Reconstructing evolution: gene transfer from plastids to the nucleus

Ralph Bock[1]* and Jeremy N. Timmis[2]

Summary

During evolution, the genomes of eukaryotic cells have undergone major restructuring to meet the new regulatory challenges associated with compartmentalization of the genetic material in the nucleus and the organelles acquired by endosymbiosis (mitochondria and plastids). Restructuring involved the loss of dispensable or redundant genes and the massive translocation of genes from the ancestral organelles to the nucleus. Genomics and bioinformatic data suggest that the process of DNA transfer from organelles to the nucleus still continues, providing raw material for evolutionary tinkering in the nuclear genome. Recent reconstruction of these events in the laboratory has provided a unique tool to observe genome evolution in real time and to study the molecular mechanisms by which plastid genes are converted into functional nuclear genes. Here, we summarize current knowledge about plastid-to-nuclear gene transfer in the context of genome evolution and discuss new insights gained from experiments that recapitulate endosymbiotic gene transfer in the laboratory. *BioEssays* 30:556–566, 2008. © 2008 Wiley Periodicals, Inc.

Introduction

Eukaryotic cells arose more than a billion years ago through endosymbiotic engulfment of free-living eubacteria. Subsequently, the enslaved bacterial endosymbionts were gradually converted into two types of DNA-containing cell organelles:

[1]Max-Planck-Institut für Molekulare Pflanzenphysiologie, Potsdam-Golm, Germany.

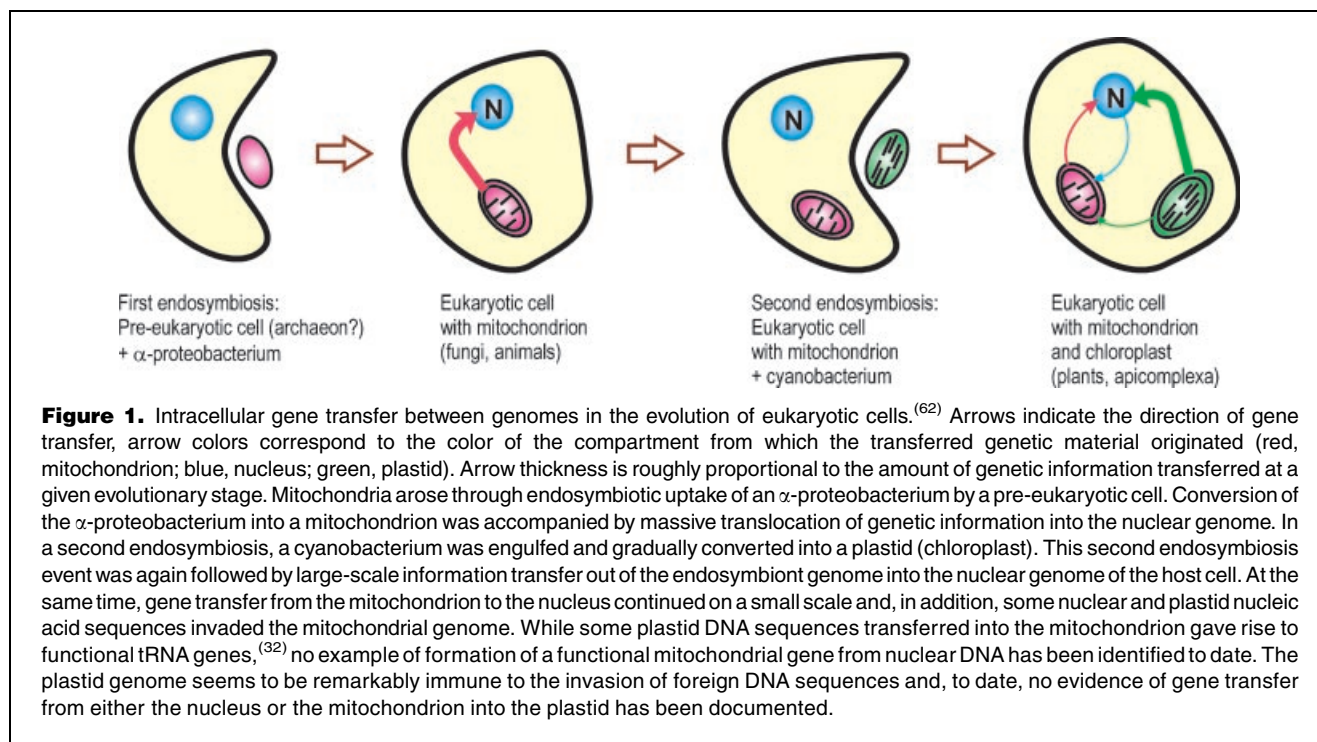[2]School of Molecular and Biomedical Science (Genetics), The University of Adelaide, Adelaide, Australia.

*Correspondence to: Ralph Bock, Max-Planck-Institut für Molekulare Pflanzenphysiologie, Am Mühlenberg 1, D-14476 Potsdam-Golm, Germany. E-mail: rbock@mpimp-golm.mpg.de

Abbreviations: BAC, bacterial artificial chromosome; bp, basepair; cDNA, copy DNA (or complementary DNA); cpDNA, chloroplast DNA; DAPI, 4′,6-diamidino-2-phenylindole; EST, expressed sequence tag; kb, kilobase pair; *mt*, mating type; norgDNA, nuclear organellar DNA; NUMT, nuclear mitochondrial DNA; NUPT, nuclear plastid DNA; ptDNA, plastid DNA; UTR, untranslated region.

the mitochondria, which stem from an α-proteobacterium, and the plastids (chloroplasts), which are derived from a cyanobacterium (Fig. 1). Early genetic and biochemical studies revealed that the genomes of plastids have been greatly diminished compared with any possible free-living ancestor. The plastid genome was shown to be far too small to encode the proteome of the organelle and it was deduced that the control of plastid biogenesis and function is massively dependent upon nuclear genes. This conclusion is fully confirmed by the 122 complete plastid genomes that have been sequenced to date (http://www.bch.umontreal.ca/ogmp/projects/other/cp_list.html). Soon after this realization, elegant early experiments[1,2,3] paved the way for the current understanding of peptide import from the nucleo-cytoplasmic genetic compartment into the organelle.[4]

An explanation of the observed nuclear control over the organelle within the framework of the equally convincing evidence for the endosymbiotic origin of mitochondria and chloroplasts (Fig. 1) required answers to several fundamental questions.[5,6] Firstly, what is the evolutionary relationship between the nuclear genes that encode organellar proteins and the ancestral prokaryote genes of similar function that have not persisted in the plastid DNA (ptDNA)? Secondly, how did the mutual dependence of the separate organellar and nucleo-cytoplasmic systems arise and what are, or were, the selective forces involved in its evolution? One answer to these questions envisaged that redundancy was removed during the early stages of endosymbiont evolution. In this scenario, many ancestral host genes adapted to produce the host product plus another of similar function that was active within the endosymbiont. The second possibility was that genes of the symbiont were physically transferred to the evolving nucleus where they were functionally activated to service the cytoplasmic compartment. In both hypotheses, genes were rapidly deleted from the endosymbiont genome. The diminution of the organelle genomes has been attributed to savings in DNA synthesis for the highly polyhaploid genomes and, though it has not been observed under experimental conditions, this process is clearly rapid in evolution as the plastids of non-photosynthetic parasitic plants, such as *Epifagus virginiana*, contain highly reduced genomes that must have lost many genes relatively recently.[7]

The nuclear control of plastid form and function results in complex regulation of activity of genes in the two compartments. The situation that exists in cells of the majority of sexually

