Chapter 29

Role of Horizontal Gene Transfer in the Evolution of Photosynthetic Eukaryotes and Their Plastids

Patrick J. Keeling

Abstract

Plastids are the organelles derived from a cyanobacterium through endosymbiosis. Unlike mitochondria, plastids are not found in all eukaryotes, but their evolution has an added layer of complexity since plastids have moved between eukaryotic lineages by secondary and tertiary endosymbiotic events. This complex history, together with the genetic integration between plastids and their host, has led to many opportunities for gene flow between phylogenetically distinct lineages. Some intracellular transfers do not lead to a protein functioning in a new environment, but many others do and the protein makeup of many plastids appears to have been influenced by exogenous sources as well. Here, different evolutionary sources and cellular destinations of gene flow that has affected the plastid lineage are reviewed. Most horizontal gene transfer (HGT) affecting the modern plastid has taken place via the host nucleus, in the form of genes for plastid-targeted proteins. The impact of this varies greatly from lineage to lineage, but in some cases such transfers can be as high as one fifth of analyzed genes. More rarely, genes have also been transferred to the plastid genome itself, and plastid genes have also been transferred to other non-plant, non-algal lineages. Overall, the proteome of many plastids has emerged as a mosaic of proteins from many sources, some from within the same cell (e.g., cytosolic genes or genes left over from the replacement of an earlier plastid), some from the plastid of other algal lineages, and some from completely unrelated sources.

Key words: Endosymbiotic gene transfer, endosymbiotic gene replacement, plastid-targeting, endosymbiosis.

1. Introduction

Mitochondria and plastids arose by the endosymbiotic uptake and retention of an alpha-proteobacterium and a cyanobacterium, respectively *(1)*. These endosymbionts were reduced in complexity and substantially integrated with their host, primarily though

Maria B. Gogarten et al. (eds.), *Horizontal Gene Transfer: Genomes in Flux, vol. 532*

DOI 10.1007/978-1-60327-853-9_29 Springerprotocols.com

the transfer of genes to the host genome and the targeting of proteins back to the endosymbiont. The processes of reduction and integration were seemingly fairly similar since the resulting two organelles now share a number of characteristics in common *(2)*, but the subsequent evolution of the two differed in several interesting ways. Mitochondria arose through a single endosymbiosis and the organelle or some derivative of it has been identified in virtually all eukaryotes (the few cases where it is not yet known for certain are most likely due to lack of evidence rather than absence), so it is now thought to have originated in the common ancestor of all known extant eukaryotes (**Fig. 29.1**), *(3)*. The evolution of plastids differed in two interesting ways. Like mitochondria, plastids also arose through a single endosymbiosis; there was some debate about the common origin of canonical plastids, and there are other cyanobacterial endosymbionts with very narrow taxonomic distribution, which may be sufficiently integrated to consider them plastids *(4–8)*, but for the purposes of this discussion these exceptions will not be considered further. However, unlike mitochondria, this endosymbiosis took place well after the origin and diversification of extant eukaryotes so that the lineage containing the endosymbiont consists only of three groups of algae, glaucophytes, red algae, green algae, and land plants *(9)*. Moreover, after the establishment of the plastid in these lineages, the organelle spread to other eukaryotes by further rounds of endosymbiosis. In "secondary" endosymbiosis, either a red or green alga was itself taken up and integrated into a new host and once again converted into an organelle (**Fig. 29.1**) *(10)*. Secondary endosymbiosis is known to have involved both green and red algae, so it must have taken place at least twice. Two lineages contain green secondary plastids, euglenids, and chlorarachniophytes, and both plastid and nuclear gene phylogenies support the conclusion that they acquired their plastids independently *(11)*. Red algal secondary plastids are known from cryptomonads, haptophytes, heterokonts, dinoflagellates, and apicomplexa. Despite this wider variety of groups and much debate (e.g. *(12–15)*), the current consensus is that these plastids originated from a single endosymbiosis event in the ancestor of these lineages, the so-called chromalveolates *(16)*. In addition to secondary endosymbiosis, additional events of endosymbiosis have taken place in dinoflagellate algae (**Fig. 29.1**). Here, the ancestral secondary red algal plastid has been lost or reduced several times and in a few lineages a new plastid has been acquired *(10, 17)*. When the new plastid is acquired from a secondary alga, they are referred to as tertiary plastids (these include plastids derived from cryptomonads, haptophytes and heterokonts); and when they are derived from a primary alga, they are called serial secondary plastids (this includes a single case where a green algal plastid has been taken up *(18)*). Of course, plastid gain is a long-term process, so

