INFORMATION TO USERS

The most advanced technology has been used to photograph and reproduce this manuscript from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book. These are also available as one exposure on a standard 35mm slide or as a 17" x 23" black and white photographic print for an additional charge.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
Mobile genetic elements in yeast mitochondrial-DNA

Wenzlau, Janet Marie, Ph.D.
The Ohio State University, 1989
MOBILE GENETIC ELEMENTS
IN YEAST MITOCHONDRIAL DNA

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By

Janet Marie Wenzlau, B. S.

* * * * *

The Ohio State University
1989

Dissertation Committee:
P. S. Perlman
C. W. Birky
C. Brietenberger
A. M. Lambowitz

Approved by

Advisor
Department of Molecular Genetics
To my parents, Tom and Nancy Wenzlau
AKNOWLEDGMENTS

I would like to thank the members of my dissertation committee, Drs. C. W. Birky, C. Breitenberger, and A. M. Lambowitz for their suggestions and comments. To Dr. Roland Saldanha, with whom much of this work was done in collaboration, I would like to express my appreciation for sharing his scientific expertise. I am grateful to Drs. Ronald Butow and Heather Conrad-Webb for their advice and support. I would like to thank Dr. Philip Perlman for giving me the opportunity to earn this degree and for helping me develop both technical and critical abilities. My stay in his laboratory was enhanced by his generosity in allowing me to represent my work at numerous scientific meetings. Finally, I would like to thank my family for their encouragement and understanding.
VITA

March 24, 1961.................Born, Appleton, Wisconsin

1983.........................B. S.,
               Denison University,
               Granville, Ohio

1983-present...............Research Assistant,
               Department of Molecular
               Genetics
               The Ohio State University

PUBLICATIONS

J. M. Wenzlau, R. J. Saldanha, R. A. Butow, and P. S.
Perlman. 1989. A latent intron-encoded maturase is also
an endonuclease needed for intron mobility. Cell 56, 421-
430.

J. M. Wenzlau and P. S. Perlman. Gene conversion of var1
alleles in the yeast mitochondrial genome. In
preparation.

Analysis of the recognition site of the aI4a-encoded

R. C. Dietrich, J. M. Wenzlau, D. Ralph, K. Jarrell, A.
Morawiec, and P. S. Perlman. 1985. Mitochondrial genome
variation in the genus Saccharomyces. Poster presentation
at the International Symposium on Achievements and
Perspectives in Mitochondrial Research, University of
Bari, Rosa Marina, Italy. September 2-6, 1985.

of var1 alleles of yeast mtDNA. Poster presentation at
the ICSABER Twelth Annual Forum of Graduate Students in
the Biological and Health Sciences, The Ohio State
University, Columbus, Ohio. April 22, 1986.


J. M. Wenzlau, R. Saldanha, R. Butow, and P. S. Perlman. 1988. A latent intron-encoded maturase is also an


FIELD OF STUDY: MOLECULAR GENETICS

vi
TABLE OF CONTENTS

DEDICATION ....................................................... ii
ACKNOWLEDGEMENTS................................................ iii
VITA........................................................................ iv
TABLE OF CONTENTS ................................................... vi
LIST OF TABLES ................................................... xii
LIST OF FIGURES ................................................... xiii

Chapter 1: INTRODUCTION ........................................ 1

I.A. Yeast mitochondrion as a model system ...................... 1

I.B. The yeast mitochondrial genome ................................ 2

   I.B.1. Yeast mitochondrial genes.............................. 2
   I.B.2. Organization of the mitochondrial genome .......... 3
   I.B.3. Variability in the yeast mitochondrial genome .... 4
   I.B.4. G+C-rich clusters ...................................... 4
   I.B.5. Introns in yeast mitochondrial genes .............. 8
   I.B.5.a. Group I introns ..................................... 10
   I.B.5.b. Group II introns .................................... 11
   I.B.6. Mitochondrial intron open reading frames ....... 12
   I.B.6.a. The maturase model ................................ 13
   I.B.6.b. The box phenomenon ................................ 14

I.C. Mitochondrial transmission genetics ......................... 15
   I.C.1. Mitochondrial genetic crosses ...................... 15
   I.C.2. Rules of mitochondrial inheritance ............... 17

I.D. Recombination events in the mitochondrial genome involving GC clusters ......... 20
   I.D.1. Petite formation ...................................... 21
   I.D.2. Petite integration .................................... 22
   I.D.3. GC clusters as mobile genetic elements .......... 23

I.E. The varl gene ................................................. 24

I.F. Transmission of the "a" GC-rich element in the varl gene ................. 25
I.G. The polarity phenomenon in the 21S rRNA gene...26
I.G.1. A site-specific intron-encoded endonuclease..........................27
I.G.2. The fitl substrate........................................28

I.H. Double strand breaks as initiators of recombination..................28

I.I. Reverse transcription and mitochondrial introns....................31
I.I.1. Some modes of reversion of mitochondrial intron mutants........32
I.I.1.a. Reversion by intron excision................................33
I.I.1.b. Other modes of reversion of intron mutations..................35
I.I.2. Reverse transcriptase activity in other fungal mitochondria...36
I.I.2.a. Senescence in Podospora anserina..........................36
I.I.2.b. Neurospora mitochondrial plasmids...........................37

I.J. Issues explored in this dissertation...........................38

Chapter II: MATERIALS AND METHODS.................................41

II.A. Strains....................................................41
II.A.1. Yeast strains........................................41
II.A.2. Recombinant plasmids.......................................41

II.B. Culture media..............................................41
II.B.1. Yeast media.............................................41
II.B.1.a. Non-selective media..................................41
II.B.1.b. Glycerol media........................................42
II.B.1.c. Drug resistance media...................................42
II.B.1.d. Minimal media..........................................42
II.B.1.e. Sporulation media......................................42
II.B.2. E. coli media.............................................42

II.C. Genetic techniques..........................................43
II.C.1. Mating..................................................43
II.C.2. Petite deletion mapping...................................44
II.C.3. Sporulation of diploids..................................44
II.C.4. Karl-mediated cytoduction................................45
II.C.5. Protoplast Fusion.........................................46
II.C.6. Ethidium bromide mutagenesis/petite induction..............46
II.C.7. TTC overlay..............................................47
II.C.8. MnCl2 mutagenesis and Adriamycin suicide treatment.........48
II.C.9. Selection of revertants...................................49
## II.C.10. Strain constructions

**49**

<table>
<thead>
<tr>
<th>II.D.</th>
<th>Colony hybridization</th>
<th>53</th>
</tr>
</thead>
<tbody>
<tr>
<td>II.E.</td>
<td>Mitochondrial DNA isolation</td>
<td>53</td>
</tr>
<tr>
<td>II.F.</td>
<td>DNA restriction site analysis</td>
<td>53</td>
</tr>
<tr>
<td>II.G.</td>
<td>Southern blot analysis</td>
<td>53</td>
</tr>
<tr>
<td>II.H.</td>
<td>Cloning of mitochondrial restriction fragments</td>
<td>54</td>
</tr>
<tr>
<td>II.I.</td>
<td>Plasmid preparation</td>
<td>55</td>
</tr>
<tr>
<td>II.J.</td>
<td>Dideoxy sequencing</td>
<td>55</td>
</tr>
<tr>
<td>II.K.</td>
<td>Site-directed mutagenesis</td>
<td>56</td>
</tr>
<tr>
<td>II.L.</td>
<td>Random <em>in vitro</em> mutagenesis</td>
<td>56</td>
</tr>
<tr>
<td>II.M.</td>
<td>Labeling of mitochondrial translation products</td>
<td>57</td>
</tr>
<tr>
<td>II.N.</td>
<td>Mitochondrial extracts</td>
<td>57</td>
</tr>
</tbody>
</table>

### Chapter III: RESULTS

**58**

| III.A. | GC cluster conversion in the \textit{varl} gene 
<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>III.A.1.</td>
</tr>
<tr>
<td>III.A.2.</td>
</tr>
<tr>
<td>III.A.3.</td>
</tr>
<tr>
<td>III.A.4.</td>
</tr>
<tr>
<td>III.A.5.</td>
</tr>
<tr>
<td>III.A.6.</td>
</tr>
<tr>
<td>III.A.7.</td>
</tr>
<tr>
<td>III.A.8.</td>
</tr>
<tr>
<td>III.A.9.</td>
</tr>
<tr>
<td>III.B.</td>
</tr>
<tr>
<td>III.B.1.</td>
</tr>
<tr>
<td>III.B.2.</td>
</tr>
</tbody>
</table>
III.B.3. Survey of mitochondrial introns or mobility............................77
III.B.4. CoxI intron aI4a is mobile in crosses...78
III.B.5. The structure of recombinants which have acquired aI4a.....................81
III.B.6. Recombinants do not have a selective advantage...........................82
III.B.7. Flanking marker co-conversion accompanies aI4a conversion.................84
III.B.8. Site-specific double-strand cleavage of recipient coxI genes...............84
III.B.9. aI4a conversion requires mitochondrial translation........................87
III.B.10. The source of the endonuclease required for aI4a conversion...............88
III.B.11. Mutations in coxI intron aI4a....................................90
III.B.12. Mutations that truncate the aI4a protein block conversion...............94
III.B.13. Sequence-specific endonuclease activity in mitochondrial extracts.....97
III.B.14. The cleavage site of the aI4a encoded endonuclease.......................98
III.B.15. Dissection of the recognition site of the aI4a endonuclease..........100

III.C. Another mechanism for intron mobility......102
III.C.1. Intron mobility through RNA intermediates..........................102
III.C.2. A strategy to assay reverse transcriptase in yeast mitochondria....105
III.C.3. C2116 is a suitable reporter intron mutation..........................106
III.C.4. Intron subtraction of aI1 and aI2 from the mutant C2116............110
III.C.5. CoxI intron 1 is sufficient for intron excision.......................113

