YEAST MITOCHONDRIAL INTRON PRODUCTS REQUIRED INTRONS FOR RNA SPLICING

George M. Church and Walter Gilbert

Biological Laboratories, Harvard University
Cambridge, Massachusetts

With the goal of investigating the functions of introns we have been studying the cytochrome b (or *box*) and *oxi-3* genes of yeast mitochondria. They are encoded in the 68 to 76 kb mitochondrial DNA (1) along with only five other known proteins and all of the ribosomal and transfer RNA's required for mitochondrial protein synthesis (see Figure 1). Although several mitochondrial genes (oli-1, oli-2, oxi-2, and nine tRNA's) have been shown by sequencing to be continuous throughout their coding regions (2-8) three genes (*box*, lg rRNA, *oxi-3*) are thought to be split. The cytochrome b and lg rRNA introns have been located by R-loop electron microscopy, blot hybridization and genetic mapping (8-13). The arguments for *oxi-3* being split are more indirect (11,14).

For the investigation of split genes this system has a number of impressive advantages over most others at present:

1) Most mitochondrially encoded RNA's and proteins have been mapped and alterations in these macromolecules are particularly easy to detect.

2) Petite deletions provide a stable way to inhibit mitochondrial protein synthesis.

3) Mutations are easily isolated and characterized by complementation and fine structure petite deletion mapping.

4) Of about 200 mutations in cytochrome b characterized to date about half affect expression of another mitochondrial gene *oxi-3*. All of these pleiotropic cytochrome b mutations affect trans-acting elements (defined by complementation properties) and many map to within the introns of the cytochrome b gene.

ISBN 0-12-633360-2

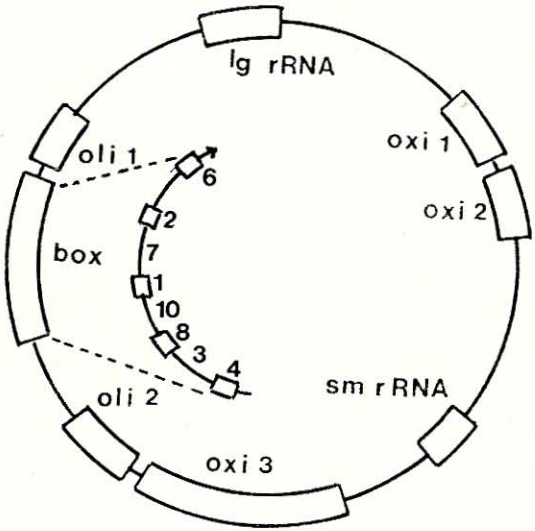

Figure 1. Genetic and physical map of the known mitochondrial genes and the cytochrome b (box) loci associated with five exons (rectangles) and three of the four introns (inner circle) .

figure 1 shows a detailed schematic of the cytochrome b (or box) gene derived from R-loop and petite deletion mapping. The coding region of 900 nucleotides is split into at least five exons (depicted as rectangles) by four large introns; the whole gene spans 8 to 9 kbp. Henceforth each locus will be refered to by its box locus number (defined by recombinational linkage in order of discovery) as well as its physical exon or intron number (in order from 5' to 3'). Thus box 7 is the third intron.

Functions within Introns

The first surprise in the study of this cytochrome b gene came from complementation between box mutants (15). Analysis of complementation was done by following the kinetics of oxygen utilization after mating two cytochrome b mutants (respiration deficient). The mutant mitochondria presumeably fuse in the zygotes soon after mixing and at 9 to 11 hours one sees respiration, if the mutants are in separate genes, due to complementation or at 17 to 19 hours due to recombination if in the same gene. A priori one expects all of the mutations within a single cytochrome b gene to be in the same

