Proc. Natl. Acad. Sci. USA
Vol. 95, pp. 14003–14005, November 1998

Commentary

Mass migration of a group I intron: Promiscuity on a grand scale

Michael W. Gray

Program in Evolutionary Biology, Canadian Institute for Advanced Research, Department of Biochemistry, Dalhousie University, Halifax, Nova Scotia B3H 4H7, Canada

Both genetically and biochemically, group I introns are rather special. At the RNA level, they are inherently autocatalytic, mediating their own removal from transcripts containing them and effecting the ligation of flanking exons (self-splicing) (1). Many group I introns are mobile elements, able to spread in genetic crosses to alleles that do not contain them via a process known as intron homing (2). Homing is initiated by a site-specific endonuclease encoded by the intron. Additionally, several group I introns specify protein cofactors (maturases) that function in the splicing of the intron RNA that encodes them (1, 2). Although a given group I intronic reading frame almost always specifies either endonuclease or maturase activity, there are a few cases known in which the encoded protein can perform both functions (3, 4).

To date, most studies of group I intron mobility have dealt with intraorganismal transfer occurring between intron$^+$ and intron$^-$ alleles of particular genes during genetic crosses. Little is known about the frequency and extent of horizontal transfer of group I introns between organisms that do not mate. Most such identified cases involve transfer into the same genome (e.g., mitochondrial) in taxa that at least belong to the same phylum (see ref. 5). There is, however, one reported instance in which interphylum (and interorganellar) transfer of group I introns appears to have occurred (6). Even so, nothing remotely approaching the extraordinary intron radiation reported by Cho *et al.* in this issue of the *Proceedings* (5) has been documented previously.

What these workers have uncovered is an explosive invasion of plant mitochondrial DNA (mtDNA) by a particular group I intron. The authors were led to the present study by a previous finding (7) of this curious intron in the gene encoding subunit 1 of cytochrome oxidase (*cox1*) in the mtDNA of an angiosperm (flowering plant), *Peperomia polybotrya*. Not only is this the sole group I intron so far reported in the mtDNA of vascular plants (in contrast to the frequent presence of group II introns in plant mtDNA), it clearly is of a different evolutionary origin than the gene in which it resides. In fact, phylogenetic evidence suggests that this intron arose recently by horizontal transfer from a fungal donor species (7). In the initial study (7), the intron was not found in *cox1* from 19 other diverse plant species, and a follow-up investigation (8) indicated that it was restricted to the single genus *Peperomia* within the order Piperales. However, the serendipitous finding of a closely related intron at the identical position in the *cox1* gene of a distantly related angiosperm (*Veronica*) led Cho *et al.* to embark on a comprehensive survey of the intron's distribution among 335 diverse genera of land plants. From this survey, the authors infer 32 separate cases of intron acquisition among the 278 genera and 281 species of angiosperm examined in the botanical equivalent of a "zoo blot." Extrapolating to angiosperms as a whole, Cho *et al.* (5) come to the startling conclusion that this intron has invaded the *cox1* gene >1,000 times among the >13,500 genera and >300,000 species of extant flowering plants.

The evidence on which this conclusion is based is of several kinds. First, the authors' Southern hybridization survey revealed clear evidence of a (presumably mtDNA-encoded) *cox1* gene in all DNA samples examined whereas the *cox1* intron was found only sporadically among the same DNAs, with little or no phylogenetic coherence in its distribution. In contrast, a nearly universal hybridization pattern was seen for 11 plant mitochondrial group II introns. These results strongly point to vertical inheritance of the group II introns from an ancestral mitochondrial genome containing them but lateral transfer of the *cox1* group I intron during the evolution of angiosperms.

A second piece of evidence for sporadic and recurrent group I intron transfer is the noncongruent phylogenetic histories of the *cox1* intron and of the organisms in which it resides. Phylogenies for 30 angiosperms whose *cox1* introns were sequenced were generated from the *cox1* coding region, from a chloroplast gene (*rbcL*), and from a data set combining both gene sequences. Phylogenetic analysis was rigorous, and clear noncorrespondence of branching order was evident in the resulting intron and organismal (i.e., gene) trees. Those introns showing greatest sequence and phylogenetic similarity were often seen to derive from plant species that are only distantly related.

A final consideration strongly favoring many separate group I intron gains is the sequence of the exons flanking the intron insertion site. Group I intron transfer is thought to proceed at the molecular level by way of a recombination/repair process initiated by a staggered double-strand break, catalyzed by the homing endonuclease, at the target site in the intron–allele (2, 9–11). The cleaved DNA strands of the recipient DNA are partially degraded, creating a gap that is filled in using information provided by the donor DNA. Any flanking exon sequence that is lost through nucleolytic degradation of the cleaved recipient DNA is converted effectively to the corresponding donor DNA sequence through this process. Thus, if flanking exons differ in sequence (as would be expected to be the case when intron$^+$ and intron$^-$ genes come from distantly related organisms), then the exon sequence in the recipient DNA will be changed to match that of the donor DNA, a phenomenon referred to as coconversion of flanking markers.

