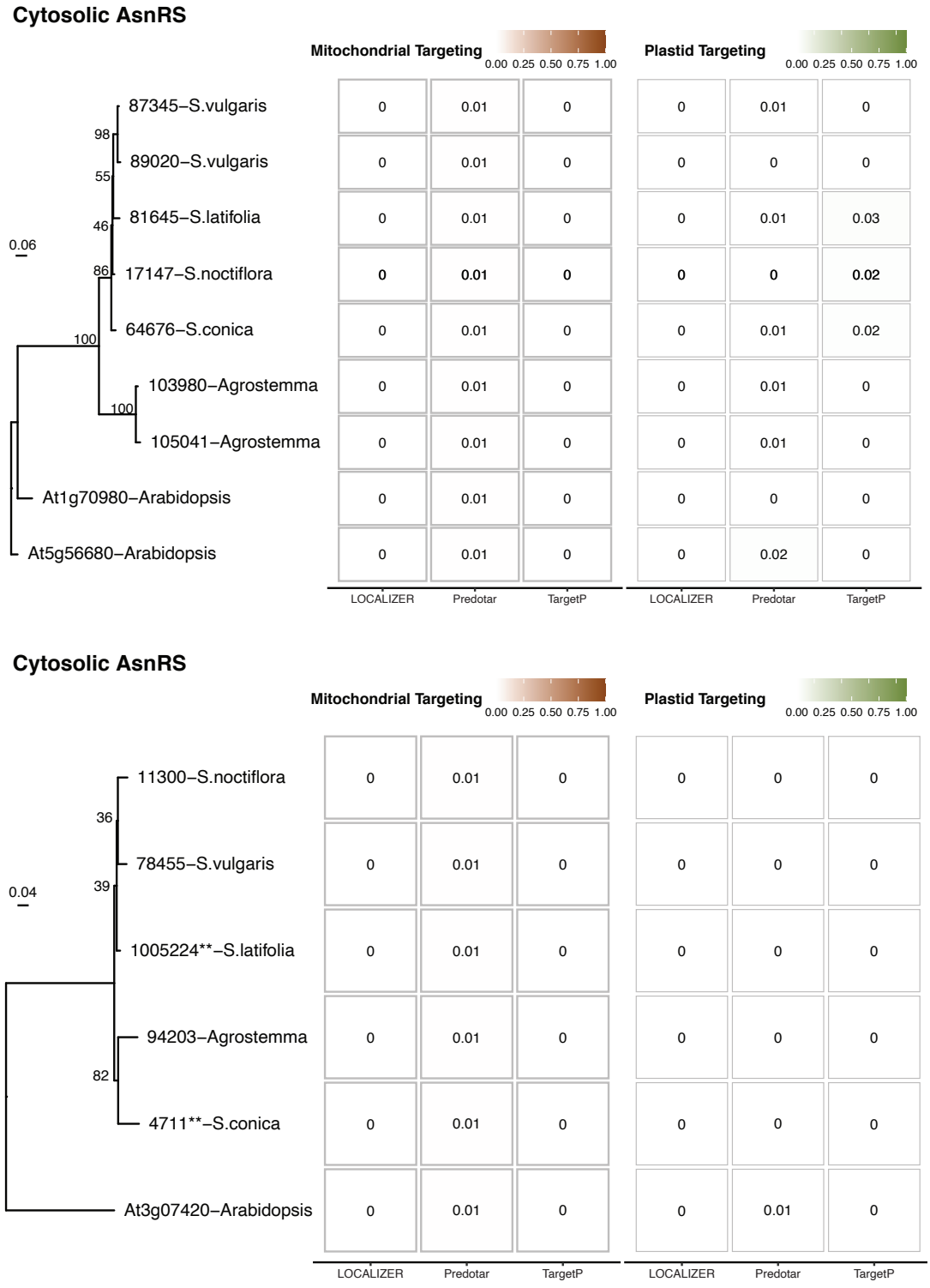


Figure 4.3| Maximum likelihood trees and targeting probabilities of cytosolic AsnRS genes in *A. thaliana* and *Sileneae*. Genes marked with ** indicate that this gene was undetected by Iso-Seq data but was identified in the nuclear assembly.

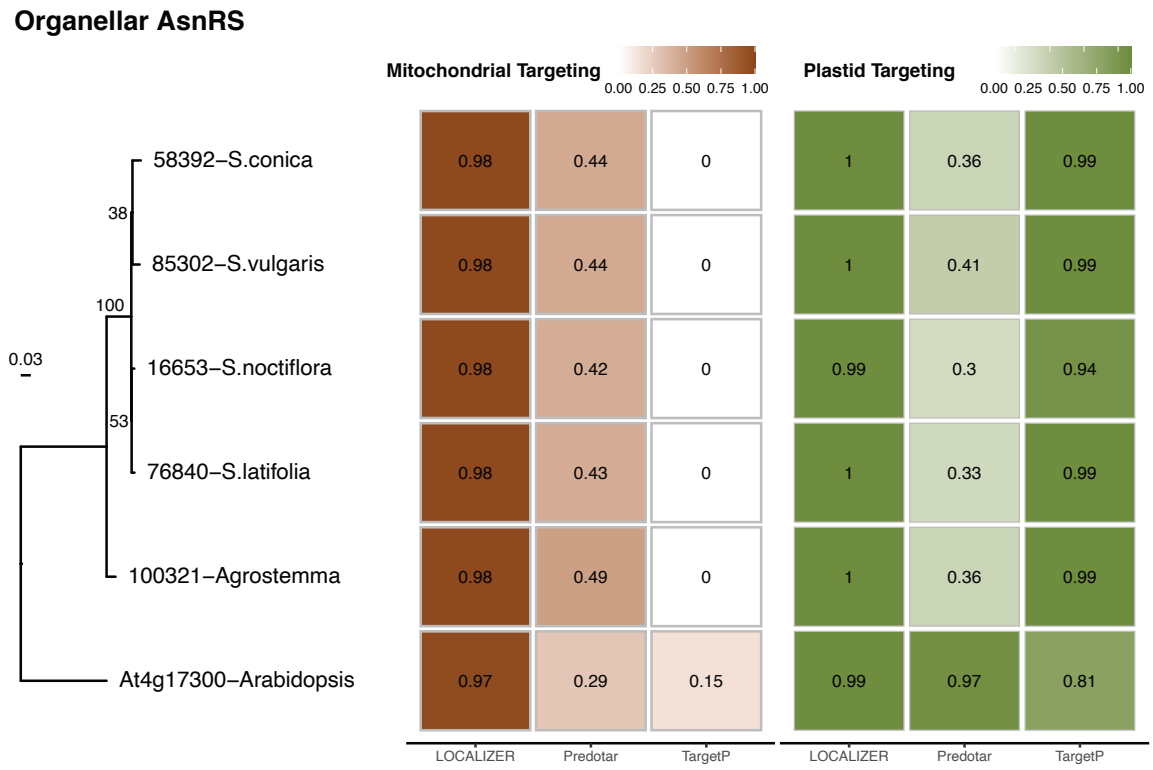


Figure 4.4| Maximum likelihood tree and targeting probabilities of organellar AsnRS genes in *A. thaliana* and *Sileneae*.

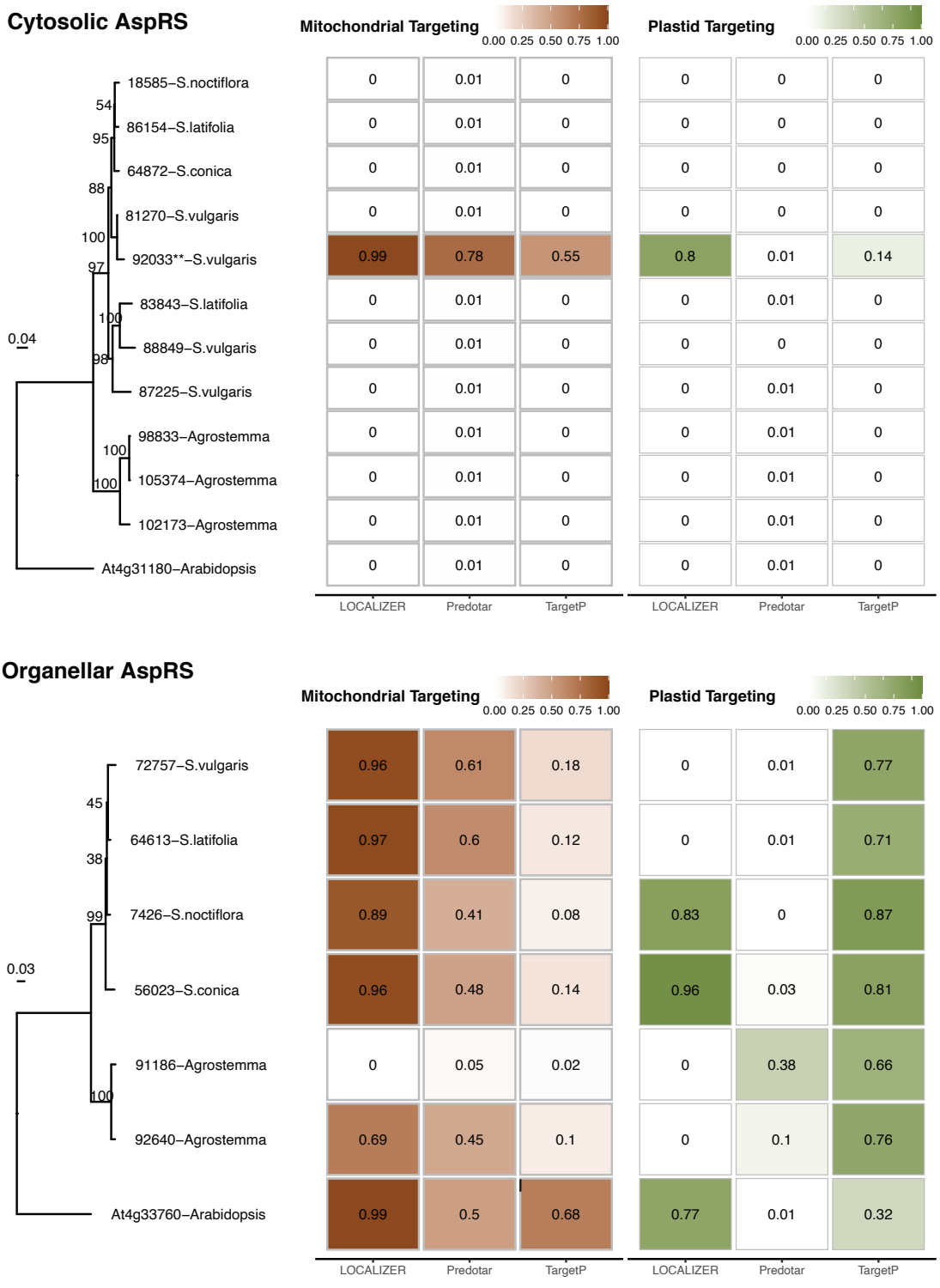


Figure 4.5| Maximum likelihood trees and targeting probabilities of cytosolic and organellar AspRS genes in *A. thaliana* and *Sileneae*. Genes marked with ** indicate that this gene was undetected by Iso-Seq data but was identified in the nuclear assembly.

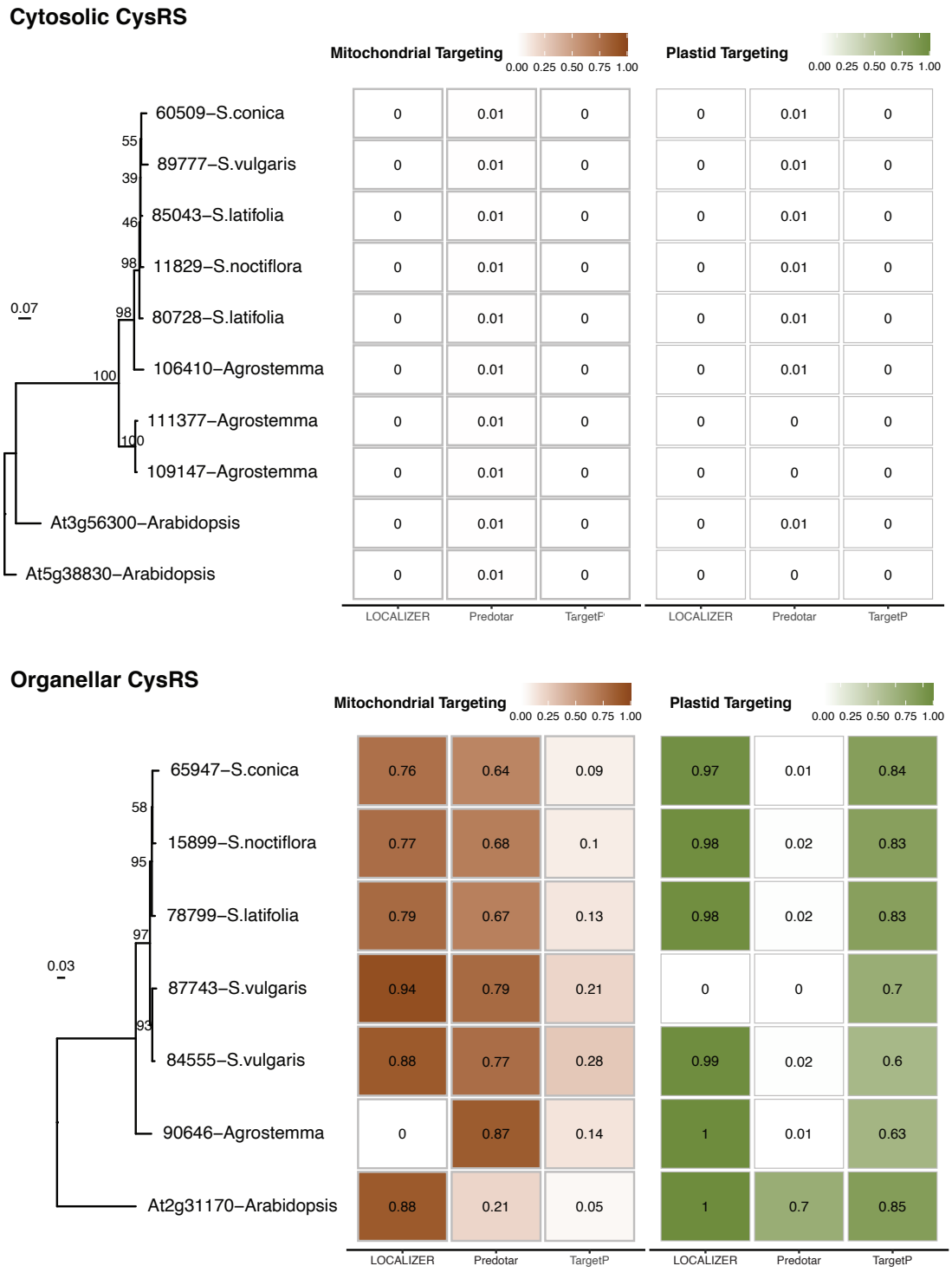


Figure 4.6| Maximum likelihood trees and targeting probabilities of cytosolic and organellar CysRS genes in *A. thaliana* and *Sileneae*.

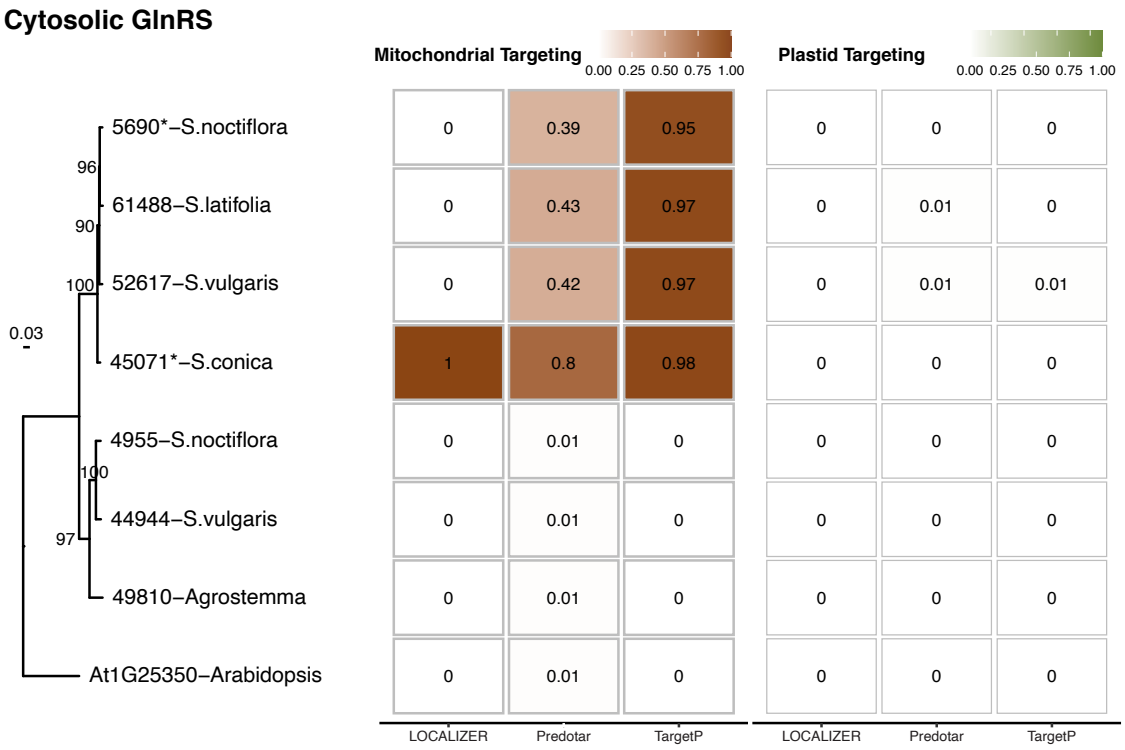


Figure 4.7| Maximum likelihood tree and targeting probabilities of cytosolic GlnRS genes in *A. thaliana* and *Sileneae*. Genes marked with an * indicate that this gene was detected in Iso-Seq data but without the 5´ terminal extension present in the nuclear assembly.

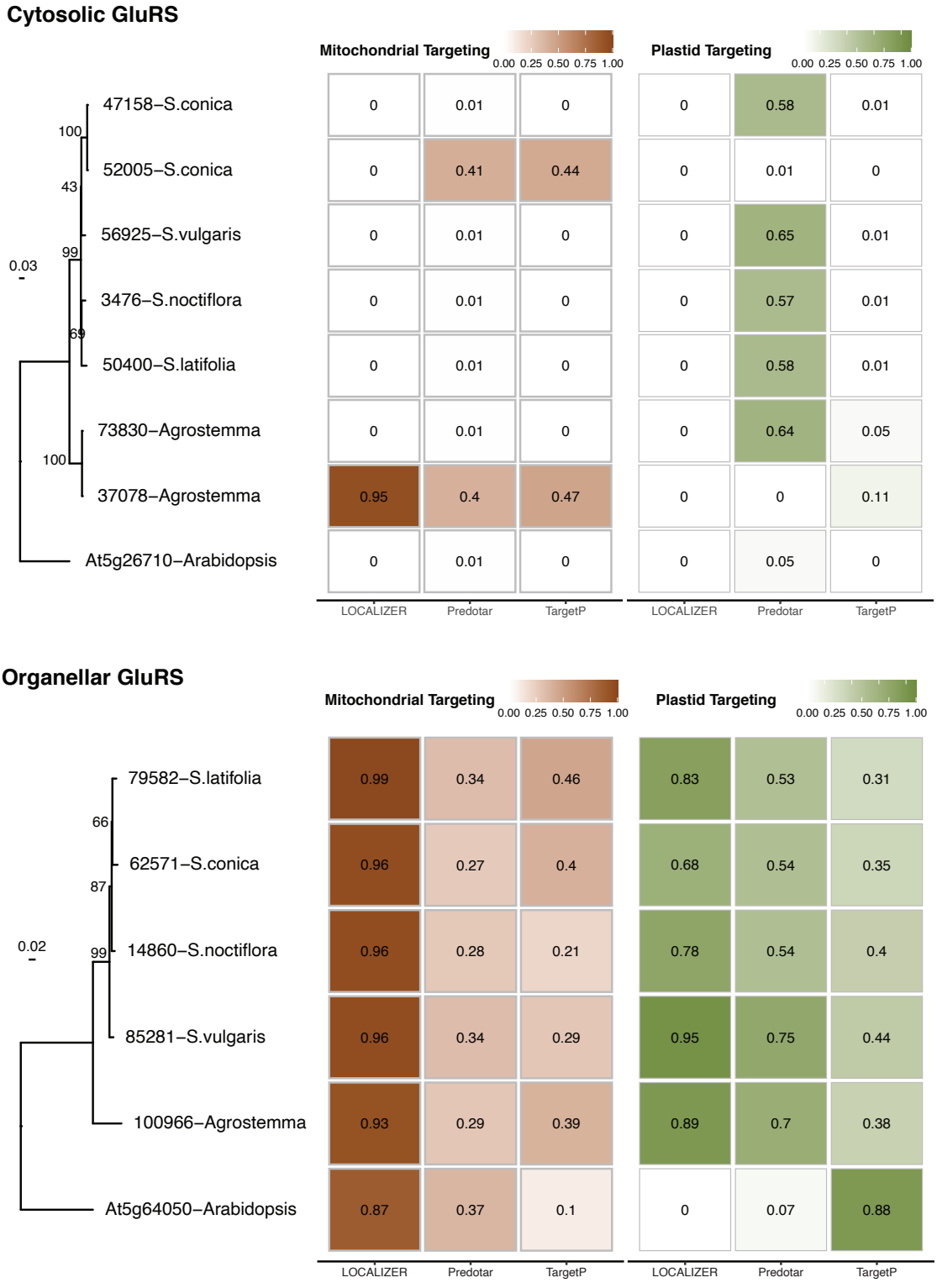


Figure 4.8| Maximum likelihood trees and targeting probabilities of cytosolic and organellar GluRS genes in *A. thaliana* and *Sileneae*

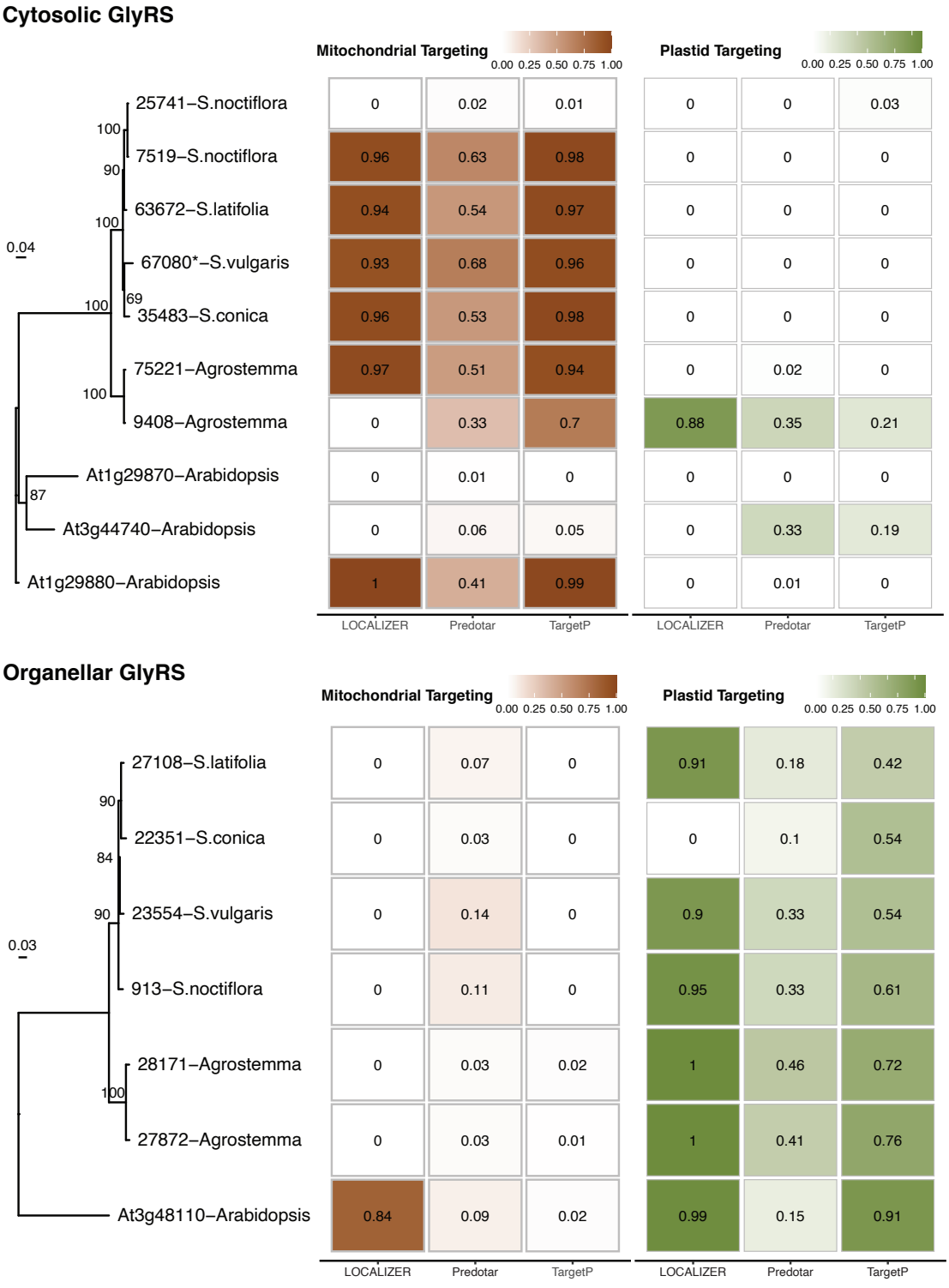


Figure 4.9| Maximum likelihood trees and targeting probabilities of cytosolic and organellar GlyRS genes in *A. thaliana* and *Sileneae*. Genes marked with an * indicate that this gene was detected in Iso-Seq data but without the 5´ terminal extension present in the nuclear assembly.

