Evolutionary constraints on the plastid tRNA set decoding methionine and isoleucine

Sibah Alkatib, Tobias T. Fleischmann, Lars B. Scharff and Ralph Bock*

Max-Planck-Institut für Molekulare Pflanzenphysiologie, Am Mühlenberg 1, D-14476 Potsdam-Golm, Germany

ABSTRACT

The plastid (chloroplast) genomes of seed plants typically encode 30 tRNAs. Employing wobble and superwobble mechanisms, most codon boxes are read by only one or two tRNA species. The reduced set of plastid tRNAs follows the evolutionary trend of organellar genomes to shrink in size and coding capacity. A notable exception is the AUN codon box specifying methionine and isoleucine, which is decoded by four tRNA species in nearly all seed plants. However, three of these four tRNA genes were lost from the genomes of some parasitic plastid-containing lineages, possibly suggesting that less than four tRNA species could be sufficient to decode the triplets in the AUN box. To test this hypothesis, we have performed knockout experiments for the four AUN-decoding tRNAs in tobacco (_Nicotiana tabacum_) plastids. We find that all four tRNA genes are essential under both autotrophic and heterotrophic growth conditions, possibly suggesting tRNA import into plastids of parasitic plastid-bearing species. Phylogenetic analysis of the four plastid tRNA genes reveals striking conservation of all those bacterial features that are involved in discrimination between the different tRNA species containing CAU anticodons.

INTRODUCTION

Sixty-one triplets of the genetic code specify 20 different amino acids. Due to relaxed base pairing (referred to as wobbling; 1) between the third position of the codon in the messenger RNA (mRNA) and the first position of the anticodon in the transfer RNA (tRNA), less than 61 tRNA species are sufficient to decode all 61 triplets. According to the wobble rules, 32 tRNA species should represent the minimum set (1).

The plastid genome (plastome) of most vascular plants harbors only 30 tRNA genes (2–4). Nine of the 16 codon boxes are decoded by two tRNA species—the expected number according to the wobble rules (Figure 1). (A codon box is defined as the set of four triplets sharing the first two nucleotides.) One codon box, the UUN box, is served by three tRNA species (Figure 1). Five codon boxes are read by only one tRNA. Four of these are family codon boxes, in which all four triplets specify the same amino acid and the fifth box contains two tyrosine codons in addition to two stop codons (UAN box; Figure 1). Knockout studies of the two plastid tRNAs for glycine in the model plant tobacco (_Nicotiana tabacum_) have demonstrated that the tRNA^{Gly}(UCC) can read all four glycine codons (GGN) using the superwobble mechanism (5). Superwobbling is based on the capability of an unmodified U in the wobble position of the anticodon of the tRNA to base pair with all four nucleotides in the third codon position of a family box (1,5). It provides a likely mechanistic explanation for all four cases, in which a family box is served by only a single tRNA gene in the plastome (Figure 1). If the superwobble mechanism would be rigorously employed, only 23 tRNA species would be needed. This is significantly less than the 30 tRNAs encoded by the plastome of most vascular plants and may suggest that there is still some leeway for further reduction of the tRNA gene number in the genome.

Interestingly, only a single codon box is read by four tRNAs: the AUN box (Figure 1). It specifies the amino acids methionine and isoleucine. When serving as a start codon for translation initiation, the AUG methionine codon is recognized by the tRNA for _N_-formyl methionine, tRNA^{fMet}(CAU), also referred to as the initiator tRNA. Internal AUG codons are read by tRNA^{Met}(CAU), the so-called elongator tRNA. Whether or not the initiator and elongator tRNAs can partially replace each other in plastids is currently unknown.

The remaining three triplets in the AUN box specify isoleucine. Although a tRNA^{Ile}(UAU) would theoretically be sufficient to read all three isoleucine codons by superwobbling, this mechanism cannot be employed because it would misread AUG codons by conventional wobbling. Elaborate strategies have evolved to allow

*To whom correspondence should be addressed. Tel: +49 331 567 8700; Fax: +49 331 567 8701; Email: rbock@mpimp-golm.mpg.de

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

Figure 1. Degeneracy of the genetic code in plastids. Shown are the 16 codon boxes and the amino acids specified by the codons in each box. White codon boxes are served by a single tRNA species, light gray boxes by two tRNAs, dark gray boxes by three tRNAs and the black box by four tRNAs. The tRNA species are indicated by the triplet their anticodon has a perfect match with.

discrimination between AUG methionine and AUA isoleucine codons at high accuracy. In most bacteria and also in plastids, the AUA codon is read by $tRNA^{Ile}(CAU)$, whose cytidine in the wobble position of the anticodon is post-transcriptionally modified to lysidine (6–10). Lysidination converts the CAU methionine anticodon into an isoleucine anticodon and effectively prevents reading of AUG triplets by $tRNA^{Ile}(CAU)$. The enzyme responsible for this modification, TilS, is essential in *Escherichia coli* (11), but has not yet been identified in plants. In *E. coli*, the specificity of decoding in the AUN box is further enhanced by another tRNA modification: acetylation of the wobble nucleotide C_{34} in the anticodon of $tRNA^{Met}(CAU)$. This modification is thought to prevent misreading of AUA isoleucine codons (12). The other two codons for isoleucine (AUU and AUC) are recognized by $tRNA^{Ile}(GAU)$ using the wobble mechanism.

Interestingly, some non-photosynthetic plastid-bearing parasites have lost one of the two *trnI* genes from their plastid genome. Gene loss concerns either *trnI-CAU* or *trnI-GAU*, but never both $tRNA^{Ile}$ species (13). This could suggest that the two $tRNA^{Ile}$ species may be able to partially replace each other. In this work, we wanted to test this hypothesis experimentally.

Remarkably, *Rhizanthella gardneri*, a parasitic orchid, has not only lost the *trnI-GAU* gene from its plastome, but also the *trnM-CAU* gene (14). Two theoretically possible scenarios could make the elongator $tRNA^{Met}(CAU)$ dispensable: (i) utilization of the $tRNA^{fMet}(CAU)$ as both initiator and elongator tRNA, or (ii) incomplete lysidine modification of the $tRNA^{Ile}(CAU)$. An unmodified fraction of this tRNA could serve as $tRNA^{Met}(CAU)$ (10,11) and, in this way, partially complement the loss of the elongator tRNA for methionine. Such a mechanism is likely to be less efficient than retention of a separate gene for $tRNA^{Met}(CAU)$, but as most chloroplast

genome-encoded gene products are involved in photosynthesis, a heterotrophic (parasitic) lifestyle greatly lowers the demand for plastid translational capacity and, therefore, could tolerate some imperfection in aminoacylation and/or decoding. Finally, it is also theoretically possible that the missing tRNA species are imported from the cytosol, although there is no experimental evidence for the existence of a plastid tRNA import pathway in any species.

To distinguish between these possible explanations for tRNA gene loss, we decided to test all tRNA genes in the AUN box for essentiality employing a reverse genetics approach in the model plant tobacco, a species in which transformation of the plastid genome is feasible (15,16). Plastid protein biosynthesis is essential in tobacco (17–19), but the plants can tolerate a dramatic reduction in translational activity, especially when grown heterotrophically on sucrose-containing synthetic medium (20,21). Therefore, tobacco plants growing heterotrophically on an exogenous carbon source should mimic non-photosynthetic parasitic plants that obtain organic carbon from their host plants.

MATERIALS AND METHODS

Plant material, growth conditions and phenotypic assays

To generate leaf material for chloroplast transformation experiments, tobacco (*Nicotiana tabacum* cv. Petit Havana) plants were raised under aseptic conditions on agar-solidified Murashige and Skoog medium (MS medium) containing 30 g/l sucrose (22). Transplastomic lines were rooted and propagated on MS medium additionally supplemented with 500 mg/l spectinomycin. For seed production and analysis of plant phenotypes, transplastomic plants were grown in soil under standard greenhouse conditions (relative humidity 55%, day temperature 25°C, night temperature 20°C, diurnal cycle 16 h light and 8 h darkness, light intensity 300–600 µE m^{-2}s^{-1}). Inheritance patterns and seedling phenotypes were analyzed by germination of surface-sterilized seeds on MS medium with spectinomycin (500 mg/l).

Construction of plastid transformation vectors

Vectors for the targeted knockout of *trnM*, *trnI-CAU* and *trnI-GAU* were constructed by inserting an *aadA* cassette (comprising the *aadA* coding region from *E. coli*, the *psbA* promoter + 5′ UTR and the *rbcL* 3′ UTR from *Chlamydomonas reinhardtii*; 21,23) into the coding region of the respective tRNA gene. The *trnfM* gene was deleted and replaced by the *aadA* cassette. All plasmids used for plastid transformation were verified by restriction analysis and DNA sequencing.

