

Plastid DNA in the nucleus

New genes for old

Mathieu Rousseau-Gueutin,^{1,*} Michael A. Ayliffe² and Jeremy N. Timmis¹

¹School of Molecular and Biomedical Science; The University of Adelaide; Adelaide, SA Australia; ²CSIRO Plant Industry; Canberra, ACT Australia

Nuclear genomes of eukaryotes are bombarded by a continuous deluge of organellar DNA which contributes significantly to eukaryote evolution. Here, we present a new PCR-based method that allows the specific amplification of nuclear integrants of organellar DNA (norgs) by exploiting recent deletions present in organellar genome sequences. We have used this method to amplify nuclear integrants of plastid DNA (nupts) from the nuclear genomes of several *Nicotiana* species and to study the evolutionary forces acting upon these sequences. The role of nupts in endosymbiotic evolution and the different genetic factors influencing the time available for a chloroplastic gene to be functionally relocated in the nucleus are discussed.

Keywords: endosymbiotic evolution, organelle, DNA transfer, functional gene transfer, functional gene relocation, nupt, numt, norg

Abbreviations: Mya, millions of years ago; nupt, nuclear integrants of plastid DNA; numt, nuclear integrants of mitochondrial DNA; norg, nuclear integrants of organelle DNA; UTR, untranslated region

Submitted: 11/12/11

Accepted: 11/14/11

<http://dx.doi.org/10.4161/psb.18762>

*Correspondence to: Mathieu Rousseau-Gueutin;
Email: mathieu.rousseau@adelaide.edu.au

Addendum to: Rousseau-Gueutin M, Ayliffe MA, Timmis JN. Conservation of plastid sequences in the plant nuclear genome for millions of years facilitates endosymbiotic evolution. *Plant Physiol* 2011; 157:2181–93; PMID:22034627; <http://dx.doi.org/10.1104/pp.111.185074>

Eukaryotic cells arose more than a billion years ago when an ancestor of the nucleated cell engulfed a free-living α -proteobacterium¹ followed by a cyanobacterium² that were gradually converted into mitochondria and chloroplasts (plastids) respectively. Since these events, there is a continuous influx of organellar DNA entering the nuclear genome.^{3–5} Organellar DNA in the nucleus is referred to as numts (nuclear integrants of mitochondrial DNA⁶) and nupts (nuclear integrants of plastid DNA⁷) or collectively as norgs (nuclear integrants of organelle DNA⁸). These norgs contribute significantly to eukaryote evolution by providing a major source of genetic diversity. They also create new genes,⁹ new nuclear exons encoding parts of novel proteins^{10,11} and novel gene regulatory elements.¹²

Interestingly, the large reduction in organelle genome size that accompanied endosymbiotic transfer of cytoplasmic organellar genes to the nucleus did not greatly change the spectrum of proteins required for function and biogenesis of the cytoplasmic organelles.⁷ Genes derived from organellar genomes are prokaryote-like and do not immediately become functional when transferred to the nuclear genome. The rare adaptation of an organellar gene in this new environment requires the acquisition of nuclear gene regulatory elements and a target peptide if the protein is to be functional within the organelle.^{13,14} The number of nuclear regulatory elements required for function of an organellar gene presumably varies since some plastid promoters (e.g., *psbA*) can function in the nucleus,^{14,15} some organellar genes encode cryptic organellar protein targeting signals¹⁶ and sequences in the 3'UTR of some plastid genes can promote cytoplasmic mRNA polyadenylation.^{13,14}

To elucidate the molecular mechanisms by which a norg-encoded gene becomes functional in the nucleus and replaces the organellar version, it is necessary to understand the evolutionary fate of norg sequences. In general, such studies have been confined to a few seed plants having both nuclear and organellar genomes sequenced^{17–22} and are likely to be compromised by large contiguous norg sequences being excluded as “contaminating” bona fide organellar DNA during nuclear genome assembly. The experimental isolation of norg sequences is greatly complicated by the presence of higher copy numbers of organellar genomes compared with nuclear genomes in most cells. Previous studies have relied upon differential

methylation²³ between the organellar and the nuclear genomes and only allowed the characterization of norgs that have recently been inserted. To circumvent these problems, we developed an innovative PCR-based method that allows the amplification of recent and older norgs (up to several millions of years old) in a range of eukaryotes. This method, presented in detail in Rousseau-Gueutin et al.,²⁴ avoids the amplification of the high-copy number extant organellar genomes and allows the specific amplification of norg sequences by