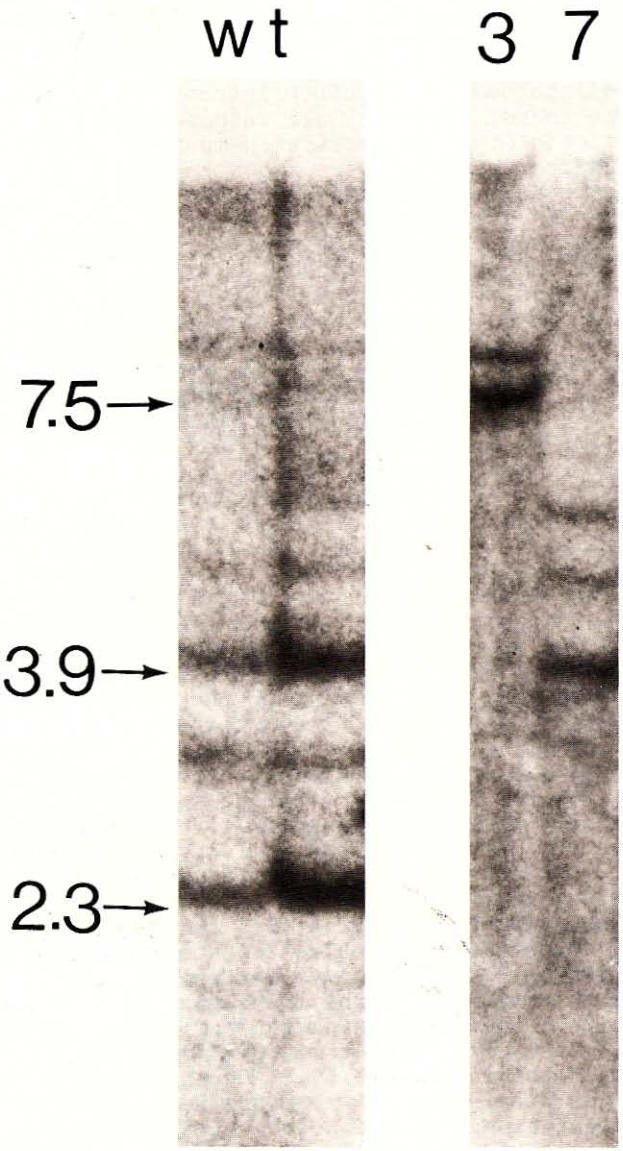

Figure 2. Analysis of mitochondrial RNA transfered to diazo paper (36) and hybridized to a 200 bp DNA fragment from the box-6 region. Wild type, box 3-2 and box 7-1 mutants are indicated. RNA sizes are in kb.

complementation group. Indeed all exon mutants (many of which terminate the cytochrome b chain) fail to complement each other in all combinations tested; intron mutants fail to complement other intron mutants alleles at the same locus. However, all intron mutants complement each other and most exon mutants. Thus box 3, box 10, and box 7 introns were shown to define three independent trans-acting elements involved in some unknown step in cytochrome b gene expression.

The next important finding came when the RNA's from the intron mutants were analyzed (12,14). figure 2 shows total mitochondrial RNA from wild type, box-3 (first intron) and box-7 (second intron) mutants separated by electrophoresis in denaturing methyl mercury agarose, blotted and hybridized with nick-translated DNA from the box-6 (fifth exon) region. The wild type pattern shows a prominent 2.3 kb RNA. Normally the 3.9 kb band is quite weak in the wild type. This preparation shows weak putative precursor RNA bands all the way up to 8.5 kb. The box-3 (first intron) mutants accumulate vast quantities of the 7.5 and 8.5 kb RNA's barely visible in wild type. The 8.5 Kb RNA may contain all introns intact. The 7.5 kb RNA probably has lost a piece of the 2kb first intron in the form of a 1 kb circular RNA (12). The box -7 (third intron) mutants accumulate the 3.9 kb RNA (which has lost the first two introns) and a subset of the larger RNA's. No 2.3 kb cytochrome b mRNA is visible in either mutant. We conclude that the trans-acting elements defined by the complementation data are required for RNA splicing.

How might the intron encoded trans-acting elements function? The RNA from these introns could act in trans either directly as base-pairing guides for splicing (figure 3) or as mRNA's for RNA splicing enzyme subunits (figure 4). We note at the top of figure 3 that in RNA guide models wild type RNA splicing normally need only use intramolecular RNA guidance (not trans acting RNA). The purely schematic bimolecular RNA complex shown at the bottom of figure 3, represents a variety of possible interactions of introns either free or as part of pre-mRNA's with sequences involved in aligning the splice junctions. Other alternative mechanisms involving recombination between mutant RNA's during splicing or otherwise has been shown to be an unlikely explanation for complementation. This was done by tagging exons with drug resistance and chain-termination markers in complementation tests (18,19). These experiments showed that only the exons from a single precursor RNA were brought together by splicing.

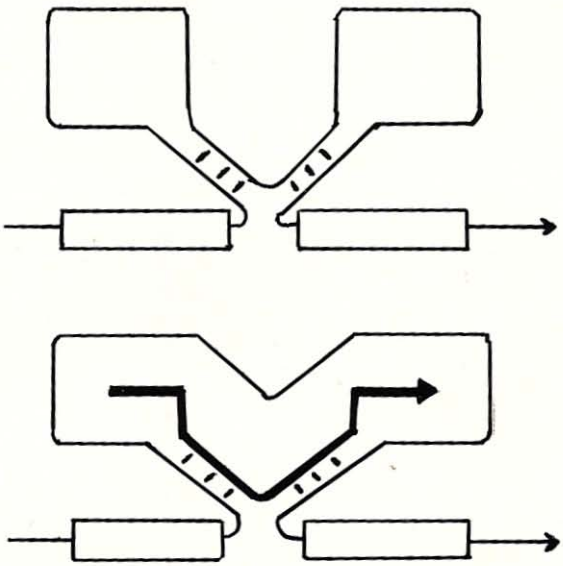

Figure 3. A typical guide RNA hypothesis for (A) wild type intramolecular RNA folding and (B) complementation of intron mutant RNA (thin outer line with rectangles as exons) by exon mutant RNA (thick inner arrow). The short regions of basepairing by intron sequences serve to align the junctions for splicing. The functional intron sequences found in the exon mutant replace the defective guide.