Of 30 intron-containing *cox1* genes sequenced by Cho *et al.* (5), 28 had three or more sequence variations within an 18-bp region immediately downstream of the intron. Strikingly, at any given site, the variations were identical, extending in a gradient 3′ to the intron insertion site; moreover, they were all silent (third position) changes at the level of Cox1 amino acid sequence. Arguing from the generally extremely low rate of sequence change in plant mtDNA (12), the absence of any selective pressure for back-mutation (because the sequence variations are phenotypically neutral), and the lack of evidence for back-mutation (which would erase the 3′ coconversion gradient), the authors draw two conclusions: that (*i*) *cox1* genes that appear closely related but whose coconversion tracts differ in length likely acquired their introns separately; and (*ii*) introns that appear closely related and have identical cocon-

© 1998 by The National Academy of Sciences 0027-8424/98/9514003-3$2.00/0
PNAS is available online at www.pnas.org.

The companion to this Commentary begins on page 14244.

version tracts nevertheless also likely arose independently, if they are present in species that are phylogenetically intermingled with taxa that do not have the intron. The latter inference follows from 24 sequenced *cox1* genes that are intron$^-$ and that show no evidence whatsoever of the 3′ coconversion variations that would be expected to persist if an intron had been present at one time but then lost. These intron$^-$ *cox1* genes, interspersed in cladograms with intron$^+$ *cox1* homologs, appear never to have harbored this particular intron.

The upshot of this analysis is that Cho *et al.* (5) effectively discount a scenario in which the current intron distribution can be explained by a single ancestral gain of the intron very early in angiosperm evolution, followed by many subsequent losses. Instead, the data are most compatible with an all-gain model, with most instances of intron acquisition being relatively recent, i.e., within the time-frame of the evolutionary diversification of individual angiosperm families.

Not surprisingly, the results reported in this study raise a number of questions, three of which are perhaps the most obvious and the most intriguing: (*i*) What accounts for the extraordinary invasiveness of this particular intron? At the heart of group I intron mobility is the target site in the intron-allele that is recognized and cleaved by the homing endonuclease. This site generally spans ≈20 bp and is located at or near the site of intron insertion (1). Therefore, a first requirement for successful horizontal intron transfer is that the target site must be highly conserved in the intronless recipient gene. Because angiosperm mitochondrial coding sequences generally diverge at an exceptionally low rate (12), there is probably a better chance of target site conservation within the broad range of angiosperm mtDNAs than among any other collection of genomes that one might identify. At the same time, because a high degree of nucleotide conservation within a protein-coding gene usually reflects functional constraints on the encoded amino acid sequence, it is important that any coconversion of flanking exons have a minimal effect on protein sequence after intron insertion. In the present case, virtually all of the intron-flanking sequence variations that were found by Cho *et al.* are neutral (silent third position) changes. Moreover, as the authors observe, coconversion tract lengths are much shorter than those typically seen for group I mobile introns, a factor that would minimize further the possibility of an adverse effect of intron homing on Cox1 function. In essence, the *cox1* target sequence recognized by this particular homing endonuclease is likely to be present in a very wide range of angiosperm species, with minimal deleterious effects apparent on Cox1 sequence and function as a consequence of intron insertion.

The authors also point out that, in this particular case, the homing endonuclease may be extremely active; splicing of the intron at the RNA level may be essentially independent of host cofactors; or splicing may depend on highly conserved, ubiquitous host factors. The latter two possibilities, in particular, would contribute further to the evidently low level of host specificity required for lateral transfer of the *cox1* group I intron.

(*ii*) How does horizontal transfer actually take place at the cellular level? For intron homing to occur, both the intron-containing donor DNA and the intronless recipient DNA must be in the same physical location; the intron$^+$ DNA must be satisfactorily transcribed; and the encoded endonuclease must be correctly translated. Because *cox1* is encoded in the mitochondrial genome, this implies that the *cox1* intron homing described by Cho *et al.* (5) takes place within mitochondria. That being the case, the mitochondrial transcription and translation systems must be able to accommodate the foreign donor DNA to allow expression of homing endonuclease function; moreover, angiosperm mitochondria must be able to carry out the recombination/repair process required for intron integration. In fact, recombination-mediated genome rearrangements are well documented in plant mitochondria, as is the evolutionary incorporation of promiscuous (e.g., chloroplast) DNA into the plant mitochondrial genome (see refs. 13 and 14). Allowing that the foreign donor DNA is of fungal mitochondrial origin (see below), this scenario raises a number of issues relating to gene expression in fungal and angiosperm mitochondria, such as differences in the genetic code used (refs. 15 and 16; see also ref. 8), the existence of C-to-U RNA editing in plant mitochondria (17), and differences in promoter sequences recognized by the fungal and plant mitochondrial transcriptional apparatus (18). Interorganism group I intron transfer obviously occurs, as compellingly demonstrated here by Cho *et al.* (5), but the details of how this happens at the cellular and molecular level are far from clear at this point.