placing a PCR primer in a region recently deleted from the organellar genome. Organellar deletion events were identified by comparison of the organellar genomes of several closely related species. We have used this method to amplify nupts of several members of the *Nicotiana* genus as plastomes sequences were available for three *Nicotiana* species and four closely related Solanaceae species. We were able to sequence seven unrearranged nupts (25 kb in total) derived from various plastomic regions which encode several

plastidic genes. The origin of the transfers and the evolutionary processes that have acted upon the nupts were then studied by sequence comparisons with the native cytoplasmic organellar genomes and those of closely related species.²⁵ These nupts were estimated to have been transferred between approximately 0.03 and 5.8 million years ago and we determined that potential protein-coding and non-coding sequences were evolving neutrally in the nuclear genome. Some of the nupts open reading frames (ORFs) were destroyed by indels

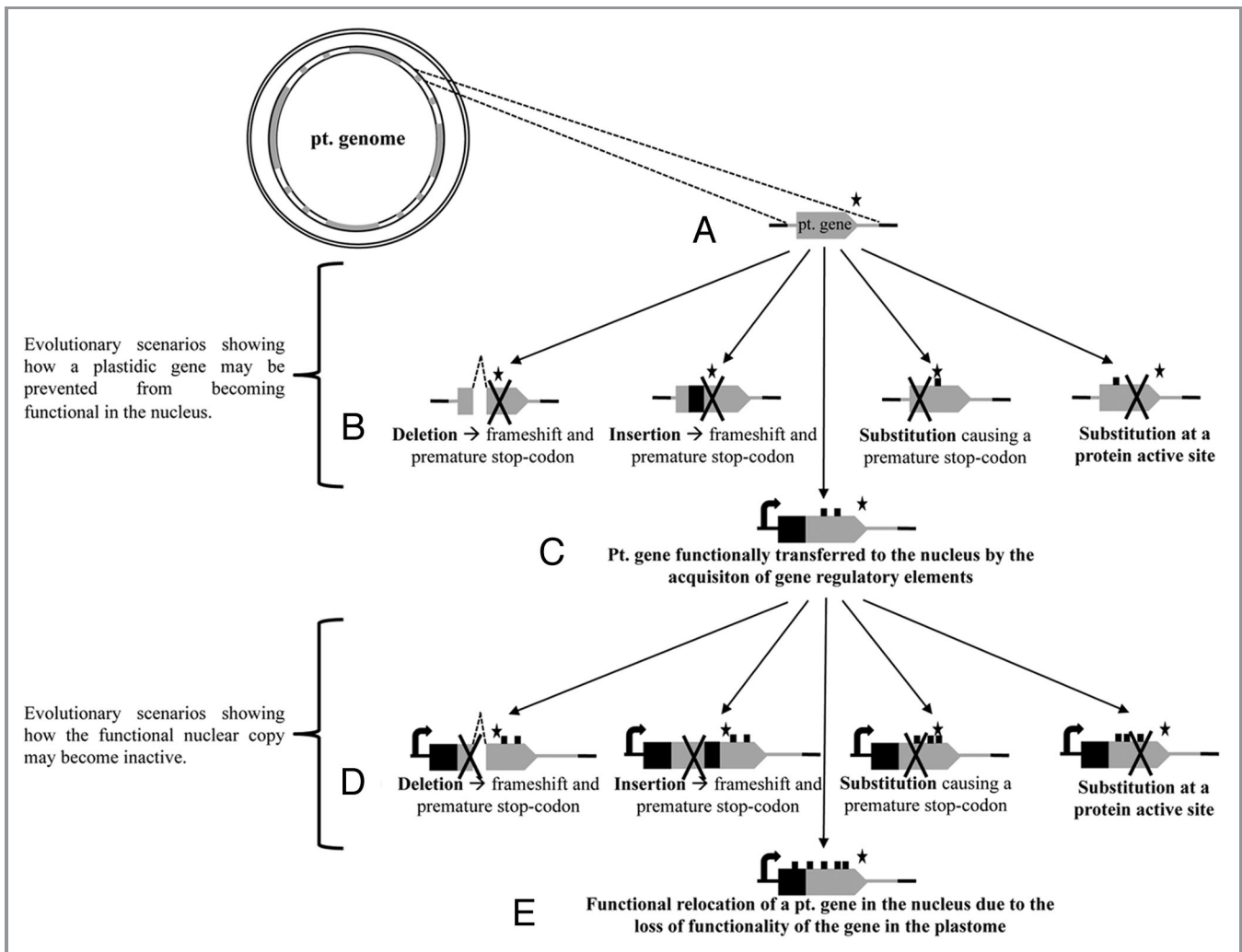


Figure 1. Possible evolution pathways of a chloroplastic gene transferred to the nucleus. (A) A plastidic region (in gray) including a plastidic gene (gray box) and some plastidic non-coding sequences (gray line) are inserted into nuclear DNA (black line). The asterisk represents the stop-codon of the potential protein-coding sequence. (B) Non-exhaustive list of the different evolutionary scenarios showing how this plastidic gene may be prevented from becoming functional in the nucleus because of indels or substitutions causing premature stop codons or because of substitution within the protein active site. Substitutions are represented by small black squares on top of the gene. (C) Gain of function when the plastidic gene in the nucleus acquires nuclear gene regulatory elements such as a promoter (black arrow) and a target peptide encoding sequence (black square). (D) Non-exhaustive list of the evolutionary scenarios that cause the subsequent inactivation of the duplicated nuclear gene, resulting in the maintenance of the plastid copy. These events will elicit repetition of the endosymbiotic transfer cycle [i.e. back to (A)]. (E) The plastidic gene functionally transferred to the nucleus relocates irrevocably to the nuclear genome because of the loss of functionality of the gene in the plastome.

