Chapter 9

Promiscuous Organellar DNA

Andrew H. Lloyd
School of Molecular and Biomedical Science, The University of Adelaide, South-Australia 5005, Australia
Institut Jean-Pierre Bourgin, UMR1318, Institut National de la Recherche Agronomique, 78026 Versailles, France

Mathieu Rousseau-Gueutin and Jeremy N. Timmis*
School of Molecular and Biomedical Science, The University of Adelaide, South-Australia 5005, Australia

Anna E. Sheppard
School of Molecular and Biomedical Science, The University of Adelaide, South-Australia 5005, Australia
Department of Evolutionary Ecology and Genetics, Zoological Institute, Christian Albrechts University of Kiel, Am Botanischen Garten 1-9, 24118 Kiel, Germany

Michael A. Ayliffe
CSIRO Division of Plant Industry, GPO Box 1600, ACT 2601, Australia

Summary 202
I. Introduction. 202
II. Organelle Genome Reduction 202
  A. Evolutionary Gene Transfer to the Nucleus 202
  B. Recent Gene Transfer Events 204
  C. Why Relocate? 205
  D. Why Retain an Organellar Genome? 206
III. Promiscuous DNA: Ongoing Organelle DNA Transfer to the Nucleus 206
  A. Organelle Sequences in Nuclear Genomes 206
  B. Evidence of Frequent Plastid and Mitochondrial DNA Transfer to the Nucleus 208
  C. Evolutionary Fate of Nuclear Located Cytoplasmic Organelle DNA (*norgs*) 210
IV. Mechanisms of Gene Transfer to the Nucleus. 210
  A. Relocation of Genetic Material 210
  B. Is There an RNA or DNA Intermediate? 211
  C. Integration into Nuclear Chromosomes 212

*Author for correspondence, e-mail: jeremy.timmis@adelaide.edu.au

R. Bock and V. Knoop (eds.), Genomics of Chloroplasts and Mitochondria, Advances in Photosynthesis and Respiration 35, pp. 201–221,
DOI 10.1007/978-94-007-2920-9_9, © Springer Science+Business Media B.V. 2012

V. Activation of a Newly Transferred Organelle Gene....213
A. Examples of Organellar Gene Activation in the Nucleus....213
B. Experimental Attempts to Detect Activation of a Chloroplast Gene After Transfer to the Nucleus....213
VI. Plastid DNA in Higher Plant Mitochondria. 214
VII. Perspective....216
Acknowledgements....216
References....217

Summary

Endosymbiotic transfer of DNA from the cytoplasmic organelles (mitochondria and chloroplasts) to the nucleus has been a major factor driving the origin of new nuclear genes, a process central to eukaryote evolution. Typically, transfer of organelle DNA to the nucleus is quickly followed by decay, deletion and rearrangement. However, in rare instances these new sequences lead to functional relocation of organelle genes to the nucleus or the generation of genes with novel function. Similar transfer of chloroplast DNA has also added to the complexity of plant mitochondrial genomes. Significantly, these processes are ongoing, making promiscuous organellar DNA an important contributor to the continued evolution of plant genomes.

I. Introduction

The nucleus, with its translational machinery in the cytoplasm, and the mitochondrion comprise two separate genetic compartments of eukaryotic cells. In plants, algae and some protist lineages, the plastid is a third genetic compartment. The mitochondria and plastids (herein referred to as "cytoplasmic organelles" or simply "organelles") have an endosymbiotic origin and are the extant descendants of once free-living α-proteobacteria and cyanobacteria, respectively. Following their incorporation into the ancestor of the eukaryote cell, the ancestors of these two cytoplasmic organelles underwent large-scale genome reduction so that their current genomes contain only 1–5% of the gene complement found in any candidate modern free-living prokaryotic relative. This genome reduction was enabled both by intracellular redundancy after cohabitation and by functional relocation of genes to the nucleus. In plants, recent functional relocation of organellar genes has been documented and transfer of non-functional DNA to the nucleus is still happening at very high frequency. As a result of the latter process, tracts of DNA that are essentially identical to regions of the extant plastid and mitochondrial genomes are found within all photosynthetic (or once photosynthetic) eukaryote nuclear genomes.

Abbreviations: CaMV – Cauliflower mosaic virus; DSB – Double strand break; GUS – β-glucuronidase; *mtpt* – Mitochondrial integrant of plastid DNA; MYA – Million years ago; NHEJ – Non-homologous end joining; *norg* – Nuclear integrant of organellar DNA; *numt* – Nuclear integrant of mitochondrial DNA; *nupt* – Nuclear integrant of plastid DNA; TAIL-PCR – Thermal asymmetric interlaced PCR; T-DNA – Transfer DNA; TOC/TIC – Translocase at the outer/inner envelope membrane of chloroplasts; TOM/TIM – Translocase at the outer/inner envelope membrane of mitochondria

II. Organelle Genome Reduction

A. Evolutionary Gene Transfer to the Nucleus

Consistent with their endosymbiotic origin, chloroplasts and mitochondria retain essentially prokaryote-like genomes. They are

separate genetic compartments in that they enclose separate transcriptional and translational machineries necessary for intraorganellar gene expression. Their respective genomes are much reduced in size when compared with those of the extant relatives of their free-living ancestors, retaining few of the ancestral protein coding genes. Mitochondrial genomes are the most reduced in size – containing only 3–67 protein-coding genes (Timmis et al. 2004), while chloroplast genomes generally encode several more proteins – around 80 in land plants and over 200 in some algae (Timmis et al. 2004).

The reduction in genome size has been, in part, due to loss of genes made redundant when the endosymbiont became resident within the eukaryote cell. However, much of the disappearance of genes from the endosymbiont's genome is due to the functional relocation of cytoplasmic organelle genes to the nuclear genome. In many cases, proteins encoded by these relocated genes retain their original role in organellar biogenesis. These, now nuclear-encoded, genes that control organelle biogenesis and function are transcribed in the nucleus, their mRNAs are translated on cytoplasmic ribosomes and the proteins are then imported into the appropriate organelle.

The acquisition of nuclear-encoded proteins from the cytoplasm required the development of sophisticated protein import machineries, most notably the TOM/TIM and TOC/TIC protein import pathways of the mitochondrion and chloroplast, respectively (Neupert 1997; Soll and Schleiff 2004). Other protein import pathways also exist, such as via the secretory pathway (Villarejo et al. 2005), but these are little understood (Millar et al. 2006; Li and Chiu 2010). The establishment of protein import mechanisms may have been the limiting step in the transition from an endosymbiont to an organelle (Cavalier-Smith and Lee 1985). Once protein import was established, proteins had a route back to the organelle, thereby enabling the transfer of organelle genes to the nucleus, cementing the genetic interdependence of the organelle and host cell. Not all genes that have relocated to the nucleus encode proteins predicted to be re-imported into the organelles (Martin et al. 2002), suggesting translocation of pathways to other compartments, protein/enzyme shuffling between pathways or the acquisition of novel non-organelle related function. Therefore, the relocation of genes from the cytoplasmic organelles to the nucleus has been a major contributor to the complexity of nuclear genomes and has given rise to many genes of novel function. Shorter stretches of organelle DNA, rather than whole genes, have also contributed to the complexity of nuclear genomes and these are observable as exonic sequences in nuclear contexts that otherwise appear unrelated to mitochondrial or plastid DNA (Noutsos et al. 2007) and as putative modifiers of gene expression (Knoop and Brennicke 1991).

With the current availability of the nucleotide sequence of well over 2,400 mitochondrial genomes and over 200 plastid genomes (NCBI 2011), it is clear that there is considerable diversity in the size of organelle genomes and the number of proteins that they encode. Animal mitochondrial genomes are relatively conserved at around 16 kb in length but much more diversity is seen in plants whose mitochondrial genomes range from 13 kb in *Polytomella capuana* [chlorophyta] (Smith and Lee 2008) to 983 kb in the seed plant species *Cucurbita pepo* (Alverson et al. 2010). The largest chloroplast genome currently known is that of *Floydiella terrestris*, a chlorophycean alga whose plastome has a length of 521 kb (Brouard et al. 2010). Whilst this is the largest sequenced to date, it encodes only 69 conserved proteins, whereas some red algae have smaller genomes that encode over 200 proteins (Reith and Munholland 1995).

The most reduced organelle genomes are found in organisms that have lost the major organelle biosynthetic pathways of oxidative phosphorylation (in the case of mitochondria) and photosynthesis (in the case of plastids). Hydrogenosomes – organelles that produce molecular hydrogen and ATP in anaerobic organisms – are highly reduced mitochondria found in diverse eukaryotes (Boxma et al. 2005). In most cases they appear to have lost

their entire genome (van der Giezen et al. 1997; Clemens and Johnson 2000). Similarly, the smallest plastid genomes are found in lineages that have lost the ability to photosynthesise. The chloroplast genome of the parasitic underground orchid, *Rhizanthella gardneri* is 59 kb in size and codes for only 20 proteins (Delannoy et al. 2011). The non-photosynthetic apicoplasts – vestigial plastids of Apicomplexan parasites – have even smaller genomes which, in the malaria parasite *Plasmodium falciparum*, is around 34 kb in size (Wilson et al. 1996). Though it has lost its photosynthetic capacity the apicoplast and its genome appear to be indispensable, offering an interesting target for antimalarial therapy (Lim et al. 2010).