Chapter IV: DISCUSSION.................................117

IV.A. GC cluster conversion in the varl gene.......117
IV.A.1. GC clusters as recombinogenic elements.................................117
IV.A.2. Comparison between GC cluster conversion and intron conversion........118
IV.A.3. Co-conversion does not accompany common cluster conversion..............120
IV.A.4. The role of double strand breaks in GC cluster conversion............122
IV.A.5. A candidate enzyme involved in GC
IV.A.6. New strategies to study GC conversion ............................................. 125

IV.B. Mobile introns .................................................................................. 126
   IV.B.1. The aI4a intron-encoded protein promotes intron transmission .... 129
   IV.B.2. Intron conversion is not limited to yeast mitochondria .......... 130
      IV.B.2.a. Comparison of features of intron conversion in other systems.. 130
      IV.B.2.b. Intron mobility in T-even phage ........................................ 132
      IV.B.2.c. A mobile intron in Physarum polycephalum .................... 134
      IV.B.2.d. Gene conversion in other systems ...................................... 135
   IV.B.3. Cleavage by the intron-encoded enzymes .................................. 137
   IV.B.4. The recognition sequence for intron-encoded endonucleases ...... 138
   IV.B.5. Intron-encoded endonucleases as restriction enzymes ............. 141
   IV.B.6. Expression of the intron-encoded endonucleases ..................... 143
   IV.B.7. The aI4a-ORF encodes two functions ....................................... 145
   IV.B.8. Introns as mobile genetic elements .......................................... 146

IV.C. Transmission of group I introns by reverse splicing ....................... 150
   IV.C.1. Models for intron excision ..................................................... 154
   IV.C.2. Evolution of mitochondrial introns .......................................... 158
   IV.C.3. Evolution of intron encoded functions .................................... 161

TABLES .............................................................................................. 164
FIGURES ............................................................................................ 172
BIBLIOGRAPHY .............................................................................. 321
### LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Yeast strains used for genetic and physical analysis</td>
<td>164</td>
</tr>
<tr>
<td>2. Clones used for physical analysis</td>
<td>166</td>
</tr>
<tr>
<td>3. CoxI intron configuration of yeast strains</td>
<td>168</td>
</tr>
<tr>
<td>4. Intron configuration of strains assayed for conversion of aI4a</td>
<td>170</td>
</tr>
<tr>
<td>FIGURE</td>
<td>PAGE</td>
</tr>
<tr>
<td>--------</td>
<td>------</td>
</tr>
<tr>
<td>1.</td>
<td>173</td>
</tr>
<tr>
<td>2.</td>
<td>175</td>
</tr>
<tr>
<td>3.</td>
<td>178</td>
</tr>
<tr>
<td>4.</td>
<td>180</td>
</tr>
<tr>
<td>5.</td>
<td>182</td>
</tr>
<tr>
<td>6.</td>
<td>184</td>
</tr>
<tr>
<td>7.</td>
<td>186</td>
</tr>
<tr>
<td>8.</td>
<td>188</td>
</tr>
<tr>
<td>9.</td>
<td>190</td>
</tr>
<tr>
<td>10.</td>
<td>192</td>
</tr>
<tr>
<td>11.</td>
<td>194</td>
</tr>
<tr>
<td>12.</td>
<td>196</td>
</tr>
<tr>
<td>13.</td>
<td>198</td>
</tr>
<tr>
<td>14.</td>
<td>200</td>
</tr>
<tr>
<td>15.</td>
<td>202</td>
</tr>
<tr>
<td>16.</td>
<td>204</td>
</tr>
<tr>
<td>17.</td>
<td>206</td>
</tr>
</tbody>
</table>
18. Schematic diagram of the varl structure of the capensis recombinant

19. Strategy used for sequence analysis of the S. capensis recombinant

20. Backcrosses demonstrate directionality of the mobile cluster

21. Varl GC clusters are converted independently

22. Co-conversion screen for recombinant varl progeny

23. Co-conversion recombinant isolates

24. Strategy for Southern blot analysis of co-conversion recombinant varl isolates

25. Southern blot analysis of co-conversion recombinants using the 5' varl probe

26. Southern blot analysis of co-conversion recombinants using the varl ORF probe

27. Varl recombinants do not arise by homologous cross overs

28. Common cluster conversion is independent of mitochondrial translation

29. Schematic diagram of a cross to monitor intron transmission

30. Transmission of the omega intron in the 21S rRNA gene

31. Transmission of intron aI5B

32. Transmission of the 3' portion of the coxI gene

33. Transmission of introns aI3g and aI4a

34. The coxI gene structure of the recombinant

35. Location of intron aI4a in the recombinant

36. Oligomycin sensitive recombinant isolates

xiv
37. Reciprocal exchange following crossing over in exon 5 ........................................ 246
38. Mixed growth tests of diploids harboring different mitochondrial genomes .................. 248
39. Norbensis sequence polymorphisms in the 3' exon flanking aI3g............................ 250
40. In vivo double strand cuts in recipient DNA molecules.............................. 252
41. Al4a conversion requires mitochondrial translation........................................ 254
42. AI4a conversion in a cross between D273-10B and S. capensis............................. 256
43. AI4a conversion in a cross between D273-10B and S. ellipsoideus........................... 258
44. AI4a conversion in a cross between S. norbensis and ID41-6/161............................ 260
45. AI4a conversion in a cross between S. norbensis and ID41-6/ 1-3............................ 262
46. AI4a conversion in a cross between S. norbensis and G-II-5g and G-II-0................. 264
47. Mutations of the aI4a reading frame .............................................................. 266
48. Truncations of the aI4a reading frame block conversion ................................. 269
49. The cleavage site of the aI4a-encoded endonuclease .......................................... 272
50. Substrate differences between the recipient, donor and artificial substrate.............. 274
51. Comparison of sequence polymorphisms in the aI4a and fit1 substrates ................. 276
52. Strategy for dissection of the recognition sequences of the aI4a-encoded endonuclease. 278
53. AI4a mutant substrate............................................................ 280
54. Amino acid sequence alignment of reverse transcriptase related sequences.................283
55. Amino acid sequence alignment of reverse transcriptase related sequences.................284
56. Coxl structure of "popout" revertants.................286
57. C2116 revertants......................................288
58. Sequencing strategy of "popout" revertants........290
59. Sequence structure of TR revertants TR1 and TR2........................................292
60. ΔaIl, aI2-C2116 glucose TR revertants..............294
61. Coxl intron structure of C2116 derivatives.........296
62. Strategy of intron subtraction for construction of strain ΔaIl, aI2-C2116.............298
63. Southern analysis to confirm the structure of ΔaIl, aI2-C2116..............................300
64. ΔaIl, aI2-C2116 revertants..........................302
65. Southern blot analysis to confirm the coxl intron structure of ΔaIl, aI2-C2116........304
66. ΔaI2-C2116 TR revertants............................306
67. Gene conversion sites recognized by recombinogenic endonucleases......................308
68. Comparison of the in vitro reaction conditions for fit1 and the αI4a endonuclease......310
69. Modes of expression of intron-encoded endonucleases......................................312
70. The fused exon/intron αI4a reading frame............314
71. Mode of intron transposition via an RNA intermediate.........................................316
72. Conserved reverse transcriptase domains in coxl aIl........................................318
73. Model for interaction between intron conversion and intron excision ................. 320
CHAPTER 1

INTRODUCTION

I.A. Yeast mitochondrion as a model system

The mitochondrion of Saccharomyces cerevisiae is a model system amenable to studies of traditional areas of molecular biology such as gene regulation, gene structure, RNA processing and recombination, since both genetic and biochemical techniques can be readily exploited. Yeast is a simple eukaryote with a relatively small mitochondrial genome (75 kb). Because yeast has a rapid generation time and simple growth requirements, ample amounts of experimental material may be acquired quickly and inexpensively. The advancement of the yeast mitochondrial genome as a genetic system was made feasible by the fact that yeast is a facultative anaerobe. Mutations which abolish the ability of yeast to grow on nonfermentable substrates such as ethanol or glycerol (whose utilization depends on a functional respiratory chain) can maintain a stable mitochondrial genome in the absence of respiration.
on fermentable substrates. Such mutants are readily isolated and characterized and have led to the elucidation of the structure, function, and expression of the majority of the mitochondrial genes.

This thesis focuses on recombination events in the yeast mitochondrial genome. The remainder of the Introduction will describe the structure of the mitochondrial genome and its participation in various recombination processes.

I.B. The yeast mitochondrial genome

I.B.1 Yeast mitochondrial genes

The mitochondrial genome of *Saccharomyces cerevisiae* specifies a subset of the components required for respiration, oxidative phosphorylation and protein synthesis, while the majority of mitochondrial proteins are encoded by the nuclear genome. Most of the mitochondrial genes encoding subunits of the inner membrane enzyme complexes have been identified by genetic analysis.

The three largest subunits of cytochrome oxidase are encoded by the *oxi3*, *oxi1*, and *oxi2* genes which are now called *coxI*, *coxII* and *coxIII*, respectively. The *cob*
gene encodes the cytochrome b apoprotein. ATPase subunits 6, 9, and 8 are encoded by the oli2, oli1 and aapl genes, respectively. The mitochondrial genome contains many genes involved in mitochondrial translation. A protein associated with the small ribosomal subunit is encoded by the varl gene. Besides the protein encoding genes, the large (21S) and small (15S) mitochondrial ribosomal RNAs, 25 tRNAs, and the tRNA synthesis locus (tsl) (an RNA molecule involved in pre-tRNA processing) are encoded on the mitochondrial genome (reviewed by Dujon, 1981; Miller et al., 1983). In addition, five free standing unassigned reading frames (URFs) have been identified (reviewed by de Zamaroczy and Bernardi, 1985). The coxI, cob, and 21S rRNA genes genes contain intervening sequences, some of which contain long open reading frames (ORFs) which are capable of encoding proteins. The functions of intron encoded proteins will be addressed in detail in later sections.

