J Mol Evol (1985) 22:243–251

Journal of Molecular Evolution

Molecular Archaeology of the Mitochondrial Genome

Robert Obar[1] and James Green[2]

[1] Department of Chemistry, Boston University, Boston, Massachusetts 02215, USA
[2] Department of Chemistry, University of Tennessee, Knoxville, Tennessee 37996-0326, USA

Summary. We develop a stepwise model for the net transfer of nucleic acid sequences between non-homologous genomes. This model is then used to explain the two major patterns in the evolutionary history of mitochondrial genomes: the gross reduction of the number of genes, and the subsequent acquisition of introns.

Key words: Molecular archaeology — Intron evolution — Mitochondria — Gene transfer — Endosymbiosis

Introduction

To determine the detailed evolutionary history of a complex genetic unit, whether a coding or regulatory nucleic acid sequence, a genome, or an entire organism, one must go beyond simple comparisons of like and unlike characteristics and construction of branching lines of descent. One successful method involves first positing a stepwise scenario of plausible steps for all transitions that involve large evolutionary discontinuities. Once such a scenario is available, the existence of analogs of proposed intermediates and the effects of natural selection on the products of each proposed step can be examined. Groups of organisms, times, and places for each step may then be envisaged; thus, this method is useful on both the molecular and organismal levels.

We have chosen here to apply this method to the problem of the coevolution of nuclear and mitochondrial genomes. This is an especially apt system to analyze, because very drastic changes have occurred in the 1 billion years since the postulated merger between proto-nucleocytoplasmic organisms and their endosymbiotic proto-mitochondria. Moreover, such intermediate forms as may have once existed have long since vanished, precluding a cladistic approach to the problem.

To avoid terminological confusion, a glossary indicating how certain terms are used here has been included at the end of this article.

Offprint requests to: R. Obar, Worcester Foundation for Experimental Biology, Shrewsbury, MA 01545

The Small Genomes of Mitochondria

Although there is a general belief that mitochondria evolved from free-living, aerobic, respiring bacteria [summarized by Gray (1983)], most of the attention given to the intermediate steps of this process has focused on the biochemistry of mitochondrial respiration and the components of the mitochondrial membrane (John and Whatley 1977; Gray and Doolittle 1982). These approaches tend to aim at providing an explanation of the evolution of metabolic integration of the organelle and the nucleocytoplasm, and to ignore the concomitant evolution of genetic integration of the two genomes. The key underlying principle in this article is that in an association, heterologous genomes become integrated only if a gene product of one partner is required for the completion of the life cycle of the second [for a fuller explanation of these topics see Margulis (1976)].

The major differences between the molecular biologies of the mitochondrion and its putative ancestor, a *Paracoccus*-like organism (John and Whatley 1977) or *Rhodopseudomonas*-like organ-

ism (Fox et al. 1980), are the greatly reduced genome size of the mitochondrion, the roles of introns (intervening sequences) and thus of RNA splicing in the gene expression of some mitochondria, and the different genetic codes used in translation.

We believe that these discontinuities evolved from the net transfer of genetic units between nucleocytoplasmic and mitochondrial genomes. By treating mitochondrial evolution as a stepwise process in the context of current understanding of the two sets of genomes involved, we will show that a coherent scenario for the evolution of eukaryotic genomes may be achieved.

Assumptions

It is reasonable to assume the following:

1. Since fossil eukaryotes at least a thousand million years old have been identified (Vidal 1984), the proto-mitochondrion first entered into a symbiosis with the proto-nucleocytoplasm before that time.
2. Each of these two partners possessed sufficient coding and enzymatic machinery to support itself as a free-living cell in its respective niche. Thus their genomic organizations, macromolecular sequences, protein compositions, and metabolic capabilities differed considerably.
3. The proto-nucleocytoplasmic genome contained introns in many of its genes before the symbiosis reached a stage of genetic integration. This assumption is supported by the evidence that the nucleocytoplasm has many archaebacterial characteristics (Gray and Doolittle 1982) and that archaebacteria have introns (Kaine et al. 1983).
4. The proto-mitochondrial genome did not contain introns, since it was a typical Gram-negative eubacterial genome.
5. There existed (and in altered form still exists) at least one mechanism to translocate genetic material between the two types of genome.

Once the association was established on metabolic grounds, i.e., when fermentative waste products of the archaebacterial host had become substrates for the metabolism of the proto–mitochondria, the unit affected by external forces such as predation and environmental variation was the symbiosis, that is, the partnership, rather than the individual symbionts. After this point external selective pressures were the same for all components of what had become a complex cell. Therefore, it is assumed that intergenomic competition within the cell was the most important force affecting the coevolution of the nuclear and mitochondrial genomes.

