J. Mol. Biol. (1977) **110**, 17–52

The Mitochondrial Genome of Wild-type Yeast Cells

V†. Genome Evolution

Ariel Prunell‡, Helena Kopecka, François Strauss and Giorgio Bernardi§

Laboratoire de Génétique Moléculaire
Institut de Biologie Moléculaire, Paris 5°, France

(*Received 6 April 1976*)

When degraded with the restriction enzymes *Hae*III or *Hpa*II, the mitochondrial DNAs from one *Saccharomyces carlsbergensis* and three genetically unrelated *S. cerevisiae* wild-type strains yielded 71 to 113 fragments ranging in molecular weight from 10^4 to 4×10^6.

Genome unit sizes, calculated by adding up the molecular weights of all fragments produced by *Hae*III, *Hpa*II and, in some cases, *Hin*dII + III and *Eco*RI, were in the 52 to 54×10^6 range for the three *S. cerevisiae* strains, whereas a value lower by about 10% was found for the *S. carlsbergensis* strain. These values are in agreement with the physical size of circular twisted yeast mitochondrial DNA, as determined by electron microscopy (Hollenberg *et al.*, 1970).

Large differences in the electrophoretic patterns of *Hae*III and *Hpa*II fragments were found among the DNAs from different *S. cerevisiae* strains; *S. cerevisiae* and *S. carlsbergensis* DNAs showed only very few bands having the same mobility. Such differences appear to originate essentially from additions and deletions in the A+T-rich spacers and to be accompanied by a large preservation of gene order. Unequal crossing-over events at the spacers seem to be the source of additions and deletions and to underlie the evolution of the mitochondrial genome of yeast.

1. Introduction

The mitochondrial genome of wild-type yeast cells contains "spacers"‖ which (1) are interspersed with the "genes"; (2) have the same average size as the "genes"; (3) form about half of the genome; (4) have a G+C level lower than 5%, whereas the "genes" have an average G+C level of 32%; (5) contain short repetitive sequences essentially consisting of short alternating (dA·T-dA·T) and non-alternating (dA,dT) sequences (Bernardi *et al.*, 1970,1972; Bernardi & Timasheff, 1970; Piperno *et al.*, 1972; Ehrlich *et al.*, 1972; Prunell & Bernardi, 1974). In such a genome organization, internal crossing-over events at homologous nucleotide sequences of "spacers" might lead to deletions affecting genetic functions essential for cell respiration and originate the

† Paper IV in this series is Prunell & Bernardi (1974).

‡ Present address: Department of Biological Chemistry, Harvard Medical School, Boston, Mass. 02115, U.S.A.

§ To whom all correspondence should be addressed.

‖ We will indicate here as "spacers" the A + T-rich stretches of mitochondrial DNA, and as "genes" the G + C-rich stretches, corresponding to mitochondrial genes with their regulatory elements (Prunell & Bernardi, 1974).

respiratory deficiency which characterizes the extremely frequent, spontaneous, cytoplasmic "petite" mutation†. An implication of this model was that crossing-over events could lead to a divergence of the mitochondrial genome with preservation of the respiratory-competent phenotype, if only nucleotide sequences non-essential for the respiration were affected. We looked therefore for differences in the nucleotide sequences in the mitochondrial DNAs of one *Saccharomyces carlsbergensis* and three genetically unrelated *S. cerevisiae* wild-type strains. Such a search was encouraged by our earlier observations of slight differences in all the chemical and physical properties of the mitochondrial DNAs from two of the *S. cerevisiae* strains investigated here (Bernardi *et al.*, 1970; Bernardi & Timasheff, 1970).

The main results of this work were reported in preliminary form (Bernardi *et al.*, 1975) at the workshop on Molecular Biology of Nucleocytoplasmic Relationships (Paris, July, 1974), at the VIIth International Conference of Yeast Genetics (University of Sussex, September, 1974), and (Bernardi, 1975) at the Xth Meeting of the Federation of European Biochemical Societies (Paris, July, 1975).

2. Materials and Methods

(a) *Yeast strains*

The wild-type *S. cerevisiae* strains A and B and their culture media have been already described by Bernardi *et al.* (1970). The *S. cerevisiae* strain DM, described elsewhere with its culture medium (Bernardi *et al.*, 1968), will be indicated here as D. The *S. carlsbergensis* NCYC 74S strain was kindly supplied by Dr C. Christiansen; culture conditions were those described by Christiansen *et al.* (1974); this strain will be indicated here as C. All cells were harvested in late exponential growth phase.

(b) *Protoplast formation*

The conversion of yeast cells to protoplasts was done according to Grivell *et al.* (1971) for *S. cerevisiae* and according to Christiansen *et al.* (1974) for *S. carlsbergensis*.

(c) *Mitochondrial DNA preparation*

Protoplasts were lysed by resuspension (at 0·5 g of starting cell wet wt/ml) in 0·15 M-NaCl–2·5% sodium dodecyl sulfate (pH 7·0), and incubation at 60°C under stirring for 1 h. The solution was then adjusted to 1 M-NaCl, kept overnight in an ice bath and centrifuged; the supernatant was precipitated with 1 vol. 95% ethanol. The subsequent procedure, i.e. dissolution of the precipitate, re-extraction with 1·5% sodium dodecyl sulfate, chloroform treatment, and purification of mitochondrial DNA by batchwise adsorption and chromatography on hydroxyapatite were performed as previously described (Bernardi *et al.*, 1972); DNA was eluted in 2 almost completely separated peaks, corresponding to nuclear and mitochondrial DNA, respectively. The early fractions of the second peak, representing about 5% of mitochondrial DNA, were discarded since they were shown to be slightly contaminated with the nuclear DNA of the first peak, by analytical density gradient centrifugation. The pure mitochondrial DNA fractions forming the bulk of the second peak were pooled, concentrated to 300 to 450 μg/ml, and dialyzed against suitable buffers for restriction enzyme digestions. The molecular weights of the mitochondrial DNA preparations used in most of the present work were in the

† This model was first presented in a lecture at the Karolinska Institute, Stockholm (October, 1969). It was subsequently presented at a number of meetings and mentioned by Piperno *et al.* (1972) and by Prunell & Bernardi (1974). A very similar model has been developed by Clark-Walker & Miklos (1974).

3·5 to $5{\cdot}2 \times 10^6$ range (Table A1 of the Appendix). Two preparations of higher molecular weight (9×10^6) obtained from strains A and B by a method to be published elsewhere (Fonty, G., Kopecka, H. & Bernardi, G., unpublished data) were also used.

(d) *Other DNA preparations used*

Simian virus 40 (SV40) DNA was prepared by selective extraction of viral DNA (Hirt, 1967) from a subcloned CV_1 cell line infected with large-plaque strain SV40. DNA from wild-type phage λ was obtained by extraction of phage particles with phenol.

(e) *Enzymes and DNA degradations*

Restriction endonucleases *Hin*dII + III, *Hpa*II and *Hae*III† were prepared from *Haemophilus influenzae*, *H. parainfluenzae* and *H. aegyptius*, respectively, according to a method developed in our laboratory (Kopecka, 1975). *Eco*RI restriction endonuclease samples from *Escherichia coli* were donated by A. Bernardi and P. Yot of this Institute.

Digestions by restriction enzymes were done in TMSH medium (10 mM-NaCl, and 6·6 mM each of Tris (pH 7·4), $MgCl_2$, 2-mercaptoethanol) for 16 h in the case of the *Haemophilus* enzymes and for 2 to 4 h in the case of *Eco*RI. All digestions were performed at 37°C and stopped with EDTA, final concn 10 mM. Completeness of the degradation was carefully checked by using different enzyme levels and different digestion times. The molecular weights of *Hpa*II digests, as obtained by sedimentation, are given in Table A1 of the Appendix.

(f) *Gel electrophoresis*

*Hpa*II and *Hae*III digests obtained from 5 μg to 20 μg of DNA in 50 to 100 μl were concentrated under reduced pressure to 30 μl. Samples were then made 1% in sodium dodecyl sulfate, incubated 30 min at 37°C, and 0·2 vol. of 60% sucrose containing 0·02% bromophenol blue was added. Electrophoresis in vertical slabs (16 cm × 40 cm × 0·4 cm) was carried out at 4°C, at a constant current of 50 mA, for 20 to 30 h until bromophenol blue reached the bottom of the gel. Alternatively 16 cm × 40 cm × 0·3 cm slabs were used at a current of 40 mA. Composite gels contained 0·5% agarose (Sigma, St. Louis, Mo.) and 2, 2·5, 3, 4 and 6% acrylamide (Eastman-Kodak, Rochester, N.Y.) associated with 5% (relative to acrylamide) *N*, *N'*-methylene bisacrylamide (Fluka AG, Buchs SG, Switzerland); 6% and 8% acrylamide gels without agarose were also used. For polymerization, *N*, *N*, *N'*, *N'*-tetramethylethylenediamine (Fluka AG) and ammonium persulfate were added to final concentrations of 0·13% and 0·066%, respectively. Electrophoresis buffer was 20 mM-sodium acetate, 2 mM-EDTA, 40 mM-Tris-acetate (pH 7·8).

Electrophoresis of *Hin*dII + III and *Eco*RI digests (obtained from 2 to 3 μg of DNA in 30 μl) was done at room temperature in vertical slabs (16 cm × 20 cm × 0·3 cm) of 0·8% agarose at a constant current of 30 mA; the resolution of low molecular weight fragments was obtained on 3% polyacrylamide/0·5% agarose gels.

Gels were stained for 2 h in electrophoresis buffer containing 2 μg of ethidium bromide/ml (Sigma) and photographed under illumination by short-wavelength ultraviolet lamps (UV Products, San Gabriel, Calif.) using Polaroid type 55 P/N or 105 films and a Polaroid MP3 camera with a Kodak no. 23A red filter. Alternatively, the ultraviolet lamps were used without short-wavelength filters; in this case the camera filter was a Kodak no. 24 filter.

The Polaroid negatives were scanned with a Joyce-Loebl (Gateshead-on-Tyne, England) microdensitometer; the final magnification from the gel to the tracing was about 6×. In the case of *Hae*III and *Hpa*II digests, the tracings for a single DNA sample on the 6 different gels used had an overall length of about 12 m.

This procedure allowed an accurate measurement of the migration distance of the fragments, which could be estimated within ±0·1 mm (on the gels). The precision on the molecular weight of fragments was equal to ±0·5% in the 0·5 to 5×10^6 range (manuscript in preparation). A precise assessment of the amount of DNA per band was obtained using a new photographic method (see the Appendix).

† Restriction endonucleases are indicated here according to Smith & Nathans (1973); *Hpa*II and *Hae*III will be indicated as *Hpa* and *Hae* in the Results and Discussion sections.

(g) *Electrophoretic mobility and molecular weight of restriction fragments*

The molecular weights of the *Hae*III and *Hpa*II fragments of mitochondrial DNAs were determined using as secondary standards the *Hpa*II fragments of DNA from strain B; in turn, these were calibrated against primary standards formed by restriction enzyme fragments from SV40 and phage λ DNA, as follows. In the high molecular weight range (15 to 1×10^6) *Eco*RI fragments of λ DNA and *Eco*RI or *Eco*RI + *Hpa*II fragments of SV40 DNA were used; for a better precision, *Eco*RI fragments of λ DNA were mixed with mitochondrial DNA B fragments and run together. Figure 1 shows these results in the form of plots of log M (molecular weight) *versus* R_F (relative mobility). In the intermediate molecular weight range (1 to $0{\cdot}1 \times 10^6$), *Hin*dII + III fragments of SV40 DNA were used as molecular weight standards (Danna & Nathans, 1971; Danna *et al.*, 1973).

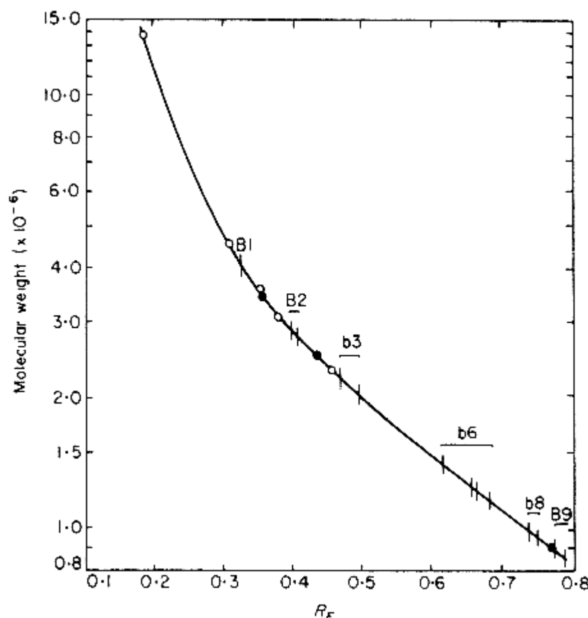

FIG. 1. Plot of molecular weights (as given by Allet *et al.*, 1973*a*,*b*; Sharp *et al.*, 1973, respectively) against relative mobilities, R_F, of fragments obtained from phage λ DNA with *Eco*RI (○), and SV40 DNA with *Eco*RI and *Eco*RI + *Hpa* (●). A molecular weight of $3{\cdot}4 \times 10^6$ for SV40 (Sharp *et al.*, 1973) ensured the best continuity between λ and SV40 fragments in the plot. Vertical bars refer to the R_F values of the bands of the *Hpa* digest of DNA from strain B; this was used as a secondary standard (see Materials and Methods). See Fig. 8 and its legend for the numbering of bands. The reference band for mobility calculation was band of $R_F = 1$ in Fig. 8.