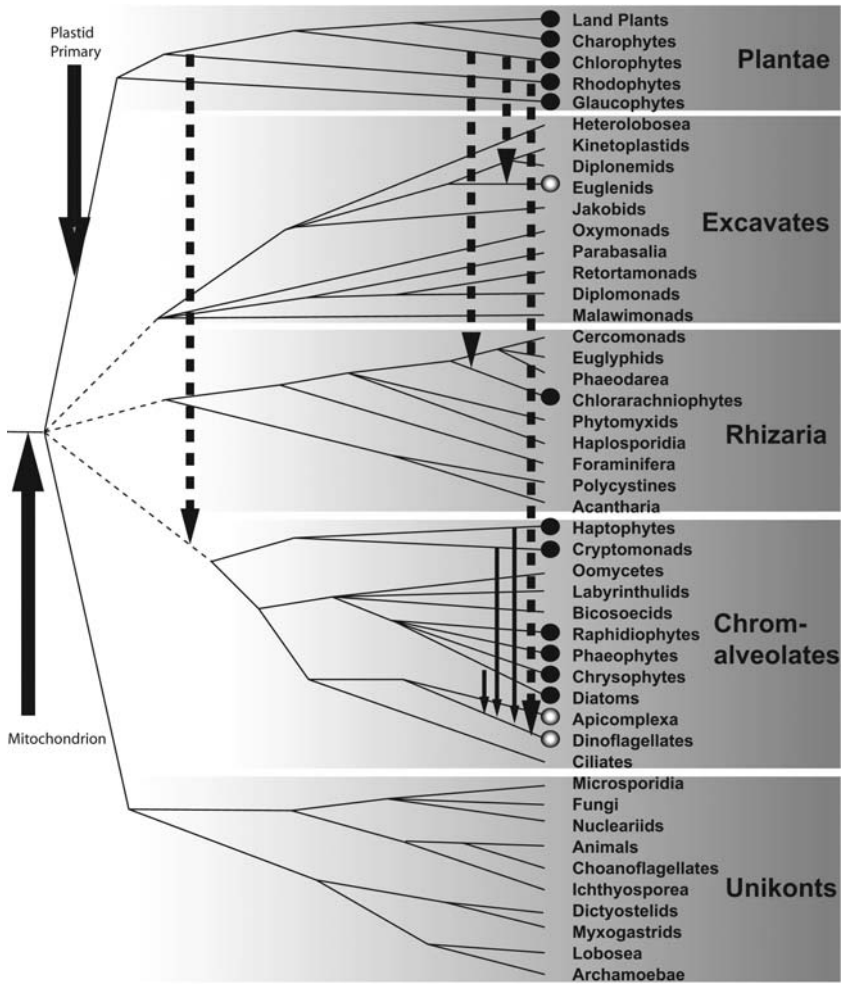

Fig. 29.1. Hypothetical tree of eukaryotes showing relationship of major lineages, including those discussed in text, and the many endosymbiotic events that explain modern plastid diversity. Many eukaryotic lineages (mostly protistan) are clustered into five major lineages, or supergroups, which are *shaded* and *named* to the *right*. Unknown relationships are indicated by polytomes, and supergroups of still contentious monophyly are indicated by the *dashed lines* at their base. Primary endosymbiosis (origin of mitochondria and primary plastids) is indicated by a *thick, solid arrow*. Secondary endosymbioses involving red or green algae are indicated by *slightly thinner, dashed arrows*. Tertiary endosymbioses in dinoflagellates are indicated by *thin, solid arrows*. Figure redrawn from *(64)*.

there are many intermediates in the spectrum between symbiont and organelle. There are also many lineages for which there is no evidence for a plastid, but which phylogenetically might be predicted to have once contained one. These may have lost the organelle or only lost photosynthesis, resulting in a cryptic plastid that has not yet been detected; in most cases the distinction cannot be made.

The evolution of photosynthetic lineages may therefore involve layers upon layers of endosymbiosis, making them genetically complex cells. The diatom plastid-containing dinoflagellate

Kryptoperidinium foliaceum is an excellent example *(19–21)*: it currently consists of five and perhaps six genome-containing compartments (two nuclei, two mitochondria, and one or perhaps two plastids), and historically is composed of no less than ten individual genetic entities (ignoring reticulations in the tree). Each of these events likely precipitated massive movements of DNA between the integrating partners as is known to have happened in the primary and secondary endosymbioses, and likely provided ample opportunity for genetic exchanges among the other component genomes as well. Overall, we now have a fairly clear picture of the evolutionary history of plastids *(10, 17)*, but the evolutionary histories of individual genes related to either the plastid or the endosymbiont are far more uncertain. Nevertheless, molecular data from a variety of algae are now abundant and this, together with our understanding of the history of plastid acquisition, combines to give us not only a set of specific expectations for the phylogenetic history of plastid derived genes but also a decent body of data to seek exceptions to those expectations.

2. Gene Transfer Related to the Establishment of Plastids

When a cyanobacterium and the ancestor of plants and primary algae integrated, a massive amount of gene transfer took place from the endosymbiont to the host nucleus *(1)*. Further, during secondary or tertiary endosymbiotic events, large-scale transfers from the endosymbiont algal nucleus to its new host nucleus also took place. There are now many such genes known from secondary endosymbioses, and several studies of how the proteins are targeted back to the organelle *(22–27)*. In the case of tertiary plastids there is comparatively little data, and the situation is more complex because all known tertiary plastids are in dinoflagellates, and therefore both the host and the endosymbiont had a plastid or plastid-containing ancestry (*see* **Section 3.2**. for more discussion on this). While these transfers are well characterized and certainly are important, they need to be distinguished from HGT in general because the circumstances surrounding these transfers are very different, as are their implications. Here, we will use the terms Endosymbiotic Gene Transfer (EGT) and Endosymbiotic Gene Replacement (EGR) to distinguish these special subsets of transfers, as described below.

2.1. Endosymbiotic Gene Transfer (EGT): Transfer of Genes Whose Products Are Targeted Back to the Organelle

The plastid genome encodes only a small fraction of the genes needed for plastid function *(2)*. Most plastid proteins are encoded by host nuclear genes, and most of these were transferred from the cyanobacterium to the host nucleus in the ancestor of primary plastids (**Fig. 29.2-A**). The protein products of these genes

are translated in the host cytoplasm and post-translationally targeted back to the organelle using a specific import pathway (**Fig. 29.2-1**), *(28)*. In secondary plastids, these genes were transferred yet again, in this case from the nucleus of the primary alga to the nucleus of the secondary one, and in tertiary plastids the same set of events occurred once more. Regardless of the number of times these genes are transferred, however, the environment in which their protein product functions remains unchanged, so while these transfers might affect the evolution of the gene (e.g. by moving to a compartment with different mutation rates), they do not so much generate novel functions or combinations of functions, but rather are a new way to organize information. They should therefore be distinguished from HGTs that yield novel combinations of proteins. These Endosymbiotic Gene Transfers (EGT) are undisputedly an important process, but will not be discussed further here.