leading to frameshifts and/or nucleotide substitutions causing premature stop codons (Fig. 1A and B). However, in several instances potential protein coding-sequences maintained intact ORFs (no indels but some substitutions). The oldest of these surviving ORFs was approximately 5.8 million years old, suggesting that lengthy periods are sometimes available for transferred chloroplast genes to gain nuclear function. This time will presumably vary depending upon the length of the ORF, the nature of its coding sequence, the conservation constraints on the amino-acid sequence, the physical location in the nuclear genome and to chance due to the stochastic nature of random mutations.

In rare cases of activation of an organellar gene in the nucleus (Fig. 1C), two functional copies will coexist in two separate genetic compartments of the cell until one became defunct. If the nuclear and organelle-encoded copies were equally

efficient, loss of functionality is presumably the result of chance mutation silencing one or the other copy.^{26,27} Prime face this would generally favor the retention of the organelle copy since there is a higher substitution rate in the nuclear genome than in the plastome.²⁸ In addition organellar genes are organized in operons and are usually uniparentally inherited, again favoring the status quo.²⁹ If the nuclear copy becomes defunct (Fig. 1D), the whole process can be repeated with another nupt. However if the organellar copy loses its functionality, then the nucleus becomes the permanent location of that gene (Fig. 1E). This explains the net diminution of cytoplasmic organellar genomes and the increased genetic influence of the nucleus. It is noteworthy that this kind of functional relocation of organellar genes to the nucleus has ceased in animals but is still occurring in plants, although other genetic effects of norg integration (e.g., creation of new exons¹⁰

or increased genetic diversity) still continue in both kingdoms.³⁰

Why have some genes remained in the cytoplasmic organelles? While some genes may remain in the plant plastome for reasons of maintaining redox balance,³¹ others could perhaps be relocated, but have so far failed to do so despite the long time available and the frequent transfer of organellar DNA. However, a few plastidic genes (e.g., *accD*, *infA*, *rpl22* and *rpl32*) have recently relocated to the nucleus of some plants (reviewed in ref. 32) and more may soon be identified as a number of essential genes have been lost from the plastome in a variety of angiosperms.³³ The process of organellar genome reduction therefore appears to be ongoing.

Acknowledgments

This research was supported under the Australian Research Council's Discovery Projects funding scheme (project number DP0986973).

