

Excision sequences in the mitochondrial genome of yeast

Claire Gaillard, François Strauss & Giorgio Bernardi

Laboratoire de Génétique Moléculaire,
Institut de Recherche en Biologie Moléculaire, Paris, France

It is well established that spontaneous cytoplasmic 'petite' mutants of *Saccharomyces cerevisiae* have mitochondrial genome units in which an excised segment of the parental wild-type genome has been tandemly amplified (Fig. 1), so that the excised segment becomes the repeat unit of the petite genome; the latter may in turn undergo further deletions leading to secondary petite genomes having shorter repeat units (see ref. 1 for a brief review). Recent investigations² on the mitochondrial genomes of several spontaneous petite mutants have shown that frequently the ends of the excised segment correspond to short sequences of the wild-type genome which are extremely rich in GC, the GC clusters; alternatively, they seem to be located in the long AT-rich stretches, the AT spacers, which form at least half of the genome. As sequence repetitions have been demonstrated in both GC clusters and AT spacers²⁻⁵, it is very likely that excision takes place by a mechanism involving illegitimate site-specific recombination events between homologous sequences, as previously postulated¹. We show here that the sequences involved in the excision of a particular spontaneous petite genome are direct nucleotide repeats located in the AT spacers.

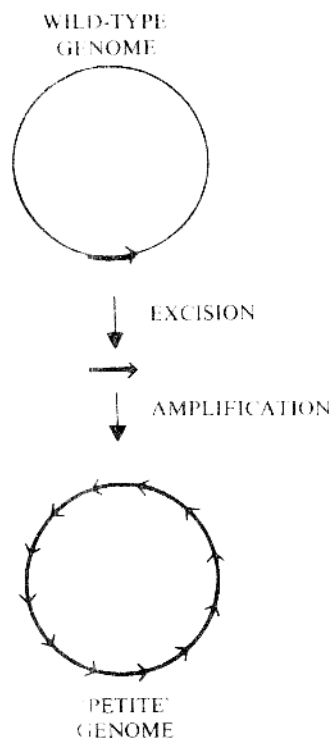


Fig. 1 Scheme of the process leading to the formation of spontaneous petite genomes. A segment of a mitochondrial genome unit from wild-type yeast cells is excised and tandemly amplified to yield the mitochondrial genome unit of a petite mutant. The excised segment from the wild-type genome becomes the repeat unit of the petite genome. This may in turn undergo further deletions leading to secondary petite genomes having simpler repeat units.

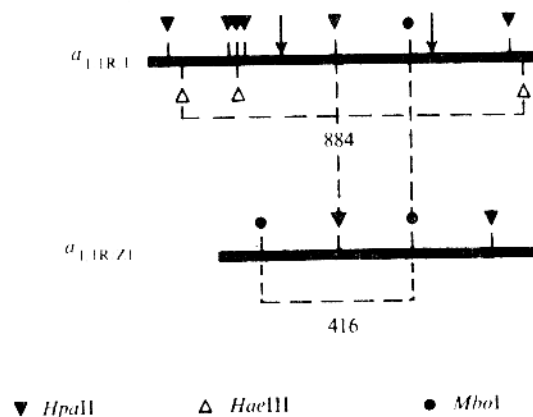


Fig. 2 Restriction enzyme maps of the repeating units of the mitochondrial genomes of two spontaneous petite mutants. The molecular weights of the repeat units are indicated (horizontal broken lines), along with the positions of *Hae*III (Δ), *Hpa*II (▼) and *Mbo*I (●) restriction sites. The vertical broken lines indicate corresponding restriction sites in the repeat units of the two genomes. Arrows indicate the approximate positions of excision sites as obtained from the data of Fig. 3.

Our approach was to sequence the repeat unit of the mitochondrial genome of a spontaneous petite mutant, $a_{1/IR/1}$, which was supposed to contain the repeat unit and the flanking sequences of another spontaneous petite mutant, $a_{1/IR/Z1}$. Both petite genomes had been previously characterized²; they originated from the same region of the parental wild-type mitochondrial genome (between map positions 27 and 46, ref. 6) and their restriction maps were known (Fig. 2); the DNA from $a_{1/IR/Z1}$ hybridised on restriction fragments from wild-type DNA corresponding to the larger *Hae*III fragments and to the two largest *Hpa*II fragments of $a_{1/IR/1}$, and shared a *Hpa*II-*Mbo*I fragment with $a_{1/IR/1}$; finally, the sequence of the repeat unit of $a_{1/IR/Z1}$ was known⁵.