For molecular weights lower than that of *Hin*dII + III fragment K of SV40 DNA (M = 136,000), a linear extrapolation of the straight line of the log M *versus* R_F relationship of SV40 fragments was used. Such a procedure may lead to an overestimate of the lowest molecular weights. M_w and M_n values of *Hae*III and *Hpa*II digests, as calculated from the molecular weight of the restriction fragments, are given in Table A1 of the Appendix†.

When a given DNA digest was run on 2 different gels having the same concentration, or on 2 different lanes of a gel where migration was not the same (the migration being usually faster at the center of the gel) R_F values were in general slightly different. A precise comparison of R_F values of the same or of different digests could be done, however, because a linear relationship of the type $R_{F1} = aR_{F2} + b$ (where 1 and 2 indicate the

† Tables with R_F and M values of the restriction enzyme bands are available upon request.

different experimental conditions) was always found between the 2 sets of R_F values. Since the parameters a and b in the above equation can be calculated from the R_F values of at least 2 fragments having the same mobility in the test and the reference digest (this situation always occurred for mitochondrial DNA digests), the R_F of all the fragments of the DNA under study could be normalized.

*Eco*RI fragments of λ DNA and *Hpa*II fragments of mitochondrial DNA from strain A (calibrated as described above) were used as molecular weight markers in the case of *Eco*RI and *Hin*dII + III digests.

(h) *Other methods*

Sedimentation velocity experiments were performed as described by Prunell & Bernardi (1973). Analytical CsCl density gradient experiments were done according to Bernardi *et al.* (1972).

3. Results

The results obtained with *Hae* and *Hpa* digests will be presented in sections (a) to (e); those concerning the *Eco*RI and *Hin*dII + III digests in section (f).

(a) *Band patterns*

The fragments produced by *Hae* or *Hpa* from the four mitochondrial DNAs investigated here are very numerous (71 to 113) and cover a wide range of molecular weights (10^4 to 4×10^6). Therefore, they could be satisfactorily separated only on a series of long slabs of different gel concentration. Figures 2 to 6 show the separations of *Hae* fragments from the four mitochondrial DNAs on 2 to 6% polyacrylamide gels containing 0·5% agarose. The Figures also show the separations of *Hpa* fragments from the DNA of strain B; these were used as secondary molecular weight standards (see Materials and Methods, section (g)). Figure 7 shows the separation of the *Hpa* fragments from the four mitochondrial DNAs on the 2% gel. The Figures are shown to give a precise idea of the resolution obtained under our experimental conditions over a wide range of fragment size, as well as of the intensity and appearance (spreading, sharpness) of the bands, and of the fragment homology (see section (e)).

Schemes of the band patterns obtained with the two enzymes on the five different gels are presented in Figures 8 to 12; 8% polyacrylamide gels were also used: these did not reveal any additional bands in the case of *Hae* digests, whereas they showed 5, 7, 8 and 10 additional bands for the *Hpa* digests from strains A, B, D and C, respectively (not shown). A large number of very small, single-stranded fragments not detectable by gel electrophoresis could be shown to be present in the digests by chromatography on hydroxyapatite (Prunell & Bernardi, 1977).

Since six gels of different concentration were used to resolve all the fragments, corresponding bands, as observed on different gels, had to be identified in order to establish the overlaps necessary for defining the fragment pattern derived from each mitochondrial DNA. This could be easily done on the basis of the molecular weights of the fragments and with the help of characteristic features in the band patterns, such as fragment clusters and multiple bands.

(b) *Multiple bands and faint bands*

(i) An inspection of Figures 2 to 7 reveals a number of bands showing higher intensities than the neighboring ones. Such multiple bands raised two problems concerning, respectively, the number and the identity in sequence of the fragments

FIGS 2 to 7. Electrophoretic patterns obtained with restriction enzyme digests of mitochondrial DNAs. The enzyme used, the DNA source and the gel composition are indicated.

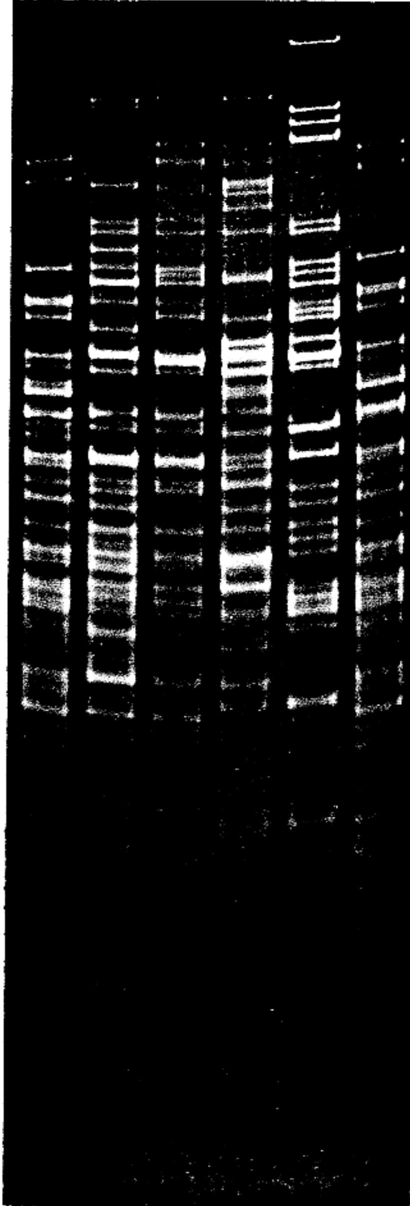

FIG. 2.

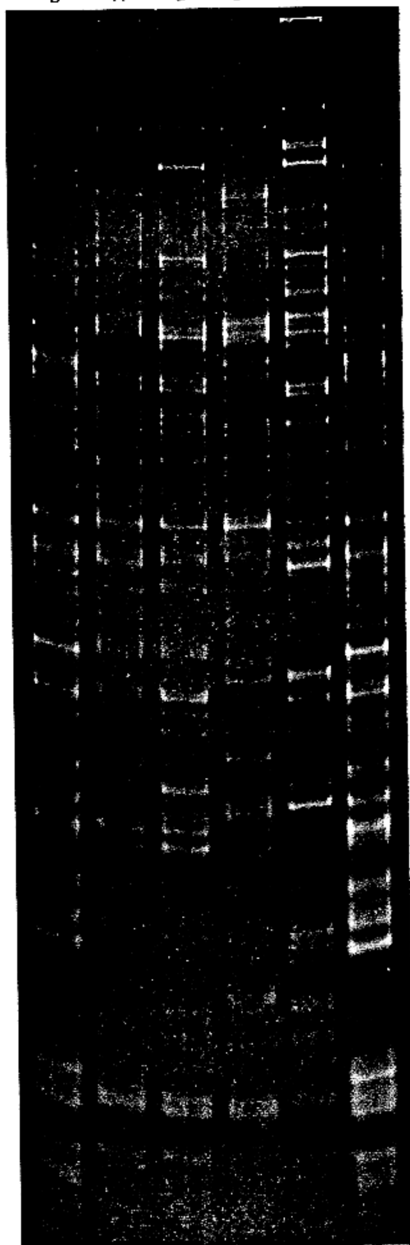

FIG. 3.

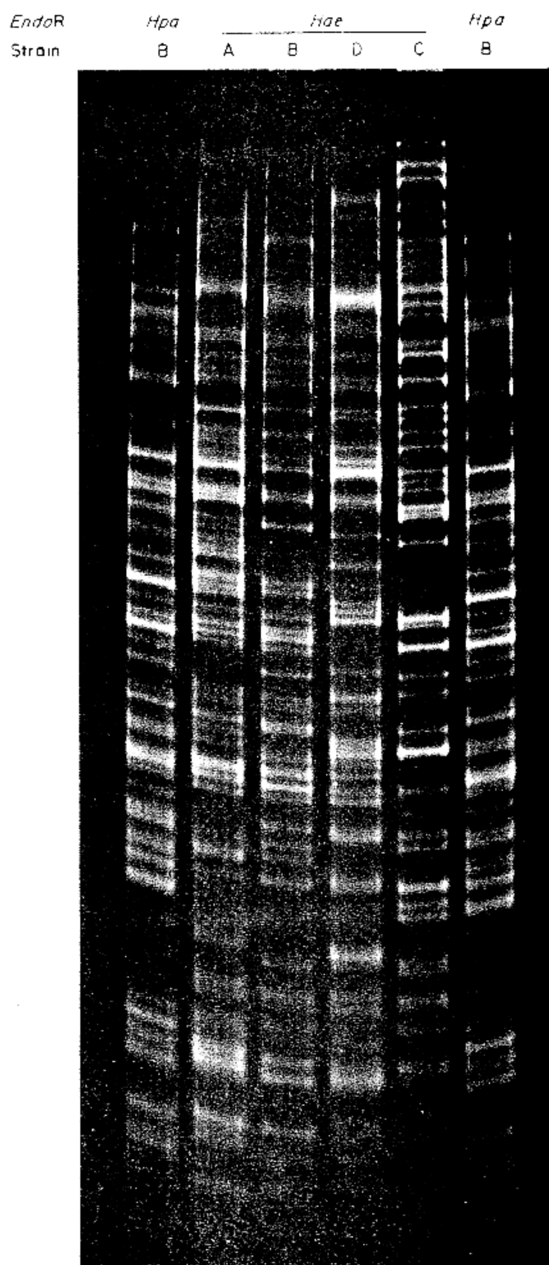

FIG. 4

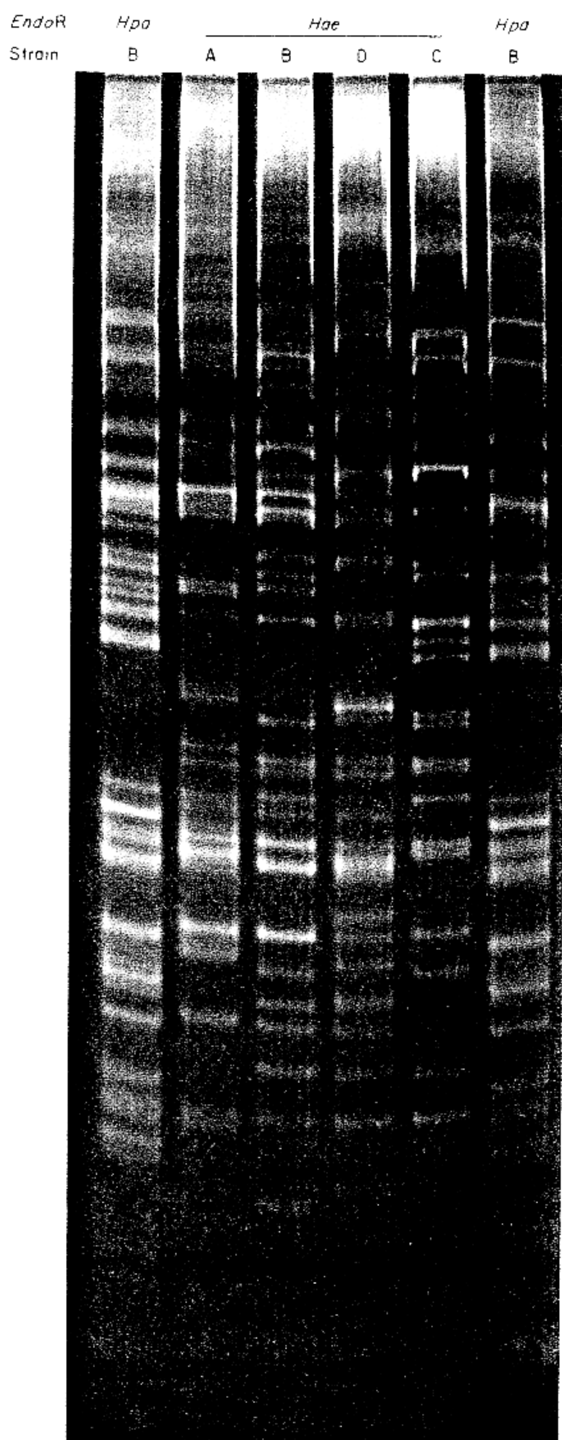

FIG. 5.

FIG. 6.