The evolutionary mechanisms and selection pressures that have driven organelle genome relocation to the nucleus are poorly understood but the presence of nuclear DNA sequences, within nearly all eukaryotic genomes, that are very similar to extant cytoplasmic organellar DNA, is clearly significant. These insertions of organelle DNA, which are the major focus of this chapter, are referred to as *numts* (nuclear integrants of mitochondrial DNA) and *nupts* (nuclear integrants of plastid DNA) or collectively as *norgs* (nuclear integrants of organellar DNA). Hazkani-Covo et al. (2010) recently instituted definitive pronunciation for these mtDNA and ptDNA integrants as "new-mights" and "new-peats", respectively.

B. Recent Gene Transfer Events

The number of genes found in plastid and mitochondrial genomes varies between species, but in all cases there are relatively few compared with the genomes of free-living prokaryotes. Therefore, it is thought that the majority of endosymbiotic gene transfer occurred early in the evolutionary history of the organelles (Timmis et al. 2004). In some lineages, including all animals where the set of mitochondrial genes is almost invariant, functional gene transfer appears to have ceased completely. In the few cases where genes are missing this is probably due to complete loss rather than transfer to the nucleus (Gissi et al. 2008). In plants, however, there is evidence of a flurry of recent functional gene transfer. Adams et al. (2000) reported 26 independent relocation events of *rps10* from the mitochondrial genomes to the nucleus amongst 277 angiosperms examined. Molecular characterisation of a number of the nuclear *rps10* genes indicated that each loss from the mitochondrial genome was likely to represent an independent transfer to the nucleus. A comparable study (Millen et al. 2001) looked at the loss of *infA* (encoding translation initiation factor I) from the chloroplast genome and discovered 24 cases of functional relocation among over 300 angiosperms. Again, molecular characterisation of nuclear *infA* genes suggested that each loss from the chloroplast genome was due to an independent transfer to the nucleus. Other elegant analyses have uncovered the complexity with which functional gene relocation has been achieved (Cusack and Wolfe 2007). A number of other genes, mainly encoding ribosomal proteins, have been lost from mitochondrial or chloroplast genomes and transferred to the nucleus in angiosperms (Rousseau-Gueutin et al. 2011), leading to considerable diversity in plant cytoplasmic organelle gene content.

In a few instances, mitochondrial genes have been replaced by nuclear genes of chloroplast origin. This is the case for *rps13* in rosids (Adams et al. 2002) and for *rpl10* in Brassicaceae and monocots (Kubo and Arimura 2010). In both cases, a nuclear gene of chloroplast origin was duplicated and one of the copies diverged so that its product is imported into mitochondria. Similarly, replacement of a chloroplast gene by a nuclear gene of mitochondrial origin has been observed. The two examples known so far are *rpl21* in *Arabidopsis thaliana* (Gallois et al. 2001) and *rps16* in *Medicago truncatula* and *Populus alba* (Ueda et al. 2008). The RPS16 protein is encoded by a single nuclear gene which is targeted to both the mitochondria and the chloroplasts. Dual targeting of ribosomal protein S16 has also been observed in

species which still retain *rps16* within the chloroplast genome. These cases may represent an intermediate step in replacing the function of an organelle gene with that of a nuclear gene (Ueda et al. 2008).

The reasons that functional gene transfer has apparently been reawakened in the angiosperms are far from clear. Given the length of time available for gene relocation prior to the invention of the seed plants, it is astonishing that *infA*, for example, was not transferred earlier, given the ease with which the event has occurred since the advent of this particular taxonomic group. It may be that the selection pressures on *infA* in other taxa were unamenable to nuclear location and this situation was reversed in angiosperms. Perhaps a change in the constraints imposed by various developmental pathways reopened means of gene relocation that were available early in evolution but were lost in the interim. The development of the gametophyte in angiosperms is one possible place to look for such modifications. There may also be other forces that contribute to the reawakening of functional gene transfer in angiosperms. For example, novel mechanisms are required to explain the hypermutation observed in *ycf4* in *Lathyrus sativus* (Magee et al. 2010). The observed hypermutation also spreads into the adjacent genes *accD* and *psaI* and their intergenic region and is of great interest because the plastid genes that show hypermutation have a recent history of relocation to the nucleus (Magee et al. 2010). It is as though sequence decay of the plastid genes has enhanced gene relocation to the nucleus.

Of course any essential gene must be functionally duplicated before either one or the other copy can be lost. Thus, a functional nuclear copy of *ycf4* is expected in *L. sativus*, but extensive efforts were unable to identify it (Magee et al. 2010). A mechanism that could cause hypermutation in specific chloroplast genes has not yet been suggested.

C. Why Relocate?

What are the possible advantages of organelle genes being located in the nucleus? The highly energetic compartments where photosynthesis and oxidative phosphorylation are carried out are clearly not the ideal environments in which to maintain genetic integrity. Proposed explanations include the high rate of oxidative stress-induced mutation within organelles (Allen and Raven 1996), genome streamlining (Selosse et al. 2001), more frequent fixation of beneficial mutations (Blanchard and Lynch 2000), avoidance of Müller's ratchet (the accumulation of mutation in asexually reproducing organelles) through the benefits of sexual recombination for elimination of deleterious mutation in nuclear genes (Lynch 1996; Martin and Herrmann 1998) and the advantages of allelic variation and meiotic recombination. These suggestions, however, seem not to apply to plant organelles which have a much lower rate of accumulation of mutations (Wolfe et al. 1987), have larger organelle genomes with more non-coding DNA (Timmis et al. 2004), and where more gene loss is observed in taxa that reproduce asexually or by self-fertilisation (Brandvain et al. 2007).

It is likely that a key factor is the unidirectional nature of transfer of genes to the nucleus. This is promoted by the high frequency translocation of gene rich organelle DNA into the nucleus and the relatively rare, or entirely absent, transfer of DNA encoding complete genes from the nucleus to the organelles. If transfer leads to two functional copies, one copy may then be lost. If the nuclear copy is lost the chloroplast gene is able to transfer again at a later stage, but if the organelle copy is lost then the nucleus becomes the permanent location of the gene establishing a 'gene-transfer ratchet' (Doolittle 1998). As long as there remains a polarity in the direction of DNA transfer, then relocation to the nucleus would be an inevitable consequence, even in the absence of a selective advantage (Berg and Kurland 2000).

From this background of selectively neutral gene transfer, the various mutational and/or selective pressures described above may contribute to the likelihood of gene transfer by altering the respective likelihoods of organelle or nuclear gene inactivation. These

pressures may have been considerably different early in evolution when the majority of transfer is likely to have occurred. The low rate of mutation in extant plant organelle genomes (Wolfe et al. 1987), presumably due to the establishment of plant-specific DNA repair and/or recombination pathways (Marechal and Brisson 2010), together with efficient gene conversion mediated by polyhaploidy (Khakhlova and Bock 2006), may well have led to a slowing in the rate of gene transfer. This could explain the differences in genome size and gene content between plant and animal mitochondrial genomes. If this is the case, it would suggest that the accumulation of mutations rather than the energetic and replicative advantage of a small organelle genome drive gene transfer to the nucleus.

D. Why Retain an Organellar Genome?

Thousands of genes have relocated from the plastids and mitochondria to the nucleus, so why do any remain given the energy outlay in maintaining all of the transcriptional and translational machinery required for the retention of alternative genetic systems. The hydrophobicity hypothesis suggests that highly hydrophobic proteins are hard to export from the cytosol to the organelles and that this precludes relocation of these genes to the nucleus (Vonheijne 1986; Daley and Whelan 2005). Counter to this, however, the chloroplast-encoded hydrophobic protein D1 can be imported from the cytosol to the chloroplast when experimentally equipped with a transit peptide and expressed from a nuclear gene (Cheung et al. 1988). In addition, several other hydrophobic organellar proteins (such as the ADP-ATP carriers) are known to be nuclear encoded (Allen 2003). The Co-location for Redox Regulation or CoRR hypothesis (Allen 2003) maintains that there is a key set of genes whose expression must be directly controlled by the redox state of their gene product or interacting electron carriers. This requires separate (organellar rather than nuclear) gene expression, as redox state is likely to vary between the many organelles within a single cell. Recently a sensor kinase has been identified that links the redox state of an electron carrier connecting the two photosystems, with chloroplast gene expression (Puthiyaveetil et al. 2008). Neither of these hypotheses, however, appear to explain the retention of genomes in non-photosynthetic plastids such as those found in parasitic plants or the apicoplasts of the Apicomplexa and several other hypotheses have been proposed to explain the situation in these cases (Barbrook et al. 2006). The 'essential tRNA' hypothesis was proposed based on the observation that the tRNA encoded by the plastid gene *trnE* is involved in tetrapyrrole biosynthesis and so may be essential even in the absence of organellar protein biosynthesis (Barbrook et al. 2006). The 'limited transfer window' hypothesis posits that organisms containing a single organelle per cell will have little opportunity for gene transfer as organelle breakdown, which may be necessary for the release of DNA, will be lethal (Barbrook et al. 2006). It may be that no single hypothesis is able to explain adequately the retention of organelle genomes in all cases and different combinations of factors may be responsible in different taxonomic groups.