For construction of vector pΔ*trnfM*, the genomic region surrounding the *trnfM* gene (corresponding to nucleotide positions 36 840–38 894 of the *N. tabacum* plastid genome; GenBank accession number NC_001879) was isolated from a previously constructed plasmid clone (5) as a 2-kb PstI/EcoRI fragment and inserted into the cloning vector pBS KS+ digested with the same enzymes. To replace the *trnfM* coding region with

the *aadA* cassette, a PCR strategy was used. PCR amplification was performed with primers P5′trnfM and P3′trnfM (Supplementary Table S1) using the plasmid pLS1 (21) as template and introducing the recognition sequences for the restriction enzymes BsrGI and HpaI with the primer sequences. The resulting amplification product was treated with BsrGI and HpaI and cloned into the PstI/EcoRI plasmid clone digested with the same enzymes, generating plastid transformation vector pΔ*trnfM*.

To construct vectors for targeted inactivation of the *trnM* gene (pΔ*trnM*-s and pΔ*trnM*-as), the *trnM*-containing region in the *N. tabacum* plastid genome (corresponding to nucleotide positions 53 613–55 934) was excised from a cloned SalI plastid DNA fragment with the restriction enzyme BamHI and inserted into cloning vector pUC18 cut with the same enzyme. This resulting plasmid clone was partially digested with ScaI to linearize it within the *trnM* gene and ligated to the *aadA* cassette excised from pLS1 (21) with SmaI. Clones for both orientations of the *aadA* cassette were identified and designated pΔ*trnM*-s (harboring the *aadA* cassette in the same transcriptional orientation as the *trnM* gene; s: sense) and pΔ*trnM*-as (containing the *aadA* in the opposite orientation; as: antisense).

To produce a vector for targeted inactivation of the *trnI-CAU* gene in the plastid genome, the previously constructed plasmid clone pSA10 (21) carrying the *trnI-CAU* genomic region (corresponding to nucleotide position 88 162–89 596) was linearized with XcmI, then treated with T4 DNA polymerase to generate blunt ends and ligated to the *aadA* cassette excised from pLS1 with SmaI. A clone carrying the *aadA* cassette in the same transcriptional orientation as *trnI-CAU* yielded plastid transformation vector pΔ*trnI-GAU*.

For construction of the vector pΔ*trnI-GAU*, the genomic region surrounding the *trnI-GAU* gene was excised from a cloned plastid DNA fragment with the restriction enzymes NheI and AatII. The fragment was treated with T4 DNA polymerase to generate blunt ends (corresponding to nucleotide positions 135 489–138 726) and inserted into the cloning vector pUC18 digested with HincII. The resulting plasmid clone was digested with BssHI and blunted by treatment with T4 DNA polymerase. This cloning procedure resulted in a deletion of 40 bp from the intron of *trnI-GAU*. Finally, the *aadA* cassette excised from pLS1 with SmaI was inserted into the disrupted *trnI-GAU* gene, generating transformation vector pΔ*trnI-GAU*.

Plastid transformation and selection of transplastomic lines

Young leaves from aseptically grown tobacco plants were bombarded with plasmid-coated 0.6-μm gold particles using a helium-driven biolistic gun (PDS1000He; BioRad). Primary spectinomycin-resistant lines were selected on plant regeneration medium containing 500 mg/l spectinomycin (23). Spontaneous spectinomycin-resistant plants were eliminated by double selection tests on medium containing both spectinomycin and streptomycin (500 mg/l each; 24). Several independent transplastomic lines were produced for each construct and subsequently subjected to three to four additional rounds of regeneration on spectinomycin-containing plant regeneration medium to enrich the transplastome and select for homoplasmy (19).

Isolation of nucleic acids and hybridization procedures

Total plant DNA was extracted from plants grown under spectinomycin selection *in vitro* by a cetyltrimethylammoniumbromide-based method (25). For RFLP analysis, DNA samples were treated with restriction enzymes, separated in 1% agarose gels and blotted onto Hybond N nylon membranes (GE Healthcare). For hybridization, $[\alpha^{32}P]dCTP$-labeled probes were produced by random priming (Multiprime DNA labeling kit; GE Healthcare). Prior to labeling, DNA fragments were purified by agarose gel electrophoresis followed by extraction from excised gel slices using the NucleoSpin Extract II kit (Macherey-Nagel). Hybridizations were performed at 65°C using standard protocols. The probes for RFLP analysis of putative transplastomic lines were prepared from restriction fragments or amplified PCR products. For analysis of the Δ*trnM* lines, the probe corresponds to nucleotide positions 37 161–37 968 of the tobacco plastid genome (NC_001879) and was excised from pΔ*trnM* with the enzymes EcoRI and SpeI. For the Δ*trnfM* lines, the probe corresponds to positions 37 663–38 076 and was prepared by PCR amplification with primers P5_pSA4F and P3_pSA4R. A probe for analysis of Δ*trnI-GAU* lines was prepared by PCR amplification with primers P5_23SF and P3_23SR (Supplementary Table S1) and covers the genomic region from nucleotide positions 133 576–136 272. For RFLP analysis of Δ*trnI-CAU* lines, the plastid genome sequence from positions 87 523–88 728 was excised from plasmid Δ*trnI-CAU* with EcoRI.

Analysis of tRNA sequences

For the analysis of the different tRNA genes, the following plastome sequences were used: *Nicotiana tabacum* (NC_001879), *Arabidopsis thaliana* (NC_000932), *Oryza sativa* cv. japonica (NC_001320), *Physcomitrella patens* (NC_005087) and *Chlamydomonas reinhardtii* (NC_005353). The gene IDs for the *trnM-CAU* sequences are: 800482 in *Nicotiana tabacum*, 1466257 in *Arabidopsis thaliana*, 3131357 in *Oryza sativa*, 2546701 in *Physcomitrella patens,* and 2716991 in *Chlamydomonas reinhardtii* (*trnM2 = trnM*). The gene IDs for *trnfM-CAU* are: 800479 in *N. tabacum*, 1466262 in *A. thaliana*, 3131344 in *O. sativa*, 2546784 in *P. patens* and 2716965 in *C. reinhardtii* (*trnM3 = trnfM*). The gene IDs for *trnI-CAU* are: 800433 in *N. tabacum*, 4042812 in *A. thaliana*, 3131361 in *O. sativa*, 2546774 in *P. patens* and 2716991 in *C. reinhardtii* (*trnM1 = trnI*). The gene IDs for *trnI-GAU* are: 800522 and 800440 in *N. tabacum*, 1466260 and 5563231 in *A. thaliana*, 3131365 and 3131379 in *O. sativa*, 2637323 and 2637324 in *P. patens* and 2716957 and 2717026 in *C. reinhardtii*. The sequences were aligned using ClustalX (26). For phylogenetic

analysis, sequences were aligned using MUSCLE (27) and phylogenetic tree construction was performed with the PHYLIP package (28).

Analysis of codon usage

Codon numbers were calculated from all protein-coding genes (excluding small, non-conserved ORFs) in the *N. tabacum* plastid genome (NC_001879) and the *Bos taurus* mitochondrial genome (NC_006853) using the Kazusa Countcodon program Version 4 (http://www. kazusa.or.jp/codon/countcodon.html). Plastids genes located in the inverted repeat region of the genome are counted twice. Codons undergoing RNA editing in tobacco plastids (29) were included in their edited form.

RESULTS

Targeted inactivation of the four plastid genome-encoded tRNA genes for the AUN codon box

The AUN codon box, comprising the four triplets that specify the amino acids methionine (AUG) and isoleucine (AUA, AUC and AUA), is the only codon box in plastids of higher plants that is read by four different tRNA species (Figure 1): the initiator tRNA^{fMet}(CAU) (*trnfM*), the elongator tRNA^{Met}(CAU) (*trnM*), tRNA^{Ile}(CAU) (*trnI-CAU*) and tRNA^{Ile}(GAU) (*trnI-GAU*). This situation somewhat contrasts the general evolutionary trend in plastids (and mitochondria) to reduce the number of organellar tRNA genes to a minimum. As some non-photosynthetic (heterotrophically growing) plastid-bearing species possess in their plastid genomes fewer than four tRNA species for the AUN box, we wanted to determine whether or not any of the four tRNA genes is dispensable. The availability of a plastid transformation technology for the seed plant tobacco (*N. tabacum*; 16,23) and the ability to grow tobacco plants heterotrophically *in vitro* even in the complete absence of photosynthesis (30–33), enabled us to address this question experimentally.

To test the four tRNA genes for essentiality, knockout alleles were constructed by cloning the corresponding regions from the tobacco plastid genome and disrupting or replacing the tRNA genes with the standard selectable marker gene for chloroplast transformation, *aadA* (Figures 2–5). The *aadA* gene encodes the enzyme aminoglycoside 3″-adenylyltransferase, which confers resistance to the aminoglycoside antibiotic spectinomycin, a potent inhibitor of chloroplast translation (23).

