Promiscuous DNA Leaps All Barriers

Jumping seems to be in the nature of genes, not just within the nucleus but between the genomes of mitochondria and chloroplasts too

According to the endosymbiont hypothesis for the origin of eukaryotic cells, mitochondria and chloroplasts arose from free-living prokaryotes that early in evolution entered and established a symbiotic relationship with a progenitor of the eukaryotic cell. Because the genomes of these organelles—mitochondria and chloroplasts—now serve only a small proportion of their genetic needs, with the remainder provided by genes in the nucleus, the hypothesis has to allow for the transfer of genes from these organelles to the nucleus since the first colonization. The first firm evidence of the movement of DNA between these various genomes—in the mitochondrion, the chloroplast, and the nucleus—has just been reported.

Toward the end of last year David Stern and David Lonsdale, of the Plant Breeding Institute, Cambridge, England, described their discovery, in maize mitochondria, of a 12-kilobase segment of DNA that is closely homologous to a section of the chloroplast genome in the same organism (*1*). The most attractive conclusion here is that the chloroplast DNA became incorporated into the mitochondrial genome after a duplication and transposition event. Commenting on the finding, John Ellis of the University of Warwick, England, designated the event as the result of "intracellular promiscuity" (*2*).

By a curious coincidence, Frances Farrelly and Ronald Butow, of the University of Texas Health Sciences Center at Dallas, invoked the same term—promiscuous DNA—to describe a similar phenomenon in yeast. They recently reported on a piece of DNA common to the mitochondrial and nuclear genomes of *Saccharomyces cerevisiae* (*3*). In this case, the DNA segment in the nucleus bears a striking signature of its mitochondrial origin. "This is clear evidence on the direction of flow of DNA between the nucleus and the mitochondrion," says Butow.

A third example comes from Roy Britten and Eric Davidson's group at the California Institute of Technology. They have found sections of mitochondrial DNA in the nuclear genome of the sea urchin *Strongylocentrotus purpuratus* (*4*). Invoking a transposition event perhaps some 25 million years ago, they too characterize the DNA as promiscuous.

The transfer at an early stage of evolution of genes from the mitochondrion to the nucleus has, as already mentioned, been inferred from the modern organelle's dependence on the nucleus for much of its genetic input. Although the figures vary between species, 90 to 95 percent of polypeptides utilized in the mitochondrion are encoded in the nucleus and imported into the organelle after manufacture in the cytoplasm. Until now experimental support for the notion of gene transfer from the nucleus to the mitochondrion early in evolution has been persuasive but indirect, such as that reported in the middle of last year by Paul van den Boogaart, John Samallo, and Etienne Agsteribbe of the State University, the Netherlands (*5*).

> The first firm evidence of the movement of DNA between these various genomes. . . .

In *Saccharomyces cerevisiae* the mitochondrial adenosinetriphosphatase (ATPase) proteolipid is encoded in the mitochondrial genome, whereas in *Neurospora crassa*, a related ascomycete, the gene is part of the nuclear genome. What Boogaart and his colleagues did was to show that the mitochondrial genome does indeed contain a copy of the gene. Surprisingly, however, the homology between the *Neurospora* genes is much less than that between the nuclear counterpart and mitochondrial ATPase genes of other organisms. The clear implication is that a duplication and transposition even shifted the gene to the nucleus, leaving behind the mitochondrial copy, whose function is presently obscure. However, these data do not rule out the reverse process—transfer from the nucleus to the mitochondrion—and so the endosymbiont hypothesis does not here derive the firm support it requires. The results from Butow's laboratory do provide this support, even though they do not involve an intact gene.

As so often happens, the most surprising results are chanced upon by sheer good fortune. So it was when Butow and Farrelly found a mosaic of four pieces of yeast mitochondrial DNA stitched together and integrated in the organism's nucleus.

One piece of the mosaic in the nucleus is part of the varl gene, whose complete product when translated is a component of the small ribosomal subunit in the mitochondrion. The gene for the equivalent protein in a number of lower organisms appears to be located in the mitochondrial genome. So, when in 1980 Butow began his search for the varl gene, he fully expected to find it in the mitochondrion.

