PNAS RESEARCH ARTICLE | EVOLUTION

# Photosynthetic demands on translational machinery drive retention of redundant tRNA metabolism in plant organelles

Rachael A. DeTar[a,1], Joanna M. Chustecki[b], Ana Martinez-Hottovy[b], Luis Federico Ceriotti[c,d], Amanda K. Broz[a], Xiaorui Lou[a], M. Virginia Sanchez-Puerta[c,d], Christian Elowsky[e], Alan C. Christensen[b], and Daniel B. Sloan[a]

Affiliations are included on p. 10.

Edited by Alice Barkan, University of Oregon, Eugene, OR; received October 25, 2024; accepted November 14, 2024

**Eukaryotic nuclear genomes often encode distinct sets of translation machinery for function in the cytosol vs. organelles (mitochondria and plastids). This raises questions about why multiple translation systems are maintained even though they are capable of comparable functions and whether they evolve differently depending on the compartment where they operate. These questions are particularly interesting in plants because translation machinery, including aminoacyl-transfer RNA (tRNA) synthetases (aaRS), is often dual-targeted to the plastids and mitochondria. These organelles have different functions, with much higher rates of translation in plastids to supply the abundant, rapid-turnover proteins required for photosynthesis. Previous studies have indicated that plant organellar aaRS evolve more slowly compared to mitochondrial aaRS in eukaryotes that lack plastids. Thus, we investigated the evolution of nuclear-encoded organellar and cytosolic aaRS and tRNA maturation enzymes across a broad sampling of angiosperms, including nonphotosynthetic (heterotrophic) plant species with reduced plastid gene expression, to test the hypothesis that translational demands associated with photosynthesis constrain the evolution of enzymes involved in organellar tRNA metabolism. Remarkably, heterotrophic plants exhibited wholesale loss of many organelle-targeted aaRS and other enzymes, even though translation still occurs in their mitochondria and plastids. These losses were often accompanied by apparent retargeting of cytosolic enzymes and tRNAs to the organelles, sometimes preserving aaRS–tRNA charging relationships but other times creating surprising mismatches between cytosolic aaRS and mitochondrial tRNA substrates. Our findings indicate that the presence of a photosynthetic plastid drives the retention of specialized systems for organellar tRNA metabolism.**

organelle gene expression | plastome | aminoacyl-tRNA synthetase | tRNA | photosynthesis

## Significance

The process of endosymbiont integration into a host to become an organelle results in a combination of gene loss, transfer to the nucleus, and retention in the organellar genome. It is unclear why some endosymbiont-derived genes are retained when a functional host counterpart exists whose gene product could simply be retargeted to the organelles. This study revealed that the photosynthetic activity in plant plastids may be responsible for retention of functionally redundant transfer RNA (tRNA) processing machinery, while mitochondria are more flexible regarding substitution with cytosolic-type enzymes. Therefore, functional constraint in the plastid is likely more important than in the mitochondria for shaping the evolution and retention of functionally redundant proteins that are dual targeted to both organelles.

The evolutionary transition from a free-living bacterium to an endosymbiont and eventually to an organelle requires the physical, biochemical, and genomic integration of two organisms. One common consequence of this process is a reduction in the genome of the endosymbiont, driven by relocation of essential genes to the nucleus via horizontal transfer or outright loss of some genes that are no longer useful after transition to an intracellular lifestyle. This trajectory has been observed repeatedly for various endosymbionts and organelles, including mitochondria, plastids, and more recently acquired intracellular bacteria such as the cyanobacterial-derived chromatophores of *Paulinella* amoeba or the many heritable endosymbionts in insects (1–6). However, there are some endosymbiotically derived genes that are recalcitrant to loss despite the presence of host counterparts with comparable functions. Redundant genes may be maintained in the organelle genomes or in the nuclear genome alongside genes encoding proteins for the same role in other cellular compartments. These observations elicit some interesting questions; 1) Why are some functionally redundant genes retained? 2) Do these redundant genes evolve differently depending on the cellular compartment where they operate?

The genes that encode the machinery required for organelle translation are widely retained, as eukaryotes maintain distinct translation systems for nuclear and mitochondrial transcripts. Translation becomes even more complicated in eukaryotes bearing three genomes. For example, plant cells contain nuclear, mitochondrial, and plastid genomes. The translation machinery associated with each of these cellular compartments includes ribosomal RNAs (rRNAs), ribosomal proteins, transfer RNAs (tRNAs), and the accompanying enzymes involved in tRNA processing and aminoacylation. The aminoacyl-tRNA synthetases (aaRS) that catalyze the "charging" of a tRNA with its cognate amino acid tend to be highly conserved, with distinct versions for function in the organelles and

Author contributions: R.A.D., L.F.C., A.K.B., M.V.S.-P., C.E., A.C.C., and D.B.S. designed research; R.A.D., J.M.C., A.M.-H., L.F.C., X.L., and C.E. performed research; R.A.D., L.F.C., and X.L. analyzed data; and R.A.D., J.M.C., A.M.-H., L.F.C., A.K.B., M.V.S.-P., C.E., A.C.C., and D.B.S. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.



[1]To whom correspondence may be addressed. Email: rachael.detar@colostate.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2421485121/-/DCSupplemental.

Published December 18, 2024.

cytosol (7). The degree of redundancy of tRNAs and their interactors between cellular compartments also varies significantly within eukaryotes. These characteristics make tRNAs and their interactors, particularly aaRS, an ideal case study for understanding cytonuclear gene redundancy.

The concerted activity of the aaRS along with other enzymes involved in posttranscriptional modification such as tRNases P and Z (end processing in tRNA maturation) and CCAses (addition of CCA tail) ensures the tRNA has the proper structure to fulfill its role in translation. Generally, organisms would need a complement of at least 20 aaRS, one for each amino acid. However, eukaryotes often have more than this minimum requirement, with some overlapping sets of enzymes targeted to different compartments (8, 9). It is not fully understood why some eukaryotes retain distinct, redundant sets of tRNA processing machinery, while others do not. However, retention of organelle-targeted aaRS correlates with the retention of tRNA genes within organelle genomes. For example, *Homo sapiens* retains 38 aaRS with nearly distinct sets for the cytosol and mitochondria (9–11). Correspondingly, all necessary tRNA genes are present in the mitochondrial genome (mitogenome) (8). In contrast, trypanosomatids do not encode any tRNAs in their mitogenomes but rather import nuclear-encoded tRNAs into the mitochondria and have a reduced set of aaRS that colocalize to the cytosol and the mitochondria (9, 12, 13). In plants, closely related species may have very different numbers of tRNAs encoded in their mitogenomes, with import of cytosolic tRNAs to compensate for losses (8, 14, 15). Notably, plants also dual-target and import an organellar set of enzymes into both plastids and mitochondria (9, 16). The co-occurrence of redundant tRNAs and aaRS could be explained by the observation that organelle-encoded tRNAs are often structurally divergent from cytosolic tRNAs and require specific enzymatic partners for recognition (8, 17).

For most eukaryotes, organellar and cytosolic aaRS are exclusively nuclear-encoded. The shared genomic location of these genes facilitates direct comparison of sequence conservation because any differences in evolutionary rate more likely reflect differences in selection, rather than other factors like compartment-specific mutation rates or effective population size (18–21). Previous studies have quantified rates of evolution for these enzymes in various animal lineages and a small number of plants. Mitochondria-specific aaRS in animals exhibit lower sequence conservation than their cytosolic counterparts (22–24). In contrast, plant organellar aaRS exhibit equivalent or greater sequence conservation compared to their cytosolic counterparts (25). Furthermore, plastid genomes (plastomes) tend to retain a set of 30 tRNA genes that is fully sufficient for translation (26, 27). Plastomes also generally have high rates of transcription (28) and translation (29) due to the abundance and rapid turnover of proteins required for photosynthesis. Thus, the disparity in evolutionary rates between animal mitochondrial aaRS and plant dual-organelle targeted aaRS may be due to the specific translational needs of the plastid. We hypothesize that aaRS and other translation-associated enzymes operating in the plastid face great functional constraints, thus promoting unusually high sequence conservation of organellar translation machinery compared to other eukaryotes. A prediction stemming from this hypothesis is that plant lineages with reduced or nonexistent plastid gene expression would exhibit more rapid evolution of organellar aaRS due to relaxed selection.

In this work, we test whether photosynthetic activity contributes to sequence conservation of organellar tRNA processing and charging enzymes by analyzing evolution of this machinery across diverse lineages of autotrophic and heterotrophic plants. It should be noted there is a gradient of heterotrophy among plants—for the purposes of this study, we use "heterotrophic" to describe plants that do not photosynthesize at all and exclusively rely on parasitism for acquisition of carbon. They do this by tapping directly into host plants or by exploiting mycorrhizal fungi (mycoheterotrophy) (30, 31). We refer to facultative autotrophs that supplement photosynthetic carbon with parasitism as "hemiparasitic." The hallmark of many heterotrophic/hemiparasitic plants is a reduced or even nonexistent plastome (32). Highly expressed genes related to photosynthesis such as the Rubisco large subunit and the Photosystem II D1 reaction center tend to be absent from plastomes in heterotrophs, thereby drastically decreasing demands on plastid gene expression. In addition, many tRNA genes have been lost from the plastomes of these plant species (32), providing a good system to examine the relationship between evolutionary redundancy of aaRS, tRNA content of the organelle genomes, and trophic lifestyle.

## Results

**Heterotrophic Plants Exhibit Extensive Loss of Organellar aaRS and tRNA Processing Enzymes but Not Ribosomal Protein Subunits.** We searched for organellar and cytosolic aaRS orthologs in a broad sampling of angiosperm species with a range of trophic lifestyles and plastome reduction (Dataset S1). For each hemiparasitic/heterotrophic species in the analysis, we included the closest fully autotrophic relative with a nuclear genome or transcriptome available, excepting the Balanophoraceae where we used hemiparasite *Santalum album* as the reference autotroph. In total, we sampled 7 heterotrophs, 8 hemiparasites, and 10 autotrophs, including reference species *Arabidopsis thaliana.* Among heterotrophs, we included the endoparasites *Rafflesia cantleyi* and *Sapria himalayana* from the family Rafflesiaceae, likely the most extreme examples of plant heterotrophy, as several studies have suggested wholesale loss of the plastome in these species (33–36).