In its biochemical repertoire the proto-nucleocytoplasm already possessed a preadaptation for the expression of "foreign" genes: a splicing apparatus necessary for the editing of noncoding sequences. The proto-mitochondrion, a typical eubacterium, could express only "nonmosaic" genes. These traits conferred upon the proto-nucleocytoplasmic genome a considerable selective edge in the net transfer process (explained below), with the effect that net transfers tended to occur in only one direction, from proto-mitochondrion to proto-nucleus. This selective advantage, however, was eventually diminished after many generations of net transfers to the proto-nucleus made the reduced genome of the proto-mitochondrion a suitable recipient for the net transfer of certain genetic elements (introns) that could take an active role in their own expression (see Facultative Mitochondrial Introns, below).

The Model

A minimal mechanism for the net transfer of a genetic element from a "donor" genome to a "recipient" genome (see Fig. 1) is described in the following series of evolutionary steps:

Step A—Duplication. A genetic element within the linear order of the donor genome is duplicated.

Step B—Use. Over time, the recipient becomes competent to utilize regularly some genetic or metabolic product of the genetic element. Note: Donor and recipient need not utilize the same product of this gene.

Step C—Transmigration. At least one copy of the (donor) genetic element moves to and is integrated into the recipient genome.

Step D—Activation. The newly acquired genetic element is transcribed and, if it encodes a protein, translated by the recipient.

Step E—Regulation. The recipient fine-tunes its regulation of the genetic element and its product(s) and synthesis and transfer of each product to its functional site. Note: In this step there is functional genetic redundancy as the genomes compete in the formation of gene product.

Step F—Selection. The donor copy of the genetic element becomes nonfunctional (due to a regulatory or functional lesion); the recipient thus assumes full control of the genetic element and its products. Note: The donor copy is now fully redundant: by definition, a pseudogene. The symbiosis is now fully obligate.

Step G—Loss. The moribund (pseudogene) copy of the genetic element is altered beyond recognition by mutation or otherwise lost from the donor genome.

These steps may be viewed as a formalization of the ideas of "levels of partner integration" (Mar-

