The origins of replication of the mitochondrial genome of yeast

Giorgio Bernardi

Recent investigations have provided information on the origin of replication of the mitochondrial genome of yeast and an explanation for the phenomenon of the suppressivity.

Three years ago, I summarized in a *TIBS* review[1] the state of our knowledge on the 'petite colonie' mutation of *Saccharomyces cerevisiae*. As shown by the pioneering work of Ephrussi and his collaborators, this mutation: (a) is characterized by an irreversible loss of respiration and by an extraordinarily high spontaneous mutation rate; and (b) is transmitted to the progeny in a non-Mendelian fashion: crosses of wild-type cells with petite mutants yield either wild-type progeny only, or both wild-type cells and petite mutants in proportions essentially dependent on the particular petite used; in the first case, the petites entering the cross are called neutral, in the second one suppressive. While the molecular basis of the mutation was understood by 1979, its details were not known then, and suppressivity still was the same elusive problem it had been for almost 25 years. The situation is now completely changed and it is possible to give here a full account of the petite mutation.

The excision-amplification process

Previous investigations (see Ref. 1 for a brief review) had established that the first event in the spontaneous cytoplasmic petite mutation is the excision of a segment from one of the 25–50 mitochondrial *genome units* (Fig. 1) of a wild-type cell; excision is then followed by a tandem amplification process in which the excised DNA segment becomes the *repeat unit* of a *defective genome unit* (Fig. 2) which replicates within the parental wild-type cell and segregates into the buds of the latter; further segregation of these defective genome units in the progeny leads to the formation of petite mutants, whose mitochondrial genome is formed by identical units.

The sequence used in the excision process have now been investigated for 17 spontaneous petite genomes[2,3]. In all cases, perfect direct repeats located in the AT spacers or in the GC clusters were found to be used as excision sequences, as predicted a long time ago (see Ref. 1).

Giorgio Bernardi is at the Laboratoire de Genetique Moleculaire, Institut de Recherche en Biologie Moleculaire, 2, Place Jussieu, 75005 Paris, France.

These results indicated that sequences of 11–12 base pairs could be used as excision sequences; longer excision sequences were also used and included both canonical and surrogate origins of DNA replication (see below); when shorter sequences were used, they were flanked by regions of patchy homology (Fig. 2).

These results, discussed in more detail elsewhere[2,3], also indicated: (a) that the excision mechanism probably involves unequal, site-specific, crossing-over events within a genome unit and that this process probably is just a special case of the very active recombination processes taking place in the mitochondrial genomes of wild-type yeast cells[4]; (b) that the highest excision rates are associated with sequences capable to form the most stable (longest and/or richest in GC) heteroduplexes; an extreme case involving two *ori* sequences (see below) was described[2]; (c) that the production of petite mutants depends not only upon the excision rate, but also upon the replication rate of the defective genome relative to the parental one, and upon its stability, namely its susceptibility to further excisions, in the cells harboring it[2]; both these parameters favor the production of petite mutants carrying mitochondrial genomes formed by short repeat units containing canonical *ori* sequences (see below). Obviously this trend can be counteracted by genetically selecting petites having a very poor replication efficiency and/or resulting from very rare events; such is the case, for the ethidium-induced petites studied in Tzagoloff's laboratory[5,6]. Here too, direct repeats were used as excision sequences, but they were found to be localized not only in non-coding sequences, but also in the open reading frames of introns and in the exons of mitochondrial genes.

Excision and amplification are general phenomena

It should be stressed that excision and amplification of defective genome units from wild-type mitochondrial genomes are general phenomena not limited to yeast. In the case of obligatory aerobes, however, defective units coexist with wild-type genome units, a certain number of which are required for respiration. The best-known examples[2] are those of senescent cultures of *Podospora anserina*, of 'ragged' mutants of *Aspergillus amstelodami*, of *poky* and *stopper* mutants of *Neurospora crassa*, of male sterile mutants in maize. Furthermore, excision and amplification do not concern only mitochondrial genomes; some bleached mutants of

