

Why are plastid genomes retained in non-photosynthetic organisms?

Adrian C. Barbrook^{1,*}, Christopher J. Howe^{1,*} and Saul Purton²

¹Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge, UK CB2 1QW

²Algal Research Group, Department of Biology, University College London, London, UK WC1E 6BT

The evolution of the plastid from a photosynthetic bacterial endosymbiont involved a dramatic reduction in the complexity of the plastid genome, with many genes either discarded or transferred to the nucleus of the eukaryotic host. However, this evolutionary process has not gone to completion and a subset of genes remains in all plastids examined to date. The various hypotheses put forward to explain the retention of the plastid genome have tended to focus on the need for photosynthetic organisms to retain a genetic system in the chloroplast, and they fail to explain why heterotrophic plants and algae, and the apicomplexan parasites all retain a genome in their non-photosynthetic plastids. Here we consider two additional explanations: the 'essential tRNAs' hypothesis and the 'transfer-window' hypothesis.

The plastid genome

The plastids of plants, algae and the apicomplexan parasites (such as *Plasmodium* and *Toxoplasma*) owe their origin to one or more symbioses between an oxygenic photosynthetic bacterium and a non-photosynthetic eukaryotic host. In some algal lineages, serial endosymbiotic events occurred in which eukaryotic cells engulfed plastid-bearing eukaryotes and retained the organelle [1]. An early event in the evolution of plastids was the loss of genetic autonomy, with the size and complexity of the endosymbiotic bacterial genome reduced dramatically by the loss of redundant genes, and the transfer of most of the remaining genes to the host nucleus. As a result, the plastids of most modern-day photosynthetic plants and algae possess a small (120–200 kb) circular genome (the 'plastome') containing ~100–250 genes. Most of these genes encode components of the photosynthetic apparatus and the transcription-translation apparatus of the plastid, although genes for other aspects of plastid metabolism such as the biosynthesis of fatty acids, pigments and amino acids are found on some plastomes, with the overall gene content higher in plastomes of red algal descent compared with those of green algae and land plants [2]. A core constituency of ~45 genes are found on the plastome of almost all photosynthetic organisms, and these genes encode many, although not all, of the subunits

of the two photosystems, the cytochrome *b₆f* complex, the ATP synthase, the bacterial-type RNA polymerase and the 70S ribosome, as well as the large subunit of ribulose-bisphosphate carboxylase (rubisco) and rRNAs and tRNAs [3]. An exceptional case is found in the peridinin-containing dinoflagellate algae. The chloroplasts of these algae have lost the conventional plastome, with only a dozen or so genes retained on small (2.5–3.0 kb) plasmids termed 'minicircles' [4]. The retained genes are a subset of those present in other plastids, and encode core subunits of the four membrane complexes of photosynthesis and rRNA; all other genes typically found on plastomes have been transferred to the dinoflagellate nucleus, or might have been lost altogether [4,5].

Plastid-bearing organisms that have lost photosynthetic function during their evolution, such as the holoparasitic plants (Figure 1), heterotrophic algae and the apicomplexan parasites (Figure 2), still retain a plastome, although its size and complexity is reduced further by the loss of most or all the photosynthetic genes. Examples of non-photosynthetic taxa include the euglenoid alga *Astasia longa* [6], which has a 73 kb genome lacking all photosynthesis genes except *rbcL*, and the holoparasitic angiosperm *Epifagus virginiana*, which has a 70 kb plastome having lost not only the photosynthesis genes but also those for the RNA polymerase, some ribosomal proteins and some tRNAs [7]. Other holoparasitic plants with plastomes as small as 20 kb [8] might have undergone even more extensive gene loss, a situation similar to that seen in plant cell cultures generated in the laboratory, where loss of photosynthetic capacity is often associated with extensive plastome deletions (e.g. [9–12]). Finally, there are the apicomplexan protozoa, which were recognized as having a photosynthetic ancestry when it became apparent that a 35 kb extrachromosomal DNA molecule was a remnant plastome [13,14]. This genome has also lost the photosynthesis genes, but retains a reduced set of genes for subunits of the RNA polymerase, the ribosome and tRNAs [15].

No photosynthetic genes, so no need for a plastome?

There have been several suggestions as to why some genes have remained in plastids despite the mass exodus to the nucleus. As discussed in Box 1, the core components of the photosynthetic complexes probably need to be synthesized within the chloroplast for one or more of the following reasons: (i) they are hydrophobic membrane proteins

Corresponding author: Purton, S. (s.purton@ucl.ac.uk).

* These authors contributed equally to this Review.

Available online 9 January 2006

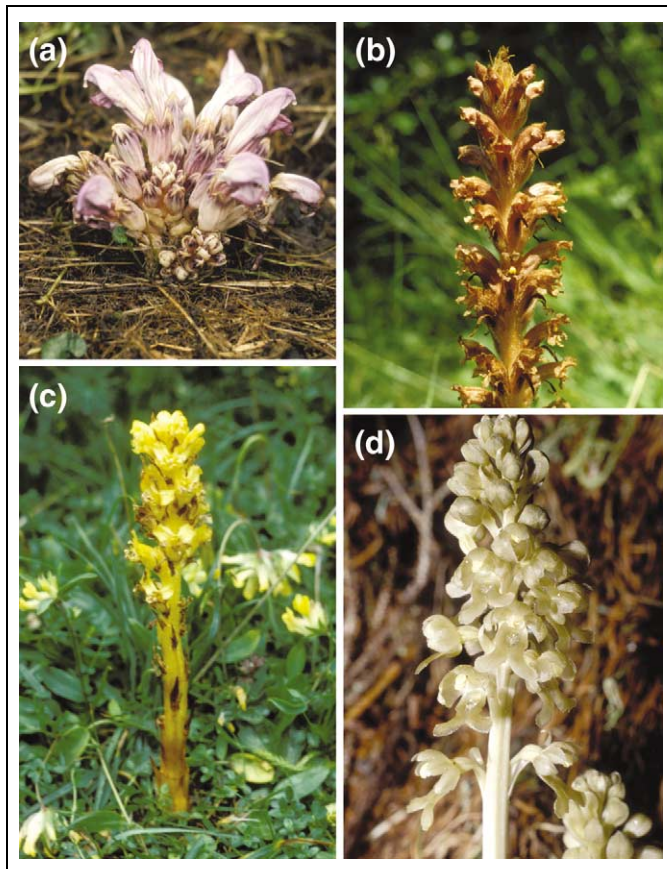


Figure 1. Examples of non-photosynthetic angiosperms: (a) purple toothwort (*Lathraea clandestina*), (b) knapweed broomrape (*Orobanche elatior*), (c) yellow broomrape (*Orobanche lutea*) and (d) birdsnest orchid (*Neottia nidus-avis*). The first three are plant parasites whereas the orchid is an epiparasite that derives its fixed carbon indirectly from autotrophic plants via a specific fungal partner [62]. Photographs kindly provided by Alan Outen.