2.2. Endosymbiotic Gene Replacement (EGR): Transfer of Genes Whose Products Are Not Targeted Back to the Organelle

In addition to the transfer of hundreds of genes for proteins destined to be targeted back to the plastid, it is now hypothesized that many other cyanobacterial genes made their way into the host genome whose products do not function in the plastid today (**Fig. 29.2-2**) *(29)*. This process, referred to as Endosymbiotic Gene Replacement (EGR), in one way could be seen as a subclass of HGT because these genes have the potential to introduce new functions and combinations of functions. However, in another way they are like EGT because they come en masse from a single source, as opposed to a slower trickle of new genes from a wide variety of sources. In some cases transferred genes can be subjected to both EGT and EGR, since differentially expressed proteins can function in either compartment (aminoacyl-tRNA synthetases provide a great example as an entire class of proteins where dual targeting is common: *(30)*). However, the full potential impact of EGR had not been obvious until recent analyses of plants concluded that the contribution of the cyanobacterium to the plant lineage was substantial *(31)*. Several examples of EGR between two eukaryotes also seem to have occurred in the secondary endosymbiosis *(32)*. Interestingly, however, a large-scale analysis of another primary plastid-containing lineage, the glaucophytes, did not reach the same conclusion *(33)*. Here, many genes of cyanobacterial origin were found, but the protein products of most were concluded to be plastid-targeted. Whether the original estimates of EGR were too high, or alternatively why two systems seem to differ when they are products of the same endosymbiosis remains to be seen, and a detailed analysis of many of the cyanobacterial genes thought to encode cytosolic proteins also remains to be performed.

2.3. The Contribution of Chlamydial Genes to the Origin of Plastids

While the plastid may be derived from a cyanobacterium, recent large-scale analyses of plastid-targeted proteins have suggested that there is also an unexpected contribution from Chlamydiales. Chlamydiales are a group of pathogenic bacteria that have no known relationship to photosynthesis or plastids. Nevertheless, an unusual number of plastid-targeted proteins have been shown to share a closer relationship to chamydial homologs than they do to cyanobacterial homologs *(34–36)*. This was first described in plants, and was originally interpreted as either HGT or as a possible link between the plant host lineage and Chlamydiales *(34)*.

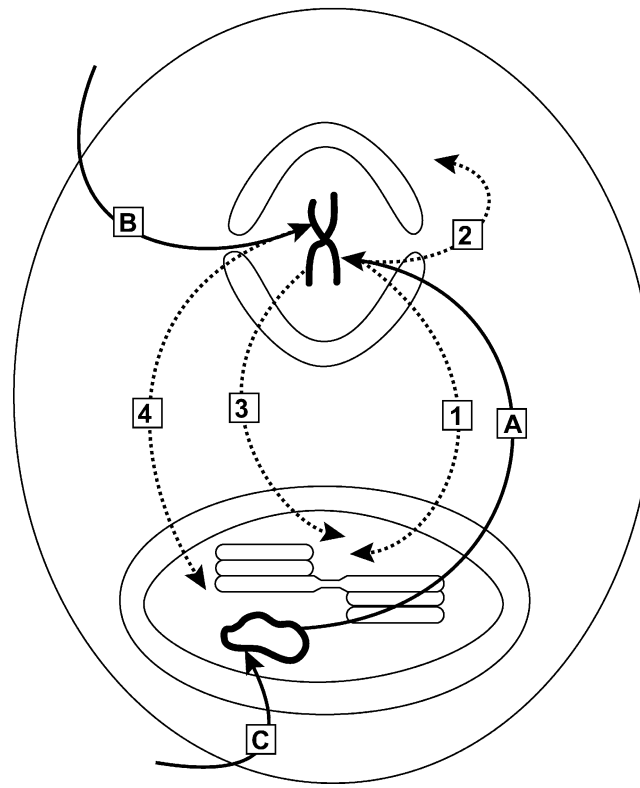

Fig. 29.2. Schematic of a primary algal cell showing some of the main types of HGT and re-targeting events known to have affected the protein makeup of the plastid and host. Gene transfers are indicated by *solid lines* labeled with letters, whereas protein-targeting is indicated by *dashed lines* labeled with numbers. (**A**) is plastid to nucleus EGT, which is widespread and common. Many protein products of these genes are targeted back to the plastid (**1**), and so the environment of the protein is unchanged. Others (**2**) may take up a function in the cytoplasm, and there is some indication that this might be common in some lineages as well. In contrast to this, the protein products of a few ancestrally nuclear genes were found to be targeted to the plastid (**3**), suggesting this process can work in both directions. (**B**) is gene transfer from an exogenous source to the nucleus. This has been well documented for numerous proteins in several algal lineages. The majority of such transfers are expected to result in cytosolic proteins (not shown), but many of the cases that are best described are proteins that are subsequently targeted to the plastid (**4**). Lastly, (**C**) is gene transfer from an exogenous source directly to the plastid genome. Such events have been documented, but are relatively rare. Not shown are the mitochondria, which could participate in the same kinds of events, or the added complexity of secondary or tertiary endosymbiosis, where genes can move from one eukaryote or its plastid to the other.

However, some of the same proteins are also found in a range of algae, and in at least some cases the genes encode plastid-targeted proteins. This was suggested to reflect a close relationship between Chlamydiales and cyanobacteria *(37)* rather than HGT;, however, this cannot explain many of the genes, where plant and algal homologs are robustly related to Chlamydiales to the exclusion of cyanobacteria *(38, 39)*. Most recently, these genes have been suggested to be relicts of a chlamydial presence in the host lineage that undertook the original primary endosymbiosis. It has been suggested that the presence of a chlamydial parasite in this host might have actually facilitated the transition of the cyanobacterium from a symbiont to an organelle, perhaps by providing essential transporters that were unlikely to have been present in a free-living cyanobacterium *(35, 39)*.

3. HGT of Genes for Plastid-Targeted Proteins

The establishment of the plastid-targeting system and the apparent ease of acquiring targeting peptides (inferred from the facts that transit peptide sequences are not highly conserved, and more importantly that they have been acquired in their thousands, repeatedly by several lineages) meant that HGT to either the plastid genome itself or to the nuclear genome could affect the plastid proteome. Here some cases will be reviewed where HGT to the nucleus or plastid, or EGT between nuclei have impacted the plastid proteome.

