EXCISION SITES IN THE GC CLUSTERS OF THE MITOCHONDRIAL GENOME OF YEAST

Giuseppe BALDACCI, Miklos de ZAMAROCZY and Giorgio BERNARDI

Laboratoire de Génétique Moléculaire, Institut de Recherche en Biologie Moléculaire, 2 Place Jussieu, 75005 Paris, France

Received 28 March 1980

1. Introduction

The mitochondrial genome units of wild-type *Saccharomyces cerevisiae* cells are characterized by the presence of: (a) long AT-spacers, which form ⩾50% of each 50×10^6 (76 000 base pairs) genome unit and consist of short alternating and non-alternating A:T sequences; (b) 100 or so short GC clusters which are embedded in the AT-spacers and are formed, in most cases, of CCGG sequences clustered with GGCC or other GC-rich sequences (reviewed [1]). The existence of a very large number of homologous or quasi-homologous sequences accounts for the enormous instability of the mitochondrial genome of yeast, which is known to generate spontaneously, at 1–5%/generation, 'petite' mutants with defective mitochondrial genomes. These genomes arise by a mechanism involving the excision of a segment from one of the 50–100 mitochondrial genome units of the parental wild-type cells, and its amplification into tandem repeat units [1]. In general, the 'petite' genome unit so formed will segregate into one of the buds of the parental wild-type cell and give rise, in a few generations time, to the mitochondrial genome of 'petite' mutant cells.

Restriction mapping of the repeat units of the mitochondrial genomes from several spontaneous 'petites' and hybridization of these genomes on restriction fragments from the parental wild-type genomes have shown that these 'petite' genomes fall into two classes [2]: (i) repeat units excised at GC clusters which appear to be more frequent; (ii) repeat units excised elsewhere, most probably in the AT-spacers. Sequence work on the repeat units of two 'petites' of class (ii) has shown that they had been excised at two pairs of direct repeats located in the AT-spacers. In both cases [3–5], the direct repeats were 13-nucleotides long, with one mismatch in one case.

Here we have studied the excision sequence of the repeat unit of a 'petite' of class (i), with the double aim of defining precisely the excision sequences used and of understanding the reasons of the preferential excision at GC clusters relative to sequences in the AT-spacers.

2. Materials and methods

The two 'petite' genomes studied here are $a_{3/1}$ and $a^*_{1/7/8}$. These genomes had been investigated in detail [2]. The methods used are in [2,3].

3. Results and discussion

The restriction maps of 'petites' $a_{3/1}$ and $a^*_{1/7/8}$ are given in fig.1. In spite of their completely independent origin from a common parental wild-type strain (A) the repeat unit of $a^*_{1/7/8}$ corresponds to the central part of the repeat unit of $a_{3/1}$; all mapped restriction sites of $a^*_{1/7/8}$ are also found on $a_{3/1}$, and exci-

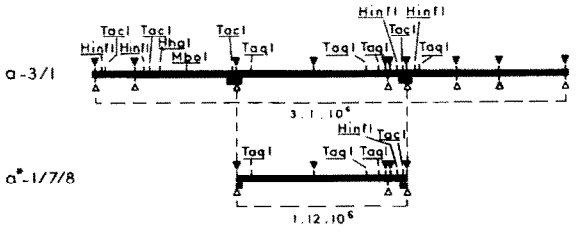

Fig.1. Restriction maps of the repeat units of the mitochondrial genomes of 'petite' mutants $a_{3/1}$ and $a^*_{1/7/8}$. Filled-in and open triangles indicate *Hpa* II and *Hae* III sites, respectively. Short horizontal bars indicate the regions which have been sequenced. Original code number subscripts follow the hyphen in fig.1 and 2.

sion of $a^*_{1/7/8}$ appears to have taken place at, or very near, two GC-clusters containing both *Hpa* II and *Hae* III sites (CCGG and GGCC) of the wild-type parental genome.

Since there is ample evidence for sequence conservation in the repeat units of spontaneous 'petites' relative to the segment of the wild-type parental genome from which they arose [2], our approach was to sequence around the putative excision sites on $a_{3/1}$ (instead of A) and inside the outermost *Hpa* II sites of $a^*_{1/7/8}$, as indicated in fig.2. The left side sequence (127 nucleotides long) of $a_{3/1}$ contains a GC-clusters 46 nucleotides long; the right-hand side sequence (101 nucleotides long) of $a_{3/1}$ contains a GC cluster 56 nucleotides long. The two clusters have a common stretch of 23 nucleotides, containing a *Tac* I (*Tha* I), a *Hpa* II and a *Hae* III site. This sequence is also found at one end of the repeat unit of $a^*_{1/7/8}$.

The boxed sequence of fig.2 defines the excision sequence of $a^*_{1/7/8}$. As in the two previous cases studied, this sequence can be visualized as starting the repeat unit of $a^*_{1/7/8}$ and being repeated just after its end. It should be pointed out, however, that the beginning of the repeat unit of $a^*_{1/7/8}$ can be put at any position within the 23 base pair repeat. The two main differences with the two previous results [3–5] are its length, 23 instead of 13 nucleotides, and that it is 80% in GC instead of 100% in AT. We suggest that these are the features which favor illegitimate site-specific recombination events at GC clusters rather than at sequences in the AT-spacers. That is, the higher stability of the heteroduplex formed at recombination seems to us to be the main reason for this preference.