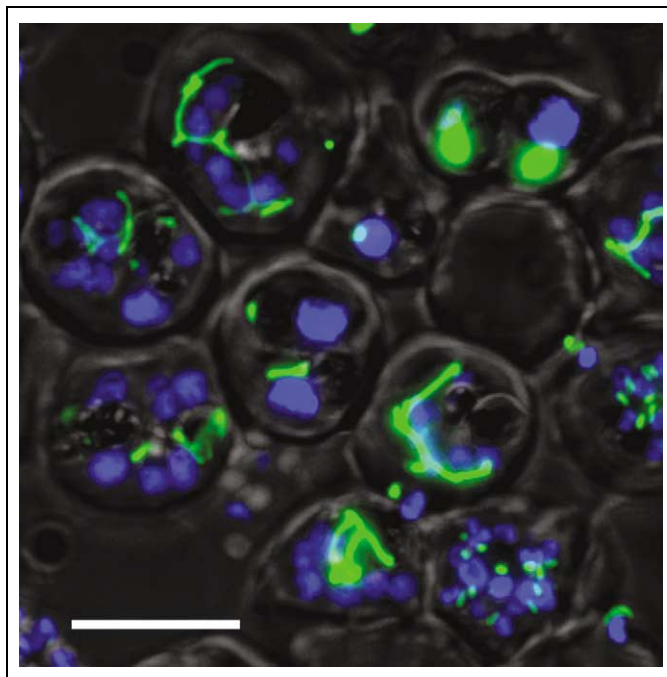


Figure 2. *Plasmodium falciparum*, the plastid-bearing malaria parasite. A fluorescence image of human erythrocytes transfected with *P. falciparum* expressing a green fluorescent protein (ACP-GFP) targeted to the plastid. The parasite nuclei are stained blue with the dye Hoechst 33342. Scale bar = 10 μ m. Figure kindly provided by Shigeharu Sato and Iain Wilson (NIMR, London, UK).

whose import from the cytosol would be problematic; (ii) their rate of synthesis can be regulated by the redox state within individual plastids; (iii) their synthesis and assembly into a complex are coupled and tightly regulated, and the transfer to the nucleus of the genes for these core components would result in the loss of this coordinated regulation. Although it is important to bear in mind that no one hypothesis for why plastid genes are retained will apply to all plant, algal and protist lineages, it is clear that the above arguments address primarily the need to keep photosynthesis genes in the organelle. Given what we know of the core set of genes retained in photosynthetic plastids, including those of dinoflagellates, once photosynthetic competence has been lost there seems to be no compelling reason why the other genes should be retained. *A priori*, if we were to remove all the photosynthesis-related genes from a typical plastome, and exclude all the genes for components of the transcription-translation apparatus, we would be left with relatively few protein-coding genes. For all these remaining genes, we can find examples of successful transfer to the nucleus (or loss of the gene altogether) amongst various lineages [3,5], and none of the genes appears to encode a particularly hydrophobic protein, or a protein that might require redox regulation from within the plastid. Consequently, one might predict that the loss of photosynthesis would allow the loss of the plastid genetic system in its entirety. This is clearly not the case. In this review, we consider alternative ideas to account for the retention of a plastome in non-photosynthetic plastid-bearing organisms: namely, the 'essential tRNAs' hypothesis and the 'limited transfer window' hypothesis. For reasons that will become clear, we need to look first at possible fates of plastid RNA genes.

Is the transfer of plastid RNA genes to the nucleus forbidden?

Although there is overwhelming evidence for the mass transfer of protein-coding genes from the plastid to the nucleus [3], there is no evidence as far as we can determine for the transfer of plastid genes encoding RNA products, whether for rRNA, tRNA, or RNA components of ribonucleoproteins such as the signal recognition particle (SRP), which is involved in targeting proteins to the thylakoid membrane [16], and ribonuclease P, which is required for tRNA 5' maturation [17]. Sequences for plastid RNA genes are indeed found in the nuclear genomes of various plants, and are clearly the result of integration of large tracts of plastid DNA [18–20], but there is no evidence that any of these genes are transcribed to yield a functional RNA product that is re-imported into the plastid. Similarly, there is no evidence of functional mitochondrial-to-nuclear transfer of RNA genes, although inter-organellar transfer of tRNA genes from plastid to mitochondrion has been documented in plants [21]. It is unclear why the successful transfer of organellar RNA genes to the nucleus does not appear to have occurred. Because nuclear expression would require the presence of the RNA products in the nucleo-cytosol before import by the organelle, one could speculate that this would be toxic to the cell because the (prokaryotic)

Box 1. Existing hypotheses – transportability, redox control, complex assembly and codon disparity

One of the first hypotheses proposed for the retention of genomes in organelles, including chloroplasts, was that some proteins would be 'untransportable' [47] into the organelle if the gene were in the nucleus. This idea was modified by Gunnar von Heijne [48], who suggested that proteins encoded within the mitochondrion contained hydrophobic stretches that would result in their export through the endoplasmic reticulum if the proteins were synthesized in the cytosol. Other features of organelle-encoded proteins have also been proposed as likely to block their import into the organelle. However, the successful import of several hydrophobic nuclear-encoded chloroplast proteins, such as the light-harvesting chlorophyll *a,b*-binding proteins, runs counter to this hypothesis. Furthermore, the PsbA and RbcL proteins can be imported into the chloroplasts of transgenic plants in which the chloroplast genes have been inserted in the nucleus and sequences encoding plastid-targeting peptides incorporated [49,50]. These relocation experiments did not give wild-type levels of activity, which might indicate inefficient import (rather than, say, inefficient assembly). It is also possible that some hydrophobic proteins are toxic if expressed in a different location from where they actually function [51].

A second hypothesis, 'CORR' or CO-location for Redox Regulation, is that genes for proteins involved in redox reactions are retained in organelles to allow efficient coupling of gene regulation to balance the redox reactions [52]. A redox imbalance (for example, in the relative activities of photosystems I and II) can therefore be rectified rapidly. By contrast, transmitting a signal to the nucleus for expression of a particular gene for a plastid redox protein would delay the response and the gene product would go to all chloroplasts or mitochondria in the cell, whether they needed it or not. The demonstrated redox regulation of chloroplast gene expression, as well as the retention of

genes for key polypeptides of the photosynthetic light reactions in the highly reduced plastid genomes of dinoflagellates are consistent with CORR. However, CORR does not readily explain the retention of a genome in plastids of organisms such as *Plasmodium*. Arguably, a plastid gene location might also be beneficial in allowing direct control of expression in response to other aspects of plastid physiology, not just redox poise [32], so we could extend the CORR hypothesis to 'CO-location for Biochemical RegulAtion' (COBRA). However, it is not obvious which *Plasmodium* plastid genes would require control in this way [13].