To knock out the *trnM* gene, two mutant alleles were constructed that differ in the orientation of the *aadA* cassette relative to the neighboring genes (Figure 3A–C). This was done because *trnM* is located between the *atpE* and *trnV-UAC* genes, which both are transcribed from the other strand of the plastid genome (Figure 3A). To exclude possible polar effects on the transcription of one of the two neighboring genes, knockout constructs for both *aadA* orientations were produced and designated p∆*trnM*-s (harboring the *aadA* cassette in the same transcriptional orientation as the *trnM* gene; s: sense) and p∆*trnM*-as (containing the *aadA* in the opposite orientation; as: antisense), respectively (Figure 3B and C).

For all other knockout constructs, the selectable marker gene was inserted in the same transcriptional orientation as the original tRNA gene (Figures 2, 4 and 5). This was sufficient, because in these cases, the tRNA gene is not located between two transcription units on the opposite strand of the plastome (Figures 2A, 4A and 5A).

The five constructed knockout alleles for the four tRNA genes of the AUN box were introduced into the tobacco plastid genome by biolistic (particle gun-mediated) transformation followed by selection of bombarded leaf pieces on spectinomycin-containing plant regeneration medium (23). Transformation of the plastid genome occurs via homologous recombination, which results in targeted replacement of the wild-type tRNA allele by the null allele disrupted with the *aadA* marker (Figures 2–5). Selected spectinomycin-resistant plant lines were subjected to three to four rounds of regeneration on antibiotic-containing medium to select against residual wild-type copies of the plastid genome and enrich the transformed genome (transplastomes; 23,24). All selection and regeneration steps were performed on sucrose-containing medium, thus facilitating the isolation of homoplasmic transplastomic mutants that are unable to grow photoautotrophically (30–32).

trnfM, *trnM*, *trnI-CAU* and *trnI-GAU* are essential plastid genes

In plastid transformation experiments, two rounds of regeneration and selection are typically sufficient to eliminate all residual wild-type copies of the plastid genome and isolate cell lines that harbor a homogeneous population of transformed genome copies (referred to as homoplasmic transplastomic lines). To test the sets of transplastomic knockout lines generated for the four tRNA genes of the AUN box for homoplasmy, the genotypes of two to three lines per construct were analyzed by restriction fragment-length polymorphism (RFLP) analysis. In all tested plants, Southern blot experiments confirmed presence of the transplastomic genome configuration expected to arise from disruption of the respective tRNA gene by homologous recombination. However, all transplastomic lines also retained copies of the wild-type genome as evidenced by the presence of a second restriction fragment that corresponded in size to the fragment seen in the wild-type control plants (Figures 2C, 3D, 4C and 5C). The ratios between the hybridization intensities of the wild-type fragment and the transplastomic fragment were remarkably similar in all lines, which is a strong indication of a balancing selection operating in the presence of spectinomycin (34). This suggests that both genome types (wild-type plastome and transplastome) are essential: the transplastome because it provides the antibiotic resistance and the wild-type genome because it expresses the essential tRNA knocked-out in the transplastome. Stable simultaneous retention of wild-type genome and transplastome is therefore observed in all reverse genetics experiments that target essential genes (17,34–36). As chloroplast translation is essential in
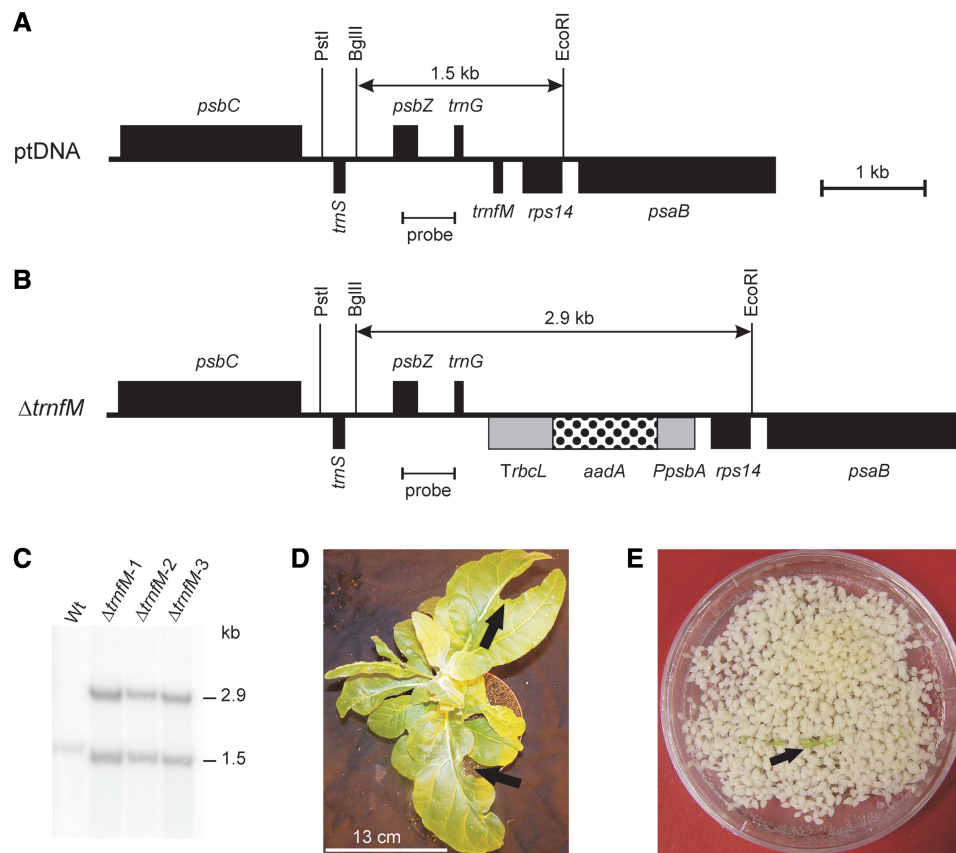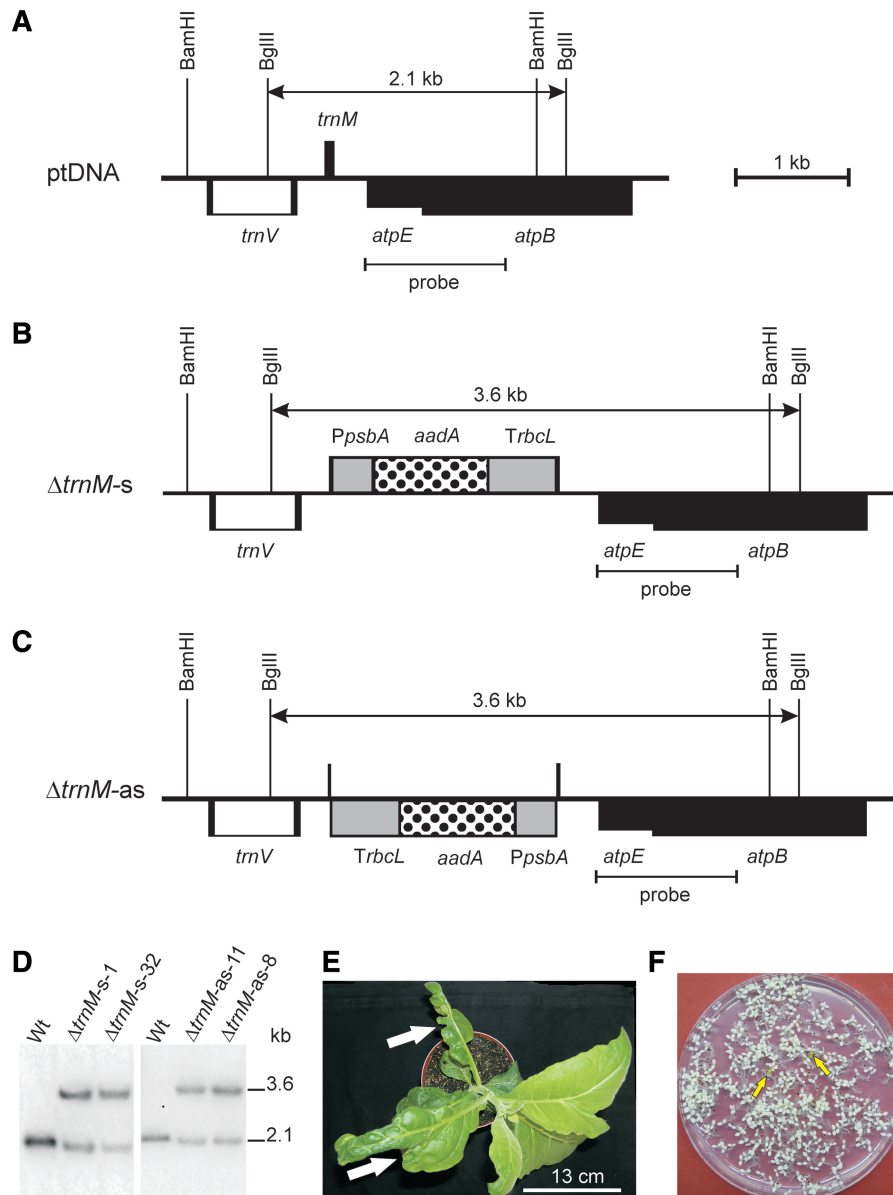
Figure 2. Targeted deletion of the plastid *trnfM* gene. (**A**) Physical map of the *trnfM*-containing region in the tobacco plastid genome (ptDNA; 56). Genes above the line are transcribed from the left to the right, genes below the line are transcribed in the opposite direction. Selected restriction sites used for cloning and RFLP analysis are indicated. The hybridization probe and the expected sizes of detected DNA fragments are also shown. (**B**) Map of the transformed plastid genome (transplastome) produced with plastid transformation vector pΔ*trnfM*. The selectable marker gene *aadA* (spotted box; 23) is driven by the plastid *psbA* promoter (P*psbA*) and the 3′-UTR from the plastid *rbcL* gene (T*rbcL*; gray boxes). (**C**) RFLP analysis of Δ*trnfM* chloroplast transformants. The transplastomic lines remain heteroplasmic and show both the 1.5 kb wild type-specific hybridization band and the 2.9 kb band diagnostic of the transplastome. Wt: wild type. (**D**) Phenotype of a typical (heteroplasmic) Δ*trnfM* transplastomic plant. Arrows point to examples of misshapen leaves. (**E**) Example of a seed assay confirming heteroplasmy of Δ*trnfM* plants and gradual loss of the transplastome in the absence of antibiotic selection. The transplastome is lost from most seedlings as evidenced by their white phenotype upon germination on spectinomycin-containing medium. The arrow points to one of the occasionally appearing green (spectinomycin-resistant) seedlings that still harbor the transplastome.