Figure 1. Intracellular gene transfer between genomes in the evolution of eukaryotic cells.[62] Arrows indicate the direction of gene transfer, arrow colors correspond to the color of the compartment from which the transferred genetic material originated (red, mitochondrion; blue, nucleus; green, plastid). Arrow thickness is roughly proportional to the amount of genetic information transferred at a given evolutionary stage. Mitochondria arose through endosymbiotic uptake of an α-proteobacterium by a pre-eukaryotic cell. Conversion of the α-proteobacterium into a mitochondrion was accompanied by massive translocation of genetic information into the nuclear genome. In a second endosymbiosis, a cyanobacterium was engulfed and gradually converted into a plastid (chloroplast). This second endosymbiosis event was again followed by large-scale information transfer out of the endosymbiont genome into the nuclear genome of the host cell. At the same time, gene transfer from the mitochondrion to the nucleus continued on a small scale and, in addition, some nuclear and plastid nucleic acid sequences invaded the mitochondrial genome. While some plastid DNA sequences transferred into the mitochondrion gave rise to functional tRNA genes, [32] no example of formation of a functional mitochondrial gene from nuclear DNA has been identified to date. The plastid genome seems to be remarkably immune to the invasion of foreign DNA sequences and, to date, no evidence of gene transfer from either the nucleus or the mitochondrion into the plastid has been documented.

reproducing organisms is that a diploid nucleus contains thousands of paired alleles on multiple linear disomic chromosomes. These chromosome pairs are separated to produce haploid gametes carrying only a single allele after separation of bivalents during meiosis. Random fertilisation then produces offspring that show characteristic Mendelian ratios when individual alleles are tracked between generations. The contrasting genetic situation in organelles is that they contain multiple haploid genomes, each with a far smaller set of genes than the nucleus and, in plants for example, there may be hundreds of polyhaploid chloroplasts in each leaf cell (Fig. 2). Therefore, although plastid genomes are genetically simple compared with the nucleus, they may comprise a large proportion of total cellular DNA in the leaf cells of some angiosperms, such as spinach (23% of total DNA, 13,000 plastid genomes per cell; Refs 8,9; Fig. 2). However, this proportion is very variable[9,10] and is smaller in many other species (e.g. 1,000 to 1,700 plastid genomes per cell in *Arabidopsis thaliana*; Ref. 11), being dependent upon the nuclear genome size, the degree of nuclear endopolyploidy, the number of plastids per cell and the number of genomes per plastid. Thus the subunit peptides of some plastid protein complexes may be encoded by both nuclear and organellar genes that differ greatly in number within the cell. For example, the large subunit of the most abundant protein on the planet, ribulose bisphosphate carboxylase/oxygenase (RuBisCO), is encoded by thousands of identical genes in the chloroplasts of each tobacco leaf cell, whereas its small subunit is encoded by low-copy-number genes in the single nucleus (Fig. 3).
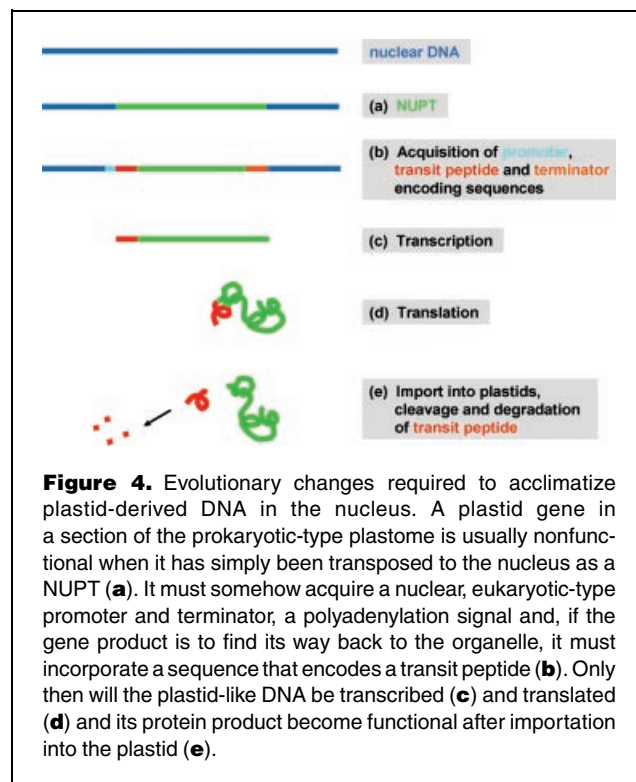

Figure 2. A leaf mesophyll cell of *Spinacea oleracea* stained with DAPI. This picture of a mature spinach leaf cell illustrates the disposition of cellular DNA that is concentrated in the genetically complex diploid nucleus (n) but DNA is also seen in the chloroplasts (cp). Multiple plastid genome copies form nucleoids, of which there are several within each chloroplast. In some species, the plastid genomes, though genetically very simple compared with the nucleus, comprise almost as much of the total cellular DNA as the nucleus because they are present in multiple copies in each organelle and because there are many chloroplasts in each leaf cell (usually more than in the cell shown here). However there is wide species variation in the proportion of cpDNA.[9,11] The chlorophyll molecules within the plastids show red fluorescence. Mitochondria are not visible in this preparation. The bar represents 10 μm.

Figure 3. Contribution of nuclear and plastid genes to the plastid proteome. The plastid genome in angiosperms contains a low number of genes (∼80) that contribute some polypeptides to the multisubunit complexes that support photosynthesis and translation. Several thousands of proteins required for chloroplast biogenesis are not encoded in the ptDNA and are provided *via* the nucleocytoplasmic genetic compartment. Precursors of the required polypeptides are directed to the chloroplast by N-terminal transit peptides that are removed on import into the organelle.

Ancestral organelle genes can be recognized in the nucleus

Gene transfer from the plastid ancestor to the nucleus has played a major role in the evolution of eukaryotes (Fig. 1). Genome-wide phylogenetic comparisons of individual nuclear genes of *Arabidopsis thaliana*[12] with representative prokaryote and eukaryotic genomes revealed that 866 of 9,368 nucleus-encoded proteins that were sufficiently conserved to allow valid comparisons showed closest similarity to proteins of the cyanobacteria. A further 834 proteins generated phylogenetic trees that contained cyanobacterial branches. Extrapolating these data to include genes that could no longer be readily recognized because of their higher rate of divergence suggested that about 4,500 of the ∼25,000 of *A. thaliana* nuclear genes (18%) were acquired from a cyanobacterial ancestor of the plastid. Only some of the nuclear genes derived from cyanobacteria were clearly recognizable as functional in chloroplast biogenesis—others had taken on novel functions elsewhere in, or outside, the cell.[5] These results not only explain the dependence of the chloroplast on nuclear genes but also implicate endosymbiotic gene transfers as a provider of DNA for successful nuclear experimentation on a massive scale and they bring into stark focus the unexpected conclusion that the endosymbiont has made large contributions to the genetic complexity of eukaryotic nuclei. The importance of the endosymbiotic contribution to the evolution of nuclear genes is confirmed by EST analyses in the

glaucophyte *Cyanophora paradoxa*, which showed that 10.8% of nuclear genes were of cyanobacterial origin, though rather fewer of them had non-plastid functions.[13] Sato et al.[14] also made estimates of endosymbiont-derived nuclear genes for *A. thaliana* (4.7%) and a red alga, *Cyanidioschyzon merolae* (12.7%) and assumed that all those revealed were plastid directed. Whether a significant proportion of genes that originated in the ancestral cyanobacterium has been reorganized to take on non-plastid roles remains to be fully clarified.

Ancient or modern transfer, or both?