Cytochrome b RNA's Found in Petite Deletion Mutants

A basic prediction of the guide RNA hypotheses is that only nuclear encoded RNA splicing enzymes need be required for cytochrome b RNA processing. In contrast, intron encoded protein hypotheses require mitochondrial protein synthesis. How might one test this critical prediction? The use of mitochondrial protein synthesis inhibitors for enough cell divisions to allow dilution of mitochondrially encoded proteins results in accumulation of drug resistant and petite deletion mutants (20). We chose to take advantage of a unique property of the petite deletion mutants. These deletions lack tRNA's or rRNA's required for mitochondrial protein synthesis and can be propagated in this state indefinitely.

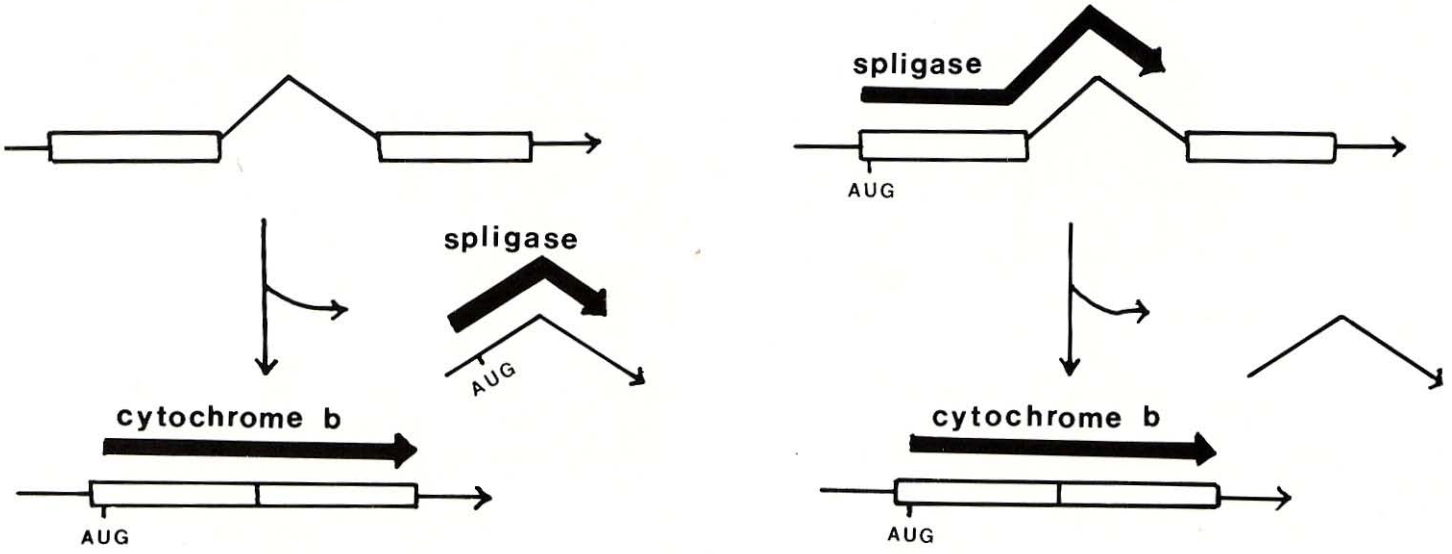

Figure 4. Intron encoded spligase hypotheses. Two representatives of a variety of translational start sites and phases are shown. Solid arrows indicate proteins encoded by exons (rectangles), introns (thin lines), or both.

Two of the petites that we have studied are depicted in figure 5; one has retained the large rRNA and the other the small rRNA gene. Both have an intact cytochrome b gene since they can restore respiratory function by recombination with a deletion of the entire box region. figure 6 shows that rRNA's in the two petites are transcribed and processed normally in the absence of mitochondrial protein synthesis (more than 60 generations after the deletion of the rRNA genes). The wild type RNA enriched for mitochondrial RNA displays, in lane C, four prominent bands: in order from the top, large cytoplasmic, large mitochodrial, small cytoplasmic, and the small mitochondrial rRNA's. The petite OIP2 has only the small mt rRNA. The petite IS2509 has normal sized large mt rRNA. Furthermore, hybridization of the IS2509 RNA with pure intron DNA (21) also indicates that splicing of the lg rRNA is normal in this petite (not shown). We conclude that all of the protein components required for the

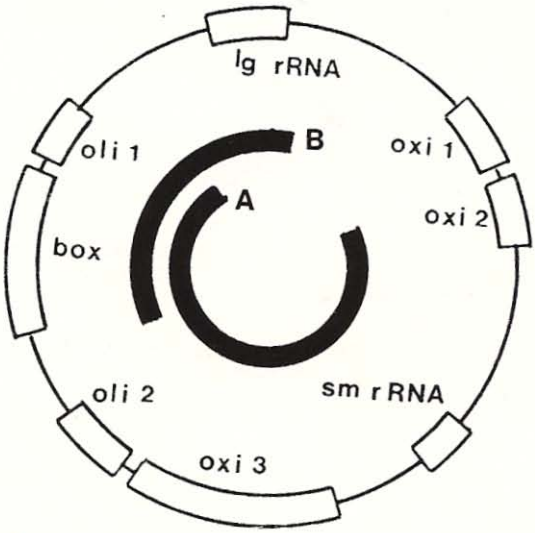

Figure 5. DNA sequences retained in petites used for analysis of rRNA and cytb RNA are indicated by the black inner curved bars. A) Petite OIP2 (37). B) petites K114-4A/I21/IS2509, 3319 (Dujon and Church unpublished).