tobacco and most other dicotyledonous plants, stable heteroplasmy has also been observed in previous experiments attempting knockout of essential components of the translational machinery (5,17–19,21,37).

Previous transplastomic studies on essential plastid genes also have provided a clear phenotypic assay that unambiguously demonstrates essentiality. Transfer of the heteroplastomic lines to soil and growth in the absence of antibiotic selection results in random segregation of the two genome types and, while segregation into homoplasmy for the wild-type plastome remains invisible, segregation into homoplasmy for the knockout of the essential gene leads to cessation of cell division and ultimately cell death. If cell death occurs early during leaf development, entire cell lines are missing from the developing organ, which results in severely misshapen leaves that lack parts of their blade (5,17,19,21). This phenotype was seen in all of our transplastomic tRNA knockout lines (Figures 2D, 3E, 4D and 5D), providing

strong evidence for essentiality of all four plastid tRNA genes in the AUN box.

Knockout of an essential gene confers a strong selective disadvantage and, therefore, the knockout allele is eventually lost upon release of the selective pressure (i.e. growth on antibiotic-free medium). Thus, typically no or only very few copies of the transplastome enter the female germline and are transmitted through seeds into the next generation (19,21,34). As expected, rapid loss of the transplastome was also observed for our knockout alleles of *trnfM*, *trnM*, *trnI-CAU* and *trnI-GAU*. When seeds were harvested from plants grown in the absence of antibiotic selection and germinated on spectinomycin-containing medium, the vast majority of the offspring was sensitive to the antibiotic, indicating complete loss of the knockout alleles. Only in rare cases, copies of the transplastome were transmitted into the next generation, as evidenced by occasional appearance of green (antibiotic-resistant) seedlings that contained
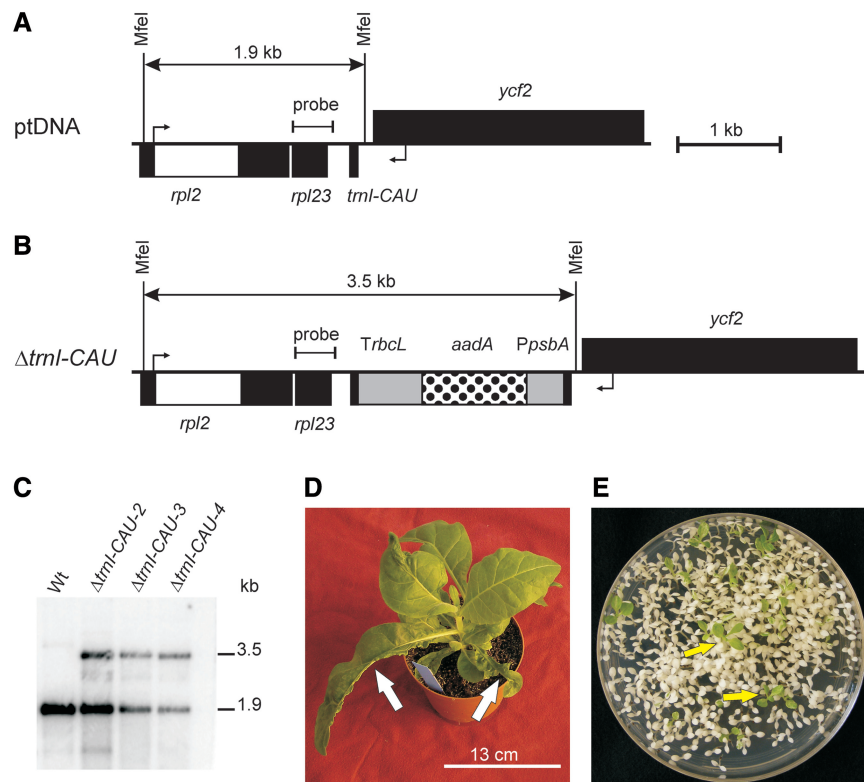
Figure 3. Targeted inactivation of the plastid *trnM* gene. (**A**) Physical map of the *trnM*-containing region in the tobacco ptDNA. Genes above the line are transcribed from the left to the right, genes below the line are transcribed in the opposite direction. Selected restriction sites relevant for cloning and RFLP analysis are indicated. The hybridization probe and the expected sizes of detected DNA fragments are also shown. Introns are represented by open boxes. (**B**) Map of the transplastome produced with plastid transformation vector pΔ*trnM*-s. The selectable marker gene *aadA* is inserted into *trnM* gene in the same transcriptional orientation. The *aadA* coding region is shown as spotted box, the expression elements driving it as gray boxes. (**C**) Map of the transplastome produced with plastid transformation vector pΔ*trnM*-as. The selectable marker gene *aadA* is inserted into *trnM* gene in the opposite (antisense) orientation. (**D**) RFLP analysis of Δ*trnM* plastid transformants. Both sets of transplastomic lines produced (Δ*trnM*-s and Δ*trnM*-as) remain heteroplasmic as evidenced by simultaneous presence of the 2.1 kb wild type-specific hybridization band and the 3.6 kb band diagnostic of the transplastome. Wt: wild type. (**E**) Phenotype of a typical Δ*trnM* transplastomic plant. Arrows point to typical misshapen leaves. (**F**) Segregation analysis of a Δ*trnM* plant. Seeds from a selfed transplastomic plant were sown on spectinomycin-containing medium. Spectinomycin sensitivity of most seedlings suggests a strong tendency to lose the transplastome in the absence of antibiotic selection. Examples of spectinomycin-resistant seedlings (that have retained the transplastome) are indicated by arrows.

a heteroplasmatic mix of wild-type plastomes and transplastomes (Figure 2E, 3F, 4E and 5E).

The two different knockout alleles constructed for *trnM-CAU* gave identical results (Figure 3), largely excluding interference with the transcription of the neighboring genes as the cause of the failure to obtain homoplasmic knockout lines. In the case of the

knockout allele for *trnI-GAU*, the selection marker was inserted into the group II intron (deleting part of it), next to one of the origins of DNA replication (*oriA*) in the tobacco plastid genome (38; Figure 5A). However, interference with DNA replication is unlikely to occur in our Δ*trnI-GAU* knockout lines, because previous mutagenesis experiments have established that *oriA* is

Figure 4. Disruption of the plastid *trnI-CAU* gene. (**A**) Physical map of the region in the tobacco plastid genome containing *trnI-CAU*. Genes above the line are transcribed from the left to right, genes below the line are transcribed in the opposite direction. The bent arrows indicate the borders of the transformation plasmid. The hybridization probe and the expected sizes of detected DNA fragments are also shown. Introns are represented by open boxes. (**B**) Map of the transformed plastid genome (transplastome) produced with plastid transformation vector pΔ*trnI-CAU*. The *aadA* coding region is shown as spotted box, the expression elements driving it as gray boxes. (**C**) RFLP analysis of Δ*trnI-CAU* transplastomic plants. All lines remain heteroplasmic as evidenced by a constant ratio of the 1.9 kb wild type-specific hybridizing fragment and the 3.5 kb fragment diagnostic of the transplastome. Wt: wild type. (**D**) Phenotype of a typical Δ*trnI-CAU* transplastomic plant. Arrows point to misshapen leaves. (**E**) Inheritance assay of a Δ*trnI-CAU* plant. Examples of spectinomycin-resistant seedlings that have retained the transplastome are indicated by arrows.

non-essential in that homoplasmic *oriA* deletion lines do not have a pronounced mutant phenotype (39,40). In contrast, insertions into the *trnI-GAU* intron cannot be brought to homoplasmy (39), indicating that loss of intron excision knocks out *trnI-GAU* function (Figure 5C–E).