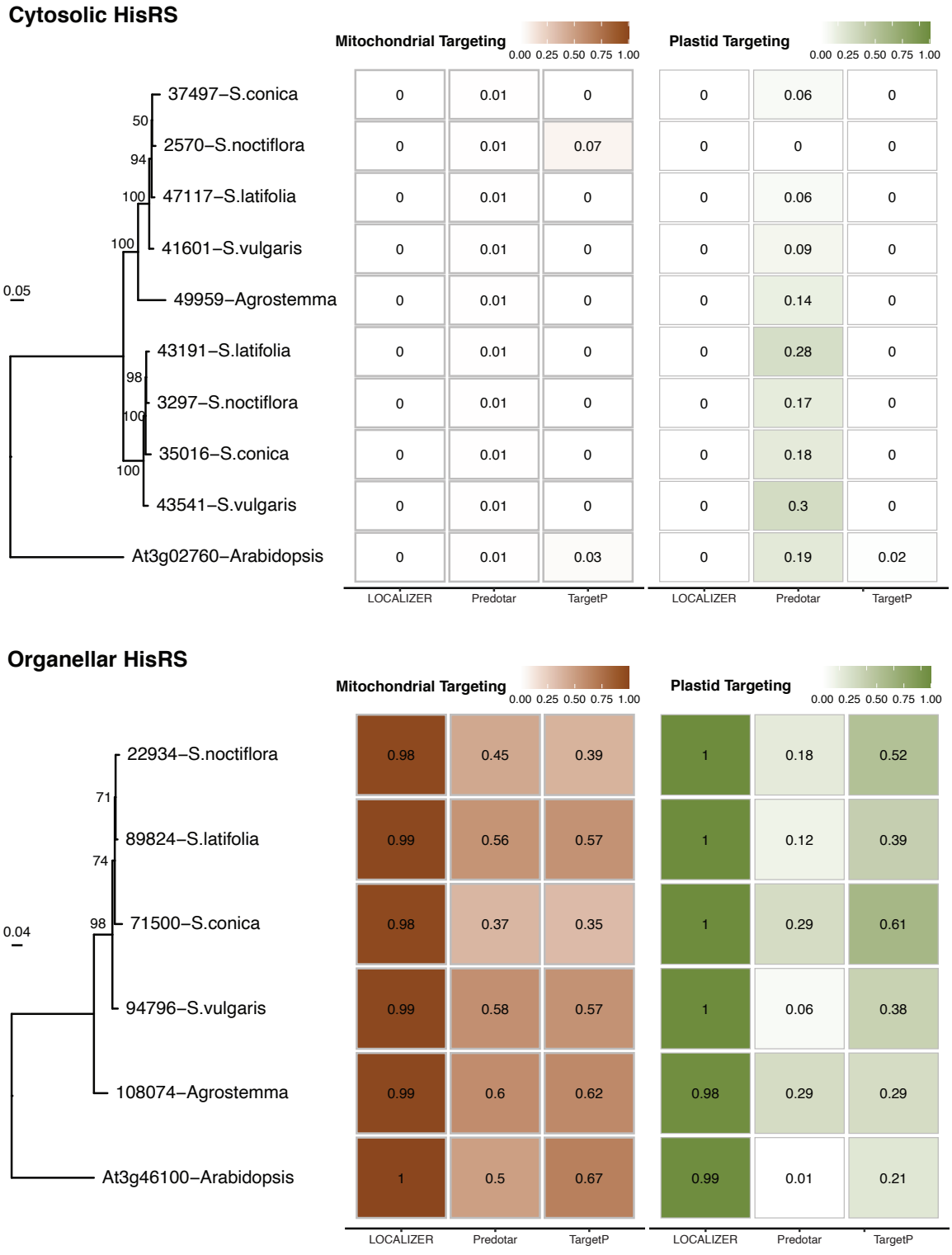


Figure 4.10| Maximum likelihood trees and targeting probabilities of cytosolic and organellar HisRS genes in *A. thaliana* and *Sileneae*

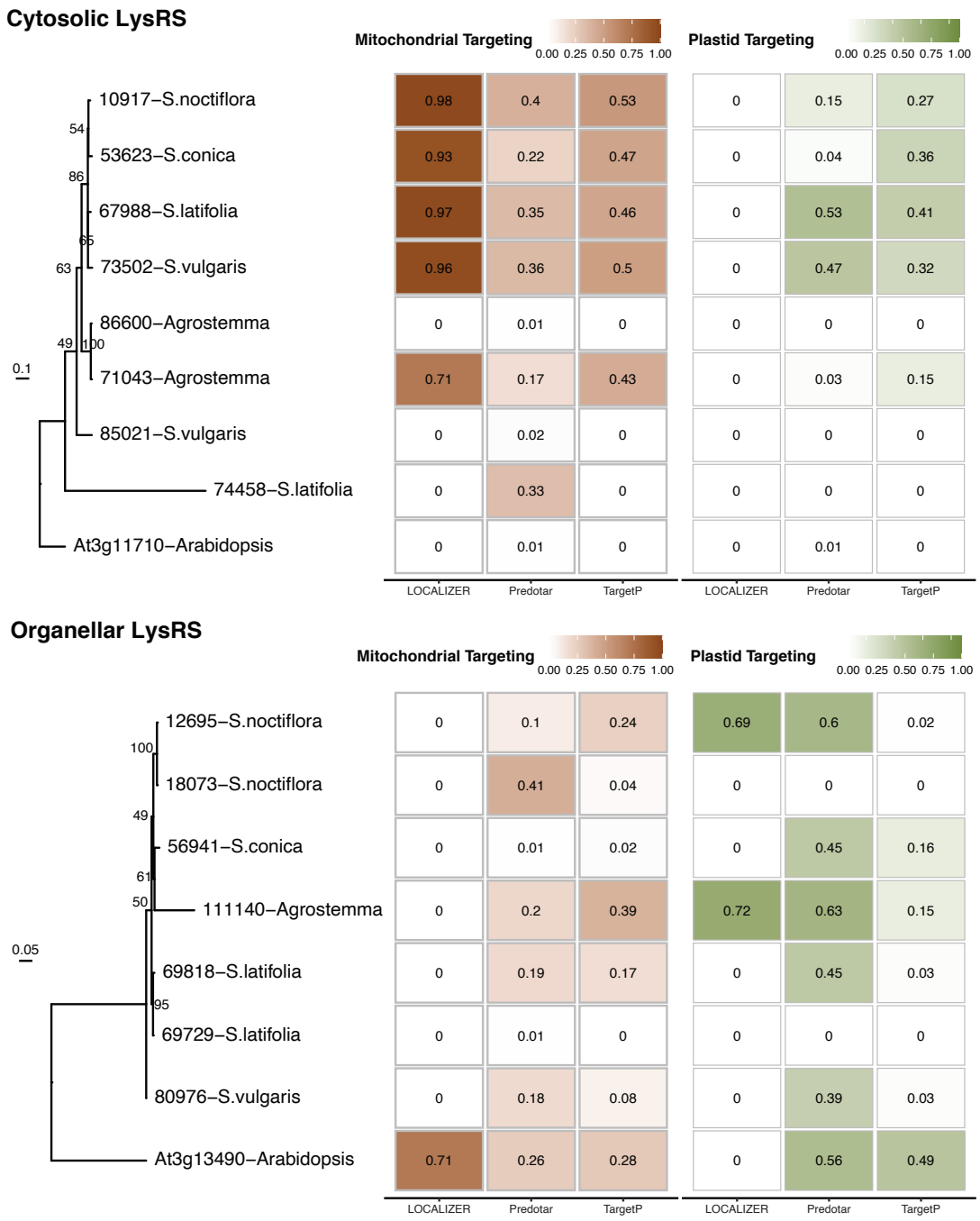


Figure 4.11| Maximum likelihood trees and targeting probabilities of cytosolic and organellar LysRS genes in *A. thaliana* and *Sileneae*

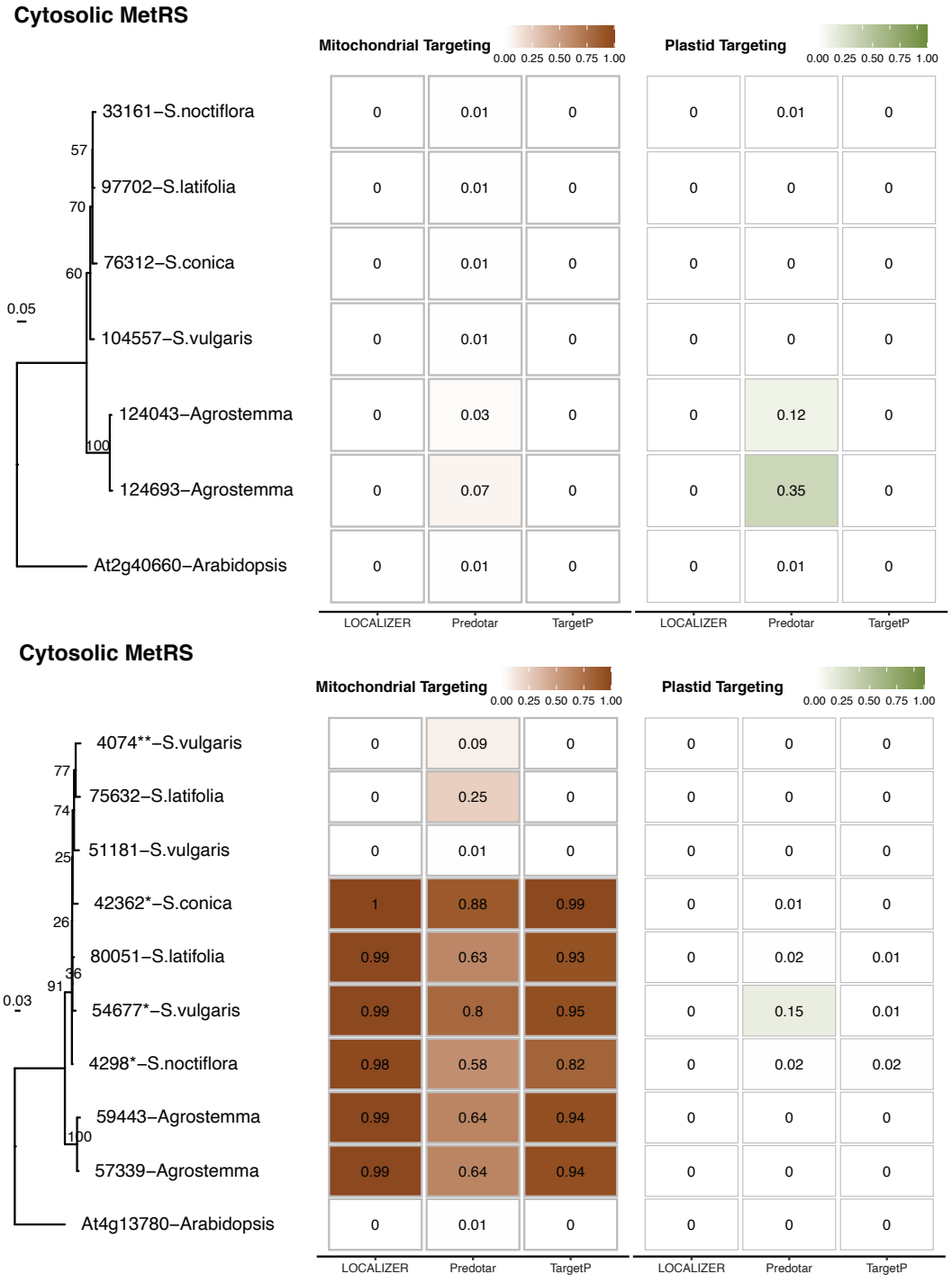


Figure 4.12| Maximum likelihood trees and targeting probabilities of cytosolic MetRS genes in *A. thaliana* and *Sileneae*. Genes marked with ** indicate that this gene was undetected by Iso-Seq data but was identified in the nuclear assembly. Genes marked with an * indicate that this gene was detected in Iso-Seq data but without the 5´ terminal extension present in the nuclear assembly.

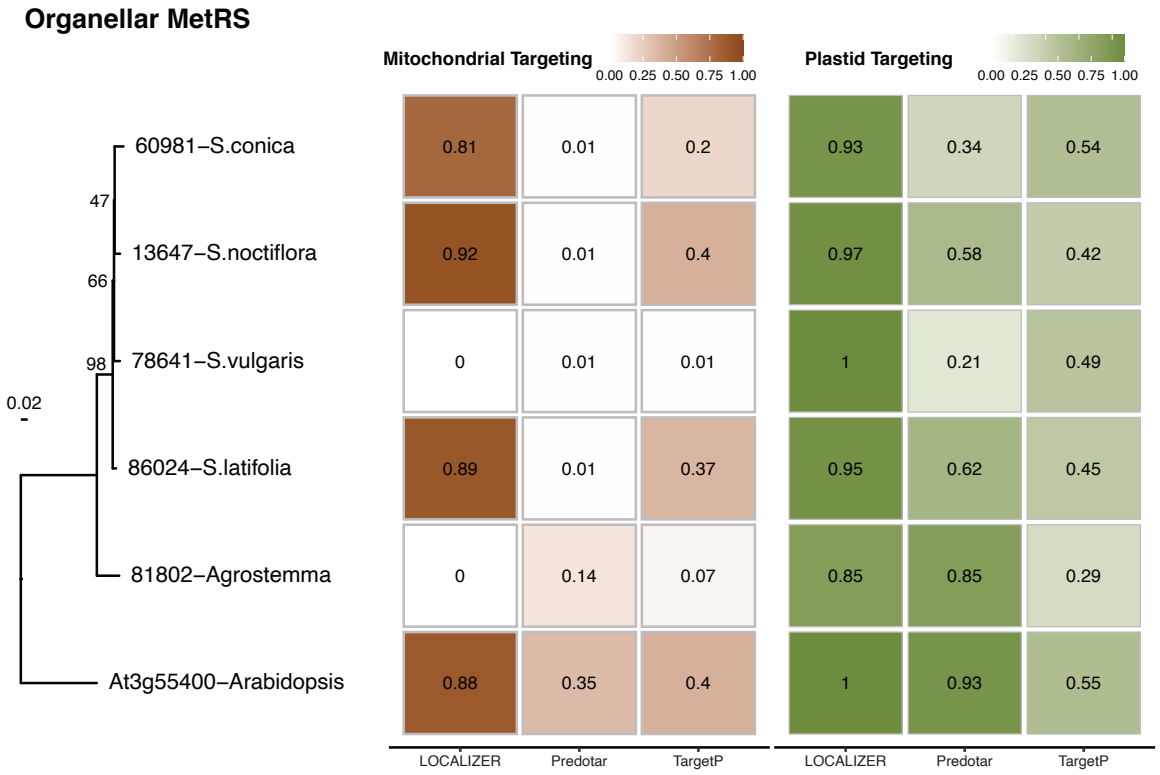


Figure 4.13| Maximum likelihood tree and targeting probabilities of organellar MetRS genes in *A. thaliana* and *Sileneae*

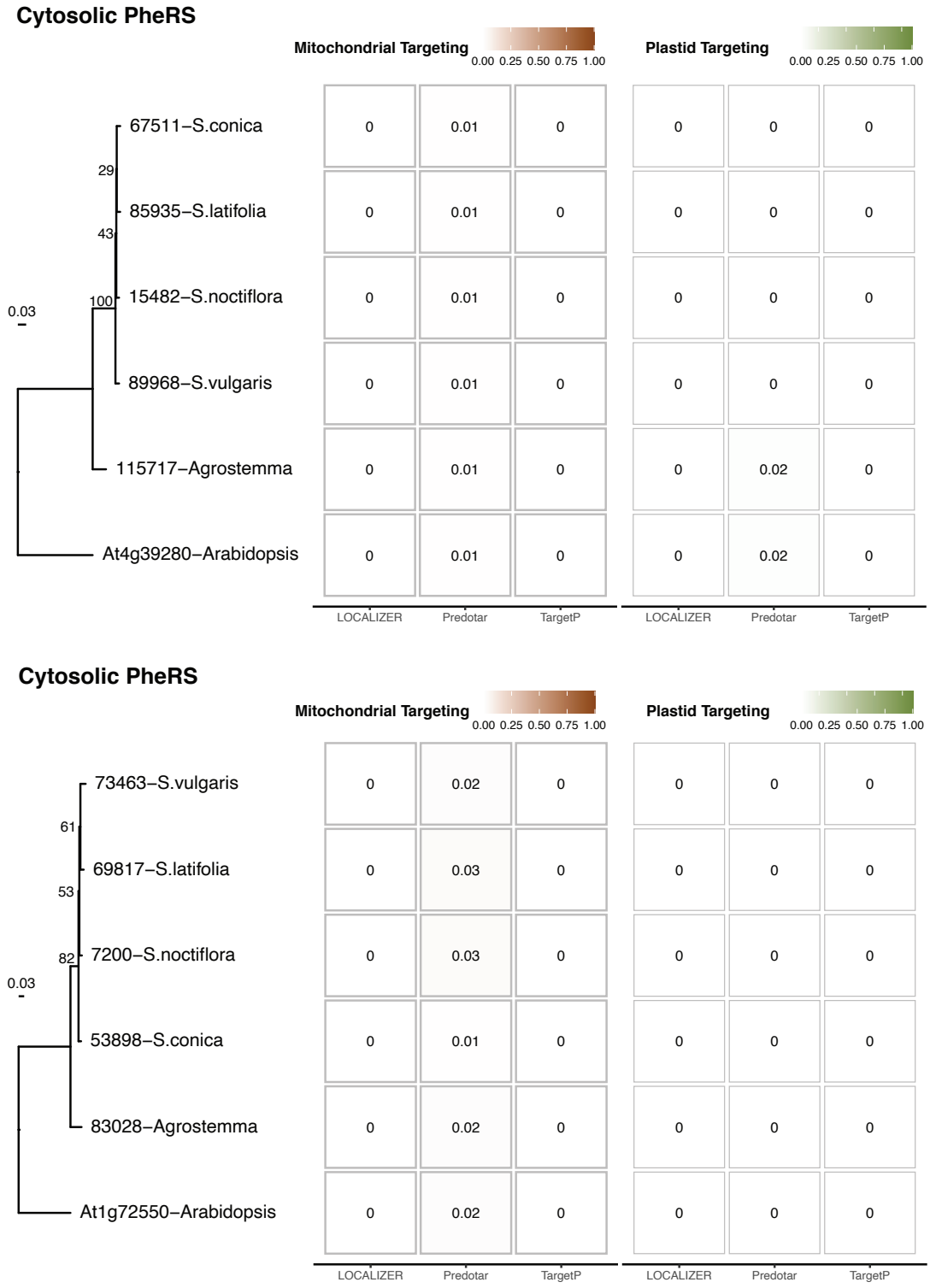


Figure 4.14| Maximum likelihood trees and targeting probabilities of cytosolic PheRS genes in *A. thaliana* and *Sileneae*

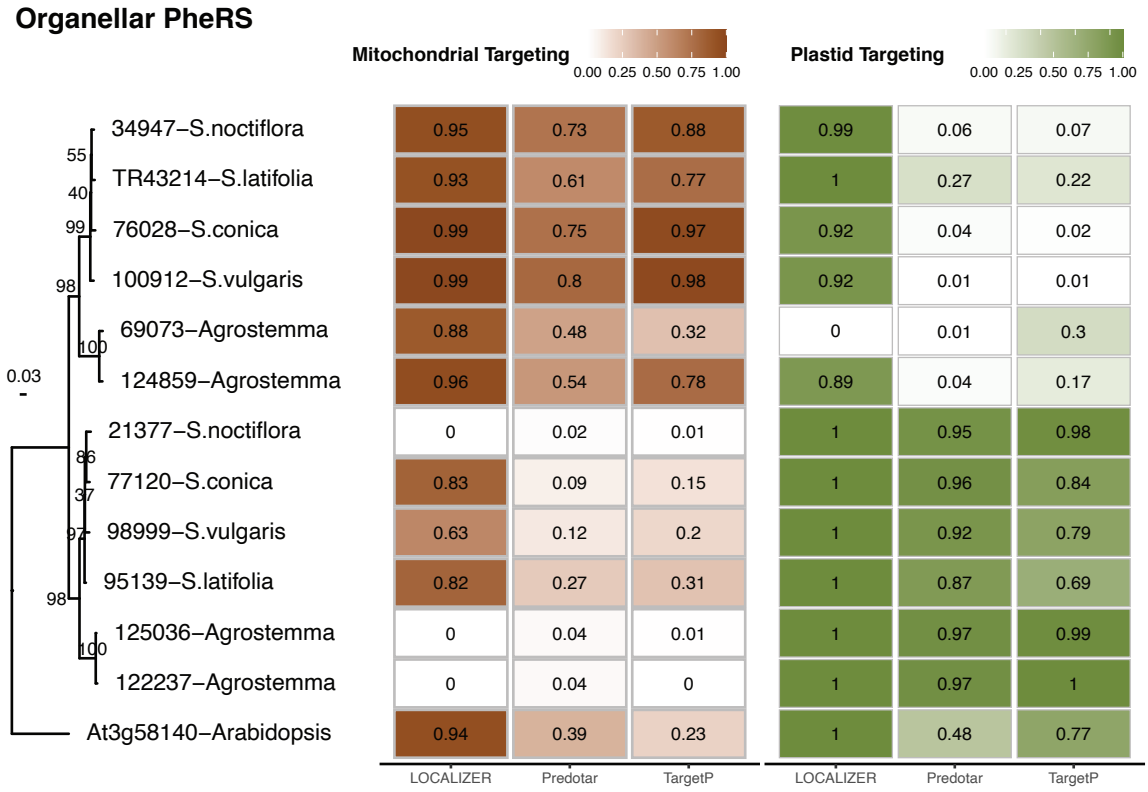


Figure 4.15| Maximum likelihood tree and targeting probabilities of organellar PheRS genes in *A. thaliana* and *Sileneae*. A Trinity (GRABHERR *et al*. 2011) assembled transcript was used for the *S. latifolia* TR43214 sequence (SLOAN *et al*. 2014b).