References

- Andersson SG, Karlberg O, Canbäck B, Kurland CG. On the origin of mitochondria: a genomics perspective. *Philos Trans R Soc Lond B Biol Sci* 2003; 358:165-77; discussion 177-9; PMID:12594925; <http://dx.doi.org/10.1098/rstb.2002.1193>
- Deusch O, Landan G, Roettger M, Gruenheit N, Kowallik KV, Allen JF, et al. Genes of cyanobacterial origin in plant nuclear genomes point to a heterocyst-forming plastid ancestor. *Mol Biol Evol* 2008; 25: 748-61; PMID:18222943; <http://dx.doi.org/10.1093/molbev/msn022>
- Thorsness PE, Fox TD. Escape of DNA from mitochondria to the nucleus in *Saccharomyces cerevisiae*. *Nature* 1990; 346:376-9; PMID:2165219; <http://dx.doi.org/10.1038/346376a0>
- Huang CY, Ayliffe MA, Timmis JN. Direct measurement of the transfer rate of chloroplast DNA into the nucleus. *Nature* 2003; 422:72-6; PMID:12594458; <http://dx.doi.org/10.1038/nature01435>
- Stegemann S, Hartmann S, Ruf S, Bock R. High-frequency gene transfer from the chloroplast genome to the nucleus. *Proc Natl Acad Sci U S A* 2003; 100: 8828-33; PMID:12817081; <http://dx.doi.org/10.1073/pnas.1430924100>
- Lopez JV, Yuhki N, Masuda R, Modi W, O'Brien SJ. Numt, a recent transfer and tandem amplification of mitochondrial DNA to the nuclear genome of the domestic cat. *J Mol Evol* 1994; 39:174-90; PMID: 7932781
- Timmis JN, Ayliffe MA, Huang CY, Martin W. Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. *Nat Rev Genet* 2004; 5: 123-35; PMID:14735123; <http://dx.doi.org/10.1038/nrg1271>
- Leister D. Origin, evolution and genetic effects of nuclear insertions of organelle DNA. *Trends Genet* 2005; 21:655-63; PMID:16216380; <http://dx.doi.org/10.1016/j.tig.2005.09.004>
- Martin W, Rujan T, Richly E, Hansen A, Cornelsen S, Lins T, et al. Evolutionary analysis of Arabidopsis, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. *Proc Natl Acad Sci U S A* 2002; 99: 12246-51; PMID:12218172; <http://dx.doi.org/10.1073/pnas.182432999>
- Noutsos C, Kleine T, Armbruster U, DalCorso G, Leister D. Nuclear insertions of organellar DNA can create novel patches of functional exon sequences. *Trends Genet* 2007; 23:597-601; PMID:17981356; <http://dx.doi.org/10.1016/j.tig.2007.08.016>
- Kleine T, Maier UG, Leister D. DNA transfer from organelles to the nucleus: the idiosyncratic genetics of endosymbiosis. *Annu Rev Plant Biol* 2009; 60:115-38; PMID:19014347; <http://dx.doi.org/10.1146/annurev.arplant.043008.092119>
- Knoop V, Brennicke A. A mitochondrial intron sequence in the 5'-flanking region of a plant nuclear lectin gene. *Curr Genet* 1991; 20:423-5; PMID: 1807833; <http://dx.doi.org/10.1007/BF00317072>
- Stegemann S, Bock R. Experimental reconstruction of functional gene transfer from the tobacco plastid genome to the nucleus. *Plant Cell* 2006; 18:2869-78; PMID: 17085684; <http://dx.doi.org/10.1105/tpc.106.046466>
- Lloyd AH, Timmis JN. The origin and characterization of new nuclear genes originating from a cytoplasmic organellar genome. *Mol Biol Evol* 2011; 28:2019-28; PMID:21252282; <http://dx.doi.org/10.1093/molbev/msr021>
- Cornelissen M, Vandewiele M. Nuclear transcriptional activity of the tobacco plastid psbA promoter. *Nucleic Acids Res* 1989; 17:19-29; PMID:2563150; <http://dx.doi.org/10.1093/nar/17.1.19>
- Ueda M, Fujimoto M, Arimura SI, Tsutsumi N, Kadowaki KI. Presence of a latent mitochondrial targeting signal in gene on mitochondrial genome. *Mol Biol Evol* 2008; 25:1791-3; PMID:18573842; <http://dx.doi.org/10.1093/molbev/msn139>
- Shahmuradov IA, Akbarova YY, Solovyev VV, Aliyev JA. Abundance of plastid DNA insertions in nuclear genomes of rice and Arabidopsis. *Plant Mol Biol* 2003; 52:923-34; PMID:14558655; <http://dx.doi.org/10.1023/A:1025472709537>
- Richly E, Leister D. NUPTs in sequenced eukaryotes and their genomic organization in relation to NUMTs. *Mol Biol Evol* 2004; 21:1972-80; PMID:15254258; <http://dx.doi.org/10.1093/molbev/msn210>
- Huang CY, Grünheit N, Ahmadinejad N, Timmis JN, Martin W. Mutational decay and age of chloroplast and mitochondrial genomes transferred recently to angiosperm nuclear chromosomes. *Plant Physiol* 2005; 138:1723-33; PMID:15951485; <http://dx.doi.org/10.1104/pp.105.060327>
- Matsuo M, Ito Y, Yamauchi R, Obokata J. The rice nuclear genome continuously integrates, shuffles, and eliminates the chloroplast genome to cause chloroplast-nuclear DNA flux. *Plant Cell* 2005; 17:665-75; PMID: 15705954; <http://dx.doi.org/10.1105/tpc.104.027706>
- Noutsos C, Richly E, Leister D. Generation and evolutionary fate of insertions of organelle DNA in the nuclear genomes of flowering plants. *Genome Res* 2005; 15:616-28; PMID:15867426; <http://dx.doi.org/10.1101/gr.3788705>
- Guo XY, Ruan SL, Hu WM, Cai D, Fan LJ. Chloroplast DNA insertions into the nuclear genome of rice: the genes, sites and ages of insertion involved. *Funct Integr Genomics* 2008; 8:101-8; PMID:17994302; <http://dx.doi.org/10.1007/s10142-007-0067-2>
- Ayliffe MA, Scott NS, Timmis JN. Analysis of plastid DNA-like sequences within the nuclear genomes of higher plants. *Mol Biol Evol* 1998; 15:738-45; PMID: 9615455
- Rousseau-Gueutin M, Ayliffe MA, Timmis JN. Conservation of plastid sequences in the plant nuclear genome for millions of years facilitates endosymbiotic evolution. *Plant Physiol* 2011; 157:2181-93; PMID: 22034627; <http://dx.doi.org/10.1104/pp.111.185074>