We found orthologs across these diverse taxa by running Orthofinder (37, 38) on protein models derived from nuclear genome or transcriptome data to find groups of likely orthologous proteins based on similarity and phylogeny, i.e., "orthogroups." Then we extracted sequences that fall into an orthogroup with known *A. thaliana* aaRS (Dataset S2). It is important to note that the number of homologous sequences found for each species in this analysis may not directly reflect the absolute number of genes or proteins due to the fragmented nature of gene models made from transcriptome data and the inclusion of two independent sets of protein models for *Balanophora fungosa* (39, 40) and *S. himalayana* (41, 42). Unexpectedly, many lineages of heterotrophs did not have any orthologs for many of the organellar aaRS. This loss of organelle-specific orthologs in heterotrophs was surprising as even species entirely lacking the plastome must still conduct tRNA aminoacylation within mitochondria. Further investigation using reciprocal-best-hit (RBH) searches with BLAST between *A. thaliana* and each species supported that these were true losses. These losses primarily occurred in purely heterotrophic plants, while hemiparasites retained most of their organellar aaRS (Fig. 1). Yet, within the heterotrophs, there was a gradient of loss. The two endophytic species from Rafflesiaceae (*R. cantleyi* and *S. himalayana*) exhibited the most striking loss of organellar aaRS, retaining only PheRS. In contrast, obligate mycoheterotroph *Monotropa hypopitys* retained more aaRS than any other heterotroph, missing only AlaRS, CysRS, ThrRS, and TrpRS (Fig. 1).

This striking pattern led us to investigate the retention of other enzymes involved in tRNA maturation and processing; see ref. 15 and Dataset S2 for a comprehensive list. Unsurprisingly, the CCAse enzyme responsible for adding CCA tails to tRNAs in all compartments was retained across all species. Two enzymes that maintain distinct cytosolic and organellar versions, tRNase Z (involved in 3′

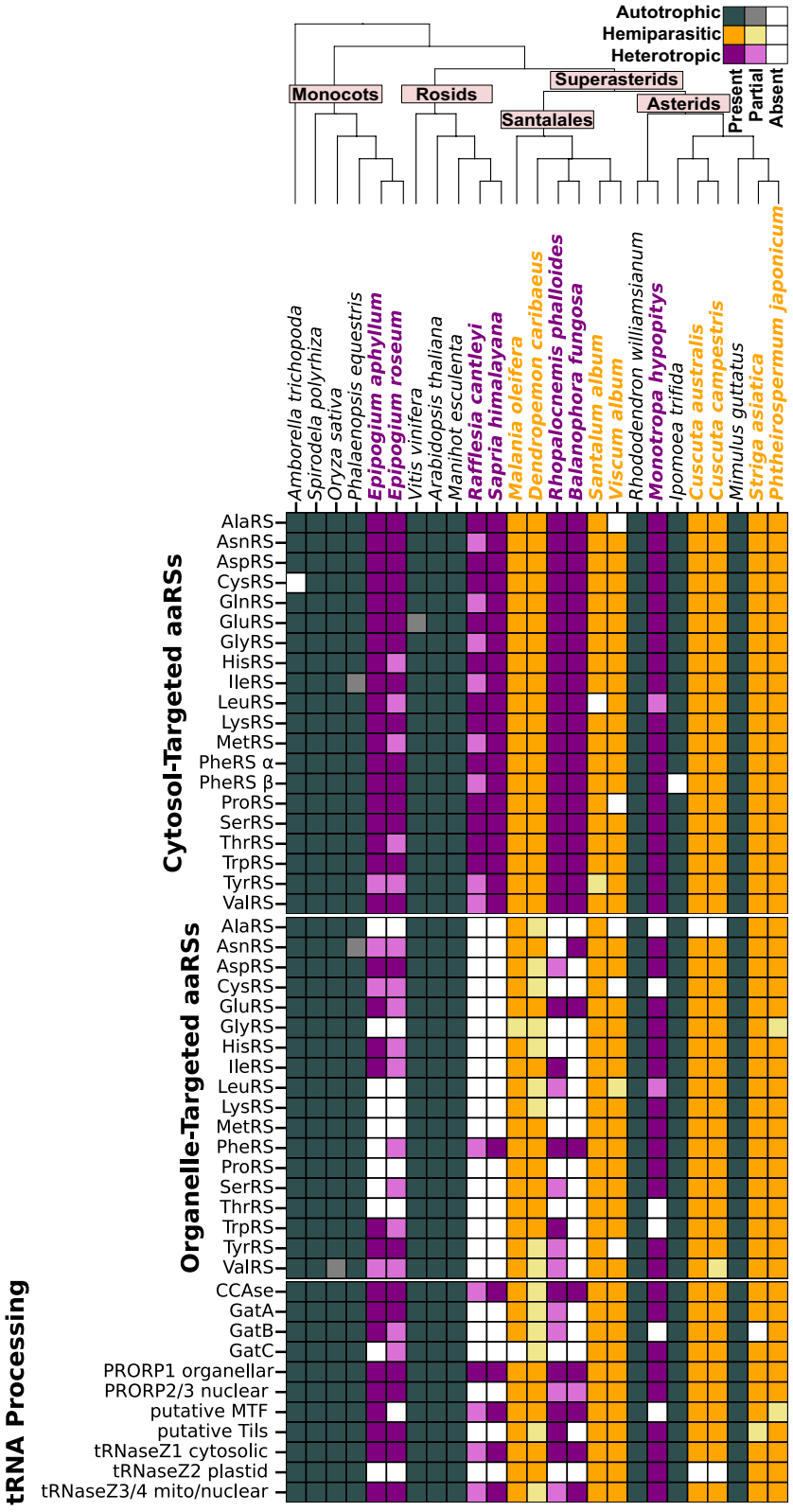


**Fig. 1.** Presence/absence matrix of selected organellar and cytosolic aaRS and tRNA-processing enzymes shows major losses of organellar enzymes in heterotrophic lineages. Hue of boxes indicates metabolism (gray = autotroph, orange = hemiparasite, purple = heterotroph), shade indicates degree of presence/absence (white = no sequence found, light colors = partial sequence found, dark colors = complete or near complete sequence found). Note that there is no organellar GlnRS as plants typically use the bacterial-type GatCAB amidotransferase system to charge organellar Gln-tRNA. ArgRS is not presented as it has no distinct organellar enzyme; gene products are either cytosol-targeted, or triple-targeted to the cytosol, mitochondria, and plastids.

end processing) and Protein Only RNase P (PRORP; required for 5′ end processing), exhibited two different trends. We found that plastid tRNase Z (tRNase Z2) (43) was missing from all heterotrophs except *M. hypopitys* and hemiparasites of the *Cuscuta* genus yet was retained in the autotrophic relatives. Conversely, the organellar PRORP1 (44, 45) appeared to be universally retained while cytosolic PRORP2/3 orthologs were missing from the Rafflesiaceae. We also investigated the retention of bacterial-like organelle-specific machinery with no

cytosolic counterpart, including the methionyl-tRNA formyltransferase (MTF), tRNA-Ile lysidine synthetase (TilS), and the bacterial-type glutamine amidotransferase (GatCAB) complex. We found orthologs for the putative *A. thaliana* MTF in all species except the heterotrophs *Epipogium roseum* and *M. hypopitys,* and putative TilS orthologs were absent only from the heterotrophs Rafflesiaceae and *B. fungosa*. GatCAB is the key complex used by organelles in land plants to accurately charge $tRNA^{Gln}$ in the absence of a dedicated organellar GlnRS (46–48). This involves a two-step process, whereby $tRNA^{Gln}$ is initially misacylated with Glu by GluRS. GatCAB then catalyzes the transamidation of Glu to Gln. Most of our test species, even heterotrophs, retained genes for at least some of the GatCAB subunits. However, Rafflesiaceae species appear to have lost all subunits in this complex.

To assess whether other functionally redundant translation machinery was lost in heterotrophic lineages, we analyzed presence/absence of nuclear-encoded riboproteins with exclusive targeting to a given compartment (*SI Appendix*, Table S1 and Dataset S2, "Riboproteins"). While there was minor loss of cytosolic, mitochondrial, and plastid subunits in all heterotrophic lines, the Rafflesiaceae lineages exhibited a dramatic reduction in plastid subunits, which is expected because these lineages appear to have dispensed with the plastid genome entirely and consequently no longer need plastid ribosomes. This finding reveals a key contrast between the fate of organellar aaRS and riboproteins in response to loss of photosynthesis and may reflect the fact that aaRS are shared between the two organelles, whereas the riboproteins are compartment-specific.

Generally, we observed reductive evolution in all heterotrophs via loss of organellar aaRS and bacterial-type enzymes for tRNA modification. This raises the question as to how gene expression is accomplished in the organelles of these species. One possibility is that cytosolic enzymes are retargeted to the organelles to functionally replace organellar enzymes.

**Heterotrophic Plants May Compensate for Loss of Organellar aaRS by Retargeting Cytosolic Counterparts.** To test for cytosolic aaRS retargeting, we applied two in silico targeting prediction programs to protein models for all cytosolic aaRS in our heterotrophs and their close autotrophic relatives (Fig. 2 and *SI Appendix*, Fig. S1). Our analysis revealed that canonically cytosolic aaRS are disproportionately retargeted to mitochondria in obligate heterotrophs compared to autotrophs ($P = 0.018$). There was also a nonsignificant trend toward retargeting to the plastid ($P = 0.11$). Targeting of mitochondrial oxidative phosphorylation (OXPHOS) components and plastid housekeeping proteins including subunits of the caseinolytic protease (Clp) complex and the acetyl-CoA carboxylase (ACC) complex involved in producing malonyl-CoA for fatty acid biosynthesis were included as known organellar positive controls (Dataset S2, "targeting controls"). Protein degradation and fatty acid biosynthesis are essential plastid functions that are still necessary in the absence of photosynthesis. Cell wall synthesis enzymes were included as negative controls. Of the controls, the only group that showed a different degree of targeting than expected was the mitochondrial OXPHOS components, where recovery of proteins with known mitochondrial targeting was low (~60%) for both autotrophs and heterotrophs (*SI Appendix*, Table S2). Thus, these in silico approaches appear to successfully (albeit imperfectly) predict subcellular localization in both heterotrophic and autotrophic plants.