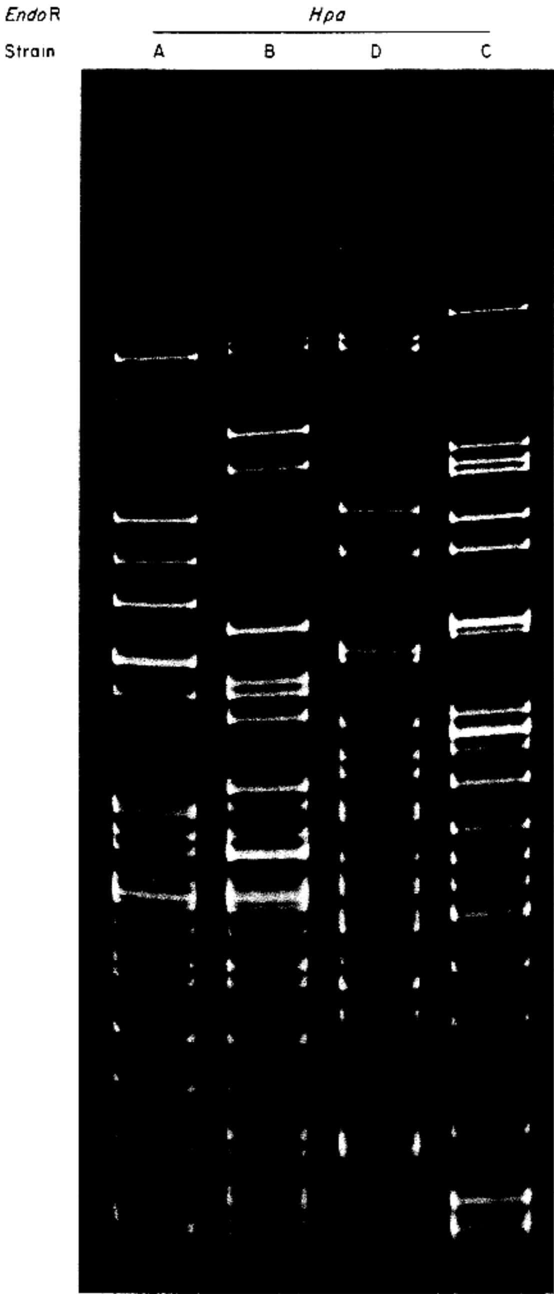

FIG. 7.

they contained. These problems were solved on the basis of two independent criteria: (1) the intensity of the bands. This was estimated for all bands by visual inspection and by densitometry of the negative pictures of the gels. For the 15 to 20 highest

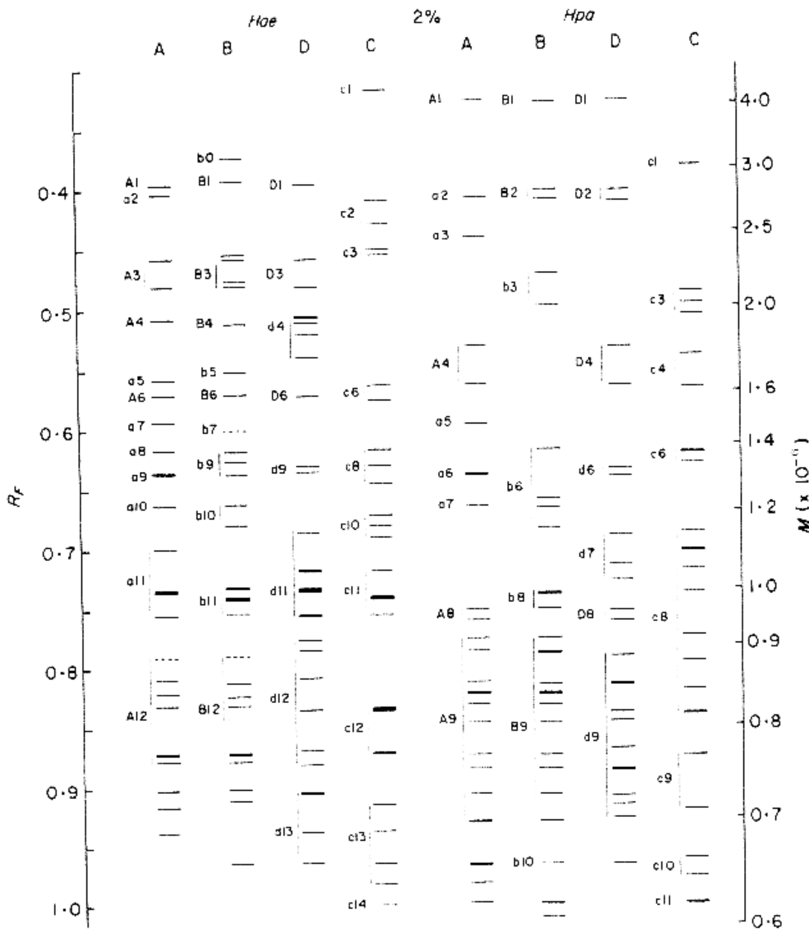

FIG. 8. Scheme of the band patterns obtained on the 2% gel. The relative mobilities and molecular weights of the fragments are indicated. Bands containing 1 fragment, 2 fragments, and more than 2 (3 or 4) fragments are indicated by a different thickness; faint bands are indicated by broken lines. Corresponding bands, or band clusters, as seen on different gels, are indicated with the same letter number combination. Capital letters indicate bands exhibiting interstrain homology (according to the criteria given in section (e) of Results), lower case letters for all other bands. Homologies between $B2_2$, $D2_2$ and a_2 of the *Hpa* pattern and between A4, B4 and $d4_2$ of the *Hae* pattern were not indicated to avoid crowding of the Figure.

This scheme, as well as those of Figs 9 to 12, only depicts the bands in the best-resolved region of the gel.

molecular weight bands a quantitative treatment was also applied (see Appendix). (2) The behavior of multiple bands on different gels. In general, bands showed slight shifts in their relative positions when run on different gels. These effects are probably due to differences in the dependence of electrophoretic mobility upon nucleotide

sequence (Zeiger *et al.*, 1972), as studied on gels of different porosity. In the case of mitochondrial DNA fragments, these effects may be associated with the presence of A + T-rich stretches of different lengths in the fragments and with their contribution to fragment flexibility. Interestingly, similar phenomena could be observed when running the same fragments on the same gel at different temperatures (Fonty *et al.*,

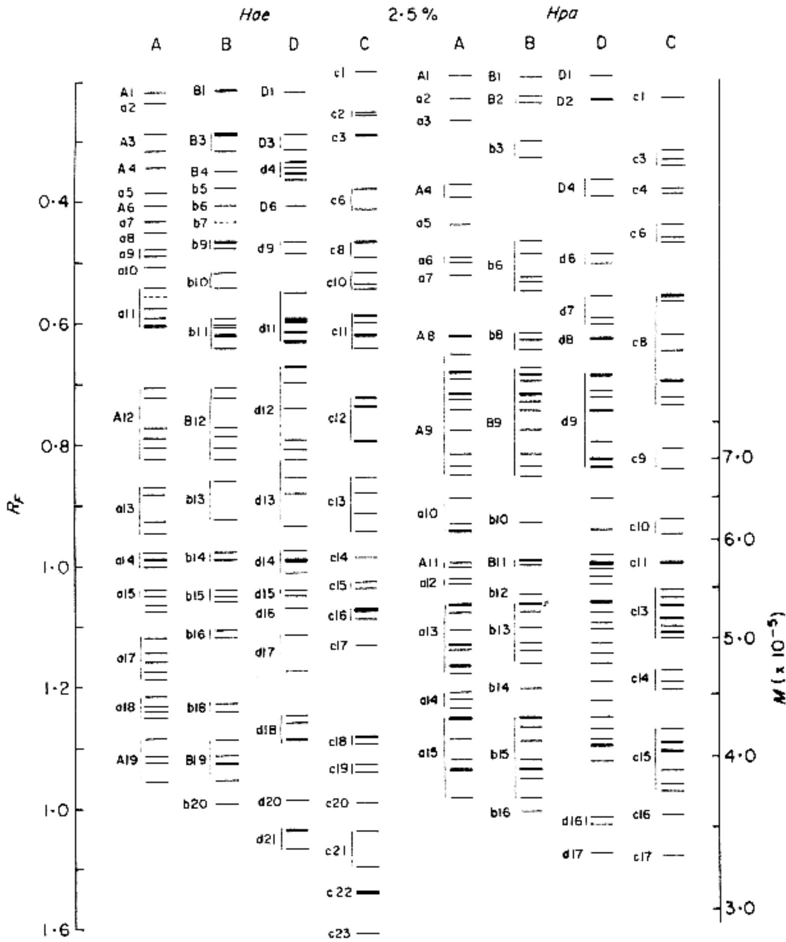

FIG. 9. Scheme of the band patterns obtained on the 2·5% gel. For other indications see legend to Fig. 8. Here, band b0 (*Hae*) of Fig. 8 has the same mobility as band B1 (*Hae*) which is now represented as a double one.

unpublished data; Allet *et al.*, 1973*a*,*b*). Because of these R_F shifts, a double band on a given gel may be resolved into a separate doublet on another gel, thus giving a clear indication on the multiplicity of the band under consideration. About 50% of the multiple bands of mitochondrial DNAs were resolved into simple bands when electrophoresis was done on another gel. Very interestingly, all the high molecular weight multiple bands of the 2% gel were either resolved, or rearranged in

clusters, on the 2·5% gel, an exception being *Hpa* band c11 which remains double on the 2, 2·5 and 3% gels; in contrast none of the low molecular weight multiple bands in the 4% to 6% gel overlaps were resolved. Therefore, most of the fragments present in the multiple bands were different in sequence; in contrast, the low molecular weight fragments present in multiple bands could be identical in sequence and correspond to duplicated sequences. The identical chromatographic behavior of some of these

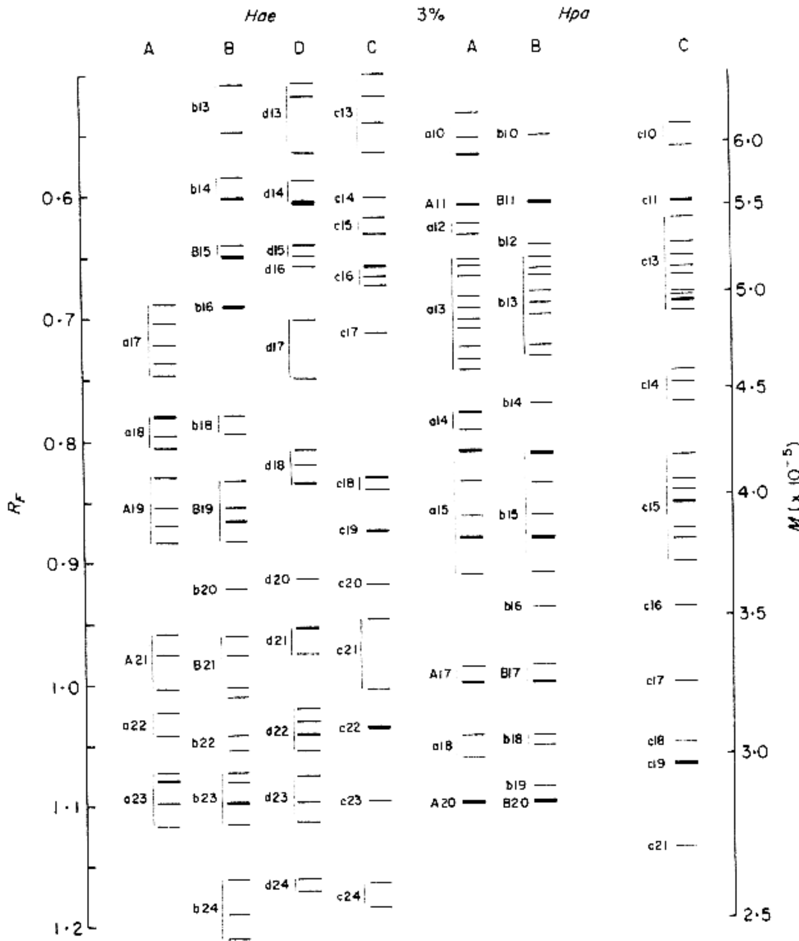

Fig. 10. Scheme of the band patterns obtained on the 3% gel. For other indications see legend to Fig. 8.

fragments on hydroxyapatite (Prunell & Bernardi, 1977) is in agreement with this idea.

(ii) Very few bands, called here faint bands, having a very much weaker intensity than expected from their molecular weights, were found in the *Hae* digests (bands b7, $A12_1$, $B12_1$ on the 2% gel; b7, $a11_2$ on the 2·5% gel; Figs 8 and 9; Figs 2 and 4). These appear to arise as the result of a specific endogenous degradation of *S. cerevisiae* DNAs (see Appendix).

(c) *Band number and fragment number*

(i) The number of bands present in the electrophoretic patterns is given in Table 1. The multiple bands which could not be resolved on any gel were counted as single ones, regardless of their intensities. Those which were resolved on a given gel, but

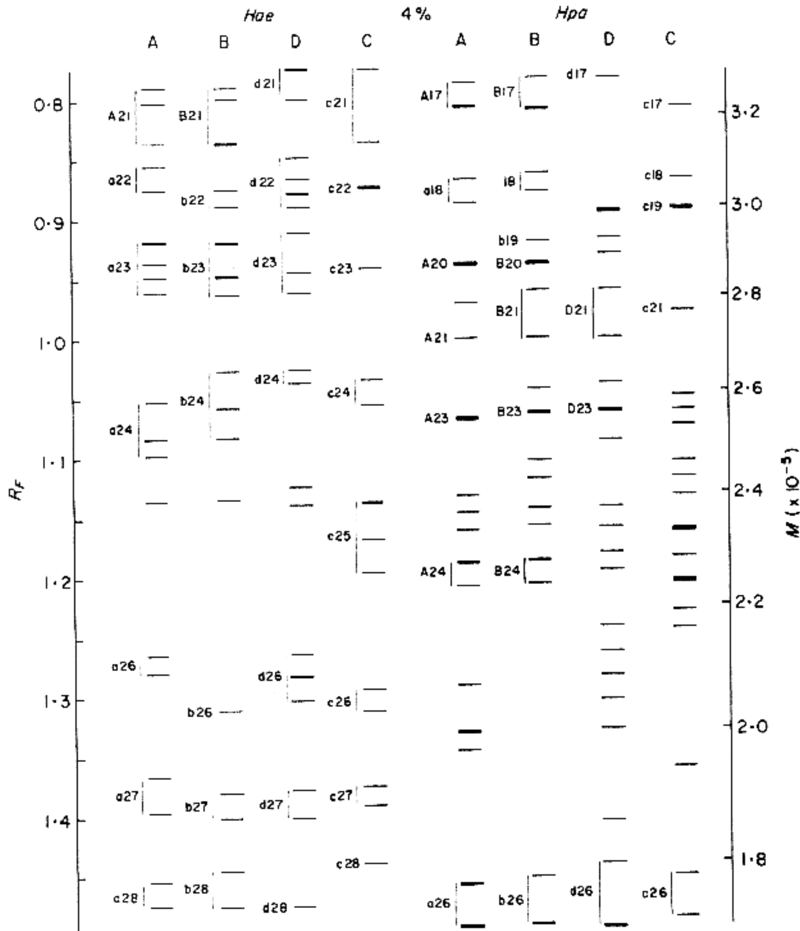

Fig. 11. Scheme of the band patterns obtained on the 4% gel. For other indications see legend to Fig. 8.

not on another one, were counted as resolved bands. The *Hpa* bands seen on the 8% gel (see section (a)) and the faint bands in the *Hae* digests were counted.