25. Hazkani-Covo E, Sorek R, Graur D. Evolutionary dynamics of large numts in the human genome: rarity of independent insertions and abundance of post-insertion duplications. *J Mol Evol* 2003; 56:169-74; PMID:12574863; <http://dx.doi.org/10.1007/s00239-002-2390-5>
26. Martin W, Herrmann RG. Gene transfer from organelles to the nucleus: how much, what happens, and Why? *Plant Physiol* 1998; 118:9-17; PMID: 9733521; <http://dx.doi.org/10.1104/pp.118.1.9>
27. Adams KL, Song K, Roessler PG, Nugent JM, Doyle JL, Doyle JJ, et al. Intracellular gene transfer in action: dual transcription and multiple silencings of nuclear and mitochondrial *cox2* genes in legumes. *Proc Natl Acad Sci U S A* 1999; 96:13863-8; PMID:10570164; <http://dx.doi.org/10.1073/pnas.96.24.13863>
28. Wolfe KH, Li WH, Sharp PM. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proc Natl Acad Sci U S A* 1987; 84:9054-8; PMID:3480529; <http://dx.doi.org/10.1073/pnas.84.24.9054>
29. Wicke S, Schneeweiss GM, dePamphilis CW, Müller KF, Quandt D. The evolution of the plastid chromosome in land plants: gene content, gene order, gene function. *Plant Mol Biol* 2011; 76:273-97; PMID:21424877; <http://dx.doi.org/10.1007/s11103-011-9762-4>
30. Hazkani-Covo E, Zeller RM, Martin W. Molecular poltergeists: mitochondrial DNA copies (numts) in sequenced nuclear genomes. *PLoS Genet* 2010; 6: e1000834; PMID:20168995; <http://dx.doi.org/10.1371/journal.pgen.1000834>
31. Allen JF. Redox control of gene-expression and the function of chloroplast genomes—an hypothesis. *Photosynth Res* 1993; 36:95-102; <http://dx.doi.org/10.1007/BF00016274>
32. Rousseau-Gueutin M, Lloyd AH, Sheppard AE, Timmis JN. Gene transfer to the nucleus. In: Bullerwell C, ed. *Organelle Genetics: Evolution of Organelle Genomes and Gene Expression*. Berlin: Springer Verlag, 2011:147-71.
33. Magee AM, Aspinall S, Rice DW, Cusack BP, Sémon M, Perry AS, et al. Localized hypermutation and associated gene losses in legume chloroplast genomes. *Genome Res* 2010; 20:1700-10; PMID:20978141; <http://dx.doi.org/10.1101/gr.111955.110>

© 2012 Landes Bioscience.
Do not distribute.