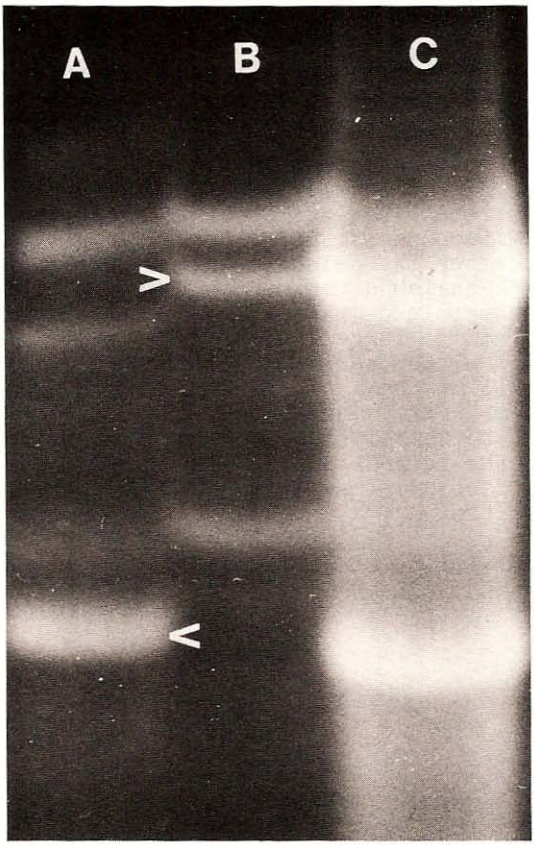

Figure 6. Petite deletion strains (A=OIP2, B= IS2509) and wild type (C) rRNA's visualized by ethidium. Normal sized small (A) and large (B) mt rRNA's are evident in the petites (white arrows).

processing for the large rRNA must be encoded in the yeast nucleus. The lanes labeled petite s in figure 7 show the pattern arising after hybridization of the petite RNA's with the box-6 (fifth exon) probe. There is a major 7.5 kb partially spliced RNA and a minor 8.5 kb RNA. No fully processed cytochrome b RNA appears at 2.3 kb (the pattern is identical to that found in box-3 mutants). So both petites fail to process cytochrome b RNA past the initial initial splicing step (which is also unaffected by all intron mutants). Although the mitochondrial ribosome or RNAs not found in either petite may be required as a cofactors for the splicing, the simplest interpretation of this result is that a mitochondrially encoded protein is required for cytochrome b RNA splicing.

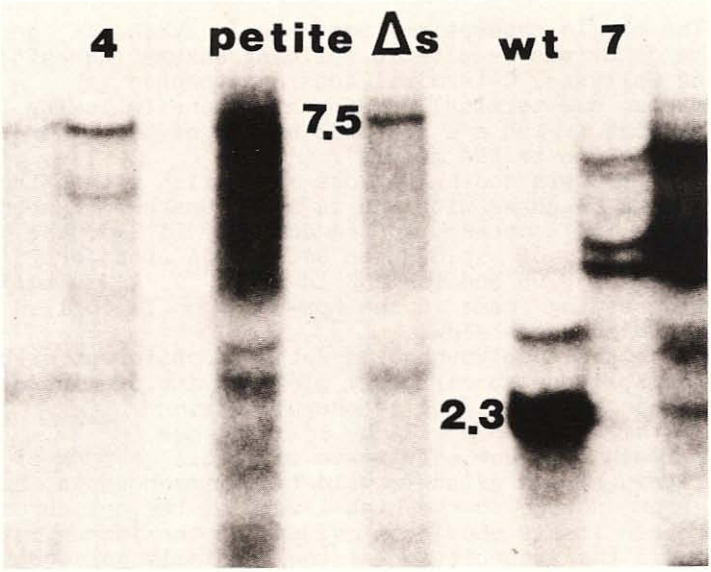

Figure 7. Cytochrome b RNA's found in the petite deletions lacking mitochondrial protein synthesis wild type and two pleiotropic box point mutants. The lanes displayed are: 4 = box 4-2 (G1334); Δs= IS2509,3319; wt = 7773A wild type mitochondrial genes; 7 = box 7-1 (G1659).

Further Predictions of the Hypothesis that Splicing Enzymes are Encoded by the Introns

These observations force us to look more closely at those models in which the trans-acting functions encoded in the cytochrome b introns would be splicing enzymes. As shown in figure 4 an initiation site for protein synthesis within the intron might be recognized in the pre-mRNA or on the free intron. Alternatively translation across a splice junction could occur either in phase or out of phase with the cytochrome b gene. The hypothesis that the active splicing enzymes are initiated in phase with and near to the cytochrome b N-terminus and then read through the splice junction, seems most consistent with a variety of additional observations on cytochrome b mutants. figure 8 briefly summarizes data on the molecular phenotypes of several exon, intron, and putative junction mutants that have been analyzed, including the sizes of the cytochrome b RNA's that accumulate (14,12), the sizes of the new proteins that appear (22-27) and the pleiotropic effects on other mitochondrial gene products: oxi-3 protein and RNA (14,22,23,25).