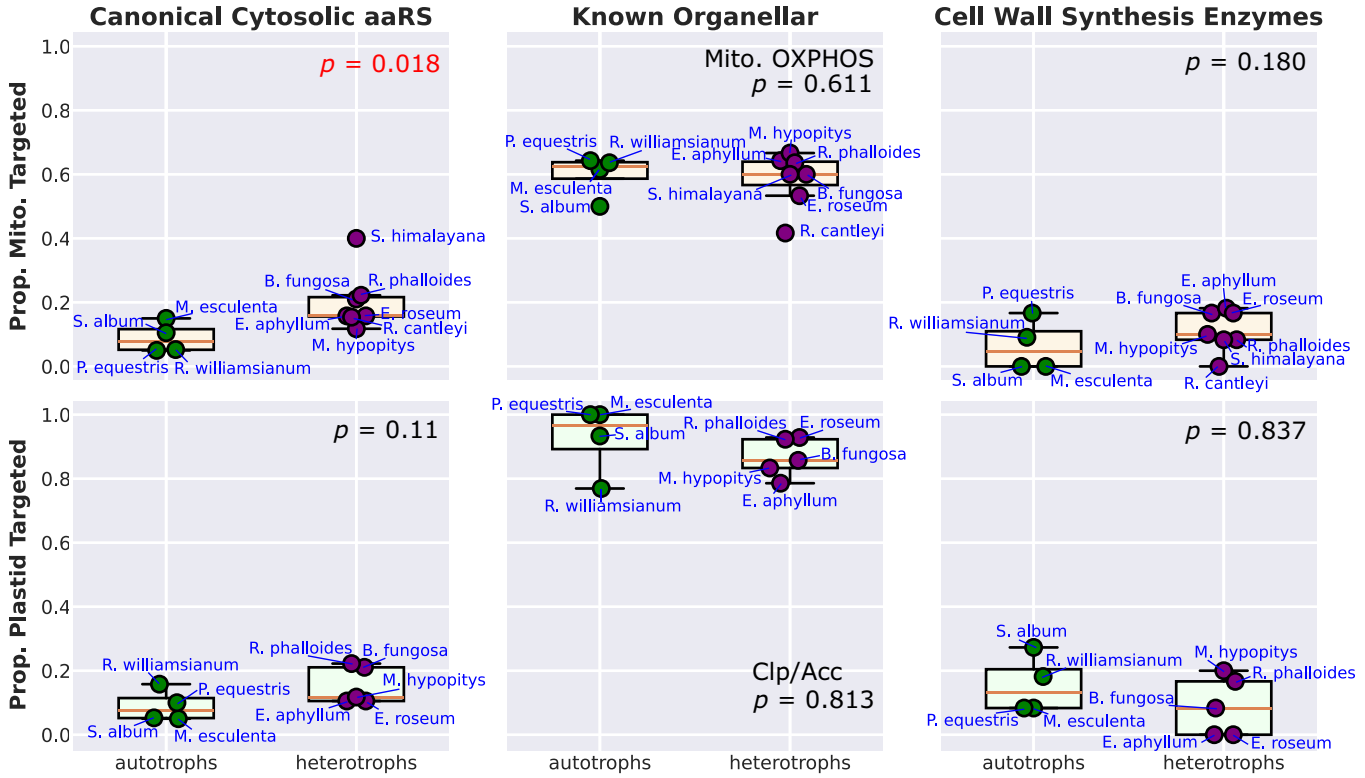


**Fig. 2.** Proportion of retargeted cytosolic aaRS is higher in heterotrophs than autotrophs based on in silico predictions. Proportion of enzymes with predicted retargeting (probability ≥ 0.5) to either mitochondria (*Top* row) or plastids (*Bottom* row) based on the outputs of LOCALIZER and/or TargetP. Autotrophs and heterotrophs are plotted in green and purple, respectively. Targeting predictions for proteins involved in mitochondrial OXPHOS are included as a positive control for mitochondrial targeting, and proteins involved in nonphotosynthesis related housekeeping roles (ACC and Clp subunits) are included as a positive control for plastid targeting. Predictions for cell-wall-synthesis associated enzymes are included as a negative control for organelle targeting. Reported *P*-values from *t* tests or Mann–Whitney *U* tests depending on distribution of the data. Note that *R. cantleyi* and *S. himalayana* are not included in plastid retargeting analysis due to suspected loss of the plastome, but they were predicted to have retargeted 7.7% and 30.0% of their cytosolic aaRS to the plastid, respectively. These species have lost too many plastid housekeeping genes to assess proportion targeted. For both species, no cell wall synthesis enzymes exhibited plastid retargeting.

We would specifically expect that a given cytosolic aaRS is more likely to be retargeted if its organellar counterpart is lost. We examined this possibility and found a clear and statistically significant correlation between loss of an organellar aaRS and retargeting of the corresponding cytosolic enzyme to one or both compartments ($P = 0.0005$, *SI Appendix*, Fig. S2).

The variable quality of gene models for each species and high false negative rate for retargeting prediction confound interpretation of retargeting for individual enzymes in each species (*SI Appendix*, Fig. S3). However, when looking only at *S. himalayana*, a species with relatively high-quality genome-based models, we observed some interesting phenomena. Cytosolic AlaRS, GlnRS, MetRS, ProRS, and TrpRS were predicted to be retargeted to mitochondria. Retargeting of cytosolic GlnRS could explain how this species copes with loss of the GatCAB complex (Fig. 1). The cytosolic AspRS, GluRS, and GlyRS were predicted to be dual targeted to both organelles, while cytosolic TyrRS and the PheRS β subunit were predicted to be targeted to the plastid despite the likely absence of a plastid genome. All other cytosolic enzymes were either not predicted to be retargeted or were missing a large portion of the N terminus. The fact that some cytosolic aaRS corresponding to lost organellar aaRS were not predicted to be retargeted to the mitochondria and that some cytosolic aaRS were targeted to a plastid presumed to lack gene expression may reflect an underestimation of mitochondrial retargeting due to the constraints of in silico predictions and the potential divergence of TP sequences in Rafflesiaceae. Indeed, an investigation of subunits for the protein import apparatus of plastids and mitochondria revealed that the Rafflesiaceae are missing some key components of the plastid import apparatus (*SI Appendix*, Fig. S6), some of which are responsible for TP binding and recognition.

To assess whether our in silico predictions hold true in vivo, we conducted targeting assays on a subset of aaRS (see Dataset S3 for sequence information) by inducing transient expression of N-terminal fusions of *S. himalayana* cytosolic aaRS transit peptides (TPs) to green fluorescent protein (GFP) in *Nicotiana benthamiana* (Fig. 3 and *SI Appendix*, Figs. S4 and S5). We used eqFP611 tagged with the TP from mitochondrial isovaleryl-CoA dehydrogenase (IVD) as a positive control for mitochondrial targeting. Infiltration with single expression vector with just IVD_FP611 was used as a negative control (*SI Appendix*, Fig. S4). We found that proteins predicted to be retargeted in silico were almost universally retargeted to one or both organelles in vivo. In some cases, putative TPs retargeted GFP to a different organelle than predicted in silico. AsnRS displayed the plastid targeting that was predicted in silico and additional mitochondrial targeting in vivo (Fig. 3). GluRS and GlyRS were predicted to be dual-targeted but exhibited targeting to only one of the two organelles. The only two TPs for which organellar GFP localization completely mismatched the in silico prediction (but still exhibited organelle targeting) were ProRS and TrpRS, which both showed strong plastid GFP fluorescence in vivo despite mitochondrial predictions. These results were consistently observed with two different cloning and expression strategies (Fig. 3 and *SI Appendix*, Fig. S5). The cytosolic AspRS TP was the only one we cloned that did not exhibit targeting to either organelle in vivo despite a clear biological expectation that it would have evolved retargeting (*SI Appendix*, Fig. S4). GFP tagged with this TP formed punctate structures in the cytosol, which may be artifactual aggregates resulting from overexpression. Generally, in vivo assays provided evidence for retargeting of cytosolic aaRS to at least one organelle and suggested that in silico prediction tools may not be adept at distinguishing plastid from mitochondrial targeting.

Generally, organellar tRNAs differ from their cytosolic counterparts in evolutionary history, sequence, and posttranscriptional modification (8). Identity elements that govern the specificity of interaction between a tRNA and its conjugate aaRS may also differ between cytosolic and organellar systems (49). However, in many systems (including plants), cytosolic tRNAs can be imported into the organelles (8, 9, 14). Thus, our observation of retargeting opens some interesting questions—are the retargeted aaRS charging organellar tRNAs? Or is there replacement of the entire organellar system of tRNA aminoacylation, including both tRNAs and aaRS? We investigated the tRNA gene content of the organelles of these species to gain insight into which tRNA substrates may be in use.

**Organelle-Encoded tRNA Retention Often, but Not Always, Mirrors Retention of Organelle-Type aaRS.** To investigate the possibility of enzyme–substrate mismatch, we cataloged the presence/absence of organellar tRNA-encoding genes (*trn*) for all sampled heterotrophic species (Fig. 4 and Dataset S4). Generally, there was strong correspondence between retention of organellar aaRS and plastome/mitogenome tRNA genes. As the Rafflesiaceae species are thought to lack plastomes entirely, they likely have no plastome-encoded tRNAs (Fig. 4). The Balanophoraceae species used in this study also have been reported to exhibit loss of all plastid tRNA genes, except for an extremely divergent *trnE* in *B. fungosa* (50–53). The Rafflesiaceae and Balanophoraceae have additionally lost most of their mitogenome-encoded tRNAs. Correspondingly, these two families also have the most reduced sets of organellar aaRS. At the opposite extreme, *M. hypopitys* retains most of its mitochondrial and plastid tRNA genes and corresponding organellar aaRS.

However, there are some incidences where a cytosolic enzyme looks to be charging an organellar tRNA due to loss of the organellar aaRS (Fig. 4). For example, the few remaining mitogenome-encoded tRNAs in the Rafflesiaceae are missing their organellar aaRS partner (except *S. himalayana* $tRNA^{Phe}$) and are likely charged by a cytosolic enzyme. Alternatively, they may be pseudogenes that have been functionally replaced by an imported tRNA. For both compartments, there are numerous instances where there is no encoded tRNA despite retention of the organellar aaRS (Fig. 4). For example, *Rhopalocnemis phalloides* retains its organellar SerRS, PheRS, and GlnRS but has lost the corresponding mitochondrial tRNAs. No retargeting of cytosolic enzymes is predicted in these cases, suggesting the organellar enzyme may be acting on an imported cytosolic tRNA.

**Testing for Accelerations in Organellar aaRS Sequence Evolution in Heterotrophs.** Although wholesale loss of many organellar aaRS in heterotrophs precludes analysis of sequence conservation, *M. hypopitys* retained the majority of organellar aaRS despite exhibiting an obligate heterotrophic lifestyle. Thus, we quantified sequence conservation of organellar and cytosolic aaRS in *M. hypopitys* and its autotrophic relative *Rhododendron williamsianum*. We built gene trees and ran relative rate analysis to test for accelerated evolution on the *M. hypopitys* branch (Dataset S5). We also directly compared branch length ratios of *M. hypopitys* to *R. williamsianum* in cytosolic and organellar aaRS (*SI Appendix*, Fig. S7 and Dataset S5). Neither analysis showed a statistically significant decrease in sequence conservation in organellar enzymes of *M. hypopitys* (*SI Appendix*, Fig. S7). However, the deep rooting of our tree on *A. thaliana* may result in poor resolution of the sequence evolution between *M. hypopitys* and *R. williamsianum*. Additionally, the transition to heterotrophy on the *M. hypopitys* branch may have occurred relatively recently compared to the split from *R. williamsianum*, thus diluting any signals of accelerated evolution.