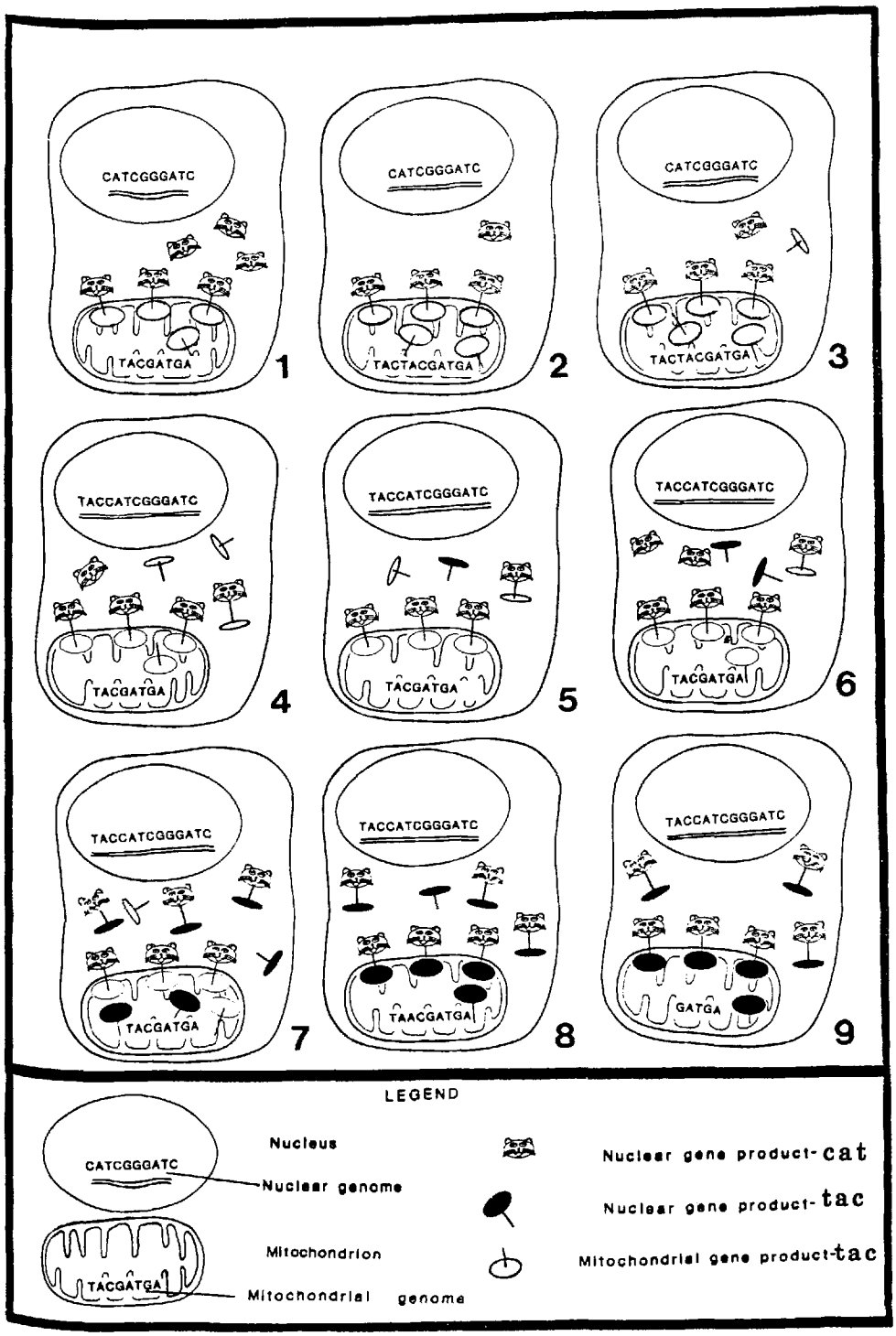

Fig. 1 (1) The nuclear gene product, "cat," and the mitochondrial gene product, "tac," are expressed and functional. (2) The "TAC" gene of the mitochondrial genome is duplicated (cf. Step A in The Model). (3) The mitochondrial gene product tac functions in the nucleocytoplasm (cf. Step B in The Model). (4) A copy of the mitochondrial gene TAC transmigrates to the nucleus (cf. Step C in The Model). (5) The nuclear gene product tac is produced (cf. Step D in The Model). (6) Coordinate regulation of the nuclear genes CAT and TAC evolves (cf. Step E in The Model). (7) The nuclear-gene-produced tac molecules outcompete the mitochondrial-gene-produced tac molecules (cf. Step F in The Model). (8) The mitochondrial copy of the TAC gene is inactivated by mutation. (9) The inactive, mutant mitochondrial gene (now TAAC) is lost (cf. Step G in The Model)

gulis 1976) and "rules for endosymbiont evolution" (Thornley and Harington 1981), which until this analysis had not been used to explain specific cases of molecular evolution.

Although each of the steps in this scenario is essential for what we have termed the "net transfer" of a genetic element, the order of the steps is not. Only the sequence: duplication (of a genetic ele-

ment) – transmigration (of duplicate copies) – competition (between copies integrated within different genomes) – loss (of redundant copy) is crucial. Clearly, each of these processes may have occurred in the absence of (and independent of) the others, but only this order will result in a net transfer.

Aspects of Donor and Recipient in Gene Transfer

A genetic element may be incorporated into a genome in two ways. If a product (RNA, peptide, or protein) of the genetic element has been previously available, the ability of the recipient to process and use this product may precede the gene's integration. An entirely new genetic element, formed by the integration (i.e., recombination) act itself, must be incorporated into a preexisting recipient transcription unit before its potential products have become available to the recipient or been subjected to natural selection. The success of such a genetic element will depend on its expressibility, and on the ease with which its products are processed and used.

For the net transfer of any protein-coding gene to occur, the recipient genome must "speak the donor's language"; in molecular terms this means the recipient must be capable of transcribing the entire gene, processing the messenger, translating it into the correct amino acid sequence, and performing any further processing and transport necessary to the function of the protein. As a result a genome with its transcription and translation closely coupled and without an RNA-splicing apparatus could rarely acquire protein-coding genes from a genome with mosaic genes, uncoupled transcription and translation, and an elaborate splicing apparatus. Thus, with respect to a given donor and a given recipient, the potential for net transfer of protein-coding genes is an all-or-none matter. This means that in the early evolution of the eukaryotic cell, the proto-nucleocytoplasm could be a successful recipient of genetic elements from the proto-mitochondrion, but not vice versa.

The selection pressures on genes for stable RNAs are different, because they produce no protein products for the scrutiny of natural selection. Genes for stable RNAs can transfer from a donor to a recipient genome only if the stable RNAs for which they code are similar enough to be recognized by the recipient protein-synthetic system. Apparently, extremely strong barriers prevent stable RNAs themselves from transferring from their own protein-synthetic system to any other (Gray and Doolittle 1982), which may be viewed as a blockage of Step B in the net transfer of the corresponding genes (see Stable RNA Genes Are Not Transferred, below).

Selection pressures governing the transfer of intervening sequence or "spacer" elements are less severe. The introduction of this type of genetic element will be selectively neutral or only mildly deleterious more often than will that of coding elements. Unless the newly acquired noncoding sequence alters the expression of other genes in the recipient genome, any deleterious effects of the transfer may not become apparent for many generations.