The single assumption that protein synthesis across splice junctions results in splicing enzyme (hereafter called spligase) C-terminal domains attached to cytochrome b N-terminal domains makes the following predictions (which are not made by any of the other intron protein or RNA models):

1) Long protein coding regions must exist within the introns in phase with the adjacent upstream cytochrome b exon (in contrast, for random DNA 18% G+C, the probability of avoiding an UAG or UAA stop for 35 codons is 0.05 and for 250 is 4×10^{-10}). A 250 amino acid reading frame in the box-3 intron is implied by data discussed below.
2) Since the complementation data is consistent with each spligase specifically cleaving its own intron, each protein would be autoregulatory: destroying its own mRNA by splicing out. In such a way only trace catalytic amounts of each protein should exist in wild type mitochondria, but overproduction to the high levels of the cytochrome b protein itself should be evident in the intron mutants due to the accumulation of the partially spliced RNA's. Exactly such levels of new proteins larger than cytochrome b are generally seen (23,24,25,27).
3) If these new proteins are read through from the cytochrome b start into the intron, they must have N-terminal homology to cytochrome b. Preliminary immunoprecipitation and peptide mapping results are consistent with this prediction (24,27).
4) The intron mutants, which have been chosen to be non leaky, must be missense or nonsense mutants in some mitochondrial protein. If the mutations are terminators, the sizes of the new overproduced proteins must be colinear with the genetic map position of the alleles. The genetic order (figure 8) of the box-3 alleles (5' to 3'): 3-1, (3-2,3-3,3-4), 3-5 fits well with a 42000 dalton spligase terminated at 30,34, and 38000 daltons (15). In this interpretation, box3-2 and 3-4 must be missense mutations.
5) Double mutants with an exon mutation upstream from an intron mutation should have detectable levels of only the short protein (terminated within the exon), as was found in all of the five such double mutants inspected (28).
6) The box-3 (first intron) mutants should accumulate the largest cytochrome b RNA's since the mRNA's for the spligases specific for subsequent steps are not matured (see figure 9). Box-10 (second intron) mutant RNA's should be one intron smaller. Box-7 (third intron) mutant RNA's still one smaller. True:the sizes are 7.5, 6.2, 3.9 kb, respectively (12,14).

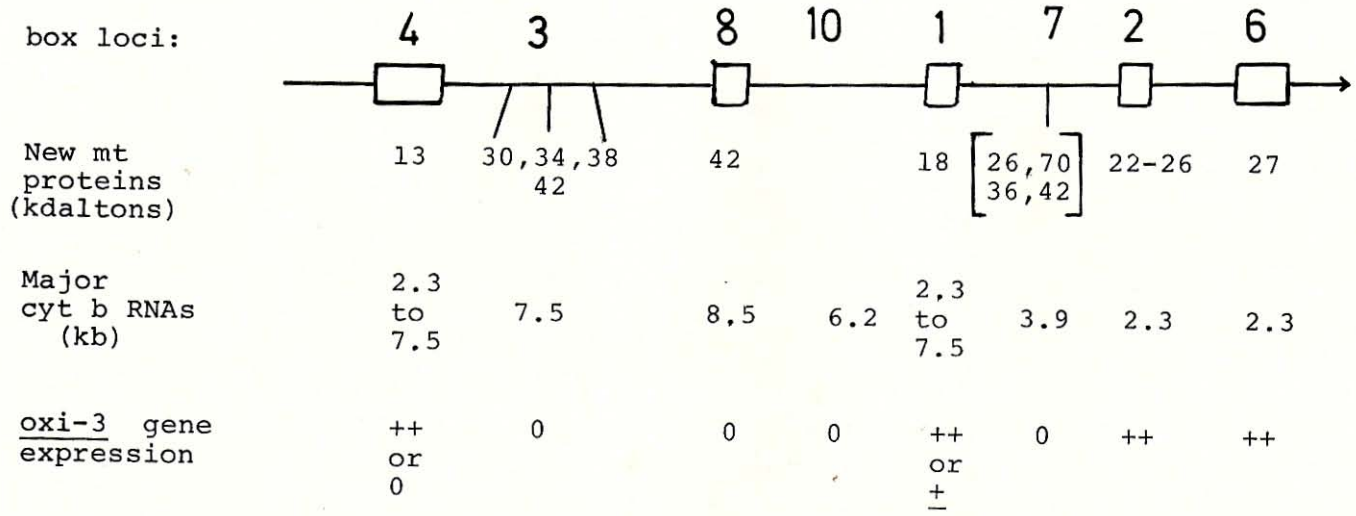

Figure 8. Summary of the sizes of the mutant cytochrome b RNA's and mitochondrial proteins and the pleiotropic effects on oxi-3 gene expression. Each column is headed by the box locus with the phenotypes below. The protein data for 14 exon (or junction) and 8 intron mutants are summarized as are the RNA data for 8 exon and 5 intron mutants. The genetic order of five alleles at three box-3 subloci (13) is indicated. Different box-4 and box-1 alleles give different RNA distributions and degree of oxi-3 expression (see text). Oxi-3 phenotype refers to disappearance of spectral band, enzymatic activity, 40000 dalton protein, and 2.3 kb oxi-3 RNA (with concomittant increase in 3.3 kb oxi-3 RNA).