III. Promiscuous DNA: Ongoing Organelle DNA Transfer to the Nucleus

A. Organelle Sequences in Nuclear Genomes

A prerequisite for the functional relocation of plastid and mitochondrial genes to the nucleus is a nucleic acid transfer mechanism. The first indications that transfer of organelle DNA to the nucleus continues today came to light about 30 years ago with the identification of nuclear sequences that are closely similar or identical to extant organelle DNA (van den Boogaart et al. 1982; Timmis and Scott 1983). The relatively recent transfer of these sequences to the nucleus was subsequently inferred from their high sequence identity (i.e. >99%) to existing organelle genes (Fig. 9.1).

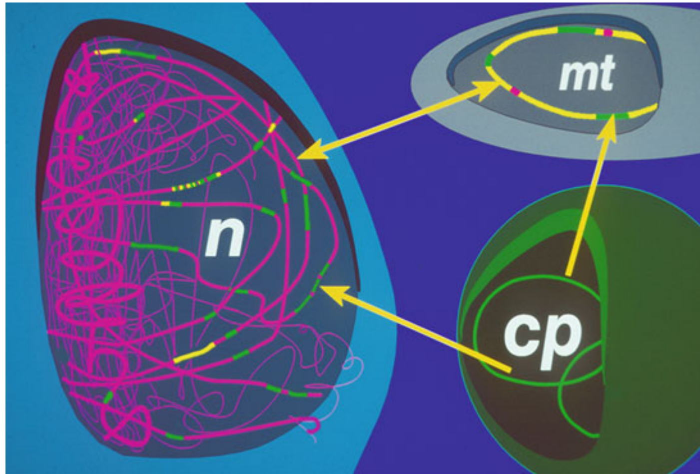

Fig. 9.1. Promiscuous DNA in the three genetic compartments of a plant cell. *n* nucleus, *mt* mitochondrion, *cp* chloroplast.

Whole genome sequencing has since revealed extensive tracts of chloroplast and/or mitochondrial DNA in the nuclear genomes of almost all eukaryotes studied (Timmis et al. 2004; Hazkani-Covo et al. 2010).

The arrangement of these sequences has been studied in detail in Arabidopsis and rice and has been found to be quite varied (Richly and Leister 2004a, b; Noutsos et al. 2005). A large proportion of the total *norg* content is found in a relatively small number of large *norgs* that can be tens or hundreds of kb in length. The remainder is found in a large number of smaller *norgs* scattered throughout the genome (Richly and Leister 2004b). Of the large *norg* loci, some are continuous sequences of chloroplast or mitochondrial origin and are clearly the result of the insertion of a single molecule, while others contain multiple fragments of DNA from diverse parts of the chloroplast or mitochondrial genome or both (Noutsos et al. 2005). Some loci of the latter type, probably represent insertions of a single contiguous fragment of organelle DNA that has since undergone deletion and/or rearrangement (Matsuo et al. 2005). However, these loci may also be formed by the insertion of multiple fragments of DNA from diverse regions of the chloroplast genome in a single event (Lloyd and Timmis 2011), or by multiple sequential insertions at a single locus (Noutsos et al. 2005). Several other *norg* loci are highly complex mosaics containing up to 80 disparate ~50–100 bp segments of the chloroplast and mitochondrial genome arranged end to end (Noutsos et al. 2005). How these loci arise is yet to be explained satisfactorily, but similar mosaics have been observed that are comprised of many short stretches of transposable element sequence (David Adelson 2011, personal communication).

Large *norg* insertions have also been observed in other species. Recently, in situ hybridization in the maize inbred line B73 identified a *nupt* that includes almost the entire 140 kb chloroplast genome on chromosome 5 (Roark et al. 2010) and a *numt* containing the majority of the 570 kb mitochondrial genome on chromosome 9 (Lough et al. 2008). These studies also showed that *numts* and *nupts* varied greatly among different inbred maize lines indicating that there have been frequent recent insertions of organelle DNA into maize nuclear genomes. Current investigations such as the 1,000 genomes projects in humans and Arabidopsis should contribute greatly to understanding the intra-species variation of *norgs* and perhaps reveal potential evolutionary ramifications.

The precise contribution of *nupt* and *numt* sequences to the nuclear genome is hard to

determine. Based on current genome assemblies it is estimated that *nupt*s and *numt*s each generally make up about 0.1–0.2% of the nuclear genome in flowering plants and significantly less in algae and the moss *Physcomitrella patens* (Table 9.1). This, however, may be the 'tip of the iceberg' as whole genome assemblies often underestimate the contribution of organelle-derived sequences to nuclear genomes. This is in large part an artefact of the elimination of seemingly 'contaminating' organelle DNA sequences – a process which must also often exclude *norgs*. An example is the honeybee genome which was initially thought to have little or no mitochondrial DNA within the nucleus (Leister 2005) but has since been found, using a different assembly, to have one of the most extensive *numt* complements (Behura 2007; Hazkani-Covo et al. 2010).

Another problem lies in the assembly of regions of the genome that contain large duplications. Chromosome 2 in Arabidopsis was initially reported to contain a 270 kb *numt* (Lin et al. 1999) but Stupar et al. (2001) later showed that this *numt* was in fact ~620 kb in length and contained several large internal duplications. The authors were only able to determine the *numt* size using fibre-FISH and showed that contig assembly using BACs tended to minimise clone length, missing large duplications. Despite this finding, this region is still only 270 kb in length in the current chromosome 2 assembly (Build 9.1, 14th Oct 2009) and recent studies (Richly and Leister 2004a; Hazkani-Covo et al. 2010) have therefore greatly underestimated total *numt* size in Arabidopsis. This problem presumably holds for *nupts* as well and will be compounded in genomes shotgun sequenced using high-throughput short-read platforms.

B. Evidence of Frequent Plastid and Mitochondrial DNA Transfer to the Nucleus

In some species it has been possible to determine experimentally the frequency with which organellar DNA moves into the nucleus. This was initially investigated in yeast by measuring the transfer of a mitochondrial plasmid to the nucleus (Thorsness and Fox 1990) which was found to occur at high frequency (~2×10^{-5} per cell per generation). Although the plasmid DNA in these first experiments was not incorporated into the nuclear chromosomes, subsequent work, also in yeast, observed integration of mitochondrial DNA at sites of nuclear double strand break repair (Ricchetti et al. 1999). With the development of chloroplast transformation in tobacco (Svab et al. 1990), similar studies became possible in higher plants. In the first of these studies (Huang et al. 2003) a selectable marker gene (*neo*), equipped for exclusive nuclear expression, was introduced into the chloroplast genome of tobacco. Transplastomic pollen was used to fertilise female wild-type plants and the resultant progeny was screened for kanamycin resistance (*neo* expression). In a large screen of 250,000 seedlings, 1 in 16,000 pollen grains were inferred to carry a copy of *neo* transferred from chloroplast DNA to the nucleus in the germline of the transplastomic male parent. A similar study measured the rate of transfer in somatic cells (Stegemann et al. 2003) and transfer was shown to occur once in approximately 5,000,000 cells. Although still relatively frequent, this was substantially less common than the transfer observed in the male germline and prompted the suggestion that degradation of the chloroplast during pollen development (associated with uniparental inheritance) may provide more opportunity for nuclear DNA transfer by liberating fragments of chloroplast DNA. This hypothesis was supported by a third study that measured the rate of gene transfer in both the female and male germline (Sheppard et al. 2008). An even greater frequency of gene transfer through the male germline was reported (1 in 11,000 pollen) which far exceeded transfer in the female germline where a single transfer event was observed in a screen of over 270,000 ovules (Sheppard et al. 2008).

In each of these screens the chloroplast gene not only transfers to the nucleus but also must integrate into the nuclear chromosomes.

Table 9.1. Current estimates of *numt* and *nupt* content in plant nuclear genomes

	Order	Family		Subfamily	Genus/species	*numt* content Kbp (%)	*nupt* content Kbp (%)
Chlorophyta	Volvocales	Chlamydomonadaceae			*Chlamydomonas reinhardtii*	2.8[a] (0.003)	2.4[b] (0.002)
Bryophyta	Funariales	Funariaceae			*Physcomitrella patens*	76[a] (0.02)	–
Liliopsida	Poales	Poaceae	BEP clade	Ehrhartoideae	*Oryza sativa ssp. indica*	409[c], 823[a] (0.16, 0.24)	804[b], 1176[d] (0.17, 0.25)
				Pooideae	*Brachypodium distachyon*	488[e] (0.14)	275[e] (0.08)
			PACCAD clade		*Sorghum bicolor*	539[a] (0.07)	–
Eudicots	Brassicales	Brassicaceae			*Arabidopsis thaliana*	198[c], 305[a] (0.16, 0.24)	35[b], 21[d] (0.03, 0.02)
		Caricaceae			*Carica papaya*	858[f] (0.23)	785[f] (0.21)
	Vitales	Vitaceae			*Vitis vinifera*	–	570[d] (0.12)
	Malpighiales	Salicaceae			*Populus trichocarpa*	–	679[d] (0.12)

For each species the *numt* and *nupt* content is given as a total length and as a percentage of the nuclear genome, – indicates not determined

[a]Hazkani-Covo et al. (2010)
[b]Richly and Leister (2004b)
[c]Richly and Leister (2004a)
[d]Arthofer et al. (2010)
[e]Vogel et al. (2010)
[f]Ming et al. (2008)

To investigate the steps in this process, Sheppard et al. (2008) introduced a *GUS* reporter gene (again designed exclusively for nuclear expression) into the chloroplast genome and leaves of the transplastomic plant were stained for GUS activity to detect cells in which the gene had transferred to the nucleus. In this instance blue staining cells represented transient expression from the nucleus/cytoplasm as well as transfer followed by stable integration into a transcriptionally active region of the nuclear genome. Interestingly, total transfer (transient and stable) was found to be 25–270-fold higher than the stable somatic transfer of *neo* detected by Stegemann et al. (2003) suggesting that most blue spots resulted from transient expression. The lack of any large and the rarity of small groups (mitotic lineages) of GUS-expressing cells suggested that few stable integrations occurred early in leaf development.