An additional argument put forward by Bill Zerges [53] proposes that the highly coordinated synthesis and assembly of the thylakoid membrane complexes from plastid-encoded subunits precludes the transfer of the genes to the nucleus because the regulatory mechanisms that control translation of one subunit in relation to its partner subunits would be lost. Furthermore, co-translation of these subunits in the cytosol and partial assembly might not allow import into the organelle.

Finally, the 'code disparity hypothesis' [54] proposes that codon modifications in some plastid genes (i.e. divergences in the genetic code; codons that are 'corrected' at the transcript level through RNA editing, or those genes that use a non-canonical start codon such as GUG) act as additional barriers to functional gene transfer. Relocation of these genes to the nucleus would thus require additional DNA mutations to restore the coding sequence to that of the universal code. Examples of possible deviations from the universal code in plastid genes have been reported only in some apicomplexans [41], although examples of non-canonical start codons are found in many plastomes. In the case of RNA editing, this process appears to be restricted to the plastomes of land plants [55] and possibly some dinoflagellates [56].

organellar RNAs would interfere with the functioning of their eukaryotic counterparts, and thus with fundamental aspects of cell biology (e.g. the assembly and functioning of the 80S ribosome). Alternatively, the nucleo-cytosol might be too hostile an environment for the RNA species, with ribonucleases degrading these 'foreign' RNAs before they are able to make it to the sanctuary of the organelle.

If we assume that transfer to the nucleus is forbidden, then there are three possible fates for organellar RNA genes during the evolution of the plastome. The first is the retention of those genes that are absolutely essential for the genetic system of the plastid, but the loss of those that are surplus to requirements. In support of this is the observation that all plastomes sequenced to date retain the major rRNA genes *rrn16* and *rrn23* (although the assignment of the *rrn16* gene in the minimalist plastome of dinoflagellates is somewhat speculative [4]) together with a set of 32 or fewer tRNA genes. For most species, it is assumed that this represents the minimal set required for translation of all 61 codons, with the 25 tRNA genes found in the highly reduced apicomplexan genome probably representing the absolute minimum [15]. The second possible fate is that a plastid gene is lost and that the RNA product is replaced by a eukaryotic counterpart from the cytosol. This is presumed to be the case for certain tRNA species in the parasitic plant *Epifagus virginiana*, where an incomplete set of 17 tRNA genes is found on the plastome [7]. A similar requirement for tRNA import is likely for the parasite *Orobancha minor* [22], and perhaps also the dinoflagellate plastids [4]. It is certainly the case that import of cytosolic tRNAs into mitochondria occurs in many species [23]. A third possible fate is that

requirement for the RNA, together with its gene, is lost. This is illustrated by the RNA components of SRP and RNase P. In some algal groups the genes for these RNAs have been retained on the plastome, whereas in higher plants and some green algae the genes are absent [16,17]. In the latter cases, it would appear that the RNA subunit of each of the ribonucleoproteins has become redundant, with the catalytic function transferred completely to the protein component [24].

So, our hypothesis is that plastid RNA genes cannot be transferred to the nucleus and retain their function. They must be kept in the plastid as long as their products are both essential for the organelle and cannot be replaced with their cytosolic orthologues. With this in mind, we can now look at why plastome loss has failed to occur in non-photosynthetic plastids.

The 'essential tRNAs' hypothesis, part 1: plastid tRNA^{Glu} is essential for haem biosynthesis

The key precursor in the biosynthetic pathway for tetrapyrroles such as haem and chlorophyll is δ -amino-laevulinic acid (ALA). In plants and algae, ALA synthesis takes place in the plastid with the first steps involving activation of glutamate by its cognate tRNA followed by conversion of the glutamyl moiety to glutamate-1-semialdehyde and then to ALA (Figure 3). The tRNA is encoded by the plastid gene *trnE*, and is unique amongst all the plastid tRNAs in that it has a role in both tetrapyrrole biosynthesis and protein biosynthesis [25]. As first pointed out by Chris Howe and Alison Smith in 1991 [26], non-photosynthetic plants would therefore need to retain *trnE*, as well as the machinery for its transcription, to