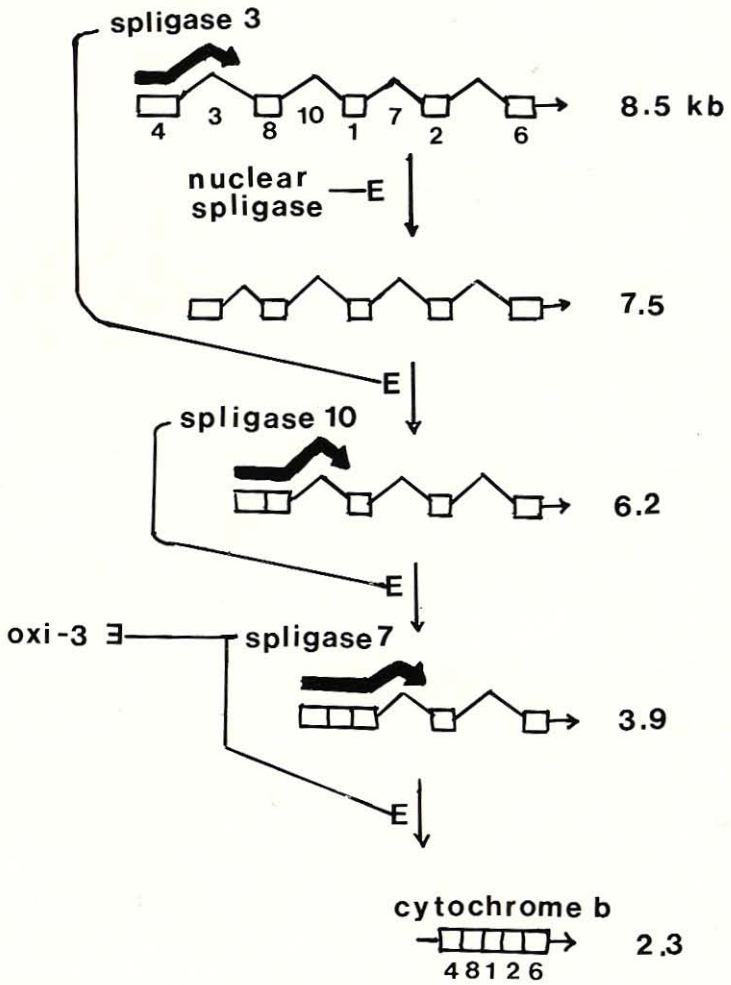

Figure 9. Detailed model for cytochrome b RNA splicing showing sequence of spligase mRNA generation and autoregulation. The RNA structures and sizes in kb are hypothetical and refer to RNA's accumulating in the mutants (figure 8).

7) Since box-7 mutants fail in some step of oxi-3 RNA processing (accumulating a 3.3 kb RNA in place of the normal 2.3 kb species), this intron must encode a product required for this process. This may be the same protein which is required for the splicing out the last two cytochrome b introns (figure 9).
8) Exon termination mutants in box-4,-8, and -1 (first,second, and third exons) but not necessarily missense mutants and not box-2 and box-6 mutants should:
   8a) accumulate cytochrome b RNA characteristic of adjacent downstream intron mutants. Box 4-2 resembles 3-2 and 3-4; 4-1 and 1-2 show abnormal levels of cytochrome b RNA precursors but detectable levels of 2.3kb mRNA (12,14, figure 7).
   8b) complement only downstream intron mutants poorly. Box 4-1,4-2 fail to complement box 3-1; box 8-1 complements 3-5 better than 10-1 and 7-1; box 1-1 complements 3-1 better than 7-1; however, box 1-2 complements box 7-1 fine (15,29).

   8c) fail in the same step of oxi-3 RNA processing as box-7 mutants. For alleles box 4-1,4-2,1-1, and 1-2 this seems to be the case (14,15, Church unpublished).

This hypothesis is an example of the general suggestion that alternate readings and splicings play a role in providing alternate proteins, that one protein's introns can be another protein's exons.

Questions Raised by the Specific Model

A few observations not obviously explained by this hypothesis deserve mention. At first glance, the chain termination mutations within the exons should fail to complement downstrean intron mutations, since they should also terminate the corresponding spligase. However, the overproduction then of the spligase messenger could permit enough read-through of a slightly leaky nonsense codon (which is often observed in other systems (30-33)) to generate a functional, enzymatic amount of spligase. In fact, in these mutants the splicing pattern is aberrant, only a reduced amount of mature mRNA but high levels of precursors.

Box 2-4 has temperature sensitive effects on oxi-3 but may lie at the junction of box-7 and box-2 (Claisse et al., 1980) or affect some function in the last intron. Far more mutations have been found within the cytochrome b exons and box-3 locus than within the box-10 or box-7 loci. This may reflect a smaller functional unit or active site (e.g. an oligonucleotide guide RNA) or peculiarities of the mutagenic or screening procedures.