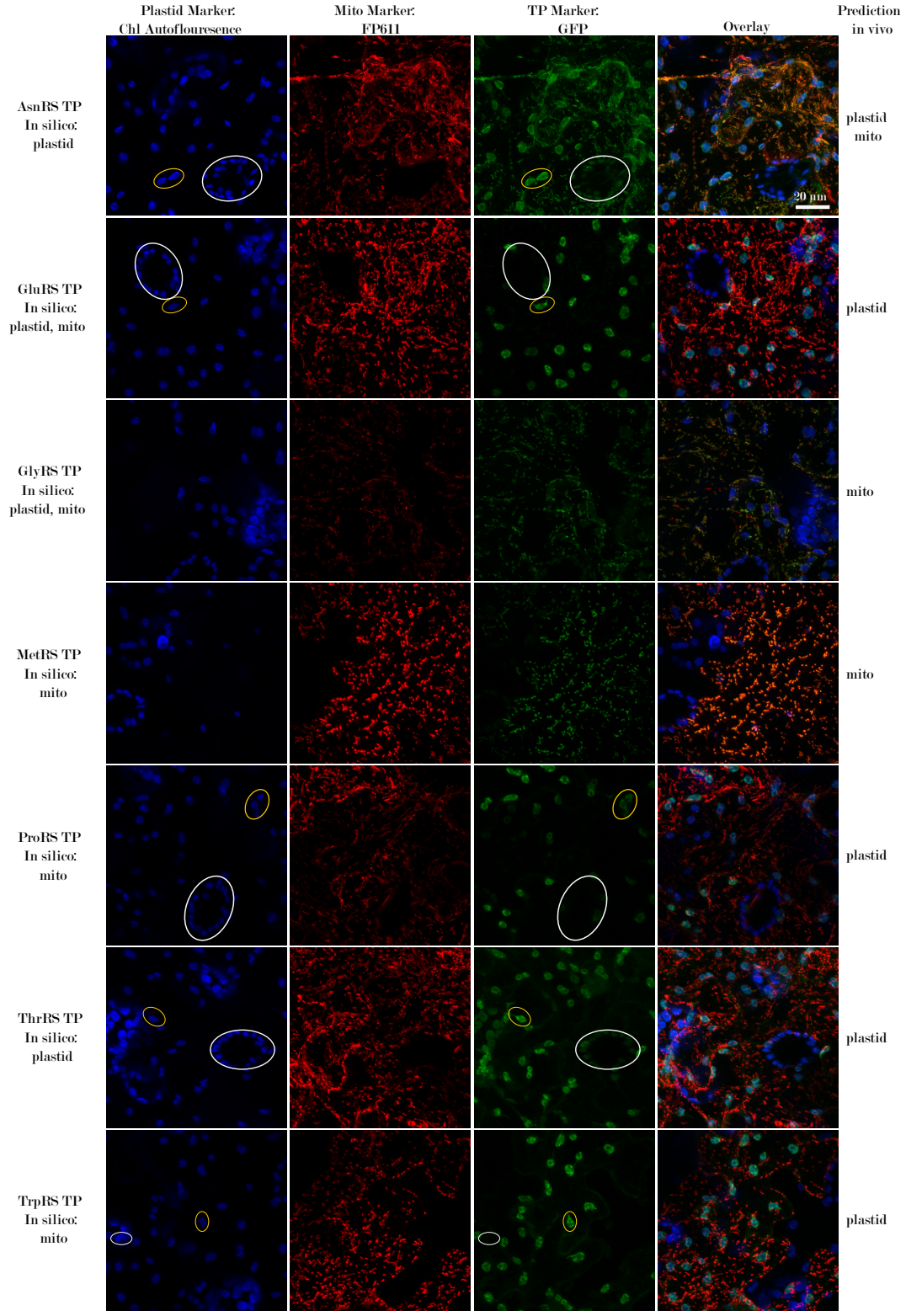


**Fig. 3.** Cytosolic-type aaRS TPs from heterotroph *S. himalayana* can retarget enzymes to organelles in *N. benthamiana*. TPs from *S. himalayana* fused to GFP (*Middle Right* column) were transiently expressed in *N. benthamiana.* Chlorophyll autofluorescence (far *Left* column) was used for colocalization with plastids. As a positive control for mitochondrial targeting, eqFP611 was tagged with the TP from mitochondrial isovaleryl-CoA dehydrogenase (*Middle Left*). The overlay of all three channels is shown on the far *Right*. Scale indicated in the *Top Right* panel is the same for all images. As a means of differentiating true GFP fluorescence from chlorophyll autofluorescence, yellow ovals are shown in panels to indicate plastids in transformed cells, whereas white ovals indicate plastids in nontransformed cells. Note the presence of GFP-fluorescing stromules in transformed plastids, an indication of true GFP signal. For more information, see *Methods* and Dataset S3.

## Discussion

**Loss of Photosynthesis Relaxes Constraints on Both Plastid and Mitochondrial Translation Machinery.** Our primary objective in this study was to test whether the high translational demands of the photosynthetic chloroplast create functional constraint on the evolution of shared translation systems in plastids and mitochondria. To do so, we investigated relative rates of protein evolution in cytosolic and organellar aaRS in heterotrophic plants compared to their photosynthetic relatives. To our surprise, heterotrophic plants underwent widespread loss of organellar aaRS and many bacterial-like enzymes involved in tRNA processing. These losses suggest an extreme relaxation of selection pressures in the absence of a photosynthesizing plastid and support our hypothesis that the translational demands in chloroplasts are the linchpin responsible for the unusually strong functional constraint on the redundant organelle-specific tRNA metabolism in plants. These correlated effects on mitochondrial translation appear to

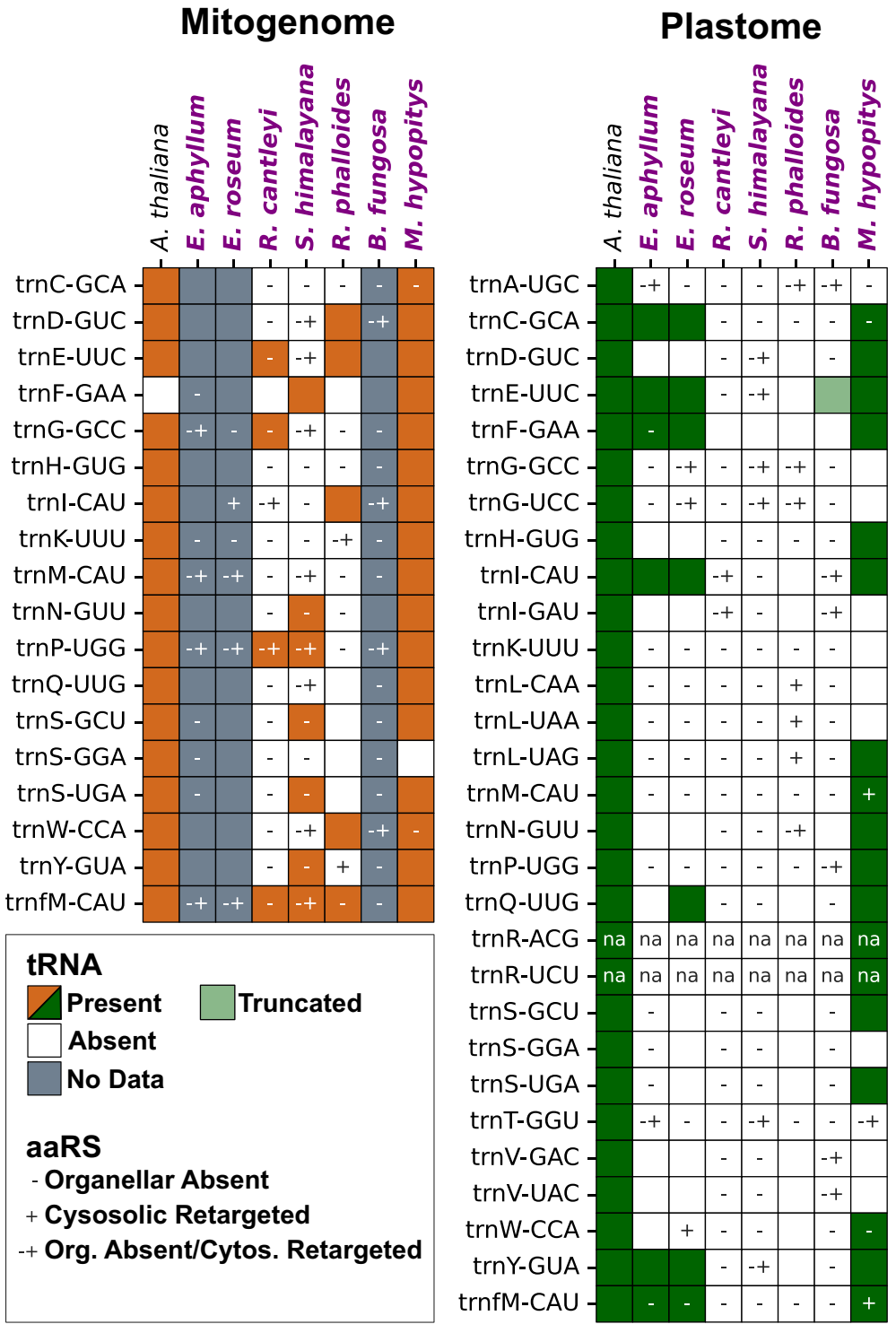


**Fig. 4.** Instances of both concordance and mismatch between organellar tRNA gene loss and aaRS loss/retargeting. Heterotrophic species are written in purple. *A. thaliana* is included for reference. *Left*: mitogenome tRNAs. *Right*: plastome tRNAs. Color of boxes indicates presence/absence of tRNA. The – and + annotation denotes loss of corresponding organellar aaRS and retargeting of cytosolic tRNA respectively as described for Fig. 2 and *SI Appendix*, Fig. S2. Retargeting predictions are specific for each organelle. Note that loss and retargeting of the corresponding aaRS was not assessed for *trnR* as there are no organelle-specific aaRS orthologs responsible for charging this tRNA. Loss of GatCAB and associated cytosolic GlnRS retargeting are presented for *trnQ*.

be stronger when the same proteins are shared by mitochondria and plastids (e.g., aaRS) because we found that the most reduced heterotrophs from Rafflesiaceae lost many plastid-specific riboproteins but not their mitochondrial-specific counterparts.

There was variation in the degree of loss across the heterotrophs, often correlated with the number of genes retained in the plastome, particularly tRNAs (see Dataset S1 for plastome/mitogenome content). For example, the most aaRS losses occurred in the Rafflesiaceae, all of which are obligate endoparasites and have likely lost their plastomes entirely (33–36, 41). The fewest aaRS losses occurred in *M. hypopitys*, which has the highest number of plastome-encoded genes of all the heterotrophic species we sampled—and notably the greatest number of retained plastid tRNAs (54). Thus, *M. hypopitys* may not be as far along the path of reductive evolution as the other species, which might provide an additional explanation as to why our comparative rate analyses on the

remaining organellar aaRS in *M. hypopitys* did not support accelerated evolution compared to cytosolic enzymes (*SI Appendix*, Fig. S7). The Balanophoraceae and *Epipogium* species exhibited intermediate degrees of aaRS and plastome gene loss, although the Balanophoraceae species used in this study have lost all plastid tRNAs save for a truncated *trnE* in *B. fungosa* that is likely retained for use in the first step of heme biosynthesis (50). An increased plastome mutation rate caused by loss of DNA repair machinery (39, 55) in the Balanophoraceae may explain accelerated plastome (tRNA) gene losses relative to nuclear (aaRS) gene loss. Generally, loss of organellar aaRS correlates well with reduction of the plastome across heterotrophs.