Much functional plastid gene transfer must have occurred soon after the first conglomerate cells were formed but, remarkably, the process continues, particularly in angiosperms. Millen et al.[15] studied the genomic location of the chloroplast translation initiation factor 1 gene (*infA*) and evidence of recurrent mobility emerged. In these experiments, cases where nuclear relocations were inferred were accompanied by various degrees of decay of the corresponding chloroplast sequence, from minor (but incapacitating) mutations to major sequence decay or comprehensive deletion. Characterization of nuclear *infA* genes provided evidence of the acquisition of de novo transit peptides rather than the hijacking of pre-existing nuclear genes encoding proteins with expedient cellular targets (Fig. 4). However, clear cases of gene hijacking exist. For example, Cusack and Wolfe[16] show how the plastid gene *rpl32* invaded an intron in the nuclear

Figure 4. Evolutionary changes required to acclimatize plastid-derived DNA in the nucleus. A plastid gene in a section of the prokaryotic-type plastome is usually nonfunctional when it has simply been transposed to the nucleus as a NUPT (**a**). It must somehow acquire a nuclear, eukaryotic-type promoter and terminator, a polyadenylation signal and, if the gene product is to find its way back to the organelle, it must incorporate a sequence that encodes a transit peptide (**b**). Only then will the plastid-like DNA be transcribed (**c**) and translated (**d**) and its protein product become functional after importation into the plastid (**e**).

gene *SODcp* in an ancestor of mangrove and poplar trees to give a chimeric *SODcp-RPL32* gene. In mangrove, differentially spliced transcripts encoding either the native SOD precursor or a truncated version of plastid-bound SOD, which also encodes the plastid RPL32, allow continued viability of the species. Both proteins use the SOD transit peptide for protein import into plastids. In poplar, a further evolutionary refinement occurred following gene duplication and divergence of *SODcp-RPL32*, which then allowed each of the two genes to dedicate themselves to different products. The apparent high frequency of functional gene transfer in angiosperms may merely reflect that the group is a strong focus for genomic analyses. An example outside angiosperms is found in a moss, *Physcomitrella patens*, where the *rpoA* gene has relocated to the nucleus compared with its chloroplast location in *Marchantia polymorpha*.[17]

Another contribution of ptDNA to nuclear genes was revealed when bioinformatic analyses discovered nuclear exons in Arabidopsis and rice that are derived from insertions of ptDNA.[18] Although nucleotide sequences were used to expose their origins, the amino acid sequences encoded by these novel nuclear exons show little resemblance to those of the organellar counterparts from which they originated. For example, the Arabidopsis gene (*At2g28820*) encodes a protein with similarity to alanine aminotransferases. Its carboxyterminal domain is clearly derived from 477 bp of the plastid gene *rpl16* to which it is 84% similar. The 73 nucleotide differences that distinguish the domain have a larger effect on the encoded amino acid sequence which is only 74% similar to the corresponding plastid protein, providing for an entirely new function.

These data suggest that gene mobility is far from rare in the plant kingdom. The regularity with which some genes that are amenable to nuclear relocation have completed the process implies that many of those that remain in the organellar genome are forced to do so by strong selection,[19] although some recalcitrant ones can be transferred experimentally.[20] The corollary of this conclusion is a paradox. Why did the transposition of genes predisposed to nuclear transfer not occur much earlier in evolution? It is as though novel mechanisms of transfer have evolved that were at first absent from the cell and the advent of their availability resulted in a recent flush of independent relocations. Such evolutionary developments could include a new system of protein import into organelles, consistent with the variety of sequences that lead to efficient targeting.

DNA per se or are RNA intermediates involved?

DNA molecules themselves could move between genetic compartments, but some studies have reported evidence for the involvement of RNA intermediates. The possibility of direct DNA transfer is founded on evidence from experiments in yeast[21,22] and on comparisons between organellar DNA and nuclear DNA of the same plant species, which show that intergenic spacers and other non-coding regions of organellar DNA are found in nuclear-transferred copies as often as are coding sequences.[23] When plant nuclear genome sequences reveal uninterrupted organelle-like DNA tracts that contain the organellar introns, tRNAs and hundreds of kilobases of organellar non-coding regions, DNA transfer seems most likely to have been involved. The possibility of RNA intermediates is based on observations of certain functional mitochondrial protein-encoding genes that are present in the mitochondrial DNA (mtDNA) in some species of angiosperms, but are located in the nucleus in others.[24,25] Mitochondrial protein-coding genes often have introns and their mRNAs are post-transcriptionally modified by RNA editing, but the nuclear copies seem to lack these organelle-specific marks suggesting they are derived from a fully processed mRNA.[26] Although the possibility that cDNA intermediates might be involved in the transfer of genes from mitochondria to the nucleus in flowering plants cannot be excluded, there are other interpretations of the same data.[27] The predominance of C-to-T (and the associated G-to-A) substitutions caused by the hypermutability of 5-methylcytosine that is abundant in plant nuclear DNAs,[28] might be misinterpreted as signs of mitochondrial editing. This class of substitution is by far the most-common mutation in long tracts of mtDNA or ptDNA in the nucleus (dubbed NUMTs and NUPTs, respectively; see below), whose formation must have been directly from organellar DNA.[28] On a broader scale, evidence that implicates RNA intermediates in nuclear transfers in eukaryotic groups other than flowering plants is so far lacking,[22] as is evidence that implicates cDNA in the transfer of plastid genes.[23]

Non-functional gene transfer: promiscuous plastid DNA in the nucleus

During evolution, most ptDNA transfer events do not result in functional nuclear genes because of the differences between the nuclear and plastid genetic environments. The prerequisite and first step in any scenario in which organellar genes relocate to, and function in, the nucleus necessarily involves migration and integration of nucleic acid molecules. After or during nucleic acid transit, additional complex modifications are required to overcome the "culture shock" for an expatriate gene as it moves from a prokaryotic to a eukaryotic compartment (Fig. 4). Experiments demonstrated fragments of extant ptDNA integrated in the maize mitochondrial genome[29] and into both mitochondrial and nuclear DNA of spinach.[30] In tobacco, it was demonstrated that very long tracts of ptDNA, some inferred to be as large in size as the entire plastid genome (plastome), were abundant in the nucleus.[31]

More-recent genome sequencing confirmed that plastid-like DNA sequences (NUPTs) are common in the nuclear genomes of photosynthetic species and other species that contain a relict plastid.[23,32,33] The diminutive nuclear genome

of *Arabidopsis* contains relatively little plastome-like DNA, but the rice nuclear genome, which is also small, contains many NUPTs totaling >800 kb (Refs 34–36). Chromosome 10 of rice alone contains 28 tracts of ptDNA greater than 80 bp in length including two large insertions of ∼131 and 33 kb (Ref. 37). These integrants showed greater than 95% sequence identity to the bona fide organellar genomes and they are assumed to reveal recent transposition events. Shotgun sequencing projects are prone to underestimate organellar DNA insertions because they may be discarded as contaminants during assembly. The BAC-by-BAC approach that first characterized the 33 kb ptDNA insert on rice chromosome 10 (Ref. 37) must be used to determine efficiently the presence or absence of organelle DNA in the nucleus. This is important as initially it was thought that norgDNAs were not present in the honeybee genome,[33] whereas recent analysis indicates that they are present, and in high copy numbers.[38]

The data pertaining to NUPTs are currently sparse as the plant genomes available for analysis were chosen specifically for their small size. Evidence suggests that norgDNAs are much more abundant in other species. For example, although we now know that NUPTs are present in Arabidopsis, Ayliffe et al.[39] were unable to detect any by the same Southern blotting method that revealed a plethora of such sequences in all other species tested. It is clear that the *Nicotiana tabacum* nucleus has a very large representation of NUPTs and we speculate that the species whose genomes are an order of magnitude larger still, have correspondingly more NUPTs.

Genomic analysis suggests that, though they may be initially integrated into the nucleus as long continuous tracts, norgDNAs are far from static and evolve rapidly via multiple substitutions, insertions, deletions and duplications.[28] Alternatively, they may have been fragmented and reassembled during insertion or there may be mixtures of existing and incoming DNA involving baroque interchromosome rearrangements that culminate in the formation of complex mosaics containing both mitochondrial and plastid DNA from many disjoined parts of the two original organelle genomes.[33]

The apparent paucity of more diverged copies is surprising and may be due to a dynamic equilibrium between ingress and egress (i.e. insertion of new copies versus decay of old copies; Sheppard and Timmis, unpublished results), or to preservation of the sequence by continuous copy correction (gene conversion) with ptDNA and mtDNA.[40] In the absence of such mechanisms, the sequence of some promiscuous DNAs may diverge to the extent that their origin cannot be recognized.