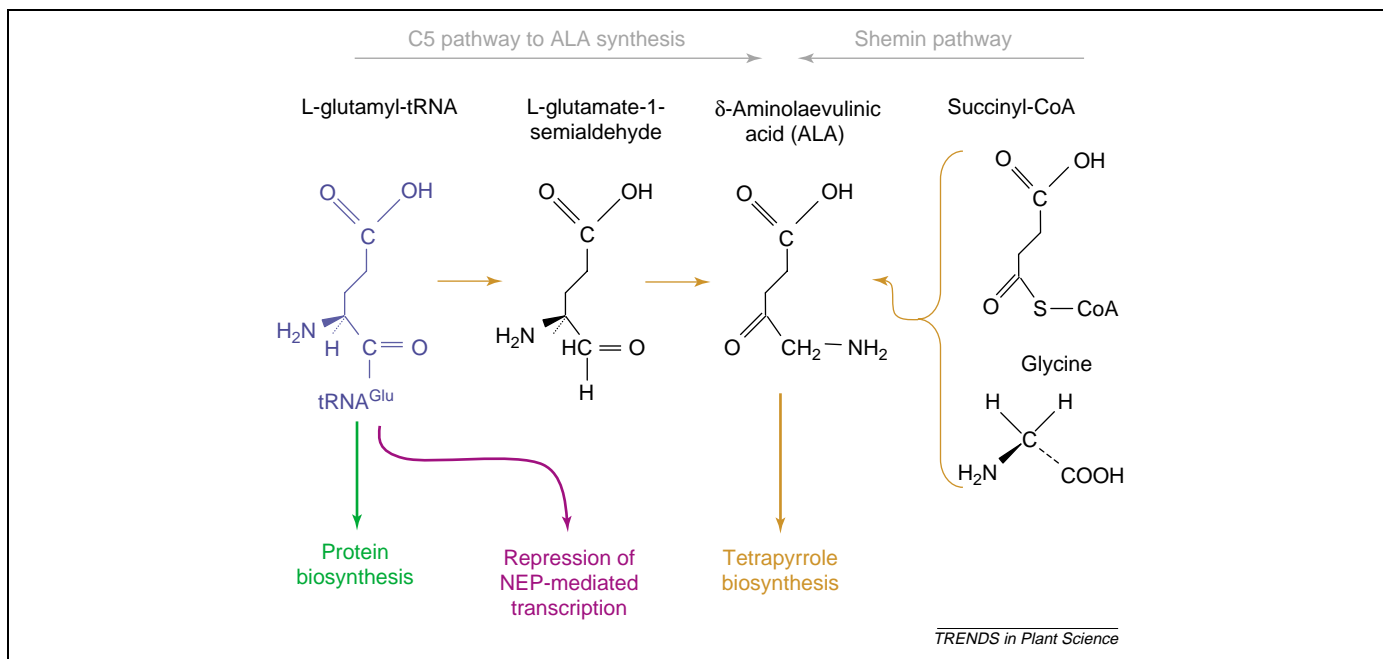


Figure 3. The multiple roles of glutamyl-tRNA in higher plant plastids. (i) In plants, algae and most bacteria, the charged tRNA (depicted in blue) is involved in the C5 synthesis pathway for δ -aminolaevulinic acid (ALA), the precursor of haem, chlorophyll and other tetrapyrroles. (In animals, fungi and many protists, ALA is made via the Shemin pathway, and does not involve glutamyl-tRNA, but rather the condensation of succinyl-CoA and glycine). (ii) The same glutamyl-tRNA is also involved in plastid protein synthesis. (iii) This particular amino acyl-tRNA is also involved in the repression of the nuclear-encoded plastid RNA polymerase (NEP) during chloroplast development [63]. Although protein synthesis and repression of NEP-mediated transcription are, in theory, dispensable in non-photosynthetic plastids, the biosynthesis of haem is not.

synthesize the haem component of mitochondrial cytochromes, P450 cytochromes and other essential oxidative enzymes. At that time, it was believed that plastid transcription was carried out exclusively by the eubacterial-type RNA polymerase encoded by the plastid *rpo* genes. Transcription of *trnE* would therefore require the expression of these *rpo* genes, together with the plastid genes for the tRNAs, rRNAs and ribosomal proteins necessary for the synthesis of the RNA polymerase. The subsequent discovery of a second, exclusively nuclear-encoded plastid RNA polymerase (a 'NEP') in angiosperms [27] suggests that all these additional plastid genes can be lost and retention of *trnE* on the plastome is all that is necessary to ensure haem biosynthesis in non-photosynthetic plants. Conversely, the apparent lack of a NEP in algal plastids [28] indicates that the whole plastid transcription-translation machinery would be necessary for haem biosynthesis in heterotrophic algae. Evidence for the essential nature of *trnE* comes from studies of plastid deletions in non-photosynthetic plant and algal mutants. In mutant lines of higher plants generated in tissue culture, large-scale deletions result in the loss of numerous plastid genes, including ribosomal genes (and consequently, loss of protein synthesis), but all deletions appear to leave *trnE* intact [9–11]. In one example, two rice cell lines were reported to retain overlapping segments of the plastid genome that were each less than 20 kb [12]. Analysis of the ~3 kb overlapping region has shown that it contains *trnE* together with four neighbouring tRNA genes, and that *trnE* is actively transcribed in these lines. By contrast, a correlation of the restriction fragments affected by 5-fluorodeoxyuridine-induced deletions in the green alga *Chlamydomonas reinhardtii* [29] with the complete plastid genome sequence [30] shows

that the deletions all appear to leave the RNA polymerase, ribosomal protein and tRNA genes intact, as predicted for an organism lacking a NEP.

We have argued above that the plastid *trnE* gene could not be relocated to the nucleus. The tRNA^{Glu} is presumably indispensable for ALA formation, but why could it not be replaced by an imported cytosolic 80S-type tRNA^{Glu}, thereby making *trnE* redundant? Because of its role in ALA synthesis, the plastid tRNA^{Glu} has to interact with glutamyl-tRNA reductase, as well as glutamyl-tRNA synthetase and elongation factor EF-Tu. It seems unlikely that the cytosolic counterpart could easily replace this dual-function tRNA; a point that is illustrated by the finding that a single base change in *trnE* from *Euglena* results in the loss of haem biosynthesis but leaves protein synthesis unaffected [31].

If *trnE* is essential for haem biosynthesis, and its product cannot be functionally replaced with a cytosolic tRNA^{Glu}, plastids must retain a genome carrying at least this single gene. This could explain the situation in non-photosynthetic plants, and leads to a prediction that the plastomes of holoparasitic plants such as *Epifagus virginiana* might eventually relinquish to the nucleus the few protein-coding genes that are essential for plastid function [7]. The remaining RNA and protein-coding genes of the translation apparatus will then become redundant and be lost, with only *trnE* maintained in the plastid on a replicating DNA minicircle transcribed by the imported NEP. Direct evidence for this process could come from sequence analysis of the highly reduced plastid genomes of other holoparasitic plant species [8,32].

The *trnE* argument does not explain the retention of a plastome in apicomplexan parasites such as *Plasmodium*. In these organisms, ALA biosynthesis does not involve

the plastidic tRNA^{Glu} [13]. Rather, synthesis occurs in the mitochondrion and involves the condensation of glycine and succinyl-CoA (Figure 3). It has been pointed out that the apicomplexan plastome possesses several protein-coding genes [principally, *clpC* and *sufB* (also known as *ycf24*)] whose products are likely to play an essential role in the plastid. As discussed by Iain Wilson [13], this in itself is not an explanation for why the plastome is retained because both genes have successfully transferred to the nucleus in plants and green algae [3]. We must invoke different arguments for why the apicomplexan plastid still has its genome.

The 'essential tRNAs' hypothesis, part 2: plastid tRNA^{fMet} is essential for mitochondrial protein synthesis

The mitochondrial genome of *Plasmodium* is remarkable in that it is the smallest organellar genome known (only 6 kb) and encodes just three components of the respiratory apparatus (the cytochrome *b* subunit of the cytochrome *bc*₁ complex, and subunits I and III of cytochrome *c* oxidase) and a fragmented set of rRNAs [15]. No tRNA genes are present on the genome and, given that the protein-coding genes need a complete set of tRNAs for their translation, it has been generally assumed that all the necessary tRNAs are imported from the cytosol. Indeed, import studies using *Toxoplasma* mitochondria indicate that this is the case for most of the tRNAs [33]. However, there is one particular tRNA, the initiator tRNA, which cannot be directly substituted by its cytosolic equivalent. In both eukaryotes and prokaryotes, translation initiation begins at a methionine codon, but involves an initiator tRNA distinct from the