Although genetic data pointed to a region of the mitochondrial genome as being important in the production of varl, sequencing data and common sense seemed to argue for a different location for the structural gene. "We knew that this varl determinant region was extremely AT-rich, with a concentrated cluster of GC's [guanine-cytosine]," explains Butow. "This is very typical of mitochondrial spacer DNA, not of mitochondrial genes. Moreover, the sequencing data, which had first been reported by another laboratory, showed no reading frame in the region."

The Dallas researchers therefore concluded that the varl determinant region in the mitochondrial genome was some kind of regulatory element. So, using part of the region as a probe, they began the search for the true gene in the nucleus. "Our first hybridization experiments lit up a fragment from the nuclear genome," remembers Butow. The experiment, it seemed, was a success. However, what appeared to be the end of one search turned out to be the beginning of a different, and unexpected, one.

In parallel with the hybrization experiments, Michael Hudspeth and Lawrence Grossman, at the University of Michigan, who had been collaborating with the Dallas team on the varl problem, obtained new sequence data on the varl region. These data showed that afterall there was a reading frame that had been missed in the earlier sequence. Moreover, the DNA codons in the sequence corresponded to the amino acid sequence of the varl protein closely enough to convince Butow that the gene he

sought was indeed in the mitochondrial genome (*6*). What, then, were the varl-like sequences in the nucleus?

One obvious answer was that there is a second copy of the varl gene in the nuclear genome. In fact, only a fragment of the gene is present. Before sequencing the nuclear segment containing the gene fragment, Butow and his colleagues decided to determine whether there was any other mitochondrial DNA associated with the varl sequences. They therefore hybridized the nuclear segment back to the complete range of mitochondrial DNA and, sure enough, picked up interaction with sequences at the 3′ end of the cytochrome b gene. This was something of a surprise as the gene for cytochrome b is not adjacent to varl on the mitochondrial genome.

Sequencing of the nuclear fragment showed it to be homologous with about 115 base pairs from within the central section of the varl gene. Upstream from this are about 123 base pairs homologous to the 3′ end of the cytochrome b gene, specifically the last 64 bases of exon 6 and a run past the translation terminator. And adjacent to this, it turned out, is a short GC-rich cluster that is homologous to a section of intron 5 of the cytochrome b gene. Completing the curious mosaic is a 29-base-pair, AT-rich sequence, 26 of which are identical to palindromic sequence that can form a hairpin loop thought to be important as part of an origin of replication in yeast mitochondrial DNA.

At first sight the mosaic of mitochondrial gene fragments appears somewhat bizarre. In fact, it can readily be explained as a product of a phenomenon typical of yeast, that is, the formation of the so-called petite phenotype.

With its high proportion of AT-rich regions and GC clusters, the yeast mitochondrial genome is especially susceptible to internal recombination, which brings together sequences normally separated from each other. As a result, large sections of the genome may be lost and the mosaic remnant, with sometimes as few as 68 base pairs, may be greatly amplified. The petite cell, which occurs at a frequency of about 1 percent per generation, is unable to manifest the normal functions carried out by mitochondria—it is respiratory-incompetent.

The mosaic that Butow and Farrelly found in the yeast nucleus is, they suggest, a typical product of the petite phenomenon. "You can see the regions of homology between the middle of the varl gene and the 3′ end of the cytochrome b gene that would bring them together as we find them in the nucleus," says Butow. "Additional recombination could bring together the short GC-rich region from intron 5 and part of exon 6."

This varl mosaic may well have no functional significance in the nucleus. Indeed, given its unusual structure, the chances of its having a function involving transcription are very slim indeed. "Its presence marks an event in history, a transfer that just happened to be stable once it was integrated into the nuclear genome. They are mitochondrial sequences fossilized in the nuclear genome. Nevertheless, the fact that they are there does demonstrate the phenomenon we always assumed happened but had not seen—that is, the transfer of genetic material from the mitochondrion to the nucleus." The mechanism of transfer, and the selection pressures under which the balance of genetic power in the cell shifted to the nucleus, remain to be established.

