



HAL
open science

Mutations in the mitochondrial split gene *COXI* are preferentially located in exons: a mapping study of 170 mutants

Pierre Netter, Sylviane Robineau, Claire Lemaire

► To cite this version:

Pierre Netter, Sylviane Robineau, Claire Lemaire. Mutations in the mitochondrial split gene *COXI* are preferentially located in exons: a mapping study of 170 mutants. *Molecular and General Genetics* MGG, 1995, 246 (4), pp.445-454. 10.1007/bf00290448 . cea-02474786

HAL Id: cea-02474786

<https://cea.hal.science/cea-02474786>

Submitted on 11 Feb 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ORIGINAL PAPER

Pierre Netter · Sylviane Robineau · Claire Lemaire

Mutations in the mitochondrial split gene *COXI* are preferentially located in exons: a mapping study of 170 mutants

Received: 18 July 1994 / Accepted: 14 September 1994

Abstract We have analysed the precise location of a large number (170) of mutations affecting the structural gene for subunit I of the cytochrome *c* oxidase complex. This gene, *COXI*, is 12.9 kb long and the major part of the sequence (i.e. 11.3 kb) is composed of introns. Several conclusions can be drawn from this study: (1) A significant proportion (84/170) of the mutations cannot be assigned to a single position within the gene by deletion mapping, in spite of clearly being located in it. These mutations are probably large deletions or multiple mutations. (2) Four mutants carry distant double mutations, which have been individually localized. (3) Eighty-two mutants have lesions that are restricted to very short regions of the gene and we therefore conclude that they are most probably due to single hits; amongst these single mutations, 41 are unambiguously located in exons and 28 in introns. This result implies that, at least in this particular split gene, the probability of selection of a mutant phenotype in an exon is, on the average, 13.3 times greater than in an intron, in spite of the existence, within most of these introns, of open reading frames specifying intronic proteins. The evolutionary significance and biological implications of these results are discussed.

Key words Yeast · Cytochrome oxidase
Mitochondria · Intronic proteins · Ribozymes

Introduction

This study focuses on the gene encoding subunit I of cytochrome *c* oxidase, which is the last complex of the mitochondrial respiratory chain. In the yeast *Saccharomyces cerevisiae*, it is composed of 12 subunits ranging

in size from 55 to 5 kDa (Taanman and Capaldi 1992). Subunits I, II and III (COXI_p, COXII_p and COXIII_p) are specified, respectively, in the mitochondrial DNA by *COXI*, *COXII* and *COXIII* genes (Slonimski and Tzagoloff 1976; Coruzzi and Tzagoloff 1979; Bonitz et al. 1980; Thalenfeld and Tzagoloff 1980). The counterparts of these three polypeptides can be found in bacteria and are apparently sufficient, in these prokaryotic organisms, to build a fully active complex (Haltia et al. 1988; Hosler et al. 1992). The other subunits present in yeast are encoded by nuclear genes (Maarse et al. 1984; Seraphin et al. 1985; Cumsky et al. 1987; Wright et al. 1984; Taanman and Capaldi 1993; Lamarche et al. 1992; Calder and McEwen 1990; Wright et al. 1986; Patterson and Poyton 1986) and are synthesized on cytoplasmic ribosomes before being imported and assembled in mitochondria. Subunit I of the cytochrome oxidase is highly conserved amongst all living organisms, from bacteria to animals and plants. This integral membrane protein exhibits 12 putative transmembrane helices (Rao and Argos 1986) and, together with subunits II and III, constitutes the catalytic core of the last respiratory mitochondrial complex.

The structural gene for subunit I, *COXI*, is located on the mitochondrial DNA and shows a mosaic organisation: it is 12877 nucleotides long but the eight exons specifying subunit I itself are only 1602 nucleotide long (corresponding to a protein of 534 amino acids) and are separated by seven introns (Bonitz et al. 1980; Hensgens et al. 1983). These introns are removed from a long RNA precursor, generating the mature mRNA for the cytochrome oxidase subunit (Carignani et al. 1986).

Such introns, which have been discovered and extensively studied in fungal mitochondria and chloroplasts (for review see Lambowitz and Belfort 1993), exhibit several very interesting properties. Their primary sequences allow them to be folded after transcription of the pre-mRNA, giving rise to two different three-dimensional structures specific to so-called group I and group II introns (Davies et al. 1982; Michel and Dujon 1983). These structured introns have ribozyme activities, lead-

Communicated by A. Kondorosi

P. Netter (✉) · S. Robineau · C. Lemaire
Centre de Génétique Moléculaire du CNRS,
Laboratoire associé à l'Université Pierre et Marie Curie,
Avenue de la Terrasse, 91198-Gif-sur-Yvette, France