3.1. HGT of Plastid Targeted Proteins in Several Algal Systems

Plastid-targeted proteins make an attractive class of proteins for detecting HGT in algae because they are relatively well sampled, and there is also a clear phylogenetic expectation that they are related to cyanobacteria-derived homologs present in green or red algal nuclear genomes. In contrast, host nuclear genes from many algal lineages are poorly sampled across eukaryotic diversity, and the closest relative of the host is often either not well known or even more poorly sampled *(10, 17)*. It is probably for these reasons, and not some intrinsic tendency of genes for plastid-targeted proteins to be transferred, that many of the first cases of HGT in algae came from genes for plastid-targeted proteins (**Fig. 29.2B**). Two groups where this is especially well examined are the chlorarachniophytes and the dinoflagellates. Chlorarachniophytes have a green algal plastid, and an analysis of expressed sequence tags (ESTs) from the model species *Bigelowiella natans* revealed that almost 20% of the genes for recognized plastid-targeted proteins were derived by HGT *(22)*. The majority of these were derived from other algae, several from red algae or lineages with red algal plastids, but others were derived from

streptophyte green algae (some might be derived from chlorophyte green algae too, but these cannot be detected since the *B. natans* plastid is itself derived from a chlorophyte). A small fraction of these genes (glyceraldehyde-3-phosphate dehydrogenase and ribulose-5-phosphate-3-epimerase) were not related to the plastid lineage at all, but rather to other bacterial lineages. Interestingly, a closely related, plastid-targeted ribulose-5-phosphate-3-epimerase has recently been found in distantly related algae with red secondary plastids *(40)*, indicating a more complex history of this transfer. The chlorarachniophytes are myxotrophic, and so it was suggested they acquired genes from their food, and the tendency for the proteins to originate from algae was most likely due to the fact that many plastid proteins only exist in phototrophs *(22)*. Similar surveys in dinoflagellates also revealed slightly fewer cases of HGT, again mostly genes derived from other algae *(23, 24, 32, 41)*. The dinoflagellate plastid is derived from a red alga, so in this case the genes most clearly originating from HGT came from green algae, with a few genes coming from other sources.

Rather than analyzing all the plastid-targeted proteins from a random sample (like ESTs), other analyses have focused on the evolutionary history of all proteins in a certain pathway. For example, the analysis of the heme biosynthetic pathway in red algae and red algal–derived plastids shows some genes have the expected phylogenetic origin, whereas others seem to be derived from the host lineages, others still are proteobacterial and inferred to be derived from the mitochondrion *(42)*. Indeed, in apicomplexan parasites this process seems to have gone one step further, as the reactions themselves are distributed among various cellular compartments *(43)*. Similarly, the shikimate pathway for aromatic amino acid biosynthesis is composed of plastid-targeted proteins that are derived not only from cyanobacteria, but also at least two other bacterial lineages, so that the majority of the pathway appears to be proteins derived from HGT *(44)*.

Many other individual cases of plastid-targeted proteins being involved in HGT have been described, but only a few will be reviewed here. In the dinoflagellates one interesting example involves a gene fusion where the context of the source gene can be inferred. Here, the dinoflagellates and their sister lineages *Oxyrrhis marina* acquired an AroB gene that is closely related to an AroB paralog found only in a subgroup of cyanobacteria and not normally in plastids *(45)*. Immediately downstream of this gene in the cyanobacterial genomes is an *O*-methyl transferase of unknown function, and in dinoflagellates this gene has been fused to the 3′ end of AroB. The *O. marina* fusion protein is cytosolic, but in several other dinoflagellates the protein encodes a plastid-targeting leader. Interestingly, the AroB and *O*-methyl transferase moieties have been split up again in one species, and here the

O-methyl transferase has acquired its own plastid-targeting leader, suggesting it acquired some function in the plastid while attached to AroB.

Another interesting example is GAPDH, where the plastid-targeted protein in most dinoflagellates is only distantly related to that of cyanobacteria and other plastids, and instead appears to be derived from a duplication of the cytosolic GAPDH *(46, 47)*. In the dinoflagellate genera *Pyrocystis* and *Akashiwo*, however, a second, apparently plastid-targeted GAPDH has been found. This protein is of the cyanobacterial type, but they are specifically related to plastid-targeted homologues from euglenid algae, which have green algal plastids as opposed to the red algal plastids of dinoflagellates *(48, 49)*. This is even more intriguing since another GAPDH has also been found in another subgroup of dinoflagellates that is related to a different class of euglenid GAPDH again, suggesting a complex history of transfers between these two distantly related algal groups *(48)*.

An apparently more ancient transfer has been described for the plastid-targeted fructose-bisphosphate aldolase of chromalveolates, perhaps originally incorporated into the cell as a cytosolic enzyme, and becoming plastid-targeted after a gene duplication event *(50)*. This case is noteworthy because, along with a similar event involving GAPDH *(47, 51)*, it lends support to the single origin of the plastid in these lineages.

Overall, the impact of HGT on plastid proteins in algae is significant in some groups, but uneven. On one hand, some genes appear to be more prone to transfer than others, for example, GAPDH, where plastid-targeted proteins have arisen by HGT in at least some dinoflagellates and the chlorarachniophytes *(10, 48, 49)*, or glucose-6-phosphate isomerase *(52)* where a variety of transfer events appear possible. One the other hand, many of the cases described here affect the same or overlapping groups of algae (in particular chlorarachniophytes and dinoflagellates), whereas other groups show little or no evidence of HGT affecting plastid proteins. The same set of proteins analyzed in the chlorarachniophyte *B. natans* was also analyzed in the green alga *Chlamydomonas reinhardtii* and, in contrast to *B. natans*, none showed any evidence for transfer in *C. reinhardtii (10)*. In addition, most of these cases are apparently relatively recent and only affect one subgroup or perhaps a few species of the algae in question (one exception being plastid-targeted fructose-bisphosphate-aldolase).