I.B.2. Organization of the mitochondrial genome

The structure of the yeast mitochondrial genome is illustrated in Figure 1. The mitochondrial genome is remarkably A+T rich. More than half of the genome is composed of intergenic DNA stretches which are almost
exclusively A+T (95%) while the coding regions average 26% G+C (de Zamaroczy and Bernardi, 1987). At irregular intervals along the A+T stretches and occasionally in coding regions or introns, short, highly G+C-rich segments called GC clusters punctuate the genome. These GC clusters comprise the majority of the G+C content of the intergenic regions and have been organized into seven families based on their primary sequences (de Zamaroczy and Bernardi, 1987).

I.B.3. Variability in the yeast mitochondrial genome

The yeast mitochondrial genome displays strain specific size variations due to the insertion or deletion of various optional sequences. Strains have been reported which are deleted for urfl, urf2 and/or urf3 (Nagley et al., 1981; Thalenfeld et al., 1983; Michel, 1984). However, there are two predominant types of optional DNA sequences. One class comprises the introns located within the genes encoding cytochrome b, cytochrome oxidase subunit I, and the 21S rRNA. The other class includes the families of G+C-rich clusters.

I.B.4. G+C-rich clusters
G+C-rich clusters are typically 30-60 base pairs (bp) long and are found scattered throughout the genome. GC clusters fall into seven different classes based on their primary sequence. A consensus sequence has been determined for three of these families and is shown in Figure 2. Also shown is a representative cluster from each of the other classes for which a consensus sequence has not yet been determined. The seven GC cluster families (called a1, a2, a3, a4, v, c, and n) are all related to specific segments of the "origins of replication" of the mitochondrial genome. The a1 family shows the highest degree of homology among its members, while families a2, a3, and a4 display reduced homology among their members and with the a1 family. Families v and c share some homology with the a2 family whereas the n family represents short GC clusters which do not fall into any of the other families. The GC clusters in the varl gene [the common cluster and the "a" cluster (see Figure 4)] are members of the v family which are typically flanked by A+T inverted or complementary inverted repeats. Several families of GC clusters are palindromic (families a1, a2, a3 and v).

Most mitochondrial genomes contain about 210-220 GC clusters located mostly in intergenic regions. The few GC clusters that are found within the coding sequences of
genes do not appear to impair gene function. GC clusters have been identified in the 15S, 21S, and var1 mitochondrial genes. In addition, GC clusters are found in many introns and in several URFs.

GC clusters are believed to have originated from the first origin of replication (oriL) for the mitochondrial genome by a series of duplication and translocation events followed by amplification events (de Zamaroczy and Bernardi, 1986). The roughly 500 bp oriL sequence of the mitochondrial genome contains three GC-rich clusters, designated A, B, and C, which are separated by AT-rich stretches. Cluster B is the inverted complement of cluster A and the two can be folded into a hairpin structure. In each of the seven ori sequences of the wild-type mitochondrial genome, clusters A, B, and C are highly conserved (De Zamaroczy et al., 1984). The various families of GC clusters show a striking resemblance to the A, B, and C clusters found in the ori sequences. The al family, in particular, shares a resemblance to the ori sequences of the mitochondrial genome.

In addition to their role in recombination (discussed below), GC clusters may retain a remnant of their original ori function. Petites devoid of ori sequences, but retaining GC clusters of the al family display a correlation between their suppressivity and the number of
al GC clusters present (Goursot et al., 1982). However, direct evidence for this function is lacking. Others have proposed that GC clusters may function in the processing of primary transcripts. The palindromic sequences found in GC clusters may form suitable structures to generate signals for cleavage by RNA processing enzymes (Tzagoloff et al., 1980). However, only circumstantial evidence exists for this notion (Simon and Faye, 1984). According to de Zamaroczy and Bernardi (1986), the key role for GC clusters may be to modulate the structure and organization of the mitochondrial genome. The genome configuration likely plays a major role in gene regulation since the specific context of promotor sequences is known to determine whether or not they are active. The conservation of GC clusters in intergenic sequences of mitochondrial genomes among wild type strains suggests an indirect function for these sequences (de Zamaroczy and Bernardi, 1985).

Although the mitochondrial genomes of most higher eukaryotes (except higher plants) are virtually devoid of optional sequences at least one other fungal mitochondrial genome is known to be littered with G+C-rich sequences, although they are probably unrelated to those in the yeast mitochondrial genome. The mitochondrial genome of *Neurospora crassa* contains similar types of GC elements
termed "Pst I clusters" distributed about the rRNA and tRNA genes. Pst I clusters are 50-100 bp averaging 70% G+C and are characterized by an 18 bp palindrome termed the "core sequence". The core sequence is usually flanked by cytosines on the 5' side and guanosines on the 3' side which would extend the stem when a hairpin structure is formed. In many cases these GC-rich clusters serve as processing signals for the maturation of mRNA transcripts. The *Neurospora crassa* mitochondrial genome typically contains 50-100 "Pst I clusters" which are relatively conserved in position (Yin et al., 1981).

The mitochondrial genomes of *Euglena gracilis* (a green alga), *Ustilago cynodontis* (a basidiomycete), and *Acanthamoeba castellani* (a protozoan), and the chloroplast genomes of *Euglena gracilis*, and *Spinacia oleracea* (a plant), are also believed to contain GC clusters (Fonty et al., 1975; Mery-Druegon et al., 1981; Schmitt et al., 1981). Furthermore, the mitochondrial genome of *Kluveromyces lactis* is known to contain GC repetitive elements which were identified by their high proportion of SacII sites (Sor and Fukahara, 1982).

I.B.5. Introns in yeast mitochondrial genes

Most intervening sequences of the yeast
mitochondrial cob, coxI, and 21S rRNA genes are optional. That is, their presence or absence appears to have no phenotypic consequence. This fact was demonstrated most convincingly by Seraphin et al., (1987) who constructed a mitochondrial genome devoid of all introns. However, the presence of intron aI4a in the coxI gene requires the presence of intron bI4 in the cob gene. Also, the presence of intron aI3B in the coxI gene of S. douglasii has been shown to be dependent upon a nuclear gene expressed (or present) in that organism but not in S. cerevisiae (Kotylak et al., 1985). Different yeast strains contain various combinations of the known sixteen introns in their mitochondrial DNA. Standard laboratory yeast strains contain up to seven introns in the coxI gene (Bonitz et al., 1980; Hensgens et al., 1983), five introns in the cob gene (Nobrega et al., 1980; Lazowska et al., 1980), and one intron in the 21S rRNA gene (Dujon, 1980). Each of these introns has been assigned to one of two groups (group I or group II) according to the presence of conserved group-specific sequences and secondary structures that are essential for splicing (Michel and Dujon, 1983). The distribution of group I and group II introns in the yeast mitochondrial genome of our standard laboratory strain is shown in Figure 3.
I.B.5.a. Group I introns

Group I introns have been identified in a variety of organisms. Group I introns include the majority of fungal mtDNA introns, nuclear rRNA introns in Physarum and Tetrahymena, introns in bacteriophages T2 and T4 of *E. coli* and SPO1 of *B. subtilis*, and a minority of chloroplast introns of higher plants but quite a few for Chlamydomonas. Representative group I introns have been shown to self-splice *in vitro* by a guanosine-initiated transesterification mechanism first described for the Tetrahymena rRNA intron (reviewed by Cech and Bass., 1986). All group I introns share a set of conserved sequence elements that form part of a preserved secondary structure required for splicing. The isolation of cis-dominant splicing defective mutants and their compensatory second site suppressors revealed that these conserved sequence elements play an essential role in establishing a productive secondary structure for these introns (Weiss-Brummer et al., 1983; Holl et al., 1985; de la Salle et al., 1982; Haldi, 1987; Perea and Jacq, 1985). Four sequence elements (P, Q, R, and S) are conserved in sequence and are capable of base pairing; P with Q, and R with S (Davies et al., 1982). Elements E and E' are conserved in their position and ability to base pair with
one another. All group I introns have an internal guide sequence (IGS) usually located near the 5' exon/intron boundary whose function is to base pair with the 5' exon terminus. The role of the 3' exon pairing with the IGS has not yet been supported by data.

I.B.5.b. Group II introns

Group II introns, while not as wide-spread as group I introns, have been identified in the mitochondrial genomes of fungi and higher plants and in chloroplast genomes of higher and lower plants. Group II introns are also capable of self-splicing; however, that splicing mechanism involves transesterification reactions that result in the formation of an intron lariat and ligated exons (Peebles et al., 1986; Van der Veen et al., 1986). They share a different set of conserved secondary sequences from group I introns. Group II introns have the consensus GUGCG at their 5' boundary, and AU or AC at their 3' boundary (Keller and Michel, 1985). Their conserved "core" secondary structure consists of 6 helical domains (designated 1-6) radiating from a central wheel (Michel and Lang, 1983). Several cis-acting mutations and compensatory second site suppressors have been described that prove some of the proposed stem/loop structures of
these domains (Schmelzer et al., 1983, 1986; Carignani et al., 1983, 1986; S. Belcher, unpublished data).

I.B.6. Mitochondrial intron open reading frames

A unique feature of both group I and group II mitochondrial introns is that some members of each group contain long open reading frames usually in frame with the preceding exon (Nobrega et al., 1980; Bonitz et al., 1980). Early studies by Slonimski et al. (1980) demonstrated the existence of a class of mutants in the cytochrome b gene that is trans-recessive in complementation tests. That is, cob exon mutants were capable of complementing in trans, cob intron 2 mutants which were found to be stop codons in the intron open reading frame (Slonimski et al., 1978; Lamouroux et al., 1978). This indicated that the intron provides a functional trans-acting splicing factor (Lazouska et al., 1980). Other genetic studies have shown that several intron ORFs encode intron-specific proteins (maturases) which (usually) direct the splicing of the intron which encode them in vivo (Lazowska et al., 1980; Anziano et al., 1982; de la Salle et al., 1982; Mahler et al., 1982; Weiss-Brummer et al., 1982; Carignani et al., 1983; Mecklenburg, 1987).
I.B.6.a. The maturase model

The maturase model suggests a role for these trans-acting factors and proposes that a hybrid maturase protein is translated from the fused upstream exons and intron reading frame. The maturase facilitates the excision of the intron from which it was encoded and thus destroys its own message (Lazowska et al., 1980). Although the maturase model was originally proposed from genetic studies of intron 2 of the \text{cob} gene, intron-encoded maturases have been demonstrated to be required for splicing of \text{cob} introns 3 and 4 and \text{coxI} introns 1, 2, and under special conditions 4 (Weiss-Brummer et al., 1982; de la Salle et al., 1982; Anziano et al., 1982; Dujardin et al., 1982; Carignani et al., 1983).