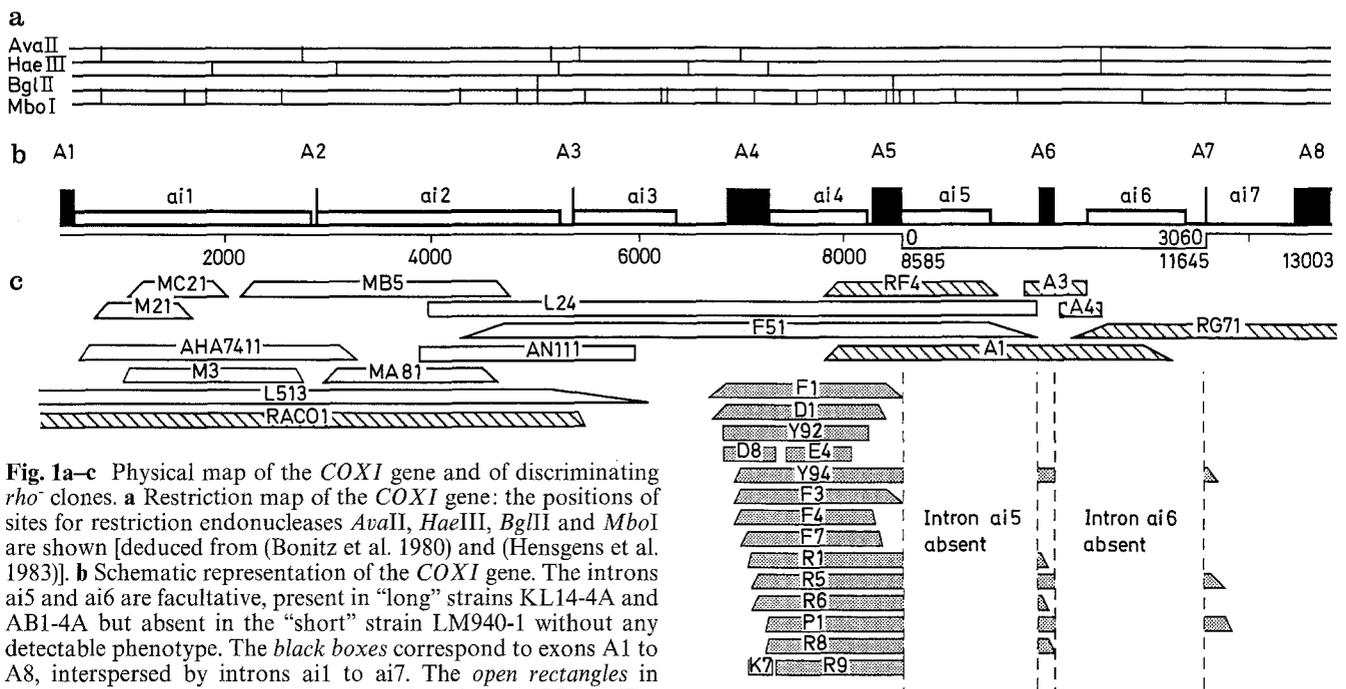


Fig. 1a-c Physical map of the *COXI* gene and of discriminating ρ^- clones. **a** Restriction map of the *COXI* gene: the positions of sites for restriction endonucleases *Ava*II, *Hae*III, *Bgl*II and *Mbo*I are shown [deduced from (Bonitz et al. 1980) and (Hensgens et al. 1983)]. **b** Schematic representation of the *COXI* gene. The introns ai5 and ai6 are facultative, present in "long" strains KL14-4A and AB1-4A but absent in the "short" strain LM940-1 without any detectable phenotype. The black boxes correspond to exons A1 to A8, interspersed by introns ai1 to ai7. The open rectangles in introns represent the intronic open reading frames (ORFs). The nucleotides are numbered from 1 to 13003, starting at position -126 of the short strain of Bonitz et al. (1980), interrupted (for exons A5 to A7) by 3060 nucleotides of the long strain from Hensgens et al. (1983). **c** Physical map of the discriminating ρ^- clones used in this study. The uncertainties due to the junction fragments are represented by slanted edges. The ρ^- clones (MC21, MB5, etc.) from strains KL14-4A and AB1-4A, in which the introns ai5 and ai6 are present, are represented in white and hatched areas respectively, whereas the ρ^- clones (F1, D1, etc.) derived from strain LM940-1, in which these introns are absent, are represented in grey. The precise junctions of some ρ^- clones have been determined by sequencing (see Table 1).

ing to in vitro self-splicing (Garriga and Lambowitz 1984; van der Horst and Tabak 1985; Peebles et al. 1986; Schmelzer and Schweyen 1986; van der Veen et al. 1986), which follows different molecular mechanisms in the two groups of introns (Michel et al. 1989a; Cech 1990). Furthermore, the sequence frequently allows the translation of intronic open reading frames (ORFs) most often in continuity with the upstream exons (for example, in the *COXI* gene, six of the seven introns exhibit such ORFs: see Fig. 1).

Mutations can be selected which abolish, in vivo, the splicing of the introns. The mutations fall into two categories: *cis*-acting mutations which modify the structure of the folded intron and, therefore, affect its ribozymic activity and *trans*-acting mutations which affect the proteins encoded by the introns.

These proteins may have various functions. (1) As just described, some are implicated, in vivo, in the RNA-splicing process itself. This so-called "maturase" activity was first described for the second and fourth introns of the mitochondrial gene *cob-box* (Lazowska et al. 1980; De La Salle et al. 1982; Banroques et al. 1986). In the *COXI* gene, this *trans* activity has been demonstrated for the first intron ai1 (Carignani et al. 1983) and,

indirectly, for the fourth one, ai4 (Dujardin et al. 1982). (2) Some of these proteins exhibit DNA-endonuclease properties and may therefore be implicated in the genomic mobility (homing) of introns. In the gene *COXI*, the third (Sargueil et al. 1991; Szczepanek et al. 1994), fourth (Schapira et al. 1993) and fifth (Moran et al. 1992) intronic ORFs have been shown to specify such endonucleases. (3) The protein encoded by the fourth intron of the *cob-box* gene stimulates recombination in various genetic systems (Kotylak et al. 1985; Goguel et al. 1989). (4) Two intronic proteins, encoded by the first and second introns, ai1 and ai2, of the *COXI* gene, show suggestive similarities with viral reverse transcriptases (Michel and Lang 1985) and the corresponding activity has recently been demonstrated biochemically (Kennell et al. 1993). It appears, therefore, that intronic proteins are able to interact with nucleic acids and sometimes exhibit a dual activity, both on RNA and DNA (maturase + endonuclease; maturase + reverse transcriptase). Their origin is still a matter of speculation.

The initial goal of this work was to characterize mutants that are unable to grow on a non-fermentable medium as a consequence of point or discrete mutations affecting subunit I of the cytochrome *c* oxidase (Netter et al. 1992; Lemarre et al. 1994). A large collection of mutants was available, which had been previously mapped in the *COXI* region (Kotylak and Slonimski 1977; Kruszewska and Slonimski 1984). For the development of the project one of our main interests was to map and characterize more precisely all these mutations in order to retain only those affecting exonic parts of the gene and therefore subunit I itself. We describe here the results of this detailed mapping analysis and the resulting hypotheses concerning the organization of this mitochondrial split gene.

Table 1 Position and sequence of the limits of seven *rho*⁻ clones used in this study. The coordinates are the same as in the Fig. 1. Note that the length of the *rho*⁻ clone R9 (nucleotides 7363 to 9964) is only 1237 bp because its parental strain, LM940-1 is devoid of the fifth intron (1365 nucleotides long)

<i>rho</i> ⁻ clone	Sequence of the terminal tandem repeat	Upstream limit	Downstream limit	Length of the <i>rho</i> ⁻ unit (bp)
Y92	AAGATATAGTCCAAAT	6803–6818	8286–8301	1483
D8	–	6827 or 6828	7386 or 7387	559
K7	TTA	7091–7093	7337–7339	246
R9	–	7363	9964	1237
E4	TTATATTA AAAA	7496–7506	8115–8125	619
A3	AAAAAAAAAAAA	9870–9879	10464–10473	594
A4	AATTAATATC	10179–10188	10578–10587	399

Materials and methods

Yeast strains

The *mit*⁻ strains have already been described in Kotylak and Slonimski (1977) and in Kruszevska and Slonimski (1984). They are all derived from 777-3A (*rho*⁺, *MATα*, *op1*, *ade1*) except for some mutants in the V series which were selected in AB1-4D (*rho*⁺, *MATα*, *op1*, *met3*, *ade1*). The *rho*⁻ clones were selected in the following strains:

AB1-4A/8 (*MATα*, *his4*) for MC21, M21, AHA7411, M3, L513, MB5, MA81, L24, F51 and AN111; KL14-4A (*MATα*, *his1*, *trp2*) for RAC01, RF4, A1, A3, A4 and RG71; LM940-1 (*MATα*, *his4*) for F1, D1, Y92, D8, E4, Y94, F3, F4, F7, R1, R5, R6, P1, R8, K7 and R9, this last-mentioned strain lacking the introns ai5 and ai6 (Slonimski and Tzagoloff 1976; Carignani et al. 1979; Netter et al. 1982a).

Media

The media used for culturing were previously described in Dujardin et al. (1980).

Crosses

Crosses between *rho*⁻ and *mit*⁻ strains were performed using auxotrophic selection of the diploids. Fresh cultures of haploid parental strains were mixed on W0 (minimal medium). After 3 days incubation at 28°C, the resulting diploids were replicated on N3 (glycerol complete medium) and checked for growth after incubation at 28°C and, eventually, at 20°C and 36°C. In the case of multiple mutations (V356, G1765, G110), disjunction of individual mutations and study of their phenotypes involved analysis of diploid clones issuing from mitotic segregation and recombination.

Preparation of mitochondrial DNA

Mitochondrial DNA was prepared as described in De La Salle et al. (1982).