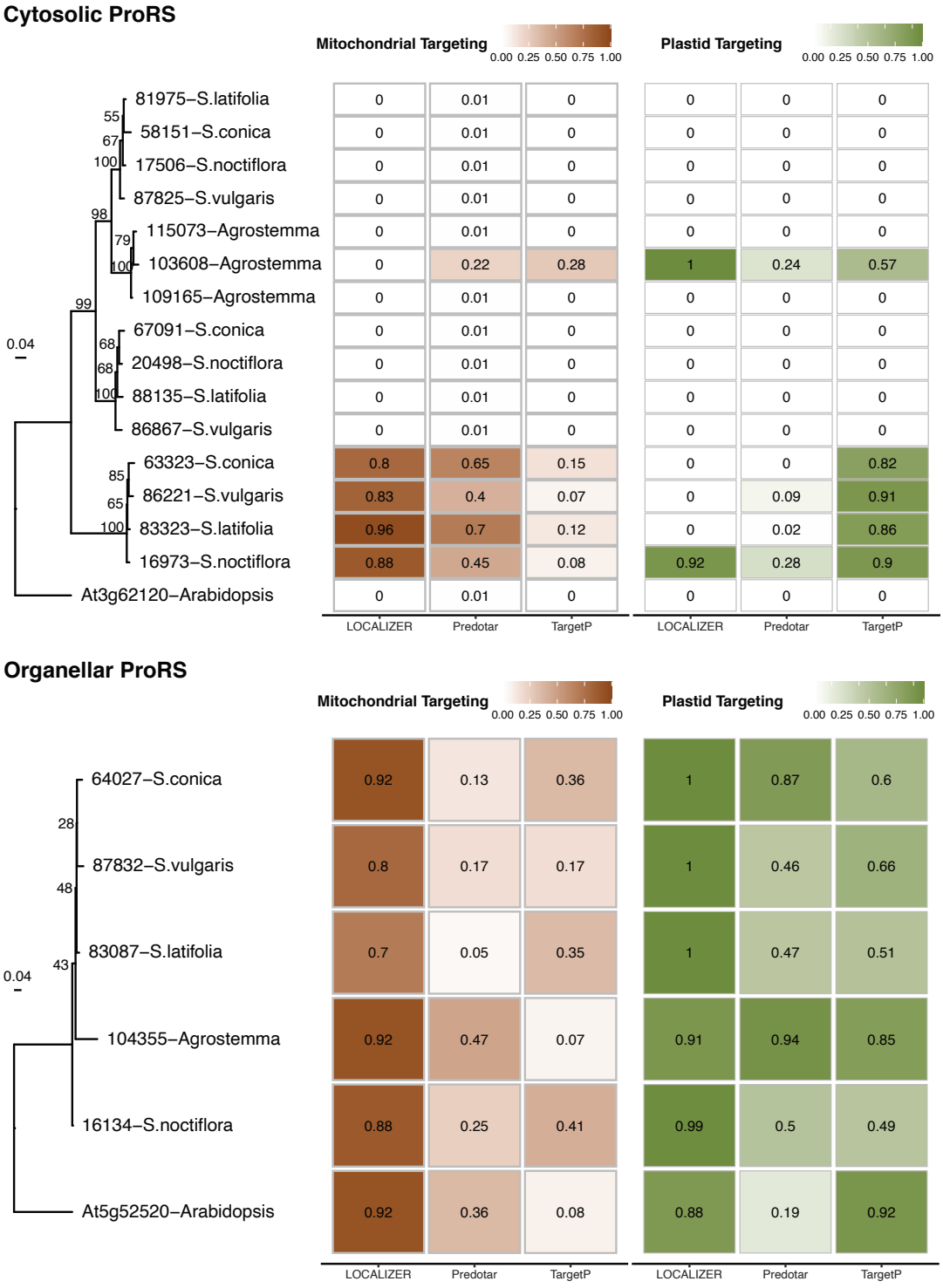


Figure 4.16| Maximum likelihood trees and targeting probabilities of cytosolic and organellar ProRS genes in *A. thaliana* and *Sileneae*

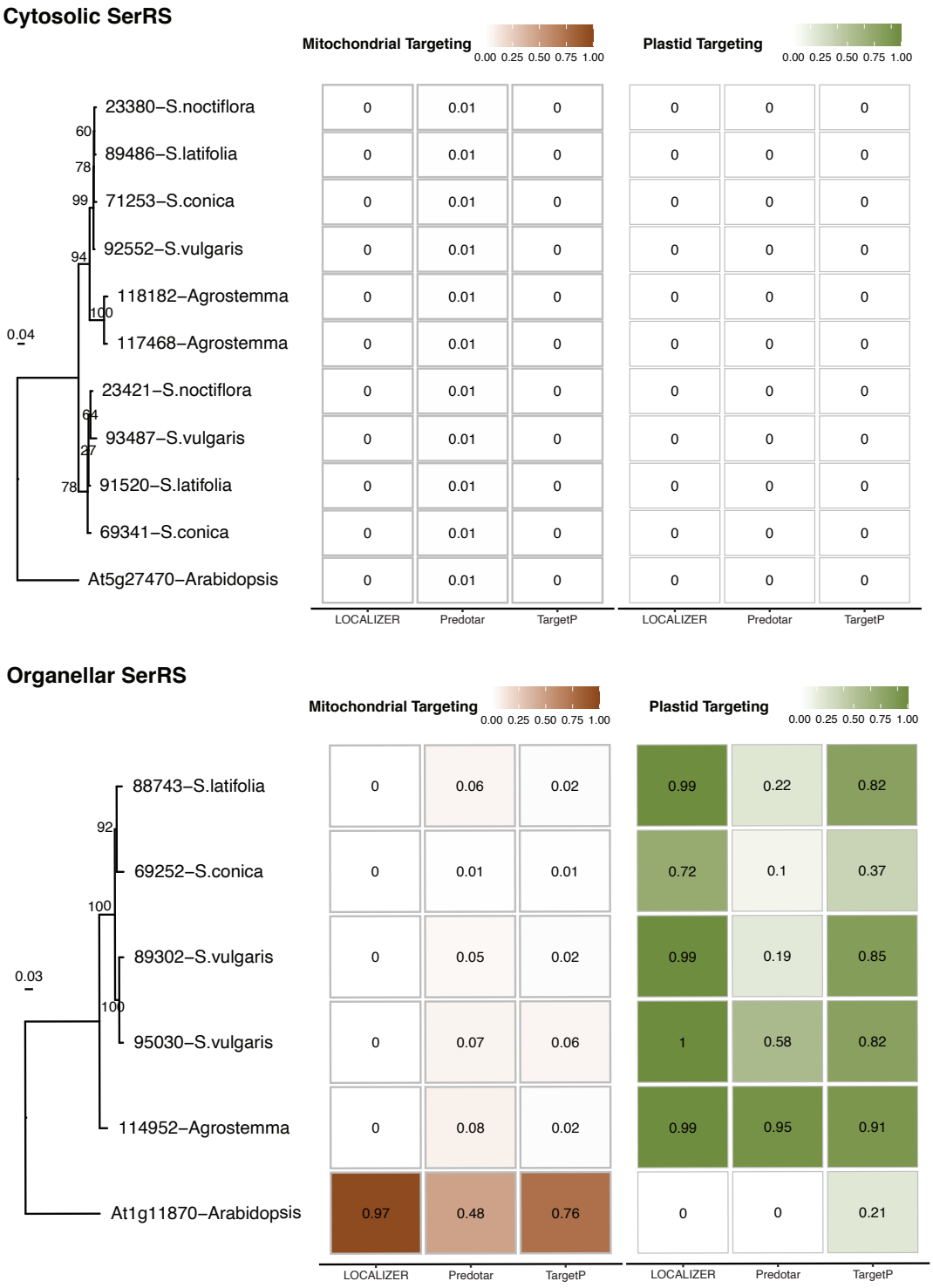


Figure 4.17| Maximum likelihood trees and targeting probabilities of cytosolic and organellar SerRS genes in *A. thaliana* and *Sileneae*

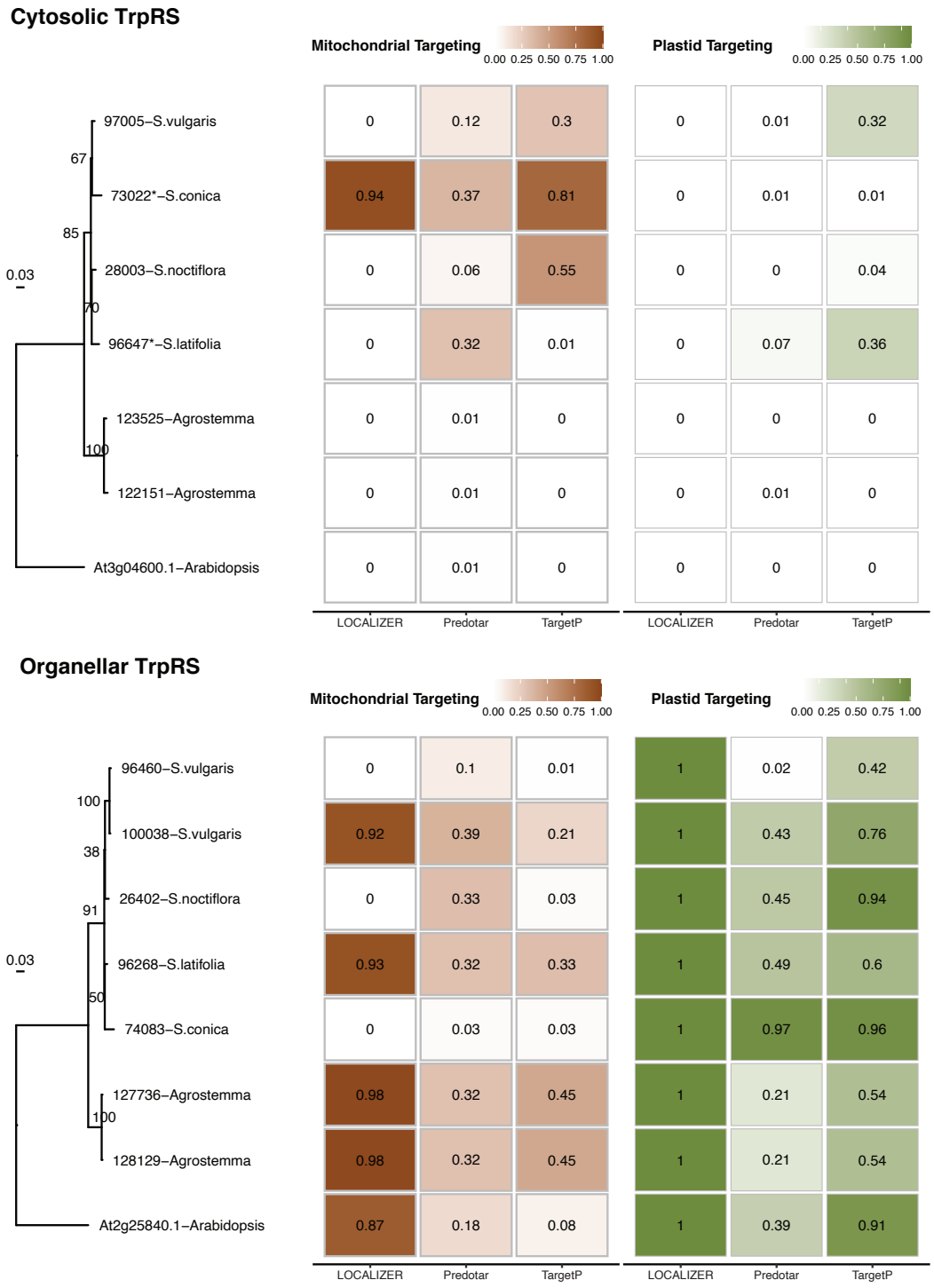


Figure 4.18| Maximum likelihood trees and targeting probabilities of cytosolic and organellar TrpRS genes in *A. thaliana* and *Sileneae*. Genes marked with an * indicate that this gene was detected in Iso-Seq data but without the 5´ terminal extension present in the nuclear assembly.

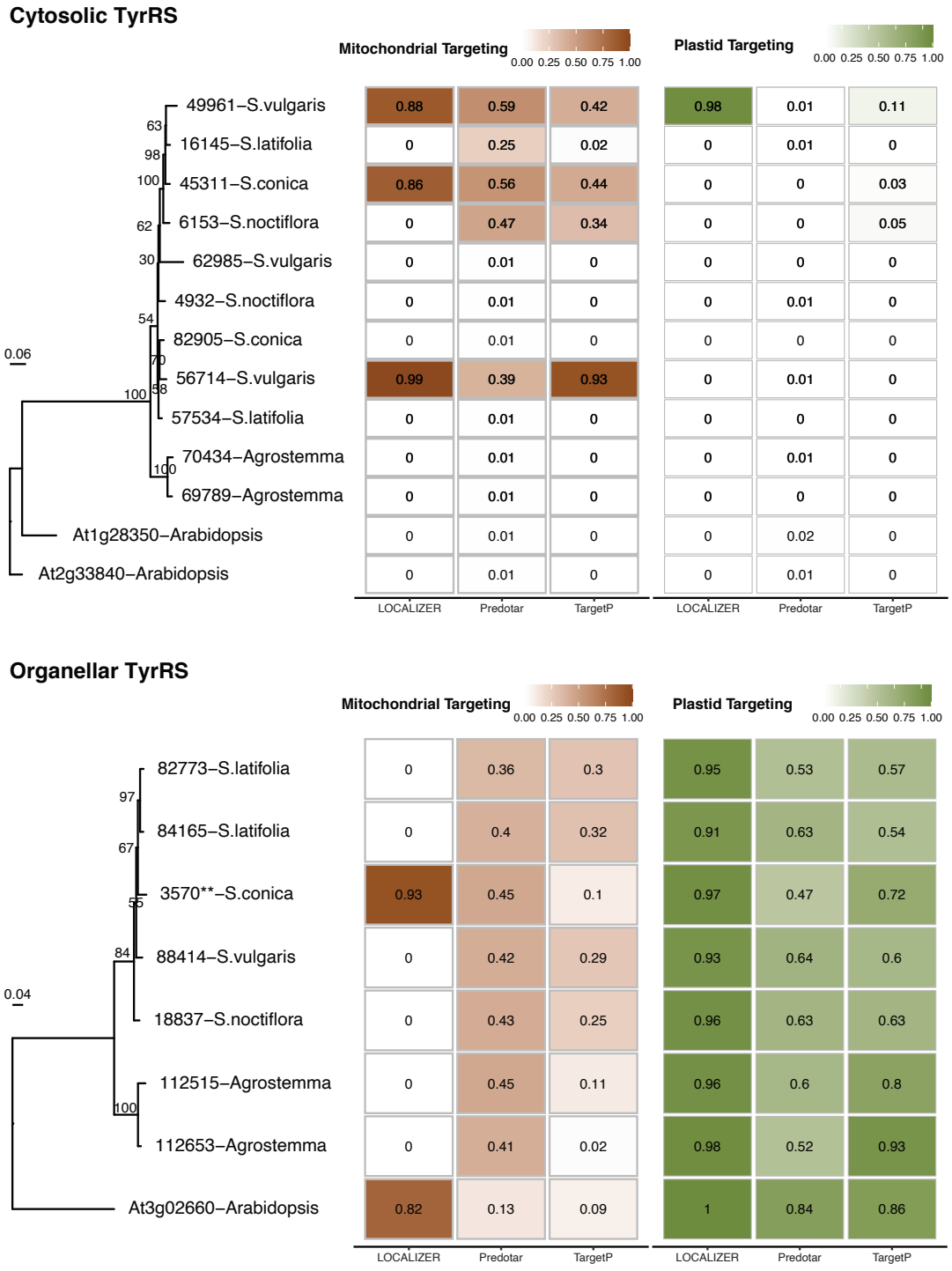


Figure 4.19| Maximum likelihood trees and targeting probabilities of cytosolic and organellar TyrRS genes in *A. thaliana* and *Sileneae*. Genes marked with ** indicate that this gene was undetected by Iso-Seq data but was identified in the nuclear assembly.

LITERATURE CITED

Adams, K. L., and J. D. Palmer, 2003 Evolution of mitochondrial gene content: gene loss and transfer to the nucleus. Mol Phylogenet Evol 29**:** 380-395.

Adams, K. L., Y. L. Qiu, M. Stoutemyer and J. D. Palmer, 2002 Punctuated evolution of mitochondrial gene content: high and variable rates of mitochondrial gene loss and transfer to the nucleus during angiosperm evolution. Proc Natl Acad Sci U S A 99**:** 9905-9912.

Adrion, J. R., P. S. White and K. L. Montooth, 2016 The Roles of Compensatory Evolution and Constraint in Aminoacyl tRNA Synthetase Evolution. Mol Biol Evol 33**:** 152-161.

Alfonzo, J. D., V. Blanc, A. M. Estévez, M. A. T. Rubio and L. Simpson, 1999 C to U editing of the anticodon of imported mitochondrial tRNATrp allows decoding of the UGA stop codon in Leishmania tarentolae. The EMBO journal 18**:** 7056-7062.

Alfonzo, J. D., and D. Söll, 2009 Mitochondrial tRNA import--the challenge to understand has just begun. Biol Chem 390**:** 717-722.

Alkatib, S., T. T. Fleischmann, L. B. Scharff and R. Bock, 2012 Evolutionary constraints on the plastid tRNA set decoding methionine and isoleucine. Nucleic Acids Res 40**:** 6713-6724.

Almagro Armenteros, J. J., M. Salvatore, O. Emanuelsson, O. Winther, G. von Heijne *et al.*, 2019 Detecting sequence signals in targeting peptides using deep learning. Life Sci Alliance 2.

Altman, S., 2007 A view of RNase P. Mol Biosyst 3**:** 604-607.

Alverson, A. J., X. X. Wei, D. W. Rice, D. B. Stern, K. Barry *et al.*, 2010 Insights into the Evolution of Mitochondrial Genome Size from Complete Sequences of Citrullus lanatus and Cucurbita pepo (Cucurbitaceae). Molecular Biology and Evolution 27**:** 1436-1448.

Anderson, S., A. T. Bankier, B. G. Barrell, M. H. de Bruijn, A. R. Coulson *et al.*, 1981 Sequence and organization of the human mitochondrial genome. Nature 290**:** 457-465.

Arimbasseri, A. G., N. H. Blewett, J. R. Iben, T. N. Lamichhane, V. Cherkasova *et al.*, 2015 RNA Polymerase III Output Is Functionally Linked to tRNA Dimethyl-G26 Modification. PLoS Genet 11**:** e1005671.

Arimura, S., J. Yamamoto, G. P. Aida, M. Nakazono and N. Tsutsumi, 2004 Frequent fusion and fission of plant mitochondria with unequal nucleoid distribution. Proc Natl Acad Sci U S A 101**:** 7805-7808.

Arita, M., T. Suematsu, A. Osanai, T. Inaba, H. Kamiya *et al.*, 2006 An evolutionary 'intermediate state' of mitochondrial translation systems found in Trichinella species of parasitic nematodes: co-evolution of tRNA and EF-Tu. Nucleic Acids Res 34**:** 5291-5299.

Arrieta-Montiel, M. P., and S. A. Mackenzie, 2011 Plant Mitochondrial Genomes and Recombination, pp. 65-82 in *Plant Mitochondria*, edited by F. Kempken. Springer New York, New York, NY.

Artimo, P., M. Jonnalagedda, K. Arnold, D. Baratin, G. Csardi *et al.*, 2012 ExPASy: SIB bioinformatics resource portal. Nucleic Acids Res 40**:** W597-603.

Ayan, G. B., H. J. Park and J. Gallie, 2020 The birth of a bacterial tRNA gene by large-scale, tandem duplication events. Elife 9.

Baker, A., and G. Schatz, 1987 Sequences from a prokaryotic genome or the mouse dihydrofolate reductase gene can restore the import of a truncated precursor protein into yeast mitochondria. Proc Natl Acad Sci U S A 84**:** 3117-3121.

Banerjee, R., S. Chen, K. Dare, M. Gilreath, M. Praetorius-Ibba *et al.*, 2010 tRNAs: cellular barcodes for amino acids. FEBS Lett 584**:** 387-395.

Berg, M., R. Rogers, R. Muralla and D. Meinke, 2005 Requirement of aminoacyl-tRNA synthetases for gametogenesis and embryo development in Arabidopsis. Plant J 44**:** 866-878.

Berglund, A. K., E. Spanning, H. Biverstahl, G. Maddalo, C. Tellgren-Roth *et al.*, 2009 Dual targeting to mitochondria and chloroplasts: characterization of Thr-tRNA synthetase targeting peptide. Mol Plant 2**:** 1298-1309.

Bergmann, P., P. Seyer, G. Burkard and J. H. Weil, 1984 Mapping of transfer RNA genes on tobacco chloroplast DNA. Plant Mol Biol 3**:** 29-36.

Bergthorsson, U., K. L. Adams, B. Thomason and J. D. Palmer, 2003 Widespread horizontal transfer of mitochondrial genes in flowering plants. Nature 424**:** 197-201.

Bermudez-Santana, C., C. S. Attolini, T. Kirsten, J. Engelhardt, S. J. Prohaska *et al.*, 2010 Genomic organization of eukaryotic tRNAs. BMC Genomics 11**:** 270.

Betat, H., Y. Long, J. E. Jackman and M. Morl, 2014 From end to end: tRNA editing at 5'- and 3'-terminal positions. Int J Mol Sci 15**:** 23975-23998.

Binder, S., A. Marchfelder and A. Brennicke, 1994 RNA editing of tRNA(Phe) and tRNA(Cys) in mitochondria of Oenothera berteriana is initiated in precursor molecules. Mol Gen Genet 244**:** 67-74.

Binder, S., A. Marchfelder, A. Brennicke and B. Wissinger, 1992 RNA editing in trans-splicing intron sequences of nad2 mRNAs in Oenothera mitochondria. J Biol Chem 267**:** 7615-7623.

Bjork, G. R., K. Jacobsson, K. Nilsson, M. J. Johansson, A. S. Bystrom *et al.*, 2001 A primordial tRNA modification required for the evolution of life? EMBO J 20**:** 231-239.

Boccaletto, P., M. A. Machnicka, E. Purta, P. Piatkowski, B. Baginski *et al.*, 2018 MODOMICS: a database of RNA modification pathways. 2017 update. Nucleic Acids Res 46**:** D303-D307.

Bonnard, G., A. Gobert, M. Arrivé, F. Pinker, T. Salinas-Giegé *et al.*, 2016 Transfer RNA maturation in Chlamydomonas mitochondria, chloroplast and the nucleus by a single RNase P protein. Plant J 87**:** 270-280.

Boore, J. L., J. R. Macey and M. Medina, 2005 Sequencing and comparing whole mitochondrial genomes of animals. Methods Enzymol 395**:** 311-348.

Brandao, M. M., and M. C. Silva-Filho, 2011 Evolutionary history of Arabidopsis thaliana aminoacyl-tRNA synthetase dual-targeted proteins. Mol Biol Evol 28**:** 79-85.

Brandvain, Y., and M. J. Wade, 2009 The functional transfer of genes from the mitochondria to the nucleus: the effects of selection, mutation, population size and rate of self-fertilization. Genetics 182**:** 1129-1139.

Brindefalk, B., J. Viklund, D. Larsson, M. Thollesson and S. G. Andersson, 2007 Origin and evolution of the mitochondrial aminoacyl-tRNA synthetases. Mol Biol Evol 24**:** 743-756.

Brown, W. M., M. George, Jr. and A. C. Wilson, 1979 Rapid evolution of animal mitochondrial DNA. Proc Natl Acad Sci U S A 76**:** 1967-1971.