Taken together, the data from our transplastomic knockout experiments provide strong genetic evidence for essentiality of both tRNAs decoding the AUG methionine codon, the initiator tRNA^{fMet}(CAU) and the elongator tRNA^{Met}(CAU), and also of both tRNAs decoding the three isoleucine codons, tRNA^{Ile}(CAU) and tRNA^{Ile}(GAU). Requirement for all four tRNA genes under both autotrophic and heterotrophic growth conditions suggests that all four tRNAs decoding the AUN box are indispensable for plastid translation.

Phylogenetic analysis of tRNAs in the AUN box

The plastid genome encodes three tRNAs with a CAU anticodon: tRNA^{fMet}(CAU), tRNA^{Met}(CAU) and tRNA^{Ile}(CAU). Our reverse genetics analysis in tobacco suggests that all three of them are essential. It seems reasonable to assume that essentiality is due to strict discrimination between the tRNA species based on structural features and/or sequence features. To gain a better understanding of the distinguishing features of the tRNAs in the AUN box, we analyzed sequences and structures of the four tRNAs (Figure 6) and compared them to the homologous tRNA genes in the plastid genomes of the dicot *Arabidopsis thaliana*, the monocot *Oryza sativa*, the moss *Physcomitrella patens* and the green alga *Chlamydomonas reinhardtii* (Supplementary Figure S1).

The initiator tRNA^{fMet}(CAU) in bacteria is characterized by a mismatch between nucleotides 1 and 72 in the acceptor stem (for numbering of nucleotide positions, see 41), an AU basepair at positions 11–24 and a GC pair at positions 12–23 in the D-loop stem. In addition, tRNA^{fMet}(CAU) has a highly conserved anticodon loop and two to three GC base pairs adjacent to it in the anticodon stem (10,42). The mismatched pair at 1-72 in the acceptor stem is also a common feature of all plastid initiator tRNAs analysed, whereas the elongator tRNAs for methionine have AU or GU pairs and the tRNAs for isoleucine have GC pairs at this position (feature 1 in Figure 6 and Supplementary Figure S1). In addition, all plastid initiator tRNAs have the conserved AU pair at 11–24 and a GC pair at 12–23, whereas the elongator tRNAs contain the same basepairs in
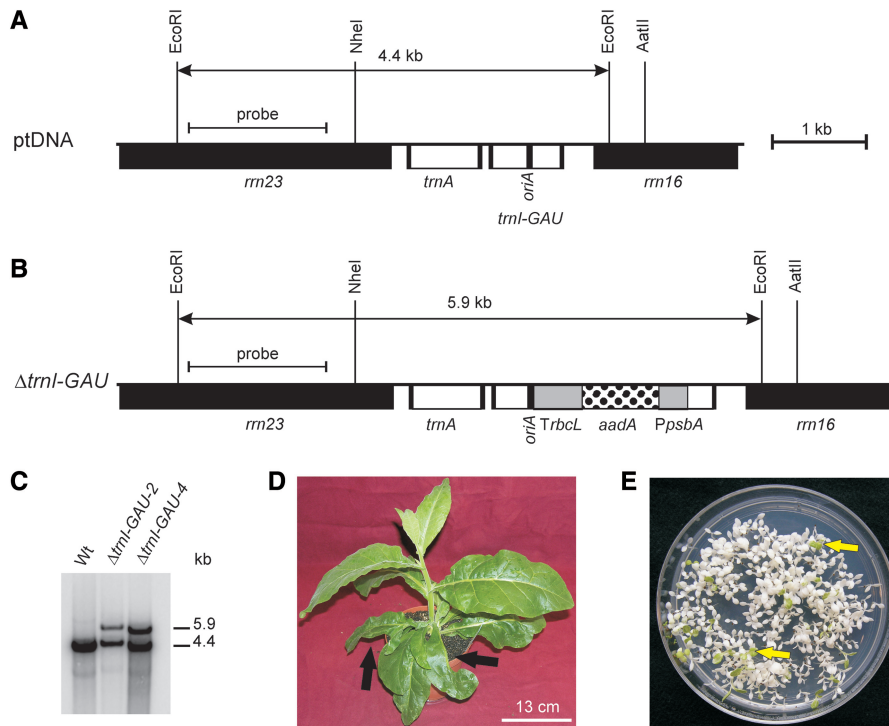
Figure 5. Targeted disruption of the plastid *trnI-GAU* gene. (**A**) Physical map of the region in the tobacco plastome containing *trnI-GAU*. Genes above the line are transcribed from the left to the right, genes below the line are transcribed in the opposite direction. Selected restriction sites used for cloning and RFLP analysis are indicated. The hybridization probe and the expected sizes of detected DNA fragments are also shown. Introns are represented by open boxes. An origin of replication located in this region is denoted as *oriA*. (**B**) Map of the transformed plastid genome produced with plastid transformation vector pΔ*trnI-GAU*. The *aadA* coding region is shown as spotted box, the expression elements driving it as gray boxes. (**C**) RFLP analysis of Δ*trnI-GAU* plastid transformants. The transplastomic lines remain heteroplasmic and show both the 4.4 kb wild type-specific hybridization band and the 5.9 kb band diagnostic of the transplastome. Wt: wild type. (**D**) Phenotype of a typical heteroplasmic Δ*trnI-GAU* plant. Misshapen leaves are indicated by arrows. (**E**) Example of a seed assay confirming heteroplasmy of Δ*trnI-GAU* plants and demonstrating loss of the transplastome in the absence of antibiotic selection. Arrows points to two green (spectinomycin-resistant) seedlings that still harbor the transplastome.

the inverted orientation (feature 3 in Figure 6 and Supplementary Figure S1). Also, the anticodon loop of the plastid *trnfM* sequences is identical to the bacterial consensus. However, the anticodon loop sequence does not serve as an unambiguous distinguishing feature, because it is also identical with all analyzed plastid tRNA^{Ile}(CAU) (Figure 6 and Supplementary Figure S1).
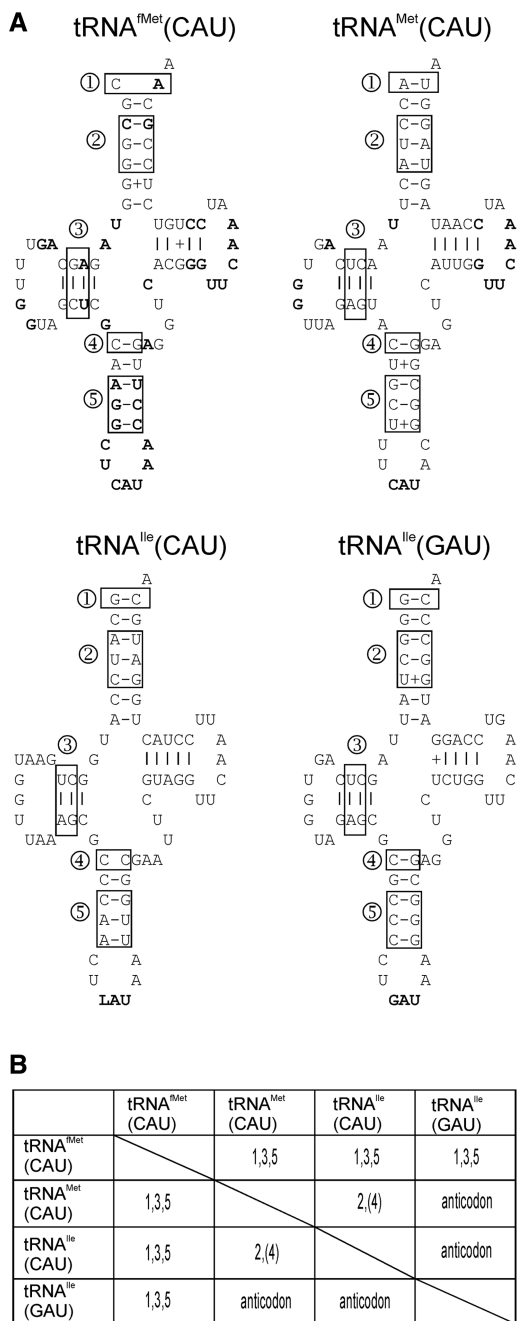
Three GC pairs present in the anticodon stem of *trnfM* in *Escherichia coli* have been described as key determinants of the initiator tRNA (43). However, not all three GC pairs are conserved in other bacterial species (42). Also in the plastid *trnfM* sequences analysed here, only two GC pairs are present in the anticodon stem (feature 5 in Supplementary Figure S1). Although some elongator methionine tRNAs also have two GC pairs in their anticodon stem, one of these basepairs is usually flipped (feature 5 in Figure 6). In several species, tRNA^{Ile}(GAU) possesses three consecutive CG pairs in its anticodon stem (feature 5 in Figure 6 and Supplementary Figure S1), but as the anticodon is different from that of the initiator methionine tRNA, this does not result in charging ambiguity.