Restriction endonuclease digestion and gel electrophoresis

The enzymatic digestions were made following the instructions of the enzyme manufacturers. The electrophoresis of DNA fragments was performed either on 1% agarose or on 6% polyacrylamide gels as described (Lewin et al. 1978).

Cloning and nucleotide sequencing

The bacterial strains and plasmids used are described in Sambrook et al. (1989). Single-stranded DNA was sequenced by the dideoxy method (Sanger et al. 1977).

Results

Physical map and sequence analysis of the *rho*⁻ strains

Some of *rho*⁻ strains used in this study have already been characterized previously (Carignani et al. 1979; Netter et al. 1982a). The others have been mapped by restriction analysis, using a large number of endonucleases. The results obtained with *Ava*II, *Hae*III, *Bgl*III and *Mbo*I are given as an example in Fig. 1a. The *rho*⁻ mutations result from the loss of the majority of the mitochondrial DNA followed by amplification of the remaining sequences, in head-to-head ("palindrome") or head-to-tail ("tandem") repeats. The limits of a *rho*⁻ DNA (Fig. 1c) are deduced from comparison of its restriction map with the *rho*⁺ wild-type map, especially by analysis of the restriction fragment corresponding to the junction of two *rho*⁻ units. Therefore, the precision of determination of these limits depends on the presence of restriction sites in the corresponding region and on its palindromic/tandem organization. For example, the right limit of L24 was determined precisely because this *rho*⁻ clone harbours the restriction site *Sac*I (coordinate 9948) but has lost *Nco*I (coordinate 9959). Its repeat unit therefore ends between these two sites, within the first 14 5' nucleotides of the sixth exon A6.

In the case of the *rho*⁻ clones Y92, D8, K7, R9, E4, A3 and A4, the precise limits of the repeating unit were determined by sequencing. All the corresponding mitochondrial DNAs were isolated by centrifugation on a CsCl gradient, digested by restriction endonucleases and subcloned in M13-derived vectors for sequence analysis (Table 1).

Y92

The mitochondrial DNA of Y92 was digested with *Bam*HI and *Hind*III (coordinates 6750 and 8198). The junction fragment was sequenced and the *rho*⁻ repetition unit (1483 nucleotides long) was shown to be limited by tandem repeats of 16 nucleotides. Therefore this *rho*⁻ clone contains the last 41 nucleotides of intron ai3, the entire exon A4 and almost all the intron ai4 except the last 48 nucleotides.

D8

The mitochondrial DNA of D8 was digested with *Nsi*I (coordinates 7138 and 7366). The sequence of the junction (...TGCTGGAGATTAT...) demonstrates a tandem organization without repeats at the ends of the repeat unit (559 nucleotides long). D8 harbours the last 16 nucleotides of intron ai3, exon A4 and the first 64 nucleotides of intron ai4.

K7

The mitochondrial DNA of K7 was digested with *Nsi*I (coordinate 7138). The *rho*⁻ repeat unit (246 nucleotides) is bordered by a short repetition of three nucleotides, TTA. Its sequence includes the right end of exon A4 and the first 15 nucleotides of intron ai4.

R9

The mitochondrial DNAs of R9 was digested with *Nsi*I (coordinate 7366), showing no terminal repeat for a unit of 1237 nucleotides. It includes the intron ai4 (except the first 39 nucleotides) and the whole of exon A4. Since it is derived from the short strain LM940-1, it lacks the intron ai5 and possesses only the first 14 nucleotides of the exon A6. This region is polymorphic between long and short strains (4/14 nucleotides being different). We believe that this region is too short and contains too many mismatches to allow any recombination events: when crossed with a long strain, this *rho*⁻ clone will be unable to transmit by recombination its short sequence of 14 nucleotides and will therefore behave genetically as if it lacks the exon A6.

E4

The mitochondrial DNA of E4 was digested with *Mbo*I (coordinates 7523, 7724, 8072, 8476 and 8511). Its repeat unit of 619 nucleotides is bordered by a repeat of 11 nucleotides (TTATATTA AAA). It is exclusively composed of sequences of intron ai4.

A3, A4

The mitochondrial DNAs of the *rho*⁻ clones A3 and A4 were digested respectively with *Sac*I (coordinates 9948) and *Alu*I (coordinates 10281). A3, which is 594 nucleotides long, is limited by two stretches of ten As, and includes the last 81 nucleotides of intron ai5, the complete exon A6 and the first 388 nucleotides of intron ai7. A4, which is 399 nucleotides long, is bordered by the tandem repeat AATTAATATC and composed exclusively of sequences derived from intron ai6.

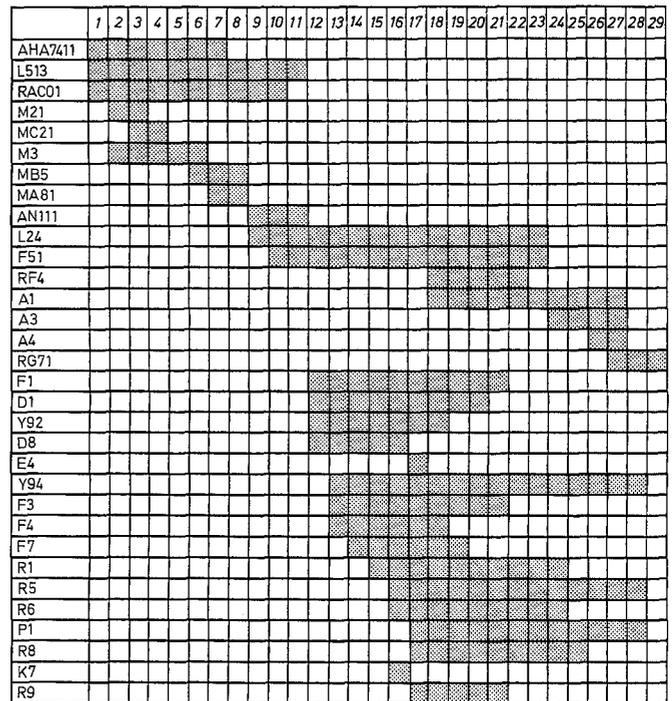


Fig. 2 Schematic representation of the *rho*⁻ clones used in this study. The numbered columns correspond to intervals, as deduced from comparison of the physical maps (Fig. 1) of the different *rho*⁻ clones listed at the left. The grey rectangles represent the intervals present in a given *rho*⁻ clone, each interval being identified by at least one discriminating clone. Note that there is no direct correlation between the physical size of a *rho*⁻ clone and the number of conserved intervals: compare, for example, D8 (five intervals, 559 nucleotides long) to E4 (one interval, 619 nucleotides)

Localization of the *mit*⁻ mutations

The comparison of the previously described *rho*⁻ clones (see Fig. 1) allowed us to subdivide the *COXI* gene into sub-regions or "intervals". An interval is defined as the shortest physical segment of a gene, which can be unambiguously shown to be present in some *rho*⁻ clones and absent in others. It is important to stress that the description of an interval is based upon the comparison of several *rho*⁻ clones and must be confirmed by genetic analysis. Figure 2 shows a schematic representation of the physical maps of the *rho*⁻ strains used here: their comparison leads to the distinction of 29 intervals covering the whole gene; most correspond to parts of individual exons or introns.