methionyl-tRNA species (tRNA^{Met-e}) used in peptide chain elongation. In eubacteria and in organelles, the initiator tRNA (encoded by *trnfm*) is used as formylmethionyl-tRNA; the attached methionine moiety being formylated by a methionyl-tRNA formyltransferase (MTF). By contrast, translation initiation on 80S ribosomes uses an initiator methionyl-tRNA without formylation [34]. Initiating translation in a mitochondrion that has lost the *trnfm* gene therefore presents a fundamental problem. One solution is seen in the trypanosomatids (which lack plastids) whose mitochondrial genomes also lack all tRNA genes and which depend on the cytosolic pool for tRNAs. Here a fraction of the imported tRNA^{Met-e} is formylated by an unusual MTF, allowing this tRNA to be used as the initiator [35]. *Plasmodium* has only a single MTF gene in the nuclear genome [33], so adopting the same solution as trypanosomatids would require MTF to be targeted both to the plastid (to formylate the initiator tRNA encoded on the plastome) and the mitochondrion (to formylate the elongator tRNA imported from the cytosol). It seems unlikely that the one enzyme would readily evolve both dual targeting capability and dual specificity for two different tRNA substrates. We suggest that the apicomplexans might have solved the need for tRNA^{fMet} in a much more straightforward way – by simply exploiting the pool of formylmethionyl-tRNA already present in the plastid. Therefore, in our model, the charged initiator tRNA is imported from the plastid into the mitochondrion and used in conjunction with the imported cytosolic tRNAs to synthesize the three core components of the respiratory apparatus (Figure 4). This is an elegant solution but has profound evolutionary consequences.

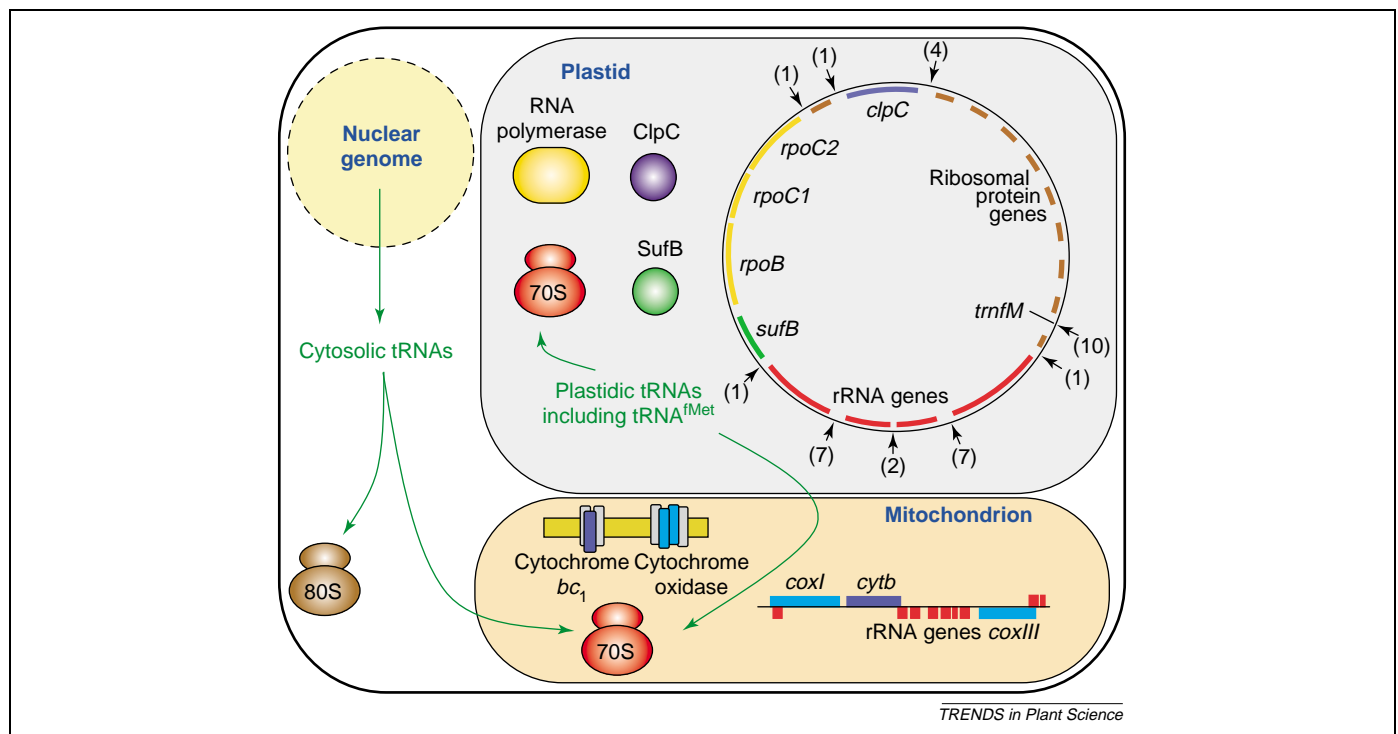


Figure 4. Organellar genetics in *Plasmodium*. Simplified versions of the plastid and mitochondrial genomes of the parasite are shown (see [15] for details). The different gene classes are colour-coded and the clusters of tRNA genes on the plastid genome are arrowed (the number of genes is in parenthesis). The products of the organellar genes are shown to the left of each genome. According to our hypothesis, the genetic system of the plastid is essential because it provides the formylmethionine-charged initiator tRNA for biosynthesis of the mitochondrially encoded components of the two cytochrome complexes, with the rest of the tRNAs being imported from the cytosol. The *clpC* and *sufB* genes have remained on the *Plasmodium* plastid genome primarily because transfer of DNA to the nucleus through plastid lysis is likely to be lethal in organisms with only a single plastid.

Unlike the situation with the higher plant *trnE*, the expression of *trnfM* in the *Plasmodium* plastid involves numerous other plastid gene products because the apicomplexan nuclear genome does not encode a plastid-targeted NEP. Transcription of *trnfM* is therefore dependent on the plastid-encoded RNA polymerase which itself must be synthesized on plastid ribosomes. Hence the plastome has retained the genes for many of these components. In principle, all such housekeeping genes could be transferred to the nucleus (and indeed several such as *rpoA* and *rpl22* have [15]); however, these genes have probably remained in the plastome for other reasons, as discussed in the next section.

This hypothesis has a beautiful irony. When the 35 kb circular DNA element that we now know to be the apicomplexan plastid genome was first described more than 30 years ago it was generally assumed to be part of the mitochondrial genetic system. Numerous subsequent studies in the 1970s and 1980s came to the same conclusion, and, when the 6 kb DNA element was identified as mitochondrial DNA, a bipartite mitochondrial genome was proposed (see [15]). Although that idea was discarded once the 35 kb element was shown to have a plastid origin and to be located in the plastid, our hypothesis resurrects the bipartite idea with the ultimate function of the plastome being to provide the essential initiator tRNA (and possibly other tRNAs) for translation of the three mitochondrial genes.