C. Evolutionary Fate of Nuclear Located Cytoplasmic Organelle DNA (norgs)

Given the constant deluge of organellar DNA entering the nuclear genome in recent evolutionary time (see also Sect. IV below), it is expected that a counterbalancing eradication of these sequences occurs to prevent continual genome expansion. This was first alluded to with the observation that, for *nupts* over 500 bp in length, there is an inverse relationship between their age (based on sequence identity to the chloroplast genome) and their size (Richly and Leister 2004b). This finding has subsequently been found to hold true for *norgs* in *Brachypodium distachyon* (Vogel et al. 2010) and *Carica papaya* (Ming et al. 2008) and suggests that insertion of large *nupts* is followed by fragmentation and deletion. Direct experimental observation of frequent deletion of about 50% of newly transferred chloroplast sequences has demonstrated the extreme instability of plastid DNA integrants in the tobacco nucleus (Sheppard and Timmis 2009). So far, it has not been possible to determine how much of the integrant is lost by recovering the sequence that remains. This *nupt* deletion occurred within 1–2 generations of insertion and it may be that more integrant loci would show instability over longer, but still evolutionarily relevant, timescales.

The deletion of organelle DNA is most unlikely to be an exact excision and partial deletion would lead to novel arrangements of organelle and nuclear DNA. The deletion may also be accompanied by other rearrangements including inversions and new insertions of organellar DNA and transposable elements (Guo et al. 2008). Richly and Leister (2004b) observed 'tight' and 'loose' clusters of organellar sequence in nuclear genomes of rice and Arabidopsis which they suggest represent progressive steps of degradation and rearrangement of large initial insertions. Deletions and other rearrangements may be part of the mutational processes that, in rare instances, lead to the activation of newly transferred genes (Bock and Timmis 2008; Lloyd and Timmis 2011).

Base substitution and indels appear to play a significant role in the evolution of *norgs*. In plants, a significant bias in C→T and G→A mutations has been observed in large recent integrants of organelle DNA (Huang et al. 2005). This mutational bias is consistent with spontaneous deamination of 5-methylcytosine inferring that these *norgs* are methylated. Studies linking the stability of *norg* sequences with methylation and chromatin structure have not yet been reported.

IV. Mechanisms of Gene Transfer to the Nucleus

A. Relocation of Genetic Material

The first step in transfer of a gene to the nucleus is the relocation of genetic material from the organelle. In general, the availability of cytoplasmic organelle nucleic acid fragments for transfection of the nucleus is likely to be made possible through loss of integrity of the organelle membrane, either through various physiological stressors or programmed degradation during development. Various environmental stress factors

and developmental stages are known to trigger programmed organelle degradation (Kundu and Thompson 2005; Stettler et al. 2009; Wada et al. 2009) and these may lead to increased ingress of organelle DNA to the nucleus. Recently, cold stress (Ruf et al. 2010) has been shown to increase the rate at which a chloroplast gene relocates to the nucleus in tobacco.

Uniparental inheritance is also implicated in leading to the presence of organelle nucleic acids in the cytoplasm that may find their way into the nucleus. In many sexually reproducing eukaryotes, only one sex contributes cytoplasmic genes to the zygote. How this uni-parental inheritance is achieved varies amongst species, but in general, the cytoplasmic organelles are degraded and/or excluded from one of the gametes or sex-specific loss of organelles occurs after fertilisation (Birky 2001). In tobacco, chloroplast genes are maternally inherited and this parallels the observation that DNA transfers from the chloroplast to the nucleus far more frequently in the male germ line than that of the female (Sheppard et al. 2008). This difference has been suggested to be due to the release of chloroplast DNA into the cytoplasm during chloroplast degradation/exclusion in the developing male gametophyte. For unicellular organisms that have only a single organelle per cell, DNA transfer is likely to be very limited, as degradation of the single but essential organelle will lead to cell death (Barbrook et al. 2006). The *Chlamydomonas reinhardtii* nuclear genome has a low *norg* content (Table 9.1) and large screens failed to detect transfer of a chloroplast gene to the nucleus (Lister et al. 2003). We considered whether transfer could occur during the diploid phase when the chloroplast is briefly duplicated but screens of many millions of germinated zygotes from reciprocal crosses between transplastomic and wild type *C. reinhardtii* failed to identify DNA transfer (unpublished results).

Further understanding of how various stress factors and modes of organelle inheritance affect chloroplast-to-nucleus DNA transfer should be an interesting area of future research in the context of the wide climatic and ecological ranges that plants have colonised. The increasing wealth of genome sequence data will pave the way for analysis of *norgs* in different ecotypes and it will be interesting to see if any relationship exists between *norg* content and environmental conditions or geographical distribution. Further understanding of how stress and organelle integrity affect endosymbiotic DNA transfer will also be of biotechnological significance in view of the desire to minimise transfer of chloroplast transgenes to the nucleus.

B. Is There an RNA or DNA Intermediate?

It is generally held that the majority of organelle nucleic acid transfer to the nucleus occurs via DNA (Timmis et al. 2004; Kleine et al. 2009), although this still remains to be experimentally proven. Some studies of the transfer of plant mitochondrial genes to the nucleus showed that nuclear copies resembled spliced, edited mRNAs and led to the suggestion that transfer was via a reverse transcribed RNA intermediate (Nugent and Palmer 1991; Grohmann et al. 1992; Adams et al. 2000). There are, however, alternative explanations, such as the elimination of introns and editing sites in mitochondrial genomes through cDNA recombination (Henze and Martin 2001), that account for these observations without involving RNA-mediated transfer. Some further evidence also suggests DNA-mediated transfer: non-coding regions of the chloroplast genome are found in nuclear genomes as abundantly as highly transcribed genic regions of the organellar genomes (Matsuo et al. 2005) and some very large nuclear insertions of organellar sequence (>100 kb) have been found (Stupar et al. 2001; The Rice Chromosome 10 Sequencing Consortium 2003), suggesting direct DNA transfer. Direct experimental evidence of RNA mediated transfer is lacking, as is the determination of the relative contributions of RNA and/or DNA mediated transfer. At least one study designed to observe transfer via an RNA intermediate failed to detect any such transfer (Sheppard et al. 2011).

C. Integration into Nuclear Chromosomes

Once the organelle nucleic acid has entered the nucleus it must be integrated into nuclear chromosomes and be included in the gametes of sexually reproducing organisms if it is to make a contribution to the evolution of the nuclear genome. It is thought that most integration of organellar DNA occurs via non-homologous end joining at sites of double strand break (DSB) repair (Kleine et al. 2009) and this has been shown to occur in yeast (Ricchetti et al. 1999). DSBs were induced in the yeast nuclear genome through expression of the rare cutting endonuclease I-SceI and insertion of mitochondrial DNA was observed in a proportion of repair events. Interestingly, in some repair events, DNA from two disparate regions of the mitochondrial genome was inserted at a single location. Similar capture of non-mitochondrial DNA has also been observed at sites of DSB repair in yeast (Haviv-Chesner et al. 2007) as well as in plant and mammalian systems (Salomon and Puchta 1998; Lin and Waldman 2001). In these studies DSBs were induced by transiently introducing a plasmid, or T-DNA, encoding a rare cutting endonuclease. This rare-cutting endonuclease cuts at a specific restriction site introduced into the nuclear genome and repair events were then analysed by PCR. Insertion of the T-DNA or plasmid DNA was often observed, as were insertions of nuclear repetitive elements such as retro-transposons and micro-satellites. While insertion of organellar DNA has so far only experimentally been observed in yeast, the fact that extra-chromosomal DNA can be captured at sites of DSB repair in plants and animals suggests this process applies more widely.