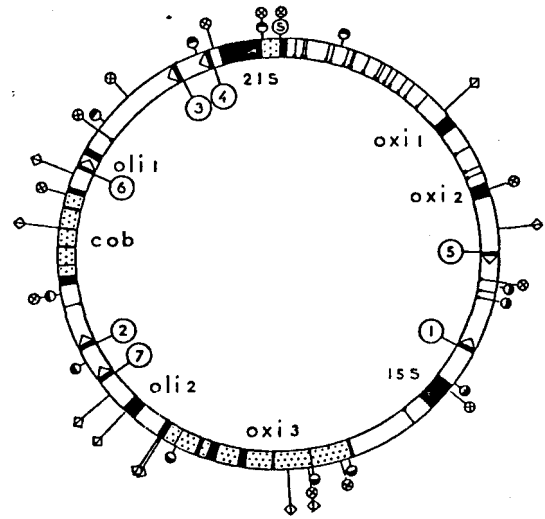

Fig. 1. Physical and genetical map of the mitochondrial genome unit of wild-type yeast (strain A). Some restriction sites are indicated: Hinc II (⊗), Hha I (●), EcoRI (◇) Sal I (Ⓢ). Circled numbers indicate the location of ori sequences 1–7 (arrowheads point in the direction cluster C to cluster A; see Fig. 3). Black and dotted areas correspond to exons and introns of mitochondrial genes, respectively. Thin radial lines indicate tRNA genes. White areas correspond to long AT spacers embedding short GC clusters. (Modified from Ref. 16.)

 0376 – 5067/82/0000 – 0000/$01.00

Euglena gracilis contain defective chloroplast genomes in which the ribosomal gene region and the origin of replication are preferentially retained; since the chloroplast genome is dispensable in *Euglena*, as is the mitochondrial genome in yeast, these mutants only contain the defective genome[7]. Finally, *Drosophila* and mammalian cells are known to contain extrachromosomal circular DNA excised from nuclear DNA[8,9], and double-minute chromosomes are known to be formed in methotrexate-resistant mouse cells[10].

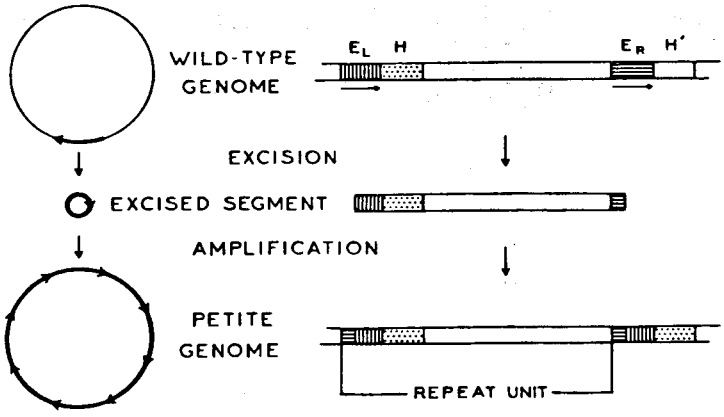

Fig. 2. (A) Scheme depicting the excision–amplification process leading to the formation of the genome of a spontaneous petite mutant. A segment of a unit of a wild-type mitochondrial genome is excised and tandemly amplified into a defective genome unit. This then replicates and segregates into the buds to form the genome of a petite mutant; the petite genome can undergo further excisions leading to the formation of secondary petite genomes. (B) Scheme showing the left and right (E_L, E_R) excision sequences as found on the parental wild-type genome region from which the repeat unit of the petite genome was excised. H, H' indicate sequences flanking the excision sequences and sharing a significant yet imperfect homology (from Ref. 3).

The canonical and the surrogate origins of replication of petite genomes

The mitochondrial genomes of the vast majority of spontaneous petites are exclusively derived from the tandem amplification of a DNA segment excised from any region of the parental wild-type genome[11]. Therefore, either the wild-type genome