Broz, A. K., G. Waneka, Z. Wu, M. Fernandes Gyorfy and D. B. Sloan, 2021 Detecting de novo mitochondrial mutations in angiosperms with highly divergent evolutionary rates. Genetics 218.

Brubacher-Kauffmann, S., L. Marechal-Drouard, A. Cosset, A. Dietrich and A. M. Duchêne, 1999 Differential import of nuclear-encoded tRNAGly isoacceptors into solanum Tuberosum mitochondria. Nucleic Acids Res 27**:** 2037-2042.

Bruce, B. D., 2001 The paradox of plastid transit peptides: conservation of function despite divergence in primary structure. Biochim Biophys Acta 1541**:** 2-21.

Burger, G., M. W. Gray, L. Forget and B. F. Lang, 2013 Strikingly bacteria-like and gene-rich mitochondrial genomes throughout jakobid protists. Genome Biol Evol 5**:** 418-438.

Bushnell, B., J. Rood and E. Singer, 2017 BBMerge - Accurate paired shotgun read merging via overlap. Plos One 12.

Calvo, S. E., and V. K. Mootha, 2010 The mitochondrial proteome and human disease. Annu Rev Genomics Hum Genet 11**:** 25-44.

Canino, G., E. Bocian, N. Barbezier, M. Echeverria, J. Forner *et al.*, 2009 Arabidopsis encodes four tRNase Z enzymes. Plant Physiol 150**:** 1494-1502.

Carrie, C., and I. Small, 2013 A reevaluation of dual-targeting of proteins to mitochondria and chloroplasts. Biochim Biophys Acta 1833**:** 253-259.

Ceci, L. R., P. Veronico, L. Siculella and R. Gallerani, 1995 Identification and mapping of trnI, trnE and trnfM genes in the sunflower mitochondrial genome. DNA Seq 5**:** 315-318.

Chan, P. P., and T. M. Lowe, 2019 tRNAscan-SE: Searching for tRNA Genes in Genomic Sequences. Methods Mol Biol 1962**:** 1-14.

Chapdelaine, Y., and L. Bonen, 1991 The wheat mitochondrial gene for subunit I of the NADH dehydrogenase complex: a trans-splicing model for this gene-in-pieces. Cell 65**:** 465-472.

Charriere, F., S. Helgadottir, E. K. Horn, D. Söll and A. Schneider, 2006 Dual targeting of a single tRNA(Trp) requires two different tryptophanyl-tRNA synthetases in Trypanosoma brucei. Proc Natl Acad Sci U S A 103**:** 6847-6852.

Charriere, F., P. O'Donoghue, S. Helgadottir, L. Maréchal-Drouard, M. Cristodero *et al.*, 2009 Dual targeting of a tRNAAsp requires two different aspartyl-tRNA synthetases in Trypanosoma brucei. J Biol Chem 284**:** 16210-16217.

Chaw, S. M., A. C. Shih, D. Wang, Y. W. Wu, S. M. Liu *et al.*, 2008 The mitochondrial genome of the gymnosperm Cycas taitungensis contains a novel family of short interspersed elements, Bpu sequences, and abundant RNA editing sites. Mol Biol Evol 25**:** 603-615.

Chen, F., Q. Tang, K. Bian, Z. T. Humulock, X. Yang *et al.*, 2016 Adaptive Response Enzyme AlkB Preferentially Repairs 1-Methylguanine and 3-Methylthymine Adducts in Double-Stranded DNA. Chem Res Toxicol 29**:** 687-693.

Chen, H. C., M. Viry-Moussaid, A. Dietrich and H. Wintz, 1997 Evolution of a mitochondrial tRNA PHE gene in A. thaliana: import of cytosolic tRNA PHE into mitochondria. Biochem Biophys Res Commun 237**:** 432-437.

Chien, C. I., Y. W. Chen, Y. H. Wu, C. Y. Chang, T. L. Wang *et al.*, 2014 Functional substitution of a eukaryotic glycyl-tRNA synthetase with an evolutionarily unrelated bacterial cognate enzyme. PLoS One 9**:** e94659.

Chimnaronk, S., M. Gravers Jeppesen, T. Suzuki, J. Nyborg and K. Watanabe, 2005 Dual-mode recognition of noncanonical tRNAs(Ser) by seryl-tRNA synthetase in mammalian mitochondria. EMBO J 24**:** 3369-3379.

Christensen, A. C., A. Lyznik, S. Mohammed, C. G. Elowsky, A. Elo *et al.*, 2005 Dual-domain, dual-targeting organellar protein presequences in Arabidopsis can use non-AUG start codons. Plant Cell 17**:** 2805-2816.

Christian, R. W., S. L. Hewitt, G. Nelson, E. H. Roalson and A. Dhingra, 2020 Plastid transit peptides-where do they come from and where do they all belong? Multi-genome and pan-genomic assessment of chloroplast transit peptide evolution. PeerJ 8**:** e9772.

Clark, W. C., M. E. Evans, D. Dominissini, G. Zheng and T. Pan, 2016 tRNA base methylation identification and quantification via high-throughput sequencing. RNA 22**:** 1771-1784.

Clary, D. O., and D. R. Wolstenholme, 1984 The Drosophila mitochondrial genome. Oxf Surv Eukaryot Genes 1**:** 1-35.

Coffin, J. W., and E. A. Cossins, 1986 Mitochondrial folates and methionyl-tRNA transformylase activity during germination and early growth of seeds. Phytochemistry 25**:** 2481-2487.

Cognat, V., G. Morelle, C. Megel, S. Lalande, J. Molinier *et al.*, 2017 The nuclear and organellar tRNA-derived RNA fragment population in Arabidopsis thaliana is highly dynamic. Nucleic Acids Res 45**:** 3460-3472.

Cognat, V., G. Pawlak, A. M. Duchene, M. Daujat, A. Gigant *et al.*, 2013 PlantRNA, a database for tRNAs of photosynthetic eukaryotes. Nucleic Acids Res 41**:** D273-279.

Costello, R., D. M. Emms and S. Kelly, 2020 Gene Duplication Accelerates the Pace of Protein Gain and Loss from Plant Organelles. Mol Biol Evol 37**:** 969-981.

Cozen, A. E., E. Quartley, A. D. Holmes, E. Hrabeta-Robinson, E. M. Phizicky *et al.*, 2015 ARM-seq: AlkB-facilitated RNA methylation sequencing reveals a complex landscape of modified tRNA fragments. Nature Methods 12**:** 879-+.

Cummings, M. P., J. M. Nugent, R. G. Olmstead and J. D. Palmer, 2003 Phylogenetic analysis reveals five independent transfers of the chloroplast gene rbcL to the mitochondrial genome in angiosperms. Curr Genet 43**:** 131-138.

Dai, Q., G. Zheng, M. H. Schwartz, W. C. Clark and T. Pan, 2017 Selective Enzymatic Demethylation of N(2) ,N(2) -Dimethylguanosine in RNA and Its Application in High-Throughput tRNA Sequencing. Angew Chem Int Ed Engl 56**:** 5017-5020.

Darty, K., A. Denise and Y. Ponty, 2009 VARNA: Interactive drawing and editing of the RNA secondary structure. Bioinformatics 25**:** 1974-1975.

de Crecy-Lagard, V., P. Boccaletto, C. G. Mangleburg, P. Sharma, T. M. Lowe *et al.*, 2019 Matching tRNA modifications in humans to their known and predicted enzymes. Nucleic Acids Res 47**:** 2143-2159.

Delage, L., A. Dietrich, A. Cosset and L. Marechal-Drouard, 2003 In vitro import of a nuclearly encoded tRNA into mitochondria of Solanum tuberosum. Mol Cell Biol 23**:** 4000-4012.

Dewe, J. M., J. M. Whipple, I. Chernyakov, L. N. Jaramillo and E. M. Phizicky, 2012 The yeast rapid tRNA decay pathway competes with elongation factor 1A for substrate tRNAs and acts on tRNAs lacking one or more of several modifications. RNA 18**:** 1886-1896.

Dietrich, A., J. H. Weil and L. Maréchal-Drouard, 1992 Nuclear-encoded transfer RNAs in plant mitochondria. Annu Rev Cell Biol 8**:** 115-131.

Doolittle, R. F., and J. Handy, 1998 Evolutionary anomalies among the aminoacyl-tRNA synthetases. Curr Opin Genet Dev 8**:** 630-636.

Doyle, J. J., and J. L. Doyle, 1987 A rapid DNA isolation procedure for small quantities of fresh leaf tissue. PHYTOCHEMICAL BULLETIN v.19(1):11-15.

Duchêne, A. M., A. Giritch, B. Hoffmann, V. Cognat, D. Lancelin *et al.*, 2005a Dual targeting is the rule for organellar aminoacyl-tRNA synthetases in Arabidopsis thaliana. Proc Natl Acad Sci U S A 102**:** 16484-16489.

Duchêne, A. M., A. Giritch, B. Hoffmann, V. Cognat, D. Lancelin *et al.*, 2005b Dual targeting is the rule for organellar aminoacyl-tRNA synthetases in Arabidopsis thaliana. Proceedings of the National Academy of Sciences of the United States of America 102**:** 16484-16489.

Duchêne, A. M., and L. Marechal-Drouard, 2001 The chloroplast-derived trnW and trnM-e genes are not expressed in Arabidopsis mitochondria. Biochem Biophys Res Commun 285**:** 1213-1216.

Duchêne, A. M., N. Peeters, A. Dietrich, A. Cosset, I. D. Small *et al.*, 2001 Overlapping destinations for two dual targeted glycyl-tRNA synthetases in Arabidopsis thaliana and Phaseolus vulgaris. Journal of Biological Chemistry 276**:** 15275-15283.

Duchêne, A. M., C. Pujol and L. Marechal-Drouard, 2009 Import of tRNAs and aminoacyl-tRNA synthetases into mitochondria. Curr Genet 55**:** 1-18.

Ebhardt, H. A., H. H. Tsang, D. C. Dai, Y. Liu, B. Bostan *et al.*, 2009 Meta-analysis of small RNA-sequencing errors reveals ubiquitous post-transcriptional RNA modifications. Nucleic Acids Res 37**:** 2461-2470.

Esseiva, A. C., A. Naguleswaran, A. Hemphill and A. Schneider, 2004 Mitochondrial tRNA import in Toxoplasma gondii. Journal of Biological Chemistry 279**:** 42363-42368.

Esser, K., U. Lüttge, W. Beyschlag and J. Murata, 2006 *Progress in Botany 68*. Springer Berlin Heidelberg.

Fey, J., A. Dietrich, A. Cosset, T. Desprez and L. Marechal-Drouard, 1997 Evolutionary aspects of "chloroplast-like" trnN and trnH expression in higher-plant mitochondria. Curr Genet 32**:** 358-360.

Forner, J., B. Weber, S. Thuss, S. Wildum and S. Binder, 2007 Mapping of mitochondrial mRNA termini in Arabidopsis thaliana: t-elements contribute to 5' and 3' end formation. Nucleic Acids Res 35**:** 3676-3692.

Fu, L., B. Niu, Z. Zhu, S. Wu and W. Li, 2012 CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics 28**:** 3150-3152.

Fuchs, R. T., Z. Sun, F. Zhuang and G. B. Robb, 2015 Bias in ligation-based small RNA sequencing library construction is determined by adaptor and RNA structure. PLoS One 10**:** e0126049.

Gandini, C. L., and M. V. Sanchez-Puerta, 2017 Foreign Plastid Sequences in Plant Mitochondria are Frequently Acquired Via Mitochondrion-to-Mitochondrion Horizontal Transfer. Sci Rep 7**:** 43402.

Ge, C., E. Spanning, E. Glaser and A. Wieslander, 2014 Import determinants of organelle-specific and dual targeting peptides of mitochondria and chloroplasts in Arabidopsis thaliana. Mol Plant 7**:** 121-136.

Geslain, R., and T. Pan, 2010 Functional analysis of human tRNA isodecoders. J Mol Biol 396**:** 821-831.

Ghifari, A. S., M. Gill-Hille and M. W. Murcha, 2018 Plant mitochondrial protein import: the ins and outs. Biochem J 475**:** 2191-2208.

Giege, R., M. Sissler and C. Florentz, 1998 Universal rules and idiosyncratic features in tRNA identity. Nucleic Acids Research 26**:** 5017-5035.

Glover, K. E., D. F. Spencer and M. W. Gray, 2001 Identification and structural characterization of nucleus-encoded transfer RNAs imported into wheat mitochondria. J Biol Chem 276**:** 639-648.

Gobert, A., B. Gutmann, A. Taschner, M. Gossringer, J. Holzmann *et al.*, 2010 A single Arabidopsis organellar protein has RNase P activity. Nat Struct Mol Biol 17**:** 740-744.

Gogakos, T., M. Brown, A. Garzia, C. Meyer, M. Hafner *et al.*, 2017 Characterizing Expression and Processing of Precursor and Mature Human tRNAs by Hydro-tRNAseq and PAR-CLIP. Cell Rep 20**:** 1463-1475.

Goodarzi, H., H. C. B. Nguyen, S. Zhang, B. D. Dill, H. Molina *et al.*, 2016 Modulated Expression of Specific tRNAs Drives Gene Expression and Cancer Progression. Cell 165**:** 1416-1427.

Goremykin, V. V., P. J. Lockhart, R. Viola and R. Velasco, 2012 The mitochondrial genome of Malus domestica and the import-driven hypothesis of mitochondrial genome expansion in seed plants. Plant J 71**:** 615-626.

Grabherr, M. G., B. J. Haas, M. Yassour, J. Z. Levin, D. A. Thompson *et al.*, 2011 Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol 29**:** 644-652.

Gray, M. W., 2012 Mitochondrial evolution. Cold Spring Harb Perspect Biol 4**:** a011403.

Greiner, S., P. Lehwark and R. Bock, 2019 OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. Nucleic Acids Res 47**:** W59-W64.

Grewe, F., P. Viehoever, B. Weisshaar and V. Knoop, 2009 A trans-splicing group I intron and tRNA-hyperediting in the mitochondrial genome of the lycophyte Isoetes engelmannii. Nucleic Acids Res 37**:** 5093-5104.

Grosjean, H., and E. Westhof, 2016 An integrated, structure- and energy-based view of the genetic code. Nucleic Acids Res 44**:** 8020-8040.

Guo, M., and P. Schimmel, 2013 Essential nontranslational functions of tRNA synthetases. Nat Chem Biol 9**:** 145-153.

Guo, W., F. Grewe, W. Fan, G. J. Young, V. Knoop *et al.*, 2016 Ginkgo and Welwitschia Mitogenomes Reveal Extreme Contrasts in Gymnosperm Mitochondrial Evolution. Mol Biol Evol 33**:** 1448-1460.

Guo, W., A. Zhu, W. Fan, R. P. Adams and J. P. Mower, 2020 Extensive Shifts from Cis- to Trans-splicing of Gymnosperm Mitochondrial Introns. Mol Biol Evol 37**:** 1615-1620.

Guo, W., A. Zhu, W. Fan and J. P. Mower, 2017 Complete mitochondrial genomes from the ferns Ophioglossum californicum and Psilotum nudum are highly repetitive with the largest organellar introns. New Phytol 213**:** 391-403.

Gutmann, B., A. Gobert and P. Giegé, 2012 PRORP proteins support RNase P activity in both organelles and the nucleus in Arabidopsis. Genes Dev 26**:** 1022-1027.

Haen, K. M., W. Pett and D. V. Lavrov, 2010a Parallel Loss of Nuclear-Encoded Mitochondrial Aminoacyl-tRNA Synthetases and mtDNA-Encoded tRNAs in Cnidaria. Molecular Biology and Evolution 27**:** 2216-2219.

Haen, K. M., W. Pett and D. V. Lavrov, 2010b Parallel loss of nuclear-encoded mitochondrial aminoacyl-tRNA synthetases and mtDNA-encoded tRNAs in Cnidaria. Mol Biol Evol 27**:** 2216-2219.

Hancock, K., and S. L. Hajduk, 1990 The mitochondrial tRNAs of Trypanosoma brucei are nuclear encoded. J Biol Chem 265**:** 19208-19215.

Hao, W., and J. D. Palmer, 2009 Fine-scale mergers of chloroplast and mitochondrial genes create functional, transcompartmentally chimeric mitochondrial genes. Proc Natl Acad Sci U S A 106**:** 16728-16733.

Hauenschild, R., L. Tserovski, K. Schmid, K. Thuring, M. L. Winz *et al.*, 2015 The reverse transcription signature of N-1-methyladenosine in RNA-Seq is sequence dependent. Nucleic Acids Res 43**:** 9950-9964.

Hecht, J., F. Grewe and V. Knoop, 2011 Extreme RNA editing in coding islands and abundant microsatellites in repeat sequences of Selaginella moellendorffii mitochondria: the root of frequent plant mtDNA recombination in early tracheophytes. Genome Biol Evol 3**:** 344-358.

Hendrickson, T. L., 2001 Recognizing the D-loop of transfer RNAs. Proc Natl Acad Sci U S A 98**:** 13473-13475.

Henley, R. Y., S. Carson and M. Wanunu, 2016 Studies of RNA Sequence and Structure Using Nanopores. Prog Mol Biol Transl Sci 139**:** 73-99.

Hoffmann, A., J. Fallmann, E. Vilardo, M. Morl, P. F. Stadler *et al.*, 2018 Accurate mapping of tRNA reads. Bioinformatics 34**:** 2339.

Holley, R. W., J. Apgar, G. A. Everett, J. T. Madison, M. Marquisee *et al.*, 1965 Structure of a Ribonucleic Acid. Science 147**:** 1462-1465.

Holzmann, J., P. Frank, E. Loffler, K. L. Bennett, C. Gerner *et al.*, 2008 RNase P without RNA: identification and functional reconstitution of the human mitochondrial tRNA processing enzyme. Cell 135**:** 462-474.

Hopper, A. K., 2013 Transfer RNA post-transcriptional processing, turnover, and subcellular dynamics in the yeast Saccharomyces cerevisiae. Genetics 194**:** 43-67.

Hopper, A. K., and R. T. Nostramo, 2019 tRNA Processing and Subcellular Trafficking Proteins Multitask in Pathways for Other RNAs. Front Genet 10**:** 96.

Hopper, A. K., and E. M. Phizicky, 2003 tRNA transfers to the limelight. Genes Dev 17**:** 162-180.

Hou, Y. M., H. Gamper and W. Yang, 2015 Post-transcriptional modifications to tRNA--a response to the genetic code degeneracy. RNA 21**:** 642-644.

Huang, C. Y., M. A. Ayliffe and J. N. Timmis, 2003 Direct measurement of the transfer rate of chloroplast DNA into the nucleus. Nature 422**:** 72-76.

Huang, S., N. L. Taylor, J. Whelan and A. H. Millar, 2009 Refining the definition of plant mitochondrial presequences through analysis of sorting signals, N-terminal modifications, and cleavage motifs. Plant Physiol 150**:** 1272-1285.

Huang, S. Q., B. Sun, Z. P. Xiong, Y. Shu, H. H. Zhou *et al.*, 2018 The dysregulation of tRNAs and tRNA derivatives in cancer. J Exp Clin Cancer Res 37**:** 101.

Huber, S. M., A. Leonardi, P. C. Dedon and T. J. Begley, 2019 The Versatile Roles of the tRNA Epitranscriptome during Cellular Responses to Toxic Exposures and Environmental Stress. Toxics 7.

Hummel, G., J. Warren and L. Drouard, 2019 The multi-faceted regulation of nuclear tRNA gene transcription. IUBMB Life 71**:** 1099-1108.

Huynen, M. A., I. Duarte and R. Szklarczyk, 2013 Loss, replacement and gain of proteins at the origin of the mitochondria. Biochimica Et Biophysica Acta-Bioenergetics 1827**:** 224-231.

Iams, K. P., J. E. Heckman and J. H. Sinclair, 1985 Sequence of histidyl tRNA, present as a chloroplast insert in mtDNA ofZea mays. Plant Mol Biol 4**:** 225-232.

Ibba, M., and C. Francklyn, 2004 Turning tRNA upside down: When aminoacylation is not a prerequisite to protein synthesis. Proc Natl Acad Sci U S A 101**:** 7493-7494.

Ibba, M., and D. Söll, 2004 Aminoacyl-tRNAs: setting the limits of the genetic code. Genes Dev 18**:** 731-738.

Iida, K., H. Jin and J. K. Zhu, 2009 Bioinformatics analysis suggests base modifications of tRNAs and miRNAs in Arabidopsis thaliana. BMC Genomics 10**:** 155.

Jackman, J. E., and J. D. Alfonzo, 2013 Transfer RNA modifications: nature's combinatorial chemistry playground. Wiley Interdiscip Rev RNA 4**:** 35-48.

Jackman, S. D., L. Coombe, R. L. Warren, H. Kirk, E. Trinh *et al.*, 2019 Largest Complete Mitochondrial Genome of a Gymnosperm, Sitka Spruce (Picea sitchensis), Indicates Complex Physical Structure. bioRxiv**:** 601104.

Jacob, D., K. Thuring, A. Galliot, V. Marchand, A. Galvanin *et al.*, 2019 Absolute Quantification of Noncoding RNA by Microscale Thermophoresis. Angew Chem Int Ed Engl 58**:** 9565-9569.

Jafari, F., S. Zarre, A. Gholipour, F. Eggens, R. K. Rabeler *et al.*, 2020 A new taxonomic backbone for the infrageneric classification of the species-rich genus Silene (Caryophyllaceae). TAXON 69**:** 337-368.

Jain, M., H. E. Olsen, B. Paten and M. Akeson, 2016 The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. Genome Biol 17**:** 239.

Janouskovec, J., D. V. Tikhonenkov, F. Burki, A. T. Howe, F. L. Rohwer *et al.*, 2017 A New Lineage of Eukaryotes Illuminates Early Mitochondrial Genome Reduction. Curr Biol 27**:** 3717-3724 e3715.

Johnston, I. G., and B. P. Williams, 2016 Evolutionary Inference across Eukaryotes Identifies Specific Pressures Favoring Mitochondrial Gene Retention. Cell Syst 2**:** 101-111.

Jordon-Thaden, I. E., A. S. Chanderbali, M. A. Gitzendanner and D. E. Soltis, 2015 Modified CTAB and TRIzol protocols improve RNA extraction from chemically complex Embryophyta. Appl Plant Sci 3.

Joyce, P. B., and M. W. Gray, 1989 Chloroplast-like transfer RNA genes expressed in wheat mitochondria. Nucleic Acids Res 17**:** 5461-5476.

Juhling, F., M. Morl, R. K. Hartmann, M. Sprinzl, P. F. Stadler *et al.*, 2009 tRNAdb 2009: compilation of tRNA sequences and tRNA genes. Nucleic Acids Res 37**:** D159-162.

Juhling, T., E. Duchardt-Ferner, S. Bonin, J. Wohnert, J. Putz *et al.*, 2018 Small but large enough: structural properties of armless mitochondrial tRNAs from the nematode Romanomermis culicivorax. Nucleic Acids Res 46**:** 9170-9180.

Kamenski, P., O. Kolesnikova, V. Jubenot, N. Entelis, I. A. Krasheninnikov *et al.*, 2007 Evidence for an adaptation mechanism of mitochondrial translation via tRNA import from the cytosol. Mol Cell 26**:** 625-637.

Kamikawa, R., T. Shiratori, K. Ishida, H. Miyashita and A. J. Roger, 2016 Group II Intron-Mediated Trans-Splicing in the Gene-Rich Mitochondrial Genome of an Enigmatic Eukaryote, Diphylleia rotans. Genome Biol Evol 8**:** 458-466.

Kan, S. L., T. T. Shen, P. Gong, J. H. Ran and X. Q. Wang, 2020 The complete mitochondrial genome of Taxus cuspidata (Taxaceae): eight protein-coding genes have transferred to the nuclear genome. BMC Evol Biol 20**:** 10.