The insertion of *norgs* has been also investigated in several bioinformatic analyses and these suggest more than one pathway for integration (Leister 2005). Some integrants show a very simple arrangement likely originating when a single organellar DNA fragment inserted at a single location. Others are much more complex and are the result of multiple fragments being inserted in a single event or multiple insertions at a single location. Organelle sequences may also insert into areas of the genome that already contain *norgs* or other repetitive sequences which also adds to the complexity of these loci. There is some evidence that organellar DNA integrates more frequently into intergenic regions in rice and Arabidopsis (Richly and Leister 2004b), in particular those containing mobile elements (Mishmar et al. 2004). Large *nupts* have also been shown to preferentially locate to pericentromeric regions in rice (Matsuo et al. 2005) which are known to be DSB hotspots (Blitzblau et al. 2007) and to contain a high density of transposable elements (Hall et al. 2006). A recent study has also linked *numt* insertion sites in yeast to origins of replication (Lenglez et al. 2010), which led the authors to suggest that these sites may be prone to DSBs resulting in high levels of insertion. These findings point toward DSB repair, possibly at sites of transposon excision (Leister 2005), as a pathway for the nuclear insertion of organellar sequences. The presence of such a DNA repair/integration mechanism would contribute significantly to the complex arrangement of organellar sequences integrating into the nuclear genome. This would be important from an evolutionary perspective as it would lead to the creation of novel sequence arrangements which, in some instances, may result in nuclear activation of the transferred organelle genes.

The cross-over in the insertion pathway and chromosomal location of *norgs* and repetitive DNA elements shows that these sequences can be dealt with in very similar ways by the nuclear DNA repair/maintenance machinery. To date, studies have focussed exclusively on either organellar DNA or transposons and other repetitive sequence. There may be significant advantage to both fields if a more unified approach is taken to investigating these areas.

Although bioinformatic analyses of *norgs* due to evolutionary transfer have added considerably to our understanding of these sequences, they are limited in that a *norg* sequence cannot usually be compared with that of the nuclear sequence prior to insertion. This makes identification of micro-homology and other indicators of NHEJ difficult to assess. Also it is impossible to

determine how much of the observed complexity of *norg* sequences is due to the primary insertion event and how much is due to subsequent fragmentation or insertion at this locus. Some partial characterisation of experimentally transferred *norgs* has been undertaken and suggests that micro-homology is involved in the insertion of these sequences (Huang et al. 2004). A fuller understanding will come with complete characterisation of de novo *norgs* and comparison with their pre-insertion sequences. This remains an important future step but is a challenging task. The reason for this is that the new integrants are often very large (dozens of kb) and the nuclear genome already contains *norg* sequences in high copy number. These, together with the superabundant cellular plastid DNA, preclude the design of primers in organellar sequences. Therefore, techniques generally used for determining the junction sequence in transgenic lines, such as genome walking, TAIL-PCR or inverse PCR, cannot be used in determining the pre-insertion site unless one of the marker genes is very close to the integrant boundary (Sheppard and Timmis 2009).

V. Activation of a Newly Transferred Organelle Gene

A. Examples of Organellar Gene Activation in the Nucleus

Only in a very few instances will transfer of organellar DNA to the nucleus lead to the functional relocation of an organelle gene. In most cases, organelle sequences transferred to the nucleus have the same fate as other non-coding DNA – freely accumulating mutation and degrading over time. The low mutation rate in plant organelle genomes means that the extant organelle genomes provide a historic reference for the sequence at the time of insertion, from which it is possible to derive many insights into the various ways in which *norg* sequences evolve that would otherwise not be possible. In a few rare cases, these sequence rearrangements and changes in base composition lead to activation of newly transferred genes. Activation, in the majority of cases, must be a multistep process and requires the acquisition of a nuclear promoter, a polyadenylation signal and, if the protein is to be targeted back to the organelle, a transit peptide or an alternative mechanism for protein targeting. Several bioinformatic studies have highlighted various means by which organellar genes have recently become activated in nuclear genomes. One such study investigated transfer of the maize gene encoding the mitochondrial protein RPS14 to the nucleus (Figueroa et al. 1999). The gene had inserted into an intron of the iron-sulphur protein subunit of succinate dehydrogenase (*sdh2*) gene and was expressed by differential splicing of the mRNA with both proteins using the SDH2 transit peptide for targeting to the mitochondria. In a similar case, the chloroplast *rpl32* gene was transferred to the nucleus in an ancestor of mangrove and poplar (Cusack and Wolfe 2007) where it inserted into an intron of the gene encoding the chloroplast superoxide dismutase (*SODcp*) to form the chimeric *SODcp-rpl32* gene. In mangrove, the SODcp protein and a SODcp amino terminus/RPL32 fusion protein are expressed from the single promoter through differential splicing. Both proteins are then targeted to the plastid using the SODcp transit peptide. In poplar, evolutionary experimentation has taken the process one step further with the duplication and subfunctionalization of the *SODcp-rpl32* gene. One copy has lost the RPL32 coding sequence and now solely encodes SODcp, the other now exclusively expresses the SODcp amino terminus/RPL32 fusion protein. There are numerous other examples of genes that have recently transferred to the nucleus in angiosperms, many of which have also hijacked transit peptides from existing nuclear encoded organellar proteins (Liu et al. 2009).

B. Experimental Attempts to Detect Activation of a Chloroplast Gene After Transfer to the Nucleus

Experimental attempts have been made to reconstruct functional gene transfer to gain a

better understanding of the diverse processes involved and the frequency with which newly transferred prokaryotic genes become activated in the nucleus. Stegemann and Bock (2006) showed functional activation of a chloroplast marker gene *aadA* that had been recently transferred to the nucleus in tobacco. In each case, *aadA* was activated, through intervening deletions, by the nearby strong CaMV 35S nuclear promoter that was integral to their experimental cassette and present in the same transcriptional polarity. In no case was activation achieved by acquisition of a native nuclear promoter and so the frequency of a 'natural' gene transfer event remains unclear.

Interestingly, they found that the *aadA* transcripts were polyadenylated despite the lack of any changes in the *psbA* 3' UTR found downstream of the *aadA* open reading frame. Examination of the *psbA* terminator revealed a sequence that matched the rather flexible AT-rich plant polyadenylation consensus sequence and this was the in vivo site of *aadA* polyadenylation. This led the authors to suggest that the AT-rich nature of plastid non-coding sequences may provide many fortuitous polyadenylation sites – greatly aiding the process of functional gene transfer. This could possibly be extended to other AT-rich regulatory motifs such as a TATA box. Indeed, the tobacco chloroplast *psbA* promoter has been shown to have weak nuclear activity that is dependent upon TATA and CAAT boxes present fortuitously (Cornelissen and Vandewiele 1989), but cryptic nuclear activity of any other chloroplast promoters remains unknown.

VI. Plastid DNA in Higher Plant Mitochondria

Most angiosperms also have large chunks of plastid DNA in their mitochondrial genome. The first report of the existence of DNA sequences that have been transferred from the chloroplast to the mitochondrion in higher plants was published three decades ago by Stern and Lonsdale (1982). In this study, they showed that a 12 kb DNA sequence present in the maize mitochondrial genome was essentially identical to part of the inverted repeat of the maize chloroplast genome. These chloroplast-derived sequences or mitochondrial plastid DNA sequences were later designated "*mtpts*" (mighty-peats) (Wang et al. 2007).

With the current availability of the nucleotide sequence of 48 plant mitochondrial genomes (NCBI 2010), it is now clear that mitochondrial genomes of seed plants are rich in sequences derived from the chloroplast. These *mtpts* constitute 1–11% of the mitochondrial genome in different species of seed plants (Table 9.2) and the transfer seems to involve random sections of the chloroplast genome (Wang et al. 2007). While no sequences homologous to chloroplast DNA have been detected in the mitochondrial genomes of bryophytes or algae (Oda et al. 1992; Turmel et al. 2003; Terasawa et al. 2007; Li et al. 2009), an insertion of chloroplast DNA has been observed in the mitochondrial genome of the lycophyte *Isoetes engelmannii* suggesting that plastid-to-mitochondrion DNA transfer phenomena began during or after the origin of vascular plants.

In vascular plants, the accumulation of *mtpts* is positively correlated with the increase of the mitochondrial genome size (Wang et al. 2007). However, the *mtpts* are not the only cause of this mitochondrial genome expansion; extra sequences also originate from the nucleus (e.g., fragments of nuclear transposable elements; Knoop et al. 1996), from other organisms (horizontal gene transfer; Goremykin et al. 2009) and also by the duplication or amplification of pre-existing sequences (Kitazaki and Kubo 2010). An interesting study (Allen et al. 2007) showed that the chloroplast DNA present in the mitochondrial genome does not only vary between species but may even vary among maize cytotypes, ranging from 16,929 bp in the maize B37 inbred line with male-sterile CMS-C cytoplasm to 29,470 bp in the male-fertile inbred line A188 (NA cytoplasm). From this study, it appears that plastid DNA can be gained and lost rapidly