3.2. Plastid-Plastid EGT in Algae with Redundant Plastids

As described in **Section 2**, the origin of tertiary plastids involves the movement of a great many genes from the nucleus of a secondary alga to the nucleus of its new host. This process is not unlike the origin of secondary plastids, but one very major difference is that all known tertiary plastids are found in dinoflagellates,

which means both the host and the incoming secondary alga have (or had) a plastid of distinct evolutionary origin. In most cases no data for plastid-targeted proteins are available from these complex systems, but EST surveys have been carried out in two related genera with haptophyte-derived plastids, *Karlodinium micrum* and *Karenia brevis (25, 53)*. Since these organisms only contain the haptophyte-derived plastid, the null expectation would be that all or nearly all their plastid-targeted proteins should also be derived from the haptophyte, but their presence in the dinoflagellate nucleus raises the intriguing possibility that the original dinoflagellate plastid may have contributed proteins too. Indeed, the *K. micrum* survey revealed that about one-third of all phylogenetically analyzable plastid-targeted proteins were leftovers from the dinoflagellate plastid, and were re-targeted to the new haptophyte plastid, suggesting the proteome of this organelle is substantially chimeric *(25)*. Analysis of *K. brevis* ESTs also revealed this, and showed that proteins from other sources were also present, as with other dinoflagellates *(53)*. Interestingly, the targeting peptides in these organisms are different from those of either dinoflagellates or haptophytes *(25)*, and it is possible that this played a role in the mixing of these proteomes. As the targeting system shifted emphasis from the characteristics common to the ancestral haptophyte plastid, a dinoflagellate-derived protein might have been as likely to co-evolve with the import system as were the haptophyte-derived proteins. It is therefore plausible that being well adapted to the environment (which would presumably favor haptophyte proteins) was not the only or even the main determinant of which protein was retained in the new plastid *(25)*. Whatever the case may be, these are special cases of EGR where new combinations of plastid proteins were the outcome.

3.3. HGT of Plastid Proteins to Prokaryotes

While the subject of this review is the impact of HGT on plastids and their hosts, it is worthwhile to discuss whether plastids have also been the donor in HGT events. Of course, many of the events described in the preceding sections involve plastid-derived genes moving between plastid lineages (e.g., plastid-targeted proteins moving between algal groups), and plastid-derived proteins taking up a function in the cytosol would have the same overall effect as an HGT event, but what about plastid genes moving to lineages and new environments? Direct transfer of plastid DNA between plants has been documented, but it has not been shown that the genes moved to the plastid genome of the recipient, or that they are expressed *(54)*. It has been demonstrated that plastid DNA could be incorporated into the genome of the naturally competent soil bacterium *Acinetobacter* under experimental conditions *(55)*, but whether such transfers have occurred in nature is uncertain. The presence of *Chlamydia*-like genes in plants and algae, which led to the idea that a *Chlamydia*-like symbiont participated

in the origin of plastids (**Section 2.3**) could also, in theory, be due to the transfer of plastid-derived genes to *Chlamydia*. This has been suggested in the case of *fabI (38)*, but in no case is the relationship of these genes to cyanobacteria so clear as to distinguish this possibility from a *Chlamydia*-to-plastid transfer *(39)*. One case involving an apparently functional gene transfer that is interesting for several reasons is the transfer of plastid-targeted FBA from red algae to the cyanobacteria *Prochlorococcus* and *Synechococcus (56)*. This is partly interesting because the recipient of the plastid gene is a cyanobacterium, which is ironic, but also because the gene has inserted beside the ancestral cyanobacterial FBA in the genome and its distribution within the genera where it is found is highly irregular, altogether suggesting a very complex history. This case is further complicated by the fact that the plastid gene in question is not actually ancestrally cyanobacterial either. Early in the evolution of the primary plastid lineage the host, cytosolic FBA duplicated and replaced the plastid-targeted analog (they are different and non-homologous classes of FBA) in the common ancestor of red and green algae *(57)*. Accordingly, the gene transferred to the cyanobacterium was not originally for a plastid-targeted protein.

4. HGT to Plastid Genomes

4.1. Direct Transfer of Genes Between Plastid Genomes

So far only nucleus-encoded genes, presumed to be cyanobacterial and presumed to be plastid targeted, have been discussed in any detail, but the plastid also has its own genome. However, HGT involving the plastid genome directly appears to be comparatively rare (**Fig. 29.2C**), in contrast to what has been found in plant mitochondria *(58)*. Many plastid genomes apparently encode unique proteins with no detectable similarity to any other protein, but the majority of these are likely not derived from HGT. One case that does appear to be transferred involves a mobile element found in the plastid-encoded *psb*A of *Euglena myxocylindracea* and a psychrophilic *Chlamydomonas (59, 60)*. In both species, phylogenetic analysis of a mobile group II intron showed it to be comparatively distant to group II introns in other plastid genomes, and instead closely related to group II introns from cyanobacteria *(59)*. The discovery of this intron in the primary green plastid of a *Chlamydomonas* and the secondary green plastid of a *Euglena* may be taken to suggest its origin in plastids is ancient, but the intron is absent in close relatives of both, and indeed in many other plastid lineages, and homing introns should be relatively likely to move. Two other cases of transfer into the plastid genome stand out as well, *rpl36* and *dnaX (61, 62)*. There

are two distinct and distantly related paralogs of *rpl36* in bacteria, and plastid genomes typically contain only one of these, embedded within a conserved operon of ribosomal proteins. In the cryptomonads and haptophytes, however, the same operon contains an *rpl36* at the same position, but unusually it is a member of the other paralogous family *(61)*. This is significant for a number of reasons, first it is interesting that the incoming paralogue was inserted at the same position as the ancestral *rpl36*, and second it is a strong evidence for a relationship between the plastids of haptophytes and cryptomonads, which has since been supported by nuclear gene phylogenies *(13,14)*. DnaX is a component of DNA polymerase responsible for repair. Neither DnaX nor any other DNA polymerase subunit has been found in a plastid genome previously, with the exception of several closely related species of the cryptomonad *Rhodomonas (62)*. Significantly, other cryptomonad plastid genomes lack the enzyme and it is not phylogenetically related to cyanobacteria, indicating it originated recently by HGT from some unidentified bacterial lineage.