Kashdan, M. A., and B. S. Dudock, 1982 The gene for a spinach chloroplast isoleucine tRNA has a methionine anticodon. J Biol Chem 257**:** 11191-11194.

Katoh, K., J. Rozewicki and K. D. Yamada, 2017 MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Brief Bioinform.

Katoh, K., and D. M. Standley, 2013 MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Molecular Biology and Evolution 30**:** 772-780.

Kietrys, A. M., W. A. Velema and E. T. Kool, 2017 Fingerprints of Modified RNA Bases from Deep Sequencing Profiles. J Am Chem Soc 139**:** 17074-17081.

Kimura, S., P. C. Dedon and M. K. Waldor, 2019 Surveying the landscape of tRNA modifications by combining tRNA sequencing and RNA mass spectrometry. bioRxiv**:** 723049.

Kirchner, S., and Z. Ignatova, 2015 Emerging roles of tRNA in adaptive translation, signalling dynamics and disease. Nat Rev Genet 16**:** 98-112.

Kitazaki, K., T. Kubo, H. Kagami, T. Matsumoto, A. Fujita *et al.*, 2011 A horizontally transferred tRNA(Cys) gene in the sugar beet mitochondrial genome: evidence that the gene is present in diverse angiosperms and its transcript is aminoacylated. Plant J 68**:** 262-272.

Knie, N., M. Polsakiewicz and V. Knoop, 2015 Horizontal Gene Transfer of Chlamydial-Like tRNA Genes into Early Vascular Plant Mitochondria. Molecular Biology and Evolution 32**:** 629-634.

Knoop, V., 2004 The mitochondrial DNA of land plants: peculiarities in phylogenetic perspective. Curr Genet 46**:** 123-139.

Knoop, V., W. Schuster, B. Wissinger and A. Brennicke, 1991 Trans splicing integrates an exon of 22 nucleotides into the nad5 mRNA in higher plant mitochondria. EMBO J 10**:** 3483-3493.

Kohn, A. B., M. R. Citarella, K. M. Kocot, Y. V. Bobkova, K. M. Halanych *et al.*, 2012 Rapid evolution of the compact and unusual mitochondrial genome in the ctenophore, Pleurobrachia bachei. Mol Phylogenet Evol 63**:** 203-207.

Kondo, K., B. Makovec, R. H. Waterston and J. Hodgkin, 1990 Genetic and molecular analysis of eight tRNA(Trp) amber suppressors in Caenorhabditis elegans. J Mol Biol 215**:** 7-19.

Kowalak, J. A., S. C. Pomerantz, P. F. Crain and J. A. McCloskey, 1993 A novel method for the determination of post-transcriptional modification in RNA by mass spectrometry. Nucleic Acids Res 21**:** 4577-4585.

Krasovec, M., M. Chester, K. Ridout and D. A. Filatov, 2018 The Mutation Rate and the Age of the Sex Chromosomes in Silene latifolia. Curr Biol 28**:** 1832-1838 e1834.

Kubo, T., S. Nishizawa, A. Sugawara, N. Itchoda, A. Estiati *et al.*, 2000 The complete nucleotide sequence of the mitochondrial genome of sugar beet (Beta vulgaris L.) reveals a novel gene for tRNA(Cys)(GCA). Nucleic Acids Res 28**:** 2571-2576.

Kuchino, Y., N. Hanyu and S. Nishimura, 1987 Analysis of modified nucleosides and nucleotide sequence of tRNA. Methods Enzymol 155**:** 379-396.

Kühn, K., U. Richter, E. H. Meyer, E. Delannoy, A. F. de Longevialle *et al.*, 2009 Phage-type RNA polymerase RPOTmp performs gene-specific transcription in mitochondria of Arabidopsis thaliana. Plant Cell 21**:** 2762-2779.

Kumar, R., L. Marechal-Drouard, K. Akama and I. Small, 1996 Striking differences in mitochondrial tRNA import between different plant species. Mol Gen Genet 252**:** 404-411.

Kunze, M., and J. Berger, 2015 The similarity between N-terminal targeting signals for protein import into different organelles and its evolutionary relevance. Front Physiol 6**:** 259.

Lama, L., J. Cobo, D. Buenaventura and K. Ryan, 2019 Small RNA-seq: The RNA 5'-end adapter ligation problem and how to circumvent it. J Biol Methods 6.

Lang, B. F., M. W. Gray and G. Burger, 1999 Mitochondrial genome evolution and the origin of eukaryotes. Annu Rev Genet 33**:** 351-397.

Lavrov, D. V., and W. Pett, 2016 Animal Mitochondrial DNA as We Do Not Know It: mt-Genome Organization and Evolution in Nonbilaterian Lineages. Genome Biol Evol 8**:** 2896-2913.

Lechner, M., W. Rossmanith, R. K. Hartmann, C. Tholken, B. Gutmann *et al.*, 2015 Distribution of Ribonucleoprotein and Protein-Only RNase P in Eukarya. Mol Biol Evol 32**:** 3186-3193.

Lee, D. W., J. K. Kim, S. Lee, S. Choi, S. Kim *et al.*, 2008 Arabidopsis nuclear-encoded plastid transit peptides contain multiple sequence subgroups with distinctive chloroplast-targeting sequence motifs. Plant Cell 20**:** 1603-1622.

Leon, P., V. Walbot and P. Bedinger, 1989 Molecular analysis of the linear 2.3 kb plasmid of maize mitochondria: apparent capture of tRNA genes. Nucleic Acids Res 17**:** 4089-4099.

Liu, Y., and Y. Chen, 2020 Mitochondrial tRNA Mutations Associated With Essential Hypertension: From Molecular Genetics to Function. Front Cell Dev Biol 8**:** 634137.

Liu, Y., J. Y. Xue, B. Wang, L. Li and Y. L. Qiu, 2011 The mitochondrial genomes of the early land plants Treubia lacunosa and Anomodon rugelii: dynamic and conservative evolution. PLoS One 6**:** e25836.

Lorenz, C., C. E. Lunse and M. Morl, 2017 tRNA Modifications: Impact on Structure and Thermal Adaptation. Biomolecules 7.

Lorenz, R., S. H. Bernhart, C. Honer Zu Siederdissen, H. Tafer, C. Flamm *et al.*, 2011 ViennaRNA Package 2.0. Algorithms Mol Biol 6**:** 26.

Lyons, S. M., M. M. Fay and P. Ivanov, 2018 The role of RNA modifications in the regulation of tRNA cleavage. FEBS Lett 592**:** 2828-2844.

Ma, X., F. Si, X. Liu and W. Luan, 2020 PRMdb: a repository of predicted RNA modifications in plants. Plant and Cell Physiology In press.

Machnicka, M. A., A. Olchowik, H. Grosjean and J. M. Bujnicki, 2014 Distribution and frequencies of post-transcriptional modifications in tRNAs. RNA Biol 11**:** 1619-1629.

Mackenzie, S. A., 2005 Plant organellar protein targeting: a traffic plan still under construction. Trends Cell Biol 15**:** 548-554.

Magallon, S., S. Gomez-Acevedo, L. L. Sanchez-Reyes and T. Hernandez-Hernandez, 2015 A metacalibrated time-tree documents the early rise of flowering plant phylogenetic diversity. New Phytol 207**:** 437-453.

Marchfelder, A., W. Schuster and A. Brennicke, 1990 In vitro processing of mitochondrial and plastid derived tRNA precursors in a plant mitochondrial extract. Nucleic Acids Res 18**:** 1401-1406.

Marechal-Drouard, L., P. Guillemaut, A. Cosset, M. Arbogast, F. Weber *et al.*, 1990 Transfer RNAs of potato (Solanum tuberosum) mitochondria have different genetic origins. Nucleic Acids Res 18**:** 3689-3696.

Marechal-Drouard, L., D. Ramamonjisoa, A. Cosset, J. H. Weil and A. Dietrich, 1993 Editing corrects mispairing in the acceptor stem of bean and potato mitochondrial phenylalanine transfer RNAs. Nucleic Acids Res 21**:** 4909-4914.

Maréchal-Drouard, L., J. Weil and A. Dietrich, 1993 Transfer RNAs and transfer RNA genes in plants. Annual review of plant biology 44**:** 13-32.

Martin, M., 2011 Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet. journal 17**:** 10-12.

Martin, R. P., J. M. Schneller, A. J. Stahl and G. Dirheimer, 1979 Import of nuclear deoxyribonucleic acid coded lysine-accepting transfer ribonucleic acid (anticodon C-U-U) into yeast mitochondria. Biochemistry 18**:** 4600-4605.

Meiklejohn, C. D., M. A. Holmbeck, M. A. Siddiq, D. N. Abt, D. M. Rand *et al.*, 2013 An Incompatibility between a Mitochondrial tRNA and Its Nuclear-Encoded tRNA Synthetase Compromises Development and Fitness in Drosophila. Plos Genetics 9.

Mercer, T. R., S. Neph, M. E. Dinger, J. Crawford, M. A. Smith *et al.*, 2011 The human mitochondrial transcriptome. Cell 146**:** 645-658.

Michaud, M., V. Cognat, A. M. Duchêne and L. Marechal-Drouard, 2011a A global picture of tRNA genes in plant genomes. Plant Journal 66**:** 80-93.

Michaud, M., V. Cognat, A. M. Duchêne and L. Marechal-Drouard, 2011b A global picture of tRNA genes in plant genomes. Plant J 66**:** 80-93.

Mireau, H., D. Lancelin and I. D. Small, 1996 The same Arabidopsis gene encodes both cytosolic and mitochondrial alanyl-tRNA synthetases. Plant Cell 8**:** 1027-1039.

Mishina, Y., and C. He, 2006 Oxidative dealkylation DNA repair mediated by the mononuclear non-heme iron AlkB proteins. J Inorg Biochem 100**:** 670-678.

Miyata, S., M. Nakazono and A. Hirai, 1998 Transcription of plastid-derived tRNA genes in rice mitochondria. Curr Genet 34**:** 216-220.

Motorin, Y., and M. Helm, 2010 tRNA stabilization by modified nucleotides. Biochemistry 49**:** 4934-4944.

Motorin, Y., and M. Helm, 2019 Methods for RNA Modification Mapping Using Deep Sequencing: Established and New Emerging Technologies. Genes 10.

Motorin, Y., S. Muller, I. Behm-Ansmant and C. Branlant, 2007 Identification of modified residues in RNAs by reverse transcription-based methods. Methods Enzymol 425**:** 21-53.

Mower, J. P., P. Touzet, J. S. Gummow, L. F. Delph and J. D. Palmer, 2007 Extensive variation in synonymous substitution rates in mitochondrial genes of seed plants. BMC Evol Biol 7**:** 135.

Mubumbila, M., E. J. Crouse and J. H. Weil, 1984 Transfer RNAs and tRNA genes of Vicia faba chloroplasts. Curr Genet 8**:** 379-385.

Murcha, M. W., B. Kmiec, S. Kubiszewski-Jakubiak, P. F. Teixeira, E. Glaser *et al.*, 2014 Protein import into plant mitochondria: signals, machinery, processing, and regulation. J Exp Bot 65**:** 6301-6335.

Murcha, M. W., S. Kubiszewski-Jakubiak, P. F. Teixeira, I. L. Gugel, B. Kmiec *et al.*, 2016 Plant-Specific Preprotein and Amino Acid Transporter Proteins Are Required for tRNA Import into Mitochondria. Plant Physiol 172**:** 2471-2490.

Nakazono, M., S. Nishiwaki, N. Tsutsumi and A. Hirai, 1996 A chloroplast-derived sequence is utilized as a source of promoter sequences for the gene for subunit 9 of NADH dehydrogenase (nad9) in rice mitochondria. Mol Gen Genet 252**:** 371-378.

Nameki, N., K. Tamura, H. Asahara and T. Hasegawa, 1997 Recognition of tRNA(Gly) by three widely diverged glycyl-tRNA synthetases. J Mol Biol 268**:** 640-647.

Notsu, Y., S. Masood, T. Nishikawa, N. Kubo, G. Akiduki *et al.*, 2002 The complete sequence of the rice (Oryza sativa L.) mitochondrial genome: frequent DNA sequence acquisition and loss during the evolution of flowering plants. Mol Genet Genomics 268**:** 434-445.

Noutahi, E., V. Calderon, M. Blanchette, N. El-Mabrouk and B. F. Lang, 2019 Rapid Genetic Code Evolution in Green Algal Mitochondrial Genomes. Mol Biol Evol 36**:** 766-783.

Novoa, E. M., M. Pavon-Eternod, T. Pan and L. Ribas de Pouplana, 2012 A role for tRNA modifications in genome structure and codon usage. Cell 149**:** 202-213.

Nurk, S., A. Bankevich, D. Antipov, A. A. Gurevich, A. Korobeynikov *et al.*, 2013 Assembling single-cell genomes and mini-metagenomes from chimeric MDA products. J Comput Biol 20**:** 714-737.

O'Donoghue, P., and Z. Luthey-Schulten, 2003 On the evolution of structure in aminoacyl-tRNA synthetases. Microbiol Mol Biol Rev 67**:** 550-573.

Onanuga, K., T. J. Begley, A. A. Chen and S. V. Ranganathan, 2017 tRNA Modification Detection Using Graphene Nanopores: A Simulation Study. Biomolecules 7.

One Thousand Plant Transcriptomes, I., 2019 One thousand plant transcriptomes and the phylogenomics of green plants. Nature 574**:** 679-685.

Ott, M., A. Amunts and A. Brown, 2016 Organization and Regulation of Mitochondrial Protein Synthesis. Annu Rev Biochem 85**:** 77-101.

Pan, T., 2018 Modifications and functional genomics of human transfer RNA. Cell Res 28**:** 395-404.

Pang, Y. L., R. Abo, S. S. Levine and P. C. Dedon, 2014a Diverse cell stresses induce unique patterns of tRNA up- and down-regulation: tRNA-seq for quantifying changes in tRNA copy number. Nucleic Acids Res 42**:** e170.

Pang, Y. L., K. Poruri and S. A. Martinis, 2014b tRNA synthetase: tRNA aminoacylation and beyond. Wiley Interdiscip Rev RNA 5**:** 461-480.

Paris, Z., and J. D. Alfonzo, 2018 How the intracellular partitioning of tRNA and tRNA modification enzymes affects mitochondrial function. IUBMB Life 70**:** 1207-1213.

Park, E. J., and T. H. Kim, 2018 Fine-Tuning of Gene Expression by tRNA-Derived Fragments during Abiotic Stress Signal Transduction. Int J Mol Sci 19.

Park, S., F. Grewe, A. Zhu, T. A. Ruhlman, J. Sabir *et al.*, 2015 Dynamic evolution of Geranium mitochondrial genomes through multiple horizontal and intracellular gene transfers. New Phytol 208**:** 570-583.

Pawar, K., M. Shigematsu, P. Loher, S. Honda, I. Rigoutsos *et al.*, 2019 Exploration of CCA-added RNAs revealed the expression of mitochondrial non-coding RNAs regulated by CCA-adding enzyme. RNA Biol 16**:** 1817-1825.

Pearce, S., C. L. Nezich and A. Spinazzola, 2013 Mitochondrial diseases: translation matters. Mol Cell Neurosci 55**:** 1-12.

Peeters, N., and I. Small, 2001 Dual targeting to mitochondria and chloroplasts. Biochim Biophys Acta 1541**:** 54-63.

Pereira, M., S. Francisco, A. S. Varanda, M. Santos, M. A. S. Santos *et al.*, 2018 Impact of tRNA Modifications and tRNA-Modifying Enzymes on Proteostasis and Human Disease. Int J Mol Sci 19.

Petersen, G., A. Cuenca, I. M. Moller and O. Seberg, 2015 Massive gene loss in mistletoe (Viscum, Viscaceae) mitochondria. Sci Rep 5**:** 17588.

Peterson, K. J., J. B. Lyons, K. S. Nowak, C. M. Takacs, M. J. Wargo *et al.*, 2004 Estimating metazoan divergence times with a molecular clock. Proc Natl Acad Sci U S A 101**:** 6536-6541.

Petrov, A. S., B. Gulen, A. M. Norris, N. A. Kovacs, C. R. Bernier *et al.*, 2015 History of the ribosome and the origin of translation. Proc Natl Acad Sci U S A 112**:** 15396-15401.

Pett, W., and D. V. Lavrov, 2015 Cytonuclear Interactions in the Evolution of Animal Mitochondrial tRNA Metabolism. Genome Biol Evol 7**:** 2089-2101.

Phizicky, E. M., and J. D. Alfonzo, 2010 Do all modifications benefit all tRNAs? FEBS Lett 584**:** 265-271.

Phizicky, E. M., and A. K. Hopper, 2010 tRNA biology charges to the front. Genes Dev 24**:** 1832-1860.

Phizicky, E. M., and A. K. Hopper, 2015 tRNA processing, modification, and subcellular dynamics: past, present, and future. RNA 21**:** 483-485.

Pons, J., P. Bover, L. Bidegaray-Batista and M. A. Arnedo, 2019 Arm-less mitochondrial tRNAs conserved for over 30 millions of years in spiders. BMC Genomics 20**:** 665.

Poodari, V. C., 2019 Direct RNA Sequencing of E. coli initiator tRNA Using the MinION Sequencing Platform, pp. UC Santa Cruz.

Potapov, V., X. Fu, N. Dai, I. R. Correa, Jr., N. A. Tanner *et al.*, 2018 Base modifications affecting RNA polymerase and reverse transcriptase fidelity. Nucleic Acids Res 46**:** 5753-5763.

Pujol, C., M. Bailly, D. Kern, L. Marechal-Drouard, H. Becker *et al.*, 2008 Dual-targeted tRNA-dependent amidotransferase ensures both mitochondrial and chloroplastic Gln-tRNAGln synthesis in plants. Proc Natl Acad Sci U S A 105**:** 6481-6485.

Pujol, C., L. Marechal-Drouard and A. M. Duchêne, 2007 How can organellar protein N-terminal sequences be dual targeting signals? In silico analysis and mutagenesis approach. J Mol Biol 369**:** 356-367.

Qiu, Y. L., and J. D. Palmer, 2004 Many independent origins of trans splicing of a plant mitochondrial group II intron. J Mol Evol 59**:** 80-89.

Quinlan, A. R., and I. M. Hall, 2010 BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26**:** 841-842.

Raina, M., and M. Ibba, 2014 tRNAs as regulators of biological processes. Front Genet 5**:** 171.

Reinbothe, S., C. Rossig, J. Gray, S. Rustgi, D. von Wettstein *et al.*, 2021 tRNA-Dependent Import of a Transit Sequence-Less Aminoacyl-tRNA Synthetase (LeuRS2) into the Mitochondria of Arabidopsis. Int J Mol Sci 22.

Rice, D. W., A. J. Alverson, A. O. Richardson, G. J. Young, M. V. Sanchez-Puerta *et al.*, 2013 Horizontal transfer of entire genomes via mitochondrial fusion in the angiosperm Amborella. Science 342**:** 1468-1473.

Rice, P., I. Longden and A. Bleasby, 2000 EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet 16**:** 276-277.

Richardson, A. O., D. W. Rice, G. J. Young, A. J. Alverson and J. D. Palmer, 2013 The "fossilized" mitochondrial genome of Liriodendron tulipifera: ancestral gene content and order, ancestral editing sites, and extraordinarily low mutation rate. BMC Biol 11**:** 29.

Rinehart, J., B. Krett, M. A. Rubio, J. D. Alfonzo and D. Söll, 2005 Saccharomyces cerevisiae imports the cytosolic pathway for Gln-tRNA synthesis into the mitochondrion. Genes Dev 19**:** 583-592.

Robinson, M. D., D. J. McCarthy and G. K. Smyth, 2010 edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26**:** 139-140.

Rodnina, M. V., K. B. Gromadski, U. Kothe and H. J. Wieden, 2005 Recognition and selection of tRNA in translation. FEBS Lett 579**:** 938-942.

Roger, A. J., S. A. Munoz-Gomez and R. Kamikawa, 2017 The Origin and Diversification of Mitochondria. Curr Biol 27**:** R1177-R1192.

Rogers, H. H., and S. Griffiths-Jones, 2014 tRNA anticodon shifts in eukaryotic genomes. RNA 20**:** 269-281.

Rojas-Benitez, D., P. C. Thiaville, V. de Crecy-Lagard and A. Glavic, 2015 The Levels of a Universally Conserved tRNA Modification Regulate Cell Growth. J Biol Chem 290**:** 18699-18707.

Ross, R., X. Cao, N. Yu and P. A. Limbach, 2016 Sequence mapping of transfer RNA chemical modifications by liquid chromatography tandem mass spectrometry. Methods 107**:** 73-78.

Rossmanith, W., 2012 Of P and Z: mitochondrial tRNA processing enzymes. Biochim Biophys Acta 1819**:** 1017-1026.

Rubio Gomez, M. A., and M. Ibba, 2020 Aminoacyl-tRNA synthetases. RNA 26**:** 910-936.

Rubio, M. A., J. J. Rinehart, B. Krett, S. Duvezin-Caubet, A. S. Reichert *et al.*, 2008 Mammalian mitochondria have the innate ability to import tRNAs by a mechanism distinct from protein import. Proc Natl Acad Sci U S A 105**:** 9186-9191.

Salazar, J. C., I. Ahel, O. Orellana, D. Tumbula-Hansen, R. Krieger *et al.*, 2003 Coevolution of an aminoacyl-tRNA synthetase with its tRNA substrates. Proc Natl Acad Sci U S A 100**:** 13863-13868.

Salinas, T., F. Duby, V. Larosa, N. Coosemans, N. Bonnefoy *et al.*, 2012 Co-evolution of mitochondrial tRNA import and codon usage determines translational efficiency in the green alga Chlamydomonas. PLoS Genet 8**:** e1002946.

Salinas, T., A. M. Duchêne and L. Marechal-Drouard, 2008 Recent advances in tRNA mitochondrial import. Trends Biochem Sci 33**:** 320-329.

Salinas, T., C. Schaeffer, L. Marechal-Drouard and A. M. Duchêne, 2005 Sequence dependence of tRNA(Gly) import into tobacco mitochondria. Biochimie 87**:** 863-872.

Salinas-Giege, T., R. Giege and P. Giege, 2015 tRNA Biology in Mitochondria. International Journal of Molecular Sciences 16**:** 4518-4559.

Salinas-Giegé, T., R. Giegé and P. Giegé, 2015 tRNA biology in mitochondria. Int J Mol Sci 16**:** 4518-4559.

Sampson, J. R., and O. C. Uhlenbeck, 1988 Biochemical and physical characterization of an unmodified yeast phenylalanine transfer RNA transcribed in vitro. Proc Natl Acad Sci U S A 85**:** 1033-1037.

Sanchez-Puerta, M. V., A. Edera, C. L. Gandini, A. V. Williams, K. A. Howell *et al.*, 2019 Genome-scale transfer of mitochondrial DNA from legume hosts to the holoparasite Lophophytum mirabile (Balanophoraceae). Mol Phylogenet Evol 132**:** 243-250.

Schmidt, O., N. Pfanner and C. Meisinger, 2010 Mitochondrial protein import: from proteomics to functional mechanisms. Nat Rev Mol Cell Biol 11**:** 655-667.

Schneider, A., 2001 Unique aspects of mitochondrial biogenesis in trypanosomatids. International journal for parasitology 31**:** 1403-1415.

Schneider, A., 2011 Mitochondrial tRNA Import and Its Consequences for Mitochondrial Translation. Annual Review of Biochemistry, Vol 80 80**:** 1033-1053.

Schwartz, S., and Y. Motorin, 2017 Next-generation sequencing technologies for detection of modified nucleotides in RNAs. RNA Biol 14**:** 1124-1137.

Selden, R. F., A. Steinmetz, L. McIntosh, L. Bogorad, G. Burkard *et al.*, 1983 Transfer RNA genes ofZea mays chloroplast DNA. Plant Mol Biol 2**:** 141-153.