A variety of biochemical traits that would serve as preadaptations for the acquisition of large numbers of new genetic elements (i.e., by a recipient genome) or for the net loss of genetic elements (i.e., by a donor) have been described in the literature, but not in an evolutionary context. We will summarize these below to shed some light on the directional nature of the transfer process. We will then look at functional and genetic preadaptations in the evolving nuclear and mitochondrial genomes, where we will show that vestiges of wholesale transfer of genetic material (both coding and noncoding) may still be observed.

Evidence for the Model

To show how this model may be applied to the nuclear and mitochondrial genomes, we point out that examples of gene duplication (Step A) (Ohno 1970) and of acquisition of the ability to use (and most likely also to regulate) the gene product of a symbiont (Step B) (Fairfield et al. 1983) are known. Transfers of functional genetic elements between genomes of symbiotic or formerly symbiotic partners (Martin and Fridovich 1981) or between coexistent organellar genomes within a single lineage (Farrelly and Butow 1983) (Steps C, D, and E) have been amply demonstrated; it is likely that many further instances will be revealed. Similarly, the presence of pseudogenes in (van den Boogart et al. 1982) and the loss of genetic material by (Borst and Grivell 1978) mitochondrial genomes (Step G) are remarkable but accepted facts.

Thus, the only uncorroborated portion of our scenario is the action of natural selection on the competition between redundant segments of the nuclear and mitochondrial genomes (Step F). This is not a serious flaw: Darwinian evolution acts on any unit capable of replication with high heritability. Therefore it acts on organelles (whether singular, like the nucleus, or part of a population, like a mitochondrion) just as it acts on animals or plants.

How well does this model for the diminution of the coding capacity of the mitochondrial genome agree with present knowledge of the organization and expression of mitochondrial genomes? We now briefly review the relevant literature in the context of this model and its predictions.

Stable RNA Genes Are Not Transferred

We have presented a general mechanism for the transfer of genetic elements of any type (and between genomes in any type of association), but the literature contains no evidence for the transfer of genes for any functional stable RNA species. All the stable RNA genes presently found in mitochondria derive from proto-mitochondrial genomes and not from proto-nucleocytoplasmic genomes; plastids also have acquired no stable RNA genes from the nucleus (Gray and Doolittle 1982). In the case of the mitochondria, at least, the requirement for the retention of a minimal set of these genes has had dramatic effects on the integrity of their genetic codes.

Stable RNA genes may be duplicated (King and Yao 1982), found as pseudogenes (Miyata et al. 1982), shuffled around a given genome, or lost altogether (Van Etten et al. 1980; Wallace 1982). They may also move between genomes, as evidenced by the region of noncoding DNA in maize mitochondria with strong homology to plastid tRNA and rRNA (Stern and Lonsdale 1982). That transferred genes for stable RNAs are not expressed means that the sequence of steps necessary for net transfer is blocked at the "use" step (Step B), at the "activation" step (Step D), at the "regulation" step (Step E), or at some combination of these. A recipient can preadapt itself for the acquisition of a protein-coding gene by finding a use for the gene product (Step B).[1] Apparently no analogous mechanism exists for assimilating the stable RNAs, which may mean that biochemical barriers make this process vastly more difficult.

In the same way that our model accounts for the transfer of protein-coding genes from mitochondria to nuclei, the "blockage" of steps in the case of stable RNA genes, discussed above, accounts for the net loss of these genes from proto-mitochondrial genomes. Mitochondrial genomes generally contain only one tRNA gene for each set of synonymous codons. When such a genome loses the ability to produce a given tRNA species, survival depends on assumption of the lost function by tRNAs with anticodons similar to the lost anticodon (in competition with each other, of course) (Garen 1968). The situation immediately following loss of a tRNA gene would be one of directed evolution of the tRNA gene pool, and would favor the expression of little-used and mutant tRNAs (i.e., hybrid products of recombinatory mismatches) (cf. Munz et al. 1982).

In other words, the loss of a tRNA gene from a given genome would be observed as the loss of the corresponding anticodon from the genetic code used by that genome. Evidence of this phenomenon is most likely to be observed in populations of genomes that contain relatively few protein-coding genes and genomes that produce gratuitous or redundant proteins.

The divergence of the mitochondrial genetic codes from those of all other types of genome and the divergence of the mitochondrial codes of different lineages from each other (Gray and Doolittle 1982) may therefore be viewed as the consequence of selection pressures for the loss of tRNA genes. Differences in codon usage (Grantham et al. 1980) and in the stabilities of different tRNA genes could have caused the loss of different tRNA genes by divergent mitochondria, followed by different compensating changes in the molecular biologies of these organelles. It is also possible that some (or all) of the short, highly structured mitochondrial tRNA genes were preadapted for eventual net loss as a result of playing some role in the net transfer of protein-coding genes to the nucleus. These mitochondrial tRNA genes may have been especially mobile genetic elements even before the mitochondrial codes diverged, despite the fact that they do not seem to be expressed by the recipient genomes.