So is there any experimental evidence for such inter-organellar genetics and the sharing of tRNAs, as depicted in Figure 4? Ultrastructural studies of *P. falciparum* during the blood stages of the life cycle of the parasite show that the plastid is always in physical contact with the mitochondrion [36,37], which could allow direct transfer of plastid tRNAs between the two organelles without passage through the cytosol. Additional evidence comes from genetic studies in *Chlamydomonas*. This alga also has a highly reduced mitochondrial genome with only three tRNA genes, none of which encodes the initiator tRNA [38]. In an elegant genetic screen, Pierre Bennoun and Monique Delosme [39] recovered chloroplast suppressors of a mitochondrial mutation. Given that one of the chloroplast suppressors also suppressed a missense mutation in the chloroplast *rbcL* gene, it is most likely that the suppressor is a mutated tRNA that is shared between the two organelles. Molecular analysis of the mitochondrial mutation and the chloroplast suppressor is required to confirm this idea, but the work does provide the first indication that plastids are able to donate tRNAs to the mitochondrion.

Although we do not yet have a complete dinoflagellate mitochondrial genome sequence, studies to date have failed to demonstrate the existence of mitochondrial genes for tRNAs [38]. Assuming there is indeed no mitochondrial *trnfM*, and given the evolutionary relationship between Apicomplexa and dinoflagellates, our hypothesis predicts the existence of a *trnfM* gene in the plastids of dinoflagellates. Recently, we have found just such a gene on one of the 'empty' minicircles in *Amphidinium operculatum* (A.C. Barbrook *et al.*, unpublished), so we predict that the encoded tRNA^{fMet} is involved in both

plastid and mitochondrial protein synthesis in this dinoflagellate.

The 'limited transfer window' hypothesis

To our minds, the need to retain the *trnfM* gene does not explain fully why the gene content of the apicomplexan plastome has remained so highly conserved. Comparison of sequenced plastomes from a range of species including *Plasmodium falciparum* [15], *Theileria parva* [40] and *Eimeria tenella* [41] reveals a highly conserved gene content even though these species represent divergent taxa within the Apicomplexa, and one might expect that different protein-coding genes would transfer to the nucleus in different lineages. One argument for such gene conservation is that the opportunity for DNA transfer is greatly reduced in protists such as the apicomplexans and *Chlamydomonas* because they possess only a single plastid per cell, and such transfers would almost always be lethal to the cell [20,42,43].

During the early stages of the endosymbiosis that gave rise to the plastid, endosymbiont division was probably not coordinated with that of the host cell, resulting in many endosymbiont cells within the host. Some of these cells presumably lysed, resulting in the spillage of bacterial DNA into the host cytoplasm and the integration of much of this DNA into the nuclear genome of the host. This high level of flux of bacterial DNA to the nucleus provided the opportunity for the creation of functional copies of essential endosymbiont genes in the nuclear genome [44]. Once the host nucleus established control of plastid division, algae with single plastids such as *Chlamydomonas* could evolve. This reduction of plastid number to one-per-cell for the major part of the cell cycle effectively closed (or severely limited) the transfer window because lysis of this plastid would be lethal to that particular cell. Other organisms with a single plastid for most, if not all, of their life cycle would similarly have a low rate of movement of DNA from plastid to nucleus. This includes the Apicomplexa.

Two lines of evidence support the idea of a marked reduction in plastid-nuclear DNA flux in single-plastid organisms. First, transgenic studies in tobacco (whose leaves contain hundreds of chloroplasts per cell) showed that marker DNA inserted into the plastome is transferred to the nucleus at a surprisingly high rate [45,46]. Conversely, similar studies in *Chlamydomonas* suggested that transfer rates in this organism are at least several orders of magnitude lower [42]. Second, analysis of plant nuclear genomes reveals extensive evidence of plastid DNA-derived sequences (NUPTs), suggesting the continuous influx of plastid DNA into plant genomes [18,19]. By contrast, NUPTs are much rarer in the nuclear genomes of *Chlamydomonas* and *Plasmodium* [20]. Therefore, the retention of genes (including the essential *clpC* and *sufB*) in the *Plasmodium* plastome might not reflect a need to have them expressed in the plastid [13], but rather an inability to get them out.

Conclusions

In addition to the existing explanations for the retention of a plastome, we propose additional explanations that

Box 2. Meeting the criteria for a plastome-less plastid

To date there is no convincing report of any plastid-bearing organism that lacks a plastome. This contrasts with the situation with the sister organelle of the plastid, the mitochondrion. Studies of anaerobic microbes reveal cryptic organelles termed hydrogenosomes or mitosomes that are believed to have evolved from mitochondria that have lost respiratory function. In most cases, these organelles appear to have also lost the organelle genome [57]. Using the various hypotheses put forward for why plastid genes are retained, including those considered in this review, we can propose a set of criteria that might allow us to discover an organism with a plastome-less plastid.

- Criterion 1: the organism must have dispensed completely with photosynthetic function. This eliminates all autotrophic plants, hemi-parasitic plants and photosynthetic algae from our search.
- Criterion 2: the organism must be able to synthesize ALA via the Shemin pathway (Figure 3) and therefore not be dependent on the plastid-encoded glutamyl-tRNA for ALA synthesis. This probably excludes all higher and lower plants (including holo-parasitic plants), as well as many algal species. Examples of plastid-bearing protists that use the Shemin pathway include *Plasmodium falciparum* [13] and the euglenoid alga *Euglena gracilis* (which uses both pathways [58]).
- Criterion 3: the organism must either possess a mitochondrially encoded formylmethionine-tRNA (e.g. higher plants [13]), or must have evolved a MTF enzyme capable of modifying a cytosolic methionyl-tRNA imported into the mitochondrion (compare with trypanosomes), or must have dispensed with the need for mitochondrial protein synthesis (hydrogenosome- or mitosome-containing protists). This criterion appears to rule out *Plasmodium* and closely related genera that possess a minimal mitochondrial genome lacking tRNA genes [22].

- Criterion 4: although not a strict criterion, the transfer-window hypothesis implies that the presence of more than one plastid per organism would greatly facilitate the transfer of any essential protein-coding genes to the nucleus, and therefore accelerate the loss of the plastome.