Reconstruction of gene transfer processes in the laboratory

To be able to decipher the precise mechanisms of gene transfer from the plastid to the nuclear genome, an experimental system is needed that facilitates the observation of gene transfer processes in real time. The phylogenetic

evidence discussed above suggested that natural transfer events can be monitored only over large evolutionary time scales. This made it clear that strong selective pressure would be needed to identify gene transfer events in laboratory experiments. The development of techniques to alter the genetic information of the plastid and, most importantly, the ability to integrate foreign genes into the plastid genome by transformation[41,42] have made it possible to design such experiments. A key experiment demonstrating the ongoing transfer of genetic material from plastids into the nucleus is illustrated in Fig. 5. By chloroplast transformation, two genes for antibiotic resistances are introduced into the plastid genome of tobacco (*N. tabacum*) plants: a spectinomycin-resistance gene (*aadA*) with plastid expression signals (a plastid promoter and 5′ untranslated region (5′ UTR) and a plastid 3′ UTR conferring mRNA stability) and a kanamycin-resistance gene (*nptII*) fused to nuclear expression signals (promoter and terminator from the 35S gene of the cauliflower mosaic virus, CaMV; Fig. 5A). Transplastomic plants containing these marker genes in their plastid genomes are resistant to spectinomycin, but not to kanamycin, because the nuclear expression signals of the *nptII* gene do not function efficiently in the organelle. Background expression can be further suppressed by insertion of a nuclear spliceosomal-type intron into the *nptII* gene which cannot be spliced out in the plastid (Fig. 5A). However, if the kanamycin-resistance gene were to jump into the nucleus, it is immediately capable of expression, allowing cell division and plant growth in the presence of kanamycin (Fig. 5B).

Using these transplastomic lines with a eukaryotic-type kanamycin-resistance gene in the plastid genome, two alternative strategies were pursued to select for DNA transfer from the plastid to the nucleus. In one approach, seeds produced by fertilization of wild-type plants with pollen from transplastomic plants were sown on synthetic medium containing kanamycin.[43] In an alternative approach, leaf explants from the transplastomic plants were subjected to kanamycin selection on a plant regeneration medium.[44] Both screens resulted in kanamycin-resistant plants at unexpectedly high frequencies. The seedling selection produced 16 antibiotic-resistant lines in a sample of 250,000 seeds (equaling one transfer event in 16,000 pollen grains). The leaf selection yielded 12 antibiotic-resistant lines from 1,200 tissue explants subjected to regeneration in the presence of kanamycin. If the data from the leaf selection are roughly converted into an estimate for the gene transfer frequency at the cellular level, this frequency is lower (approximately one transfer event per 5 million somatic cells) than in pollen grains. It should not be formally excluded that this discrepancy can be attributed to differences in the experimental design and/or the in vitro selection procedures used. For example, the transplastomic lines used in the pollen experiment have two copies of the *nptII* per plastome, the lines used in the leaf experiment
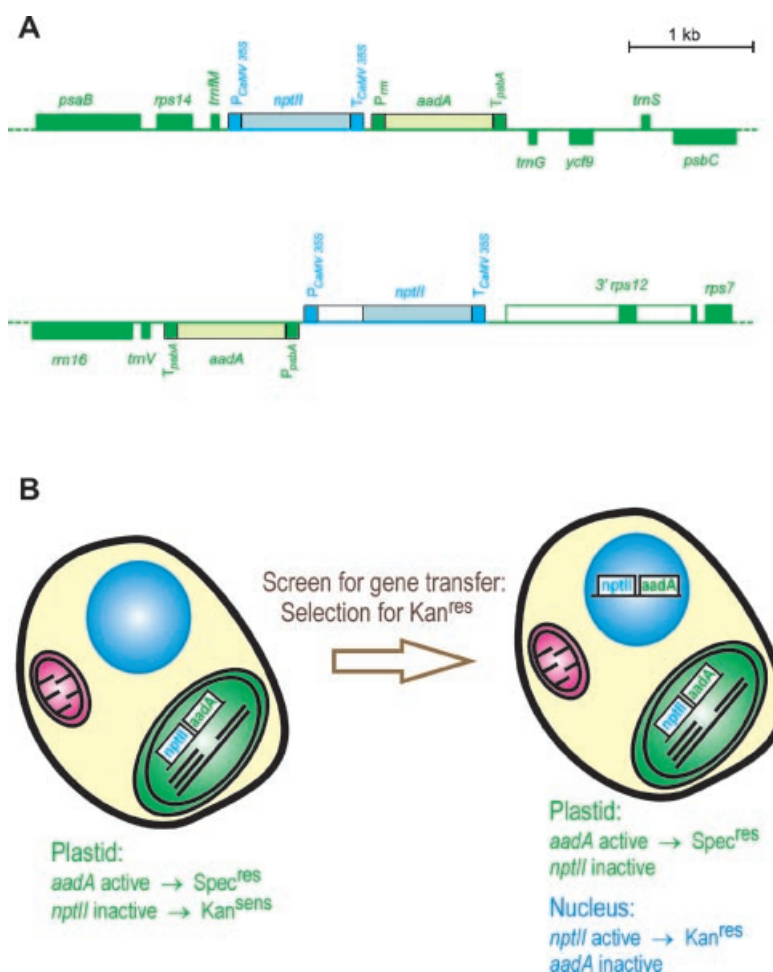
Figure 5. Experimental reconstruction of gene transfer from the plastid to the nuclear genome. **A**: Transgenic chloroplast genomes were produced to facilitate selection for gene transfer events in two independent studies.[43,44] Genes above the line are transcribed from the left to the right; genes below the line are transcribed in the opposite direction. Introns are shown as open boxes. For selection of chloroplast transformants, both studies employed the spectinomycin-resistance gene *aadA* tethered to plastid-specific (i.e. prokaryotic-type) expression signals: a promoter taken from the plastid ribosomal RNA operon (P*rrn*) or the *psbA* gene (P*psbA*) followed by a Shine-Dalgarno sequence mediating translation initiation, and the 3′ untranslated region from the *psbA* gene (T*psbA*) which folds into a stable stemloop-type RNA secondary structure conferring transcript stability. In contrast, the kanamycin-resistance gene *nptII* was provided with nuclear (i.e. eukaryotic-type) expression signals: the strong promoter and terminator from the 35S transcript of the cauliflower mosaic virus (CaMV). In one of the studies,[43] a nuclear spliceosomal intron was additionally incorporated into the *nptII* cassette to reduce background expression of the nuclear-type *nptII* gene in the chloroplast. The two chloroplast transformation constructs also differed in the targeting region of the plastid genome: the antibiotic-resistance genes were either incorporated into the intergenic spacer between the *trnV* gene and the *rps12-rps7* operon in the inverted repeat region of the tobacco plastid genome[43] or were targeted to the spacer between two tRNA genes (*trnfM* and *trnG*) in the large single copy region.[44] **B**: Selection for gene transfer events from the plastid genome to the nucleus. Integration of the two selectable marker gene cassettes into the plastid genome produces cells that are resistant to spectinomycin (because the *aadA* gene is driven by plastid-specific expression signals), but sensitive to kanamycin (because the nuclear expression signals of the *nptII* gene are not efficiently recognized by the prokaryotic gene expression machinery of the plastid). Acquisition of kanamycin resistance requires movement of the *nptII* gene from the plastid to the nuclear genome. Thus, subjecting seedlings[43] or leaf tissue explants[44] to selection on kanamycin-containing medium allows identification of gene transfer events. Note that in Ref. 44, after gene transfer in somatic cells, the plastids of the cell still harbor both resistance genes. This is because the genes integrated into the nucleus are likely to come from a single disintegrated plastid leaving all other plastids and plastid genome copies in the cell unchanged. Due to the maternal mode of plastid inheritance, the transgenic chloroplasts can be eliminated by crosses to wild-type plants in which the transplastomic plants serve as pollen donor (see Fig. 6). When seedlings were screened,[43] all but one instance of kanamycin resistance was selected after the elimination of the transplastome by backcrossing to female wild-type plants.