In summary, nearly all of the features conserved in bacterial initiator tRNAs (42) are also conserved in plastid *trnfM* genes. The only exception is a mutation of G_{26} to A in the *trnfM* from rice. However, as neither G_{26} nor A_{26}

can base pair with the conserved A_{44} (Figure 6 and Supplementary Figure S1), this mutation is likely to be neutral.

tRNA^{Ile}(CAU) needs to unambiguously distinguished from tRNA^{Met}(CAU), because the cytidine in the wobble position of tRNA^{Ile}(CAU) must be modified to lysidine. This post-transcriptional modification converts the methionine anticodon of tRNA^{Ile}(CAU) into an anticodon decoding the isoleucine triplet AUA (6–9). Two distinguishing features of tRNA^{Ile}(CAU) have been described in the literature: (i) the three base pairs at positions 3–70, 4–69 and 5–68 in the acceptor stem (10,11), and (ii) unpaired or weakly paired nucleotides at 27–43 in the anticodon stem (11,42). The distinguishing features described in *Escherichia coli*, CG pairs at 4–69 and 5–68 (11), are only partly conserved in the plastid tRNA^{Ile}(CAU) (feature 2 in Figures 6 and Supplementary Figure S1). In contrast, all tRNA^{Met}(CAU) genes have a UA and an AU pair at these two positions. In an analysis of 234 bacterial genomes, a purine–pyrimidine pair at position 3–70 and frequently occurring CG pairs at 4–69 and 5–68 in the acceptor stem were identified as characteristic features of tRNA^{Ile}(CAU) (11). All analyzed plastid tRNA^{Ile}(CAU) sequences meet these criteria (feature 2 in Figure 6 and

A

B

	tRNA^{fMet} (CAU)	tRNA^{Met} (CAU)	tRNA^{Ile} (CAU)	tRNA^{Ile} (GAU)
tRNA^{fMet} (CAU)		1,3,5	1,3,5	1,3,5
tRNA^{Met} (CAU)	1,3,5		2,(4)	anticodon
tRNA^{Ile} (CAU)	1,3,5	2,(4)		anticodon
tRNA^{Ile} (GAU)	1,3,5	anticodon	anticodon	

Figure 6. Characteristic sequence features of the analyzed plastid tRNAs. (**A**) Secondary structures of the *Nicotiana tabacum* tRNA^{Met}(CAU), tRNA^{fMet}(CAU), tRNA^{Ile}(CAU) and tRNA^{Ile}(GAU). The tRNAs were folded using the ARAGORN webservice (57). The − indicates canonical base pairing, + denotes a GU base pair. The anticodons and the nucleotides in tRNA^{Met}(CAU) and tRNA^{fMet}(CAU), which are conserved (according to 42) are shown in bold. Additional characteristic sequence features are boxed and numbered (see text for details). The lysidine modification in the anticodon of tRNA^{Ile}(CAU) is indicated by an L. (**B**) List of distinguishing features that allow discrimination between the four tRNA species. Features are numbered as in (A). Features that were described as discriminatory in bacteria, but may not be so in plastids are in parentheses (see text for details). Cases in which the anticodon is probably the main discriminatory feature are also indicated.

Supplementary Figure S1). The unpaired or weakly paired (GU) nucleotides at position 27–43 in the acceptor stem (11,42) could potentially distinguish tRNA^{Ile}(CAU) from tRNA^{Met}(CAU) in tobacco, but not in the other analyzed species. The lack of conservation of this feature in plastid tRNA^{Ile}(CAU) suggests that the anticodon stem may not be the key discriminator between tRNA^{Ile}(CAU) and tRNA^{Met}(CAU). Instead, the differences in the acceptor stems (feature 2 in Figure 6 and Supplementary Figure S1; 10) are more likely to confer unambiguous tRNA recognition by the tRNA(Ile)-lysidine synthetase (11).

DISCUSSION

The use of wobble and superwobble interactions in codon–anticodon recognition enables plastids to carry out translation with a reduced set of tRNAs (1,5). In the AUN codon box, superwobbling by a hypothetical tRNA^{Ile}(UAU) cannot be employed, because no discrimination between methionine and isoleucine would be possible (Figure 7). This problem persists if the box were decoded by classical wobbling, which would involve a tRNA species with a GAU anticodon (which does exist) and a tRNA species with an UAU anticodon (which does not exist). To distinguish between AUA Ile and AUG Met codons, plastids utilize the same strategy as most bacteria: lysidination of cytidine C_{34} in the tRNA^{Ile}(CAU) (6–9). This modification confers specificity to the AUA triplet and effectively prevents decoding of the Met codon AUG (Figure 7). Our finding that all three elongator tRNAs reading the AUN box are essential in plastids (Figures 3D–F, 4C–E and 5C–E) provides strong evidence for conservation of the bacterial mechanism for ensuring decoding specificity.

In *C. reinhardtii*, the three plastid tRNAs with a CAU anticodon are all annotated as tRNA^{Met}(CAU) (*trnM1, trnM2* and *trnM3*; 44). This is likely to be incorrect, as has also been noted in a recently published comprehensive survey of plant tRNA genes (45). The analysis of conserved structural features (Figure 6) and a phylogenetic analysis of the three tRNA genes (Supplementary Figures S1 and S2) clearly identify the misannotated *trnM1* as *trnI-CAU* and the misannotated *trnM3* as

Amino acid	N$_3$	tRNA anticodon				Amino acid
		Super-wobble	Wobble	Bacteria / plastids	Mito-chondria	
Ile	U					Ile
Ile	C		G	G	G	Ile
Ile	A	(U)	(U)	C*		Met
Met	G			C	C**	Met

Figure 7. Codon recognition in the AUN codon box in bacteria, plastids and metazoan mitochondria. tRNA species are denoted by the first nucleotide of their anticodon (which pairs with the third codon position, N$_3$). Isoleucine-specifying triplets are shown as white boxes, methionine-specifying triplets as gray boxes. Note that AUA specifies Ile in bacteria and plastids, but Met in metazoan mitochondria. Naturally existing tRNA species are indicated for bacteria, plastids and mitochondria. Hypothetical tRNA species whose existence would lead to decoding ambiguities (due to forbidden wobble and superwobble interactions) are shown in parentheses. C* = lysidine, C** = 5-formylcytidine.

trnfM-CAU. *trnM2* is the only true *trnM-CAU* gene and should be renamed as such.

A theoretical possibility of reducing the tRNA set in the AUN box would be partial lysidine modification of tRNA^{Ile}(CAU). This would allow the same gene to be used for production of an unmodified tRNA charged with methionine (decoding AUG triplets) and a modified tRNA charged with isoleucine (reading AUA triplets). However, essentiality of both the *trnM-CAU* and the *trnI-CAU* gene (Figures 3D–F and 4C–E) shows that this strategy is not used. Possible reasons could lie in the specificities of the aminoacyl-tRNA synthetases and/or a high catalytic activity of the lysidine synthetase leaving insufficient amounts of unmodified (cytidine-containing) tRNA molecules to serve as tRNA^{Met}(CAU).

Another possibility to reduce the tRNA set in the AUN box would be a tRNA^{Met} that serves as both initiator and elongator tRNA. The initiator tRNA is highly conserved in all three domains of live (42,46) and its gene is not lost from the reduced plastomes of any of the non-photosynthetic plastid-containing parasitic species investigated to date (13,14). However, metazoan mitochondria contain a single methionine tRNA, which is used as both initiator and elongator tRNA. This unusual tRNA^{Met}(CAU) combines several typical features of initiator and elongator tRNAs suggesting co-evolution with translation initiation and elongation factors as well as the methionyl-tRNA transformylase and the methionyl-tRNA synthetase (47). In *Nicotiana tabacum*, both tRNA^{fMet}(CAU) and tRNA^{Met}(CAU) are essential (Figure 2C–E and 3D–F). This is because, in contrast to the tRNA^{Met}(CAU) in metazoan mitochondria, the typical sequence features enabling discrimination of the initiator and elongator tRNAs are conserved in the two plastid Met tRNAs (Figure 6). Therefore, the two plastid tRNA species cannot replace each other.