**Loss of Organellar aaRS Is Associated with Retargeting of Cytosolic aaRS.** The wholesale loss of organellar aaRS in heterotrophs was unexpected, as most nonphotosynthetic plants still contain a minimal plastome to support heme and fatty acid biosynthesis (32, 56) and fully intact mitogenomes (Dataset S1). This conundrum prompted us to investigate whether cytosolic aaRS are retargeted to organelles in heterotrophs to compensate for gene loss, as has been shown for aaRS/tRNAs in mitochondria of other plant species (57). Retargeting a protein to the organelles generally involves obtaining an N-terminal TP sequence that can be recognized by organellar protein import complexes (58–60). We found that many cytosolic aaRS in the heterotrophs had N-terminal sequences predicted to target to one or both organelles (Figs. 2 and 3 and *SI Appendix*, Figs. S2–S5). Therefore, the import of cytosolic aaRS likely compensates for the loss of organellar aaRS and even the loss of the GatCAB complex in Rafflesiaceae via retargeting of cytosolic GlnRS.

There was not perfect correspondence between loss of organellar aaRS and predicted retargeting of cytosolic aaRS. However, the extent of retargeting may have been underestimated for two main reasons. First, TP sequence composition and protein import machinery vary considerably by plant lineage (61), making in silico predictions more challenging in uncharacterized species. Thus, some organelle-imported cytosolic aaRS may eschew the use of typical TPs that are key for targeting prediction in silico (62). Indeed, some organellar proteins are known to not rely at all on N-terminal TPs for import (63, 64). TP sequence divergence could also be accelerated by alteration of the sequence or subunit composition of protein import machinery. Notably, our presence/absence analysis of plastid and mitochondrial protein import machinery revealed some critical subunits for plastid protein import are lost in the heterotrophs, especially the Rafflesiaceae, which could influence which sequences are recognized as TPs (*SI Appendix*, Fig. S6). For example, Toc159, one of the key outer envelope import receptors involved in TP binding and recognition, is missing from the Rafflesiaceae (65). This subunit specializes in binding and import of proteins involved in photosynthesis, whereas Toc132, which is retained in these species, specializes in "nongreen" protein import (66, 67). The Rafflesiaceae are also missing Tic236, the protein that spans the space between the envelope membranes to physically connect the inner and outer envelope (68). This raises interesting questions about the nature of plastid protein import and the ultrastructure of envelope membranes in Rafflesiaceae. Other heterotrophs in addition to the Rafflesiaceae are also missing some subunits present in *A. thaliana*; *Epipogium* species, *M. hypopitys,* and the Rafflesiaceae lack one or more Tic20 and Tic22 paralogs. The Tic22 protein family is highly conserved and expressed, with a hypothesized role in outer membrane maturation in addition to its role in protein import (69, 70). Study of Tic-22 loss-of-function mutants revealed that these proteins are not essential, at least in *A. thaliana* (69, 71). Tic20 forms the structural basis of the inner pore of the translocon, with Tic20-I/Tic20-IV being the essential isoforms for plant survival (65, 72–74). Similar trends have also recently been found by Guo et al. (75). These alterations to the composition of the TIC-TOC complex may reduce the specificity of plastid protein targeting predictions. In contrast, the heterotrophic lines retained almost all TIM-TOM subunits forming the mitochondrial translocon. The divergence of the TIC-TOC complex in some heterotrophs may result in accelerated evolution of TPs, thus, targeting predictions made based on characteristics of photosynthetic plants may not be fully accurate.

Second, uncertain protein models for some of our species may have led to true N-terminal sequences being absent or incomplete. Retargeting of cytosolic aaRS may occur through two different mechanisms; 1) alternative splicing or alternative transcription/translation start sites within one gene resulting in cytosolic and organellar isoforms and 2) duplicated genes within the genome encoding cytosolic and organellar paralogs. Our ability to investigate these two mechanisms was limited as we generally only had transcriptome or genome information but not both for each species. However, observations of transcriptomes assembled with consideration for alternative splice isoforms indicate that both mechanisms are potentially at play (Dataset S6). For example, *R. phalloides* has two separate paralogs for cytosolic ThrRS, one of which is predicted to be mitochondrial-targeted (Rphallo_DN1669_c0_g1) while the other is not (Rphallo_DN1669_c0_g2). Conversely, *R. phalloides* TyrRS is encoded by a single locus (Rphallo_DN204_c0_g1) with a multitude of splice isoforms, some of which are putatively mitochondrial-targeted. Genome sequencing and long-read mRNA sequencing would allow for more thorough analysis of which mechanism is most common in each species.

While almost all the TPs that we investigated with in vivo microscopy assays exhibited targeting to mitochondria and/or plastids, the specific organelle sometimes differed from the in silico prediction. The ProRS and TrpRS results in particular were unexpected as the in silico predictions were strongly mitochondrial (Dataset S3), but we observed clear plastid localization in vivo in two separate rounds of imaging with different vectors (Fig. 3 and *SI Appendix*, Figs. S4 and S5). There were other TPs where targeting predictions were only partly supported by GFP localization experiments, including GluRS and GlyRS which exhibited targeting to only one of the two predicted organelles, and AsnRS which was only predicted to have plastid localization but additionally exhibited mitochondrial localization in vivo.

These results may reflect the fact that target peptide prediction models like those used in this study do not have complete precision or sensitivity for a given subcellular compartment even when applied to benchmark datasets. This is not surprising, as changing just a few amino acids in a TP can alter protein targeting from singular to dual targeting in vivo (76). Furthermore, predictions made by a classifier trained on a broad sampling of TPs from diverse plant species cannot accurately predict every TP for an individual species, since species-specific features may not be represented enough to be detected by the model. Even if we assume our microscopy data are a perfect reflection of "true" targeting in the native system, we might expect some inaccuracies in our targeting predictions just based on the sensitivity and precision of the prediction algorithms (TargetP-2.0 and LOCALIZER). Of the two models, TargetP-2.0 specifically identifies mitochondrial and plastid TPs with 85 to 86% sensitivity and 87 to 90% precision and differentiates proteins lacking TPs with over 98% sensitivity and precision (77, 78). The publication describing LOCALIZER does not directly quantify the tool's ability to distinguish proteins with no TP but has 60 to 73% sensitivity and 60 to 80% precision for mitochondrial and plastid predictions respectively, counting proteins with predicted dual-targeting as positive hits for both plastid and mitochondria.

The artificial nature of heterologous expression in *N. benthamiana* can also create artifacts. We can see that in some cases predictions for a given TP can change when it is fused to GFP as opposed to the native aaRS sequence (Dataset S3). This implies that more than just the N-terminal extension of a protein determines targeting and may create some artifacts in assessing targeting based on TP-GFP fusions. Furthermore, alternative splicing or translation initiation sites can be a mechanism of alternative targeting (79–81). Alternative translation initiation may require RNA cis regulatory factors i.e., 5′ UTR sequences not captured in our heterologous expression system (82). Additionally, the import of a heterologously expressed TP-tagged GFP can change based on the age of leaves during infiltration, the efficiency of protein folding, protein expression levels, and truncation of an internal targeting sequence (83–85). As comparison of the in silico and heterologous expression assays is effectively combining the errors created by both methods, we would not expect perfect concordance between the methods regarding the specific identity of the retargeted organelle. Rather, these methods are complementary tools to assess general retargeting to the endosymbiotic organelles rather than differentiating between TPs specifically targeting mitochondria or plastids from those that are dual targeted. An additional method for future investigation of targeting to plastids or mitochondria would be to subfractionate tissue from one or more of these heterotrophic plant species and conduct comparative high-throughput proteomics on each fraction.

**Potential Rewiring of aaRS–tRNA Interactions.** Retargeting of cytosolic aaRS led us to investigate organelle genome tRNA content and whether there was substrate-enzyme mismatch in the organelles, i.e., instances where the cytosolic enzyme was interacting with an organellar tRNA or vice versa. Our interpretations presume that loss of organellar tRNA genes is compensated by import of the corresponding cytosolic tRNAs (9, 14, 86). We found that the retention of organellar aaRS in heterotrophs generally matched the retention of plastid tRNAs but not necessarily mitochondrial tRNAs (Fig. 4). This is further supported by observations in hemiparasite *Viscum album*, which has lost most of its mitochondrial tRNAs, yet retains most organellar aaRS in conjunction with a nearly full set of plastid tRNAs (32, 87, 88). Substrate-enzyme mismatch was rare within the plastid and usually involved loss of a plastid tRNA yet continued presence of the corresponding organellar aaRS (e.g., see *R. phalloides* where all plastid tRNAs are lost but a handful of organellar aaRS remain). There were also instances of organellar aaRS apparently charging imported cytosolic tRNAs in the mitochondria (e.g., *trnF* in *R. cantleyi and R. phalloides*). More surprisingly, however, there were cases where a mitochondrial tRNA was retained despite loss of the corresponding organellar aaRS and apparent retargeting of the cytosolic enzyme, particularly in the Rafflesiaceae. This type of mismatch is unexpected because cytosolic aaRS are thought to generally be more discriminating enzymes than their mitochondrial counterparts (8, 57, 89–91), making it more difficult to transition to a new tRNA substrate. However, our results suggest some cytosolic enzymes may be interacting with bacterial-type tRNAs in these heterotrophic species, although we cannot rule out the possibility that cytosolic tRNAs are being redundantly imported. Indeed, our findings raise fascinating questions about the process of functionally replacing tRNAs and interacting enzymes in these dynamic systems (57, 92), which could be explored with comparative and biochemical approaches in closely related species still undergoing these transitions. Given that parasitic plant lineages tend to be fast evolving across all three genomes (93), the cytosolic aaRS in these species may have evolved lower specificity and the capability to charge organellar tRNAs. Overall, our findings suggest that the plastid tRNA content, and not the mitochondrial tRNA content, drives retention of organellar aaRS, again placing the photosynthetic plastid as a central factor in the retention of functionally redundant translation machinery.