Table 9.2. Mtpt content determined for mitochondrial genomes of seed plants

	Order	Family		Subfamily	Genus/species	Mitochondrial genome bp	Cp. sequence in the mitochondrion bp (%)	Length of Cp. fragments bp	Reference
Cycadophyta	Cycadales	Cycadaceae			*Cycas taitungensis*	414,903	18,113 (4.4)	–	Chaw et al. (2008)
Liliopsida	Poales	Poaceae	BEP clade	Ehrhartoideae	*Oriza sativa ssp. japonica*	490,520	22,593 (6.3)	32–6,653	Notsu et al. (2002)
				Pooideae	*Triticum aestivum*	452,528	13,455 (3)	27–4,239	Ogihara et al. (2005)
			PACCAD clade		*Zea mays ssp. mays (CMS-C)*	739,719	16,929 (2.3)	34–2,220	Allen et al. (2007)
					Zea mays ssp. mays (CMS-S)	557,162	20,780 (3.7)	34–3,726	Allen et al. (2007)
					Zea mays ssp. mays (CMS-T)	535,825	23,669 (4.4)	34–3,739	Allen et al. (2007)
					Zea mays ssp. mays (fertile cytotype: NA)	701,046	29,470 (4.2)	34–3,756	Allen et al. (2007)
					Zea mays ssp. mays (fertile cytotype: NB)	569,630	26,239 (4.6)	28–12,592	Allen et al. (2007), Clifton et al. (2004)
Eudicots	Brassicales	Brassicaceae			*Arabidopsis thaliana*	366,924	3,958 (1.1)	30–930	Unseld et al. (1997)
					Brassica napus	221,853	7,950 (3.6)	43–2,181	Handa (2003)
		Caricaceae			*Carica papaya*	476,890	–	106 2,495	Ming et al. (2008)
	Caryophyllales	Caryophyllaceae			*Silene latifolia*	253,413	2,462 (1)	43–588	Sloan et al. (2010)
		Chenopodiaceae			*Beta vulgaris*	368,799	(2.1)	25–3,366	Kubo et al. (2000)
	Cucurbitales	Cucurbitaceae			*Citrullus lanatus*	379,236	22,779 (6)	–	Alverson et al. (2010)
					Cucurbita pepo	982,833	113,347 (11)	92–18,534	Alverson et al. (2010)
	Solanales	Solanaceae	Nicotianoideae		*Nicotiana tabacum*	430,597	9,942 (2.5)	–	Sugiyama et al. (2005)
	Vitales	Vitaceae			*Vitis vinifera*	773,279	68,237 (8.8)	62–9,106	Goremykin et al. (2009)

For each species, the size of the mitochondrial genome, the total size of the chloroplast sequences present in the mitochondrion (given in bp and as a percentage of the mitochondrial genome size) and their size range are presented. – indicates not determined

from the mitochondrial genome. However, most of the variation in plastid DNA amount among the newly sequenced maize genomes is due to only 10 out of the 45 segments of plastid origin and only three major differences account for much of the variation in plastid DNA content.

In most cases, the predicted protein-coding *mtpts* present in the mitochondrial genome are assumed to be non-functional based on the presence of frameshift mutations and indels. However, for a minority of the *mtpt* sequences, mitochondrial functions have been discovered. For example, a tRNA sequence of plastid origin functions as a tRNA for mitochondrial translation (Kanno et al. 1997; Miyata et al. 1998). Indeed, a number of chloroplast-derived tRNA genes are transcribed and processed to mature tRNAs in mitochondria. However, the acquisition of function does not seem to be immediate, since at least one mitochondrial tRNA gene presenting 100% identity to the native plastid gene is not transcribed (Miyata et al. 1998). Amongst seed plant species, both the number and the type of mitochondrial tRNA genes replaced by a chloroplast-derived tRNA gene vary (Miyata et al. 1998). A second known impact of the *mtpt* sequences is as a source of promoters for mitochondrial genes, as demonstrated for the rice mitochondrial *nad9* gene (Nakazono et al. 1996). In this study, it was determined that the transcription of the *nad9* gene is initiated in a chloroplast-derived sequence that is located in a region upstream of the mitochondrial *nad9* gene.

In addition to the known positive impact of some chloroplast-derived sequences on mitochondrial gene function, it has recently been observed that sequences of chloroplast origin may be used in gene conversion events within the mitochondrial genome (Hao and Palmer 2009). This study reports that an internal segment (ranging from 14 to 78 bp in different species) of mitochondrial *atp1* (encoding the alpha subunit of ATP synthase) has been replaced with a plastid *atpA* sequence in a number of angiosperms belonging to diverse families. The plastid *atpA* sequence is found within a region of mitochondrially located chloroplast DNA and it seems that independent conversions occurred by intra-mitochondrial genome recombination, probably occurring well after the integration of the chloroplast *atpA* genes. The resulting "chimeric" genes – composed of mitochondrial and chloroplast sequences – might still be functional, since *atp1* has an intact open reading frame, is presumed to be a single copy gene in the mitochondrion and is not known to have been transferred to the nucleus in any angiosperm.

VII. Perspective

The constant integration of organellar DNA has had profound consequences in the evolution of eukaryote nuclear genomes (Timmis et al. 2004; Kleine et al. 2009). The ingress of DNA is followed by decay, deletion and rearrangement of these sequences, which leads to novel sequence combinations. In rare instances, these new sequences can lead to the functional relocation of organelle genes to the nucleus or the generation of genes with novel function. This process is of great evolutionary interest as it has been a major pathway for the generation of new genes in eukaryote nuclear genomes. It is also of great interest both to plant biotechnologists and the wider public in assessing the level of transgene containment provided by chloroplast transformation (Ruf et al. 2007).

Acknowledgements

We thank the Australian Research Council for financial support (Grants DP0667006 and DP0986973). This chapter is not intended to be an exhaustive review of the topic. Therefore we apologise for omitting citation and discussion of some important research contributions in the interest of brevity. We thank David Adelson for discussion.

References

Adams KL, Daley DO, Qiu YL, Whelan J, Palmer JD (2000) Repeated, recent and diverse transfers of a mitochondrial gene to the nucleus in flowering plants. Nature 408:354–357

Adams KL, Daley DO, Whelan J, Palmer JD, Mollier P, Hoffmann B, Debast C, Small I (2002) Genes for two mitochondrial ribosomal proteins in flowering plants are derived from their chloroplast or cytosolic counterparts. Plant Cell 14:931–943

Allen JF (2003) The function of genomes in bioenergetic organelles. Philos Trans R Soc Lond B Biol Sci 358:19–37

Allen JF, Raven JA (1996) Free-radical-induced mutation vs redox regulation: costs and benefits of genes in organelles. J Mol Evol 42:482–492

Allen JO, Fauron CM, Minx P, Roark L, Oddiraju S, Lin GN, Meyer L, Sun H, Kim K, Wang CY et al (2007) Comparisons among two fertile and three male-sterile mitochondrial genomes of maize. Genetics 177:1173–1192

Alverson AJ, Wei XX, Rice DW, Stern DB, Barry K, Palmer JD (2010) Insights into the evolution of mitochondrial genome size from complete sequences of *Citrullus lanatus* and *Cucurbita pepo* (Cucurbitaceae). Mol Biol Evol 27:1436–1448

Arthofer W, Schuler S, Steiner FM, Schlick-Steiner BC (2010) Chloroplast DNA-based studies in molecular ecology may be compromised by nuclear-encoded plastid sequence. Mol Ecol 19:3853–3856

Barbrook AC, Howe CJ, Purton S (2006) Why are plastid genomes retained in non-photosynthetic organisms? Trends Plant Sci 11:101–108

Behura SK (2007) Analysis of nuclear copies of mitochondrial sequences in honeybee (*Apis mellifera*) genome. Mol Biol Evol 24:1492–1505

Berg OG, Kurland CG (2000) Why mitochondrial genes are most often found in nuclei. Mol Biol Evol 17:951–961

Birky CW (2001) The inheritance of genes in mitochondria and chloroplasts: laws, mechanisms, and models. Annu Rev Genet 35:125–148

Blanchard JL, Lynch M (2000) Organellar genes: why do they end up in the nucleus? Trends Genet 16: 315–320

Blitzblau HG, Bell GW, Rodriguez J, Bell SP, Hochwagen A (2007) Mapping of meiotic single-stranded DNA reveals double-strand-break hotspots near centromeres and telomeres. Curr Biol 17: 2003–2012

Bock R, Timmis JN (2008) Reconstructing evolution: gene transfer from plastids to the nucleus. Bioessays 30:556–566

Boxma B, de Graaf RM, van der Staay GW, van Alen TA, Ricard G, Gabaldon T, van Hoek AH, Moon-van der Staay SY, Koopman WJ, van Hellemond JJ et al (2005) An anaerobic mitochondrion that produces hydrogen. Nature 434:74–79

Brandvain Y, Barker MS, Wade MJ (2007) Gene co-inheritance and gene transfer. Science 315:1685

Brouard JS, Otis C, Lemieux C, Turmel M (2010) The exceptionally large chloroplast genome of the green alga *Floydiella terrestris* illuminates the evolutionary history of the Chlorophyceae. Genome Biol Evol 2:240–256

Cavalier-Smith T, Lee JJ (1985) Protozoa as hosts for endosymbioses and the conversion of symbionts into organelles. J Protozool 32:376–379

Chaw SM, Shih ACC, Wang D, Wu YW, Liu SM, Chou TY (2008) The mitochondrial genome of the gymnosperm *Cycas taitungensis* contains a novel family of short interspersed elements, Bpu sequences, and abundant RNA editing sites. Mol Biol Evol 25: 603–615

Cheung AY, Bogorad L, Vanmontagu M, Schell J (1988) Relocating a gene for herbicide tolerance: a chloroplast gene is converted into a nuclear gene. Proc Natl Acad Sci USA 85:391–395