Each *mit*⁻ mutant was crossed with discriminating *rho*⁻ clones and the restoration of a fully respiratory-competent phenotype was checked for the resulting diploids. For 80 mutants, the phenotypic studies gave results compatible with membership of a single interval (see Table 2). These mutants were therefore considered as single mutations. The majority of the intervals correspond to purely exonic or intronic sequences. For intervals covering intron/exon junctions (i.e. nos. 6, 10, 12, 16, 19 and 28) the position of the mutations and their exon or intron membership was established by sequence

Table 2 The 29 intervals of the *COXI* gene. The mutations located in each interval are indicated. Many intervals correspond to sub-segments of introns or exons which are indicated in the second column. For intervals 6, 10, 12, 19 and 28, the position of the mutations was established by sequence analysis (Netter et al. 1992; Lemarre et al. 1994; G. Carignani, E. Bergantino, S. Robineau and P. Netter, unpublished work; C. Lemaire, S. Robineau and P. Netter, unpublished work) except for mutations denoted with a \$. The last category (interval 0) corresponds to the mutations which cannot be assigned unambiguously to a given interval (multiple mutations or large deletions)

Interval	Localization	Mutations
1	Intron 1	G2399, V293
2	Intron 1	G3010
3	Intron 1	V94
4	Intron 1	V284
5	Intron 1	V281, V221, V375
6		G2166, V277 (seq. = intron 2)
7	Intron 2	G1194
8	Intron 2	W165
9	Intron 2	G110-1
10		V202, V231 (seq. = exon 3), V264\$
11	Intron 3	G2027, G3013
12		G1099, G1765-1 (seq. = exon 4)
13	Exon 4	G205, G588, G2567, V112, V217, V312, V356-1
14	Exon 4	G1003, G2276, V32, V106, V203, V207, V380
15	Exon 4	G481, G1547, V130
16		G247, G2508 (seq. = exon 4), V356-2\$
17	Intron 4	G192, G476-1, G1765-2, G2190, V274
18	Intron 4	V226
19		G291, G450, G2443, G3015, W113 (seq. = exon 5)
20	Exon 5	V124, V243, V256
21	Exon 5	G110-2, V402, W164, W201
22	Intron 5	V36, V237, V241
23	Intron 5	V249, V370
24	Exon 6	G1979, V128 (seq. = exon 6)
25	Exon 6	V255, V282, W109, V379\$
26	Intron 6	G421
27	Intron 6	V38
28		G2084 (seq. = exon 7)
29	Intron 7 and exon 8	G341, G433, G1375, G2029, G2124, G2312, G2394, V13, V19, V95, V209, V220, V238, V290, V300, V393, V400
0	Deletions or multiple mutations	G20, G74, G390, G656, G689, G698, G789, G728, G922, G1055, G1160, G1181, G1186, G1264, G1273, G1287, G1305, G1734, G1859, G1921, G2133, G2220, G2342, G2343, G2354, G2469, G2496, V3, V16, V40, V41, V55, V73, V105, V129, V212, V213, V214, V215, V218, V222, V223, V228, V229, V230, V233, V240, V242, V244, V252, V258, V261, V262, V268, V271, V275, V279, V280, V289, V299, V307, V308, V310, V315, V334, V337, V338, V339, V351, V355, V358, V359, V374, V376, V377, V378, V381, V387, V383, V385, V391, V394, V395, V396

analysis (Netter et al. 1992; Lemarre et al. 1994; G. Carignani, E. Bergantino, S. Robineau and P. Netter, unpublished work; C. Lemaire, S. Robineau and P. Netter, unpublished work) except for the interval 29.

It should be noted that, in some instances, the quantitative results of the genetic analysis allowed some ambiguity in the assignment of a given mutation to a single interval. These ambiguities are due to recombination disturbance at the limits of a *rho*⁻ repetition unit, called a border effect. These mutations are indicated by a \$ symbol in Table 2 and Fig. 3 and will not be considered in the Discussion.

Mutants G110, G1765, V356 and G476, after random crosses and analysis of mitotic segregation, were shown to be due to double mutations named G110-1 and -2 for G110, G1765-1 and -2 for G1765 and so on, localized in different intervals. Individual mutations exhibited in most cases their own phenotypic properties. Nevertheless, a fully wild-type phenotype was only obtained when all mutations were corrected, for example in crosses with long *rho*⁻ clones covering the two affected intervals.

Finally, for 84 mutants unambiguously located in the *COXI* region (after crosses with the long *rho*⁻ clone KL14-4A/RP131 which covers the entire *COXI* gene (Kotylak and Slonimski 1977), we were unable to identify single mutations in a given interval. We believe that these mutations (grouped in interval 0, Table 2) are due to multiple distant and stringent mutations or to large deletions. For example, in *mit* × *mit*⁻ crosses, mutant G1859 was shown to be a deletion starting within the sixth intron and extending leftward at least to the fourth exon.

Discussion

Exonic or intronic mutations

Amongst 170 *mit*⁻ mutants mapped to the *COXI* region, 86 have been shown to result from single mutations located within one of the 29 intervals, as deduced from crosses with the panel of *rho*⁻ clones used here and in previous publications (Carignani et al. 1983; Netter et