The identity of retained organellar aaRS versus those substituted by cytosolic counterparts was seemingly nonrandom. GluRS and PheRS were the two mostly highly retained organellar aaRS across these diverse lineages of plants (Fig. 1). Organellar PheRS is likely retained because the cytosolic PheRS is unusual in that it is a heteromeric complex with independently encoded α and β subunits (94). Therefore, functional replacement of the organellar PheRS would require two independent acquisitions of TPs, as well as stoichiometric import and assembly within the organelles. Such a coordinated set of changes may be unlikely to evolve (24, 57). The one exception was *E. aphyllum,* for which we could not find an organellar PheRS. However, this may be an artifact of low expression, as protein models for this species were curated from RNA-seq data. As $tRNA^{Glu}$ is the most frequently retained tRNA in heterotrophic plant plastomes due to its additional role in the first step of heme biosynthesis (56, 95, 96), it is not surprising that GluRS was among the most retained organellar aaRS in this study. Intriguingly, the *Rafflesiaceae* are the only species that lost the organellar GluRS, as well as plastid *trnE*. Yet, genes encoding heme synthesis enzymes are retained in these species (94), opening interesting questions into how they carry out the first step in the pathway. In *S. himalayana*, we found retargeting of the cytosolic GluRS to the plastid, which along with presumed import of the cytosolic $tRNA^{Glu}$, could functionally replace the first step in the heme biosynthesis pathway. This has the interesting implication that the glutamyl-tRNA reductase catalyzing the next step in heme biosynthesis is nondiscriminating toward cytosolic tRNAs. The organellar GluRS is also necessary for misaminoacylation of $tRNA^{Gln}$ prior to transamidation of Glu to Gln by the GatCAB complex (97). Cytosolic GlnRS is retargeted to the mitochondria in *S. himalayana*, thus allowing for direct aminoacylation of $tRNA^{Gln}$ assuming there is also import of cytosolic $tRNA^{Gln}$. This likely precipitated the loss of the GatCAB complex and negated the need for a nondiscriminating organellar GluRS.

At the other extreme, there was repeated loss of organellar AlaRS, which was absent from all heterotrophs and the hemiparasitic *Cuscuta* species, possibly because plastomes of heterotrophs are often AT-rich and thus have underrepresentation of amino acids with GC-rich codons like Ala, Gly, and Pro (51, 52). Thus, the corresponding organellar aaRS and tRNAs may be subject to decreased functional constraint and be less likely to be retained. This hypothesis is supported by the fact that we also see frequent loss of organellar GlyRS and ProRS in heterotrophs (Fig. 1).

Our results indicate that imported cytosolic aaRS and tRNAs are sufficient to sustain mitochondrial translation in plants, supporting the hypothesis that the selective pressure to keep organellar aaRS/tRNAs is maintained by translational demands associated with photosynthesis. Yet other eukaryote lineages that never bore plastids often still retain organelle-specific tRNAs and aaRS for their mitochondria (8, 9, 22, 24). This is even more puzzling when considering that various eukaryotic lineages have jettisoned nearly all their mitochondrial tRNAs/aaRS, including cnidarians, ctenophores, some fungi, and trypanosomes (12, 24, 98). One possible explanation is the logical parallel to our finding that photosynthesis governs plastid translation rates and maintains gene redundancy; eukaryotic lineages with high respiratory rates may require high mitochondrial gene expression and, therefore, retain mitochondria-specific translation machinery. Indeed, mitochondrial gene expression tracks well with respiratory rate and cellular activity across different tissues and metabolic states (99–101). This connection to cell energetics would also explain why mitochondrial tRNAs/aaRS are dispensable in parasitic eukaryotes like trypanosomes and heterotrophic plants, which

presumably have lower energy demands. By studying nonmodel plant species with unusual physiology and life histories, we were able to link plastid bioenergetic processes to the evolution of both organellar and nuclear genomes. The sequencing of mitogenomes and characterization of metabolism in nonmodel eukaryotes will provide exciting opportunities to conduct similar studies to better understand how mitochondrial metabolism effects gene redundancy and elucidate the principles governing endosymbiotic integration.

## Methods

See *SI Appendix* for more detailed methods.

**Data Acquisition and Ortholog Search.** Gene models for species used in the study were collected from publicly available data sources (Dataset S1). Orthofinder v2.5.5 (37) was run on all sets of protein models to find groups of likely orthologous proteins. Orthogroups containing proteins of interest including cytosolic/organellar aaRS, tRNA-processing enzymes, riboproteins, and organelle protein import machinery characterized in *A. thaliana* (Dataset S2) were extracted and checked for completeness via RBH against all protein models for a given species using BLASTP 2.15.0+ (102) and custom Python/Bash scripts. For CysRS, tRNase Z, and PRORP, sequences for both cytosolic and organellar proteins were clustered together in one orthogroup likely due to recent shared ancestry. For these genes, cytosolic vs. organellar versions were differentiated by building maximum likelihood trees (*SI Appendix*, Figs. S8–S10). using RAxML v 8.2.12 (103) after aligning sequences using MAFFT v7.525 (104) trimming with trimAl v1.5 (105). Once we had a curated set of protein sequences for all species, they were checked for completeness of length and subsequent visualization with a custom Python script. A "full-length" ortholog was defined as being at least 50% as long as the orthologous *A. thaliana* protein.

**In Silico Targeting Predictions.** All transcripts for cytosolic aaRS, a "positive control" set of genes involved in mitochondrial OXPHOS/plastid housekeeping and a "negative control" set of genes involved in cell wall synthesis were run-through targeting prediction programs LOCALIZER-1.0.5 (78) and TargetP-2.0 (77).

High confidence retargeting was defined as a > 0.5 probability/likelihood of targeting for both LOCALIZER and TargetP-2.0. Moderate confidence was defined as a > 0.5 probability/likelihood for one of the algorithms. Low confidence retargeting was defined as neither algorithm providing a probability/likelihood >0.5. For subsequent analysis of ratios (Fig. 2 and *SI Appendix*, Figs. S1 and S2), any ortholog with a "moderate" or better confidence was counted as retargeted. Note that *R. cantleyi* and *S. himalayana* proteins were excluded from predictions of targeting to plastids as these species likely lack plastomes. All analyses including plotting and statistics were carried out using custom Bash, Python, and R scripts available via Dryad (https://doi.org/10.5061/dryad.6hdr7sr5x).

**In Vivo Localization of *S. himalayana* aaRS TPs via Tobacco Infiltration and Confocal Microscopy.** TPs were chosen based on having an at least moderate probability (see above) of targeting a protein to either the mitochondria or plastid and/or having an N-terminal extension preceding conserved sequence shared with *A. thaliana*. We selected either the sequence predicted to be a TP by TargetP-2.0 plus 10 residues or the entire N-terminal overhang plus the first 10 residues aligning with the functional portion of the *A. thaliana* protein. These amino acid sequences were then converted into DNA sequences using EMBOSS Backtranseq (106) with the *Nicotiana tabacum* codon usage table (Dataset S3). These sequences were then used to generate constructs for *Agrobacterium* transformation using a redesigned dual-expression vector containing both the mitochondrial control IVD-FP611 and GFP fused to the C-terminus of our TP of interest. Some TPs were also originally cloned into the single-gene expression vector described in ref. 57. Transformation into *N. benthamiana* and confocal microscopy was constructed as described in ref. 57. Constructs are available from Addgene; see Dataset S3 for ID numbers.

**Mitogenome Assembly.** Prior to tRNA analysis, the mitogenomes of *S. himalayana* (SRR629601) and *R. cantleyi* (SRR629613) were assembled de novo from raw reads available from the NCBI SRA database. Reads were first trimmed using FASTP (107). SPAdes v.3.15.4 (108) was then used to assemble trimmed reads de novo.

**tRNA Analysis.** Excepting the Rafflesiaceae, for which we assembled mitogenomes de novo, the GenBank (.gbk) files for mitogenomes and plastomes from species of interest were downloaded from NCBI (Dataset S1). A custom Perl script was used to extract annotated tRNAs from the gbk files. tRNAscan-SE-2.0.9 (109, 110) was run on downloaded genomes (parameters -BQH), and results were cross-checked with preexisting GenBank annotations (Dataset S4). For *S. himalayana* and *R. cantleyi*, the same tRNAscan-SE-2.0.9 pipeline was used to find tRNA sequences. tRNA presence/absence patterns were plotted along with the corresponding aaRS presence/absence/retargeting data using a custom Python script. Scripts are available via Dryad (https://doi.org/10.5061/dryad.np5hqc009). The highly divergent and truncated trnE for *B. fungosa* was added based on prior publication (50).

**Rate Analysis.** Primary protein sequences for cytosolic and organellar aaRS from heterotroph *M. hypopitys* and related autotroph *R. williamsianum* from the original Orthofinder search were used for comparison of rates, with *A. thaliana* orthologs forming the outgroup. CysRS, LeuRS, ThrRS, and TrpRS were excluded from the analysis because the organellar ortholog was missing or too short to create informative gene trees in one or both test species. Predicted TPs were trimmed from sequences using output of TargetP-2.0. Sequences were then aligned using MAFFT v7.490 (104) and trimmed manually in MEGA 11 (111). Maximum likelihood trees were built using RAxML v 8.2.12 (103). The resulting trees were then evaluated for topology, see *SI Appendix, Supplemental Methods* for more information. Trees were rooted to the *A. thaliana* ortholog(s), then branch lengths to the last shared node from each *H. monotropa* and *R. williamsianum* ortholog were extracted from the trees. Then, the $\log_{10}$-transformed ratio of the *H. monotropa*/*R. williamsianum* branches was plotted, and mean branch lengths for cytosolic and organellar enzymes were calculated. Statistical difference in mean branch ratio in cytosolic vs. organellar enzymes was tested using a *t* test. Tajima's relative rate tests were run in MEGA 11 using *A. thaliana* as the outgroup. Data are available in Dataset S5. Scripts, alignments, and maximum likelihood trees are available via Dryad (https://doi.org/10.5061/dryad.tb2rbp05r).

**Data, Materials, and Software Availability.** Some study data are available and have been deposited in Dryad (112–115). The data can be accessed through the following DOIs: https://doi.org/10.5061/dryad.0cfxpnw7p, https://doi.org/10.5061/dryad.tb2rbp05r, https://doi.org/10.5061/dryad.np5hqc009, and https://doi.org/10.5061/dryad.6hdr7sr5x.

**ACKNOWLEDGMENTS.** We thank Jessica Warren for helpful discussion and feedback on this project and Anna Pratt for helpful advice on the cloning approach. We thank and acknowledge the contributions of undergraduate researchers Xiaorui Lou and Janna Novak. This work was supported by grants MCB-2322154, MCB-2048407, and MCB-1933590 from the NSF and utilized the Rocky Mountain Advanced Computing Consortium Summit supercomputer (NSF ACI-1532235 and ACI-1532236). R.A.D. was supported by an NSF Postdoctoral Research Fellowship in Biology (IOS-2208908). J.M.C. was supported by the University of Nebraska Foundation.