Earlier reports have suggested that a plastome might be lacking in certain parasitic plants such as *Rafflesia* [8]. However, this suggestion is based on negative results when screening for various plastid genes; we suggest that a search for *trnE* would yield a positive result. Similarly, it has been argued that the apicomplexan *Cryptosporidium parvum* lacks a plastome [59]. The recent genome project for *C. parvum* [60] has confirmed this, but also indicates that the parasite lacks a plastid. It is likely that the common ancestor of all apicomplexans possessed a plastid [61] but that this was secondarily lost early in the lineage leading to *C. parvum*. Interestingly, this protist has also dispensed with a mitochondrial genome, thereby removing one of the potential barriers to plastome loss – the need to supply tRNA^{fMet} to the mitochondrion (criterion 3). It is possible therefore that a search of other apicomplexans that lack respiratory function could reveal species that have lost their plastome but retained the plastid compartment.

Alternatively, it might prove fruitful to search amongst the non-photosynthetic species of euglenoid algae because three of the four criteria (1, 2 and 4) appear to be met. What is not known at present is whether euglenoid mitochondrial genomes possess a *trnFM* gene [38]. *Astasia longa*, the one non-photosynthetic euglenoid alga studied so far, does have a plastome, possibly because it has retained rubisco activity and therefore needs to maintain the *rbcl* gene encoding the hydrophobic subunit of this CO₂-fixing enzyme [6].

address specifically the situation in non-photosynthetic organisms. Different explanations will apply to different lineages, and in some lineages, more than one will be relevant, as illustrated in Box 2. Our hypotheses are testable. We predict that the sequencing of plastomes from additional holoparasite species will reveal highly reduced genomes in which *trnE* is the only gene that is always present. Similarly, plant cell lines carrying plastome deletions will always retain *trnE*. Molecular-genetic and organelle-fractionation studies using *Chlamydomonas* or apicomplexan species should show that the initiator tRNA, and possibly other tRNAs, are shared between the plastid and mitochondrion. Finally, the sequencing of the nuclear genomes of organisms containing a single plastid should always reveal a low abundance of NUPTs.

Acknowledgements

Work in our laboratories is supported by the Leverhulme Trust and the Biotechnology and Biological Sciences Research Council. We thank Alan Outen, Iain Wilson and Shigeharu Sato for providing the photographs used in Figures 1 and 2.

References

- Palmer, J.D. (2003) The symbiotic birth and spread of plastids: how many times and whodunit? *J. Phycol.* 39, 4–11
- Grzebyk, D. *et al.* (2003) The Mesozoic radiation of eukaryotic algae: the portable plastid hypothesis. *J. Phycol.* 39, 259–267
- Martin, W. *et al.* (2002) 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
- Koumandou, V.L. *et al.* (2004) Dinoflagellate chloroplasts – where have all the genes gone? *Trends Genet.* 20, 261–267
- Bachvaroff, T.R. *et al.* (2004) Dinoflagellate expressed sequence tag data indicate massive transfer of chloroplast genes to the nuclear genome. *Protist* 155, 65–78
- Gockel, G. and Hachtel, W. (2000) Complete gene map of the plastid genome of the nonphotosynthetic euglenoid flagellate *Astasia longa*. *Protist* 151, 347–351
- Wolfe, K.H. *et al.* (1992) Function and evolution of a minimal plastid genome from a nonphotosynthetic parasitic plant. *Proc. Natl. Acad. Sci. U. S. A.* 89, 10648–10652
- Nickrent, D.L. *et al.* (1997) Do non-asterid holoparasitic flowering plants have plastid genomes? *Plant Mol. Biol.* 34, 717–729
- Day, A. and Ellis, T.H.N. (1984) Chloroplast DNA deletions associated with wheat plants regenerated from pollen: possible basis for maternal inheritance in chloroplasts. *Cell* 39, 359–368
- Abe, T. *et al.* (2002) Large deletions in chloroplast DNA of rice calli after long-term culture. *J. Plant Physiol.* 159, 917–923
- Cahoon, A.B. *et al.* (2003) Maize BMS cultured cell lines survive with massive plastid gene loss. *Curr. Genet.* 44, 104–113
- Chiba, T. *et al.* (1996) Transcription of tRNA genes from a large-scale plastid DNA deletion clearly reveals the action of nuclear-encoded RNA polymerase in the plastid. *J. Plant Physiol.* 148, 652–656
- Wilson, R.J.M. (2005) Parasite plastids: approaching the endgame. *Biol. Rev. Camb. Philos. Soc.* 80, 129–153
- Howe, C.J. (1992) Plastid origin of an extrachromosomal DNA molecule from *Plasmodium*, the causative agent of malaria. *J. Theor. Biol.* 158, 199–205
- Wilson, R.J.M. and Williamson, D.H. (1997) Extrachromosomal DNA in the Apicomplexa. *Microbiol. Mol. Biol. Rev.* 61, 1–16
- Rosenblad, M.A. and Samuelsson, T. (2004) Identification of chloroplast signal recognition particle RNA genes. *Plant Cell Physiol.* 45, 1633–1639
- de la Cruz, J. and Vioque, A. (2003) A structural and functional study of plastid RNAs homologous to catalytic bacterial RNase P RNA. *Gene* 321, 47–56
- Matsuo, M. *et al.* (2005) The rice nuclear genome continuously integrates, shuffles, and eliminates the chloroplast genome to cause chloroplast-nuclear DNA flux. *Plant Cell* 17, 665–675