(ii) The number of all fragments, taking into account band multiplicities, is also given in Table 1. This number was higher in *Hpa* (107 to 113) compared to *Hae* digests (71 to 84), the differences being mainly due to the presence in the former of small fragments resolved on the 6 and 8% gels. The number of fragments released by the same enzyme from different DNAs was in most cases not the same.

(iii) The sums of the molecular weights of all fragments derived by *Hae* or *Hpa* hydrolysis (Table 1) provide estimates of the unit sizes of the mitochondrial genomes of each strain.

(iv) The contributions of unresolved fragments to such unit sizes are given in Table 1. This corresponds to the difference between the estimates of the genome unit

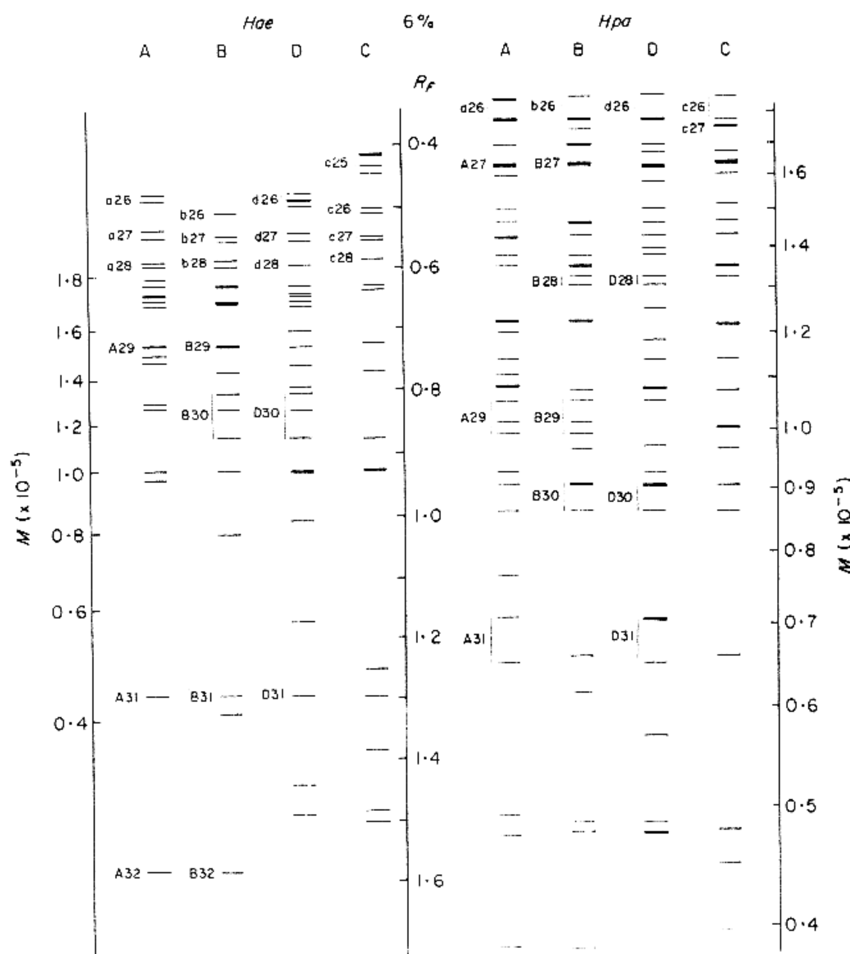

FIG. 12. Scheme of band patterns obtained on the 6% gel. The different relationships between R_F and M for *Hpa* and *Hae* digests are due to the fact that 0·5% agarose was present in the gel used for *Hpa* digest but not in that used for *Hae* digest.

sizes, as derived from the fragments (see (ii), above) and from the bands (see (i), above), respectively; the latter one clearly is a minimum estimate.

(d) *Size distribution of restriction fragments*

The fragment size distributions, as judged from gel electrophoresis, are shown in Figure 13. Data are presented in the form of plots of fragment number *versus* ln M. Plots were normalized in order to allow comparisons of different digests with the

TABLE 1

Hae *and* Hpa *restriction fragments and unit sizes of four yeast mitochondrial genomes*

	Enzyme	Strains A	B	D	C
Number of bands†	*Hae*	81	73	75	63
	Hpa	93	91	97	91
Number of fragments‡	*Hae*	84	81	83	71
	Hpa	112	107	113	107
Genome unit size, ΣM_i§	*Hae*	52	55	52	50
	Hpa	53	52	52	49
Contribution of "unresolved" fragments to genome unit size ‖	*Hae*	0·9	2·9	4·5	4·4
	Hpa	6·1	5·1	6·4	3·8

† This is the number of all "resolved" bands (see sections (b) and (c) of Results).

‡ Based on the band multiplicities given in the schemes (Figs 8 to 12).

§ M_i is the molecular weight of the fragments ($\times 10^{-6}$), as determined by gel electrophoresis. Values of 52, 50 and 46×10^6 were obtained for strains A, B and C from *Hin*dII + III digests; a value of 54×10^6 was found for strain A from *Eco*RI digest.

‖ This is the difference between the genome unit size, as calculated from the fragments, and the genome size, as calculated from the bands. Values are $\times 10^{-6}$.

distributions which should arise if mitochondrial DNAs were broken down by a random degradation into the number of fragments experimentally found. Such random distributions are described by:

$$\frac{\mathrm{d}N(M)}{\mathrm{d}M} = \frac{1}{M_\mathrm{n}} \mathrm{e}^{-M/M_\mathrm{n}}$$

or, in terms of ln M, by:

$$f(M) = \mathrm{d}N/\mathrm{d} \ln M = \frac{M}{M_\mathrm{n}} \mathrm{e}^{-M/M_\mathrm{n}},$$

where N is the number of fragments, M their molecular weights and M_n the number average molecular weight of all the fragments. In contrast with the random distribution curves, all experimental curves show a similar, characteristic multimodality. Two remarkable features of the distribution curves are (1) the great similarity of the profiles obtained with both hydrolysates from each strain, and (2) the shift towards lower molecular weights of the *Hpa* curve compared to the *Hae* curve; this is particularly evident in difference curves (not shown). The first finding points to a clustering of *Hpa* and *Hae* sites; the second one is due to the fact that the *Hpa* digest contains more small fragments than the *Hae* digest (see section (c) (ii)) and also to the fact that in over half of *Hae*, *Hpa* site clusters the *Hpa* sites are external to the *Hae* sites (Prunell & Bernardi, 1977).

(c) *Interstrain comparison of band patterns*

The problem of fragment homology among DNAs from different yeast strains was approached in two different ways.

(i) Fragments from different strains were considered homologous, i.e. to have the same nucleotide sequences (1) when they had the same electrophoretic mobility even

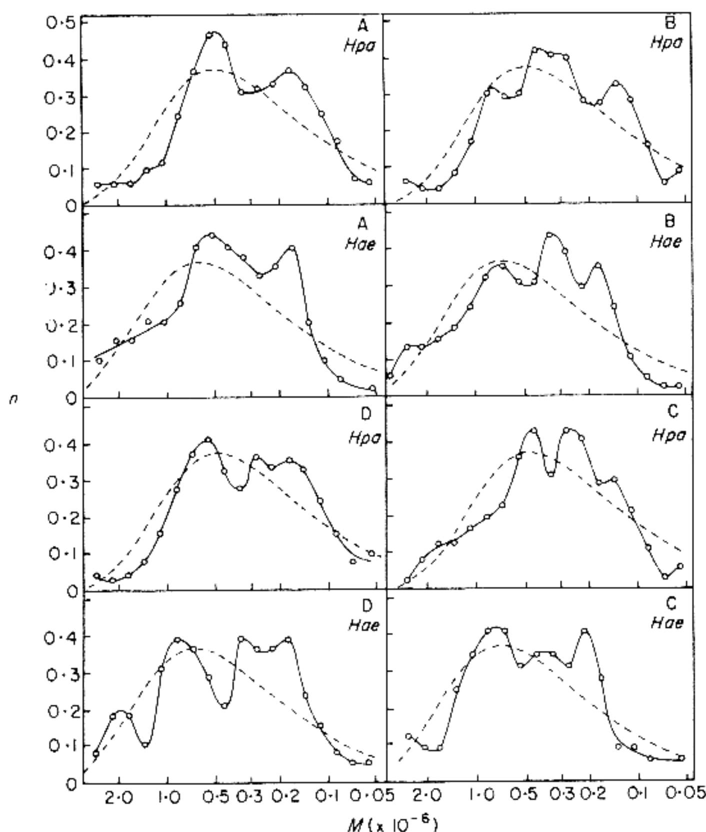

FIG. 13. Semilogarithmic plot of n *versus* M for the restriction enzyme fragments obtained from the mitochondrial DNAs.

$n = (1/N_0) \times (\Delta N/\Delta \ln M)$, where N_0 is the total number of fragments, ΔN the number of fragments between $\ln M$ and $\ln M + \Delta \ln M$. Intervals overlapping each other by 50% were taken in order to increase the number of points. The value of $\Delta \ln M = \Delta M/M$, used in the plot, was 46%; values of 23 and 92% gave similar results. The broken lines correspond to random distributions calculated for the number of fragments (experimentally found) for each digest.

on different gels (a condition which obviously could not be checked for bands outside overlaps); for instance, in the *Hpa* digests band A1 had the same mobility as bands B1 and D1 on both the 2% and the 2·5% gels; (2) when they had comparable yields (see Appendix), and/or released upon double-digestion fragments having the same mobility (Prunell & Bernardi, 1977). These criteria concerned a relatively small number of bands.

(ii) For an overall estimate of interstrain fragment homology, the following procedure was used. The percentage of the number, or of the weight, of the restriction fragments from a given DNA coinciding, within a certain $\Delta M/M$ range, with the fragments of another DNA was plotted against $\Delta M/M$ (Fig. 14). For each DNA pair, a similar plot was calculated for fragments produced by random breakage; these plots (broken lines in Fig. 14) provided an estimate of what might be called the "random homology", namely the chance of having molecular weight coincidence, within a certain $\Delta M/M$ range, in fragments having no sequence correlation.