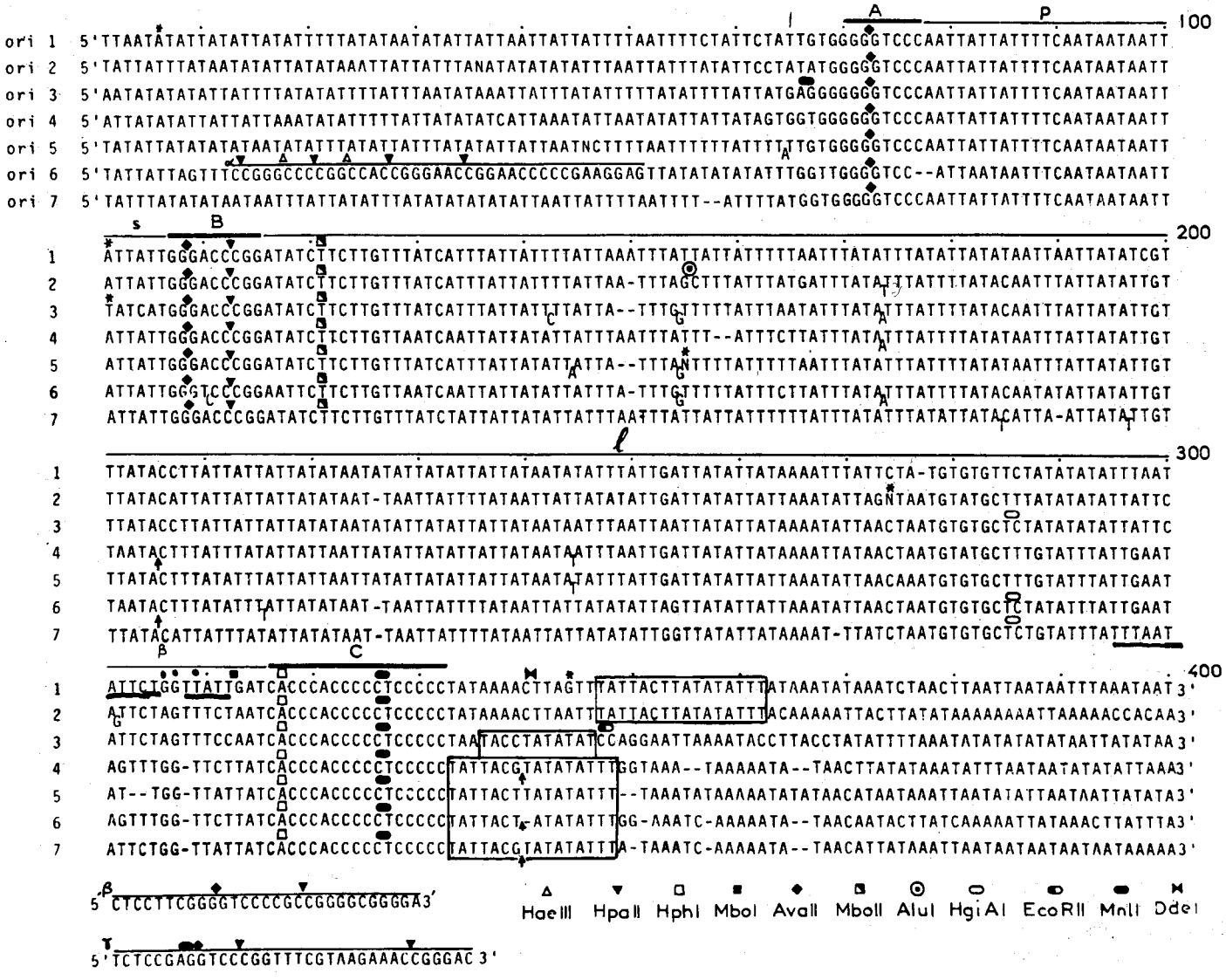

Fig. 3. Primary structure of the ori *sequences and their flanking sequences. Thick lines indicate GC clusters A, B, and C, thin lines AT regions* p, s *and* l. *The positions and the sequences of extra GC clusters β and γ are given (as well as that of GC cluster α, which is located outside the* ori *sequence. Restriction sites are indicated by the symbols shown. The sequences homologous to initiation transcription sites at the right of cluster C, are indicated by boxes.* (From Refs 16 and 24; G. Baldacci and G. Bernardi submitted, and unpublished results of M. de Zamaroczy.)

contains several origins of replication and at least one of them is present on the excised segment[12], or sequences other than the origins of replication of the wild-type genome are used as surrogate origins of replication. In fact, both situations have been found to occur, although with very different frequencies.

Considering that the first explanation was the more likely one, when we first sequenced[13] the repeat units of two petite genomes excised from the same region of the wild-type genome, we looked for an origin of replication in the segment shared by them and found a region characterized by two short GC clusters, A and B, flanking a palindromic AT sequence, *p*, and a short AT segment, *s*; and one long GC cluster, C, separated from B by a long AT segment, *l* (see Fig. 3). The potential secondary structure of the A–B region, the primary structure of cluster C and the general arrangement of the whole region are remarkably similar to those found in other mitochondrial origins of replication (Refs 14 and 15; Fig. 4).