- 19 Noutsos, C. *et al.* (2005) Generation and evolutionary fate of insertions of organelle DNA in the nuclear genomes of flowering plants. *Genome Res.* 15, 616–628
- 20 Richly, E. and Leister, D. (2004) NUPTs in sequenced eukaryotes and their genomic organization in relation to NUMTs. *Mol. Biol. Evol.* 21, 1972–1980
- 21 Marienfeld, J. *et al.* (1999) The mitochondrial genome of *Arabidopsis* is composed of both native and immigrant information. *Trends Plant Sci.* 4, 495–502
- 22 Lohan, A.J. and Wolfe, K.H. (1998) A subset of conserved tRNA genes in plastid DNA of nongreen plants. *Genetics* 150, 425–433
- 23 Schneider, A. and Maréchal-Drouard, L. (2000) Mitochondrial tRNA import: are there distinct mechanisms? *Trends Cell Biol.* 10, 509–513
- 24 Schuenemann, D. *et al.* (1998) A novel signal recognition particle targets light-harvesting proteins to the thylakoid membranes. *Proc. Natl. Acad. Sci. U. S. A.* 95, 10312–10316
- 25 Kumar, A.M. *et al.* (1996) Glutamyl-transfer RNA: at the crossroad between chlorophyll and protein biosynthesis. *Trends Plant Sci.* 1, 371–376
- 26 Howe, C.J. and Smith, A.G. (1991) Plants without chlorophyll. *Nature* 349, 109
- 27 Hess, W.R. and Borner, T. (1999) Organellar RNA polymerases of higher plants. *Int. Rev. Cytol.* 190, 1–59
- 28 Smith, A.C. and Purton, S. (2002) The transcriptional apparatus of algal plastids. *Eur. J. Phycol.* 37, 301–311
- 29 Myers, A.M. *et al.* (1982) Mutants of *Chlamydomonas reinhardtii* with physical alterations in their chloroplast DNA. *Plasmid* 7, 133–151
- 30 Maul, J.E. *et al.* (2002) The *Chlamydomonas reinhardtii* plastid chromosome: islands of genes in a sea of repeats. *Plant Cell* 14, 2659–2679
- 31 Stange-Thomann, N. *et al.* (1994) A point mutation in *Euglena gracilis* chloroplast tRNA^{Glu} uncouples protein and chlorophyll biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* 91, 7947–7951
- 32 Bungard, R.A. (2004) Photosynthetic evolution in parasitic plants: insight from the chloroplast genome. *BioEssays* 26, 235–247
- 33 Esseiva, A.C. *et al.* (2004) Mitochondrial tRNA import in *Toxoplasma gondii*. *J. Biol. Chem.* 279, 42363–42368
- 34 Rajbhandary, U.L. (1994) Initiator transfer RNAs. *J. Bacteriol.* 176, 547–552
- 35 Tan, T.H.P. *et al.* (2002) Eukaryotic-type elongator tRNA^{Met} of *Trypanosoma brucei* becomes formylated after import into mitochondria. *Proc. Natl. Acad. Sci. U. S. A.* 99, 1152–1157
- 36 Hopkins, J. *et al.* (1999) The plastid in *Plasmodium falciparum*. Asexual blood stages: a three-dimensional ultrastructural analysis. *Protist* 150, 283–295
- 37 van Dooren, G.G. *et al.* (2005) Development of the endoplasmic reticulum, mitochondrion and apicoplast during the asexual life cycle of *Plasmodium falciparum*. *Mol. Microbiol.* 57, 405–419
- 38 Gray, M.W. *et al.* (2004) Mitochondria of protists. *Annu. Rev. Genet.* 38, 477–524
- 39 Bennoun, P. and Delosme, M. (1999) Chloroplast suppressors that act on a mitochondrial mutation in *Chlamydomonas reinhardtii*. *Mol. Gen. Genet.* 262, 85–89
- 40 Gardner, M.J. *et al.* (2005) Genome sequence of *Theileria parva*, a bovine pathogen that transforms lymphocytes. *Science* 309, 134–137
- 41 Cai, X. *et al.* (2003) Apicoplast genome of the coccidian *Eimeria tenella*. *Gene* 321, 39–46
- 42 Lister, D.L. *et al.* (2003) DNA transfer from chloroplast to nucleus is much rarer in *Chlamydomonas* than in tobacco. *Gene* 316, 33–38
- 43 Selosse, M. *et al.* (2001) Reducing the genome size of organelles favours gene transfer to the nucleus. *Trends Ecol. Evol.* 16, 135–141
- 44 Doolittle, W.F. (1998) You are what you eat: a gene transfer ratchet could account for bacterial genes in eukaryotic nuclear genomes. *Trends Genet.* 14, 307–311
- 45 Huang, C.Y. *et al.* (2003) Direct measurement of the transfer rate of chloroplast DNA into the nucleus. *Nature* 422, 72–76
- 46 Stegemann, S. *et al.* (2003) High-frequency gene transfer from the chloroplast genome to the nucleus. *Proc. Natl. Acad. Sci. U. S. A.* 100, 8828–8833
- 47 Bogorad, L. (1975) Evolution of organelles and eukaryotic genomes. *Science* 188, 891–898
- 48 von Heijne, G. (1986) Why mitochondria need a genome. *FEBS Lett.* 198, 1–4
- 49 Cheung, A.Y. *et al.* (1988) Relocating a gene for herbicide tolerance: a chloroplast gene is converted into a nuclear gene. *Proc. Natl. Acad. Sci. U. S. A.* 85, 391–395
- 50 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. U. S. A.* 91, 1969–1973
- 51 Daley, D.O. and Whelan, J. (2005) Why genes persist in organelle genomes. *Genome Biol.* 6, 110. DOI:10.1186/gb-2005-6-5-110 (<http://genomebiology.com/>)
- 52 Allen, J.F. (2003) The function of genomes in bioenergetic organelles. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 358, 19–37
- 53 Zerges, W. (2002) Does complexity constrain organelle evolution? *Trends Plant Sci.* 7, 175–182
- 54 de Grey, A.D.N.J. (2005) Forces maintaining organellar genomes: is any as strong as genetic code disparity or hydrophobicity? *BioEssays* 27, 436–446
- 55 Tsudzuki, T. *et al.* (2001) Comparative analysis of RNA editing sites in higher plant chloroplasts. *J. Mol. Evol.* 53, 327–332
- 56 Zauner, S. *et al.* (2004) Substitutional editing of transcripts from genes of cyanobacterial origin in the dinoflagellate *Ceratium horridum*. *FEBS Lett.* 577, 535–538
- 57 van der Giezen, M. and Tovar, J. (2005) Degenerate mitochondria. *EMBO Rep.* 6, 525–530
- 58 Iida, K. *et al.* (2002) Evaluation of two biosynthetic pathways to δ -aminolevulinic acid in *Euglena gracilis*. *Eur. J. Biochem.* 269, 291–297
- 59 Zhu, G. *et al.* (2000) *Cryptosporidium parvum* appears to lack a plastid genome. *Microbiology* 146, 315–321
- 60 Abrahamsen, M.S. *et al.* (2004) Complete genome sequence of the apicomplexan *Cryptosporidium parvum*. *Science* 304, 441–445
- 61 Lang-Unnasch, N. *et al.* (1998) Plastids are widespread and ancient in parasites of the phylum Apicomplexa. *Int. J. Parasitol.* 28, 1743–1754
- 62 Leake, J.R. (2004) Myco-heterotroph/epiparasitic plant interactions with ectomycorrhizal and arbuscular mycorrhizal fungi. *Curr. Opin. Plant Biol.* 7, 422–428
- 63 Hanaoka, M. *et al.* (2005) Glutamyl-tRNA mediates a switch in RNA polymerase use during chloroplast biogenesis. *EMBO Rep.* 6, 545–550

Free journals for developing countries

The WHO and six medical journal publishers have launched the Access to Research Initiative, which enables nearly 70 of the world's poorest countries to gain free access to biomedical literature through the Internet.

Gro Harlem Brundtland, director-general for the WHO, said that this initiative was 'perhaps the biggest step ever taken towards reducing the health information gap between rich and poor countries'.

See <http://www.healthinternetwork.net> for more information.