only one (Fig. 5A). However, a higher gene transfer frequency in pollen would be compatible with the idea that plastid degradation triggers ptDNA transfer to the nucleus.[32] In tobacco and most other angiosperms, plastids are maternally inherited and largely excluded from pollen transmission.[45] One of the cytological mechanisms implicated in the elimination of plastids from sperm cells during pollen grain maturation is plastid disintegration (for review see Ref. 46). Decaying plastids could release huge amounts of ptDNA (Fig. 2) into the cytosol, thus providing ample material for incorporation into the nuclear genome at high frequency. If this is so, it is tempting to speculate that natural selection may have favored uniparental inheritance simply because it increases organellar DNA transfer to the nucleus in addition to, or rather than, the other explanations that have been offered for its existence.[47,48]

It also should be noted that the data obtained from the selection experiments give rather conservative estimates of the gene transfer frequencies: all those transfer events that did not include the entire selectable kanamycin marker gene went undetected in the screens. Also, integrations into unfavorable genomic locations, such as transcriptionally incompetent heterochromatic regions, may not have allowed for sufficiently strong expression of the *nptII* to give rise to kanamycin-resistant lines. Thus it seems entirely possible that the real frequencies are even higher than the ones measured in the laboratory experiments.

The mapping of the transferred ptDNA pieces and the analysis of integration sites in the nuclear genome[49] revealed several interesting insights into the mechanisms of gene transfer. First, large pieces of ptDNA (between 6 and more than 22 kb) were found to be transferred largely excluding involvement of RNA/cDNA intermediates and lending support to the idea that direct DNA-mediated transfer represents the prevailing mechanism of nucleic acid transfer from the plastid to the nucleus. Second, some NUPT loci displayed a complex structure and were composed of multiple transferred ptDNA fragments. This may indicate that more than one copy of the plastome is released at the onset of a transfer event, which makes organelle destruction a prime candidate process initiating gene transfer. Finally, microhomologies of 2 to 5 bp were found at the integration sites, possibly suggesting that NUPT integration proceeds via microhomology-mediated illegitimate recombination.[49]
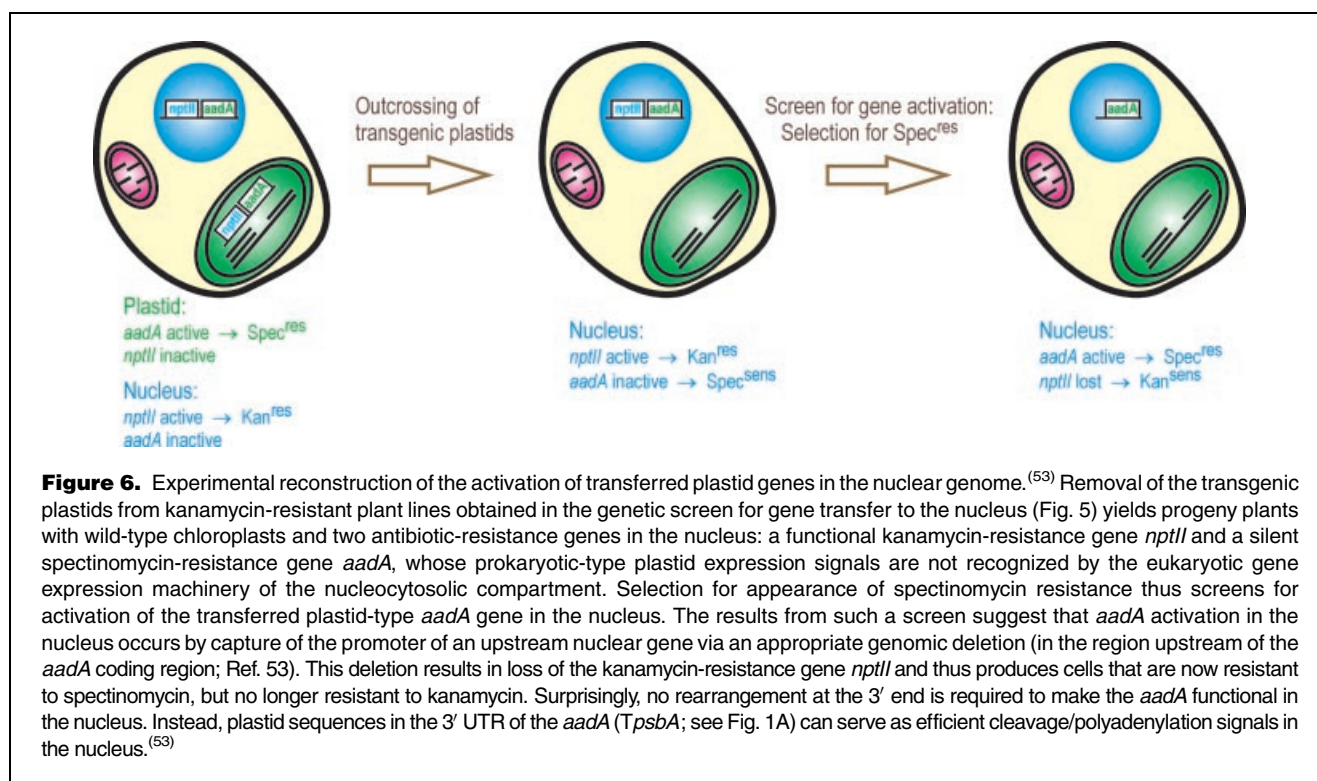
It is insightful to note that attempts to reproduce the tobacco gene somatic cell transfer experiments in the haploid unicellular green alga *Chlamydomonas reinhardtii* have failed.[50] This seems surprising at first sight, but there may be a simple explanation: unlike angiosperms (Fig. 2), *Chlamydomonas* has only a single large chloroplast. As discussed above, it is reasonable to assume that the release of DNA from the chloroplast does not occur via an active export process, but rather requires chloroplast decay, as it could occur, for example, after irreversible photooxidative damage. If the cell has only a single chloroplast, its lysis would be a lethal event, explaining why gene transfer from the plastid to the nucleus is not observed in *Chlamydomonas*. Nevertheless, many short NUPTs certainly occur in *Chlamydomonas*,[51] suggesting that a search for transfer during sexual reproduction, and the associated destruction of the *mt⁻* plastid DNA,[52] may be fruitful.

The surprisingly high frequency of DNA escape from plastids to the nucleus has profound implications, not only for the experimental reconstruction of eukaryotic genome evolution, but also for understanding nuclear genome stability and dynamics. For example, the rate of DNA transfer measured in tobacco exceeds by far the number of cells in a single leaf, leading to the somewhat shocking conclusion that the cells within a single leaf of a plant are genetically heterogeneous and differ in their patterns of NUPT (and presumably also NUMT) integrants. Thus, plastid-to-nuclear gene transfer must be considered as a major source of intraspecific and even intraorganismic genetic variation.[39] In addition, the high frequency of gene transfer should regularly result in the disruption of nuclear genes (or their *cis*-regulatory elements) or in the generation of chimeric genes by in-frame fusion of a nuclear gene with an incoming plastid gene. To what extent such integration events cause phenotypically significant somatic mutations and/or new functional gene variants is yet to be learnt but, in the extreme, DNA transfer from organelles to the nucleus could turn out to be more important to genome evolution and represent a much more significant source of genetic variation than, for example, transposons or retrotransposons.