Clemens DL, Johnson PJ (2000) Failure to detect DNA in hydrogenosomes of *Trichomonas vaginalis* by nick translation and immunomicroscopy. Mol Biochem Parasit 106:307–313

Clifton SW, Minx P, Fauron CMR, Gibson M, Allen JO, Sun H, Thompson M, Barbazuk WB, Kanuganti S, Tayloe C, Meyer L, Wilson RK, Newton KJ (2004) Sequence and comparative analysis of the maize NB mitochondrial genome. Plant Physiol 136:3486–3503

Cornelissen M, Vandewiele M (1989) Nuclear transcriptional activity of the plastid *psbA* promoter. Nucleic Acids Res 17:19–29

Cusack BP, Wolfe KH (2007) When gene marriages don't work out: divorce by subfunctionalization. Trends Genet 23:270–272

Daley DO, Whelan J (2005) Why genes persist in organelle genomes. Genome Biol 6:110

Delannoy E, Fujii S, des Francs CC, Brundrett M, Small I (2011) Rampant gene loss in the underground Orchid Rhizanthella gardneri highlights evolutionary constraints on plastid genomes. Mol Biol Evol 28:2077–2086

Doolittle WE (1998) You are what you eat: a gene transfer ratchet could account for bacterial genes in eukaryotic nuclear genomes. Trends Genet 14: 307–311

Figueroa P, Gomez I, Holuigue L, Araya A, Jordana X (1999) Transfer of *rps14* from the mitochondrion to

the nucleus in maize implied integration within a gene encoding the iran-sulphur subunit of succinate dehydrogenase and expression by alternative splicing. Plant J 18:601–609

Gallois JL, Achard P, Green G, Mache R (2001) The *Arabidopsis* chloroplast ribosomal protein L21 is encoded by a nuclear gene of mitochondrial origin. Gene 274:179–185

Gissi C, Iannelli F, Pesole G (2008) Evolution of the mitochondrial genome of Metazoa as exemplified by comparison of congeneric species. Heredity 101:301–320

Goremykin VV, Salamini F, Velasco R, Viola R (2009) Mitochondrial DNA of *Vitis vinifera* and the issue of rampant horizontal gene transfer. Mol Biol Evol 26:99–110

Grohmann L, Brennicke A, Schuster W (1992) The mitochondrial gene encoding ribosomal protein S12 has been translocated to the nuclear genome in *Oenothera*. Nucleic Acids Res 20:5641–5646

Guo XY, Ruan SL, Hu WM, Ca DG, Fan LJ (2008) Chloroplast DNA insertions into the nuclear genome of rice: the genes, sites and ages of insertion involved. Funct Integr Genomics 8:101–108

Hall AE, Kettler GC, Preuss D (2006) Dynamic evolution at pericentromeres. Genome Res 16:355–364

Handa H (2003) The complete nucleotide sequence and RNA editing content of the mitochondrial genome of rapeseed (*Brassica napus* L.): comparative analysis of the mitochondrial genomes of rapeseed and *Arabidopsis thaliana*. Nucleic Acids Res 31:5907–5916

Hao WL, Palmer JD (2009) Fine-scale mergers of chloroplast and mitochondrial genes create functional, transcompartmentally chimeric mitochondrial genes. Proc Natl Acad Sci USA 106: 16728–16733

Haviv-Chesner A, Kobayashi Y, Gabriel A, Kupiec M (2007) Capture of linear fragments at a double-strand break in yeast. Nucleic Acids Res 35: 5192–5202

Hazkani-Covo E, Zeller RM, Martin W (2010) Molecular poltergeists: mitochondrial DNA copies (numts) in sequenced nuclear genomes. PLoS Genet 6:e1000834

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

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

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

Huang CY, Grunheit 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

Kanno A, Nakazono M, Hirai A, Kameya T (1997) Maintenance of chloroplast-derived sequences in the mitochondrial DNA of gramineae. Curr Genet 32:413–419

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

Kitazaki K, Kubo T (2010) Cost of having the largest mitochondrial genome: evolutionary mechanism of plant mitochondrial genome. J Bot 2010. doi:10.1155/2010/620137

Kleine T, Maier UG, Leister D (2009) DNA transfer from organelles to the nucleus: the idiosyncratic genetics of endosymbiosis. Annu Rev Plant Biol 60:115–138

Knoop V, Brennicke A (1991) A mitochondrial intron sequence in the 5′-flanking region of a plant nuclear lectin gene. Curr Genet 20:423–425

Knoop V, Unseld M, Marienfeld J, Brandt P, Sunkel S, Ullrich H, Brennicke A (1996) *copia-, gypsy-* and LINE-like retrotransposon fragments in the mitochondrial genome of *Arabidopsis thaliana*. Genetics 142:579–585

Kubo N, Arimura S (2010) Discovery of the *rpl10* gene in diverse plant mitochondrial genomes and its probable replacement by the nuclear gene for chloroplast RPL10 in two lineages of angiosperms. DNA Res 17:1–9

Kubo T, Nishizawa S, Sugawara A, Itchoda N, Estiati A, Mikami T (2000) The complete nucleotide sequence of the mitochondrial genome of sugar beet (*Beta vulgaris* L.) reveals a novel gene for tRNA(Cys) (GCA). Nucleic Acids Res 28:2571–2576

Kundu M, Thompson CB (2005) Macroautophagy versus mitochondrial autophagy: a question of fate? Cell Death Differ 12:1484–1489

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

Lenglez S, Hermand D, Decottignies A (2010) Genome-wide mapping of nuclear mitochondrial DNA sequences links DNA replication origins to chromosomal double-strand break formation in *Schizosaccharomyces pombe*. Genome Res 20: 1250–1261

Li HM, Chiu CC (2010) Protein transport into chloroplasts. Annu Rev Plant Biol 61:157–180

Li LB, Wang B, Liu Y, Qiu YL (2009) The complete mitochondrial genome sequence of the Hornwort *Megaceros aenigmaticus* shows a mixed mode of

conservative yet dynamic evolution in early land plant mitochondrial genomes. J Mol Evol 68: 665–678
Lim L, Linka M, Mullin KA, Weber APM, McFadden GI (2010) The carbon and energy sources of the non-photosynthetic plastid in the malaria parasite. FEBS Lett 584:549–554
Lin YF, Waldman AS (2001) Capture of DNA sequences at double-strand breaks in mammalian chromosomes. Genetics 158:1665–1674
Lin XY, Kaul SS, Rounsley S, Shea TP, Benito MI, Town CD, Fujii CY, Mason T, Bowman CL, Barnstead M et al (1999) Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*. Nature 402:761–768
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
Liu SL, Zhuang Y, Zhang P, Adams KL (2009) Comparative analysis of structural diversity and sequence evolution in plant mitochondrial genes transferred to the nucleus. Mol Biol Evol 26: 875–891
Lloyd AH, Timmis JN (2011) The origin and characterization of new nuclear genes originating from a cytoplasmic organellar genome. Mol Biol Evol 28:2019–2028
Lough AN, Roark LM, Kato A, Ream TS, Lamb JC, Birchler JA, Newton KJ (2008) Mitochondrial DNA transfer to the nucleus generates extensive insertion site variation in maize. Genetics 178:47–55
Lynch M (1996) Mutation accumulation in transfer RNAs: molecular evidence for Muller's ratchet in mitochondrial genomes. Mol Biol Evol 13: 209–220
Magee AM, Aspinall S, Rice DW, Cusack BP, Semon M, Perry AS, Stefanovic S, Milbourne D, Barth S, Palmer JD, Gray JC, Kavanagh TA, Wolfe KH (2010) Localized hypermutation and associated gene losses in legume chloroplast genomes. Genome Res 20:1700–1710
Marechal A, Brisson N (2010) Recombination and the maintenance of plant organelle genome stability. New Phytol 186:299–317
Martin W, Herrmann RG (1998) Gene transfer from organelles to the nucleus: how much, what happens, and why? Plant Physiol 118:9–17
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 99:12246–12251
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
Millar AH, Whelan J, Small I (2006) Recent surprises in protein targeting to mitochondria and plastids. Curr Opin Plant Biol 9:610–615
Millen RS, Olmstead RG, Adams KL, Palmer JD, Lao NT, Heggie L, Kavanagh TA, Hibberd JM, Giray JC, Morden CW, Calie PJ, Jermiin LS, Wolfe KH (2001) Many parallel losses of *infA* from chloroplast DNA during angiosperm evolution with multiple independent transfers to the nucleus. Plant Cell 13: 645–658
Ming R, Hou SB, Feng Y, Yu QY, Dionne-Laporte A, Saw JH, Senin P, Wang W, Ly BV, Lewis KLT et al (2008) The draft genome of the transgenic tropical fruit tree papaya (*Carica papaya* Linnaeus). Nature 452:991–996
Mishmar D, Ruiz-Pesini E, Brandon M, Wallace DC (2004) Mitochondrial DNA-like sequences in the nucleus (NUMTs): insights into our African origins and the mechanism of foreign DNA integration. Hum Mutat 23:125–133
Miyata S, Nakazono M, Hirai A (1998) Transcription of plastid derived tRNA genes in rice mitochondria. Curr Genet 34:216–220
Nakazono M, Nishiwaki S, Tsutsumi N, Hirai A (1996) A chloroplast-derived sequence is utilized as a source of promotes sequences for the gene for subunit 9 of NADH dehydrogenase (*nad9*) in rice mitochondria. Mol Gen Genet 252:371–378
NCBI (2011) NCBI eukaryotae organelles list. http://www.ncbi.nlm.nih.gov/genomes/GenomesHome.cgi?taxid=2759&hopt=html. Accessed 25 Mar 2011
Neupert W (1997) Protein import into mitochondria. Annu Rev Biochem 66:863–917
Notsu Y, Masood S, Nishikawa T, Kubo N, Akiduki G, Nakazono M, Hirai A, Kadowaki K (2002) The complete sequence of the rice (*Oryza sativa* L.) mitochondrial genome: frequent DNA sequence acquisition and loss during the evolution of flowering plants. Mol Genet Genomics 268:434–445
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
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
Nugent JM, Palmer JD (1991) RNA-mediated transfer of the gene *coxII* from the mitochondrion to the