An *ori* sequence like the one just described was found in almost all the mitochondrial genomes of spontaneous petite mutants. Restriction mapping and hybridization of petite genomes on restriction fragments of wild-type genomes[16] provided evidence for the existence of seven such *ori* sequences in the mitochondrial genome of wild-type cells. The primary structure of the *ori* sequences shows that they are extremely homologous, particularly in their GC clusters; some of them, *ori* 4, 6, and 7, contain additional GC clusters, β and γ, identical in sequence and position (Fig. 3). All these *ori* sequences have been precisely localized and oriented on the physical map of the wild-type genome (Fig. 1).

It should be noted: (a) that some *ori* sequences display one orientation on the wild-type genome, and some the opposite; (b) that *ori* 2 and 7 and *ori* 3 and 4 are close to each other and tandemly oriented; (c) that *ori* 4 is absent in a wild-type strain (G. Faugeron-Fonty, personal communication); (d) that *ori* sequences containing the γ cluster have been found only once (*ori* 4) or not at all (*ori* 6, *ori* 7) in extensive screenings of spontaneous petite genomes.

Ori° petites, lacking a canonical *ori* sequence, have also been found, although very rarely. An investigation of the mitochondrial genomes of eight such *ori*° petites[17] has revealed that their repeat units contain, instead of canonical *ori* sequences, one or more *ori*^{s} sequences. These 44-nucleotide long surrogate origins of replication are a subset of GC clusters characterized by a potential secondary fold with two sequences ATAG and GGAG inserted in AT spacers; these sequences are followed by two AT base pairs, a GC stem (broken in the middle and in most cases also near the base, by non-paired nucleotides) and a terminal loop (Fig. 5). This structure is reminiscent of that of GC clusters A and B from canonical *ori* sequences (Fig. 4). Like the latter, *ori*^{s} sequences are present in both orientations, are located in intergenic regions and can be used as excision sequences when tandemly oriented. *Ori*^{s} sequences are homologous with many other subsets of GC clusters (one of these subsets, the *ori*^{s}-like sequences, is shown in Fig. 5) some or all of which might perhaps also act as surrogate origins of replication, possibly still less efficiently than *ori*^{s} sequences.

TABLE I. Replication and transcription of petite genomes

A)	*Ori*$^+$ petites	Suppressivity[a]	Transcription	B)	*Ori*$^-$ petites	Suppressivity	Transcription
	ori 1	>95%	+		*ori* 1 A$^-$	80%	+
	ori 2	>95%	+		*ori* 1 C$^-$	n.d.	–
	ori 3	85%	±		*ori* 3 C$^-$	< 5%	–
	ori 4	[b]	–	C)	*Ori*° petites		
	ori 5	90%	+		a-15/4/1/10/3	~ 1%	–
	ori 6	[b]	n.d.		a-3/1/B4	[c]	–
	ori 7	[b]	n.d.	D)	*Ori*^{r} petite	< 5%	n.d.

[a] Values found for petite genomes having repeat units ~900 (*ori* 1, 2) or 1800 (*ori* 3, 5) base pairs long.
[b] *Ori* 4 was only found once, *ori* 6 and 7 were never found alone in the extensive screenings of spontaneous petite genomes.
[c] Diploid.

The replication of petite genomes and the phenomenon of suppressivity

A functional evidence that *ori* sequences are indeed involved in the replication of the mitochondrial genome came from crosses of spontaneous petites, characterized in their mitochondrial genome and their suppressivity, with wild-type cells[16,18,19]. Crosses of spontaneous, highly suppressive petites having mitochondrial genomes formed by very short repeat units (400–900 base pairs) with wild-type cells produced diploids which harbored only the unaltered mitochondrial genomes of the parental petite, which was called supersuppressive[18,19]. When petites with different degrees of suppressivity were used in the crosses, the genomes of diploid petite progeny had restriction maps identical to those of the parental haploid petites. A very few exceptions were found and these corresponded to new excision processes affecting one of the parental genomes.