Functional activation of transferred plastid transgenes in the nucleus

As a consequence of the plastid's prokaryotic origin, transcription of plastid genes as well as RNA processing and translation are controlled by prokaryotic-type regulatory sequence elements, which do not function in the eukaryotic environment of the nucleus. In addition to eukaryotic-type promoter elements, nuclear genes contain specific sequences that direct faithful mRNA 3′ end cleavage and polyadenylation, which are essential processing steps for transcript stability and translation. Therefore, any piece of plastid DNA arriving in the nucleus is unlikely to be expressed (Fig. 4). Consistent with this expectation, the chloroplast selectable marker gene *aadA*, although co-transferred with the kanamycin-resistance gene *nptII*, was found to be inactive in the gene transfer lines selected as described above (Figs. 5 and 6). Consequently, the selected gene transfer lines were sensitive to spectinomycin (and the structurally related aminoglycoside antibiotic streptomycin). The structure of the nuclear locus resembles the landing of a transferred plastid gene, the *aadA*, downstream of a functional nuclear gene, the *nptII* (Fig. 5A upper panel; Fig. 6). To test whether or not transferred silent plastid genes can give rise to

Figure 6. Experimental reconstruction of the activation of transferred plastid genes in the nuclear genome.[53] Removal of the transgenic plastids from kanamycin-resistant plant lines obtained in the genetic screen for gene transfer to the nucleus (Fig. 5) yields progeny plants with wild-type chloroplasts and two antibiotic-resistance genes in the nucleus: a functional kanamycin-resistance gene *nptII* and a silent spectinomycin-resistance gene *aadA*, whose prokaryotic-type plastid expression signals are not recognized by the eukaryotic gene expression machinery of the nucleocytosolic compartment. Selection for appearance of spectinomycin resistance thus screens for activation of the transferred plastid-type *aadA* gene in the nucleus. The results from such a screen suggest that *aadA* activation in the nucleus occurs by capture of the promoter of an upstream nuclear gene via an appropriate genomic deletion (in the region upstream of the *aadA* coding region; Ref. 53). This deletion results in loss of the kanamycin-resistance gene *nptII* and thus produces cells that are now resistant to spectinomycin, but no longer resistant to kanamycin. Surprisingly, no rearrangement at the 3′ end is required to make the *aadA* functional in the nucleus. Instead, plastid sequences in the 3′ UTR of the *aadA* (T*psbA*; see Fig. 1A) can serve as efficient cleavage/polyadenylation signals in the nucleus.[53]

functional nuclear genes, a second genetic screen was designed to select for activation of the transferred *aadA* gene in the nuclear genome. To this end, three gene transfer lines were selected in which the entire *nptII* and *aadA* gene cassettes were successfully transferred to the nucleus. In order to facilitate selection for *aadA* functionalization in the nucleus, the transgenic plastids were crossed out and replaced by wild-type plastids (Fig. 6). These plant lines were then subjected to a second large-scale screen by exposing leaf explants to plant regeneration medium containing spectinomycin.[53] Spectinomycin and streptomycin-resistant lines with a functionally activated *aadA* gene in the nucleus could indeed be obtained, demonstrating that DNA-mediated gene transfer can give rise to functional nuclear genes when followed by suitable mutations or rearrangements in the nuclear genome. Selection from over 5,500 leaf explants resulted in the isolation of eight lines that showed transcriptional activation of the transferred *aadA* gene in the nucleus,[53] a frequency that is not much lower than the frequency of plastid-to-nuclear DNA transfer measured in the leaf explant system.[44]

Characterizing the molecular changes that had occurred in these lines revealed that the acquisition of gene function was always associated with the capture of the promoter of an upstream nuclear gene. In all lines, rearrangements had placed the *aadA* gene under the control of the *nptII* promoter. This was accomplished by deletions occurring either in short directly repeated sequences (presumably via microhomology-mediated illegitimate recombination) or, alternatively, after the

rejoining of adjacent DNA double-strand breaks. As a result of these deletions, *nptII* gene function was knocked out such that acquisition of spectinomycin resistance was always accompanied by loss of kanamycin resistance[53] (Fig. 6). Interestingly, when the downstream regions of the activated *aadA* were analyzed to determine if and how the genes might have acquired mRNA 3′ end cleavage and polyadenylation signals, no evidence of rearrangements was found in any of the eight lines obtained in the screen. Instead, mapping of the cleavage and polyadenylation site in the *aadA* transcripts indicated that the signals needed for mRNA 3′ end formation had been provided by the plastid 3′ UTR downstream of the *aadA* coding region (i.e. the 3′ UTR from the plastid *psbA* gene; Fig. 5A, upper panel). Inspection of the sequence in the *psbA* 3′ UTR that was utilized as cleavage and polyadenylation site revealed that it matched the loose consensus sequence for mRNA 3′ end formation in plants, which is highly AU rich. AT richness is commonly found in plastid genomes, particularly in UTRs and intergenic regions. This raises the intriguing possibility that the lack of the need to acquire specific sequences for faithful mRNA 3′ end formation could be responsible for the high success rate of gene transfer during evolution by limiting the requirements for gene functionalization in the nucleus to promoter capture.[53]

Obviously, the frequency of functional gene transfer would be lower if, additionally, it required acquisition of a transit peptide sequence for re-routing of the protein product into the plastid (Fig. 4). However, somewhat counter intuitively,

re-import has not been the rule in evolution and the proteins encoded by many genes that were transferred from the cyano-bacterial endosymbiont to the *A. thaliana* nucleus are now targeted to cellular compartments other than the plastid.[12]

In summary, the design of efficient screens for gene transfer from the plastid genome to the nucleus has made it possible to reproduce events in the laboratory that normally were expected to occur only over large evolutionary time scales. The successful experimental reconstruction of DNA transfer and gene functionalization in the nucleus, together with analyses of the molecular mechanisms involved in these processes, have provided important insights into how plastid gene functions can be taken over by the nucleus, and how organellar DNA may take on other diverse genetic roles within the cell. Clearly, these insights go far beyond what has been inferred from comparative genomics and phylogenetic analyses and thus highlight the great potential of experimental evolution.

Why have genes been retained in the plastid?

Given the high frequency of gene transfer and the successful translocation of more than 95% of the genes of the cyanobacterial endosymbiont into the nucleus, an intriguing question is why plastids have retained a remnant genome at all. In view of the existence of mitochondria without a genome (e.g. many of the so-called hydrogenosomes; Refs 54–56) and the absence of a genome from many other membrane-bound cell organelles (e.g. peroxisomes, lysosomes, glyoxisomes), there seems to be no obvious reason why plastids would need a genome. In fact, since plastids are asexual genetic systems and excluded from sexual recombination, the location of genes in the plastid should even pose a serious disadvantage. This is because asexual reproduction is thought to result in the accumulation of mutations over time, a hypothesis known as Muller's ratchet.[57] Since the vast majority of mutations are deleterious, any asexual genetic system is expected to suffer a continuous decline in fitness. However, in spite of their asexual reproduction, plastid genomes have even lower mutation rates than nuclear genomes.[58] A recent study has provided experimental evidence that may resolve this conundrum: the plastid's high polyhaploidy, together with the presence of gene conversion as an efficient mutation-correcting mechanism, counteracts the detrimental effects of Muller's ratchet and keeps mutation rates in plastid genomes tolerably low.[59] This suggests a molecular link between asexual reproduction, high genome copy numbers and low mutation rates and thus could also explain why there may be no strong selective pressure driving genes out of present-day plastids into the nucleus.