Author affiliations: [a]Department of Biology, Colorado State University, Fort Collins, CO 80523; [b]School of Biological Sciences, University of Nebraska-Lincoln, Lincoln, NE 68588; [c]Instituto de Biología Agrícola de Mendoza, Universidad Nacional de Cuyo, Consejo Nacional de Investigaciones Científicas y Técnicas, Facultad de Ciencias Agrarias, Chacras de Coria, Mendoza M5528AHB, Argentina; [d]Facultad de Ciencias Exactas y Naturales, Universidad Nacional de Cuyo, Ciudad de Mendoza, Mendoza M5502JMA, Argentina; and [e]Department of Agronomy and Horticulture, University of Nebraska-Lincoln, Lincoln, NE 68588

1. A. N. Khachane, K. N. Timmis, V. A. P. Martins Dos Santos, Dynamics of reductive genome evolution in mitochondria and obligate intracellular microbes. *Mol. Biol. Evol.* **24**, 449–456 (2007).
2. P. J. Keeling, J. D. Palmer, Horizontal gene transfer in eukaryotic evolution. *Nat. Rev. Genet.* **9**, 605–618 (2008).
3. W. Martin *et al.*, Evolutionary analysis of Arabidopsis, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 12246–12251 (2002).
4. E. C. M. Nowack *et al.*, Endosymbiotic gene transfer and transcriptional regulation of transferred genes in Paulinella chromatophora. *Mol. Biol. Evol.* **28**, 407–422 (2011).
5. N. A. Moran, J. P. McCutcheon, A. Nakabachi, Genomics and evolution of heritable bacterial symbionts. *Annu. Rev. Genet.* **42**, 165–190 (2008).
6. J. P. McCutcheon, N. A. Moran, Extreme genome reduction in symbiotic bacteria. *Nat. Rev. Microbiol.* **10**, 13–26 (2012).

7. V. Rajendran, P. Kalita, H. Shukla, A. Kumar, T. Tripathi, Aminoacyl-tRNA synthetases: Structure, function, and drug discovery. *Int. J. Biol. Macromol.* **111**, 400–414 (2018).
8. T. Salinas-Giegé, R. Giegé, P. Giegé, TRNA biology in mitochondria. *Int. J. Mol. Sci.* **16**, 4518–4559 (2015).
9. A. M. Duchêne, C. Pujol, L. Maréchal-Drouard, Import of tRNAs and aminoacyl-tRNA synthetases into mitochondria. *Curr. Genet.* **55**, 1–18 (2009).
10. L. Bonnefond *et al.*, Toward the full set of human mitochondrial aminoacyl-tRNA synthetases: Characterization of AspRS and TyrRS. *Biochemistry* **44**, 4805–4816 (2005).
11. E. Tolkunova, H. Park, J. Xia, M. P. King, E. Davidson, 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 (2000).
12. S. Shikha, R. Brogli, A. Schneider, N. Polacek, TRNA biology in trypanosomes. *Chimia (Aarau)* **73**, 395–405 (2019).
13. T. H. P. Tan, R. Pach, A. Crausaz, A. Ivens, A. Schneider, tRNAs in Trypanosoma brucei: Genomic organization, expression, and mitochondrial import. *Mol. Cell Biol.* **22**, 3707–3717 (2002).
14. J. M. Warren *et al.*, Rapid shifts in mitochondrial tRNA import in a plant lineage with extensive mitochondrial tRNA gene loss. *Mol. Biol. Evol.* **38**, 5735–5751 (2021).
15. J. M. Warren, D. B. Sloan, Interchangeable parts: The evolutionarily dynamic tRNA population in plant mitochondria. *Mitochondrion* **52**, 144–156 (2020).
16. A.-M. Duchêne *et al.*, Dual targeting is the rule for organellar aminoacyl-tRNA synthetases in Arabidopsis thaliana. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 16484–16489 (2005).
17. K. Fujishima, A. Kanai, tRNA gene diversity in the three domains of life. *Front. Genet.* **5**, 142 (2014).
18. M. Lynch, B. Koskella, S. Schaack, Mutation pressure and the evolution of organelle genomic architecture. *Science* **1979**, 1727–1730 (2006).
19. S. Schaack, E. K. H. Ho, F. MacRae, Disentangling the intertwined roles of mutation, selection and drift in the mitochondrial genome. *Philos. Trans. R. Soc. Lond. B: Biol. Sci.* **375**, 20190173 (2020).
20. K. Crosby, D. R. Smith, Does the mode of plastid inheritance influence plastid genome architecture? *PLoS One* **7**, e46260 (2012).
21. K. H. Wolfe, W.-H. Li, P. M. Sharp, Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proc. Natl. Acad. Sci. U.S.A.* **84**, 9054–9058 (1987).
22. J. R. Adrion, P. S. White, K. L. Montooth, The roles of compensatory evolution and constraint in aminoacyl trna synthetase evolution. *Mol. Biol. Evol.* **33**, 152–161 (2016).
23. F. S. Barreto *et al.*, Genomic signatures of mitonuclear coevolution across populations of Tigriopus californicus. *Nat. Ecol. Evol.* **2**, 1250–1257 (2018).
24. W. Pett, D. V. Lavrov, Cytonuclear interactions in the evolution of animal mitochondrial tRNA metabolism. *Genome Biol. Evol.* **7**, 2089–2101 (2015).
25. D. B. Sloan, R. A. DeTar, J. M. Warren, Aminoacyl-tRNA synthetase evolution within the dynamic tripartite translation system of plant cells. *Genome Biol. Evol.* **15**, evad050 (2023).
26. S. Wicke, G. M. Schneeweiss, C. W. dePamphilis, K. F. Müller, D. Quandt, The evolution of the plastid chromosome in land plants: Gene content, gene order, gene function. *Plant Mol. Biol.* **76**, 273–297 (2011).
27. M. Michaud, V. Cognat, A. M. Duchêne, L. Maréchal-Drouard, A global picture of tRNA genes in plant genomes. *Plant J.* **66**, 80–93 (2011).
28. E. S. Forsythe *et al.*, Organellar transcripts dominate the cellular mRNA pool across plants of varying ploidy levels. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2204187119 (2022).
29. J. Mergner *et al.*, Mass-spectrometry-based draft of the Arabidopsis proteome. *Nature* **579**, 409–414 (2020).
30. J. Těšitel, Functional biology of parasitic plants: A review. *Plant Ecol. Evol.* **149**, 5–20 (2016).
31. L. Teixeira-Costa, C. C. Davis, Life history, diversity, and distribution in parasitic flowering plants. *Plant Physiol.* **187**, 32–51 (2021).
32. S. Wicke, J. Naumann, Molecular evolution of plastid genomes in parasitic flowering plants. *Adv. Bot. Res.* **85**, 315–347 (2018).
33. D. R. Smith, S. R. Asmail, Next-generation sequencing data suggest that certain nonphotosynthetic green plants have lost their plastid genomes. *N. Phytol.* **204**, 7–11 (2014).
34. H. T. Li *et al.*, Plastid phylogenomic insights into relationships of all flowering plant families. *BMC Biol.* **19**, 232 (2021).
35. J. Molina *et al.*, Possible loss of the chloroplast genome in the parasitic flowering plant Rafflesia lagascae (Rafflesiaceae). *Mol. Biol. Evol.* **31**, 793–803 (2014).
36. D. L. Nickrent, Y. Ouyang, R. J. Duff, C. W. Depamphilis, Do nonasterid holoparasitic flowering plants have plastid genomes? *Plant Mol. Biol.* **34**, 717–729 (1997).
37. D. M. Emms, S. Kelly, OrthoFinder: Phylogenetic orthology inference for comparative genomics. *Genome Biol.* **20**, 238 (2019).
38. D. M. Emms, S. Kelly, OrthoFinder: Solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol.* **16**, 157 (2015).
39. M. I. Schelkunov, M. S. Nuraliev, M. D. Logacheva, Genomic comparison of non-photosynthetic plants from the family Balanophoraceae with their photosynthetic relatives. *PeerJ* **9**, e12106 (2021).
40. X. Chen *et al.*, Balanophora genomes display massively convergent evolution with other extreme holoparasites and provide novel insights into parasite–host interactions. *Nat. Plants* **9**, 1627–1642 (2023), 10.1038/s41477-023-01517-7.
41. L. Cai *et al.*, Deeply altered genome architecture in the endoparasitic flowering plant Sapria himalayana Griff. (Rafflesiaceae). *Curr. Biol.* **31**, 1002–1011.e9 (2021).
42. X. Guo *et al.*, The *Sapria himalayana* genome provides new insights into the lifestyle of endoparasitic plants. *BMC Biol.* **21**, 134 (2023).
43. G. Canino *et al.*, Arabidopsis encodes four tRNase Z enzymes. *Plant Physiol.* **150**, 1494–1502 (2009).
44. B. Gutmann, A. Gobert, P. Giegé, PRORP proteins support RNase P activity in both organelles and the nucleus in Arabidopsis. *Genes Dev.* **26**, 1022–1027 (2012).
45. A. Gobert *et al.*, A single Arabidopsis organellar protein has RNase P activity. *Nat. Struct. Mol. Biol.* **17**, 740–744 (2010).
46. C. Pujol *et al.*, 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 (2008).
47. A. Nagao, T. Suzuki, T. Katoh, Y. Sakaguchi, T. Suzuki, Biogenesis of glutaminyl-mt tRNA Gln in human mitochondria. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 16209–16214 (2009).
48. L. Echevarría *et al.*, Glutamyl-tRNAGln amidotransferase is essential for mammalian mitochondrial translation in vivo. *Biochem. J.* **460**, 91–101 (2014).
49. R. Giegé, G. Eriani, The tRNA identity landscape for aminoacylation and beyond. *Nucleic Acids Res.* **51**, 1528–1570 (2023).
50. L. F. Ceriotti, M. E. Roulet, M. V. Sanchez-Puerta, Plastomes in the holoparasitic family Balanophoraceae: Extremely high AT content, severe gene content reduction, and two independent genetic code changes. *Mol. Phylogenet. Evol.* **162**, 107208 (2021).
51. H. J. Su *et al.*, Novel genetic code and record-setting AT-richness in the highly reduced plastid genome of the holoparasitic plant Balanophora. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 934–943 (2019).
52. M. I. Schelkunov, M. S. Nuraliev, M. D. Logacheva, Rhopalocnemis phalloides has one of the most reduced and mutated plastid genomes known. *PeerJ* **7**, e7500 (2019).
53. X. Chen *et al.*, Comparative plastome analysis of root-and stem-feeding parasites of santalales untangle the footprints of feeding mode and lifestyle transitions. *Genome Biol. Evol.* **12**, 3663–3676 (2019).
54. M. D. Logacheva, M. I. Schelkunov, V. Y. Shtratnikova, M. V. Matveeva, A. A. Penin, Comparative analysis of plastid genomes of non-photosynthetic Ericaceae and their photosynthetic relatives. *Sci. Rep.* **6**, 30042 (2016).
55. L. F. Ceriotti, L. Gatica-Soria, M. V. Sanchez-Puerta, Cytonuclear coevolution in a holoparasitic plant with highly disparate organellar genomes. *Plant Mol. Biol.* **109**, 673–688 (2022).
56. A. C. Barbrook, C. J. Howe, S. Purton, Why are plastid genomes retained in non-photosynthetic organisms? *Trends Plant. Sci.* **11**, 101–108 (2006).
57. J. M. Warren *et al.*, Rewiring of aminoacyl-tRNA synthetase localization and interactions in plants with extensive mitochondrial tRNA gene loss. *Mol. Biol. Evol.* **40**, msad163 (2023).
58. B. Bölter, En route into chloroplasts: Preproteins' way home. *Photosynth. Res.* **138**, 263–275 (2018).
59. P. Chotewutmontri, K. Holbrook, B. D. Bruce, Plastid protein targeting: Preprotein recognition and translocation. *Int. Rev. Cell Mol. Biol.* **330**, 227–294 (2017).
60. M. W. Murcha *et al.*, Protein import into plant mitochondria: Signals, machinery, processing, and regulation. *J. Exp. Bot.* **65**, 6301–6335 (2014).
61. R. W. Christian, S. L. Hewitt, G. Nelson, E. H. Roalson, A. Dhingra, 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 (2020).
62. S. Reinbothe *et al.*, tRNA-dependent import of a transit sequence-less aminoacyl-tRNA synthetase (Leurs2) into the mitochondria of Arabidopsis. *Int. J. Mol. Sci.* **22**, 3808 (2021).
63. W. Neupert, J. M. Herrmann, Translocation of proteins into mitochondria. *Annu. Rev. Biochem.* **76**, 723–749 (2007).
64. M. H. Chen, L. F. Huang, H. M. Li, Y. R. Chen, S. M. Yu, Signal peptide-dependent targeting of a rice α-amylase and cargo proteins to plastids and extracellular compartments of plant cells. *Plant Physiol.* **135**, 1367–1377 (2004).
65. L. G. L. Richardson, D. J. Schnell, Origins, function, and regulation of the TOC-TIC general protein import machinery of plastids. *J. Exp. Bot.* **71**, 1226–1238 (2020).
66. H. Inoue, C. Rounds, D. J. Schnell, The molecular basis for distinct pathways for protein import into Arabidopsis chloroplasts. *Plant Cell* **22**, 1947–1960 (2010).
67. S. Bischof *et al.*, Plastid proteome assembly without Toc159: Photosynthetic protein import and accumulation of N-Acetylated plastid precursor proteins. *Plant Cell* **23**, 3911–3928 (2011).
68. Y. L. Chen *et al.*, TIC236 links the outer and inner membrane translocons of the chloroplast. *Nature* **564**, 125–129 (2018).
69. A. R. Kasmati *et al.*, Evolutionary, molecular and genetic analyses of Tic22 Homologues in Arabidopsis thaliana chloroplasts. *PLoS One* **8**, e63863 (2013).
70. J. Tripp *et al.*, Structure and conservation of the periplasmic targeting factor Tic22 protein from plants and cyanobacteria. *J. Biol. Chem.* **287**, 24164–24173 (2012).
71. M. Rudolf *et al.*, In vivo function of Tic22, a protein import component of the intermembrane space of chloroplasts. *Mol. Plant* **6**, 817–829 (2013).
72. A. R. Kasmati, M. Töpel, R. Patel, G. Murtaza, P. Jarvis, Molecular and genetic analyses of Tic20 homologues in Arabidopsis thaliana chloroplasts. *Plant J.* **66**, 877–889 (2011).
73. E. Kovács-Bogdán, J. P. Benz, J. Soll, B. Bölter, Tic20 forms a channel independent of Tic110 in chloroplasts. *BMC Plant Biol.* **11**, 133 (2011).
74. H. Liu, A. Li, J. D. Rochaix, Z. Liu, Architecture of chloroplast TOC–TIC translocon supercomplex. *Nature* **615**, 349–357 (2023).
75. X. Guo, H. Wang, D. Lin, Y. Wang, X. Jin, Cytonuclear evolution in fully heterotrophic plants: Lifestyles and gene function determine scenarios. *BMC Plant Biol.* **24**, 989 (2024).
76. C. Carrie, J. Whelan, Widespread dual targeting of proteins in land plants: When, where, how and why. *Plant Signal. Behav.* **8**, e25034 (2013).
77. J. J. A. Armenteros *et al.*, Detecting sequence signals in targeting peptides using deep learning. *Life Sci. Alliance* **2**, e201900429 (2019).
78. J. Sperschneider *et al.*, LOCALIZER: Subcellular localization prediction of both plant and effector proteins in the plant cell. *Sci. Rep.* **7**, 44598 (2017).
79. M. Sharma, B. Bennewitz, R. B. Klösgen, Rather rule than exception? How to evaluate the relevance of dual protein targeting to mitochondria and chloroplasts. *Photosynth. Res.* **138**, 335–343 (2018).
80. L. Xu, C. Carrie, S. R. Law, M. W. Murcha, J. Whelan, Acquisition, conservation, and loss of dual-targeted proteins in land plants. *Plant Physiol.* **161**, 644–662 (2013).
81. D. B. Sloan, J. M. Warren, A. M. Williams, S. A. Kuster, E. S. Forsythe, Incompatibility and interchangeability in molecular evolution. *Genome Biol. Evol.* **15**, evac184 (2023).
82. Y. R. Li, M. J. Liu, Prevalence of alternative AUG and non-AUG translation initiators and their regulatory effects across plants. *Genome Res.* **30**, 1418–1433 (2020).
83. S. K. Tanz, I. Castleden, I. D. Small, A. Harvey Millar, Fluorescent protein tagging as a tool to define the subcellular distribution of proteins in plants. *Front. Plant Sci.* **4**, 214 (2013).
84. J. Jeong, B. Moon, I. Hwang, D. W. Lee, GREEN FLUORESCENT PROTEIN variants with enhanced folding are more efficiently imported into chloroplasts. *Plant Physiol.* **190**, 238–249 (2022).
85. M. Sharma, B. Bennewitz, R. B. Klösgen, Dual or not dual?–Comparative analysis of fluorescence microscopy-based approaches to study organelle targeting specificity of nuclear-encoded plant proteins. *Front. Plant Sci.* **9**, 1350 (2018).
86. M. W. Murcha *et al.*, Plant-specific preprotein and amino acid transporter proteins are required for tRNA import into mitochondria. *Plant Physiol.* **172**, 2471–2490 (2016).
87. E. Skippington, T. J. Barkman, D. W. Rice, J. D. Palmer, 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–E3524 (2015).
88. G. Petersen, A. Cuenca, I. M. Møller, O. Seberg, Massive gene loss in mistletoe (Viscum, Viscaceae) mitochondria. *Sci. Rep.* **5**, 17588 (2015).