Harington and Thornley (1982) emphasize the problem of explaining the evolution of transport of translation products to organelles without special receptors, a mechanism involved in Step E above; to this may be added the related problem of explaining what aspects of existing transport mechanisms have allowed genes that code for mRNA, but not genes for tRNA or rRNA to undergo net transfer to the nucleus. A potential solution to both questions may come from studying a case of genetic elements moving to the mitochondria: organellar introns.

Intron Evolution

Three general classes of RNA-splicing reactions are known (Cech 1983), corresponding to three general types of intervening sequence:

1. Nuclear tRNA introns, which are excised by specific endonucleases; the excision reaction is independent of the sequence and structure of the introns, but dependent on the structure of the exons.

2. Nuclear and plastid mRNA introns, the excision of which requires the presence of specific sequences within the introns, small nuclear RNA cofactors, and specific endonucleases.

3. Nuclear rRNA and mitochondrial mRNA and rRNA introns, which play a prominent role in their own excision. In at least one case (discussed below) a nuclear rRNA intron can be excised in the absence of endonucleolytic proteins (Zaug et al. 1983). In-

[1] For an example of malaria parasites doing just this with host superoxide dismutase, see Fairfield et al. (1983).

trons of this type have been further classified as belonging to Group I or Group II (see, for example, Michel et al. 1982).

All introns tend to have compact secondary and tertiary structures, and, by definition, participate in RNA-splicing reactions—traits that may be considered preadaptations for Steps C, D, and G in our model.

The potential for net transfer of intron sequences must be assessed separately for introns of the three classes. By using the framework of our model we see that net transfer of an intron of the first class will occur only when the intron is inserted at the "correct" site of the tRNA (or tRNA-like) recipient sequence and the required endonucleases are available previous to the transmigration (Step C).

Net transfer of introns of the second class similarly requires preavailability of endonucleases and snRNAs. It is rather difficult to conceive of a selective advantage for the recipient genome to acquire these splicing endonucleases and/or small RNAs unless the recipient already harbored introns of one or both of these classes.

In contrast, net transfer of introns of the third class requires no such preadaptations; all the information needed for the use of the transferred intron may be contained within the intron at the time of transfer. This special trait accounts for many of the singular aspects of the mitochondrial introns: their mobility, their generally "facultative" status, their high degree of conservation, and their ability to become "obligate" or "stuck" and encode portions of maturase proteins.

Facultative Mitochondrial Introns

Three types of evidence indicate that in their present evolutionary place, most mitochondrial introns are facultative.

First, if one compares the sequences of a common mitochondrial gene, such as that for cytochrome b, from a variety of organisms, such as *Homo sapiens, Saccharomyces cerevisiae, Aspergillus nidulans,* and *Neurospora crassa* (Burke et al. 1984), a variety of gene organizations is found. Despite the high degree (>50%) of similarity among the sequences of the encoded proteins, there are two introns in the *S. cerevisiae* gene (short form) (Borst and Grivell 1981a, b), two in the *N. crassa* gene, one in the *A. nidulans* gene, and none in the animal gene. Although the *A. nidulans* intron is homologous in sequence and location to intron 2 of the *S. cerevisiae* gene (short form), the two introns in the *N. crassa* gene are located at different sites from those of the introns in the *S. cerevisiae* gene, and in different reading frames.

Second, in an elegant set of experiments, Labouesse and Slonimski (1983) used intronless revertant strains and novel intronless constructions to demonstrate that all the introns in the *S. cerevisiae* cytochrome b gene may be removed without impairing the function of the resultant cytochrome b protein.[2]

Third, it is highly significant that these introns all share several long conserved structural and coding sequences (Waring et al. 1983), and many code for at least portions of maturase proteins that assist in intron removal.

Mitochondrial introns exhibit splicing intermediates (Tabak, et al. 1984) and/or in vitro self-splicing abilities (Garriga and Lambowitz 1984) similar to those of the self-splicing *Tetrahymena* introns. In vitro self-splicing of different mitochondrial introns depends on protein cofactors to different extents, even in the same reaction mixtures (Garriga and Lambowitz 1984).

Clearly, if these introns have no functions other than the processing of intron-containing mRNAs, the interruption of mitochondrial genes by introns could not have preceded the existence of the ability to process the mosaic RNA transcripts. Since the maturase enzymes require internal intron sequences for processing, there would have been no selective advantage for the mitochondrial genome having a processing system prior to its acquisition of introns. We therefore conclude that both introns and processing capability must have been acquired in the same generation.