5. Conclusion/ Outlook

The plastid proteome is much more complex than simply a residual collection of cyanobacterial proteins, and its relationship to the host is also more complicated than originally envisioned. The close connection between the two lineages has provided them with ample opportunity for proteins to move into a new environment, and this has happened many times. Similarly, genes for plastid-targeted proteins have moved between algal lineages many times, some genes showing a propensity to do this. The plastid genome seems resistant to HGT, so most of the changes to its proteome have occurred via the host nucleus, but even plastid genomes have been invaded by new genes on a few occasions. Overall, the plastid proteome is doubtlessly a variable mosaic of proteins from several sources, although the degree to which this is the case is still unclear. Genome-wide surveys in recent years have radically changed our view of the plastid proteome, so perhaps the biggest change we can expect in the years to come will emerge from analyzing plastid proteomics with an eye for invasion of foreign genes *(63)*. Such data will reveal which proteins are actually located in the organelle, something we infer with varying degrees of confidence now but do not really know in most cases. Without a more reliable catalog of plastid proteins, it is difficult to conclude what actually was the contribution of the host to the plastid proteome and to be certain of the plastid location of non-cyanobacterial genes derived from HGT.

Acknowledgments

Work in the Keeling lab on HGT is supported by a grant from the Natural Sciences and Engineering Research Council of Canada and a grant to the Centre for Microbial Diversity and Evolution from the Tula Foundation.

References

1. Gray, M. W., Burger, G., Lang, B. F. (1999) Mitochondrial evolution. *Science* **283**, 1476–81.
2. Gray, M. W. (1999) Evolution of organellar genomes. *Curr Opin Genet Dev* **9**, 678–87.
3. Williams, B. A. P., Keeling, P. J. (2003) Cryptic organelles in parasitic protists and fungi. *Adv Parasitol* **54**, 9–67.
4. Nowack, E. C., Melkonian, M., Glockner, G. (2008) Chromatophore genome sequence of *Paulinella* sheds light on acquisition of photosynthesis by eukaryotes. *Curr Biol* **18**, 410–8.
5. Keeling, P. J., Archibald, J. M. (2008) Organelle evolution: What's in a name? *Curr Biol* **18**, R345–7.
6. Palmer, J. D. (2003) The symbiotic birth and spread of plastids: How many times and whodunit? *J Phycol* **39**, 1–9.
7. Larkum, A. W., Lockhart, P. J., Howe, C. J. (2007) Shopping for plastids. *Trends Plant Sci* **12**, 189–95.
8. Keeling, P. J. (2004) The diversity and evolutionary history of plastids and their hosts. *Am J Bot* **91**, 1481–93.
9. Rodriguez-Ezpeleta, N., Brinkmann, H., Burey, S. C., Roure, B., Burger, G., Loffelhardt, W., Bohnert, H. J., Philippe, H., Lang, B. F. (2005) Monophyly of primary photosynthetic eukaryotes: Green plants, red algae, and glaucophytes. *Curr Biol* **15**, 1325–30.
10. Archibald, J. M., Keeling, P. J. (2002) Recycled plastids: A green movement in eukaryotic evolution. *Trends Genet* **18**, 577–84.
11. Rogers, M. B., Gilson, P. R., Su, V., Mcfadden, G. I., Keeling, P. J. (2007) The complete chloroplast genome of the chlorarachniophyte *Bigelowiella natans*: Evidence for independent origins of chlorarachniophyte and euglenid secondary endosymbionts. *Mol Biol Evol* **24**, 54–62.
12. Burki, F., Shalchian-Tabrizi, K., Minge, M., Skjaeveland, A., Nikolaev, S. I., Jakobsen, K. S., Pawlowski, J. (2007) Phylogenomics reshuffles the eukaryotic supergroups. *PLoS ONE* **2**, e790.
13. Hackett, J. D., Yoon, H. S., Li, S., Reyes-Prieto, A., Rummele, S. E., Bhattacharya, D. (2007) Phylogenomic analysis supports the monophyly of cryptophytes and haptophytes and the association of rhizaria with chromalveolates. *Mol Biol Evol* **24**, 1702–13.
14. Patron, N. J., Inagaki, Y., Keeling, P. J. (2007) Multiple gene phylogenies support the monophyly of cryptomonad and haptophyte host lineages. *Curr Biol* **17**, 887–91.
15. Leigh, J. W., Susko, E., Baumgartner, M., Roger, A. J. (2008) Testing congruence in phylogenomic analysis. *Syst Biol* **57**, 104–15.
16. Cavalier-Smith, T. (2004) Chromalveolate diversity and cell megaevolution: Interplay of membranes, genomes and cytoskeleton, in *Organelles, genomes and eukaryotic evolution* (Hirt, R. P., Horner D., ed.) Taylor and Francis, London 71–103.
17. Delwiche, C. F. (1999) Tracing the thread of plastid diversity through the tapestry of life. *Am Nat* **154, Supplement**, S164–77.
18. Watanabe, M. M., Suda, S., Inouye, I., Sawaguchi, I., Chihara, M. (1990) *Lepidodinium viride* gen. et sp. nov. (Gymnodiniales, Dinophyta), a green dinoflagellate with a chlorophyll *a*- and *b*-containing endosymbiont. *J Phycol* **26**, 741–51.
19. Chesnick, J. M., Hooistra, W. H., Wellbrock, U., Medlin, L. K. (1997) Ribosomal RNA analysis indicates a benthic pennate diatom ancestry for the endosymbionts of the dinoflagellates *Peridinium foliaceum* and *Peridinium balticum* (Pyrrhophyta). *J Eukaryot Microbiol* **44**, 314–20.
20. Inagaki, Y., Dacks, J. B., Doolittle, W. F., Watanabe, K. I., Ohama, T. (2000) Evolutionary relationship between dinoflagellates bearing obligate diatom endosymbionts: Insight into tertiary endosymbiosis. *Int J Syst Evol Microbiol* **50 Pt 6**, 2075–81.
21. Imanian, B., Keeling, P. J. (2007) The dinoflagellates *Durinskia baltica* and *Kryptoperidinium foliaceum* retain functionally overlapping mitochondria from two evolu-