Several novel large proteins found in some pleiotropic mutants are still unexplained. These may represent post-translational modification, new reading frames, or oxi-3 encoded proteins (whose mRNA's have accumulated for reasons outlined above). We would suggest that these other proteins are not translated from random non-coding regions but like the hypothetical fused splicing enzymes in figure 9 have functions in wild type, some possibly unrelated to RNA processing.

Some intriguing directions for inquiry are suggested by this hypothesis of fused splocing enzymes encoded in introns: Are the intron protein functions sufficient for splicing or are nuclear proteins and/or mitochondrial guide RNA's also required? Are the three intron spligases homologous? If so, was a special transposition mechanism used to generate a domain triplication (34)? Of what use is an N-terminal cytochrome b domain for a splicing enzyme? Perhaps membrane binding? How many proteins in this and other systems translate across splice junctions and read through terminator codons normally? So far papovavirus early genes (35) represent examples of the former, TMV and MuLV polymerase genes the latter (30-32). How many other uncommon proteins lie hidden in the "extra" DNA of the yeast mitochondrion whose genome is five times larger than that of other mitochondria? Finally, if the cytochrome b mRNA processing really requires four separate RNA splicing recognition subunits how does one interpret the apparent universality of nuclear splicing activities?

ACKNOWLEDGMENTS

We thank B.Dujon and H.Fukuhara for gifts of petite strains and advice; P.P. Slonimski, A. Halbreich, W. Herr, P. Lomedico, R. Morimoto, and D. Wirth for provoking discussions. This work was supported by National Institutes of Health grant GM09541 to W.G. G.M.C. is supported by Training Grant GM07598-02. W.G. is an American Cancer Society Professor of Molecular Biology.

Claude Jacq, Jaga Lazowska, and Piotr P. Slonimski have determined the DNA sequence of part of the intron containing box-3 mutations and find a continuous reading frame of at least 170 amino acids. Box-3-5 is a terminator in this frame, while box 3-2 is a double missense mutation. They have suggested that box-3 might encode a protein involved in RNA splicing. (Comptes rendus Acad. Sci. Paris 290).

REFERENCES

1. Borst, P. and Grivell, L.A. (1978) Cell 15, 705.
2. Fox, T.D. (1979) Proc. Nat. Acad. Sci. USA 76, 6534.
3. Corruzi,G. and Tzagoloff, A. (1979) J. Biol. Chem. 254, 9324.
4. Li,M. and Tzagoloff, A. (1979) Cell 18, 47.
5. Hensgens, L.A.M., Grivell, L.A., Borst, P. and Bos, J.L. (1979) Proc. Nat. Acad. Sci. USA 76, 1663.
6. Macino, G. and Tzagoloff, A. (1979) J. Biol. Chem. 254,4617.
7. Miller, D.L., Martin, N.C., Pham, H.D., and Donelson, J.E. (1979) J. Biol. Chem. 254, 11735.
8. Grivell, L.A., Arnberg, A.C., Boer, P.H., Borst, P., Bos, J.L., Van Bruggen, E.F.J., Groot, G.S.P., Hecht, N.B., Hensgens, L.A.M., Van Ommen, G.J.B., and Tabak, H.F. (1979) in Extrachromosomal DNA: ICN-UCLA Symposia on Molecular and Cellular Biology, vol 15 ,Cummings, D., Borst, P., Dawid, I., Weissman, S., and Fox, C.F., eds. (New York:Academic Press), in press.
9. Bos, J.L., Heyting, C., Borst, P., Arnberg, A.C., and Van Brugger, E.F.J. (1978) Nature 275, 336.
10. Faye, G., Dennebouy, N., Kujawa, C., and Jacq, C. (1979) Mol. Gen. Genet. 168, 101.
11. Van Ommen, G.-J.B., Groot, G.S.P., and Grivell, L.A. (1979) Cell 18, 511.
12. Halbreich, A., Pajot, P., Foucher, M., Grandchamp, C., and Slonimski, P. (1980) Cell in press.
13. Slonimski, P.P., Pajot, P., Jacq, C., Foucher, M., Perrodin, G., Kochko, A., and Lamouroux, A. (1978b) in Biochemistry and Genetics of Yeast, Bacila, M., Horecker, B.L., and Stoppani, A.W.N., eds. (New York: Academic Press), pp. 339.
14. Church, G.M., Slonimski, P., and Gilbert, W. (1979) Cell 18,1209.
15. Slonimski, P.P., Claisse, M.L., Foucher,M., Jacq, C., Kochko, A., Lamouroux, A., Pajot, P., Perrodin, G., Spyridakis, A., and Wambier-Kuppel, M.L. (1978a) in Biochemistry and Genetics of Yeast, Bacila, M., Horecker, B.L., and Stoppani, A.W.N., eds. (New York: Academic Press), pp. 391.
16. Lerner, M.R., Boyle, J.A., Mount, S.M., Wolin, S.L., Steitz, J.A. (1980) Nature 283, 220.
17. Murray, V. and Holliday, R. (1979) FEBS Lett. 106, 5.
18. Kochko, A., Colson, A.M., Lamouroux, A., Pajot, P., and Slonimski, P.P. in preparation.
19. Dujon, B. (1979) Nature 282, 777.
20. Williamson, D.H., Maroudas, N.G., and Wilkie, D. (1971) Molec. Gen. Genet. 111, 209.