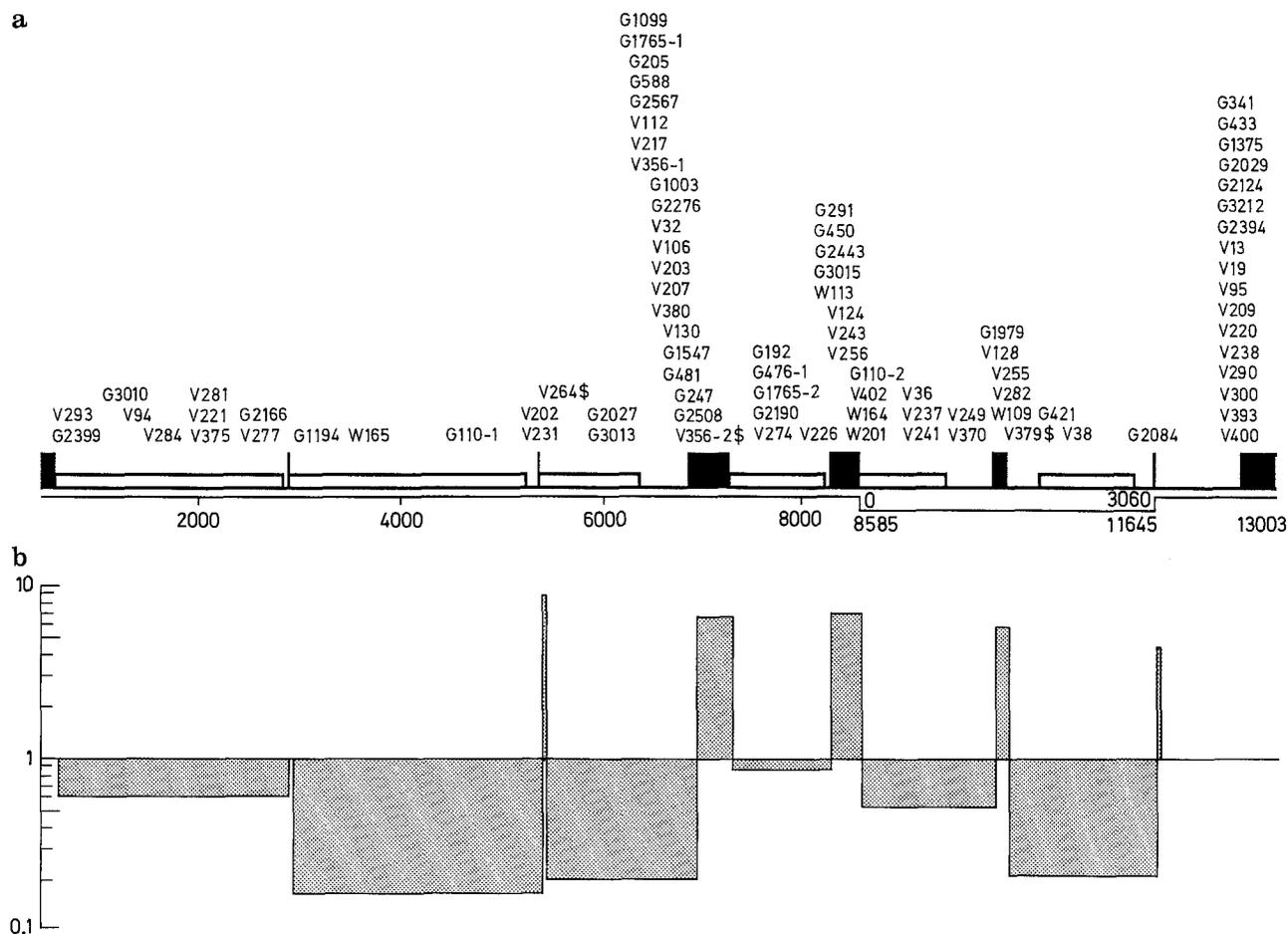


Fig. 3a,b Mapping of the *mit* mutations of the *COXI* gene. **a** Distribution of the 89 mutations within the 29 intervals deduced from the physical maps of the discriminating *rho*⁻ clones (see Fig. 1 and Table 2). **b** Deviation ($n_{\text{obs}}/n_{\text{calc}}$) of the probability of appearance of a mutation in exonic or intronic regions of the gene from expected value. The values are calculated in Table 3 (right column). Symbols used in **a** are as given in the legend to Fig. 1a

al. 1982a; see Table 2). The remaining 84 mutants, which cannot be assigned a single physical location in spite of belonging to the *COXI* gene, are probably due to multiple mutations or small deletions. These mutants will not be considered further in the Discussion, but we must keep in mind that by eliminating them, we obviously introduce a bias in favour of the single mutations, i.e. nucleotide substitutions, additions or deletions.

The *COXI* gene is 12.9 kb long; therefore the number of single mutations corresponds to a mean frequency of one mutation every 150 nucleotides. It is apparent (see Fig. 3) that the positions of the mutations, which have been clearly assigned to exons or introns (except for the interval 29), are not randomly distributed: 41 mutations have been localized in the 1140 nucleotides of exons A1 to A7 (i.e. on average, one mutation every 28 nucleotides in exons) whereas only 28 mutations have been identified in the 10388 nucleotides of introns ai1 to ai6 (one mutation every 371 nucleotides in introns). The probability of the appearance of a mutation is always

higher in exonic and always lower in intronic regions. One can express this bias by the ratio of the observed number of mutations to the expected number (assuming a constant probability along the gene of one mutation every 150 nucleotides; see above). The values of the bias for the different regions of the gene, given in the Table 3, are always less than 1 for introns and greater than 1 for exons. This bias could be partially explained by the difference in GC content in the exons (30.7%) and the introns (22%) of *COXI*, if one remembers that the published *mit* mutations due to nucleotide substitutions in the mitochondrial DNA of *S. cerevisiae* frequently correspond to GC → AT transitions or GC → TA transversions. However, this cannot account for the major portion of the bias, which establishes that the selection of a *mit* mutation is 13.3 (371/28) times more probable in an exon than in an intron.

It must be noted that the initial screening of the *mit* mutants is only based upon the inability of the yeast cell to use glycerol as carbon source. This selection procedure introduces a bias in favour of the mutations leading to a stringent respiratory-deficient phenotype. It tends to eliminate the leaky or partial mutants and, of course, those exhibiting phenotypes unrelated to respiration. In order to explain the difference between mutation frequencies in exons and introns, we must therefore analyse the targets which, in both cases, are able to confer a respiratory-negative phenotype when mutated.

Table 3 The bias in appearance of *mit* mutations in exons vs introns of the *COXI* gene. The length of each exon or intron of the gene is indicated in the second column. The physical map does not permit us to discriminate unambiguously between the last intron ai7 and exon A8, which are therefore considered together (ai7+A8 = interval 29). The third column lists the number of mutations observed in each exon or intron, derived from the results of the Table 2. The fourth column gives the expected number of mutations if the probability of appearance of a *mit* mutation is constant along the gene (one mutation every 150 nucleotides). The numbers are obtained by dividing the length of the segment (exonic or intronic) by 150. The fifth column gives the bias of appearance of *mit* mutations in exons and introns, as a ratio of the observed to calculated numbers of mutations (see the Discussion)

Exon/ Intron	Length (in bp)	n_{obs} (number of observed mutations)	n_{calc} (number of calculated mutations)	Bias ($n_{\text{obs}}/n_{\text{calc}}$)
A1	169	0	1.1	–
ai1	2448	10	16.3	0.61
A2	36	0	0.24	–
ai2	2515	3	16.8	0.18
A3	35	2	0.23	8.7
ai3	1514	2	10.1	0.20
A4	480	21	3.2	6.6
ai4	1010	6	6.7	0.89
A5	252	12	1.7	7.1
ai5	1365	5	9.0	0.55
A6	135	5	0.92	5.6
ai6	1536	2	10.2	0.20
A7 (ai7 + A8)	33 (879 + 470)	1 (17)	0.22 (9.3)	4.5 –

Exonic targets

Exonic sequences specify almost exclusively the peptide encoded by the gene (in our case COXI_p, the first subunit of the cytochrome oxidase complex). Short exonic sequences (complementary to guide sequences in group I introns, intron binding sequences IBS1 and IBS2 in group II introns) are implicated in the splicing of flanking introns, and mutations affecting these sequences have been described (Perea and Jacq 1985); however such mutations affect only a small number of the exonic nucleotides. Therefore, in exons, the targets of mutations are mainly the codons specifying the amino acids of COXI_p. The appearance of a mutant phenotype depends upon the importance of the residue in a given position and of the severity of the amino acid change, which must eliminate the activity of the respiratory complex IV.

There is an excess of mutations in exons (bias of 8.7, 6.6, 7.1, 5.6 and 4.5 for exons A3, A4, A5, A6 and A7 respectively; see Table 3) but their number seems to be roughly proportional to the exon length, except for the exon A1 described below. These results suggest that the corresponding segments of the protein (i.e. subunit I of the cytochrome oxidase) contribute equally to the activity of the complex and can be similarly affected by mutations abolishing its functional properties. If this rule extends to the C-terminal part of the protein, we could

predict that, amongst the 17 mutations roughly located in the (ai7 + A8) region, a majority should be located in the exon A8. In fact, taking into account the length of this exon (470 nucleotides) and the mean frequency of mutations in the exonic part of the gene (1 every 28 nucleotides), one expects $470/28 = 16.7$ mutations. Using the same calculation for intron ai7, only 2.3 mutations ($879 \text{ nucleotides}/371$) are predicted.