The nucleotide sequence was determined, using the method of Maxam and Gilbert⁷, on both mitochondrial DNA prepared from $a_{1/IR/1}$ and its *Mbo*I repeat unit inserted in the *Bam*HI site of pBR322 plasmid and amplified in a *rec*⁻ *Escherichia coli* strain (HVC46; C600 *leu*⁻, *thi*⁻, *lac*⁻, *tonA*, *supE44*, *hsdR*, *hsdM*). The same sequence was found in both cases (Fig. 3). The repeat unit of $a_{1/IR/1}$ is 884 nucleotides long and contains, between positions 278 and 693, a 416-nucleotide segment which is identical to the sequence of the repeat unit of $a_{1/IR/Z1}$ (ref. 5), except for two base pair changes, which are indicated by an asterisk in Fig. 3. The most important feature of the sequence of $a_{1/IR/Z1}$ is that the initial nonanucleotide of the repeat unit of $a_{1/IR/Z1}$, AATAATATA, is repeated just after the end of the repeat unit; if allowance is made for one A·T to T·A change, these direct repeats are, in fact, 13 nucleotides long. This situation is reminiscent of those found on both sides of insertion sequence IS1 and of transposon Tn9 in *E. coli*, where two direct repeats of 9 base pairs were found^{8,9}. It provides evidence consistent with the excision model proposed previously, and raises the interesting possibility that segments like $a_{1/IR/Z1}$ can be transposed onto other mitochondrial genome units and contribute to the changes in length observed in genome units belonging to different strains¹⁰. Incidentally, the data do not allow one to decide whether the repeat unit of $a_{1/IR/Z1}$ was excised from the parental wild-type genome or from $a_{1/IR/1}$, although the latter explanation seems the more likely. This is not an essential point, however, as it is known that the same excision process is operative in wild-type and petite genomes².

Several other features of the sequence of $a_{1/IR/1}$ are worth mentioning: (1) two direct AT repeats, 16 nucleotides long, containing the nonanucleotides flanking the $a_{1/IR/Z1}$ sequence,

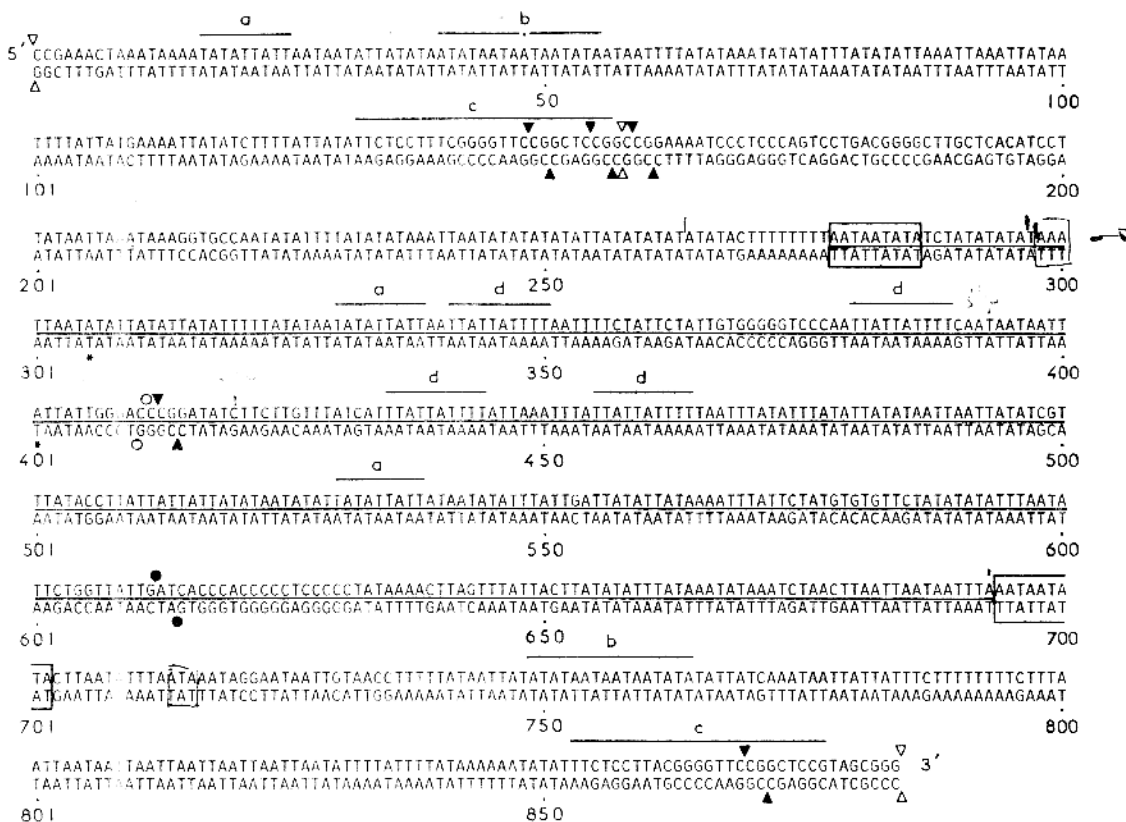


Fig. 3 Nucleotide sequence of the repeat unit of the mitochondrial genome of spontaneous petite mutant $a_{1/IR/1}$. Cutting sites of *Hae*III (Δ), *Hpa*II (\blacktriangle), *Mbo*I (\bullet) and *Mbo*II (\circ) are indicated. The sequence between positions 278 and 693 is identical to that of $a_{1/IR/Z1}$ (ref. 15) except for asterisked base pairs 306 and 401, which are replaced by G-C and C-G, respectively, in $a_{1/IR/Z1}$. Boxed sequences correspond to the nonanucleotides flanking the sequence of $a_{1/IR/1}$. Lines indicate repeated sequences (see text).