A third theoretical possibility of reducing the number of plastid tRNAs decoding the AUN box would be changes in the genetic code. In metazoan mitochondria, both AUG and AUA specify methionine and are read by a tRNA^{Met}(CAU) carrying a modified cytidine in the wobble position (5-formylcytidine; 47). This enabled the reduction of the number of tRNAs for isoleucine by making the tRNA^{Ile}(CAU) reading AUA triplets dispensable (Figure 7). Such a change in the genetic code must be accompanied by compensatory mutations in the protein-coding genes, at least in those cases where the isoleucine residues are important for protein function. Metazoan mitochondria have a very low number of protein-coding genes (for example, only 13 genes in *Bos taurus* mitochondria) and consequently, the number of affected AUA codons is low (only 216 in *B. taurus*; Supplementary Table S2). Although also the number of plastid genes decreased dramatically during evolution, tobacco plastids still retain 87 protein-coding genes with altogether 454 AUA codons. It is tempting to speculate that a low number of affected codons increases the probability for changes in the genetic code. However, codon reassignment has also occurred in the nuclear genome of some yeast (*Candida*) species, where CUG triplets specify serine instead of leucine (48–50). As the *Candida* nuclear

genome contains thousands of genes, this argues against a reduction in gene number being a prerequisite for modifications of the genetic code.

In vitro, tobacco cells can tolerate a severe impairment of plastid translation, to the point that photosynthesis is reduced to extremely low levels and growth is severely retarded (20, 21). The inviability of homoplasmic knock-out cells for the four tRNA genes in the AUN box even on sucrose-containing medium (that abrogates any need for photosynthetic carbon fixation; 31,32) strongly indicates that no net protein biosynthesis can occur in the absence of any of the four tRNA species. Furthermore, our finding that all tRNA genes in the AUN box are essential even under heterotrophic conditions demonstrates that the evolutionary loss of individual tRNA genes from the plastomes of non-photosynthetic parasitic species (13) is not an indication of non-essentiality. Essentiality of these genes also suggests that, at least in tobacco, loss of these tRNAs cannot be complemented by tRNAs imported from the cytosol.

In view of the essential function of all tRNA species in the AUN box, the loss of tRNA genes from the plastomes of parasitic lineages remains puzzling. No indications for an adaptation of the codon usage to the reduced tRNA set were found. For example, *Rhizanthella gardneri*, which lacks the plastid-encoded tRNA^{Met}(CAU) and tRNA^{Ile}(GAU), shows no corresponding changes in codon usage (14).

In theory, the tRNA gene loss from plastomes of non-photosynthetic lineages could be explained by gene transfer to the nucleus and tRNA reimport into plastids (14,51,52) or import of nuclear-encoded eukaryotic tRNAs (as demonstrated to occur in mitochondria; 53, 54). The mechanistic problem with this explanation is that, thus far, no convincing evidence for RNA import into plastids has been obtained. As tRNA gene loss from the plastid genome has occurred in parasites representing widely different taxonomic groups (13), the additional assumption would need to be made that gene transfer to the nucleus and evolution of a tRNA import capacity occurred several times independently (or, alternatively, a cryptic ancestral import pathway exists that can be activated upon demand). An accelerated transfer of plastid genes to the nucleus of parasitic plants has recently been proposed for ribosomal protein genes and was explained by the relaxed requirement for plastid translation in the absence of photosynthesis (21). However, in the absence of any experimental evidence for tRNA import into plastids, possible alternative explanations that could make individual tRNAs dispensable (such as, the partial lysidine modification of the CAU anticodon discussed above) currently should not be dismissed. However, modified decoding is unlikely to explain all cases of tRNA gene loss, especially not those affecting a tRNA species that reads an entire codon box (14).

Under the assumptions that the standard genetic code is used, no unusual tRNA modification patterns occur and no tRNA import exists, the plastid tRNA set of the parasitic non-photosynthetic green alga *Helicosporidium sp.* (55) appears to exemplify the minimum set of tRNA genes that is achievable by maximal utilization of

wobbling and superwobbling. For most amino acids, only one tRNA gene is present in the *Helicosporidium* plastome, but, consistent with our results reported here, all four tRNAs for the AUN box are conserved: tRNA^{fMet}(CAU), tRNA^{Met}(CAU), tRNA^{Ile}(CAU) and tRNA^{Ile}(GAU).

In summary, our data reported here provide strong genetic and phylogenetic evidence that the nature of the genetic code prevents a reduction of the plastid tRNA set for the AUN box. To faithfully differentiate between methionine and isoleucine, superwobbling must be avoided (thus excluding existence of a tRNA-UAU) and wobbling is restricted to tRNA^{Ile}(GAU) (reading the iso-leucine codons AUU and AUC). The molecular mechanisms compensating for the tRNA gene loss from the plastid genomes of non-photosynthetic parasites remain to be elucidated.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Tables 1–2 and Supplementary Figures 1–2.

ACKNOWLEDGEMENTS

FUNDING

REFERENCES

1. Crick,F.H.C. (1966) Codon-anticodon pairing: the wobble hypothesis. *J. Mol. Biol.*, **19**, 548–555.
2. Sugiura,M. and Wakasugi,T. (1989) Compilation and comparison of transfer RNA genes from tobacco chloroplasts. *Crit. Rev. Plant Sci.*, **8**, 89–101.
3. Osawa,S., Jukes,T.H., Watanabe,K. and Muto,A. (1992) Recent evidence for evolution of the genetic code. *Microbiol. Rev.*, **56**, 229–264.
4. Bock,R. (2007) Structure, function, and inheritance of plastid genomes. *Top. Curr. Genet.*, **19**, 29–63.
5. Rogalski,M., Karcher,D. and Bock,R. (2008) Superwobbling facilitates translation with reduced tRNA sets. *Nat. Struct. Mol. Biol.*, **15**, 192–198.
6. Harada,F. and Nishimura,S. (1974) Purification and characterization of AUA specific isoleucine transfer ribonucleic acid from Escherichia coli B. *Biochemistry*, **13**, 300–307.
7. Kashdan,M.A. and Dudock,B.S. (1982) The gene for a spinach chloroplast isoleucine tRNA has a methionine anticodon. *J. Biol. Chem.*, **257**, 11191–11194.
8. Muramatsu,T., Yokoyama,S., Horie,N., Matsuda,A., Ueda,T., Yamaizumi,Z., Kuchino,Y., Nishimura,S. and Miyazawa,T. (1988) A novel lysine-substituted nucleoside in the first position of the anticodon of minor isoleucine tRNA from Escherichia coli. *J. Biol. Chem.*, **263**, 9261–9267.
9. Muramatsu,T., Nishikawa,K., Nemoto,F., Kuchino,Y., Nishimura,S., Miyazawa,T. and Yokoyama,S. (1988) Codon and amino-acid specificities of a transfer RNA are both converted by a single post-transcriptional modification. *Nature*, **336**, 179–181.
10. Silva,F.J., Belda,E. and Talens,S.E. (2006) Differential annotation of tRNA genes with anticodon CAT in bacterial genomes. *Nucleic Acids Res.*, **34**, 6015–6022.
11. Ikeuchi,Y., Soma,A., Ote,T., Kato,J., Sekine,Y. and Suzuki,T. (2005) Molecular mechanism of lysidine synthesis that determines tRNA identity and codon recognition. *Mol. Cell*, **19**, 235–246.
12. Ikeuchi,Y., Kitahara,K. and Suzuki,T. (2008) The RNA acetyltransferase driven by ATP hydrolysis synthesizes N4-acetylcytidine of tRNA anticodon. *EMBO J.*, **27**, 2194–2203.
13. Krause,K. (2011) Piecing together the puzzle of parasitic plant plastome evolution. *Planta*, **234**, 647–656.
14. Delannoy,E., Fujii,S., des Francs-Small,C.C., Brundrett,M. and Small,I. (2011) Rampant gene loss in the underground orchid Rhizanthella gardneri highlights evolutionary constraints on plastid genomes. *Mol. Biol. Evol.*, **28**, 2077–2086.
15. Maliga,P. (2004) Plastid transformation in higher plants. *Annu. Rev. Plant Biol.*, **55**, 289–313.
16. Maliga,P. and Bock,R. (2011) Plastid biotechnology: food, fuel, and medicine for the 21st century. *Plant Physiol.*, **155**, 1501–1510.
17. Ahlert,D., Ruf,S. and Bock,R. (2003) Plastid protein synthesis is required for plant development in tobacco. *Proc. Natl Acad. Sci. USA*, **100**, 15730–15735.
18. Rogalski,M., Ruf,S. and Bock,R. (2006) Tobacco plastid ribosomal protein S18 is essential for cell survival. *Nucleic Acids Res.*, **34**, 4537–4545.
19. Rogalski,M., Schöttler,M.A., Thiele,W., Schulze,W.X. and Bock,R. (2008) Rpl33, a nonessential plastid-encoded ribosomal protein in tobacco, is required under cold stress conditions. *Plant Cell*, **20**, 2221–2237.
20. Allison,L.A., Simon,L.D. and Maliga,P. (1996) Deletion of rpoB reveals a second distinct transcription system in plastids of higher plants. *EMBO J.*, **15**, 2802–2809.
21. Fleischmann,T.T., Scharff,L.B., Alkatib,S., Hasdorf,S., Schöttler,M.A. and Bock,R. (2011) Nonessential plastid-encoded ribosomal proteins in tobacco: a developmental role for plastid translation and implications for reductive genome evolution. *Plant Cell*, **23**, 3137–3155.
22. Murashige,T. and Skoog,F. (1962) A revised medium for rapid growth and bio assays with tobacco tissue culture. *Physiol. Plant.*, **15**, 473–497.
23. Svab,Z. and Maliga,P. (1993) High-frequency plastid transformation in tobacco by selection for a chimeric aadA gene. *Proc. Natl Acad. Sci. USA*, **90**, 913–917.
24. Bock,R. (2001) Transgenic chloroplasts in basic research and plant biotechnology. *J. Mol. Biol.*, **312**, 425–438.
25. Doyle,J.J. and Doyle,J.L. (1990) Isolation of plant DNA from fresh tissue. *Focus*, **12**, 13–15.
26. Thompson,J.D., Gibson,T.J., Plewniak,F., Jeanmougin,F. and Higgins,D.G. (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.*, **25**, 4876–4882.
27. Edgar,R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.*, **32**, 1792–1797.
28. Felsenstein,J. (1989) PHYLIP - Phylogeny Inference Package (Version 3.2). *Cladistics*, **5**, 164–166.
29. Kahlau,S., Aspinall,S., Gray,J.C. and Bock,R. (2006) Sequence of the tomato chloroplast DNA and evolutionary comparison of solanaceous plastid genomes. *J. Mol. Evol.*, **63**, 194–207.
30. Kanevski,I. and Maliga,P. (1994) Relocation of the plastid rbcL gene to the nucleus yields functional ribulose-1,5-bisphosphate