Sequeira-Mendes, J., I. Araguez, R. Peiro, R. Mendez-Giraldez, X. Zhang *et al.*, 2014 The Functional Topography of the Arabidopsis Genome Is Organized in a Reduced Number of Linear Motifs of Chromatin States. Plant Cell 26**:** 2351-2366.

Shigematsu, M., S. Honda, P. Loher, A. G. Telonis, I. Rigoutsos *et al.*, 2017 YAMAT-seq: an efficient method for high-throughput sequencing of mature transfer RNAs. Nucleic Acids Research 45.

Simpson, A. M., Y. Suyama, H. Dewes, D. A. Campbell and L. Simpson, 1989 Kinetoplastid Mitochondria Contain Functional Transfer-Rnas Which Are Encoded in Nuclear-DNA and Also Contain Small Minicircle and Maxicircle Transcripts of Unknown Function. Nucleic Acids Research 17**:** 5427-5445.

Sinn, B. T., and C. F. Barrett, 2019 Ancient mitochondrial gene transfer between fungi and the orchids. Mol Biol Evol.

Sissler, M., J. Putz, F. Fasiolo and C. Florentz, 2005 Mitochondrial Aminoacyl-tRNA Synthetases, pp. in *Madame Curie Bioscience Database [Internet]*. Landes Bioscience, Austin (TX).

Skippington, E., T. J. Barkman, D. W. Rice and J. D. Palmer, 2015 Miniaturized mitogenome of the parasitic plant Viscum scurruloideum is extremely divergent and dynamic and has lost all nad genes. Proc Natl Acad Sci U S A 112**:** E3515-3524.

Sloan, D. B., A. J. Alverson, J. P. Chuckalovcak, M. Wu, D. E. McCauley *et al.*, 2012a Rapid Evolution of Enormous, Multichromosomal Genomes in Flowering Plant Mitochondria with Exceptionally High Mutation Rates. Plos Biology 10.

Sloan, D. B., A. J. Alverson, H. Storchova, J. D. Palmer and D. R. Taylor, 2010a Extensive loss of translational genes in the structurally dynamic mitochondrial genome of the angiosperm Silene latifolia. Bmc Evolutionary Biology 10.

Sloan, D. B., A. J. Alverson, H. Storchova, J. D. Palmer and D. R. Taylor, 2010b Extensive loss of translational genes in the structurally dynamic mitochondrial genome of the angiosperm Silene latifolia. BMC Evol Biol 10**:** 274.

Sloan, D. B., A. J. Alverson, M. Wu, J. D. Palmer and D. R. Taylor, 2012b Recent acceleration of plastid sequence and structural evolution coincides with extreme mitochondrial divergence in the angiosperm genus Silene. Genome Biol Evol 4**:** 294-306.

Sloan, D. B., K. Muller, D. E. McCauley, D. R. Taylor and H. Storchova, 2012c Intraspecific variation in mitochondrial genome sequence, structure, and gene content in Silene vulgaris, an angiosperm with pervasive cytoplasmic male sterility. New Phytol 196**:** 1228-1239.

Sloan, D. B., B. Oxelman, A. Rautenberg and D. R. Taylor, 2009 Phylogenetic analysis of mitochondrial substitution rate variation in the angiosperm tribe Sileneae. BMC Evol Biol 9**:** 260.

Sloan, D. B., D. A. Triant, N. J. Forrester, L. M. Bergner, M. Wu *et al.*, 2014a A recurring syndrome of accelerated plastid genome evolution in the angiosperm tribe Sileneae (Caryophyllaceae). Mol Phylogenet Evol 72**:** 82-89.

Sloan, D. B., D. A. Triant, M. Wu and D. R. Taylor, 2014b Cytonuclear interactions and relaxed selection accelerate sequence evolution in organelle ribosomes. Mol Biol Evol 31**:** 673-682.

Sloan, D. B., J. M. Warren, A. M. Williams, Z. Wu, S. E. Abdel-Ghany *et al.*, 2018 Cytonuclear integration and co-evolution. Nat Rev Genet 19**:** 635-648.

Sloan, D. B., and Z. Wu, 2014 History of plastid DNA insertions reveals weak deletion and at mutation biases in angiosperm mitochondrial genomes. Genome Biol Evol 6**:** 3210-3221.

Small, I., K. Akashi, A. Chapron, A. Dietrich, A. M. Duchêne *et al.*, 1999 The strange evolutionary history of plant mitochondrial tRNAs and their aminoacyl-tRNA synthetases. Journal of Heredity 90**:** 333-337.

Small, I., N. Peeters, F. Legeai and C. Lurin, 2004 Predotar: A tool for rapidly screening proteomes for N-terminal targeting sequences. Proteomics 4**:** 1581-1590.

Smith, A. M., R. Abu-Shumays, M. Akeson and D. L. Bernick, 2015 Capture, Unfolding, and Detection of Individual tRNA Molecules Using a Nanopore Device. Front Bioeng Biotechnol 3**:** 91.

Soprano, A. S., J. H. C. Smetana and C. E. Benedetti, 2018 Regulation of tRNA biogenesis in plants and its link to plant growth and response to pathogens. Biochim Biophys Acta Gene Regul Mech 1861**:** 344-353.

Sperschneider, J., A. M. Catanzariti, K. DeBoer, B. Petre, D. M. Gardiner *et al.*, 2017 LOCALIZER: subcellular localization prediction of both plant and effector proteins in the plant cell. Sci Rep 7**:** 44598.

Stamatakis, A., 2014 RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30**:** 1312-1313.

Stupar, R. M., J. W. Lilly, C. D. Town, Z. Cheng, S. Kaul *et al.*, 2001 Complex mtDNA constitutes an approximate 620-kb insertion on Arabidopsis thaliana chromosome 2: implication of potential sequencing errors caused by large-unit repeats. Proc Natl Acad Sci U S A 98**:** 5099-5103.

Su, D., C. T. Chan, C. Gu, K. S. Lim, Y. H. Chionh *et al.*, 2014 Quantitative analysis of ribonucleoside modifications in tRNA by HPLC-coupled mass spectrometry. Nat Protoc 9**:** 828-841.

Suzuki, T., and K. Miyauchi, 2010 Discovery and characterization of tRNAIle lysidine synthetase (TilS). FEBS Lett 584**:** 272-277.

Takeuchi, N., L. Vial, M. Panvert, E. Schmitt, K. Watanabe *et al.*, 2001 Recognition of tRNAs by Methionyl-tRNA transformylase from mammalian mitochondria. J Biol Chem 276**:** 20064-20068.

Tan, T. H., R. Pach, A. Crausaz, A. Ivens and A. Schneider, 2002 tRNAs in Trypanosoma brucei: genomic organization, expression, and mitochondrial import. Mol Cell Biol 22**:** 3707-3717.

Taylor, Z. N., D. W. Rice and J. D. Palmer, 2015 The Complete Moss Mitochondrial Genome in the Angiosperm Amborella Is a Chimera Derived from Two Moss Whole-Genome Transfers. PLoS One 10**:** e0137532.

Telonis, A. G., Y. Kirino and I. Rigoutsos, 2015 Mitochondrial tRNA-lookalikes in nuclear chromosomes: could they be functional? RNA Biol 12**:** 375-380.

Telonis, A. G., P. Loher, Y. Kirino and I. Rigoutsos, 2014 Nuclear and mitochondrial tRNA-lookalikes in the human genome. Front Genet 5**:** 344.

Telonis, A. G., P. Loher, R. Magee, V. Pliatsika, E. Londin *et al.*, 2019 tRNA Fragments Show Intertwining with mRNAs of Specific Repeat Content and Have Links to Disparities. Cancer Res 79**:** 3034-3049.

Theologis, A., J. R. Ecker, C. J. Palm, N. A. Federspiel, S. Kaul *et al.*, 2000 Sequence and analysis of chromosome 1 of the plant Arabidopsis thaliana. Nature 408**:** 816-820.

Timmis, J. N., M. A. Ayliffe, C. Y. Huang and W. Martin, 2004 Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. Nat Rev Genet 5**:** 123-135.

Tolkunova, E., H. Park, J. Xia, M. P. King and E. Davidson, 2000 The human lysyl-tRNA synthetase gene encodes both the cytoplasmic and mitochondrial enzymes by means of an unusual alternative splicing of the primary transcript. J Biol Chem 275**:** 35063-35069.

Torres, A. G., D. Pineyro, M. Rodriguez-Escriba, N. Camacho, O. Reina *et al.*, 2015 Inosine modifications in human tRNAs are incorporated at the precursor tRNA level. Nucleic Acids Res 43**:** 5145-5157.

Torres, A. G., O. Reina, C. Stephan-Otto Attolini and L. Ribas de Pouplana, 2019 Differential expression of human tRNA genes drives the abundance of tRNA-derived fragments. Proc Natl Acad Sci U S A 116**:** 8451-8456.

Trewick, S. C., T. F. Henshaw, R. P. Hausinger, T. Lindahl and B. Sedgwick, 2002 Oxidative demethylation by Escherichia coli AlkB directly reverts DNA base damage. Nature 419**:** 174-178.

Tserovski, L., V. Marchand, R. Hauenschild, F. Blanloeil-Oillo, M. Helm *et al.*, 2016 High-throughput sequencing for 1-methyladenosine (m(1)A) mapping in RNA. Methods 107**:** 110-121.

Turmel, M., C. Otis and C. Lemieux, 2013 Tracing the evolution of streptophyte algae and their mitochondrial genome. Genome Biol Evol 5**:** 1817-1835.

Urbonavicius, J., Q. Qian, J. M. Durand, T. G. Hagervall and G. R. Bjork, 2001 Improvement of reading frame maintenance is a common function for several tRNA modifications. EMBO J 20**:** 4863-4873.

Valach, M., G. Burger, M. W. Gray and B. F. Lang, 2014 Widespread occurrence of organelle genome-encoded 5S rRNAs including permuted molecules. Nucleic Acids Res 42**:** 13764-13777.

Vandivier, L. E., Z. D. Anderson and B. D. Gregory, 2019 HAMR: High-Throughput Annotation of Modified Ribonucleotides. Methods Mol Biol 1870**:** 51-67.

Varre, J. S., N. D'Agostino, P. Touzet, S. Gallina, R. Tamburino *et al.*, 2019 Complete Sequence, Multichromosomal Architecture and Transcriptome Analysis of the Solanum tuberosum Mitochondrial Genome. Int J Mol Sci 20.

Vermeulen, A., S. A. McCallum and A. Pardi, 2005 Comparison of the global structure and dynamics of native and unmodified tRNAval. Biochemistry 44**:** 6024-6033.

Vinogradova, E., T. Salinas, V. Cognat, C. Remacle and L. Marechal-Drouard, 2009 Steady-state levels of imported tRNAs in Chlamydomonas mitochondria are correlated with both cytosolic and mitochondrial codon usages. Nucleic Acids Res 37**:** 1521-1528.

von Braun, S. S., A. Sabetti, P. J. Hanic-Joyce, J. Gu, E. Schleiff *et al.*, 2007 Dual targeting of the tRNA nucleotidyltransferase in plants: not just the signal. J Exp Bot 58**:** 4083-4093.

Wang, D., M. Rousseau-Gueutin and J. N. Timmis, 2012 Plastid sequences contribute to some plant mitochondrial genes. Mol Biol Evol 29**:** 1707-1711.

Wang, D., Y. W. Wu, A. C. Shih, C. S. Wu, Y. N. Wang *et al.*, 2007 Transfer of chloroplast genomic DNA to mitochondrial genome occurred at least 300 MYA. Mol Biol Evol 24**:** 2040-2048.

Wang, X. C., H. Chen, D. Yang and C. Liu, 2018 Diversity of mitochondrial plastid DNAs (MTPTs) in seed plants. Mitochondrial DNA A DNA Mapp Seq Anal 29**:** 635-642.

Wang, Z., M. Gerstein and M. Snyder, 2009 RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 10**:** 57-63.

Warren, J. M., T. Salinas-Giege, G. Hummel, N. L. Coots, J. M. Svendsen *et al.*, 2020 Combining tRNA sequencing methods to characterize plant tRNA expression and post-transcriptional modification. RNA Biol**:** 1-15.

Warren, J. M., T. Salinas-Giegé, G. Hummel, N. L. Coots, J. M. Svendsen *et al.*, 2021 Combining tRNA sequencing methods to characterize plant tRNA expression and post-transcriptional modification. RNA Biology 18**:** 64-78.

Warren, J. M., T. Salinas-Giegé, G. Hummel, N. L. Coots, J. M. Svendsen *et al.*, 2019 Combining tRNA sequencing methods to characterize plant tRNA expression and post-transcriptional modification. bioRxiv**:** 790451.

Warren, J. M., and D. B. Sloan, 2020a Hopeful monsters: Unintended sequencing of famously malformed mite mitochondrial tRNAs reveals widespread expression and processing of sense-antisense pairs. bioRxiv**:** 2020.2009.2008.286963.

Warren, J. M., and D. B. Sloan, 2020b Interchangeable parts: The evolutionarily dynamic tRNA population in plant mitochondria. Mitochondrion 52**:** 144-156.

Watanabe, Y., T. Suematsu and T. Ohtsuki, 2014 Losing the stem-loop structure from metazoan mitochondrial tRNAs and co-evolution of interacting factors. Front Genet 5**:** 109.

Weber, F., A. Dietrich, J. H. Weil and L. Maréchal-Drouard, 1990 A potato mitochondrial isoleucine tRNA is coded for by a mitochondrial gene possessing a methionine anticodon. Nucleic Acids Res 18**:** 5027-5030.

Werner, S., L. Schmidt, V. Marchand, T. Kemmer, C. Falschlunger *et al.*, 2020 Machine learning of reverse transcription signatures of variegated polymerases allows mapping and discrimination of methylated purines in limited transcriptomes. Nucleic Acids Res 48**:** 3734-3746.

Williams, A. M., M. W. Itgen, A. K. Broz, O. G. Carter and D. B. Sloan, 2020 Long-read transcriptome and other genomic resources for the angiosperm Silene noctiflora. bioRxiv**:** 2020.2008.2009.243378.

Williams, E. J., C. Pal and L. D. Hurst, 2000a The molecular evolution of signal peptides. Gene 253**:** 313-322.

Williams, M. A., Y. Johzuka and R. M. Mulligan, 2000b Addition of non-genomically encoded nucleotides to the 3'-terminus of maize mitochondrial mRNAs: truncated rps12 mRNAs frequently terminate with CCA. Nucleic Acids Res 28**:** 4444-4451.

Wilusz, J. E., 2015a Controlling translation via modulation of tRNA levels. Wiley Interdiscip Rev RNA 6**:** 453-470.

Wilusz, J. E., 2015b Removing roadblocks to deep sequencing of modified RNAs. Nat Methods 12**:** 821-822.

Wissinger, B., W. Schuster and A. Brennicke, 1991 Trans splicing in Oenothera mitochondria: nad1 mRNAs are edited in exon and trans-splicing group II intron sequences. Cell 65**:** 473-482.

Wu, Z., J. M. Cuthbert, D. R. Taylor and D. B. Sloan, 2015 The massive mitochondrial genome of the angiosperm Silene noctiflora is evolving by gain or loss of entire chromosomes. Proc Natl Acad Sci U S A 112**:** 10185-10191.

Xing, F., S. L. Hiley, T. R. Hughes and E. M. Phizicky, 2004 The specificities of four yeast dihydrouridine synthases for cytoplasmic tRNAs. J Biol Chem 279**:** 17850-17860.

Xu, L., C. Carrie, S. R. Law, M. W. Murcha and J. Whelan, 2013 Acquisition, conservation, and loss of dual-targeted proteins in land plants. Plant Physiol 161**:** 644-662.

Xue, J. Y., Y. Liu, L. Li, B. Wang and Y. L. Qiu, 2010 The complete mitochondrial genome sequence of the hornwort Phaeoceros laevis: retention of many ancient pseudogenes and conservative evolution of mitochondrial genomes in hornworts. Curr Genet 56**:** 53-61.

Yarham, J. W., J. L. Elson, E. L. Blakely, R. McFarland and R. W. Taylor, 2010 Mitochondrial tRNA mutations and disease. Wiley Interdiscip Rev RNA 1**:** 304-324.

Yogev, O., and O. Pines, 2011 Dual targeting of mitochondrial proteins: mechanism, regulation and function. Biochim Biophys Acta 1808**:** 1012-1020.

Yona, A. H., Z. Bloom-Ackermann, I. Frumkin, V. Hanson-Smith, Y. Charpak-Amikam *et al.*, 2013 tRNA genes rapidly change in evolution to meet novel translational demands. Elife 2**:** e01339.

Youvan, D. C., and J. E. Hearst, 1979 Reverse transcriptase pauses at N2-methylguanine during in vitro transcription of Escherichia coli 16S ribosomal RNA. Proc Natl Acad Sci U S A 76**:** 3751-3754.

Yu, C. H., J. Y. Liao, H. Zhou and L. H. Qu, 2008 The rat mitochondrial Ori L encodes a novel small RNA resembling an ancestral tRNA. Biochem Biophys Res Commun 372**:** 634-638.

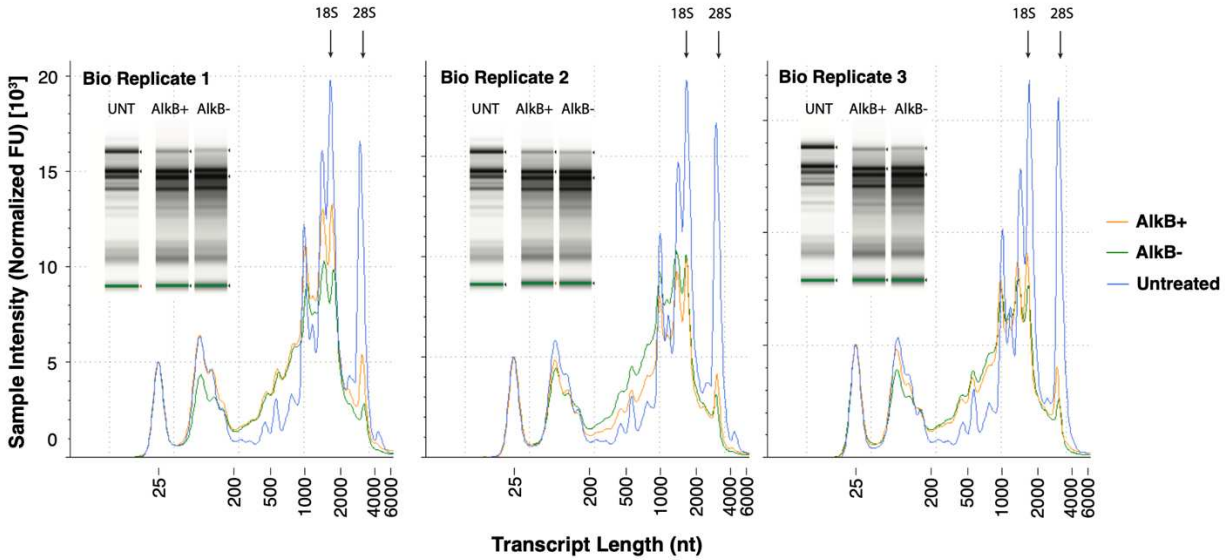
Zhang, J., and A. R. Ferre-D'Amare, 2016 The tRNA Elbow in Structure, Recognition and Evolution. Life (Basel) 6.

Zhao, L., H. Zhang, M. V. Kohnen, K. Prasad, L. Gu *et al.*, 2019 Analysis of Transcriptome and Epitranscriptome in Plants Using PacBio Iso-Seq and Nanopore-Based Direct RNA Sequencing. Front Genet 10**:** 253.

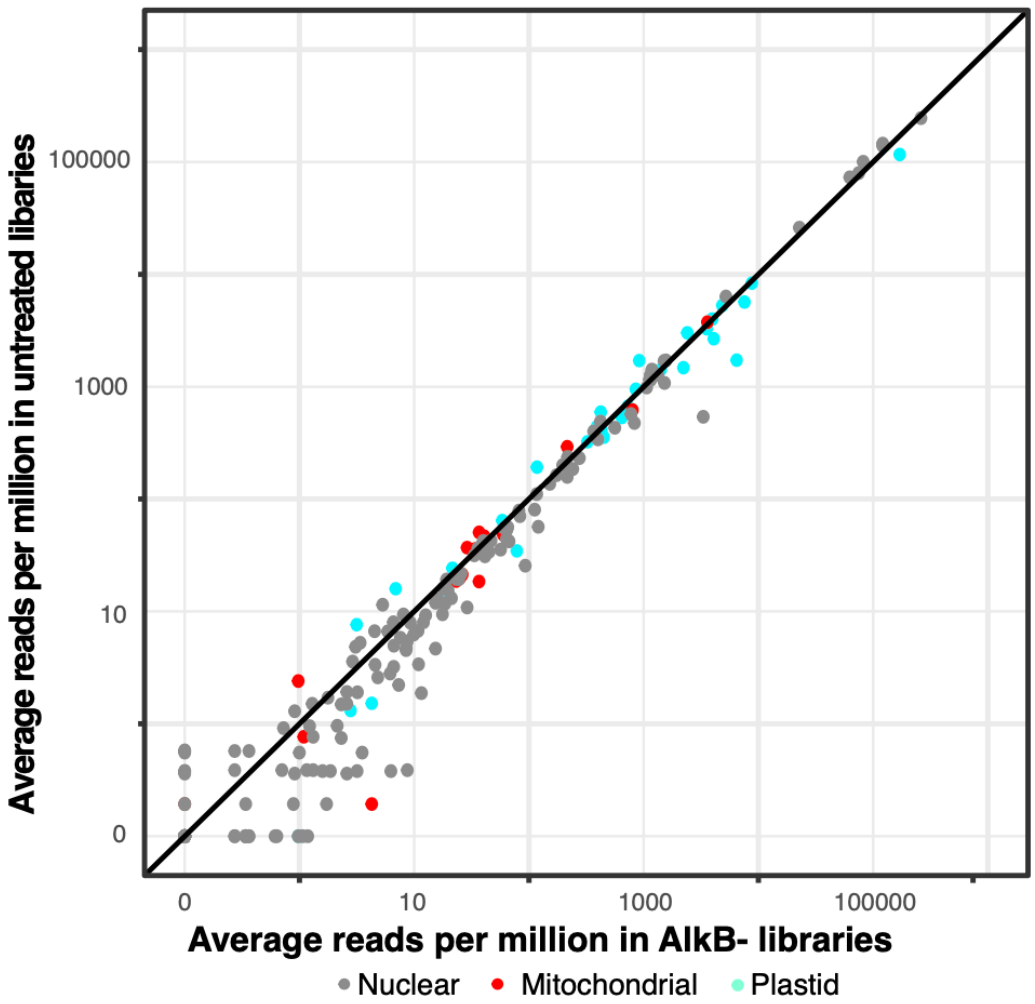
Zheng, G., Y. Qin, W. C. Clark, Q. Dai, C. Yi *et al.*, 2015 Efficient and quantitative high-throughput tRNA sequencing. Nat Methods 12**:** 835-837.

Zhong, J., C. Xiao, W. Gu, G. Du, X. Sun *et al.*, 2015 Transfer RNAs Mediate the Rapid Adaptation of Escherichia coli to Oxidative Stress. PLoS Genet 11**:** e1005302.