(*iii*) What is the evolutionary pathway of spread of the *cox1* group I intron among angiosperms? The fungal ancestry of the *cox1* intron is evident from the fact that it is more closely related to a group of fungal mitochondrial introns (some of which do not have the same insertion site in *cox1* and one that is even located in a different gene) than to positionally equivalent and homologous *cox1* introns present in the early-arising land plant *Marchantia polymorpha* (liverwort) and the chlorophyte alga *Prototheca wickerhamii*. To account for the current broad distribution of the intron throughout angiosperms, the authors consider two models: (*i*) many or all of the donors of the intron were nonplants (presumably fungi), in which case the donors themselves must all have been closely related; or (*ii*) a great many plant-to-plant lateral transfers occurred subsequent to one or a few fungal transfers. Although these two scenarios make distinctly different phylogenetic predictions, the current data are not sufficient to distinguish between them. Generating the information base for doing so, and for deducing the timing of transfer and actual donor–recipient identities, will be "a daunting task," in the authors' words. *A priori*, model *i* draws support from the observation (7) that, in nearly all plant species, vesicular mycorrhizal (VAM) fungi are known to grow in close association with root cells, as obligate inter- and intracellular symbionts (see ref. 19). Plant-to-plant transmission, on the other hand, presumably would require some sort of vector (viruses, bacteria, insects, etc.).

The work reported by Cho *et al.* (5) is significant in a broader context. First, it provides a valuable lesson in the virtues and rewards of comprehensive taxonomic sampling when biological questions of this type are addressed. The picture of *cox1* intron distribution developed by surveying several hundreds of plant species (5) turns out to be radically different than the one that initially emerged when only a few tens of species were examined (7, 8). Second, this study raises the issue of the frequency with which interspecies DNA (and hence gene) transfer occurs in the wild. It remains to be seen whether the results reported by Cho *et al.* are directly relevant to concerns about lateral gene transfer from transgenic crop plants to wild relatives (e.g., ref. 20) or even nonrelatives. Nevertheless, the results make it clear that plants acquire foreign DNA by lateral transfer considerably more frequently in nature than we might have suspected. The corollary is that gene flow across plant breeding barriers, not as readily monitored as intron flow, also may be of a greater magnitude than we currently appreciate.

1. Saldanha, R., Mohr, G., Belfort, M. & Lambowitz, A. M. (1993) *FASEB J.* **7,** 15–24.
2. Lambowitz, A. M. & Belfort, M. (1993) *Annu. Rev. Biochem.* **62,** 587–622.
3. Henke, R. M., Butow, R. A. & Perlman, P. S. (1995) *EMBO J.* **14,** 5094–5099.
4. Ho, Y., Kim, S.-J. & Waring, R. B. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 8994–8999.
5. Cho, Y., Qiu, Y.-L., Kuhlman, P. & Palmer, J. D. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 14244–14249.

6. Turmel, M., Côté, V., Otis, C., Mercier, J.-P., Gray, M. W., Lonergan, K. M. & Lemieux, C. (1995) *Mol. Biol. Evol.* **12,** 533–545.
7. Vaughn, J. C., Mason, M. T., Sper-Whitis, G. L., Kuhlman, P. & Palmer, J. D. (1995) *J. Mol. Evol.* **41,** 563–572.
8. Adams, K. L., Clements, M. J. & Vaughn, J. C. (1998) *J. Mol. Evol.* **46,** 689–696.
9. Szostak, J. W., Orr-Weaver, T. L., Rothstein, R. J. & Stahl, F. W. (1983) *Cell* **33,** 25–35.
10. Clyman, J. & Belfort, M. (1992) *Genes Dev.* **6,** 1269–1279.
11. Belfort, M. & Perlman, P. S. (1995) *J. Biol. Chem.* **270,** 30237–30240.
12. Wolfe, K. H., Li, W.-H. & Sharp, P. M. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 9054–9058.
13. Hanson, M. R. & Folkerts, O. (1993) *Int. Rev. Cytol.* **141,** 129–172.
14. Wolstenholme, D. R. & Fauron, C. M.-R. (1995) in *The Molecular Biology of Plant Mitochondria*, eds. Levings, C. S., III & Vasil, I. K. (Kluwer, Dordrecht, The Netherlands), pp. 1–59.
15. Fox, T. D. (1987) *Annu. Rev. Genet.* **21,** 67–91.
16. Osawa, S., Jukes, T. H., Watanabe, K. & Muto, A. (1992) *Microbiol. Rev.* **56,** 229–264.
17. Gray, M. W. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 8157–8159.
18. Tracy, R. L. & Stern, D. B. (1995) *Curr. Genet.* **28,** 205–216.
19. Bonfante-Fasolo, P. (1984) in *VA Mycorrhizae*, eds. Powell, C. L. & Bagyaraj, D. J. (CRC, Boca Raton, FL), pp. 5–33.
20. Bergelson, J., Purrington, C. B. & Wichmann, G. (1998) *Nature (London)* **395,** 25.