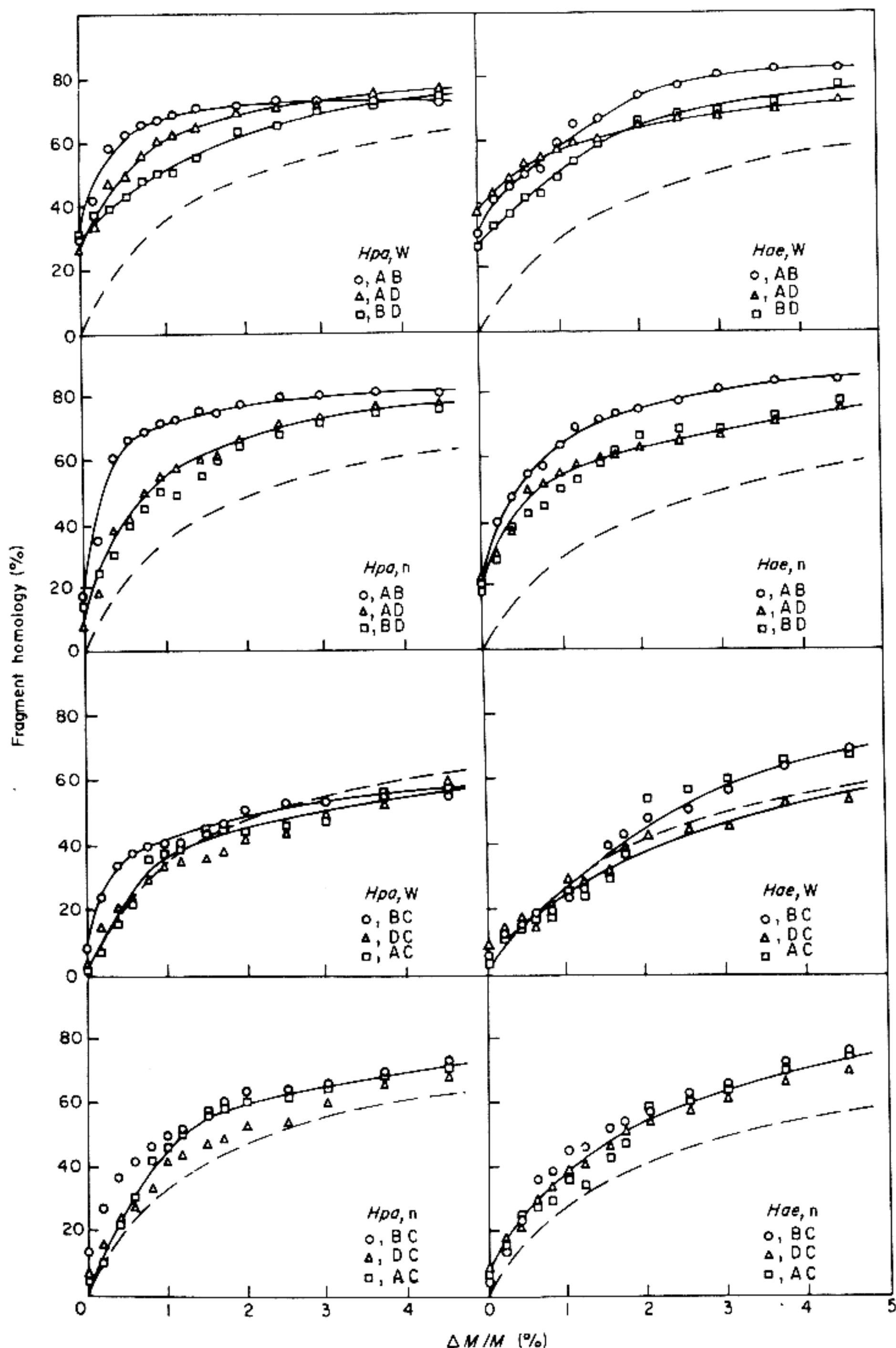

FIG. 14. Plot of the computed fragment homologies (%) *versus* $\Delta M/M$ with the 4 strains and both enzymes. The broken lines correspond to the statistical homologies. n and W indicate that the number and the weight homologies, respectively, are concerned.

In order to calculate the "random homology", a computer procedure was used to generate 200 samples of a random distribution of N_0 fragments. These samples correspond to a random cleavage of the circular genome in N_0 points. The length of the genome is taken as 1, and the positions of the cleavage points are given by random numbers generated by the computer and comprised between 0 and 1. After the N_0 cuts are introduced, the difference between the positions of two neighbor cuts, multiplied by the genome unit size, corresponds to the fragment size. The 200 fragment distributions are arranged in 100 pairs. The distributions in each pair are compared, the homology calculated with the procedure described above, and averaged for all pairs. Two values of N_0, 111 and 85, were used corresponding to the *Hpa* and *Hae* case, respectively. Random homologies by weight or by number were found to be the same.

(iii) The difference curves between experimental and "random homology" (not shown) exhibited maxima at an abscissa value of zero, or very near zero ($\Delta M/M < 0.5\%$) because of experimental error. The fact that the "random homology" curves tend to zero when $\Delta M/M$ tends to zero, whereas the experimental curves have, as a rule, a finite value and the difference curves a maximum value at (or near) $\Delta M/M = 0$, indicates that the intuitive criterion that fragments having the same electrophoretic mobility are homologous is correct on average. The values of the maxima of the difference curves represent, therefore, the percentage of fragment homology, by number or by weight; such values are given in Table 2. Estimates from *Hae* and *Hpa*

TABLE 2

Interstrain fragment homology†

Strains	*Hpa*		*Hae*	
	w	n	w	n
A–B	40	43	32	35
A–D	28	20	37	29
B–D	29	16	23	22
A–C	0	7	0	10
B–C	10	19	6	14
D–C	0	6	0	11

† Values in this Table are % homologies by fragment weight (w), or fragment number (n). Values were calculated as described in the text.

are in fair agreement with each other. The data clearly indicate that fragment homologies are much higher among *S. cerevisiae* strains than between *S. cerevisiae* and *S. carlsbergensis*. In the latter case, the homology by weight was non-significant and the very limited homology by number only concerned the smallest fragments.

(*f*) *Band patterns of* HindII + III *and* EcoRI *digests*

Figure 15 shows the fragment patterns obtained on 0·8% agarose gels with *Hind*II + III and *Eco*RI digests of mitochondrial DNAs. Fragments having molecular weights lower than about $1{\cdot}5 \times 10^6$ were resolved on a 3% polyacrylamide/0·5% agarose gel (not shown). Figure 16 presents the fragment patterns as obtained on both gels with all the digests; Table 3 gives the molecular weights of all the fragments.

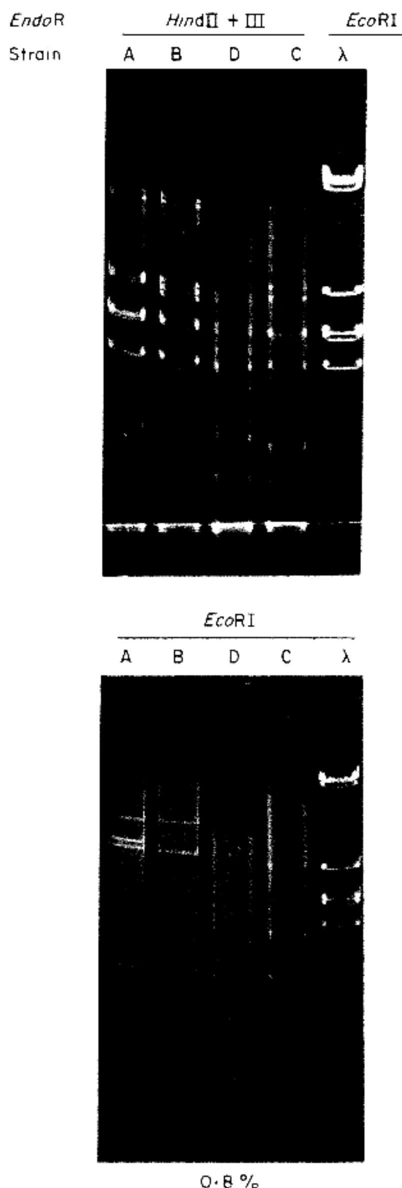

FIG. 15.

(i) *Hind*II + III digests. The genome unit sizes were estimated by summing the molecular weights of all fragments and by taking into account the existence of some double bands in the digests (Fig. 16). The values arrived at were 52, 50 and 46×10^6 for strains A, B and C, respectively (Table 3). The good general agreement of the genome unit size estimates, as obtained from *Hind*II + III and *Hae* or *Hpa*, indicates that no large *Hind*II + III fragments were missed from the gels. The absence of any background material above the top *Hind*II + III bands for DNAs which showed such material above *Eco*RI bands of even higher molecular weight, confirms this

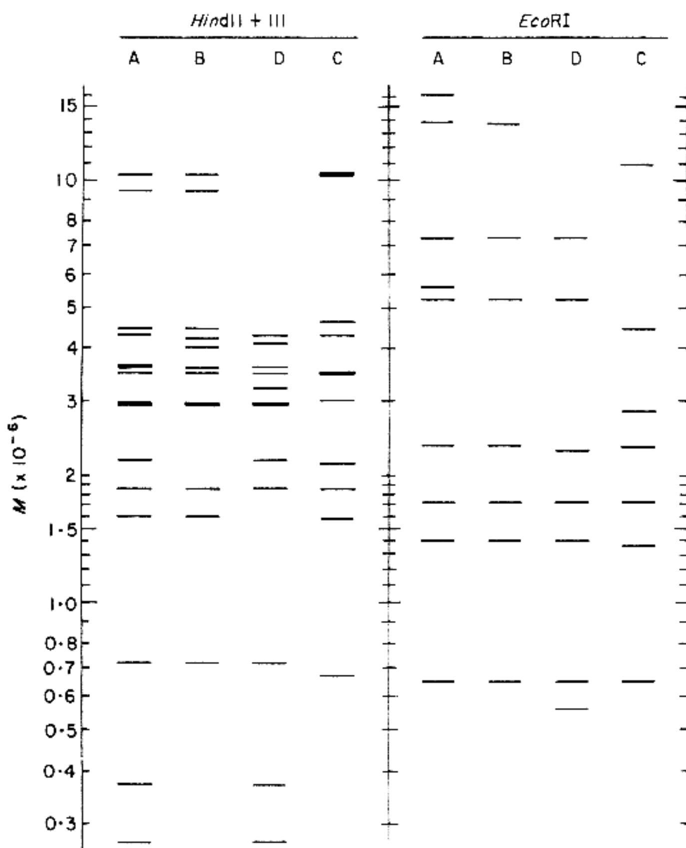

FIG. 16. Scheme of the band patterns obtained with *Hin*dII + III and *Eco*RI digests of mitochondrial DNAs.

Results obtained on 0·8% agarose, and 3% polyacrylamide/0·5% agarose were used. DNAs from strains A, D and C were those of Table A1 of the Appendix; the higher molecular weight preparation from strain A was used: the DNA preparation from strain B used here had a molecular weight of about 9×10^6. Thicker lines indicate double bands. The top bands of *Hin*dII + III digest of DNA from strain D, and of *Eco*RI digests of DNAs from strains B, D and C are missing (see text).

conclusion. In the case of DNA from strain D (which had a lower initial molecular weight than the other DNAs; see legend to Fig. 16) probably two large molecular weight fragments were missed.

(ii) *Eco*RI digests. In this case, the estimate of the genome unit size of strain A, 54×10^6, was in quite satisfactory agreement with the results obtained with the other three enzymes. In the other cases, clearly one (strains B and C) or two (strain D) fragments of high molecular weights were missed.

(iii) Band homology in *Hin*dII + III and *Eco*RI digests. The *Hin*dII + III and *Eco*RI restriction patterns obtained with the *S. cerevisiae* DNAs were characterized by the fact that most bands had either the same or a similar mobility. Several *S. carlsbergensis* DNA bands, particularly in the low molecular weight range, also showed mobilities identical or similar to those of *S. cerevisiae* DNAs. Because of the very

TABLE 3

Molecular weight of Hind*II + III and* Eco*RI restriction enzyme fragments of yeast mitochondrial DNAs*†

Strains	*Hind*II + III A	B	D	C‡	*Eco*RI A	B	D	C‡
	10·3	10·3		<u>10·3</u>	16			
	9·4	9·4			13·6	13·6		
	4·45	4·45		4·60				10·8
	4·30	4·22	4·30	4·30	7·3	7·3	7·3	
		4·05	4·10		5·6			
					5·2	5·2	5·2	
	<u>3·60</u>	3·60	3·60					4·42
								2·90
					2·35	2·35	2·30	2·35
	3·50	3·50	3·50	<u>3·50</u>	1·72	1·72	1·72	1·72
			3·22		1·40	1·40	1·40	1·36
	<u>2·95</u>	<u>2·95</u>	<u>2·95</u>	3·00	0·65	0·65	0·65	0·65
	2·17		2·17	2·12			0·56	
	1·85	1·85	1·85	1·85				
	1·60	1·60		1·58				
	0·72	0·72	0·72	0·67				
	0·37		0·37					
	0·27		0·27					
Genome unit size	52	50		46	54			

† All values are given $\times 10^{-6}$. Underlined figures refer to double bands. Figures referring to bands having identical or similar mobilities are put on the same lines (see text). Two bands were supposed to be missed in the *Hind*II + III and *Eco*RI digests of DNA from strain D; one band was missed in the *Eco*RI digest of DNA from strains B and C.

‡ Slightly different values for the DNA of the same *S. carlsbergensis* strain have been published by Christiansen & Christiansen (1976) and by Sanders *et al.* (1976) after this work was completed. The classification of the top *Hind*II + III band as a double band rests on the yield of this band (see Appendix) and on results obtained on another DNA preparation of higher molecular weight; this classification is confirmed by the independent work of Sanders *et al.* (1975).

small number of fragments produced by *Hind*II + III or *Eco*RI, fragment homology could not be estimated using the statistical approach outlined in the preceding section, and was, therefore, taken as equal to the percentage of DNA present in bands having the same mobility in different digests. In the *Hind*II + III digests, the (weight) homology values of the DNA pairs A–B, A–C and B–C (where no bands are missing) were 82, 42 and 33%, respectively, in sharp contrast to the *Hae* and *Hpa* estimates of 32 and 42% for the A–B pair, and of 0% for the A–C and B–C pairs (Table 2). This apparent paradox is largely due to the fact that *Hind*II + III fragments showing identical mobilities (and, therefore, considered homologous according to the criterion mentioned above) may be formed by *Hae* or *Hpa* fragments some of which have different mobilities in different strains. An example of such "perfect" compensations is given by the $10{\cdot}3 \times 10^6$ fragment of *S. carlsbergensis* DNA

which can only be formed by *Hae* or *Hpa* fragments that do not have identical mobilities in *S. carlsbergensis* and *S. cerevisiae* DNAs. "Imperfect" length compensations, on the other hand, may account for the large number of fragments having similar yet not identical mobilities in different digests. The overall similarity of the *Hin*dII + III and *Eco*RI restriction patterns exhibited by different DNAs has an important implication as far as gene order in the corresponding genomes is concerned (see Discussion, section (b) (ii)).