There are two clear correlations between the *ori* sequence of the petite used in the cross and its degree of suppressivity. First,

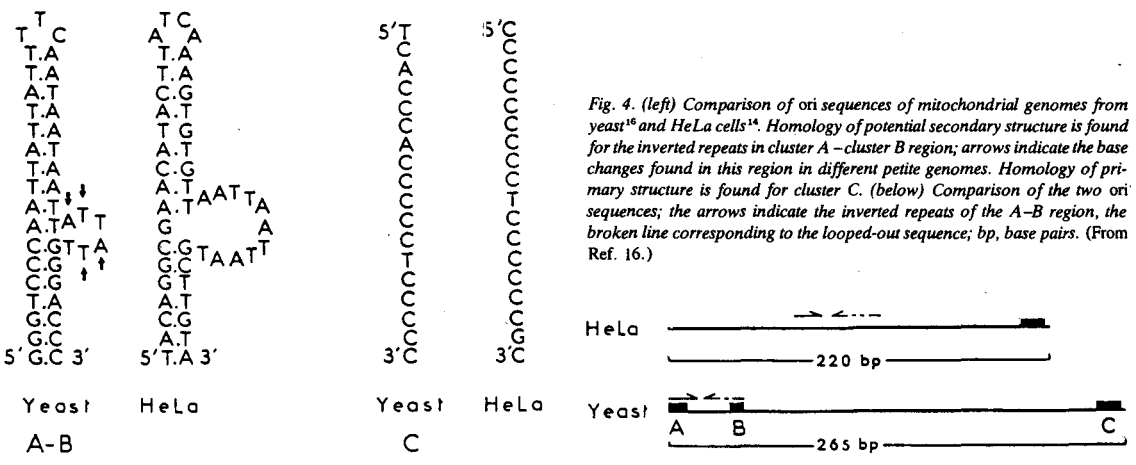

Fig. 4. (left) Comparison of ori *sequences of mitochondrial genomes from yeast*[16] *and HeLa cells*[14]. *Homology of potential secondary structure is found for the inverted repeats in cluster A – cluster B region; arrows indicate the base changes found in this region in different petite genomes. Homology of primary structure is found for cluster C. (below) Comparison of the two* ori *sequences; the arrows indicate the inverted repeats of the A–B region, the broken line corresponding to the looped-out sequence; bp, base pairs.* (From Ref. 16.)

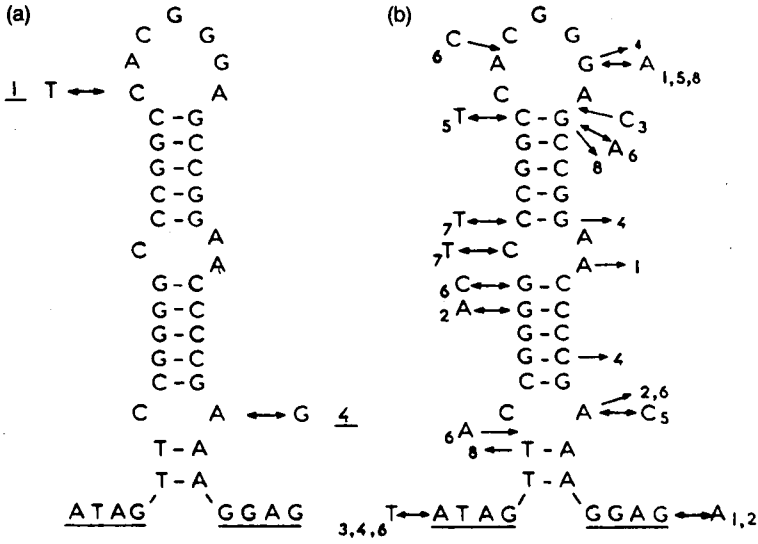

Fig. 5. Potential secondary structure of: (a) the ori^s *sequences; and (b) the* ori^s*-like sequences. All sequences are drawn in the same orientation ATAG → GGAG. Double-headed arrows indicate base exchanges, arrows pointing towards, or away from, the structure indicate insertions and deletions, respectively. Numbers indicate the* ori^s*, or* ori^s*-like, sequences presenting these changes.* (From Ref. 17.)

all other things being comparable (namely, the intact state of the *ori* sequence and the total amount of mitochondrial DNA per cell), the lower the overall density of *ori* sequences on the genome units, the lower the supressivity. Second (see Table I), partial or total deletion of the *ori* sequences, or their rearrangement, affects the suppressivity. (a) *Ori*$^-$ petites, in which the *ori* sequence is partially deleted, show a decreased suppressivity relative to *ori*$^+$ petites carrying intact *ori* sequences. The loss of cluster C with its flanking sequence has a much more dramatic effect than the loss of cluster A; (b) *ori*^{s} petites, which show an inverted orientation of two *ori* sequences within the same repeat units (the latter having, in turn, an alternate inverted and tandem orientation) have a very low degree of suppressivity. (c) ori^0 petites, which lack the *ori* sequence altogether but contain *ori*^{s} sequences instead, have low to minimal degrees of suppressivity.