There are four obligatory steps (A, C, F, and G) in the above model, and it follows that the genetic elements most likely to undergo net transfer are those most likely to bypass the nonobligatory steps. An acquired element that interrupts a recipient genome but contains all the information necessary for its own excision would exempt the recipient genome from fulfilling Steps B, D, and E. This hypothetical genetic element very closely resembles three recently discovered genetic elements: (a) the mitochondrial plasmid of *Podospora anserina,* which is both the intron of the cytochrome c oxidase (subunit I) gene (Osiewacz and Esser 1984) and the infective agent causing senescence (during which it is integrated into the *Podospora* nuclear genome) (Wright and Cummings 1983); (b) the mitochondrial plasmid of *N. crassa,* which has been termed a "free-standing intron" (Nargang et al. 1984); and (c) the intervening sequence of the (nucleus-encoded) large rRNA of *Tetrahymena thermophila* and *Tetrahymena pigmentosa,* which specifically and autocatalytically

[2] In this vein it should be noted that intronless nuclear genes are rarely functional; generally they atrophy (Wilde et al. 1982) (also see Step F in The Model, above).

excises itself from the rRNA transcript (Zaug et al. 1983). This last intron is absent from some species of *Tetrahymena* (and may therefore also be considered facultative), and has been shown to share several highly conserved sequences with all known mitochondrial introns (Waring et al. 1983), as well as with at least one of the mitochondrial plasmids (Nargang et al. 1984).

These relationships allow all these sequences to be placed in the third class of introns described above. They also indicate a common origin and a common method of splicing, and suggest that some sequences ancestral to these introns moved fluidly between nuclei (where they would be forced to compete with introns of the first two classes) and mitochondria. The absence of any introns from, e.g., animal mitochondria, the selective advantages self-excising introns could have for a small, dwindling genome, and the relatedness of all the observed mitochondrial introns strongly suggest that these mobile elements did not originate in the mitochondria.

The recently discovered *Neurospora* and *Podospora* mitochondrial plasmids are undoubtedly similar in structure and function(s) to the most recent common ancestor of the extant *Tetrahymena* large-subunit rRNA introns and the mitochondrial introns. A likely site for such a sequence to have first been inserted into a recipient mitochondrial genome would have been in a large rRNA gene (e.g., Faye et al. 1979; Hahn et al. 1979), because the presence of an intron in the large rRNA transcripts of mitochondria may confer a selective advantage in the formation of the mature gene product, as first suggested by Apirion (1983). The adaptability and transferability of the *Tetrahymena* intron/fungal plasmid make it an excellent prototype for our model of genetic-element transfer. Many of the experimental observations already mentioned concern the functional divergence of the transferred sequences, which in the case of fungal mitochondria may reveal some especially interesting things about the nature of the (recipient) genomes prior to and during the net transfer process.

Divergent Mitochondrial Introns

Introns in the nuclear rRNA genes are capable of effecting their own excision in the absence of other protein or nucleic acid cofactors. Maturase-encoding mitochondrial introns must be translated for the splicing reactions to occur. How could the splicing of these two very different types of genes have diverged from a common method of splicing?

A clue may lie in an observation of Borst and Grivell (1981a, b). When a maturase is encoded by adjacent intronic and exonic portions of a gene, mutations in the exon do not affect the function of the maturase product, whereas mutations in the intron do. If the exonic portion of the gene (which also encodes another protein) does not code for an essential portion of the maturase protein, then the intronic portion of the maturase gene must have evolved the maturase-encoding function.

We suggest that after the ancestral self-excising intron/plasmid sequence (with no coding function) became a heritable part of a mitochondrial cistron, the mosaic mRNA was translated by mitochondrial ribosomes regularly enough for the hybrid protein product to evolve into a functional maturase under the influence of natural selection. Once this type of maturase-encoding mosaic had evolved, there was no longer a strict requirement for the intron to process itself; the facultative, autocatalytic intron had reached an evolutionary dead end as an obligate part of a protein-encoding gene.

We believe that after most of the essential elements of the mitochondrial genome had been transferred to other genomes or lost, the reduced genome that remained was preadapted for preservation by the introduction of "free-standing," autoexcising intronic sequences. The subsequent stepwise evolution of certain of these sequences into maturase-encoding genes, or incompetent self-splicers, may be viewed as the nonconcerted evolution of a family of genetic elements. The ancestral genetic element from which this family evolved did not originate in the mitochondrial genome, but it is unclear just how and where it might have arisen.