21. Morimoto, R. and Dujon, B. (1980) in preparation.
22. Alexander, N.J., Vincent, R.D., Perlman, P.S., Miller, D.H., Hanson, D.K., and Mahler, H.R. (1979) J. Biol. Chem. 254, 2471.
23. Claisse, M.L., Spyridakis, A., Wambier-Kuppel, M.L., Pajot, P., and Slonimski, P.P. (1978) in Biochemistry and Genetics of Yeast, Bacila, M., Horecker, B.L., and Stoppani, A.W.N., eds. (New York: Academic Press), pp. 369.
24. Claisse, M., Slonimski, P.P., Johnson,J., and Mahler, H. (1980) Molec. Gen. Genet. in press.
25. Haid, A., Schweyen, R.J., Bechmann, H., Kaudewitz, F., Solioz,M., and Schatz,G. (1979) Eur. J. Biochem. 94, 451.
26. Hanson, D.K., Miller, D.H., Mahler, H.R., Alexander, N.J., and Perlman, P.S. (1979) J. Biol. Chem. 254, 2480.
27. Solioz, M. and Schatz, G. (1979) J. Biol. Chem. 254, 9331.
28. Mahler, H.R. and Perlman, P.S. (1979) in Extrachromosomal DNA: ICN-UCLA Symposia on Molecular and Cellular Biology, vol 15 ,Cummings, D., Borst, P., Dawid, I., Weissman, S., and Fox, C.F., eds. (New York:Academic Press), in press.
29. Lamouroux, A., Pajot, P., Kochko, A., Colson, A.M., and Slonimski, P.P. in preperation.
30. Pelham, H.R. (1978) Nature 272, 469.
31. Phillipson, L., Andersson, P., Olshevsky, U., Weinberg, R., and Baltimore, D. (1978) Cell 13,189.
32. Murphy, E.C., Kopchick, J.J., Watson, K.F., and Arlinghaus,R.B. (1978) Cell 13, 359.
33. Geller, A.I. and Rich, A. (1980) Nature 283, 41.
34. Gilbert, W. (1978) Nature 271, 501.
35. Rigby, P. (1979) Nature 282, 781.
36. Alwine, J.C., Kemp, D.J. and Stark, G.R. (1977) Proc. Nat. Acad. Sci. USA 74, 5350.
37. Lewin, A., Morimoto, R., Rabinowitz, M., and Fukuhara, H. (1978) Molec. Gen. Genet. 163, 257.

DISCUSSION

I. HERSKOWITZ: Is it possible that some of your complementation results with different oxi-3 mutants are due to recombination rather than complementation?

G.M. CHURCH: I have not done these experiments myself so I can't give an expert opinion on this, but this complementation has been done with a number of genes. For example, when you pair cytochrome B with oxi-1, 2 and 3 mutants you see complementation. For all pairs of exon mutants that have been

looked at, you see only the recombinations not complementation. Those are the only internal controls that I know of. In every case, time points are taken every half hour and screened for recombinants.

R. BUTOW: Do you know if the pattern of splicing that you see in a given petite mutant is independent of the nuclear background?

G.M. CHURCH: I have looked at diploids where the nuclear background has been changed to introduce the OP-1+ allele for the box mutants and those look the same, but I haven't looked at effects on petites yet.

A. HALBREICH: There are a few things I wanted to correct here. First of all, there is a distinct difference between complementation and recombination in the sense that any pairwise combination of two mutants will eventually restore wild type phenotype by recombination, but only the four complementation groups have been seen. Furthermore, by using the petite mutants to complement box mutants, one can make the distinction between those mutants which are complemented and those that can complement. In this way it can be seen that no exon mutant can be complemented. They can only complement intron mutants. The same conclusion can be derived from the experiments on the combinations of diuron resistance with exon and intron mutations, namely that in all cases of complementation, all the exons from one parental gene with the mutated intron are expressed. In other words, the only thing that you can do in complementation is to overcome an intron mutation and allow splicing to occur. In this way, one can say with certainty that box-8 is not complemented; it complements box-3 mutations but does not complement box-10 and box-7 mutations.

G.M. CHURCH: I agree.

A. HALBREICH: The other point of criticism, on your proposed spligase concerns the fact that box-4 mutants have a stop codon inside that exon; consequently they cannot translate beyond that exon into the next intron, nevertheless, most of them are processed normally. Similarly, the box-1 mutant,

cannot be translated beyond box-1 and therefore cannot possibly make the last spligase that you eluded to. In conclusion, there is no possibility for this kind of a model.

G.M. CHURCH: I'll basically just address the box-1 situation. As I've mentioned, certain box-1 alleles, in fact two out of the original three, are pleiotropic in terms of gene expression, which indicates that they are, in fact, affecting a downstream function. In addition, one of them complements box-7-1 poorly. I agree in the sense that this is a weak point in the theory.