Interestingly, the set of genes that has been retained in the chloroplast genome appears not to be the result of chance. For example, a conspicuous pattern seen in the encoding of membrane protein complexes is that the core subunits (e.g. the reaction center proteins of the two photosystems) are encoded by the plastid genome whereas most peripheral subunits are encoded by the nuclear genome. Many different hypotheses have been entertained to explain why a specific set of genes has been retained in the plastome.[60] One of them posits that nuclear encoding of the core subunits of membrane protein complexes potentially could be deleterious because the complexes could, at least partially, assemble in non-target membranes, such as the plasma membrane, the endoplasmic reticulum or the vacuolar membrane. Another hypothesis proposes that these proteins are difficult to import into the plastid due to their high hydrophobicity and/or unwieldy three-dimensional structure. Yet another hypothesis, termed 'co-location of gene and gene product for redox regulation of gene expression' (CORR) postulates a selective advantage conferred by redox-regulated expression of reaction center proteins directly in the plastid. Such a feedback regulation by redox signals would allow the individual chloroplast to control autonomously the accumulation of those protein complexes that are the key to environmental adaptation. It is conceivable that, even within a single cell, different chloroplasts have different needs for the protein complexes of the photosynthetic electron transport chain (photosystem II, cytochrome b_6f complex, photosystem I, ATP synthase). Plastid encoding of core subunits whose availability limits complex assembly could provide a direct way to regulate these genes by redox signals from the electron transport chain and thus to adjust complex stoichiometries at the level of the individual plastid.[60] It is probably fair to say that none of these hypotheses is supported by more than circumstantial evidence. Also, it should be noted that they are not necessarily mutually exclusive. It seems entirely possible that a combination of chance and the selective advantages or disadvantages of plastid versus nuclear location would account for the composition of the set of genes found in present-day plastids.

Another intriguing question asks whether the current status, in the most-gene-attenuated functional plastomes, reflects a more or less final stage of the evolutionary optimization of endosymbiosis in that all those genes that are better off in the chloroplast are still there while all others have been transferred to the nucleus. Obviously, this question is difficult to answer, but the possibility to relocate genes between the plastid and the nuclear genomes by genetic transformation techniques[20] should allow the design of suitable experiments to test whether or not we are now witnessing an evolutionary static state of affairs.

Conclusions and outlook

Genomics, bioinformatics and experimental data have provided compelling evidence for gene transfer from the plastid to the nuclear genome. In evolution, the translocation of thousands of genes from the cyanobacterial endosymbiont to the nucleus has shaped decisively the plant nuclear genome. Moreover, DNA transfer from the plastid to the nucleus is an

ongoing process that keeps providing new raw material for genome evolution and represents a constant source of intraspecific genetic variation. The rapid progression of whole-genome sequencing is expected to provide new insights into the evolutionary dynamics of plastid-to-nucleus gene transfer. It will be particularly interesting to search the genomes for evolutionary intermediates of functional gene transfer processes, including pseudogenes in plastid genomes and plastid genes in the nucleus that show signs of incipient functional activation.[18]

The experimental demonstration that the transfer of large pieces of plastid DNA into the nucleus occurs at high frequency raises a number of intriguing prospects. First, it will be interesting to see to what extent these transfer events contribute to the production of new functional genes in the nucleus. Second, the steady inflow of genetic material from plastids (and mitochondria) into the nuclear genome raises the question of how the nuclear genome deals with these large amounts of incoming DNA. It seems plausible to assume that mechanisms must be in place to eliminate efficiently most of the transferred DNA, which otherwise would cause rampant enlargement of the nuclear genome and severely compromise genome stability. One possible mechanism could be recombination at small direct repeats. Microhomologies at insertion sites of plastid sequences in the nuclear genome[49] and at the breakpoints of deletions resulting in the functional activation of transferred plastid genes,[53] suggest that microhomology-mediated illegitimate recombination represents a major mechanism underlying genome rearrangements in the nucleus. Thus it is possible that most gene transfer events are reversed by the very same microhomology-mediated recombination event that caused the integration of the plastid DNA into the nuclear genome. Whether or not plastid DNA sequences are particularly good substrates for this recombination type is unknown, but high AT content and lack of cytosine methylation in plastid DNA could help the genome surveillance machinery in the nucleus to identify promiscuous plastid sequences reliably and to eliminate most of them successfully. Finally, it will be interesting to test whether or not the frequency of gene transfer and the frequency of functional activation of transferred plastid genes in the nucleus are dependent on environmental conditions, such as exposure to biotic and abiotic stresses. Recent work has demonstrated that, in yeast, microhomology-mediated illegitimate recombination is stimulated by ionizing radiation and other DNA-damaging treatments.[61] Thus, it is conceivable that certain stresses, such as UV light stress or pathogen infection, could not only trigger the release of plastid DNA into the nucleocytosolic compartment by causing plastid damage and decay, but could also stimulate the integration of the released DNA into the nuclear genome.

Last but not least, we believe that the experimental reconstruction of key evolutionary processes, such as gene transfer from endosymbiotic organelles to the genome of their host cell, has important educational and societal implications. The past few years have witnessed a bizarre increase in attempts to discredit evolutionary theory and deny the existence of compelling scientific evidence for the evolution of life on earth. It is hoped that, with the growing number of evolutionary processes that can be reproduced in the laboratory, these dogmatic and ideological arguments will lose more and more ground such that public understanding and acceptance of evolutionary theory will concomitantly increase.

Acknowledgments

References

1. Highfield PE, Ellis RJ. 1978. Synthesis and transport of the small subunit of chloroplast ribulose bisphosphate carboxylase. Nature 271:420–424.
2. Dobberstein B, Blobel G, Chua N-H. 1977. In vitro synthesis and processing of a putative precursor for the small subunit of ribulose-1-5-bisphosphate carboxylase of Chlamydomonas reinhardtii. Proc Natl Acad Sci USA 74:1082–1085.
3. Bedbrook JR, Smith SM, Ellis RJ. 1980. Molecular cloning and sequencing of cDNA encoding the precursor to the small subunit of chloroplast ribulose-1,5-bisphosphate carboxylase. Nature 287:692–697.
4. Soll J, Schleiff E. 2004. Protein import into chloroplasts. Nature Rev Mol Cell Biol 5:198–208.
5. Martin W, Herrmann RG. 1998. Gene transfer from organelles to the nucleus: how much, what happens, and why? Plant Physiol 118:9–17.
6. Martin W, Stoebe B, Goremykin V, Hansmann S, Hasegawa M, Kowallik KV. 1998. Gene transfer to the nucleus and the evolution of chloroplasts. Nature 393:162–165.
7. Wolfe KH, Morden CW, Palmer JD. 1991. Ins and outs of plastid genome evolution. Curr Op Genet Dev 1:523–529.
8. Lawrence ME, Possingham JV. 1986. Microspectrofluorometric measurement of chloroplast DNA in dividing and expanding leaf cells of Spinacia oleracea. Plant Physiol 81:708–710.
9. Bendich AJ. 1987. Why do chloroplasts and mitochondria contain so many copies of their genome? BioEssays 6:279–282.
10. Li W, Ruf S, Bock R. 2006. Constancy of organellar genome copy numbers during leaf development and senescence in higher plants. Mol Gen Genomics 275:185–192.
11. Zoschke R, Liere K, Börner T. 2007. From seedling to mature plant: Arabidopsis plastidial genome copy number, RNA accumulation and transcription are differentially regulated during leaf development. Plant J 50:710–722.
12. Martin W, Rujan T, Richly E, Hansen A, Cornelsen S, 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 USA 99:12246–12251.
13. Reyes-Prieto A, Hackett JD, Soares MB, Bonaldo MF, Bhattacharya D. 2006. Cyanobacterial contribution to algal nuclear genomes is primarily limited to plastid functions. Curr Biol 16:2320–2325.
14. Sato N, Ishikawa M, Fujiwara M, Sonoike K. 2005. Mass identification of chloroplast proteins of endosymbiont origin by phylogenetic profiling based on organism-optimized homologous protein groups. Genome Inform 16:56–68.
15. Millen RS, Olmstead RG, Adams KL, Palmer JD, Lao NT, et al. 2001. Many parallel losses of infA from chloroplast DNA during angiosperm evolution with multiple independent transfers to the nucleus. Plant Cell 13:645–658.
16. Cusack BP, Wolfe KH. 2007. When gene marriages don't work out: divorce by subfunctionalization. Trends Genet 23:270–272.

17. Sugiura C, Kobayashi Y, Aoki S, Sugita C, Sugita M. 2003. Complete chloroplast DNA sequence of the moss Physcomitrella patens: evidence for the loss and relocation of rpoA from the chloroplast to the nucleus. Nucleic Acids Res 31:5324–5331.