The first exon A1 constitutes an exception: no mutation was observed whereas six ($169/28 = 6.0$) were expected within its 169 nucleotides, if one applies the values for exonic sequences. This result could indicate that the N-terminal part of the protein, which includes the first and second transmembrane helices (Rao and Argos 1986), plays a less important role in the activity of the complex and can harbour many silent nucleotide changes. Furthermore, if one excludes a priori, possible reinitiation by the translation apparatus, this result indicates that mutations that generate nonsense codons (directly or by a frame-shift) are rare, at least in the first exon.

Intronic targets

Intronic sequences carry a complex mixture of structural and functional information. Firstly, from 200 to 600 intronic nucleotides are devoted to the three-dimensional constitution of the catalytic core of the ribozyme (Davies et al. 1982; Michel and Dujon 1983). Secondly, in addition to this structural information, many introns harbour long ORFs, some of which encode reverse transcriptases, DNA-endonucleases and/or RNA-maturases (see the Introduction). The conservation of these ORFs is a good argument in favour of the reality of their supposed functions and suggests that the ORFs are not simply pseudogenes. One would then expect a high density of structural and/or functional mutations in such pleiotropic regions; this is not observed. Several factors can be invoked to account for this apparent discrepancy.

(1) Structural (*cis*-acting) mutants of group I yeast mitochondrial introns, selected *in vivo* on the basis of their respiratory-deficient phenotype, have been shown to affect only a very limited number of residues, mainly those constituting what is now known to be the guanosine-binding site of the ribozyme (De La Salle et al. 1982; Weiss-Brummer et al. 1982; Netter et al. 1982b; Michel et al. 1989b). It seems, therefore, that the effective size of the ribozyme target, deduced from *in vivo* selection of *cis*-acting mutants in yeast mitochondria, is much smaller than predicted by the structural models and by *in vitro* experiments. In the *Tetrahymena* group I intron, systematic mutagenesis of the ribozyme residues led to the conclusion that the majority of single and double substitutions significantly affected the rate of *in vitro* self-splicing of the intron (Couture et al. 1990). This apparent discrepancy between *in vivo* and *in vitro* results is not a general rule: in experiments per-

formed on the intron *td* of bacteriophage T4 thymidylate synthase, a majority of the mutations introduced in conserved positions of this intron led to both splicing defects *in vitro* and to a detectable [thy] phenotype *in vivo* (Belfort et al. 1987). This variable susceptibility to nucleotide changes between these two biological systems must reflect different requirements for the synthesis or the expression of the final product, i.e. subunits of the mitochondrial respiratory complexes in yeast and T4 thymidylate synthase in *E. coli*.

(2) The functions encoded by the introns must also be clearly distinguished. It has been demonstrated that DNA endonucleases are encoded by the third, fourth and fifth introns (Sargueil et al. 1991; Schapira et al. 1993; Moran et al. 1992), and reverse transcriptase by the first one (Kennell et al. 1993). However it has not been shown whether the loss of reverse transcriptase or DNA endonuclease activities affects RNA splicing and gives rise to a respiratory-deficient phenotype. It is most probable that the *mit* phenotype is due solely to the inactivation of RNA maturases.

In the *COXI* gene, *trans*-complementable *mit* mutations affecting the splicing process and therefore, most probably, the RNA maturase itself, have been demonstrated only in the first intron *ai1* (mutants G3010, V94, V284, V281, V221 and V375; see Carignani et al. 1983). These *ai1* maturase mutants are unable to splice both the *ai1* and *ai2* intronic sequences from the long mRNA precursor, suggesting that the *ai1* maturase is implicated in the splicing of the first two introns of the gene. In addition, variant strains devoid of the genomic sequence of the first intron still correctly splice *ai2* intron from the pre-mRNA (Carignani et al. 1986). This result provides indirect evidence that the product of the ORF of the second intron (the sequence of which is very similar to the ORF of the first intron), also has RNA maturase properties.

Finally, in intron *a4*, indirect evidence of a transactive product implicated in the RNA splicing has been obtained for mutant G1765-2 (Dujardin et al. 1983) and for suppressors of *cob-box* mutations. In fact, the protein encoded by the fourth intron of *COXI* is 65% similar to the protein encoded by the fourth intron of the *cob-box* gene, which possesses RNA maturase activity (Banroques et al. 1986). When this latter activity is lost by mutation, several single mutations affecting the protein encoded by intron *ai4* can lead to the activation of latent maturase activity, in spite of the different primary structure of the two peptides (Dujardin et al. 1982; P. Netter and S. Robineau, unpublished work).

Thus there is evidence that the ORFs of the first, second and fourth introns encode proteins related to RNA splicing. To explain why clear-cut transactive *mit* mutants were found only in the first intron, we must assume that amino acid changes in the second and fourth intronic maturase-like proteins lead to a very leaky or discrete phenotype undetectable by our screening procedure. It is conceivable that the ORFs in introns *ai3*, *ai5* and *ai6* could also encode discrete RNA mat-

urase activities. Some sequence similarities (such as the LAGLI-DADG motif) between these proteins and bona fide maturases could argue in favour of this hypothesis. It is then possible to imagine that a minimal selective advantage, associated with the correct splicing of the introns, could be sufficient to account for the maintenance of open reading frames, in spite of this not being detectable by laboratory experiments. The possibility also cannot be excluded that the same kind of undetectable selective advantage is associated with other activities of intronic ORFs, such as reverse transcriptase, recombinase or endonuclease.

Intronic proteins

The region of the mitochondrial genome studied here may be represented as a discontinuous exonic gene (composed of eight exons and encoding COXI_p subunit) interspersed with "intronic genes". We must then try to explain why the probability of appearance of mutations in these intronic genes is much lower than in the exonic gene. The respiratory complexes are composed of essential and highly specialized proteins, which are the result of a very long and stringent process of selection. Thus it is difficult to imagine that they could acquire new functions as a result of simple mutations. The very high level of sequence conservation in these proteins in the whole of the living world, from bacteria to man, illustrates this high degree of specialization. Conversely, many single amino acid changes in these proteins have been shown to lead to respiratory-deficient phenotypes (Netter et al. 1992; Meunier et al. 1993; Lemarre et al. 1994; C. Lemaire, S. Robineau and P. Netter, unpublished work; for a review on cytochrome *b*, see Degli Esposti et al. 1993). The critical need to maintain a fully active respiratory chain is illustrated by the natural overproduction of the mRNA for the corresponding proteins. We have frequently observed, in studying respiratory-competent pseudo-wild revertants of *mit* mutations, that a very low level of mRNA is sufficient to maintain a normal synthesis of these proteins (data not shown). In the wild-type strain, the overproduction of mRNA constitutes a natural protection against mutations and could explain the small number of RNA maturase mutants: a splicing defect leading to a respiratory-deficient phenotype would be obtained only in the case of (multiple) mutations completely abolishing the maturase activities.