4. Discussion†

(a) *Unit size and physical size of the mitochondrial genome*

(i) *Unit size*

The sum of the molecular weights of all fragments derived by *Hae* or *Hpa* hydrolysis (Table 1) provides an estimate of the unit size of the mitochondrial genomes under investigations: this is equal to 49 to 53 $\times 10^6$ for the different strains studied here. It should be noted that (1) the *Hae* estimates should be increased by $0{\cdot}8 \times 10^6$ and the *Hpa* estimates by 1×10^6, since these amounts of DNA were present in the digests as small, single-stranded oligonucleotides detectable by hydroxyapatite chromatography, but not by gel electrophoresis (Prunell & Bernardi, 1977); (2) the *Hin*dII + III and *Eco*RI results concerning strain A are in good agreement with *Hae* or *Hpa* data (compare Tables 1 and 3); (3) the genome unit size of strain B, as estimated from the *Hae* digest, is much higher than those obtained from *Hpa* or *Hin*dII + III digests: in all likeliness, this overestimate is due to the wrong assessment of multiplicity in a high molecular weight band; (4) the discrepancy between genome unit size estimates of *S. carlsbergensis* as obtained from *Hae* or *Hpa versus Hin*dII + III might also be due to errors in band multiplicity assessments in *Hae* or *Hpa* patterns; if so, the real number of fragments originated by *Hae* (which is of importance; see Prunell & Bernardi, 1976) is comprised between the number of bands, 63, and the number of fragments, 71 (Table 1).

A very significant difference in unit size exists between the mitochondrial genomes of the *S. cerevisiae* strains used here and that of the *S. carlsbergensis* strain; this has been recently confirmed by the electron microscopic work of Christiansen & Christiansen (1976). Significant differences also exist among the genome unit sizes of the *S. cerevisiae* strains; the case of DNAs from strains A and B, as analyzed by *Hin*dII + III, and of some recombinant DNAs from A × B crosses (Fonty *et al.*, unpublished data) is particularly meaningful because of the identical mobility of most fragments. It is important to remark that, while the unit sizes of the wild-type mitochondrial genomes examined here differ from each other, the size range is rather narrow; considering the similarity of properties strongly dependent upon the amount of A+T-rich spacers (such as G + C content, buoyant density, melting and circular dichroism spectra) in DNAs derived from different strains (Bernardi *et al.*, 1970; Bernardi & Timasheff, 1970), this indicates (1) that the total amount of spacer sequences is relatively constant in different strains and kept under some control, and suggests (2) that most of the mitochondrial genes are indispensable for respiratory competence.

† This Discussion is best followed if the conclusions reached in the following paper (Prunell & Bernardi, 1977) are known.

(ii) *Genome unit size and physical size; homogeneity and complexity of the mitochondrial genome*

The good agreement of the genome unit sizes of Table 1 with the physical size, 50×10^6, estimated by electron microscopy (Hollenberg *et al.*, 1970), has three important implications: (1) it unequivocally confirms the validity of the electron microscopy results and the fact that the 25-μm circular supercoiled DNA molecules indeed correspond to mitochondrial genome units: such molecules, released by osmotic shock from mitochondria directly on the electron microscope grid (Hollenberg *et al.*, 1970), could never be observed in other laboratories nor could they be isolated. This conclusion (Bernardi *et al.*, 1975) was also reached by Borst (1975) and then confirmed by the very recent data of Morimoto *et al.* (1975) and Christiansen & Christiansen (1976). (2) It rules out not only the unlikely possibility that mitochondrial genes are distributed over more than one physical unit (a situation existing in some plant viruses), but also the existence of a "discrete" heterogeneity, namely of a small number of different genomes present in comparable amounts. In fact, it indicates that the mitochondrial genome of a given yeast strain is quite homogeneous in terms of nucleotide sequences (strictly speaking in terms of distribution of restriction sites). It should be stressed, however, (a) that some small sequence duplications are strongly suggested by the lack of resolution on different gels of multiple bands in the low molecular weight range and by the same chromatographic behavior on hydroxyapatite of the fragments contained in such double bands; these fragments appear to lack spacers and to correspond to gene fragments (Prunell & Bernardi, 1977); (b) that other larger sequence duplications are possible (see section (b), (iii)); (c) that neither the agreement between genome unit size and physical size nor the stoichiometry of the fragments (see Appendix) rule out the possibility that a certain percentage of spontaneous wild-type mutant genomes contaminate the predominant genome; in fact, such mutants would behave like the spontaneous "petite" mutants which are present in our wild-type cell cultures, namely they would affect the stoichiometry of fragments below our detection level and contribute to the background smear; these spontaneous wild-type mutants are, in fact, likely to be present and to provide the material for the divergence and evolution of the mitochondrial genome (see section (b), below). (3) The general conclusion on the sequence homogeneity of the mitochondrial genome from a given strain seems to suggest the existence of a "unique" (Britten & Kohne, 1968) nucleotide sequence in mitochondrial DNA or, in other words, a perfect correspondence between the "complexity" (Britten & Kohne, 1968) of the mitochondrial genome and its physical size. This would, indeed, be the case if the restriction sites were randomly distributed over the mitochondrial genome. The actual situation is, however, quite different. As far as *Hind*II + III and *Eco*RI are concerned, these enzymes insert too few breaks to be of any help; on the other hand, the very numerous *Hae* and *Hpa* sites, G-G-C-C and C-C-G-G (K. Murray & A. Morrison, unpublished results; Garfin & Goodman, 1974), are completely absent in the A+T-rich spacers and are very strongly clustered in particular segments of the mitochondrial genome (Prunell & Bernardi, 1977). Our data do not say anything, therefore, about the existence of a "unique" nucleotide sequence in mitochondrial DNA and, clearly, are perfectly compatible with the existence of repetitive sequences inside the enzyme-resistant restriction fragments. It should be pointed out here that the previous claim of a "unique" nucleotide sequence in mitochondrial DNA (Hollenberg *et al.*, 1970; Christiansen *et al.*, 1974), resting on the agreement between

the kinetic complexity and the physical size of the mitochondrial genome, is extremely controversial because of the large number of unresolved problems which affect the renaturation kinetics of this DNA (Christiansen *et al.*, 1971,1974; Prunell & Bernardi, 1977).

(b) *Homology and divergence of the mitochondrial genome*

(i) *Sequence homology and restriction fragment homology in mitochondrial genomes*

As shown in Table 2, *Hae* and *Hpa fragment homology* (by weight) between *S. cerevisiae* and *S. carlsbergensis* DNAs is equal to zero for the DNA pairs A–C and D–C, and lower than 10% for the pair B–C. In apparent full contradiction with these results, a number of findings indicate a high level of *sequence homology* in the mitochondrial DNAs investigated here: (1) all genomes released approximately the same number of *Hae* and *Hpa* fragments (Table 1) and, therefore, contained approximately the same number of restriction sites (or restriction site clusters; Prunell & Bernardi, 1977); this is so much more significant as such a number is extremely high compared to statistical expectations (see section (ii) (2), below). In all cases, *Hpa* fragments were more numerous than *Hae* fragments, revealing a common bias in all DNAs. (2) The size distributions of the restriction fragments were similar (Fig. 13), an indication that *Hae* and *Hpa* sites (or site clusters) are similarly distributed in the genomes. (3) In the case of strains B and C, *Hpa* fragments of similar size, as separated by chromatography on agarose columns, have a similar base composition (Prunell & Bernardi, 1977), suggesting that fragments having similar lengths may be homologous in sequence. A stronger indication of sequence homology for fragments of similar size from the DNAs of strains B and C comes from their similar chromatographic behavior on hydroxyapatite columns (Prunell & Bernardi, 1977).

All these results are not surprising in view of the essential identity of all physical and chemical properties (buoyant density, G+C level, melting, pyrimidine isostichs, A+T-rich spacers, restriction site clusters, G+C-rich clusters) shared by the DNAs under consideration and are in general agreement with the finding (Groot *et al.*, 1975) that mitochondrial DNAs from the same *S. carlsbergensis* strain used here and a *S. cerevisiae* strain show 100% homology, by DNA–DNA hybridization, with no mismatch in the hybrids. An explanation of the apparent paradox represented by the fact that DNAs from *S. carlsbergensis* and *S. cerevisiae* (to take the extreme case) show 100% sequence homology and no *Hae* or *Hpa* fragment homology is given in the following two subsections.

(ii) *Origin of differences in the fragment patterns*

Two, not mutually exclusive, possibilities can explain the origin of the differences found in the fragment patterns exhibited by the mitochondrial DNAs investigated here: (1) point mutations and/or base modifications (methylation, glucosylation) *at* the restriction sites. (2) Deletions and/or additions *between* restriction sites.

The first possibility can be ruled out as a mechanism accounting for more than a minor part of the changes on the basis of the following arguments: (1) practically all fragments of *S. carlsbergensis* DNA have electrophoretic mobilities different from those of the *S. cerevisiae* DNAs. This implies that at least half of the restriction sites are different in the two strains. Assuming a random point-mutation mechanism and a single base-pair change per two sites, this means at least one mutated nucleotide out

of eight (the number of nucleotides present in two sites) in the mitochondrial DNAs. Such a situation should give rise to a measurable mismatch in the DNA–DNA hybrids ($\Delta t_m \simeq 10$ deg.C), whereas none was observed (Groot *et al.*, 1975). Obviously, the degree of mismatch would be even larger if the restriction sites were changed by an addition/deletion mechanism *at* the sites instead of a point-mutation mechanism. An additional remark is that the large number of point mutations required to bring about the observed pattern changes would certainly affect rRNA and tRNA genes, which represent more than 10% of the genome, and also lead to an inacceptably large divergence time (over 10^8 years, at a point-mutation rate of 10^{-7} per generation) between strains which show 100% homology in their nuclear genomes as well (Groot *et al.*, 1975). (2) Since the number and the size distribution of the restriction fragments is very similar in DNAs from different strains (Table 1 and Fig. 13), the disappearance of "old" restriction sites should be accompanied by the formation (caused by the same point-mutation mechanism) of an approximately equal number of "new" restriction sites having a very similar distribution on the genome. This very unlikely explanation becomes totally unacceptable if we consider: first, that the number of *Hae* or *Hpa* sites in mitochondrial DNAs is more than 20 and 40 times, respectively, the number, 5, expected for a statistical DNA having the same base composition (Prunell & Bernardi, 1977; Appendix), and that a random mutation process cannot preserve such a number since it will randomize the nucleotide sequences; and, second, that restriction sites are clustered in mitochondrial DNA (Prunell & Bernardi, 1977), a finding which implies that several point mutations at neighboring sites are needed to change the restriction fragment patterns. (3) As far as base modifications *at* the restriction sites are concerned, glucosylation is non-existent in yeast mitochondrial DNA (Bernardi *et al.*, 1970), and methylation can be ruled out as an explanation for the lack of fragment homology. In fact, recombinant mitochondrial genomes, such as those studied in diploids issued from A×B crosses (Bernardi, 1975; Fonty *et al.*, unpublished data), show *Hpa* patterns where different combinations of the patterns shown by the DNAs of the parent strains are represented; this suggests that no strain-specific modifications took place at any of the restriction sites present on either DNA.

We conclude, therefore, that the largely predominant mechanism underlying the changes in restriction site distributions is one involving deletions and additions between such sites.

(iii) *Localization of deletions and additions in the restriction fragments*

It should be recalled here that all *Hae* fragments and most *Hpa* fragments contain a gene and a spacer element flanked by G+C-rich segments containing *Hae* and *Hpa* site clusters. Under these circumstances, the important question concerns which one of these elements is the seat of deletions and additions. It is obvious that such phenomena cannot affect the genes themselves since in the case of the DNAs from *S. carlsbergensis* and *S. cerevisiae* all allelic genes should have different lengths. Deletions and additions must, therefore, affect either the A+T-rich spacers or the G+C-rich segments (these comprise the *Hae*, *Hpa* site clusters and the G+C-rich clusters; Prunell & Bernardi, 1977). In both cases, one could expect an absence of mismatch in the DNA–DNA hybrids because of precise pairing of the gene sequences present in the fragments and of foldback pairing in the extra lengths of A+T-rich spacers or G+C-rich segments, both of which are known to contain hyphenated palindromic

sequences (Ehrlich *et al.*, 1972; Prunell & Bernardi, 1977). It is very likely that the A+T-rich spacers are the seat of changes in length of restriction fragments for the following reasons: (1) the A+T-rich spacers represent 50% of the genome, the G+C-rich segments less than 10%; if the number of these two elements is the same, as seems to be the case (Prunell & Bernardi, 1976), the G+C-rich segments are just too short to allow for the changes in length observed; in the two cases where length changes in sequence-homologous fragments could be assessed, these were much longer than the average length of G+C-rich segments (Prunell & Bernardi, 1977). (2) The variations in the unit sizes of mitochondrial genomes of different strains may reach values of at least 4×10^6; these large changes can be easily understood in terms of additions and deletions in the 25×10^6 of A+T-rich spacers whereas they are much too large for the 3×10^6 or so of G+C-rich segments; in addition, one-third of the DNA in these segments (the *Hpa* site clusters) is known not to change in amount in the DNAs from strains B and C (Prunell & Bernardi, 1977). The alternative explanation that changes in genome size are due to deletions and additions of very long genome segments encompassing a number of restriction fragments does not explain the changes in length of all restriction fragments. (3) If the hypothesis that the G+C-rich segments correspond to regulatory sequences is correct (Prunell & Bernardi, 1977), it is obvious that these sequences cannot be altered without functional losses; the preservation of the amount of *Hpa* clusters in strains B and C, mentioned above is in keeping with this concept; in contrast, the A+T-rich spacers are not likely to play the same role nor to be under the same evolutionary constraints.