These results provide a molecular basis for a replicative advantage being the explanation for suppressivity. This hypothesis can be traced back to Ephrussi *et al.*[20], but received a particular attention[21,22] after the work of Mills *et al.*[23] on the *in vitro* replication of QβRNA. It is quite possible that a replicative advantage also explains amplification of the original monomeric genomes of petites.

The *ori* sequences as transcription initiation sites

Three recent results on the transcription initiation sites in petite genomes are also relevant[24].

(a) Transcription initiation efficiency parallels replication efficiency. Petite genomes containing some canonical *ori* sequences (*ori* 1, 2, 5, and, to a lesser extent, *ori* 3) are transcribed very actively; others, containing *ori* 4, or deleted in their C clusters (but not those deleted in A clusters), or lacking canonical *ori* sequences (*ori*$_0$ petites), are not (Table I). Likewise, *ori* sequences containing a γ cluster (*ori* 4, 6 and 7; two of these are in tandem with *ori* 3 and *ori* 2, respectively) are probably not very efficient in DNA replication, as suggested by the fact that they are very rarely or never found in extensive screenings of spontaneous petites and may even be absent (*ori* 4) in some wild-type genomes.

(b) Since transcriptionally active *ori* sequences (see above) are present in both orientations on the wild-type genome, it is very likely that both strands are transcribed, although the non-sense strand appears to be transcribed less accurately or more slowly. Hybridization experiments with separated DNA strands have identified the template strand used in transcription as the strand which contains the oligonucleotide stretch of cluster C.

This conclusion supports previous independent evidence[25–27] and puts the transcription of the mitochondrial genome of yeast in line with that of the animal mitochondrial genome. Similarly, replication might proceed unidirectionally from some *ori* sequences (possibly *ori* 2 and 5 for one strand, and *ori* 1 and 3 for the other). If this is the case, the replication of the mitochondrial genome of wild-type yeast cells would be analogous with the replication of unicircular dimers of the mammalian mitochondrial genome[28].

(c) Transcription initiates next to the oligopyrimidine stretch of cluster C, at a sequence (Fig. 3) very homologous (Fig. 6) to the transcription initiation sequences of rRNA genes[29], and proceeds in from cluster C to cluster A. As already mentioned, the insertion of cluster γ in the middle of this sequence (as in *ori* 4) or the loss of cluster C and of this sequence (as in *ori* C$^-$ petite genomes) is accompanied by a loss of transcriptional activity. This suggests that *ori* 2 and *ori* 5 might be among the initiation sites used for transcribing the sense strand, *ori* 1 and *ori* 3 among those used for the transcription of the other strand. A small number of other sequences largely homologous to the transcription initiation sites of tRNA genes have also been found. These might play a role in the multipromotor transcription[30] of the mitochondrial genome of wild-type cells, postulated years ago[12].

In conclusion, the investigations outlined in this review provide answers to the questions raised many years ago by the work of Ephrussi in the petite mutation. In fact, they have done more than this, since they have opened the way to a fine analysis of replication, recombination and expression in the mitochondrial genome of yeast, and have shed some light on the general problem of genome evolution (see Ref. 31 for a recent discussion on the latter point).

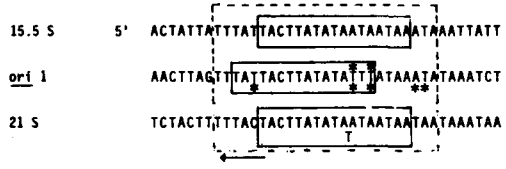

Fig. 6. Comparison of the transcription initiation sequences of ori *1 and of the 15.5s and 21s rRNA genes*[27]*. Solid-line boxes indicate the transcription initiation sequences (as read on the coding strand to ensure consistency with Fig. 3); the broken-line box indicates the region of homology among the three sequences; asterisks indicate base differences, the arrows the start of the tRNA transcripts.* (From Ref. 24.)