Another feature of interest in the sequences investigated is that $a^*_{1/7/8}$ differs from $a_{3/1}$ in two basepair changes and in the deletion of one base pair. These changes have occurred in the common parental genome over the time (several years) separating the excision of the two 'petite' genomes because they exist on all repeat units, indicating that the changes preceded the amplification process. Other changes may have occurred in the two 'petite' genomes after their formation, but they are not detectable because they are not present in all repeat units.

In conclusion, the present investigations:

(1) Provide sequence information on two GC clusters, fitting with our predictions of sequence homology in these sequences and are similar to those for three other GC clusters [6];
(2) Confirm our prediction on the use of such clusters in the excision of 'petite' genomes [2,7];
(3) Appear to account for the higher frequency of excision at GC clusters *versus* AT-spacers.

Finally, it should be mentioned that the 'physiological' role of GC cluster, and excision sequences in general, has been discussed in [5], and that the sequences of $a^*_{1/7/8}$ and $a_{3/1}$ contain an origin of replication [8].

Acknowledgement

The authors wish to thank Mr Alain Huyard for his help in some experiments.

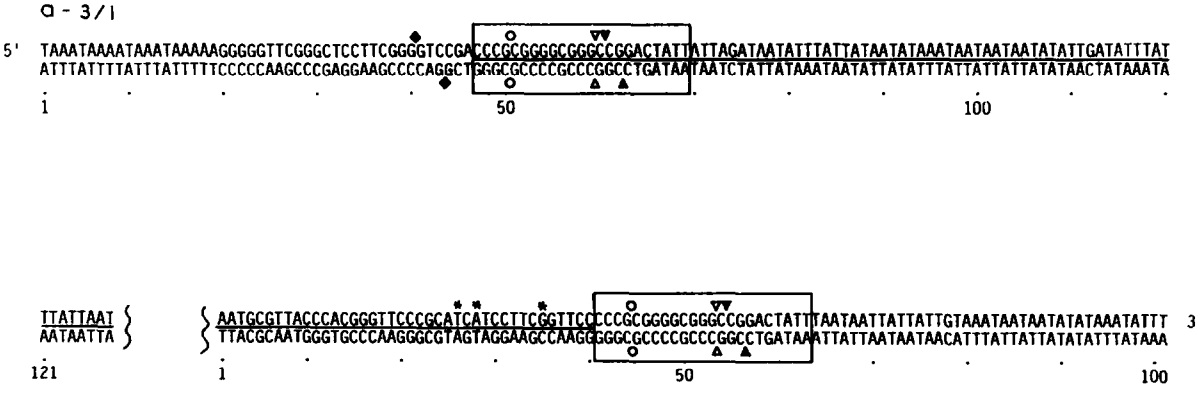

Fig.2. Primary structure around the putative excision sequences used in the formation of the repeat unit of $a^*_{1/7/8}$. Lozenges indicate an *Ava* II site, the open circle a *Tac* I site, filled-in and open triangles *Hpa* II and *Hae* III sites, respectively. The $a^*_{1/7/8}$ sequence is identical to that of $a_{3/1}$ in the region indicated by the continuous line, except for two base pair changes (A:T and G:C replace T:A and A:T) and one deletion (G:C) at the positions indicated by asterisks. (See Text.)

References

[1] Bernardi, G. (1979) Trends Biochem. Sci. 4, 197–201.
[2] Faugeron-Fonty, G., Culard, F., Baldacci, G., Goursot, R., Prunell, A. and Bernardi, G. (1979) J. Mol. Biol. 134, 493–537.
[3] Gaillard, C., Strauss, F. and Bernardi, G. (1980) Nature 283, 218–220.
[4] Bernardi, G., Baldacci, G., Culard, F., Faugeron-Fonty, G., Gaillard, C., Goursot, R., Strauss, F. and De Zamaroczy, M. (1980) FEBS Symp. DNA (Sponar, J. and Zadrazil, S. eds) Pergamon, New York, in press.
[5] Bernardi, G., Baldacci, G., Culard, F., Faugeron-Fonty, G., Gaillard, C., Goursot, R., Strauss, F. and De Zamaroczy, M. (1980) in: Mobilization and Reassembly of Genetic Information (Scott, W. A. et al. eds) Academic Press, New York, in press.
[6] Tzagoloff, A., Macino, G., Nobrega, M. and Li, M. (1979) in: Extrachromosomal DNA (Cummings, D. J. et al. eds) pp. 339–355, Academic Press, New York.
[7] Bernardi, G., Prunell, A., Fonty, G., Kopecka, H. and Strauss, F. (1976) in: The Genetics Function of Mitochondrial DNA (Saccone, C. and Kroon, A. M. eds) pp. 185–198, Elsevier/North-Holland, Amsterdam, New York.
[8] De Zamaroczy, M., Baldacci, G. and Bernardi, G. (1979) FEBS Lett. 108, 429–432.