In connection with the localization of additions and deletions in the A+T-rich spacers, it is worth recalling (see section (i), above) that *Hae* or *Hpa* fragments of *S. cerevisiae* and *S. carlsbergensis* DNA sharing common sequences have similar, yet not identical, electrophoretic mobilities, a finding which indicates that additions and deletions are short, relative to fragment size. This explains how compensations between additions and deletions keep within a narrow range not only the unit size of the diverged genomes, but also a number of *Hin*dII + III fragments. The length compensations found in the latter have an important implication, namely that gene order is largely preserved in the genomes examined; in fact if extensive gene translocations had taken place in these genomes one would not find that most of the fragments have either the same or a similar size, whereas this is easily explained by compensations in spacer lengths in otherwise homologous genome segments.

The variability in length of A+T-rich spacers in the mitochondrial genomes of different strains implies that duplicated genes present in a given genome may be found on *Hae* or *Hpa* fragments of different lengths because of the different length of the neighboring spacers.

(iv) *Mechanisms underlying deletions and additions in the restriction fragments*

Several mechanisms may conceivably be responsible for the deletion-addition events occurring in the A+T-rich spacers of the mitochondrial genome of yeast. The basic mechanism which we favor is one involving unequal crossing-over events in the A+T-rich spacers. This is being investigated in our laboratory by sequence analysis of spacers from cloned fragments of mitochondrial DNAs from different strains. At the present time, we can only mention a number of possibilities concerning the unequal crossing-over events. These might occur: (1) within single spacer elements, leading to local deletions; according to this explanation, the genomes under

investigation have diverged from a common ancestral genome because of the different localizations and extents of deletions; (2) at different spacer elements belonging to the same genome unit: such a process leads to the deletion of a genome segment and the formation of a "novel" spacer formed by the fusion of the "old" ones: this process should, however, essentially lead to "petite" mutants, as originally suggested†; (3) at spacer elements of different genome units: (a) if spacer elements flanking non-homologous genes are involved, this leads to exchanges of non-homologous genome segments; the most frequent result of this process will be the formation of defective genomes (with both deletions and additions of genome segments) and therefore of "petite" mutants; it is conceivable that deletions only affecting genes non-essential for cell respiration, as well as gene duplications and translocations, may be caused by this process, which leads in any case to additions and deletions in the spacers where crossing-over events took place; (b) if spacer elements flanking homologous genes are involved, exchanges of spacer sequences or of homologous genome segments will occur; in the latter case, additions and deletions in the spacers can be the result of either an exchange accompanied by segment inversion, or of incorrect pairing of nucleotide sequences in the spacers flanking similarly oriented segments; interestingly this situation may arise not only if there is a general possibility of crossing-over events anywhere in the spacers (as implicitly supposed above), but also if this possibility is limited to spacers flanking homologous genes.

Two of the mechanisms just described ((2) and (3) (a)) predominantly lead to the formation of defective genomes. This stresses the fundamental analogy of the mechanisms underlying the "petite" mutation and the evolution of the wild-type genome, unequal crossing-over events at the repetitive nucleotide sequences being responsible for both phenomena. In this connection, it should be mentioned that recombinational events seem also to take place in the G+C-rich segments, in which case, however, they systematically lead to "petite" mutant genomes (manuscript in preparation).

The other two mechanisms ((1) and (3) (b)) can account for the divergence of wild-type genomes. Both of them preserve gene order, in agreement with the indication given by *Hin*dII + III fragment patterns. It is most likely, however, that mechanism (1) is probably hindered by topological contraints and/or lack of adequate intra-spacer sequence homology and that mechanism (3) (b) is the predominant mechanism responsible for genome divergence. In fact, the restriction patterns of diverged mitochondrial genomes are identical to those of mitochondrial genomes of diploids issued from A×B crosses in which case clearly intergenomic recombination occurs (Bernardi, 1975; Fonty *et al.*, unpublished data). A very important feature of the recombination of different genome units is their extremely high rate (compared to point-mutation rate) not only in zygotes (Bernardi, 1975; Fonty *et al.*, unpublished data), but also in vegetative cells (Williamson & Fennell, 1974).

(v) *General implications*

In summary, what has been shown in the present work is that, in an interspersed system of genes and internally repetitive (Ehrlich *et al.*, 1972; and manuscript in preparation) spacers, evolution goes about essentially by recombination, this process being several orders of magnitude faster than point mutation. This conclusion is very interesting because it may apply to the nuclear genome of eukaryotes which is also

† See footnote to p. 18.

made up of genes and interspersed repetitive sequences. It is conceivable that recombination processes taking place at a high rate at such sequences have played a very important role in the evolution of eukaryotes. A key feature of evolution by recombination at interspersed repetitive sequences is that it can generally affect the genome essentially by translocations, inversions, duplications and deletions of genes whereas it is very conservative as far as gene nucleotide sequences are concerned: only the much slower process of point mutation is operative on these. A major difference between prokaryotes and eukaryotes might precisely be that eukaryotes developed an interspersed system of repetitive sequences and this allowed them to evolve much more rapidly than prokaryotes.

We wish to thank Dr Alberto Bernardi and Dr Pierre Yot for the gift of *Eco*RI samples, Dr E. Cassuto for the gift of λ DNA, M. François Goutorbe for his help in the calculation of fragment homology, Drs S. D. Ehrlich and C. V. Lusena for critical reading of this paper, Miss Louise Marty for the preparation of SV40 DNA, Miss Martine Brient for typing this manuscript and M. Philippe Breton for the drawing and photographic work.

REFERENCES

Allet, B., Jeppesen, P. G. N., Katagiri, K. J. & Delius, H. (1973*a*). *Nature (London)*, **241**, 120–122.

Allet, B., Katagiri, K. J. & Gesteland, R. F. (1973*b*). *J. Mol. Biol.* **78**, 589–600.

Bernardi, G. (1975). In *Fed. Eur. Biochem. Soc. Proc. Tenth Meeting*, Paris, July 1975 (Bernardi, G. & Gros, F., eds), vol. 38, pp. 41–56. North-Holland, Amsterdam.

Bernardi, G. & Timasheff, S. N. (1970). *J. Mol. Biol.* **48**, 43–52.

Bernardi, G., Carnevali, F., Nicolaieff, A., Piperno, G. & Tecce, G. (1968). *J. Mol. Biol.* **37**, 493–505.

Bernardi, G., Faurès, M., Piperno, G. & Slonimski, P. P. (1970). *J. Mol. Biol.* **48**, 23–42.

Bernardi, G., Piperno, G. & Fonty, G. (1972). *J. Mol. Biol.* **65**, 173–189.

Bernardi, G., Prunell, A. & Kopecka, H. (1975). In *Molecular Biology of Nucleocytoplasmic Relationships* (Puiseux-Dao, S., ed.), pp. 85–90. Elsevier, Amsterdam.

Borst, P. (1975). In *Molecular Biology of Nucleocytoplasmic Relationships* (Puiseux-Dao, S., ed.), pp. 81–84. Elsevier, Amsterdam.

Britten, R. J. & Kohne, D. E. (1968). *Science*, **161**, 529–540.

Christiansen, C., Leth Bak, A., Stenderup, A. & Christiansen, G. (1971). *Nature New Biol.* **231**, 176.

Christiansen, C., Christiansen, G. & Leth Bak, A. (1974). *J. Mol. Biol.* **84**, 65–82.

Christiansen, G. & Christiansen, C. (1976). *Nucleic Acids Res.* **3**, 465–476.

Clark-Walker, G. D. & Miklos, G. L. G. (1974). *Genet. Res.* **24**, 43–57.

Danna, K. & Nathans, D. (1971). *Proc. Nat. Acad. Sci., U.S.A.* **68**, 2913–2917.

Danna, K. J., Sach, G. H. & Nathans, D. (1973). *J. Mol. Biol.* **78**, 363–376.

Ehrlich, S. D., Thiery, J. P. & Bernardi, G. (1972). *J. Mol. Biol.* **65**, 207–212.

Garfin, D. E. & Goodman, H. M. (1974). *Biochem. Biophys. Res. Commun.* **59**, 108–116.

Grivell, L. A., Reijnders, L. & Borst, P. (1971). *Biochim. Biophys. Acta*, **247**, 91–103.

Groot, G. S. P., Flavell, R. A. & Sanders, J. P. M. (1975). *Biochim. Biophys. Acta*, **378**, 186–194.

Hirt, B. (1967). *J. Mol. Biol.* **26**, 365–369.

Hollenberg, C. P., Borst, P. & Van Bruggen, E. F. J. (1970). *Biochim. Biophys. Acta*, **209**, 1–15.

Kopecka, H. (1975). *Biochim. Biophys. Acta*, **391**, 109–120.

Morimoto, R., Lewin, A., Hsu, H.-J., Rabinowitz, M. & Fukuhara, H. (1975). *Proc. Nat. Acad. Sci., U.S.A.* **72**, 3868–3872.

Piperno, G., Fonty, G. & Bernardi, G. (1972). *J. Mol. Biol.* **65**, 191–205.

Prunell, A. & Bernardi, G. (1973). *J. Biol. Chem.* **248**, 3433–3440.

Prunell, A. & Bernardi, G. (1974). *J. Mol. Biol.* **86**, 825–841.

Prunell, A. & Bernardi, G. (1977). *J. Mol. Biol.* **110**, 53–74.

Sanders, J. P. M., Borst, P. & Weijers, P. J. (1975). *Mol. Gen. Genet.* **143**, 53–64.
Sharp, P. A., Sugden, B. & Sambrook, J. (1973). *Biochemistry*, **12**, 3055–3063.
Smith, H. & Nathans, D. (1973). *J. Mol. Biol.* **81**, 419–423.
Williamson, D. H. & Fennell, D. J. (1974). *Mol. Gen. Genet.* **131**, 193–207.
Zeiger, R. S., Salomon, R., Dingham, C. W. & Peacock, A. C. (1972). *Nature (London)*, **238**, 65–69.

Appendix

Yield of Restriction Fragments from Yeast Mitochondrial DNA

Ariel Prunell, François Goutorbe, François Strauss and Giorgio Bernardi

(a) *Preliminary considerations*

Yeast mitochondrial DNA can only be prepared, so far, in a more or less degraded form. DNA preparations having molecular weights of only 3·5 to $5{\cdot}2 \times 10^6$ (the genome unit size being about 50×10^6) were used in most of this work (Table A1). In spite of such extensive degradation, all the *Hae* and *Hpa* fragments could be

Table A1

Molecular weights of the mitochondrial DNAs before and after restriction enzyme degradation

Initial DNAs	A†	A†	Strains B	D	C
$s_{20,w}$	26	17·9	18·3	20·5	20·9
M_w (from sedimentation)‡	9	3·5	3·6	4·9	5·2
M_w (from eqn (A9))	9	3·4	2·2	4·3	4·2
Degraded DNA (*Hpa*)					
$s_{20,w}$		8·6	8·6	8·6	9·7
M_w (from sedimentation)§		0·48	0·48	0·48	0·71
M_w (from electrophoresis)‖		1·14	1·28	1·27	1·07
M_n (from electrophoresis)		0·47	0·52	0·50	0·53
M_w/M_n (from electrophoresis)		2·4	2·5	2·5	2·0
Degraded DNA (*Hae*)					
M_w (from electrophoresis)‖		1·18	1·35	1·20	1·53
M_n (from electrophoresis)		0·62	0·68	0·62	0·71
M_w/M_n (from electrophoresis)		1·9	2·0	1·9	2·2

† Values ($\times 10^{-6}$) for 2 different preparations are given.

‡ M_w was calculated using the relationship of Richards & Bernardi (Bernardi & Sadron, 1964; Prunell & Bernardi, 1973). The relationship of Studier (1965) would lead to molecular weights higher by 50%.

§ M_w was calculated using the relationship of Prunell & Bernardi (1973).

‖ Using the well-known relationships:

$$M_w = \Sigma M_i^2 / \Sigma M_i$$

$$M_n = \Sigma M_i / \Sigma_i.$$

M_i being the molecular weight of the fragment i as determined by gel electrophoresis.

detected since the size of the largest fragments was equal to only 4×10^6 in these digests, whereas for *Eco*RI and *Hind*II + III higher molecular weight preparations had to be used to obtain satisfactory results (see Results, section (f) in the main text). Expectedly, however, the amount of DNA in the *Hae* and *Hpa* bands was not proportional to the molecular weights of the corresponding fragments, but showed a relative decrease with increasing molecular weight of the fragments. This effect, already evident upon inspection of Figures 2, 3 and 7 of the main text, is clearly demonstrated by a comparison of the molecular weights of *Hpa* digests, as obtained from their sedimentation coefficient with those calculated from the weight-average molecular weight of the fragments as determined from gel electrophoresis data (Table A1). Such a comparison indicates that the former always are lower than the latter by a factor of at least 2 for the *S. cerevisiae* DNAs. The higher yield of the highest molecular weight fragments in the *Hpa* digest of *S. carlsbergensis* DNA (see section (c), below) accounts for the relatively higher molecular weight (as estimated from sedimentation) of this digest.