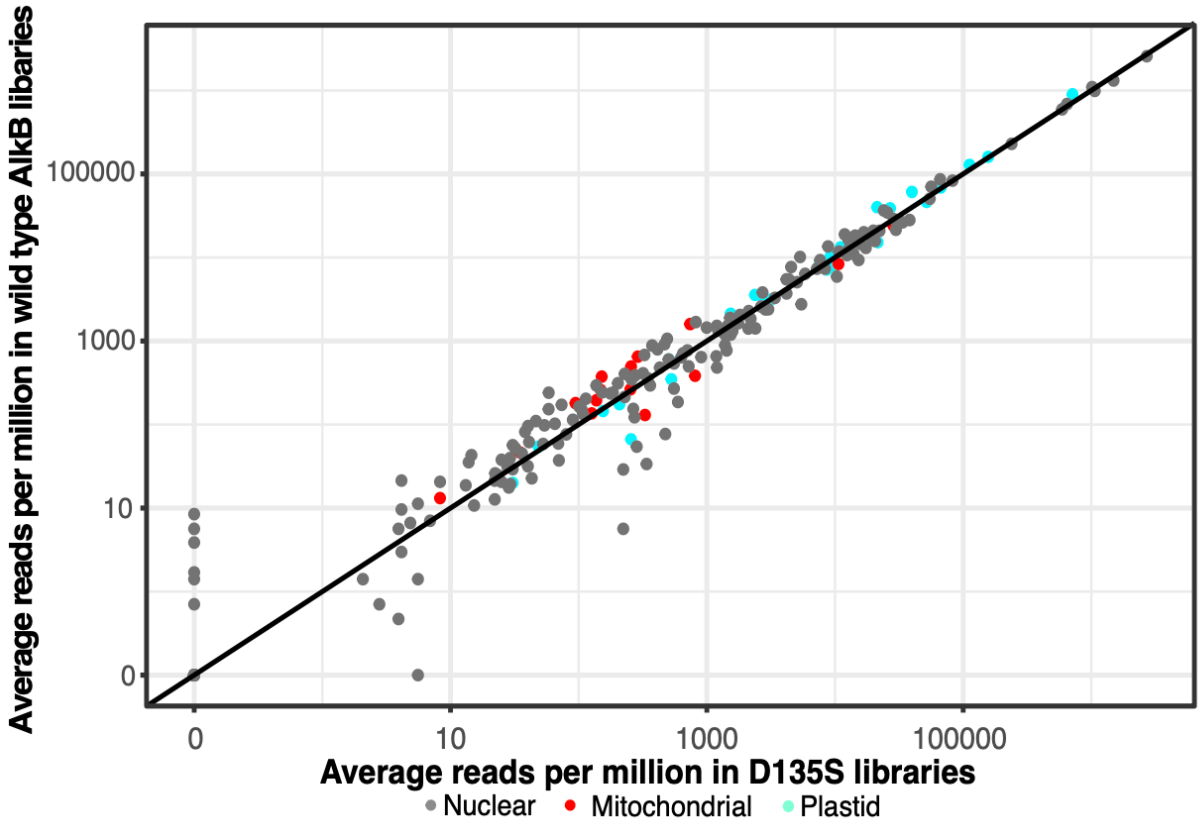
APPENDIX: SUPPLEMENTARY FIGURES



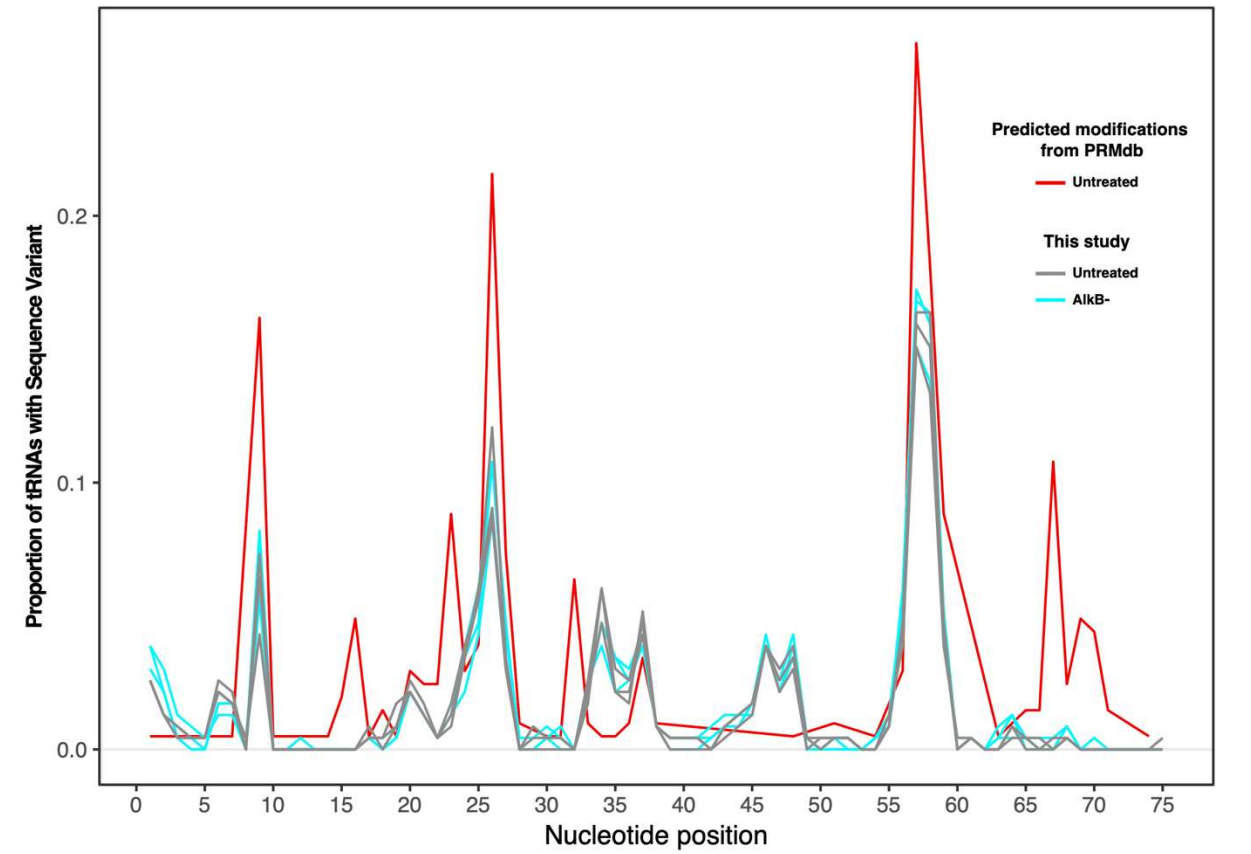
Supplementary Figure 2.1| Size distribution of RNA samples used in this study. Samples were analyzed using TapeStation 2200 and show the relative abundance on the y-axis and the transcript length on the x-axis. Treatment is indicated by color.



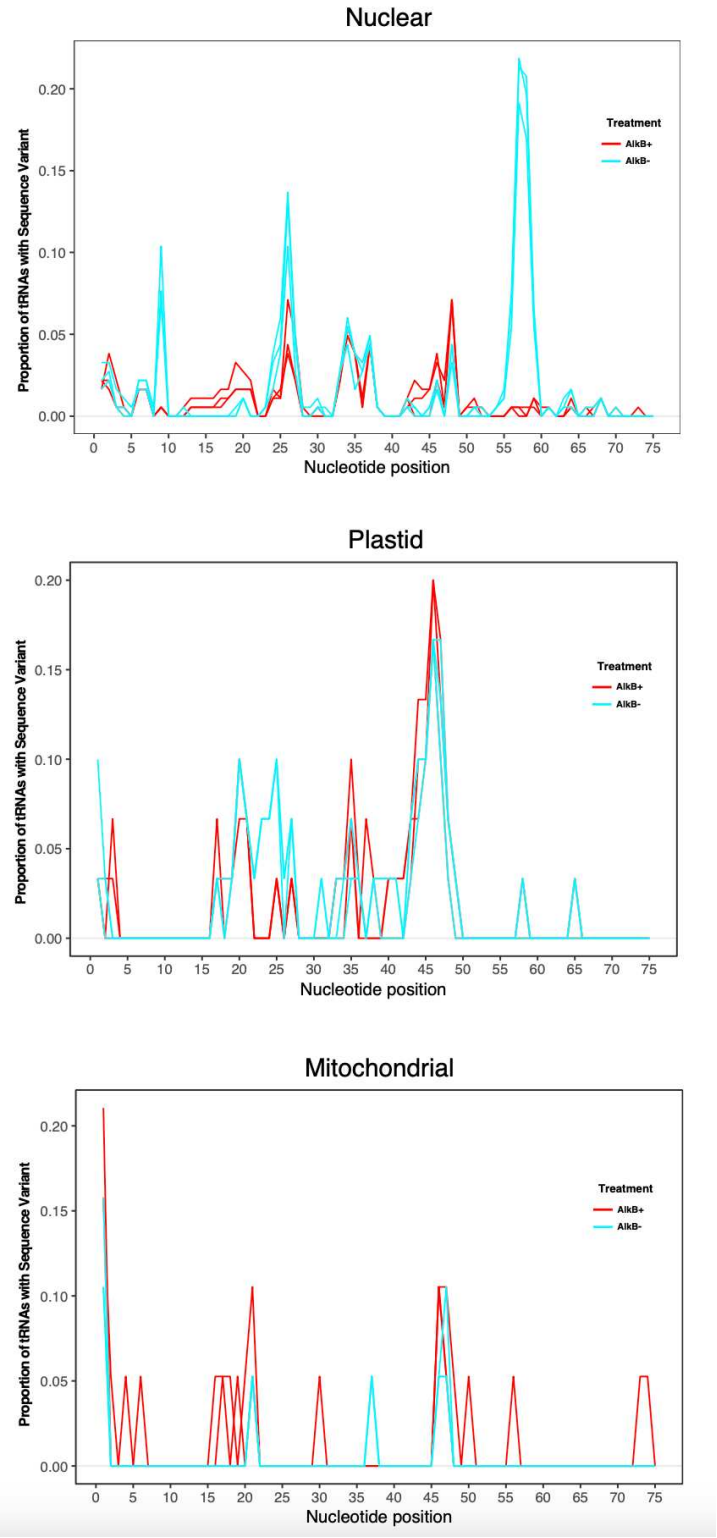
Supplementary Figure 2.2| Average read counts per million across three biological replicates are shown for each unique *A. thaliana* tRNA reference sequence for untreated libraries and AlkB- (buffer only) libraries.



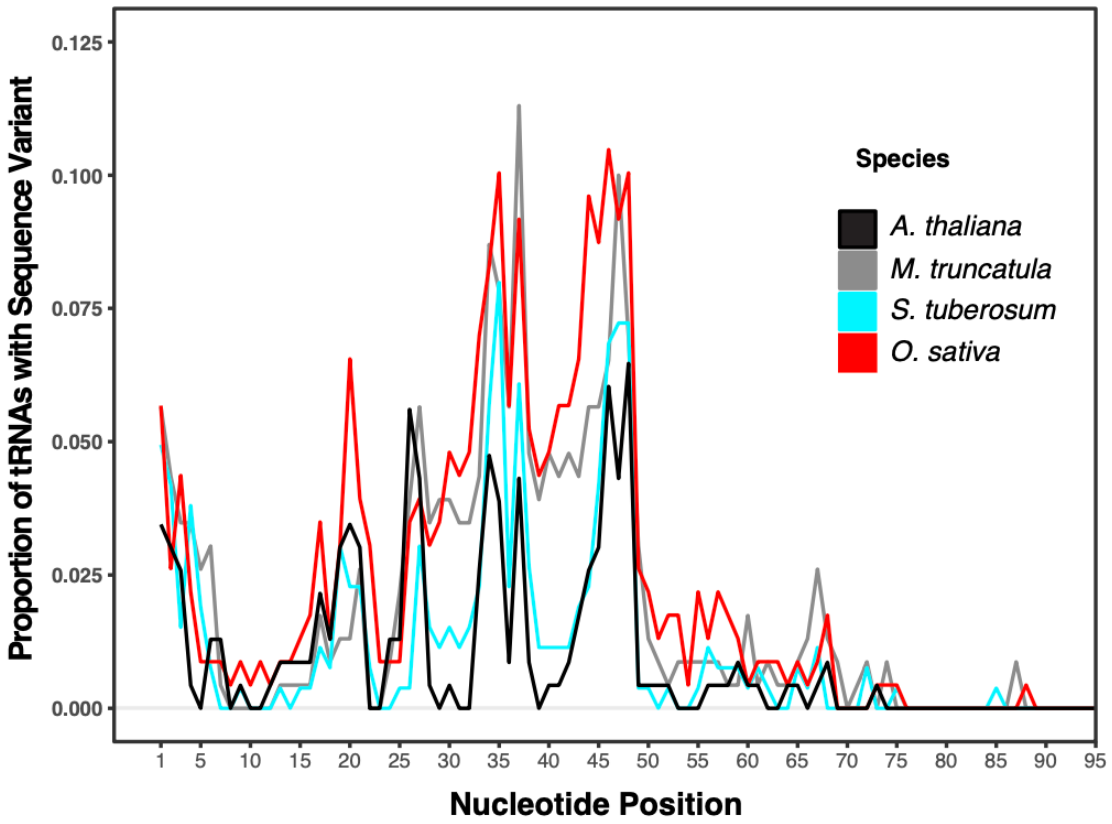
Supplementary Figure 2.3| Average read counts per million across three biological replicates are shown for each unique *A. thaliana* tRNA reference sequence for wild type AlkB and D135S AlkB libraries.



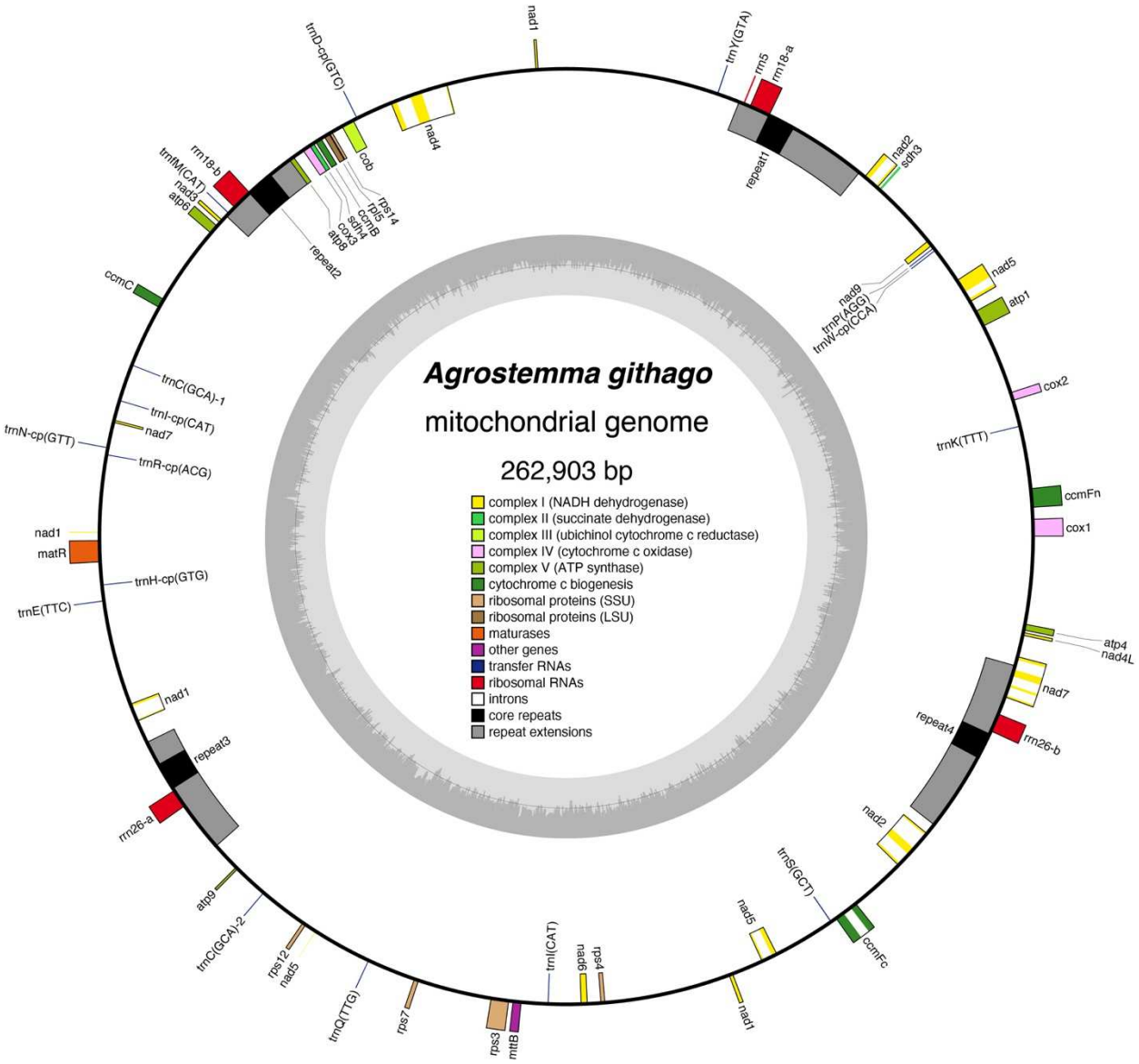
Supplementary Figure 2.4| Comparison of modification indexes from this analysis and the modifications reported in PRMdb: a repository of predicted RNA modifications in plants, Ma et al., 2020. Modification indexes show the proportion of all *A. thaliana* tRNA reference sequences with a misincorporation/deletion at each nucleotide position. Gray lines represent untreated libraries, and teal lines represent AlkB- libraries from this analysis. Only based that were misread greater than 30% of the time were reported in this study.



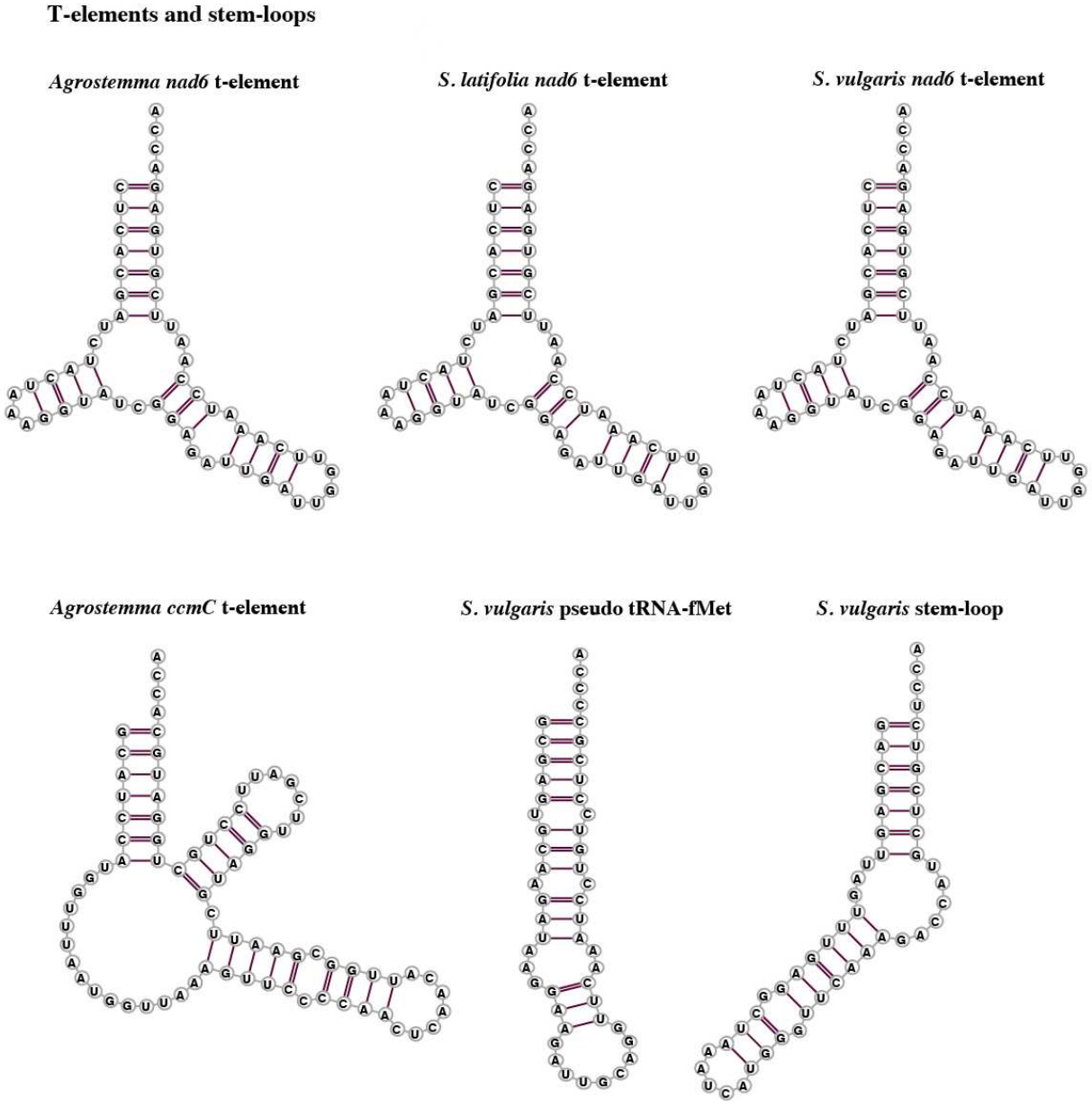
Supplementary Figure 2.5| tRNA modification indexes by genomic compartment. A position was considered modified if ≥30% of the mapped reads differed from the reference sequence and the sequence was detected by more than five reads. Only AlkB+ treated(red) and AlkB negative (teal) libraries are shown.



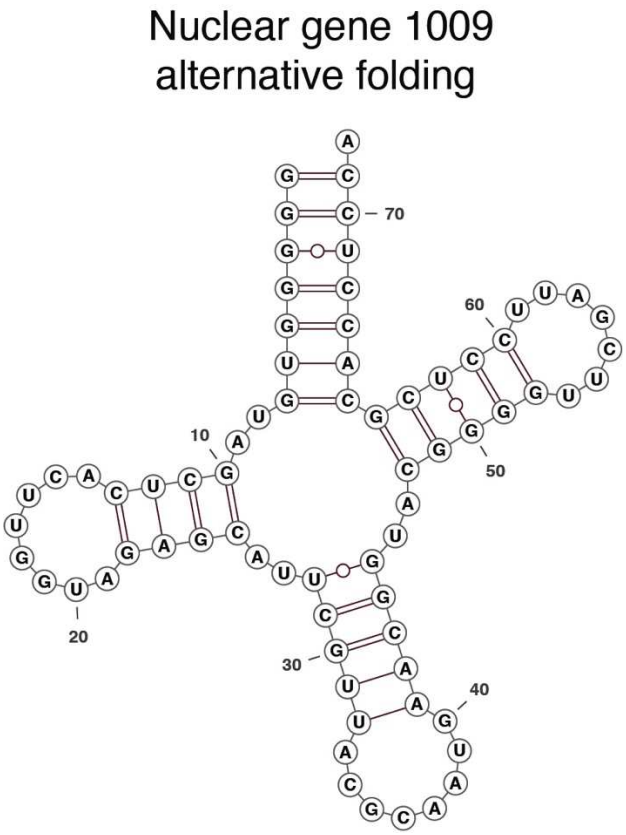
Supplementary Figure 2.6| A tRNA modification index showing the proportion of *A. thaliana, M. truncatula, S. tuberosum, and O. sativa* tRNA reference sequences with a misincorporation/deletion at each nucleotide position. A position was considered modified if ≥30% of the mapped reads differed from the reference sequence and the sequence was detected by more than five reads. The modification index for *A. thaliana* was produced from the AlkB+ (rep 1) library.