are found at positions 40-55 and 749-764 (these are labelled b in Fig. 3); (2) two GC-rich repeats, 25 nucleotides long with only one A-T to T-A change, are found at positions 132-156 and 853-877 (these are labelled c in Fig. 3); the first one is part of a 65-nucleotide stretch very rich in GC (GC=65%), which accounts for the higher GC level of $a_{1/IR/1}$ (GC=16.2%) compared with $a_{1/IR/Z1}$ (GC=14.2%); (3) a 20-nucleotide sequence formed by five repeats of the sequence TTAA is found at positions 809-828; (4) symmetrical sequences are found in the GC cluster at position 150, in agreement with previous predictions¹⁰ and findings^{3,4}; (5) other sequence features of $a_{1/IR/1}$ are common with $a_{1/IR/Z1}$ and have already been mentioned⁵; the most remarkable one is a 47-nucleotide stretch (369-415) formed by a palindrome 23 nucleotides long (378-400) and a small symmetrical sequence TTATT (402-406), which are flanked by two symmetrical GC clusters (369-377 and 407-415).

The main conclusion of the present work is that crossing-over involving short homologous sequences is indeed the most likely mechanism underlying the excision of spontaneous petite genomes. Sequences used in the excision of other spontaneous petite genomes are being investigated in our laboratory. Some of the genomes under consideration are also encompassed by $a_{1/IR/1}$, and it will be interesting to see which other sequences of $a_{1/IR/1}$ are involved in excision. We expect to find among them some or the other repeats of $a_{1/IR/1}$.

Another point of interest is that the presence of direct repeats in its sequence has not caused any instability of the $a_{1/IR/1}$ repeat unit as propagated in a *rec⁻* *E. coli*; it should be stressed, however, that monomers of the repeat unit were cloned and that oligomers were not tested.

Finally, an inspection of the sequence of $a_{1/IR/1}$ fails to reveal the presence of any gene, a conclusion confirmed by the present knowledge of codons used for the synthesis of mitochondrially encoded proteins^{3,11} and of tRNA gene sequences¹². The only

function left in this petite genome (as well as in $a_{1/IR/Z1}$) is replication. As the excision process leading to the formation of spontaneous petite genomes is extremely conservative as far as the excised sequence is concerned² (in contrast to that induced by ethidium, where sequence rearrangements are frequent¹³), each repeat unit of $a_{1/IR/1}$ should contain a site used for the initiation of DNA replication. A good candidate is the 80-nucleotide sequence centred around position 412. This sequence, (which contains the remarkable stretch, 369-415, mentioned above), is present not only in the two overlapping genomes $a_{1/IR/1}$ and $a_{1/IR/Z1}$, but also in another one¹⁴, b, whose 875-nucleotide repeat unit is derived² from an entirely different region of the parental wild-type genome (map positions 64-76, ref. 6). If a site for the initiation of DNA replication is present in each repeat unit of these petite genomes, a high replication rate may be expected. This indeed seems to be the case, as the three petites under consideration are 'supersuppressive', namely petites whose mitochondrial genomes compete out those of wild-type cells in crosses and become the only ones transmitted to the progeny¹⁴.

Received 10 September; accepted 31 October 1979.

- Bernardi, G. *Trends biochem. Sci.* **4**, 197-201 (1979).
- Faugeron-Ponty, G. *et al.* *J. molec. Biol.* (in the press).
- Cosson, J. & Tzagoloff, A. *J. Biol. Chem.* **254**, 42-43 (1979).
- Macino, G. & Tzagoloff, A. *Proc. natn. Acad. Sci. U.S.A.* **76**, 131-135 (1979).
- Gaillard, C. & Bernardi, G. *Molec. gen. Genet.* **174**, 335-337 (1979).
- Sanders, J. P. M., Helying, C., Verbeet, M. P., Meijink F. C. P. W. & Borst, P. *Molec. gen. Genet.* **157**, 239-261 (1977).
- Maxam, A. & Gilbert W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 560-564 (1977).
- Calos, M. P., Johnson, L. & Miller, J. H. *Cell* **13**, 411-418 (1978).
- Grindley, N. D. F. *Cell* **13**, 419-426 (1978).
- Prunell, A. & Bernardi, G. *J. molec. Biol.* **110**, 53-74 (1977).
- Hengens, J., A. M., Grivell, L. A., Borst, P. & Bus, J. L. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1663-1667 (1979).
- Martin, N. C., Miller, D. & Donelson, J. E. *J. supramolec. Struct. Suppl.* **3**, 132 (1979).
- Lewin, A., Morimoto, R., Rabinowitz, M. & Fukuhara, H. *Molec. gen. Genet.* **163**, 257-275 (1978).
- Goursot, R., de Zamato, M., Baldacci, G. & Bernardi, G. *Molec. gen. Genet.* (in the press).