Conclusions

Nonsymbiotic explanations of mitochondrial and nuclear molecular-biological features are not associated with any clear stepwise models. When endosymbiotic origin is assumed, these features become understandable results of a trend toward "intimacy between strangers" consisting of gene duplication, net transfer, integration, and preservation by natural selection. The presence of redundant genetic information led to specialization and emergence of new functions. Fusion of once independent genomes has been limited by the original historical constraints on the partners—an archaebacterium with introns, and a respiring eubacterium that lacked them. When we understand the evolutionary history of the organellar genomes in the context of the organisms within which these genomes evolved their peculiarities, we shall be able to reconstruct the functions and mechanisms of action of the elaborately detailed differences in genome organization made evident by the powerful techniques of molecular biology.

Furthermore, by outlining some of the steps it becomes clear that a richer understanding of the events of mitochondrial evolution may be obtained from an analysis of the roles of competition and natural selection in the observed evolutionary trends. The use of this method leads to the conclusions that virtually all the recent revelations about mitochondrial molecular biology are consistent with and derivable from the idea that the mitochondrion and the nucleus originated as independent genetic entities, and that the functions of a gene and its products, and not merely the intracellular location of the gene, supply valuable clues to the evolutionary stories of the genomes involved.

Acknowledgements. We would like gratefully to thank the following people for the editorial efforts and intellectual input they contributed to this paper: Dr. Lynn Margulis, Dr. Scott Mohr, Dr. Richard Laursen, Dr. Betsey Dyer, Mr. David Bermudes, Ms. Donna Mehos, and Mr. Stuart Brown, as well as Mr. Steven Alexander, for artistic expertise. This work was supported in part by a grant from the U.S. National Science Foundation (PCM82-03004).

Glossary

Donor: A discrete genetic system that undergoes the net loss of a genetic element or elements.

Exon: A discrete, continuous nucleic acid sequence that constitutes one of the coding regions of a cistron, with respect to a specific gene product.

Facultative: Not essential for the survival of a genome and its genetic system.

Genetic element: An independent nucleic acid sequence; the unit of net transfer between genomes.

Intron (intervening sequence): A discrete, continuous nucleic acid sequence that interrupts the coding regions of a cistron, with respect to a specific gene product.

Maturase: An RNA processing enzyme, often containing translation products of mosaic genes.

Mosaic: A nucleic acid sequence made up of intronic and exonic regions.

Obligate: Essential for the survival of a genome and its genetic system.

Preadaptation: A trait selected for under one set of conditions that increases the probability of survival under changed conditions.

Processing: Specific excision of sequences from an RNA.

Proto-mitochondrion: A *Paracoccus*-like genetic system, ancestral to extant mitochondrial systems, that entered into a symbiosis with a proto-nucleocytoplasmic bacterium.

Proto-nucleocytoplasm: An archaebacterium-like microbial genetic system, ancestral to the modern-day nucleocytoplasmic system, that entered into a symbiosis with a proto-mitochondrion.

Recipient: A discrete genetic system that undergoes the net gain of a genetic element or elements.

snRNA: Small nuclear RNA. Some of these play a part in the processing of nuclear (and presumably plastid) RNA transcripts.

Stable RNA: Transfer RNA (tRNA), ribosomal RNA (rRNA), or their unprocessed precursors. Note: For simplicity, no other types of RNA (such as stable messenger RNA) are included in this definition.

References

Apirion D (1983) RNA processing in a unicellular microorganism: implications for eukaryotic cells. Prog Nucleic Acid Res Mol Biol 30:1–40

Borst P, Grivell L (1978) The mitochondrial genome of yeast. Cell 15:705–723

Borst P, Grivell L (1981a) One gene's intron is another gene's exon. Nature 289:439–440

Borst P, Grivell L (1981b) Small is beautiful—portrait of a mitochondrial genome. Nature 290:443–444

Burke JM, Breitenberger C, Heckman JE, Dujon B, RajBhandary UL (1984) Cytochrome b gene of *Neurospora crassa* mitochondria. J Biol Chem 259:504–511

Cech TR (1983) RNA splicing: three themes with variations. Cell 34:713–716

Fairfield AS, Meshnick SR, Eaton JW (1983) Malaria parasites adopt host cell superoxide dismutase. Science 221:764–766

Farrelly F, Butow RA (1983) Rearranged mitochondrial genes in the yeast nuclear genome. Nature 301:296–301

Faye G, Dennebouy N, Kujawa C, Jacq C (1979) Inserted sequence in the mitochondrial 23S ribosomal RNA gene of the yeast *Saccharomyces cerevisiae*. Mol Gen Genet 168:101–109

Fox GE, Stackebrandt E, Hespell RB, Gibson J, Maniloff J, Dyer TA, Wolfe RS, Balch WE, Tanner RS, Magrum LJ, Zablen LB, Blakemore R, Gupta R, Bonen L, Lewis BJ, Stahl DA, Luehrsen KR, Chen KN, Woese CR (1980) The phylogeny of prokaryotes. Science 209:457–463