More recently Banroques et al. (1986) demonstrated that a trans-recessive mutation in \text{cob} intron 4 could be rescued by nuclear transformation with a plasmid containing the \text{cob} intron 4 reading frame fused to a mitochondrial import signal. The precise biochemical role of maturases in splicing is unknown, but they likely stabilize the intron secondary structure required for splicing since many introns are capable of self-splicing \textit{in vitro}.
I.B.6.b. The box phenomenon

Genetic studies have shown that many intron mutations of the cob gene (cob-box) have a pleiotropic phenotype: they prevent not only the synthesis of cytochrome b but also that of subunit I of cytochrome oxidase. The block in gene expression results from a processing defect of both the cob and coxI mRNAs. The splicing of cob intron 4 and coxI intron 4 is dependent upon the cob I4 maturase (Dhwale et al., 1982; Weiss-Brummer et al., 1982; Banroques et al., 1987). CoxI intron 4 shares a striking sequence homology with cob I4 and has been shown to encode a latent maturase activity (Dujardin et al., 1982; Herbert et al., 1988).

The mitochondrial suppressor mutation, mim2-l, suppresses cob I4 trans-recessive mutants. Genetic and sequence analysis demonstrates that mim2-l suppression is due to a single G to A transition in the coxI I4 intron open reading frame. This substitution replaces a glutamic acid codon (in the wild type) with a lysine codon in the mim2-l suppressor. The mim2-l mutation induces a latent maturase activity in the product of the aI4a reading frame which can substitute for a deficient bI4 maturase (Dujardin et al., 1982).
A nuclear mutant, NAM2, can also induce the production of the aI4a-encoded latent maturase. NAM2 was isolated as a nuclear suppressor of a mutation that inactivates the bI4 maturase. By employing double mutants, the suppressor activity of NAM2 has been shown to be dependent on the integrity of the aI4a open reading frame (Dujardin et al., 1983). The NAM2 gene product is the mitochondrial leucyl tRNA synthetase (Herbert et al., 1988). A synthetase had previously been implicated in splicing by Akins and Lambowitz (1987) who demonstrated the mitochondrial tyrosyl tRNA synthetase from Neurospora crassa is involved in the splicing of some group I mitochondrial introns. Other nuclear genes of yeast have since been identified which are involved directly in the splicing of specific mitochondrial introns (Hill et al., 1985; McGraw and Tzagoloff, 1983; Faye and Simon, 1982, 1983; Pillar et al., 1983; Kreike et al., 1986).

I.C. Mitochondrial transmission genetics

I.C.1. Mitochondrial genetic crosses

A typical mitochondrial cross involves two homoplasmic, haploid yeast strains of opposite mating types, complementing nuclear auxotrophic markers, and
distinct mitochondrial genotypes. Recombination between markers in the yeast mitochondrial genome (such as resistance to drugs or polymorphic restriction enzyme sites) may be measured by analyzing the diploid progeny of a cross. Haploid yeast cells exist in either of two mating types, a or a. Both mating types synthesize mating pheromones which affect cells of the opposite mating type to allow fusion of a pair of cells to form a diploid zygote. Upon cell fusion, the cytoplasms mix rapidly to form a heteroplasm which allows the mitochondria to fuse permitting the 25 or so mitochondrial DNA molecules from each haploid cell to associate and recombine. MtDNA molecules in heteroplasmic cells (which contain parental and recombinant mitochondrial genomes) segregate rapidly during vegetative growth (mitosis). At each cell division a subset of the population of mitochondrial DNA molecules segregates into a bud so that after approximately 20 generations, nearly all of the diploid cells are homoplasmic. The diploid progeny of a cross may then be scored for their mitochondrial genotypes. Genetic markers, such as resistance or sensitivity to antibiotics, may be scored on differential media while physical markers may be scored by isolating mitochondrial DNA from a sample of the total progeny of the cross and employing restriction analysis (reviewed by Dujon, 1981).
I.C.2. Rules of mitochondrial inheritance

Most studies of mitochondrial genetics have utilized antibiotic resistant mutants distributed about the mitochondrial genome. Standard yeast strains are incapable of growth on non-fermentable carbon sources supplemented with drugs such as chloramphenicol, erythromycin, paromomycin, oligomycin, or antimycin. These drugs either interfere with mitochondrial translation or expression of mitochondrial genes (reviewed by Dujon, 1983).

Mitochondrial genetics can be treated in terms of population genetics at the intracellular level. Dujon and Slonimski first noted the similarity of mitochondrial crosses to bacteriophage crosses and developed a model using that analogy (Dujon 1974, 1976; Dujon et al., 1974, 1976; Dujon and Slonimski 1976). The model assumes that random pairing of mitochondrial DNA molecules occurs in the zygote and postulates that the output of a cross reflects the relative proportion of the two parental mtDNA molecules contributed by the parents of the zygote (the input ratio). The output ratio is variable from cross to cross, but is reproducible for a given cross. This suggests that input varies from strain to strain,
presumably due to nuclear genotype. The bias (output ratio) of a cross can be modified by pre-treatment of the parental cells with cycloheximide or α-factor or by varying the ploidy of one parent of the cross such that an unequal input of mitochondrial genomes occurs (Dujon et al., 1975; Goldthwaite et al., 1974; Birky, 1975; Boker et al., 1976; Dujardin et al., 1978; Gunge, 1976; Birky et al., 1978). Under such circumstances of biased input, uniparental zygotes are generated which transmit only mitochondrial alleles from the majority parent in the cross. Some uniparental zygotes are also seen in unbiased crosses, but much fewer (Birky et al., 1978). The observation of uniparental inheritance cannot be explained by the phage analogy model.

The model also proposes multiple rounds of random pairing and recombination among mtDNA molecules. The maximum frequency of recombination between two unlinked mitochondrial loci is 25%. Because pairing is random, recombination between identical genomes occurs, resulting in undetectable recombination events; hence, the low upper limit of 25% recombination.

Yeast mitochondrial DNA is highly recombinogenic; markers over 2000 bp apart cannot be ordered in these point crosses and so appear to be unlinked. The recombination frequency for unlinked nuclear genes is 50%.
According to the Phage Analogy Model, when a mitochondrial cross is biased, the frequency of recombination is reduced since the probability of pairing between the two genetically distinct mtDNA molecules is reduced. Furthermore, the frequency of double recombinants should greatly exceed that predicted by single recombinants when random pairing occurs (reviewed by Dujon, 1983). All of the above predictions have been confirmed by genetic studies (Dujon and Slonimski, 1976).

One aspect of the model has been disputed by mitochondrial segregation studies (Callen, 1974; Strausberg and Perlman, 1978; Waxman and Birky, 1982). A "panmictic" intracellular population of mtDNA molecules, where random pairing and recombination between different parental mtDNA molecules can ensue, may not exist. Slow cytoplasmic mixing in zygotes prevents a random distribution of mtDNA molecules from each parent. Pedigree studies conducted by Strausberg and Perlman (1978) and Waxman and Birky (1982) demonstrated incomplete cytoplasmic mixing prior to first bud formation. The majority of lateral buds dissected from zygotes are homoplasmic and have a lower frequency of recombinant genotypes suggesting only one type of mitochondrial genome is accessible. A much higher proportion of medial buds is heteroplasmic; such buds have a higher frequency
of recombinant genotypes suggesting that the central zone of a zygote is more likely to contain both types of mtDNA molecules. Furthermore, zygotes whose first buds are medial generate clones that have a higher recombination frequency than zygotes whose first buds are lateral (Waxman and Birky, 1982). Thus the low recombination frequency observed in mitochondrial genomes (25%) also reflects recombination of a subset of mtDNA molecules in the zygote. The rate of segregation is directly influenced by slow cytoplasmic mixing which occurs first in the central region of the zygote. Consequently, the relationship between bias and the recombination frequency in crosses may be altered. Apparently a large proportion of the mtDNA in the zygote enters each first bud so cytoplasmic partitioning likely plays an important role in recombination (Strausberg and Perlman, 1978).

Specific portions of the genome display non-reciprocal recombination. Unidirectional gene conversion of defined alleles of the 21S rRNA and var1 genes results in the preferential transmission of these alleles to the diploid progeny. These specialized recombination phenomena are now fairly well understood and will be addressed in detail in later sections.

I.D. Recombination events in the mitochondrial genome
involving GC clusters

I.D.1. Petite formation

The mitochondrial genome is highly active in recombination. One of the classes of optional mtDNA sequences, GC clusters, has been implicated in several types of recombination events. Petite mutations form by intramolecular recombination crossovers that require only short stretches of sequence homology. Studies to characterize the endpoints of petite deletion mutants indicate that intramolecular recombination between short repeats of the wild-type genome is responsible for the formation of most petite mutants. Some mitochondrial genomes of spontaneous petites have been shown to arise by recombination via excision between GC clusters of the intergenic A+T spacer regions (de Zamaroczy et al., 1983). Other petite mutants have been characterized that are believed to arise by abortive replication since their repeat units are too short to be generated by intramolecular circularization and excision (reviewed by Dujon and Belcour, 1989). Petite mutant genomes consist of varying sizes of subgenomic fragments of yeast mtDNA that are reiterated in either a "head-to-head" or "head-to-tail" fashion to create large molecules where the
retained sequence is repeated. Inverted-repeat petite molecules are believed to occur by intrastrand pairing and recombination at short inverted repeats, while the repeated unit in direct repeat petite molecules is flanked by short direct repeats. Petites usually retain a small region of the genome, but are respiratory deficient due to the deletion of genes required for respiration or mitochondrial translation (Morimoto et al., 1975; Kreigl and Bos, 1977). Petite deletions occur at a high (but variable) spontaneous frequency; usually 1-2% petite mutants per cell per generation. Ethidium bromide is commonly used to enhance the rate of petite formation, although a variety of other agents are also efficient (Slonimski et al., 1968; Dujon 1981).