On the contrary, it seems that the intronic proteins are less specialized and more primitive. They frequently have conserved dual functions: the protein encoded by the first and second introns of *COXI* accumulates RNA maturase and reverse transcriptase activities (Carignani et al. 1983; Kennell et al. 1993); the protein encoded by the fourth intron of the *cob-box* gene shows DNA recombinase and RNA maturase properties (De La Salle et al. 1982; Kotylak et al. 1985; Banroques et al. 1986; Goguel et al. 1989). It is conceivable that this phe-

nomenon constitutes a sign of the invasion of ribozymic introns by intronic protein genes, either by reverse transcription or by DNA homing, associated with selection pressure in favour of the acquisition of RNA splicing ability (Lambowitz and Belfort 1993). As a consequence of the conservation of these multiple functions, the intronic proteins may have been unable to reach a high level of specialization in each distinct activity. This hypothesis is favoured by the finding that mutants defective in maturase activities are rarely due to discrete amino acid changes but, in the majority of cases, to nonsense (direct or frameshift) mutations leading to the loss of a large part of the molecule (Lazowska et al. 1980; Jacq et al. 1982; De La Salle et al. 1982). We propose, therefore, that splicing-related activities of intronic proteins are largely unaffected by single amino acid replacements and that most nucleotide changes in intronic genes are functionally silent, whereas exonic proteins like COXI_p are highly susceptible to such replacements. This differential sensitivity to amino acid changes would explain the very unequal distribution of the mutations observed in COXI gene.

Acknowledgements This work was supported by EC contract no. SCI 000476. We thank F. Michel and C. J. Herbert for suggestions and critical reading of the manuscript.

References

- Banroques J, Delahodde A, Jacq C (1986) A mitochondrial RNA maturase gene transferred to the yeast nucleus can control mitochondrial mRNA splicing. *Cell* 46:837–844
- Belfort M, Chandry PS, Pedersen-Lane J (1987) Genetic delineation of functional components of the group I intron in the phage T4 *td* gene. *Cold Spring Harbor Symp Quant Biol* 52:181–192
- Bonitz SG, Coruzzi G, Thalenfeld BE, Tzagoloff A, Macino G (1980) Assembly of the mitochondrial membrane system: physical map of the *oxi3* locus of yeast mitochondrial DNA. *J Biol Chem* 255:11927–11941
- Calder KM, McEwen JE (1990) Nucleotide sequence of the gene encoding cytochrome *c* oxidase subunit VII from *Saccharomyces cerevisiae*. *Nucleic Acids Res* 18:1632
- Carignani G, Dujardin G, Slonimski P (1979) Petite deletion map of the mitochondrial *oxi3* region in *S. cerevisiae*. *Mol Gen Genet* 167:301–308
- Carignani G, Groudinsky O, Frezza D, Schiavon E, Bergantino E, Slonimski P (1983) An mRNA maturase is encoded by the first intron of the mitochondrial gene for the subunit I of cytochrome oxidase in *S. cerevisiae*. *Cell* 35:733–742
- Carignani G, Netter P, Bergantino E, Robineau S (1986) Expression of the mitochondrial split gene coding for cytochrome oxidase subunit I in *S. cerevisiae*: RNA splicing pathway. *Curr Genet* 11:55–63
- Cech TR (1990) Self-splicing of group I introns. *Annu Rev Biochem* 59:543–568
- Coruzzi G, Tzagoloff A (1979) Assembly of the mitochondrial membrane system: DNA sequence of subunit 2 of yeast cytochrome oxidase. *J Biol Chem* 254:9324–9330
- Couture S, Ellington A, Gerber A, Cherry M, Doudna J, Green R, Hanna M, Pace U, Rajapogal J, Szostak J (1990) Mutational analysis of conserved nucleotides in a self-splicing group I intron. *J Mol Biol* 215:345–358
- Cumsky MG, Trueblood CE, Ko C, Poyton RO (1987) Structural analysis of two genes encoding divergent forms of yeast cytochrome *c* oxidase subunit V. *Mol Cell Biol* 7:3511–3519
- Davies RW, Waring RB, Ray JA, Brown TA, Scazzocchio C (1982) Making ends meet: a model for RNA splicing in fungal mitochondria. *Nature* 300:719–724
- De La Salle H, Jacq C, Slonimski PP (1982) Critical sequences within mitochondrial introns: pleiotropic mRNA maturase and *cis*-dominant signals of the *box* intron controlling reductase and oxidase. *Cell* 28:721–732
- Degli Esposti M, De Vries S, Crimi M, Ghelli A, Patarnello T, Meyer A. (1993) Mitochondrial cytochrome *b*: evolution and structure of the protein. *Biochim Biophys Acta* 1143:243–271
- Dujardin G, Jacq C, Slonimski PP (1982) Single base substitution in an intron of oxidase gene compensates splicing defects of the cytochrome *b* gene. *Nature* 298:628–632
- Dujardin G, Pajot P, Groudinsky O, Slonimski PP (1980) Long range control circuits within mitochondria and between nucleus and mitochondria. I. Methodology and phenomenology of suppressors. *Mol Gen Genet* 179:469–482
- Garriga G, Lambowitz AM (1984) RNA splicing in *Neurospora* mitochondria: self-splicing of a mitochondrial intron *in vitro*. *Cell* 39:631–641
- Goguel V, Bailone A, Devoret R, Jacq C (1989) The *bI4* RNA mitochondrial maturase of *S. cerevisiae* can stimulate intrachromosomal recombination in *E. coli*. *Mol Gen Genet* 216:70–74
- Haltia T, Puustinen A, Finel M (1988) The *Paracoccus denitrificans* cytochrome *aa3* has a third subunit. *Eur J Biochem* 172:543–546
- Hensgens LAM, Bonen L, de Haan M, Horst G, Grivell L (1983) Two intron sequences in yeast mitochondrial COX I gene: homology among URF containing introns and strain dependant variation in flanking exons. *Cell* 32:379–389
- Hosler J, Fetter J, Tecklenburg M, Espe M, Lerma C, Ferguson-Miller S (1992) Cytochrome *aa3* of *Rhodobacter sphaeroides* as a model for mitochondrial cytochrome *c* oxidase. *J Biol Chem* 267:24264–24272
- Jacq C, Pajot P, Lazowska J, Dujardin G, Claisse M, Groudinsky O, De La Salle H, Grandchamps C, Labouesse M, Gargouri A, Guiard B, Spiridakis A, Dreyfus M, Slonimski P (1982) Role of introns in the yeast cytochrome *b* gene: *cis*- and *trans*-acting signal, intron manipulation, expression and intergenic communications. In: Slonimski P, Borst P, Attardi G (eds) *Mitochondrial Genes*. (Cold Spring Harbour Symposium on Mitochondrial Genes) Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, pp 155–183
- Kennell JC, Moran JV, Perlman PS, Butow RA, Lambowitz AM (1993) Reverse transcriptase activity associated with maturase-encoding group II introns in yeast mitochondria. *Cell* 73:133–146
- Kotylak Z, Slonimski PP (1977) Mitochondrial mutants isolated by a new method based upon the use of the nuclear mutation *opl1*. In: Bandlow W, Schweyen RJ, Wolf K, Kandewitz F (eds) *Mitochondria 1977*. de Gruyter, Berlin. pp 83–89
- Kotylak Z, Lazowska J, Hawthorne DC, Slonimski PP (1985) Intron encoded proteins in mitochondria: key elements of gene expression and genomic evolution. In: Quagliariello E, Slater EC, Palmieri F, Sacconne C, Kroon AM (eds) *Achievements and perspectives in mitochondrial research*, vol 2. Elsevier Amsterdam, pp 1–20
- Kruszewska A, Slonimski P (1984) Mitochondrial and nuclear mitoribosomal suppressors that enable misreading of ochre codons in yeast mitochondria. *Curr Genet* 9:11–19
- Lamarche AEP, Abate MI, Chan SHP, Trumpower BL (1992) Isolation and characterization of COX12, the nuclear gene for a previously unrecognized subunit of *Saccharomyces cerevisiae* cytochrome-*c* oxidase. *J Biol Chem* 267:22473–22480
- Lambowitz AM, Belfort M (1993) Introns as mobile elements. *Annu Rev Biochem* 62:587–622
- Lazowska J, Jacq C, Slonimski PP (1980) Sequence of introns and flanking exons in wild-type and *box3* mutants of cytochrome *b*