Garen A (1968) Sense and nonsense in the genetic code. Science 160:149–159

Garriga G, Lambowitz AM (1984) RNA splicing in *Neurospora* mitochondria: self-splicing of a mitochondrial intron in vitro. Cell 38:631–641

Grantham R, Gautier C, Gouy M, Mercier R, Pare A (1980) Codon catalog usage and the genome hypothesis. Nucleic Acids Res 8(1):r49–r62

Gray MW (1983) The bacterial ancestry of plastids and mitochondria. BioScience 33:693–699

Gray MW, Doolittle WF (1982) Has the endosymbiont hypothesis been proven? Microbiol Rev 46:1–42

Hahn U, Lazarus CM, Lunsdorf H, Kuentzel H (1979) Split gene for mitochondrial 24S ribosomal RNA of *Neurospora crassa*. Cell 17:191–200

Harington A, Thornley AL (1982) Biochemical and genetic consequences of gene transfer from endosymbiont to host genome. J Mol Evol 18:287–292

John P, Whatley FR (1977) *Paracoccus denitrificans* and the evolutionary origin of the mitochondrion. Adv Botan Res 4: 51–115

Kaine B, Gupta R, Woese CR (1983) Putative introns in tRNA genes of prokaryotes. Proc Nat Acad Sci USA 80:3309–3312

King BO, Yao M (1982) Tandemly repeated hexanucleotide at *Tetrahymena* rDNA free end is generated from a single copy during development. Cell 31:177–182

Labouesse M, Slonimski PP (1983) Construction of novel cytochrome b genes in yeast mitochondria by subtraction or addition of introns. EMBO J 2:269–276

Margulis L (1976) Genetic and evolutionary consequences of symbiosis. Exp Parasitol 39:277–349

Martin JP Jr, Fridovich I (1981) Evidence for a natural gene transfer from the ponyfish to its bioluminescent bacterial symbiont *Photobacter leiognathi*. J Biol Chem 256:6080–6089

Michel F, Jacquier A, Dujon B (1982) Comparison of fungal mitochondrial introns reveals extensive homologies in RNA secondary structure. Biochimie 64:867–881

Miyata T, Kikuno R, Ohshima Y (1982) A pseudogene cluster in the leader region of the *Euglena* chloroplast 16S–23S rRNA genes. Nucleic Acids Res 10:1771–1780

Munz P, Amstutz H, Kohli J, Leupold U (1982) Recombination between dispersed serine tRNA genes in *Schizosaccharomyces pombe*. Nature 300:225–231

Nargang FE, Bell JB, Stohl LL, Lambowitz AM (1984) The DNA sequence and genetic organization of a *Neurospora* mitochondrial plasmid suggest a relationship to introns and mobile elements. Cell 38:441–453

Ohno S (1970) Evolution by gene duplication. Springer-Verlag, Berlin

Osiewacz HD, Esser K (1984) The mitochondrial plasmid of *Podospora anserina:* a mobile intron of a mitochondrial gene. Curr Genet 8:299–305

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

Tabak HF, Van der Horst G, Osinga KA, Arnberg AC (1984) Splicing of large ribosomal precursor RNA and processing of intron RNA in yeast mitochondria. Cell 39:623–629

Thornley AL, Harington A (1981) Diploidy and sex as the selective advantages for retaining genes transferred from mitochondrial and plastid ancestors in the nuclear genome. J Theor Biol 91:515–523

van den Boogart 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 Etten RA, Walberg MW, Clayton DA (1980) Precise localization and nucleotide sequence of the two mouse mitochondrial rRNA genes and three immediately adjacent novel tRNA genes. Cell 22:157–170

Vidal G (1984) The oldest eukaryotic cells. Sci Am 250(2): 48–57

Wallace DC (1982) Structure and evolution or organelle genomes. Microbiol Rev 46:208–240

Waring RB, Scazzocchio C, Brown TA, Davies RW (1983) Close relationship between certain nuclear and mitochondrial introns. J Mol Biol 167:595–605

Wilde CD, Crowther CE, Cowan NJ (1982) Diverse mechanisms in the generation of human β-tubulin pseudogenes. Science 217:549–552

Wright RM, Cummings DJ (1983) Integration of mitochondrial gene sequences within the nuclear genome during senescence in a fungus. Nature 302:86–88

Zaug AJ, Grabowski PJ, Cech TR (1983) Autocatalytic cyclization of an excised intervening sequence RNA is a cleavage–ligation reaction. Nature 301:578–583

Received April 26, 1985/Revised May 30, 1985