I.D.2. Petite integration

Preferential recombination between GC clusters has been documented where integration of a petite genome into a wild-type mitochondrial genome occurs. Studies designed to analyze homologous recombination between wild-type and petite genomes employ heteroplasmic strains. Some petites which arise in strains that are respiratory deficient allow the strain to grow on non-fermentable carbon sources because they create specific gene rearrangements which are
PLEASE NOTE:

Page(s) not included with original material and unavailable from author or university. Filmed as received.
1985; Weiller, 1987). However, for some families of G+C-rich clusters, the flanking sequences are conserved, suggesting that unique target sites exist for these clusters. This latter feature of a subset of GC-rich clusters is not typical of transposons.

I.E. The varl gene

All forms of the varl gene in strains of Saccharomyces cerevisiae (shown schematically in Figure 4) contain a 46 bp palindromic GC cluster termed the common cluster (cc). It is located 186 bp downstream of the start of the varl ORF. The "a" insert is identical in sequence to the common cluster but is in an inverted orientation 160 bp downstream of the common cluster. The "a" insert is found in combination with the common cluster in some strains, but no strains have been reported that contain only the "a" GC-rich element. Both elements are flanked by two additional A residues such that the insert is in-frame for translation. Both the "a" insert and common cluster are flanked by identical pure A+T-rich sequences; 9 bp on one side and 11 bp on the other. An identical GC cluster is found between the olil and serine tRNA genes. The A+T-rich sequence flanking this cluster is identical as well, but is inverted relative to the orientation of those in
the \textit{varl} gene. Other insertions, in locations bl and b2 (see Figure 4), consist of varying numbers of asparagine codons inserted into strings of asparagine (AAT) codons. The bl insertion contains four asparagine residues while the b2 insert extends the reading frame by six additional asparagine codons. "Partial" bl elements, or blp inserts, have been described which add two asparagine codons at the same site as bl elements (Hudspeth et al., 1984). The b elements are found in various combinations either alone or along with the GC cluster elements described above.

I.F. Transmission of the "a" GC-rich element in the \textit{varl} gene

In crosses between strains that differ in their "a" and "b" \textit{varl} determinants, progeny arise that synthesize nonparental (recombinant) sizes of \textit{varl} protein (Strausberg et al., 1978). For instance in crosses of the type "a\textsuperscript{+}b\textsuperscript{+}" x "a\textsuperscript{-}b\textsuperscript{-}" 25-30\% of the progeny were found to be the nonparental recombinant "a\textsuperscript{+}b\textsuperscript{-}". This transfer of the "a" element formally resembles asymmetric gene conversion since the optional GC cluster is preferentially inserted into \textit{varl} alleles that lack the cluster. GC cluster transmission appears to be a site-specific process since strains that contain only the common cluster never
acquire the cluster at the "a" or olil target sites in vegetatively growing cells.

I.G. The polarity phenomenon in the 21S rRNA gene

Specialized recombination is not restricted to GC clusters in the mitochondrial genome. Studies of the optional omega (w) intron in the 21S rRNA gene revealed that in crosses between strains that have (w⁺) and lack (w⁻) the intron, virtually all of the progeny are w⁺. The quantitative conversion of the omega intron is accompanied by conversion of flanking DNA sequences, the so-called polarity of conversion. Several mutations in the exon sequences of the 21S rRNA gene confer resistance to inhibitors of mitochondrial translation. These point mutations and some small non-recombinogenic GC-rich clusters in the exon sequences serve as flanking markers and show a gradient of conversion such that the frequency of flanking marker co-conversion varies inversely with their distance from the omega intron (Dujon et al., 1986; Zinn and Butow, 1985). Several mutants have been isolated from w⁻ strains that are deficient for the "polarity of recombination" observed at the 21S rRNA gene. Such mutants have been termed omega neutral (wⁿ), and result from single base pair substitutions within 3 bp flanking
the intron insertion site (Jacquier and Dujon, 1985).

I.G.1. A site-specific intron-encoded endonuclease

The 1143 bp omega intron contains a free-standing ORF that can encode a protein of 235 amino acids. Genetic studies have demonstrated that the intron-encoded protein product is required for conversion of \( w^- \) to \( w^+ \) alleles in crosses (Macreadie et al., 1985; Jacquier and Dujon, 1985). In zygotes of crosses between \( w^+ \) and \( w^- \) strains a transient double-strand break is detected at the site of intron insertion which may initiate conversion of the intron. This double-strand cut in \( w^- \) mtDNA occurs so long as one parental strain is capable of mitochondrial translation. Mutations in the intron open reading frame that alter the protein product abolish \( w^+ \) conversion and the appearance of the double-strand break in \( w^- \) mtDNA during mating (Macreadie et al., 1985; Jacquier and Dujon, 1985). The \( w^+ \) ORF was later shown to encode a site-specific endonuclease (fitl) that recognizes and cleaves the sequence at the intron insertion site \textit{in vitro} (Jacquier and Dujon, 1986). Since mitochondrial ORF products are synthesized in such small amounts, a universal code equivalent of the omega ORF was engineered in \textit{E. coli} and shown to direct the specific double-strand
cut in cloned DNA containing the \( w^- \) site (Colleaux et al., 1986).

I.G.2. The \( \text{fit}1 \) substrate

Employing the omega protein synthesized in \textit{E. coli} in \textit{in vitro} assays, Colleaux et al., (1986) demonstrated that the double-strand break produces a 4 bp staggered cut leaving free 3' hydroxyl overhangs. Detailed characterization of the recognition and cleavage site by site directed mutagenesis has shown that it is situated asymmetrically over 18 bp that spans the intron insertion site (Colleaux et al., 1988). The sequence constraints of the 18 bp substrate are quite stringent since at least one substitution at each of the 18 bp interferes with cleavage of that substrate (Colleaux et al., 1988).

I.H. Double strand breaks as initiators of recombination

Double strand breaks have been shown to be highly recombinogenic in yeast. The finding that experimentally introduced double strand breaks efficiently target integration of hybrid (\textit{E. coli}/yeast) plasmids into the yeast genome supported the view that double strand breaks are repaired by a recombination mechanism (Orr-Weaver and
Szostak, 1983). Although transformants can be obtained with closed circular plasmid DNA, cutting a plasmid in a region of homology with the recipient genome greatly enhances the efficiency with which the plasmid becomes integrated (Orr-Weaver et al., 1981). All integrations occur at the genomic locus homologous with the region of the plasmid containing the double strand break (even when another region of homology is contained on the plasmid) indicating that DNA ends interact directly with homologous sequences. When gaps are created within the yeast sequences of the transforming plasmid, they are efficiently repaired using homologous genomic sequences as a template during plasmid integration. Because repair of a double strand gap involves the transfer of genetic information from one DNA duplex to another, conversion (the non-reciprocal transfer of information from one DNA duplex or another) occurs with gap repair (Orr-Weaver and Szostak, 1983).

Although several classical models have been proposed to account for the recombination events seen in meiotic studies (Holliday, 1964; Meselson and Radding, 1975), data gathered from the studies of yeast transformation described above led to the proposal of the double-strand-break-repair model for recombination.

The critical features of the double-strand-break-
repair model are shown in Figure 5. A double-strand cleavage in the recipient DNA duplex by a double-strand endonuclease initiates the recombination process. As the break is enlarged, (5'-3') exonucleases generate 3' single-strand termini which can invade the homologous region of the donor duplex to act as a primer. Repair synthesis ensues using the donor duplex as a template. Another round of repair synthesis is initiated from the other 3' end using the displaced strand as a template. The model accounts for the observed properties of meiotic recombination and plasmid integration: heteroduplex formation in flanking regions, double strand breaks as initiators of genetic rearrangements, and gene conversion of flanking DNA sequences. According to the model, conversion is the result of repair of the gap by copying from both strands of the homologous duplex. Thus, polarity results from variable extents of gapping; the chance that a marker will be included in the gap decreases with increasing distance from the site of the double strand cut.

Double strand cuts occur naturally in the process of mating type interconversion and at the ARG4 locus (Strathern et al., 1982; Sun et al., 1989; Nichols et al., 1989). Homothallic (HO) yeast strains switch from one mating type to another (a or ) by a recombination event
that introduces either a or information from silent loci into an expression locus termed MAT. This interconversion event involves a site-specific endonuclease (the product of the HO gene) that creates a double strand cleavage near the Y/Z junction of the MAT locus. This double strand cut initiates the recombination process which allows mating type switching to occur.

The ARG4 locus is also known to engage in conversion. Flanking DNA sequences show a gradient of conversion from one end of ARG4 to the other. This polarity phenomenon suggests that there is an initiation site for ARG4 conversion, and the probability of flanking sequences being co-converted decreases with increasing distance from this site. Studies by Nichols et al., (1989) approximated the location of the ARG4 initiation site to within 270 bp by constructing strains deleted for DNA segments that show a high frequency of co-conversion. Furthermore, Sun et al., (1989) demonstrated that a double strand break is generated in the ARG4 initiation region contained on a plasmid early in meiosis. Plasmids with the ARG4 initiation region deleted do not experience a double strand cleavage.