nucleus during flowering plant evolution. Cell 66:473–481

Oda K, Yamato K, Ohta E, Nakamura Y, Takemura M, Nozato N, Akashi K, Kanegae T, Ogura Y, Kohchi T, Ohyama K (1992) Gene organization deduced from the complete sequence of liverwort *Marchantia polymorpha* mitochondrial DNA: a primitive form of plant mitochondrial genome. J Mol Biol 223:1–7

Ogihara Y, Yamazaki Y, Murai K, Kanno A, Terachi T, Shiina T, Miyashita N, Nasuda S, Nakamura C, Mori N, Takumi S, Murata M, Futo S, Tsunewaki K (2005) Structural dynamics of cereal mitochondrial genomes as revealed by complete nucleotide sequencing of the wheat mitochondrial genome. Nucleic Acids Res 33:6235–6250

Puthiyaveetil S, Kavanagh TA, Cain P, Sullivan JA, Newell CA, Gray JC, Robinson C, van der Giezen M, Rogers MB, Allen JF (2008) The ancestral symbiont sensor kinase CSK links photosynthesis with gene expression in chloroplasts. Proc Natl Acad Sci USA 105:10061–10066

Reith M, Munholland J (1995) Complete nucleotide sequence of the *Porphyra purpurea* chloroplast genome. Plant Mol Biol Rep 13:333–335

Ricchetti M, Fairhead C, Dujon B (1999) Mitochondrial DNA repairs double-strand breaks in yeast chromosomes. Nature 402:96–100

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

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

Roark LM, Hui Y, Donnelly L, Birchler JA, Newton KJ (2010) Recent and frequent insertions of chloroplast DNA into maize nuclear chromosomes. Cytogenet Genome Res 129:17–23

Rousseau-Gueutin M, Lloyd AH, Sheppard AE, Timmis JN (2011) Gene transfer to the nucleus. In: Bullerwell CE (ed) Organelle genetics: evolution of organelle genomes and gene expression. Springer, Berlin, pp 147–171

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

Ruf S, Braune S, Endries P, Hasse C, Stegemann S, Bock R (2010) Plastid transmission, gene transfer and the impact of the environment. ISCGGE, Maynooth

Salomon S, Puchta H (1998) Capture of genomic and T-DNA sequences during double-strand break repair in somatic plant cells. EMBO J 17:6086–6095

Selosse M, Albert B, Godelle B, Berg OG, Kurland CG (2001) Reducing the genome size of organelles favours gene transfer to the nucleus. Trends Ecol Evol 16:135–141

Sheppard AE, Timmis JN (2009) Instability of plastid DNA in the nuclear genome. PLoS Genet 5:e1000323

Sheppard AE, Ayliffe MA, Blatch L, Day A, Delaney SK, Khairul-Fahmy N, Li Y, Madesis P, Pryor AJ, Timmis JN (2008) Transfer of plastid DNA to the nucleus is elevated during male gametogenesis in tobacco. Plant Physiol 148:328–336

Sheppard AE, Panagiotis M, Lloyd AH, Day A, Ayliffe MA, Timmis JN (2011) Introducing an RNA editing requirement into a plastid-localised transgene reduces but does not eliminate functional gene transfer to the nucleus. Plant Mol Biol 76:299–309

Sloan DB, Alverson AJ, Storchova H, Palmer JD, Taylor DR (2010) Extensive loss of translational genes in the structurally dynamic mitochondrial genome of the angiosperm *Silene latifolia*. BMC Evol Biol 10:e274

Smith DR, Lee RW (2008) Mitochondrial genome of the colorless green alga *Polytomella capuana*: a linear molecule with an unprecedented GC content. Mol Biol Evol 25:487–496

Soll J, Schleiff E (2004) Protein import into chloroplasts. Nat Rev Mol Cell Biol 5:198–208

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

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

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

Stettler M, Eicke S, Mettler T, Messerli G, Hortensteiner S, Zeeman SC (2009) Blocking the metabolism of starch breakdown products in *Arabidopsis* leaves triggers chloroplast degradation. Mol Plant 2:1233–1246

Stupar RM, Lilly JW, Town CD, Cheng Z, Kaul S, Buell CR, Jiang JM (2001) Complex mtDNA constitutes an approximate 620-kb insertion on *Arabidopsis thaliana* chromosome 2: implication of potential sequencing errors caused by large-unit repeats. Proc Natl Acad Sci USA 98:5099–5103

Sugiyama Y, Watase Y, Nagase M, Makita N, Yagura S, Hirai A, Sugiura M (2005) The complete nucleotide sequence and multipartite organization of the tobacco mitochondrial genome: comparative analysis of mitochondrial genomes in higher plants. Mol Genet Genomics 272:603–615

Svab Z, Hajdukiewicz P, Maliga P (1990) Stable transformation of plastids in higher plants. Proc Natl Acad Sci USA 87:8526–8530

Terasawa K, Odahara M, Kabeya Y, Kikugawa T, Sekine Y, Fujiwara M, Sato N (2007) The mitochondrial genome of the moss *Physcomitrella patens* sheds new light on mitochondrial evolution in land plants. Mol Biol Evol 24:699–709

The Rice Chromosome 10 Sequencing Consortium (2003) In-depth view of structure, activity, and evolution of rice chromosome 10. Science 300: 1566–1569

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

Timmis JN, Scott NS (1983) Sequence homology between spinach nuclear and chloroplast genomes. Nature 305:65–67

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

Turmel M, Otis C, Lemieux C (2003) The mitochondrial genome of *Chara vulgaris*: insights into the mitochondrial DNA architecture of the last common ancestor of green algae and land plants. Plant Cell 15:1888–1903

Ueda M, Nishikawa T, Fujimoto M, Takanashi H, Arimura S, Tsutsumi N, Kadowaki K (2008) Substitution of the gene for chloroplast RPS16 was assisted by generation of a dual targeting signal. Mol Biol Evol 25:1566–1575

Unseld M, Marienfeld JR, Brandt P, Brennicke A (1997) The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366,924 nucleotides. Nat Genet 15:57–61

van den Boogaart P, Samallo J, Agsteribbe E (1982) Similar genes for a mitochondrial ATPase subunit in the nuclear and mitochondrial genomes of *Neurospora crassa*. Nature 298:187–189

van der Giezen M, Sjollema KA, Artz RR, Alkema W, Prins RA (1997) Hydrogenosomes in the anaerobic fungus *Neocallimastix frontalis* have a double membrane but lack an associated organelle genome. FEBS Lett 408:147–150

Villarejo A, Buren S, Larsson S, Dejardin A, Monne M, Rudhe C, Karlsson J, Jansson S, Lerouge P, Rolland N, von Heijne G, Grebe M, Bako L, Samuelsson G (2005) Evidence for a protein transported through the secretory pathway *en route* to the higher plant chloroplast. Nat Cell Biol 7:1224–1231

Vogel JP, Garvin DF, Mockler TC, Schmutz J, Rokhsar D, Bevan MW, Barry K, Lucas S, Harmon-Smith M, Lail K et al (2010) Genome sequencing and analysis of the model grass *Brachypodium distachyon*. Nature 463:763–768

Vonheijne G (1986) Why mitochondria need a genome. FEBS Lett 198:1–4

Wada S, Ishida H, Izumi M, Yoshimoto K, Ohsumi Y, Mae T, Makino A (2009) Autophagy plays a role in chloroplast degradation during senescence in individually darkened leaves. Plant Physiol 149: 885–893

Wang D, Wu YW, Shih ACC, Wu CS, Wang YN, Chaw SM (2007) Transfer of chloroplast genomic DNA to mitochondrial genome occurred at least 300 MYA. Mol Biol Evol 24:2040–2048

Wilson RJM, Denny PW, Preiser PR, Rangachari K, Roberts K, Roy A, Whyte A, Strath M, Moore DJ, Moore PW, Williamson DH (1996) Complete gene map of the plastid-like DNA of the malaria parasite *Plasmodium falciparum*. J Mol Biol 261: 155–172

Wolfe KH, Li WH, Sharp PM (1987) Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. Proc Natl Acad Sci USA 84:9054–9058