References

1 Bernardi, G. (1979) *Trends Biochem. Sci.* 4, 197–201
2 Marotta, R., Colin, Y., Goursot, R. and Bernardi, G. (1982) *EMBO J.* 5, 529–534

3 de Zamaroczy, M., Faugeron-Fonty, G. and Bernardi, G. (1982) *Proc. Natl Acad. Sci. U.S.A.* (submitted)
4 Fonty, G., Goursot, R., Wilkie, D. and Bernardi, G. (1978) *J. Mol. Biol.* 119, 215–235
5 Bonitz, G. S., Coruzzi, G., Thalenfeld, B. E. and Tzagoloff, A. (1980) *J. Biol. Chem.* 255, 11927–11941
6 Thalenfeld, B. E. and Tzagoloff, A. (1980) *J. Biol. Chem.* 255, 6173–6180
7 Heizmann, P., Doly, J., Hussein, Y., Nicolas, P., Nigon, V. and Bernardi, G. (1981) *Biochim. Biophys. Acta* 633, 412–415
8 Flavell, A. J. and Ish-Horowicz, D. (1981) *Nature* 292, 591–595
9 Calabretta, B., Robberson, D. L., Barrera-Saldana, H. A., Lambrou, T. P. and Saunders, F. G. (1982) *Nature* 296, 219–224
10 Kaufman, R. J., Alt, F. W., Kellems, R. F. and Schimke, R. T. (1978) *Science* 202, 1051–1055
11 Faugeron-Fonty, G., Culard, F., Baldacci, G., Goursot, R., Prunell, A. and Bernardi, G. (1979) *J. Mol. Biol.* 134, 493–537
12 Prunell, A. and Bernardi, G. (1977) *J. Mol. Biol.* 110, 53–74
13 Gaillard, C. and Bernardi, G. (1979) *Mol. Gen. Genet.* 174, 335–337
14 Crews, S., Ojala, D., Posakony, J., Nishiguchi, J. and Attardi, G. (1979) *Nature* 277, 192–198
15 Gillum A. M. and Clayton D. A. (1979) *J. Mol. Biol.* 135, 353–368
16 de Zamaroczy, M., Marotta, R., Faugeron-Fonty, G., Goursot, R., Mangin, M., Baldacci, G. and Bernardi, G. (1981) *Nature* 292, 75–78
17 Goursot, R., Mangin, M. and Bernardi, G. (1982) *EMBO J.* 6, 705–711
18 de Zamaroczy, M., Baldacci, G. and Bernardi, G. (1979) *FEBS Letters* 108, 429–432
19 Goursot, R., de Zamaroczy, M., Baldacci, G. and Bernardi, G. (1980) *Current Genet.* 1, 173–176
20 Ephrussi, B., Jakob, H. and Grandchamp, S. (1966) *Genetics* 54, 1–29
21 Rank, G. H. (1970) *Can. J. Genet. Cytol.* 12, 129–136
22 Rank, G. H. (1970) *Can. J. Genet. Cytol.* 12, 340–346
23 Mills, D. R., Peterson, R. L. and Spiegelman, S. (1967) *Proc. Natl Acad. Sci. U.S.A.* 58, 217–224
24 Baldacci, G. and Bernardi, G. *EMBO J.* (in press)
25 Li, M. and Tzagoloff, A. (1979) *Cell* 18, 47–53
26 Coruzzi, G., Bonitz, S. G., Thalenfeld, B. E. and Tzagoloff, A. (1981) *J. Biol. Chem.* 256, 12782–12787
27 Beilharz, M. W., Cobon, G. S. and Nagley, P. (1982) *Nucleic Acid Res.* 10, 1051–1070
28 Clayton, D. A. (1982) *Cell* 28, 693–705
29 Osinga, K. A. and Tabak, H. F. (1982) *Nucleic Acid Res.* 10, 3617–3626
30 Levens, D., Ticho, B., Ackerman, E. and Rabinowitz, M. (1981) *J. Biol. Chem.* 256, 5226–5232
31 Bernardi, G. (1982) in *Mitochondrial Genes* (Slonimski, P. P. *et al.*, eds), pp. 269–278, Cold Spring Harbor, New York