I.I. Reverse transcription and mitochondrial introns
Intron transfer by site-specific recombination need not be the only mechanism by which stable intron transfer can occur. For instance, Woodson and Cech (1989) have demonstrated that self-splicing is reversible, and that introns might insert themselves into novel exonic environments. Intron movement by reverse splicing requires that a DNA copy of the newly synthesized RNA be generated (presumably by reverse transcription) and incorporated into the genome by homologous recombination.

Regions of most group II mitochondrial intron ORFs of various fungi share a distinct homology to retroviral reverse transcriptase (Michel and Lang, 1985; Xiong and Eickbush, 1988). The hypothesis supporting that these ORFs produce proteins with reverse transcriptase activity are genetic studies which have demonstrated that some mutations in group I and group II introns are capable of reverting to growth on non-fermentable media by precise deletion of the impaired intron at a relatively high frequency (Jacq et al., 1982; Gargouri et al., 1983; Perea and Jacq, 1985; Hill et al., 1985; Seraphin et al., 1987). The generation of this class of revertants suggests recombination with a cDNA copy of mature message.

I.I.I. Some modes of reversion of mitochondrial intron mutants
I.I.1.a. Reversion by intron excision

A rare class of "pseudo-wild type" revertants of a cis-dominant mitochondrial mutation located in the 3' portion of intron bI4 (mutant G2590) were the first intron excision revertants reported. Restriction mapping of such revertants revealed that they contained a deletion of roughly 2000 bp. Electron microscopic analysis of heteroduplexes of mtDNA obtained from the wild type and deleted revertants revealed two loops which corresponded to introns bI4 and bI5. Further restriction mapping was consistent with the clean excision of bI4 and bI5 along with maintenance of the reading frame of the cob gene. The so-called "pseudo-wild type" phenotype of the revertants resulted from the simultaneous appearance of a NAM nuclear suppressor (discussed previously) to compensate for the loss of the bI4 maturase which facilitates splicing of coxI intron aI4a (Jacq et al., 1982).

Genetic studies aimed at studying the splicing process of bI4, an mRNA maturase-encoding intron, employed mutant G2457 which consists of a G to A transition located 2 bp upstream from the 5' intron/exon junction. The base change presumably affects the internal guide sequence
interaction since proteins accumulate in the mutant which are typical of mutants blocked in the splicing of b14. Molecular analysis conducted on revertants capable of growing on non-fermentable media revealed (among several classes of revertants) two groups of revertants where cytochrome b introns have been deleted. One revertant class had only cob introns 4 and 5 deleted, while the other class had all the introns in the cob gene deleted. In both cases the introns appear to be cleanly excised. Since deletion of cob intron 4 should prevent the splicing of the fourth intron of coxI, this class of revertants has also simultaneously acquired a nuclear NAM2 mutation (Perea and Jacq, 1985).

Studies designed to detect intron-DNA deletions employed a cis-dominant, RNA splicing-deficient mutation in the ORF of cytochrome b intron 2. In revertants examined for the presence of cytochrome b introns, two classes were observed; those that had lost two introns (bI1 and bI2) and those that had lost three introns (bI1, bI2, and bI3). Multiple deletions appear to be more frequent than single deletions in these studies. This "pop-out" phenomenon is quite general as such revertants have been characterized from several splicing deficient mutants (both cis-dominant and trans-recessive) in the first, second, and third introns of the cytochrome b gene.
The frequency of revertants with "processed genes" occurs between $0.2 \times 10^{-8}$ to $2 \times 10^{-8}$ per cell, which is typical for the frequency of some true back mutations or suppressors (Gargouri et al., 1983).

A trans-recessive mutant of intron aII in the coxI gene (C1036), has also been shown to revert by intron excision (P. Q. Anziano, unpublished data). The C1036 mutation consists of six single site changes in the aII intron open reading frame; four are missense mutations, one is a silent mutation, and one truncates the intron-encoded protein. Three classes of glycerol growing revertants were characterized; those deleted for aII alone, those deleted for both aII and aI2, and those deleted for aII, aI2, and aI3. The introns in each class of revertant were shown to be cleanly excised by sequence analysis.

I.I.I.l.b. Other modes of reversion of intron mutations

Pseudorevertants of other mitochondrial point mutations which impair the maturases encoded by the bI4 and aII intron open reading frames have been characterized and shown to be heteroplasmic cells which contain two types of mitochondrial genomes, both of which are deficient. In each case, the trans-acting mutation is suppressed by the expression of a hybrid intron reading
frame generated by a recombination event to create a "fusion-maturase" between two closely related intron encoded proteins. In the former case, a petite genome having a bI4/aI4 fusion intron is generated by intramolecular recombination events involving various pairs of repeated sequences within those introns. Suppression of the all maturase mutation involves the formation of a mit− genome containing an aI1/aI2 fusion intron. This recombination mechanism likewise uses short conserved sequences in each intron to generate the deleted mit− genome. Such pseudorevertants can be maintained so long as selective pressure is imposed by maintaining them on media containing non-fermentable carbon sources.

I.I.2. Reverse transcriptase activity in other fungal mitochondria

I.I.2.a. Senescence in Podospora anserina

Other evidence for a relationship between reverse transcription and intron excision is found in Podospora anserina. Senescence (ageing) is associated with the propagation of a mitochondrial plasmid (a-senDNA) that is a precisely excised group II intron from the cytochrome oxidase subunit I gene (Osiewacz and Esser, 1984). The
reading frame from this intron encodes a protein similar to reverse transcriptase. Nuclear/mitochondrial extracts prepared from middle aged cultures of *Podospora anserina* contain a DNA-polymerase activity with characteristics of reverse transcriptase which can not be detected in young mycelia (Steinhilber and Cummings, 1986). Thus, the enzyme might be involved in the process of senescence.

I.I.2.b. Neurospora mitochondrial plasmids

The Mauriceville and Varkud mitochondrial plasmids of *Neurospora crassa* contain sequence elements characteristic of group I introns (Nargang et al., 1984; Akins et al., 1988). The plasmids encode a protein that contains short blocks of amino acids which are typical of reverse transcriptases. Recently, a reverse transcriptase activity associated with the Mauriceville and Varkud mitochondrial plasmid carrying strains has been characterized in mitochondrial lysates and ribonucleoprotein (RNP) preparations from these strains. This reverse transcriptase activity is highly specific for endogenous plasmid RNA in RNP preparations and transcribes full-length minus-strand DNAs starting at the 3' end of the plasmid transcript. These plasmids likely employ a mechanism of reverse transcription which utilizes tRNA-
like structures at the 3' ends of the plasmid transcripts (Kuiper and Lambowitz, 1988). The plasmids may represent retro-elements that are confined to an autonomous DNA form.

I.J. Issues explored in this dissertation

The research described in this dissertation addresses several aspects of recombination of the mitochondrial genome, but concentrates on two major, overlapping themes: gene conversion events and intron mobility. The original goal of my dissertation research was to make a comprehensive comparison between GC cluster conversion and intron mobility in crosses. The processes of GC cluster and intron conversion had appeared similar, since they both employ unidirectional gene conversion. The discovery of new var1 alleles, and the identification of new polymorphic DNA markers, enabled me to determine that GC cluster conversion proceeds by a mechanism that does not involve flanking marker co-conversion. GC cluster conversion was found to be independent both of the orientation of the cluster, as well as mitochondrial translation. Thus GC cluster conversion and intron mobility are distinct processes and the experiments in the following sections illustrate that they are clearly
directed by different molecular mechanisms.

Since only one intron in the mitochondrial genome was known to be capable of conversion, other optional group I introns were assayed for mobility. Several optional introns in the coxI gene were tested for their ability to engage in gene conversion. These experiments allowed me to validate the features of intron conversion that were first uncovered with studies of the omega intron in the 21S rRNA gene. Intron aI4a was found to be efficiently transmitted to genes lacking it. Mutations that truncate the intron-encoded protein abolish intron transmission demonstrating a function for the intron product in conversion. The intron was found to encode a site-specific endonuclease which recognizes defined exon sequences and cleaves them, thus initiating the process of gene conversion. Since the endonuclease activity has been characterized by in vitro studies using mitochondrial extracts, the recognition sequence for this enzyme has been dissected using site-directed mutagenesis.

In keeping with the notion of intron mobility, other mechanisms have been investigated. Mutants defective in splicing mitochondrial introns have been shown to be capable of reverting by clean excision of the impaired intron (often together with adjacent introns) implying recombination with a cDNA of mature message. The reading
frames of group II introns aI1 and aI2 contain a distinct homology to reverse transcriptase. Special mitochondrial genome constructs were generated to examine the dependence of reversion by intron excision on the presence of introns aI1 and aI2. I found that there is a distinct relationship between the presence of introns aI1 and aI2 and the capacity to revert via intron excision. Intron aI1 was found to be sufficient to confer the ability to revert by this "pop-out" mechanism. These two opposing intron-encoded functions that have been examined in this dissertation, insertion and excision, likely exist in a dynamic equilibrium to generate the diverse intron configurations of mitochondrial genes observed.

Interestingly, the protein products of introns aI1, aI2, and aI4a are also known to function as maturases that facilitate splicing. The dual function nature of these proteins, and the implications of intron mobility as they relate to evolution are discussed.
CHAPTER II
MATERIALS AND METHODS

II.A. Strains

II.A.1. Yeast strains: Table 1 lists the yeast strains used in this dissertation. Strains are listed according to their nuclear and mitochondrial genomes.

II.A.2. Recombinant plasmids: Table 2 lists the recombinant plasmids used in this dissertation. All clones were maintained in host strains DH5αF', JM101, JM103, CJ236, or HB101.

II.B. Culture Media

II.B.1. Yeast: Solid media is prepared by adding 2% agar.

II.B.1.a. Non-selective growth media contains 1% yeast extract, 1% bactopeptone, and 2% glucose (YPD), 10% glucose (10% YPD), or 2% galactose (Y-gal).
II.B.1.b. Yeast glycerol media contains 1% yeast extract, 1% bactopeptone, and 2% glycerol (YG).