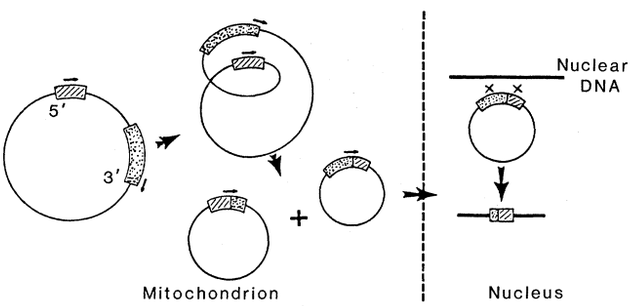
Meanwhile, varl, the gene for which Butow began his search, turns out to be as interesting in itself as does its fragmented echo in the nucleus (*6*). For a start, its use of codons is much more akin to that in free-standing and intron-encoded open reading frames than it is to codon use in mitochondrial genes. And, with a composition of virtually 90 percent AT, it is the most AT-rich gene known in nature. Moreover, about one-third of its GC residues are concentrated in a single 46-base-pair cluster. "The structure is typical of what is considered to be spacer DNA in mitochondrial genomes," says Butow. "We think the gene evolved as an open reading frame within an AT-rich region. This primitive gene encodes a protein, of which 30 percent of its amino acids are asparagine, which binds with ribosomal RNA."

Butow points to the 46-base-pair GC cluster in varl and notes the presence of an identical sequence near to, but outside the coding region of, an ATPase gene. "It gives you the idea that the varl gene has 'recruited' the sequence, if I can use such a phrase, to become part of its coding structure. In one place you have a sequence that codes for something while in another place the same sequence is noncoding."

The homology between the sequences in the middle of varl and at the 3′ end of the cytochrome b gene inspire the same interpretation of "sequence recruitment," suggests Butow. "In varl, the sequence is a coding region whereas in the cytochrome b gene it overlaps a coding and a noncoding region. This is quite extraordinary."

Butow and his colleagues point out that some 30 percent of the yeast mitochondrial genome is as yet uncharted for coding sequences, because its high proportion of AT residues appears to exclude this possibility. There may indeed be other primitive and extraordinary genes lurking there.

It is perhaps apposite that the search for a primitive gene should lead, fortuitously, to the first positive signature of the ancient process of gene transfer from the mitochondrion to the nucleus.



Petite DNA

Homologous regions in the varl gene (stripes) and the cytochrome b gene (stippled) allow recombination with the loss of large segments of mtDNA. Parts of the petite DNA so formed might then be incorporated into the nuclear genome.

This result, together with those on maize and sea urchin, surely portends the discovery of many more examples of the transposition of genes and gene fragments between the various cellular genomes. It is evident too that this phenomenon was not restricted to the distant times of the Pre-Cambrian: DNA promiscuity is a constant and continuous process through all evolutionary time. And, as Davidson and his colleagues point out, the extraordinary degree to which the nuclear genome is a patchwork of DNA transpositions, from both within the nucleus and without, is only now beginning to become fully evident.—**ROGER LEWIN**

References

1. D. B. Stern and D. M. Lonsdale, *Nature (London)* **299**, 698 (1982).
2. J. Ellis, *ibid.*, p. 678.
3. F. Farrelly and R. A. Butow, *ibid.* **301**, 296 (1983).
4. H. T. Jacobs *et al.*, *J. Cell. Biochem.* **6** (Suppl.), 817 (1982); H. T. Jacobs, J. W. Posakony, J. W. Grula, J. W. Roberts, Ji-Hou Xin, R. J. Britten, E. H. Davidson, *J. Mol. Biol.*, in press.
5. P. van den Boogaart, J. Samallo, E. Agsteribbe, *Nature (London)* **298**, 187 (1982).
6. M. E. S. Hudspeth, W. M. Ainley, D. S. Shumard, R. A. Butow, L. I. Grossman, *Cell* **30**, 617 (1982).