The effect of the initial degradation of mitochondrial DNA on fragment yield was therefore studied using the approach outlined in the following section.

(b) *Restriction fragment yield from a randomly degraded DNA*

If the DNA degradation preceding the restriction enzyme digestion is random, the yield of the DNA fragments released by the restriction enzyme can be calculated from the molecular weights of the intact and of the degraded DNA. The probability R_i for a given restriction fragment of N_i nucleotides to be intact is given by:

$$R_i = (1 - p)^{N_i - 1}, \tag{A1}$$

where p is the probability for any bond to be broken; p can be expressed as the percentage of bonds broken in the intact DNA:

$$p = \rho / N_0, \tag{A2}$$

where N_0 is the number of nucleotides in the intact molecule. ρ, the number of random breaks, can be calculated according to Charlesby (1954):

$$M_0 / M_w = \rho^2 / 2\ (e^{-\rho} + \rho - 1), \tag{A3}$$

where M_w is the weight average molecular weight of the randomly degraded DNA and M_0 is the molecular weight of the intact DNA of N_0 nucleotides. Equation (A3) can be approximated by:

$$M_0 / M_w = \rho^2 / 2\ (\rho - 1)\ . \tag{A4}$$

Equation (1) can be approximated by:

$$R_i = e^{-\rho \times M_i / M_0}, \tag{A5}$$

where M_i is the molecular weight of a fragment containing N_i nucleotides.

R_i is the yield of fragment i and can be physically expressed as the ratio of the amount of DNA in the corresponding band to the amount that would be obtained if intact DNA had been used. Figure A1 shows the straight lines relating log R_i to M_i for different M_w values of the degraded DNA and for $M_0 = 50 \times 10^6$. It can be seen from Figure A1 that, for instance, the yield of a fragment having a molecular weight of 4×10^6 (corresponding to the largest fragments obtained by *Hae* or *Hpa* in this work) is equal to 20% if the molecular weight of the degraded DNA is 5×10^6. If Q_i is the

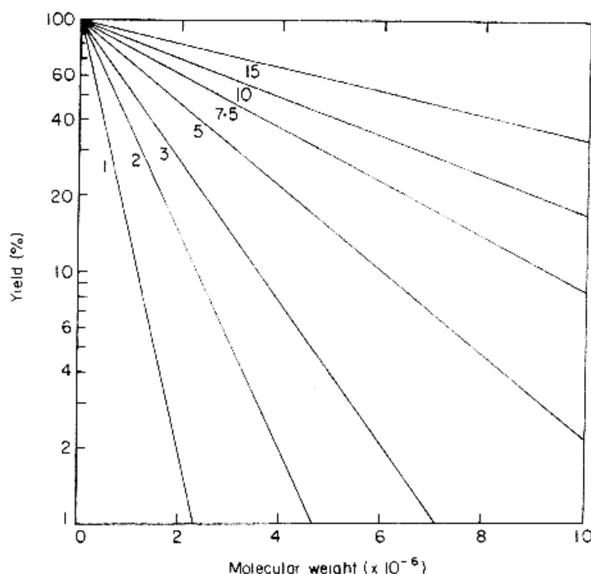

FIG. A1. Plots of the yield of the fragments, on a logarithmic scale, *versus* their molecular weight for different values of the molecular weight, M_W, of the randomly degraded starting DNA.

amount of DNA, per fragment, in band i, and M_i the corresponding molecular weight, the ratio Q_i/M_i is proportional to the yield, R_i, of the fragment, i.e.:

$$Q_i/M_i = \alpha R_i \,. \tag{A6}$$

Equation (A5) shows that the slope, β, of log Q_i/M_i *versus* M_i is:

$$\beta = -\rho/M_0 = -1/M_n \,, \tag{A7}$$

where M_n is the number average molecular weight of the DNA of weight average M_w, and that the ordinate at the origin, b, is:

$$b = \log_e \alpha \,. \tag{A8}$$

Equation (A8) allows us to derive α and therefore the absolute yield of the fragments from equation (A6).

It is possible to derive M_w from M_0; using equations (A4) and (A7) one obtains:

$$M_w = -\frac{2}{\beta 2}\left(\beta + \frac{1}{M_0}\right) \simeq 2\ M_n \tag{A9}$$

The M_w value calculated from this equation corresponds to the average size of the DNA molecules which are the enzyme target.

It is clear from equation (A9) that $1/M_0$, which is equal to 0·02, can be neglected compared to β whose absolute value is comprised here between 0·2 and 0·9 (see Results, section (d)). This prevents the use of equation (A9) to calculate M_0 from β and M_w.

(c) *Quantitative measurement of DNA amount in bands*

The treatment just described requires a precise knowledge of the amount of DNA present in the gel electrophoresis bands. A photographic procedure for this quantitation was therefore developed. This essentially requires the experimental set-up

described in Materials and Methods, section (d), of the main text, except that Kodak Ektapan (4 in $\times$ 5 in) films were used instead of Polaroid films. The blackening curve, i.e. the relationship between the intensity of the light hitting the film and the blackening of the film, was determined by illuminating a film through a step tablet having zones of different transmission. The resulting negative was scanned with a Joyce-Loebl microdensitometer and the blackening curve was obtained by plotting the pen deflections of the microdensitometer against the transmission of the zones of the step tablet. Densitometric tracings of DNA bands were analyzed by converting the pen deflections corresponding to the slices into which DNA peaks were cut into fluorescence intensities using the blackening curve. After baseline subtraction, the intensities corresponding to all slices of a peak were added together to give the total fluorescence of each band which was shown to be directly proportional to the amount of DNA. A detailed presentation of this method will be given elsewhere (Prunell *et al.*, unpublished data).

(d) Hae *and* Hpa *fragment yield from yeast mitochondrial DNAs*

Figure A2 shows plots of fragment yield *versus* the molecular weight of the fragments for the 15 to 20 bands of highest molecular weight. These results lead to a number of interesting conclusions.

(1) Loss of fragments having molecular weights higher than those corresponding to the top bands of Figures 2 and 7 of the main text can be ruled out since DNAs from strains A and B having higher molecular weights (9×10^6) did not show any additional bands of higher molecular weights, but simply higher yields of the top bands; data for DNA from strain A are shown in Fig. A2 (squares); a similar quantitative treatment was not done for the DNA from strain B.

(2) Multiple bands, as defined in Results, section (b) of the main text, fitted the straight line only after dividing the amount of DNA by the band multiplicity, confirming our assessment of the multiplicity.

(3) Some points showed a reproducible deviation from the straight line (Fig. 2). For instance, bands A1 and a3 (*Hpa*) appear to have yields approximately four times lower than expected. The yield became close to the expected value when the DNA preparation of higher molecular weight (9×10^6) was used (Fig. A2, squares), showing that a preferential breakdown of DNA is responsible for this phenomenon. Similar effects have been seen for bands A3 (*Hae*), and (not shown) for the faint band $A12_1$ (*Hae*).

(4) In the case of strains B and D, bands homologous to those just mentioned showed the same behavior, the deviation being stronger when the starting molecular weight was lower. This phenomenon provides an additional criterion of fragment homology between different strains. In contrast, similar deviations were not seen in the case of *S. carlsbergensis* DNA.

(5) The plots of Figure A2 permit us to calculate the target size of the DNAs for the restriction enzymes using equation (A9). The values so calculated (Table A1) are in general agreement with the molecular weight of the starting DNAs.

(6) The yields considered here were calculated assuming that the yield of the smallest bands is 100% (Fig. A2) and not on the basis of the amount of DNA loaded on the gel. As a consequence, the presence of contaminating DNA (like nuclear DNA and DNA from spontaneous "petite" and wild-type mutants (see Discussion, section

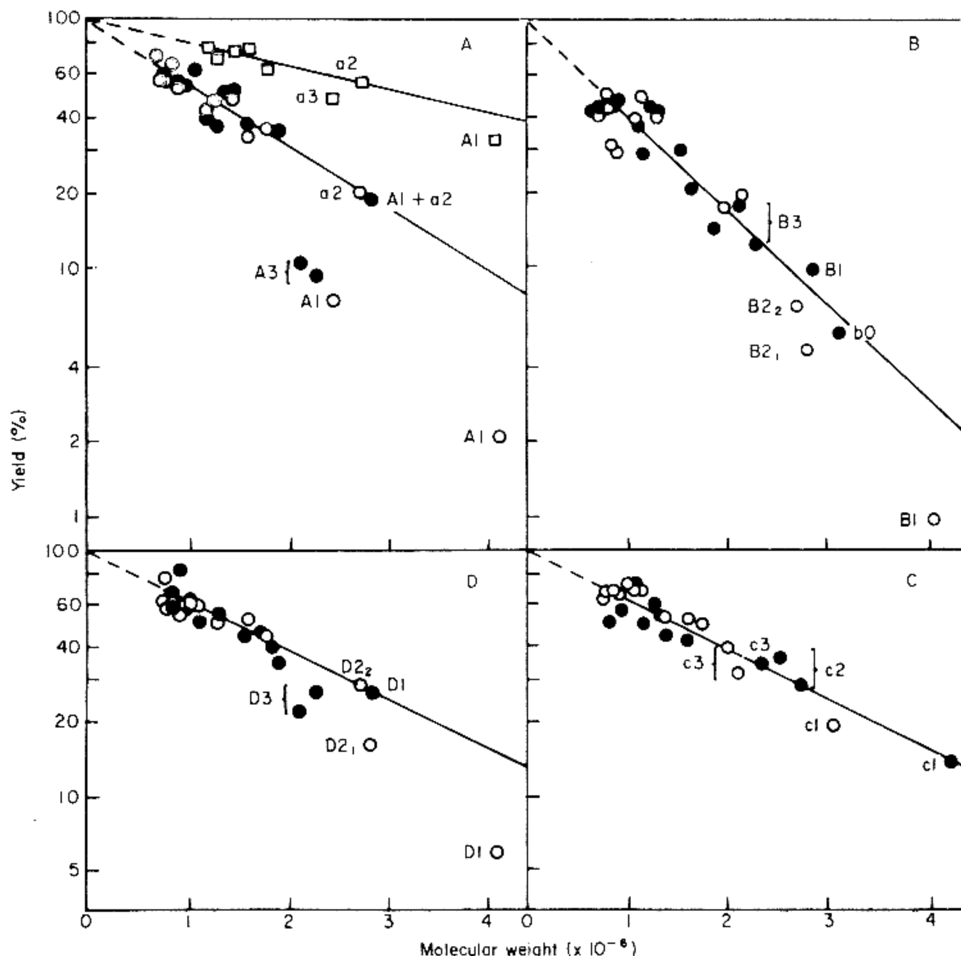

FIG. A2. Plot of the fragment yield, on a logarithmic scale, from the four mitochondrial DNAs degraded by *Hpa* (○) and *Hae* (●) restriction endonuclease, *versus* the molecular weight of the fragments. (□) Refer to the fragments obtained by *Hpa* degradation of mitochondrial DNA from strain A whose molecular weight was higher (9×10^6). The yield was calculated from equation (A6). In order to obtain the α values, Q_i/M_i was plotted *versus* M first, and α determined from equation (A8) using a least-squares procedure.

All Q_i/M_i values were then divided by α and replotted. Some points concern multiple bands or incompletely separated bands. In these cases the Q_i values were divided by the corresponding number of fragments. In the calculation of the α values, the points which deviate too strongly from the line were not taken into account, namely A1, a3 (*Hpa*) and A3 (*Hae*), B1 and B2$_1$ (*Hpa*), D1, D2$_1$ (*Hpa*) and D3 (*Hae*). The doublets in B3 (*Hae*), Fig. 8 of the main text, were considered as double bands.

(a) (ii), in the main text) contributing to the background smear does not affect the estimation of the fragment yield.

(e) *Endogenous degradation of mitochondrial DNA*

The fact that the target size is in general agreement with the molecular weight of the starting DNAs (Table A1) indicates that the overall degradation occurring during the preparation of the mitochondrial DNA can be considered as a random one.

Nevertheless, the observation of a specific endogenous breakage which is superimposed on the random breakage and which induces lower yields of some bands and the appearance of faint bands in the DNA from the *S. cerevisiae* strains, indicates the presence of a highly specific DNase in yeast. This activity, possibly localized in the mitochondria of *S. cerevisiae* strains, appears to be different from previously described ones, including the one recently reported (Zeman & Lusena, 1975) to attack preferentially the A+T-rich spacers. In *S. carlsbergensis* the specific activity is absent or much lower than in *S. cerevisiae*. This may be an explanation for the greater ease of preparation of high molecular weight DNA from *S. carlsbergensis*.

REFERENCES

Bernardi, G. & Sadron, C. (1964). *A. Baselli Conference on Nucleic Acids and Their Role in Biology*, pp. 62–80. Istituto Lombardo, Milan.

Charlesby, A. (1954). *Proc. Roy. Soc., ser. A*. **224**, 120–128.

Prunell, A. & Bernardi, G. (1973). *J. Biol. Chem.* **248**, 3433–3440.

Studier, F. W. (1965). *J. Mol. Biol.* **11**, 373–390.

Zeman, L. & Lusena, C. (1975). *Eur. J. Biochem.* **57**, 561–567.