II.B.1.c. Drug resistance markers of various strains were scored by replica plating to YG medium containing 3 mg/ml of chloramphenicol, 1 mg/ml of erythromycin, or 2 μg/ml of oligomycin. Drugs are dissolved in ethanol and added to cooled media.

II.B.1.d. Yeast minimal medium contains 0.67% yeast nitrogen base and 2% glucose (MD). Yeast auxotrophic markers are accommodated by adding 75 mg/l of the appropriate amino acid or growth factor.

II.B.1.e. Sporulation of diploids is induced on solid pre-sporulation medium containing 2% potassium acetate, 0.25% yeast extract, and 0.1% glucose. Cells are then transferred to liquid sporulation medium which contains 2% potassium acetate.

II.B.2. E. coli: Solid medium is prepared by adding 1.5% agar. Rich medium contains 1% bactotryptone, 0.5% yeast extract, and 0.5% NaCl (YT). The pH of YT media is adjusted to 7.2. Supplements are added after cooling: 100 mg/l of ampicillin and/or 30 mg/l of chloramphenicol.
Matings for GC cluster conversion and intron transmission experiments were done in patches on solid YPD medium. Approximately $5 \times 10^7$ cells of each mating type were patched together on solid YPD medium. A very low estimate of the frequency of mating (1%) would indicate that no fewer than $1 \times 10^6$ zygotes would result from such a cross. Patched matings were suspended in minimal medium and grown to stationary phase to expand the population of diploid zygotes. The overnight culture was used to inoculate another minimal medium culture which was also grown to stationary phase. An aliquot of these pooled diploids was expanded in liquid Y-gal medium for labeling of mitochondrial translation products or in YPD for isolating mitochondrial DNA. This type of mating ensures that a substantial number of zygotes give rise to the recombination events examined genetically and biochemically. In the experiments described in the following sections, cells examined from expansion of the second cycle of minimal selection are referred to as "pooled diploid progeny". For the *in vivo* double strand cleavage experiments strains were grown to late logarithmic phase. Cultures were mixed and mated
synchronously by the method of Sena (1982). Mating mixtures were examined under a microscope to ensure a substantial number (at least 20%) of zygotes were present.

II.C.2. Petite Deletion Mapping

Our laboratory maintains both a set of well characterized \textit{mit-} mutations (mitochondrial point mutations or small deletions), and a set of petite mutants which retain a given segment of the mitochondrial genome. Strains to be tested were suspended in sterile water in microtiter wells and stamped onto confluent lawns of testers on YPD plates. Strains were permitted to mate for 2 days on YPD medium incubating at 30°C and were then replica plated to both minimal media (supplemented with the appropriate amino acid if required) to select for diploids, and YG to select for respiration competent cells. Replica plates were allowed to grow for at least 24 hours at 30°C before being scored for mating and growth on glycerol media.

II.C.3. Sporulation of Diploids

Sporulation of diploids was stimulated by growth on solid presporulation medium for 2 days at 30°C. Cells
were then transferred to liquid sporulation medium and incubated with shaking for 3 days at 30°C. The cell suspension was monitored with a light microscope for the presence of 4-spored asci (tetrads). Tetrads were digested in 10% glusulase and random spores were isolated following sonication. Haploid spores were scored for auxotrophic markers on amino acid supplemented MD media. The mating type of each spore was determined by its ability to grow on MD media after crossing them to strains of known mating type.

II.C.4. Karl-mediated cytoduction

Transfer of mitochondrial genomes from one nuclear background to another was accomplished by employing the karl dominant nuclear mutation, which inhibits fusion of nuclei in zygotes (Conde and Fink, 1976). In a typical "kar switch" the recipient nuclear background is a rho-zero derivative and either the mitochondrial donor or the recipient carries the karl mutation. Strains are mated in patches for 8-12 hours on solid YPD media and subcloned onto minimal medium supplemented with appropriate growth factors to select for the desired nucleus. Cytoductants are distinguished from rho-zero parents and prototrophic diploids by growth on various media and/or by matings to proper tester strains.
II.C.5. Protoplast Fusion

The mitochondrial genomes of various homothallic diploid yeast species of the genus *Saccharomyces* were transferred into rho-zero haploid nuclear backgrounds using a modification of the protoplast fusion procedure of Sipiezki and Ferenczy (1977). Spheroplasts of each strain were fused with 40% PEG and allowed to regenerate in solid YPD with 1.2 M sorbitol added. Fusion products were obtained by screening for auxotrophic markers and counter-selecting on YG medium. If the recipient rho-zero nuclear background carries an ade2- nuclear auxotrophy, respiratory competent fusion products will appear as red colonies on YPD media and may be identified by direct visual inspection.

II.C.6. Ethidium Bromide Mutagenesis/Petite Induction

Ethidium bromide (EB) mutagenesis was used to create petite and rho-zero derivatives of rho+ strains. Rho-zero derivatives were generated by submitting the cells to three passages of growth in YPD medium supplemented with 50 mg/ml of ethidium bromide. Subclones were then mated to various mit- tester strains to ensure the candidate rho-zero isolates are incapable of restoring respiratory
competency to known deletion mutants.

Strains submitted to petite mutagenesis were innoculated to an optical density (O.D.) of 0.2 (at 600nm) and grown 2 hours to allow the cells to reach a logarithmic growth phase. About 0.2 mls of log phase cells were added to 2 mls of YPD with 250 μl of a solution containing 0.4 mg/ml of EB. Cells were incubated at 30°C and at various time points up to 8 hours aliquots were removed and used to innoculate fresh YPD cultures to allow the petite genomes to segregate. Each culture representing a time point of the mutagenesis was then subcloned on YPD medium to generate single colonies. The desired petites were identified by crossing candidates to appropriate tester strains. Often a particular time point is enriched for the petites of choice depending on the strain employed in the mutagenesis.

II.C.7. TTC Overlay

The suppressivity of petites derived by ethidium bromide mutagenesis was determined by mating them to a rho+ tester strain on solid YPD medium overnight, selecting for diploids by subcloning on minimal medium, and performing the TTC overlay procedure. 2, 3, 5 triphenol tetrazolium chloride (TTC) is converted to an insoluble red formazan in the presence of respiring cells.
TTC buffer containing 1.0% TTC in 0.5 M KH$_2$PO$_4$, pH 7.0 (25 ml) is added to 100 ml of cooled 2.0% agar and poured onto a minimal plate containing colonies to be tested. The distinctive red color indicative of respiratory positive cells appears in 15-30 minutes.

II.C.8. MnCl$_2$ Mutagenesis and Adriamycin Suicide Treatment

Mitochondrial mutations (mit- mutants) were induced by MnCl$_2$ mutagenesis. Cells to be mutagenized are grown to stationary phase and used to inoculate fresh YPD medium to allow cells to enter a logarithmic phase of growth. After 2 hours at 30°C, MnCl$_2$ was added to a final concentration of 20 mM and permitted to mutagenize for 5 hours. Cells were washed once in 10 mM MgCl$_2$ and once in sterile water and a dilution series of an aliquot is plated as a 5 hour sample on 10% YPD. The rest of the sample was allowed to mutagenize overnight at 30°C and the cells were then washed and various dilutions plated as before.

An aliquot of the mutagenized cells was resuspended in YG medium and allowed to grow at 30°C for 30 minutes. This allows cells to respire, but not divide. Adriamycin (Adria Doxorubicin-HCl) was added to a final concentration of 50 μg/ml. The cells were incubated at 30°C and at
various time points aliquots were removed to plate a dilution series on 10% YPD media. Adriamycin selectively kills respiring cells. The desired mit- mutations were identified by mating them to a series of tester strains.

II.C.9. Selection of Revertants

Various revertants of mit- mutations were selected by plating a lawn of roughly $10^9$ cells on YG solid medium and incubating them for 10-14 days at 30°C. Typically, temperature resistant (TR) revertants appear first as large colonies on the lawn while temperature sensitive (TS) revertants are smaller and slower growing. If the mit- mutation harbors an ade- nuclear auxotrophy, revertants may be isolated by plating a lawn of $10^9$ cells on YPD solid media. Revertants appear as red colonies after about 14 days. The specific phenotypes of revertants were characterized by testing subclones for growth at different temperatures on various media.

II.C.10. Strain Constructions

A set of strains having specific mitochondrial configurations was constructed to be used in a screen to isolate nuclear pet- mutants that affect either the splicing of coxI intron 5g, the splicing of all group II
introns, or the splicing of all mitochondrial introns. Although this project is in progress (by a different researcher) these strains have been employed in studies of intron conversion in the mitochondrial genome and their construction will thus be described here.

The goal of the construction was to obtain a haploid yeast strain retaining cox1 I5g as the only group II intron in its mitochondrial genome. Strain D273-10B (Table 3, line 4) contains the "cob short" form of the cytochrome b gene: that is, it has only bI4 and bI5 and is missing bI1 (the only group II intron in the gene), bI2, and bI3. This strain also has the "oxi short" form of the cytochrome oxidase subunit I gene: it has introns aI1, aI2, aI3, aI4, and aI5g.