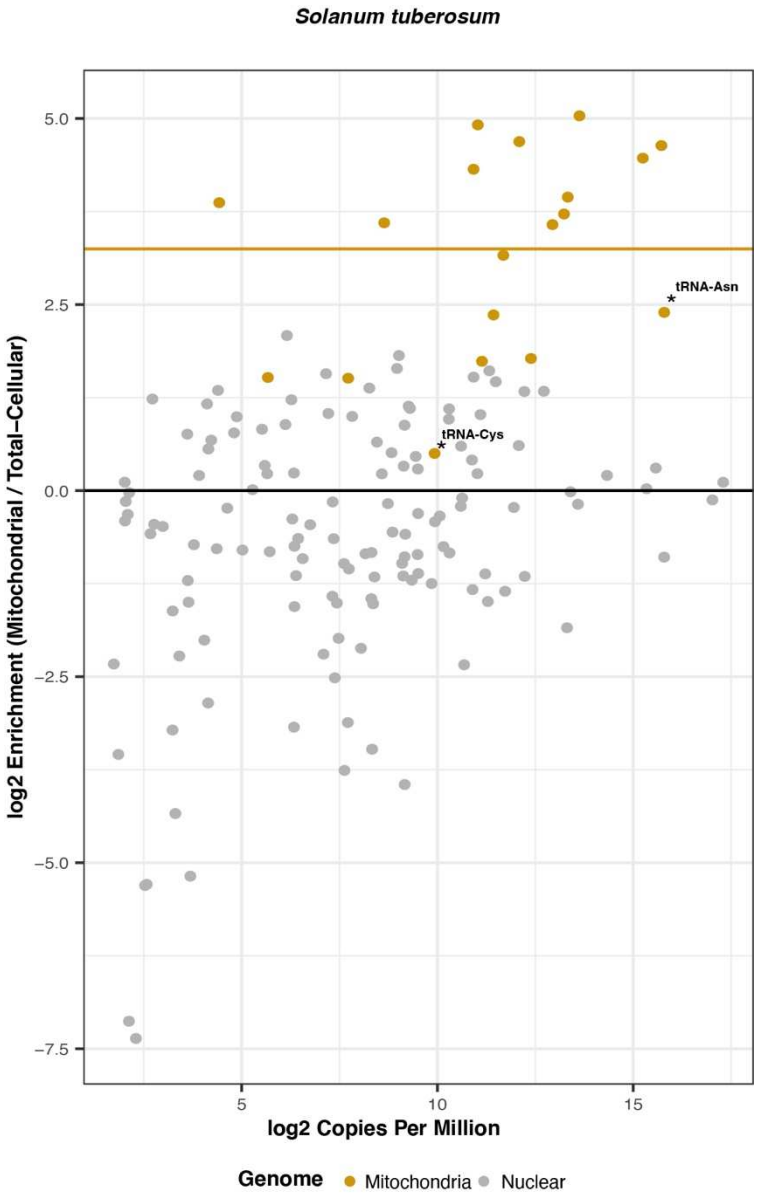
Supplementary Figure 3.1| Mitochondrial genome map of *Agrostemma githago*. This "master circle" represents only one possible configuration of the mitogenome due to multiple recombing repeats. Boxes inside and outside the circle correspond to the clockwise and anticlockwise strands, respectively. The black regions of the repeats designate the identical sequence shared by all four copies. The flanking gray boxes represent extensions to these repeat regions, which are present in two or more (but not all four) copies. The inner gray track represents GC content. Figure was generated with OGDraw v1.3.1(GREINER *et al*. 2019)



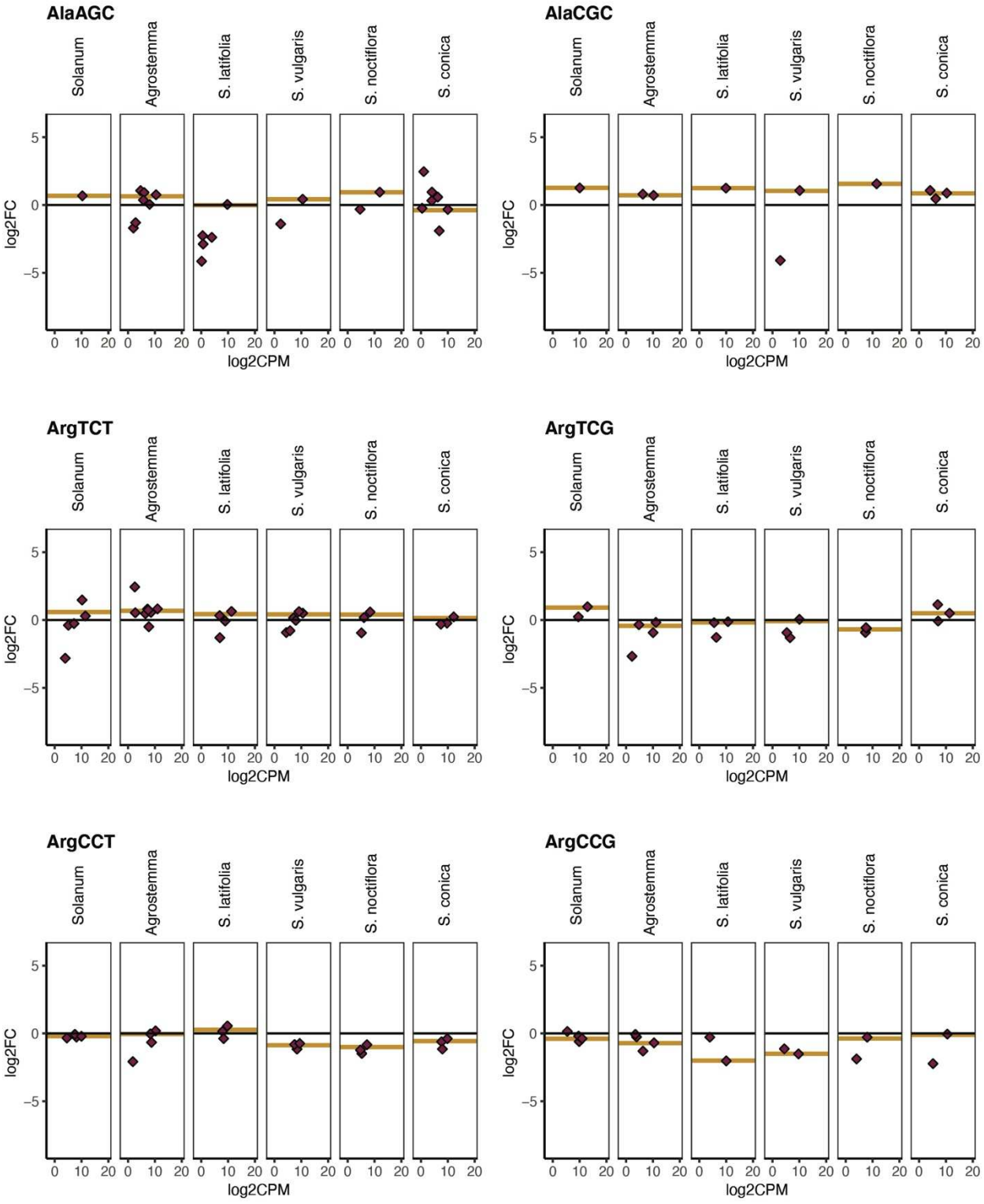
Supplementary Figure 3.2| Predicted foldings of some of the most highly detected stem-loop and t-element structures. Folding predictions were done using RNAfold (LORENZ *et al*. 2011) with a maximum free energy model, and diagrams were created with VARNA (ver. 3.9, (DARTY *et al*. 2009)).

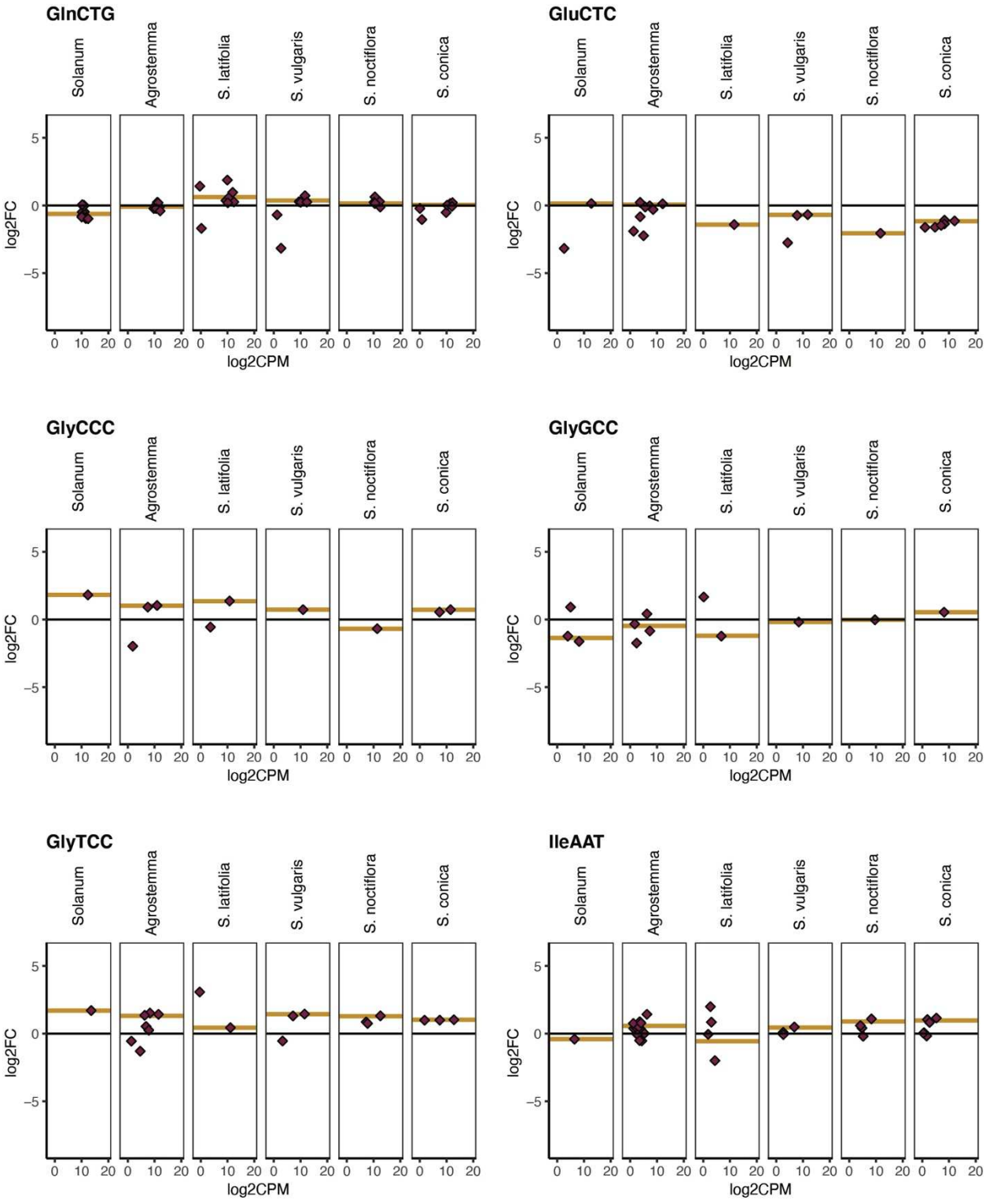


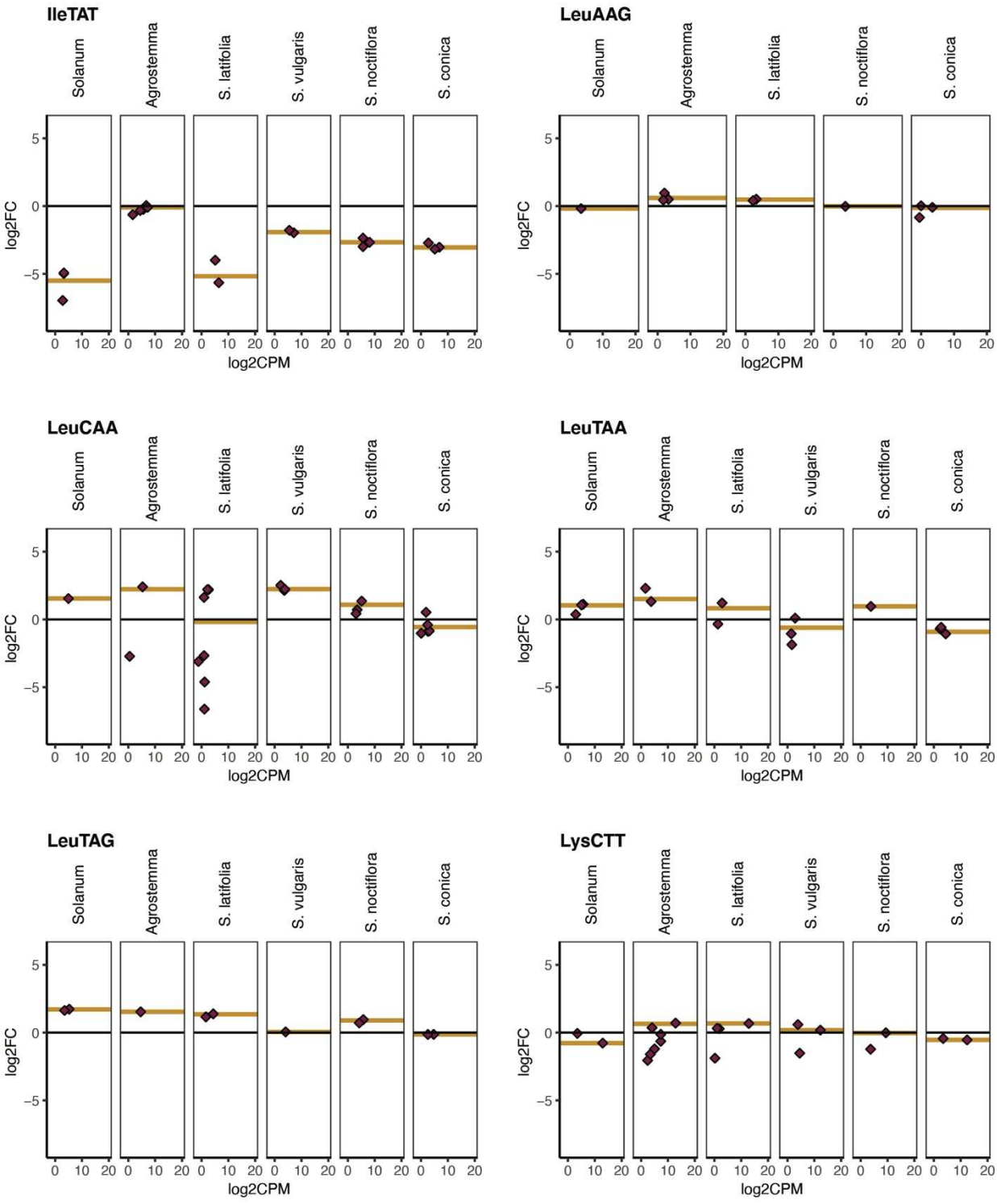
Supplementary Figure 3.3| Alternative folding prediction of the mitochondrially-enriched, nuclear-encoded tRNA in *S. vulgaris*. Structure presented represents the maximum free energy folding model using the program the program RNAfold (LORENZ *et al*. 2011) and diagram was created using the program VARNA ver. 3.9 (DARTY *et al*. 2009).

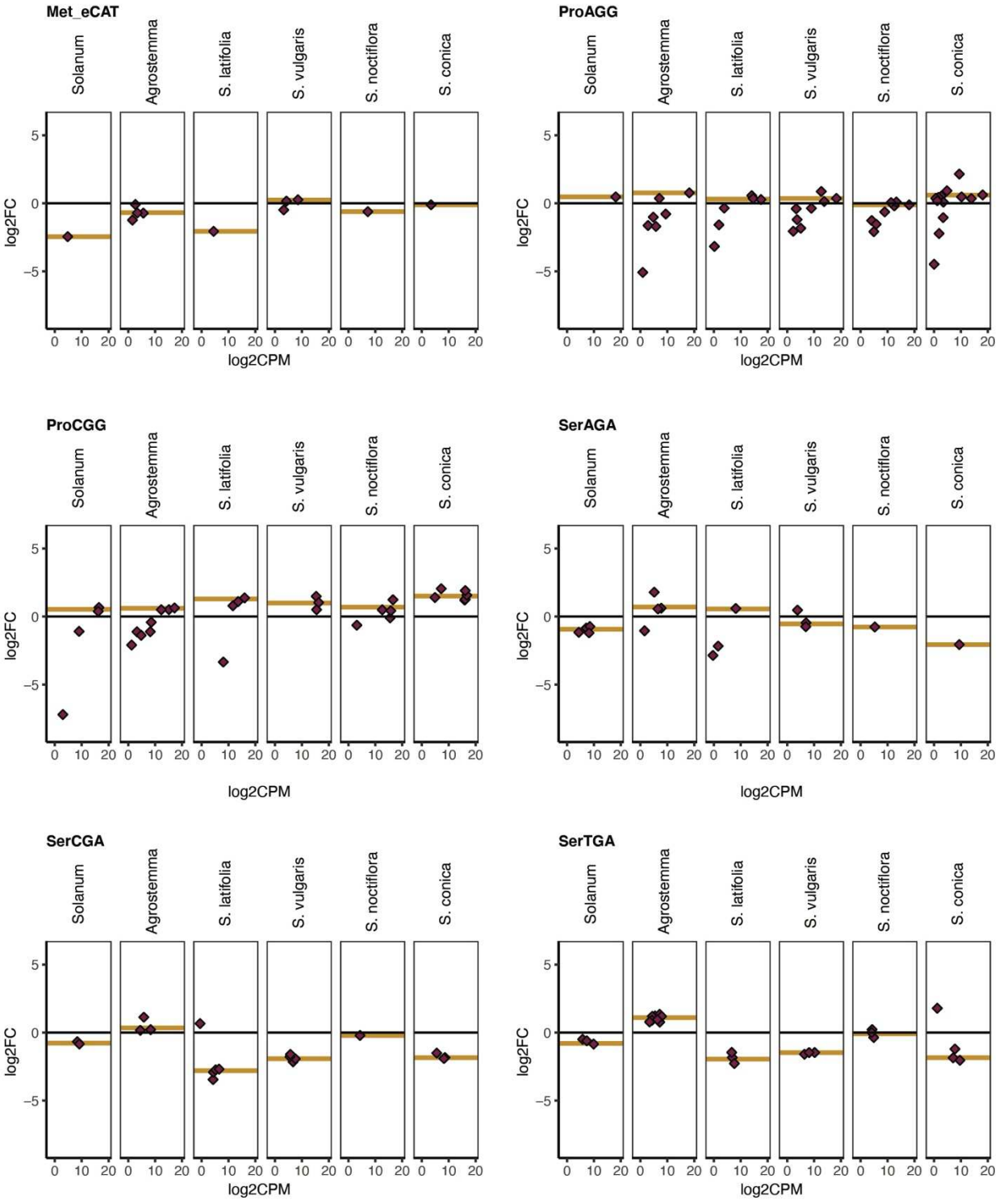


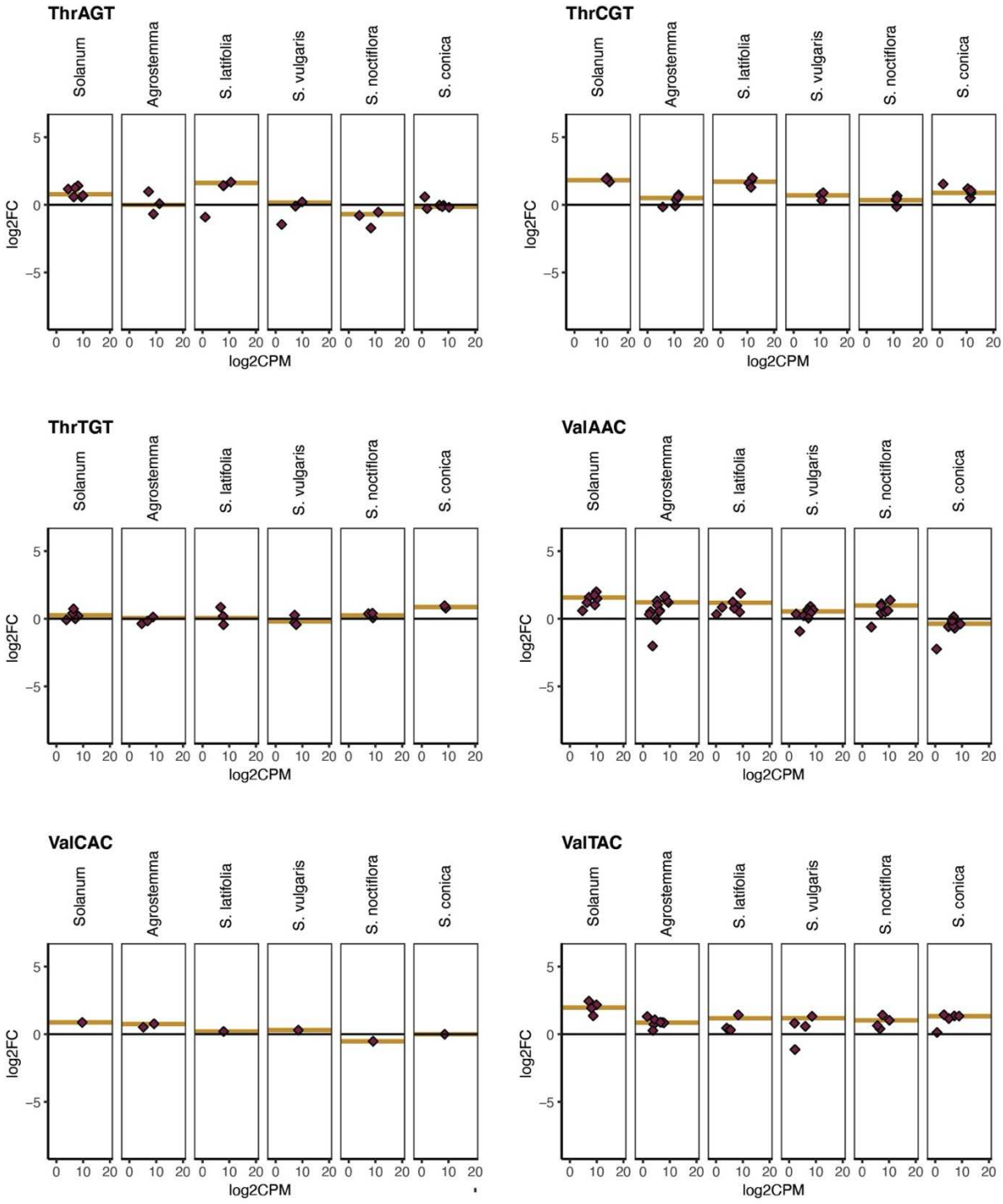
Supplementary Figure 3.5| Enrichment of mitochondrially encoded and nuclear-encoded tRNAs in the mitochondrial isolate relative to total-cellular sample in *S. tuberosum*. Each dot represents a unique tRNA sequence. The heavy gold line represents the average enrichment of mitochondrial-encoded genes. The genomic origin of the tRNA is indicated by color, with gray being nuclear-encoded and gold being mitochondrial. The two labeled mt-tRNAs (Cys and Asn) are identical in sequence to the plastid-encoded tRNA counterparts.


AlaAGC
AlaCGC
ArgTCT
ArgTCG
ArgCCT
ArgCCG
Solanum
Agrostemma
S. latifolia
S. vulgaris
S. noctiflora
S. conica
log2FC
log2CPM
5
0
−5
0
10
20


GlnCTG
GluCTC
GlyCCC
GlyGCC
GlyTCC
IleAAT
Solanum
Agrostemma
S. latifolia
S. vulgaris
S. noctiflora
S. conica
log2FC
log2CPM


IleTAT
LeuAAG
LeuCAA
LeuTAA
LeuTAG
LysCTT
Solanum
Agrostemma
S. latifolia
S. vulgaris
S. noctiflora
S. conica
log2FC
log2CPM


Met_eCAT
ProAGG
ProCGG
SerAGA
SerCGA
SerTGA
Solanum
Agrostemma
S. latifolia
S. vulgaris
S. noctiflora
S. conica
log2FC
log2CPM



Supplementary Figure 3.5| Enrichment and expression of individual cytosolic tRNA genes The y-axis is enrichment in $\log_2$ fold change in mitochondrial isolates versus total-cellular samples. The x-axis is expression level in counts per million on a $\log_2$ scale. Points represent individual reference sequences, and gold lines represent the average enrichment of all nuclear-encoded tRNAs with that anticodon weighted by expression (see Methods).