tionarily distinct lineages. *BMC Evol Biol* **7**, 172.
22. Archibald, J. M., Rogers, M. B., Toop, M., Ishida, K., Keeling, P. J. (2003) Lateral gene transfer and the evolution of plastid-targeted proteins in the secondary plastid-containing alga *Bigelowiella natans*. *Proc Natl Acad Sci USA* **100**, 7678–83.
23. Bachvaroff, T. R., Concepcion, G. T., Rogers, C. R., Herman, E. M., Delwiche, C. F. (2004) Dinoflagellate expressed sequence tags data indicate massive transfer of chloroplast genes to the nuclear genome. *Protist* **155**, 65–78.
24. Hackett, J. D., Yoon, H. S., Soares, M. B., Bonaldo, M. F., Casavant, T. L., Scheetz, T. E., Nosenko, T., Bhattacharya, D. (2004) Migration of the plastid genome to the nucleus in a peridinin dinoflagellate. *Curr Biol* **14**, 213–8.
25. Patron, N. J., Waller, R. F., Keeling, P. J. (2006) A tertiary plastid uses genes from two endosymbionts. *J Mol Biol* **357**, 1373–82.
26. Patron, N. J., Waller, R. F., Archibald, J. M., Keeling, P. J. (2005) Complex protein targeting to dinoflagellate plastids. *J Mol Biol* **348**, 1015–24.
27. Ralph, S. A., Van Dooren, G. G., Waller, R. F., Crawford, M. J., Fraunholz, M. J., Foth, B. J., Tonkin, C. J., Roos, D. S., Mcfadden, G. I. (2004) Tropical infectious diseases: Metabolic maps and functions of the *Plasmodium falciparum* apicoplast. *Nat Rev Microbiol* **2**, 203–16.
28. Mcfadden, G. I. (1999) Plastids and protein targeting. *J Eukaryot Microbiol* **46**, 339–46.
29. Martin, W., Stoebe, B., Goremykin, V., Hansmann, S., Hasegawa, M., Kowallik, K. V. (1998) Gene transfer to the nucleus and the evolution of chloroplasts. *Nature* **393**, 162–5.
30. Duchene, A. M., Giritch, A., Hoffmann, B., Cognat, V., Lancelin, D., Peeters, N. M., Zaepfel, M., Marechal-Drouard, L., Small, I. D. (2005) Dual targeting is the rule for organellar aminoacyl-trna synthetases in *Arabidopsis thaliana*. *Proc Natl Acad Sci USA* **102**, 16484–9.
31. Martin, W., Rujan, T., Richly, E., Hansen, A., Cornelsen, S., Lins, T., Leister, D., Stoebe, B., Hasegawa, M., Penny, D. (2002) Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. *Proc Natl Acad Sci USA* **99**, 12246–51.
32. Li, S., Nosenko, T., Hackett, J. D., Bhattacharya, D. (2006) Phylogenomic analysis identifies red algal genes of endosymbiotic origin in the chromalveolates. *Mol Biol Evol* **23**, 663–74.
33. Reyes-Prieto, A., Hackett, J. D., Soares, M. B., Bonaldo, M. F., Bhattacharya, D. (2006) Cyanobacterial contribution to algal nuclear genomes is primarily limited to plastid functions. *Curr Biol* **16**, 2320–5.
34. Stephens, R. S., Kalman, S., Lammel, C., Fan, J., Marathe, R., Aravind, L., Mitchell, W., Olinger, L., Tatusov, R. L., Zhao, Q., Koonin, E. V., Davis, R. W. (1998) Genome sequence of an obligate intracellular pathogen of humans: *Chlamydia trachomatis*. *Science* **282**, 754–9.
35. Tyra, H. M., Linka, M., Weber, A. P., Bhattacharya, D. (2007) Host origin of plastid solute transporters in the first photosynthetic eukaryotes. *Genome Biol* **8**, R212.
36. Royo, J., Gimez, E., Hueros, G. (2000) CMP-KDO synthetase: A plant gene borrowed from gram-negative eubacteria. *Trends Genet* **16**, 432–3.
37. Brinkman, F. S., Blanchard, J. L., Cherkasov, A., Av-Gay, Y., Brunham, R. C., Fernandez, R. C., Finlay, B. B., Otto, S. P., Ouellette, B. F., Keeling, P. J., Rose, A. M., Hancock, R. E., Jones, S. J., Greberg, H. (2002) Evidence that plant-like genes in *Chlamydia* species reflect an ancestral relationship between Chlamydiaceae, cyanobacteria, and the chloroplast. *Genome Res* **12**, 1159–67.
38. Ryall, K., Harper, J. T., Keeling, P. J. (2003) Plastid-derived type II fatty acid biosynthetic enzymes in chromists. *Gene* **313**, 139–48.
39. Huang, J., Gogarten, J. P. (2007) Did an ancient chlamydial endosymbiosis facilitate the establishment of primary plastids? *Genome Biol* **8**, R99.
40. Rogers, M. B., Watkins, R. F., Harper, J. T., Durnford, D. G., Gray, M. W., Keeling, P. J. (2007) A complex and punctate distribution of three eukaryotic genes derived by lateral gene transfer. *BMC Evol Biol* **7**, 89.
41. Waller, R. F., Patron, N. J., Keeling, P. J. (2006) Phylogenetic history of plastid-targeted proteins in the peridinin-containing dinoflagellate *Heterocapsa triquetra*. *Int J Syst Evol Microbiol* **56**, 1439–47.
42. Obornik, M., Green, B. R. (2005) Mosaic origin of the heme biosynthesis pathway in photosynthetic eukaryotes. *Mol Biol Evol* **22**, 2343–53.
43. Sato, S., Clough, B., Coates, L., Wilson, R. J. (2004) Enzymes for heme biosynthesis are found in both the mitochondrion and plastid of the malaria parasite *Plasmodium falciparum*. *Protist* **155**, 117–25.
44. Richards, T. A., Dacks, J. B., Campbell, S. A., Blanchard, J. L., Foster, P. G., Mcleod,