18. Noutsos C, Kleine T, Armbruster U, DalCorso G, Leister D. 2007. Nuclear insertions of organellar DNA can create novel patches of functional exon sequences. Trends Genet 23:597–601.

19. Allen JF, Puthiyaveetil S, Ström J, Allen CA. 2005. Energy transduction anchors genes in organelles. BioEssays 27:426–435.

20. Kanevski I, Maliga P. 1994. Relocation of the plastid rbcL gene to the nucleus yields functional ribulose-1,5-bisphosphate carboxylase in tobacco chloroplasts. Proc Natl Acad Sci USA 91:1969–1973.

21. Thorsness PE, Fox TD. 1990. Escape of DNA from mitochondria to the nucleus in Saccharomyces cerevisiae. Nature 346:376–379.

22. Thorsness PE, Weber ER. 1996. Escape and migration of nucleic acids between chloroplasts, mitochondria and the nucleus. Int Rev Cytol 165:207–234.

23. Timmis JN, Ayliffe MA, Huang CY, Martin W. 2004. Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. Nature Rev Genet 5:123–136.

24. Adams KL, Daley DO, Qiu Y-L, Whelan J, Palmer JD. 2000. Repeated, recent and diverse transfer of a mitochondrial gene to the nucleus in flowering plants. Nature 408:354–357.

25. Adams KL, Palmer JD. 2003. Evolution of mitochondrial gene content: gene loss and transfer to the nucleus. Mol Phyl Evol 29:380–395.

26. Adams KL, Qiu Y-L, Stoutemeyer M, Palmer JD. 2002. Punctuated evolution of mitochondrial gene content: High and variable rates of mitochondrial gene loss and transfer to the nucleus during angiosperm evolution. Proc Natl Acad Sci USA 99:9905–9912.

27. Henze K, Martin W. 2001. How do mitochondrial genes get into the nucleus. Trends Genet 17:383–387.

28. Huang CY, Grünheit N, Ahmadinejad N, Timmis JN, Martin W. 2005. Mutational decay and age of chloroplast and mitochondrial genomes transferred recently to angiosperm nuclear chromosomes. Plant Physiol 138:1723–1733.

29. Stern DB, Lonsdale DM. 1982. Mitochondrial and chloroplast genomes of maize have a 12-kilobase DNA sequence in common. Nature 299:698–702.

30. Timmis JN, Scott NS. 1983. Spinach nuclear and chloroplast DNAs have homologous sequences. Nature 305:65–67.

31. Ayliffe MA, Timmis JN. 1992. Tobacco nuclear DNA contains long tracts of homology to chloroplast DNA. Theor Appl Genet 85:229–238.

32. Bock R. 2006. Extranuclear inheritance: Gene transfer out of plastids. Prog Bot 67:75–98.

33. Leister D. 2005. Origin, evolution and genetic effects of nuclear insertions of organelle DNA. Trends Genet 21:655–663.

34. Noutsos C, Richly E, Leister D. 2005. Generation and evolutionary fate of insertions of organelle DNA in the nuclear genomes of flowering plants. Genome Res 15:616–628.

35. Matsuo M, Ito Y, Yamauchi R, Obokata J. 2005. The rice nuclear genome continuously integrates, shuffles, and eliminates the chloroplast genome to cause chloroplast-nuclear DNA flux. Plant Cell 17:665–675.

36. Richly E, Leister D. 2004. NUMTs in sequenced eukaryotic genomes. Mol Biol Evol 21:1081–1084.

37. Yuan Q, Hill J, Hsiao J, Moffat K, Ouyang S, et al. 2002. Genome sequencing of a 239-kb region of rice chromosome 10L reveals a high frequency of gene duplication and a large chloroplast DNA insertion. Mol Genet Genomics 267:713–720.

38. Pamilo P, Viljakainen L, Vihavainen A. 2007. Exceptionally high density of NUMTs in the honeybee genome. Mol Biol Evol 24:1340–1346.

39. Ayliffe MA, Scott NS, Timmis JN. 1998. Analysis of plastid DNA-like sequences within the nuclear genomes of higher plants. Mol Biol Evol 15:738–745.

40. Clifton SW, Minx P, Fauron CM-R, Gibson M, Allen JO, et al. 2004. Sequence and comparative analysis of the maize NB mitochondrial genome. Plant Physiol 136:3486–3503.

41. Bock R. 2001. Transgenic chloroplasts in basic research and plant biotechnology. J Mol Biol 312:425–438.

42. Maliga P. 2004. Plastid transformation in higher plants. Annu Rev Plant Biol 55:289–313.

43. Huang CY, Ayliffe MA, Timmis JN. 2003. Direct measurement of the transfer rate of chloroplast DNA into the nucleus. Nature 422:72–76.

44. Stegemann S, Hartmann S, Ruf S, Bock R. 2003. High-frequency gene transfer from the chloroplast genome to the nucleus. Proc Natl Acad Sci USA 100:8828–8833.

45. Ruf S, Karcher D, Bock R. 2007. Determining the transgene containment level provided by chloroplast transformation. Proc Natl Acad Sci USA 104:6998–7002.

46. Bock R. 2007. Structure, function, and inheritance of plastid genomes. Top Curr Genet 19:29–63.

47. Allen JF. 1996. Separate sexes and the mitochondrial theory of aging. J theor Biol 180:135–140.

48. Burt A, Trivers R. 2006. Genes in conflict: the biology of selfish genetic elements. London, England: The Belknap Press of Harvard University Press, Cambridge, USA.

49. Huang CY, Ayliffe MA, Timmis JN. 2004. Simple and complex nuclear loci created by newly transferred chloroplast DNA in tobacco. Proc Natl Acad Sci USA 101:9710–9715.

50. Lister DL, Bateman JM, Purton S, Howe CJ. 2003. DNA transfer from chloroplast to nucleus is much rarer in Chlamydomonas than in tobacco. Gene 316:33–38.

51. Richly E, Leister D. 2004. NUPTs in sequenced eukaryotes and their genomic organization in relation to NUMTs. Mol Biol Evol 21:1972–1980.

52. Nishimura Y, Misumi O, Kato K, Inada N, Higashiyma T, et al. 2002. An mt$^+$ gamete-specific nuclease that targets mt$^-$ chloroplasts during sexual reproduction in C. reinhardtii. Genes and Dev 16:1116–1128.

53. Stegemann S, Bock R. 2006. Experimental reconstruction of functional gene transfer from the tobacco plastid genome to the nucleus. Plant Cell 18:2869–2878.

54. Tielens AGM, Rotte C, van Hellemond JJ, Martin W. 2002. Mitochondria as we don't know them. Trends Biochem Sci 27:564–572.

55. van der Giezen M, Tovar J. 2005. Degenerate mitochondria. EMBO Rep 6:525–530.

56. Hackstein JHP, Tjaden J, Huynen M. 2006. Mitochondria, hydrogenosomes and mitosomes: products of evolutionary tinkering. Curr Genet 50:225–245.

57. Muller HJ. 1964. The relation of recombination to mutational advance. Mutat Res 1:2–9.

58. Wolfe KH, Li W-H, Sharp PM. 1987. Rates of nucleotide substitutions vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. Proc Natl Acad Sci USA 84:9054–9058.

59. Khakhlova O, Bock R. 2006. Elimination of deleterious mutations in plastid genomes by gene conversion. Plant J 46:85–94.

60. Allen JF. 2003. The function of genomes in bioenergetic organelles. Phil Trans R Soc Lond B 358:19–38.

61. Chan CY, Kiechle M, Manivasakam P, Schiestl RH. 2007. Ionizing radiation and restriction enzymes induce microhomology-mediated illegitimate recombination in Saccharomyces cerevisiae. Nucleic Acids Res 35:5051–5059.

62. Gray MW, Burger G, Lang BF. 1999. Mitochondrial evolution. Science 283:1476–1481.