- reveals an interlaced splicing protein coded by an intron. *Cell* 22:333–348
- Lemarre P, Robineau S, Colson AM, Netter P (1994) Sequence analysis of three deficient mutants of cytochrome oxidase subunit I of *S. cerevisiae* and their revertants. *Curr Genet* 26(5–6):546–552
- Lewin A, Morimoto R, Rabinowitz M, Fukuhara H (1978) Restriction enzyme analysis of mitochondrial DNAs of petite mutants of yeast. *Mol Gen Genet* 163:257–275
- Maarse AC, Van LA, Riezman H, Gregor I, Schatz G, Grivell LA (1984) Subunit IV of yeast cytochrome *c* oxidase: cloning and nucleotide sequencing of the gene and partial amino acid sequencing of the mature protein. *EMBO J* 3:2831–2837
- Meunier B, Coster F, Lemarre P, Colson AM (1993) Insight into topological and functional relationships of cytochrome-*c* oxidase subunit I of *S. cerevisiae* by means of intragenic complementation. *FEBS Lett* 321:159–162
- Michel F, Dujon B (1983) Conservation of RNA secondary structures in two intron families including mitochondrial-, chloroplast- and nuclear encoded members. *EMBO J* 2:33–38
- Michel F, Lang BF (1985) Mitochondrial class II introns encode proteins related to the reverse transcriptases of retroviruses. *Nature* 316:641–643
- Michel F, Umesono K, Okesi H (1989a) Comparative and functional anatomy of group II catalytic introns – a review. *Gene* 82:5–30
- Michel F, Hanna M, Green R, Bartel D, Szostak J (1989b) The guanosine binding site of *Tetrahymena* ribozyme. *Nature* 339:391–395
- Moran JV, Wernette CM, Mecklenburg KL, Butow RA, Perlman PS (1992) Intron 5 alpha of the *COXI* gene of yeast mitochondrial DNA is a mobile group I intron. *Nucleic Acids Res* 20:4069–4076
- Netter P, Carignani G, Jacq C, Groudinsky O, Clavillier L, Slonimski PP (1982a) The cytochrome oxidase subunit I split gene in *Saccharomyces cerevisiae*: genetic and physical studies of the mtDNA segment encompassing the 'cytochrome *b*-homologous' intron. *Mol Gen Genet* 188:551–559
- Netter P, Jacq C, Carignani G, Slonimski PP (1982b) Critical sequences within mitochondrial introns: *cis*-dominant mutations of the "cytochrome-*b*-like" intron of the oxidase gene. *Cell* 28:733–738
- Netter P, Robineau S, Sirand-Pugnet P, Fauvarque MO (1992) The unusual reversion properties of a mitochondrial mutation in the structural gene of subunit I of cytochrome oxidase of *Saccharomyces cerevisiae* reveal a probable histidine ligand of the redox center. *Curr Genet* 21:147–151
- Patterson TE, Poyton RO (1986) *COX8*, the structural gene for yeast cytochrome *c* oxidase subunit VIII. DNA sequence and gene disruption indicate that subunit VIII is required for maximal levels of cellular respiration and is derived from a precursor which is extended at both its NH₂ and COOH termini. *J Biol Chem* 261:17192–17197
- Peebles CL, Perlman PS, Mecklenburg KL, Petrillo ML, Tabor JH, Jarrell KA, Cheng HL (1986) A self-splicing RNA excises an intron lariat. *Cell* 44:213–223
- Perea J, Jacq C (1985) Role of the 5' hairpin structure in the splicing accuracy of the fourth intron of the yeast *cob*-box gene. *EMBO J* 4:3281–3288
- Rao JKM, Argos P (1986) A conformational preference parameter to predict helices in integral membrane proteins. *Biochim Biophys Acta* 869:197–214
- Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York
- Sanger F, Nicklen S, Coulson A (1977) DNA sequencing with chain terminating inhibitors. *Proc Natl Acad Sci USA* 74:5463–5467
- Sargueil B, Delahodde A, Hatat D, Tian GL, Lazowska J, Jacq C (1991) A new specific DNA endonuclease activity in yeast mitochondria. *Mol Gen Genet* 225:340–341
- Schapiro M, Desdouets C, Jacq C, Perea J (1993) *I-SceIII* an intron-encoded DNA endonuclease from yeast mitochondria. Asymmetrical DNA binding properties and cleavage reaction. *Nucleic Acids Res* 21:3683–3689
- Schmelzer C, Schweyen RJ (1986) Self-splicing of group II introns in vitro: mapping of the branch point and mutational inhibition of lariat formation. *Cell* 46:557–565
- Seraphin B, Simon M, Faye G (1985) Primary structure of a gene for subunit 5 of the cytochrome oxidase from *S. cerevisiae*. *Curr Genet* 9:435–439
- Slonimski P, Tzagoloff A (1976) Localization in yeast mitochondrial DNA of mutations expressed in a deficiency of cytochrome oxidase and/or coenzyme QH₂-cytochrome *c* reductase. *Eur J Biochem* 61:27–41
- Szczepanek T, Macadre C, Meunier B, Lazowska J (1994) Two homologous introns from related *Saccharomyces* species differ in their mobility. *Gene* 139:1–7
- Taanman JW, Capaldi RA (1992) Purification of yeast cytochrome *c* oxidase with a subunit composition resembling the mammalian enzyme. *J Biol Chem* 267:22481–5
- Taanman JW, Capaldi RA (1993) Subunit VIa of yeast cytochrome *c* oxidase is not necessary for assembly of the enzyme complex but modulates the enzyme activity. Isolation and characterization of the nuclear-coded gene. *J Biol Chem* 268:18754–18761
- Thalenfeld BE, Tzagoloff A (1980) Assembly of the mitochondrial membrane system. Sequence of the *oxi 2* gene of yeast mitochondrial DNA. *J Biol Chem* 255:6173–6180
- van der Horst G, Tabak HF (1985) Self-splicing of yeast mitochondrial ribosomal and messenger RNA precursors. *Cell* 40:759–766
- van der Veen R, Arnberg AC, van der Horst G, Bonen L, Tabak HF, Grivell LA (1986) Excised group II introns in yeast mitochondria are lariats and can be formed by self-splicing in vitro. *Cell* 44:225–234
- Weiss-Brummer B, Rodel G, Schweyen R, Kaudewitz F (1982) Expression of the split gene *cob* in yeast: evidence for a precursor of a maturase protein translated from intron 4 and preceding exons. *Cell* 29:527–536
- Wright RM, Dircks LK, Poyton RO (1986) Characterization of *COX9*, the nuclear gene encoding the yeast mitochondrial protein cytochrome *c* oxidase subunit VIIa. Subunit VIIa lacks a leader peptide and is an essential component of the holoenzyme. *J Biol Chem* 261:17183–17191
- Wright RM, KO C, Cumsy MG, Poyton RO (1984) Isolation and sequence of the structural gene for cytochrome *c* oxidase subunit 6 from *S. cerevisiae*. *J Biol Chem* 259:15401–15407