R., Roberts, C. W. (2006) Evolutionary origins of the eukaryotic shikimate pathway: Gene fusions, horizontal gene transfer, and endosymbiotic replacements. *Eukaryot Cell* **5**, 1517–31.
45. Waller, R. F., Slamovits, C. H., Keeling, P. J. (2006) Lateral gene transfer of a multigene region from cyanobacteria to dinoflagellates resulting in a novel plastid-targeted fusion protein. *Mol Biol Evol* **23**, 1437–43.
46. Fagan, T., Woodland Hastings, J., Morse, D. (1998) The phylogeny of glyceraldehyde-3-phosphate dehydrogenase indicates lateral gene transfer from cryptomonads to dinoflagellates. *J Mol Evol* **47**, 633–9.
47. Fast, N. M., Kissinger, J. C., Roos, D. S., Keeling, P. J. (2001) Nuclear-encoded, plastid-targeted genes suggest a single common origin for apicomplexan and dinoflagellate plastids. *Mol Biol Evol* **18**, 418–26.
48. Takishita, K., Ishida, K., Maruyama, T. (2003) An enigmatic GAPDH gene in the symbiotic dinoflagellate genus *Symbiodinium* and its related species (the order Suessiales): Possible lateral gene transfer between two eukaryotic algae, dinoflagellate and euglenophyte. *Protist* **154**, 443–54.
49. Fagan, T. M., Hastings, J. W. (2002) Phylogenetic analysis indicates multiple origins of chloroplast glyceraldehyde-3-phosphate dehydrogenase genes in dinoflagellates. *Mol Biol Evol* **19**, 1203–7.
50. Patron, N. J., Rogers, M. B., Keeling, P. J. (2004) Gene replacement of fructose-1,6-bisphosphate aldolase (FBA) supports a single photosynthetic ancestor of chromalveolates. *Eukaryot Cell* **3**, 1169–75.
51. Harper, J. T., Keeling, P. J. (2003) Nucleus-encoded, plastid-targeted glyceraldehyde-3-phosphate dehydrogenase (GAPDH) indicates a single origin for chromalveolate plastids. *Mol Biol Evol* **20**, 1730–5.
52. Grauvogel, C., Brinkmann, H., Petersen, J. (2007) Evolution of the glucose-6-phosphate isomerase: The plasticity of primary metabolism in photosynthetic eukaryotes. *Mol Biol Evol* **24**, 1611–21.
53. Nosenko, T., Bhattacharya, D. (2007) Horizontal gene transfer in chromalveolates. *BMC Evol Biol* **7**, 173.
54. Park, J. M., Manen, J. F., Schneeweiss, G. M. (2007) Horizontal gene transfer of a plastid gene in the non-photosynthetic flowering plants *Orobanche* and *Phelipanche* (Orobanchaceae). *Mol Phylogenet Evol* **43**, 974–85.
55. De Vries, J., Herzfeld, T., Wackernagel, W. (2004) Transfer of plastid DNA from tobacco to the soil bacterium *Acinetobacter* sp. by natural transformation. *Mol Microbiol* **53**, 323–34.
56. Rogers, M. B., Patron, N. J., Keeling, P. J. (2007) Horizontal transfer of a eukaryotic plastid-targeted protein gene to cyanobacteria. *BMC Biol* **5**, 26.
57. Gross, W., Lenze, D., Nowitzki, U., Weiske, J., Schnarrenberger, C. (1999) Characterization, cloning, and evolutionary history of the chloroplast and cytosolic class I aldolases of the red alga *Galdieria sulphuraria*. *Gene* **230**, 7–14.
58. Bergthorsson, U., Adams, K. L., Thomason, B., Palmer, J. D. (2003) Widespread horizontal transfer of mitochondrial genes in flowering plants. *Nature* **424**, 197–201.
59. Sheveleva, E. V., Hallick, R. B. (2004) Recent horizontal intron transfer to a chloroplast genome. *Nucleic Acids Res* **32**, 803–10.
60. Odom, O. W., Shenkenberg, D. L., Garcia, J. A., Herrin, D. L. (2004) A horizontally acquired group II intron in the chloroplast psbA gene of a psychrophilic *Chlamydomonas*: In vitro self-splicing and genetic evidence for maturase activity. *RNA* **10**, 1097–107.
61. Rice, D. W., Palmer, J. D. (2006) An exceptional horizontal gene transfer in plastids: Gene replacement by a distant bacterial paralog and evidence that haptophyte and cryptophyte plastids are sisters. *BMC Biol* **4**, 31.
62. Khan, H., Parks, N., Kozera, C., Curtis, B. A., Parsons, B. J., Bowman, S., Archibald, J. M. (2007) Plastid genome sequence of the cryptophyte alga *Rhodomonas salina* CCMP1319: Lateral transfer of putative DNA replication machinery and a test of chromist plastid phylogeny. *Mol Biol Evol* **24**, 1832–42.
63. Van Wijk, K. J. (2004) Plastid proteomics. *Plant Physiol Biochem* **42**, 963–77.
64. Keeling, P. J., Palmer, J. D. (2008) Horizontal gene transfer in eukaryotic evolution. *Nat Rev Genet* **9**, 605–18.