89. A. Fender *et al.*, Loss of a primordial identity element for a mammalian mitochondrial aminoacylation system. *J. Biol. Chem.* **281**, 15980–15986 (2006).
90. Y. Kumazawa, H. Himeno, K.-I. Miura, K. Watanabe, Unilateral aminoacylation specificity between bovine mitochondria and eubacteria. *J. Biochem.* **109**, 421–427 (1991).
91. L. Bonnefond, M. Frugier, R. Giegé, J. Rudinger-Thirion, Human mitochondrial TyrRS disobeys the tyrosine identity rules. *RNA* **11**, 558–562 (2005).
92. I. Small *et al.*, The strange evolutionary history of plant mitochondrial tRNAs and their aminoacyl-tRNA synthetases. *J. Hered.* **90**, 333–337 (1999).
93. L. Bromham, P. F. Cowman, R. Lanfear, Parasitic plants have increased rates of molecular evolution across all three genomes. *BMC Evol. Biol.* **13**, 126 (2013).
94. M. Safro, N. Moor, O. Lavrik, "Phenylalanyl-tRNA synthetases" in *Madame Curie Bioscience Database* (Landes Bioscience, 2013).
95. R. Tanaka, K. Kobayashi, T. Masuda, Tetrapyrrole metabolism in Arabidopsis thaliana. *Arabidopsis Book* **9**, e0145 (2011), 10.1199/tab.0145.
96. A. Schon *et al.*, The RNA required in the first step of chlorophyll biosynthesis is a chloroplast glutamate transfer-RNA. *Nature* **322**, 281–284 (1986).
97. A. Schon, C. G. Kannangara, S. Gough, D. Sollt, Protein biosynthesis in organelles requires misaminoacylation of tRNA. *Nature* **331**, 187–190 (1988).
98. K. M. Haen, W. Pett, D. V. Lavrov, Parallel loss of nuclear-encoded mitochondrial aminoacyl-tRNA synthetases and mtDNA-encoded tRNAs in cnidaria. *Mol. Biol. Evol.* **27**, 2216–2219 (2010).
99. I. Martin, O. Vinas, T. Mampel, F. Villarroya, Effects of cold environment on mitochondrial genome expression in the rat: Evidence for a tissue-specific increase in the liver, independent of changes in mitochondrial gene abundance. *Biochem. J.* **296**, 231–234 (1993).
100. R. F. Hevner, M. T. T. Wong-Riley, Mitochondrial and nuclear gene expression for cytochrome oxidase subunits are disproportionately regulated by functional activity in neurons. *J. Neurosci.* **13**, 1805–1819 (1993).
101. E. A. Amiott, J. A. Jaehning, Mitochondrial transcription is regulated via an ATP "sensing" mechanism that couples RNA abundance to respiration. *Mol. Cell* **22**, 329–338 (2006).
102. C. Camacho *et al.*, BLAST+: Architecture and applications. *BMC Bioinf.* **10**, 421 (2009).
103. A. Stamatakis, RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).
104. K. Katoh, K. Misawa, K.-I. Kuma, T. Miyata, MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **30**, 3059–3066 (2002).
105. S. Capella-Gutiérrez, J. M. Silla-Martínez, T. Gabaldón, trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972–1973 (2009).
106. F. Madeira *et al.*, Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Res.* **50**, W276–W279 (2022).
107. S. Chen, Y. Zhou, Y. Chen, J. Gu, Fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884–i890 (2018).
108. A. Bankevich *et al.*, SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **19**, 455–477 (2012).
109. P. P. Chan, T. M. Lowe, tRNAscan-SE: Searching for tRNA genes in genomic sequences. *Methods Mol. Biol.* **1962**, 1–14 (2019).
110. P. P. Chan, B. Y. Lin, A. J. Mak, T. M. Lowe, TRNAscan-SE 2.0: Improved detection and functional classification of transfer RNA genes. *Nucleic Acids Res.* **49**, 9077–9096 (2021).
111. K. Tamura, G. Stecher, S. Kumar, MEGA11: Molecular evolutionary genetics analysis version 11. *Mol. Biol. Evol.* **38**, 3022–3027 (2021).
112. R. A. DeTar, Dataset and code "Finding orthologs for aminoacyl tRNA synthetases in parasitic plants." Dryad. https://doi.org/10.5061/dryad.0cfxpnw7p. Updated version deposited on 18 October 2024.
113. R. A. DeTar, Dataset and code "Evolutionary rate analysis of aminoacyl tRNA synthetases." Dryad. https://doi.org/10.5061/dryad.tb2rbp05r. Deposited 2 August 2023.
114. R. A. DeTar, Dataset and code "Organellar tRNAs in parasitic plant species." Dryad. https://doi.org/10.5061/dryad.np5hqc009. Updated version deposited on 18 October 2024.
115. R. A. DeTar, Dataset and code "In silico subcellular targeting predictions for cytosolic aminoacyl tRNA-synthetases (aaRS) in parasitic plants." Dryad. https://doi.org/10.5061/dryad.6hdr7sr5x. Updated version deposited on 18 October 2024.