carboxylase in tobacco chloroplasts. *Proc. Natl Acad. Sci. USA*, **91**, 1969–1973.

31. Ruf,S., Kössel,H. and Bock,R. (1997) Targeted inactivation of a tobacco intron-containing open reading frame reveals a novel chloroplast-encoded photosystem I-related gene. *J. Cell Biol.*, **139**, 95–102.

32. Hager,M., Biehler,K., Illerhaus,J., Ruf,S. and Bock,R. (1999) Targeted inactivation of the smallest plastid genome-encoded open reading frame reveals a novel and essential subunit of the cytochrome b6f complex. *EMBO J.*, **18**, 5834–5842.

33. Hager,M., Hermann,M., Biehler,K., Krieger-Liszkay,A. and Bock,R. (2002) Lack of the small plastid-encoded PsbJ polypeptide results in a defective water-splitting apparatus of photosystem II, reduced photosystem I levels, and hypersensitivity to light. *J. Biol. Chem.*, **277**, 14031–14039.

34. Drescher,A., Ruf,S., Calsa,T. Jr, Carrer,H. and Bock,R. (2000) The two largest chloroplast genome-encoded open reading frames of higher plants are essential genes. *Plant J.*, **22**, 97–104.

35. Shikanai,T., Shimizu,K., Ueda,K., Nishimura,Y., Kuroiwa,T. and Hashimoto,T. (2001) The chloroplast cplP gene, encoding a proteolytic subunit of ATP-dependent protease, is indispensable for chloroplast development in tobacco. *Plant Cell Physiol.*, **42**, 264–273.

36. Kode,V., Mudd,E.A., Iamtham,S. and Day,A. (2005) The tobacco plastid accD gene is essential and is required for leaf development. *Plant J.*, **44**, 237–244.

37. Legen,J., Wanner,G., Herrmann,R.G., Small,I. and Schmitz-Linneweber,C. (2007) Plastid tRNA genes trnC-GCA and trnN-GUU are essential for plant cell development. *Plant J.*, **51**, 751–762.

38. Kunnimalaiyaan,M. and Nielsen,B.L. (1997) Fine mapping of replication origins (oriA and oriB) in Nicotiana tabacum chloroplast DNA. *Nucleic Acids Res.*, **25**, 3681–3686.

39. Mühlbauer,S.K., Lössl,A., Tzekova,L., Zou,Z. and Koop,H.-U. (2002) Functional analysis of plastid DNA replication origins in tobacco by targeted inactivation. *Plant J.*, **32**, 175–182.

40. Scharff,L.B. and Koop,H.-U. (2007) Targeted inactivation of the tobacco plastome origins of replication A and B. *Plant J.*, **50**, 782–794.

41. Sprinzl,M., Horn,C., Brown,M., Ioudovitch,A. and Steinberg,S. (1998) Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res.*, **26**, 148–153.

42. Marck,C. and Grosjean,H. (2002) tRNomics: analysis of tRNA genes from 50 genomes of Eukarya, Archaea, and bacteria reveals anticodon-sparing strategies and domain-specific features. *RNA*, **8**, 1189–1232.

43. Seong,B.L. and RajBhandary,U.L. (1987) Escherichia coli formylmethionine tRNA: mutations in GGGCCC sequence conserved in anticodon stem of initiator tRNAs affect initiation of protein synthesis and conformation of anticodon loop. *Proc. Natl Acad. Sci. USA*, **84**, 334–338.

44. Maul,J.E., Lilly,J.W., Cui,L., dePamphilis,C.W., Miller,W., Harris,E.H. and Stern,D.B. (2002) The Chlamydomonas reinhardtii plastid chromosome: islands of genes in a sea of repeats. *Plant Cell*, **14**, 2659–2679.

45. Michaud,M., Cognat,V., Duchêne,A.M. and Maréchal-Drouard,L. (2011) A global picture of tRNA genes in plant genomes. *Plant J.*, **66**, 80–93.

46. Grosjean,H., de Crécy-Lagard,V. and Marck,C. (2010) Deciphering synonymous codons in the three domains of life: co-evolution with specific tRNA modification enzymes. *FEBS Lett.*, **584**, 252–264.

47. Helm,M., Brulé,H., Friede,D., Giegé,R., Pütz,D. and Florentz,C. (2000) Search for characteristic structural features of mammalian mitochondrial tRNAs. *RNA*, **6**, 1356–1379.

48. Ueda,T., Suzuki,T., Yokogawa,T., Nishikawa,K. and Watanabe,K. (1994) Unique structure of new serine tRNAs responsible for decoding leucine codon CUG in various Candida species and their putative ancestral tRNA genes. *Biochimie*, **76**, 1217–1222.

49. Santos,M.A.S. and Tuite,M.F. (1995) The CUG codon is decoded in vivo as serine and not leucine in Candida albicans. *Nucleic Acids Res.*, **23**, 1481–1486.

50. Osawa,S., Muto,A., Jukes,T.H. and Ohama,T. (1990) Evolutionary changes in the genetic code. *Proc. R. Soc. Lond.*, **241**, 19–28.

51. Morden,C.W., Wolfe,K.H., dePamphilis,C.W. and Palmer,J.D. (1991) Plastid translation and transcription genes in a non-photosynthetic plant: intact, missing and pseudo genes. *EMBO J.*, **10**, 3281–3288.

52. Wolfe,K.H., Morden,C.W., Ems,S.C. and Palmer,J.D. (1992) Rapid evolution of the plastid translational apparatus in a nonphotosynthetic plant: loss or accelerated sequence evolution of tRNA and ribosomal protein genes. *J. Mol. Evol.*, **35**, 304–317.

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

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

55. de Koning,A.P. and Keeling,P.J. (2006) The complete plastid genome sequence of the parasitic green alga Helicosporidium sp. is highly reduced and structured. *BMC Biol.*, **4**, 1–10.

56. Shinozaki,K., Ohme,M., Tanaka,M., Wakasugi,T., Hayashida,N., Matsubayashi,T., Zaita,N., Chunwongse,J., Obokata,J., Yamaguchi-Shinozaki,K. *et al.* (1986) The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. *EMBO J.*, **5**, 2043–2049.

57. Laslett,D. and Canbäck,B. (2004) ARAGORN, a program for the detection of transfer RNA and transfer-messenger RNA genes in nucleotide sequences. *Nucleic Acids Res.*, **32**, 11–16.