The first necessary step in the construction of a strain containing only one group II intron was the removal of the w+ intron from the 21S rRNA gene. Previous studies have shown that syn- or pet- mutants affecting mitochondrial translation have unstable mitochondrial genomes (Sherman, 1963). Thus, a pet- mutant that affects the splicing of all introns would disrupt mitochondrial translation since it would be incapable of splicing the w+ intron. The elimination of the w+ intron was accomplished by crossing the D273-10B strain (Table 3, line 4) to a strain containing the w- allele of the 21S rRNA gene (Table 3, line 1). A chloramphenicol marker present in
the 21S rRNA gene of the \textit{w}^- allele facilitated selection for recombinants now lacking the \textit{w}^+ intron. Initial screens for these recombinants were done with colony hybridizations and confirmation of the absence of the \textit{w}^+ intron was accomplished by Southern analysis (Southern, 1975). The D273-10B/\textit{w}^- strain (Table 1) contains three group II introns; \textit{aII}, \textit{aI2}, and \textit{aII}g. The removal of introns \textit{aII}, and \textit{aI2} took advantage of the earlier finding that the \textit{coxi} gene of \textit{S. capensis} (Table 4, line 2) lacks \textit{aII} and \textit{aI2} (Ralph, 1986). Petite derivatives of \textit{S. capensis} were induced by EB mutagenesis and those retaining the \textit{coxi} exon 1-3 portion of the genome were indentified by standard petite deletion mapping. They were subsequently crossed to the D273-10B/\textit{w}^- derivative and respiratory sufficient recombinants missing \textit{aII} and \textit{aI2} were selected by screening colony hybridizations with intron specific probes following enrichment for recombinants by their acquisition of auxotrophic markers carried by the petites. The structure of this derivative (D273-10B/GII-5g) (Table 3, line 5) was confirmed by Southern analysis with a variety of probes, some of which were \textit{aII}, \textit{aI2}, and \textit{aI3} specific. The \textit{coxi} gene of this strain contains \textit{aII}g as the only group II intron as well as group I introns \textit{aI3a} and \textit{aI3g}.

Since the original objective was to obtain nuclear mutants which affect the splicing of at least group II
introns, the desired mutants can be detected by substituting the mitochondrial genome in each pet-
candidate with a rho+ genome that contains no group II introns. Pet- mutants that are respiratory deficient with a mitochondrial genome containing one group II intron but can be rescued by a mitochondrial genome that contains no group II introns (via karl-mediated cytoduction) can be identified. This screen depends on a mitochondrial genome which lacks aI5g. S. ellipsoideus (Table 4, line 3) lacks intron aI5g but has other group II introns. Petite mutants were generated from S. ellipsoideus and mapped as before to the exon 5-6 region. The previously described strain having only one group II intron (D273-10B/GII-5g) was mutagenized with MnCl2 followed by an adriamycin suicide enrichment to generate a mit- in the aI5g region. Mutants mapped to this region were crossed with the S. ellipsoideus petite and respiratory sufficient recombinants now having the "group II-less" mitochondrial configuration were identified by Southern analysis (Table 3, line 6).

This set of strains was among several others employed in studies to identify the source of the mitochondrial endonuclease that participates in intron (aI4a) transmission. The construction of strains used in studies of intron excision will be described in the Results section.
II.D. Colony Hybridizations

Yeast colony hybridizations were performed as described by Zinn and Butow (1985). Bacterial colony hybridizations were performed according to Grunstein and Hogness (1975).

II.E. Mitochondrial DNA Isolation

Crude mitochondrial DNA was isolated from small scale yeast cultures essentially by the minilysate procedure of Jacquier and Dujon (1985). Purified mitochondrial DNA was prepared by the procedure of Bingham and Nagley (1983).

II.F. DNA Restriction Site Analysis

Restriction enzyme digestions were performed as recommended by the enzyme supplier (Bethesda Research Laboratories). Gel electrophoresis was conducted as described by Maniatis et al. (1982).

II.G. Southern Blot Analysis

Restriction fragments separated by gel electrophoresis were transferred by the method of Southern
(1975) or electrophoretically to Hybond-N membranes (Amersham) or to nitrocellulose filters. Membranes or filters being hybridized to oligonucleotide probes were pre-hybridized in 6 X SSC containing 10 X Denhardt's, 0.5% SDS, and 0.5 mg/ml of denatured calf thymus DNA at 65°C. Oligonucleotide probes were allowed to hybridize to membranes for a minimum of 2 hours at 42°C. Blots were washed 3 times for 15 minutes each at 50°C in 6 X SSC.

Hybridization using other probes was performed in a solution containing 50% deionized formamide, 25 mM Na2PO4; pH 6.5, 5 X SSC, 1.0% SDS, 0.1% Ficoll, 0.1% polyvinylpyrrolidone, 0.1% BSA, and 1.0 mg/ml of denatured calf thymus DNA. Denatured probes were hybridized overnight at 42°C and washed twice each for 15 minutes at either 25°C or 42°C in 2 X SSC with 0.1% SDS added and .2 X SSC with 0.1% SDS added. All filters were exposed to Kodak XAR-5 X-ray film.

Oligonucleotides were end-labeled with 32P-ATP using polynucleotide kinase according to Maniatis et al. (1982). Other probes were generated by the Multiprime DNA Labeling System (Amersham).

II.H. Cloning of Mitochondrial Restriction Fragments

Mitochondrial DNA restriction fragments were isolated from preparative "low-melt" agarose gels (FMC) by melting
the agarose slab at 65°C for 10 minutes, phenol extracting, and ethanol precipitating the isolated DNA fragment. Vector DNA was cleaved with restriction enzymes, phenol extracted, and ethanol precipitated. Ligations were performed according to Maniatis et al. (1982) and transformation into various host strains was accomplished by CaCl₂ transformation described by Mandel and Higa (1970). Various clones were subcloned into M13 vectors for sequence analysis as instructed in the BRL M13 Cloning and Dideoxy Sequencing Manual.

II.I. Plasmid Preparation

Plasmid DNA was prepared on a small scale according to the method of Holmes and Quigley (1981). Large scale plasmid preparations were conducted according to Birnboim and Doly (1979).

II.J. Dideoxy Sequencing

Preparation of M13 single stranded template and dideoxy sequencing was performed according to the BRL Cloning and Dideoxy Sequencing Manual. The Sequenase Sequencing Strategy (United States Biochemical Corporation) was employed for sequencing some single and double stranded templates.
II.K. Site-Directed Mutagenesis

Oligonucleotide-directed mutagenesis was performed according to the method of Kunkel et al. (1986). Single stranded DNA from various plasmids to be mutagenized was isolated following transformation into *E. coli* strain CJ236 (dut<sup>-</sup>, ung<sup>-</sup>). After annealing the mutant oligonucleotide to the DNA, second strand synthesis was performed according to Ti-Zi Su and El-Genely (1988). The reaction was then transformed into competent *E. coli* JM101 or DH5αF<sup>+</sup> cells. Transformants were identified by dideoxy sequencing.

II.L. Random *in vitro* mutagenesis

A mutagenic oligonucleotide was generated using mixes adjusted to yield a substitution probability of 0.1 per base at each position of the 18 bp to be substituted. This would generate a Poisson distribution of 2 substitutions per synthesis. EcoRI and BamHI ends were generated to facilitate cloning of the mutant oligonucleotides. The oligonucleotides were annealed and extended with Klenow and dNTPs to yield a "double" site. This material was then cleaved with EcoRI and cloned into the EcoRI/SmaI sites of a Bluescript vector. After
transformation, clones containing a recognition site with only a few substitutions were identified by colony hybridization using an oligonucleotide containing the wild-type recognition sequence as a probe. The nature of each positive clone was then determined by dideoxy sequencing and each mutant was tested in a cleavage assay.

II.M. Labeling of Mitochondrial Translation Products

Mitochondrial proteins were labeled in vivo with $^{35}$S-$\text{SO}_4$ according to the procedure of Douglas et al. (1979).

II.N. Mitochondrial Extracts

Yeast mitochondrial extracts used in this dissertation were prepared by Dr. Roland Saldanha as described in Wenzlau et al., (1989).
CHAPTER III

RESULTS

III.A. GC cluster conversion in the Varl Gene

III.A.1. Background

The mitochondrial genome of Saccharomyces cerevisiae displays an unusual compartmentalization of its base composition. More than half of the genome is composed of DNA stretches which are almost exclusively A+T (95%) (de Zamaroczy and Bernardi, 1987). However, each mitochondrial genome is punctuated by about 200 short G+C-rich sequences (26% G+C) which are related, but vary in primary structure. These GC clusters, most of which reside in non-coding regions, have been organized into seven families (de Zamaroczy and Bernardi, 1986).

The 1186 base pair (bp) varl gene is remarkably A+T-rich and 27% of its G's and C's are clustered in the 46 bp palindromic common cluster ("cc"). Some varl genes also contain the "a" cluster within the varl reading frame 160 bp downstream from the common cluster. Members of the
family of "b" inserts are present in some strains in various combinations with other elements. More than a dozen strain-specific forms of the varl gene are reflected in distinct sizes of the varl protein which range from 40-44 kilodaltons (kd). The varl protein is associated with the mitochondrial ribosome, but these protein size polymorphisms have no obvious phenotypic consequence. The locations of these various optional inserts in the varl gene are shown in Figure 4.

Conservation of sequences flanking the various GC cluster families suggests that they may be acceptor sites for GC cluster transmission (Weiller, 1987; de Zamaroczy and Bernardi, 1986). The demonstrated conversion of the "a" element supports the notion of "targeted" GC cluster transmission (Strausberg and Butow, 1981).

The "a" GC cluster of the varl gene and the mobile introns of the 21S rRNA and coxI genes engage in unidirectional gene conversion whereby alleles lacking the optional sequence acquire it in crosses. However, conversion of the "a" insert into recipient varl alleles is somewhat less efficient than conversion of the omega and aI4a introns where transmission of the optional sequence is nearly quantitative. Furthermore, the issue of polarity associated with G+C-rich cluster conversion has not been resolved due to the lack of flanking markers
in the varl gene. To determine whether this superficial similarity between these processes implies related mechanisms new studies of the varl system were conducted that permit a much more comprehensive comparison with intron conversion.

III.A.2. A novel form of the varl gene

With the goal of uncovering suitable DNA polymorphisms in the varl region, I surveyed the varl gene of several species of the genus Saccharomyces. Characterization of mitochondrial protein profiles of several species revealed novel variants of the varl protein (see Figure 6). A subset of the Saccharomyces species was then examined by restriction site analysis to reveal DNA polymorphisms in the varl region. The varl allele of S. capensis was inspected most closely by cloning a fragment spanning the varl region and sequencing the entire gene along with its flanking regions. The strategy utilized to determine the primary sequence of the S. capensis varl allele is shown in Figure 7. This allele of the varl gene was found to lack both coding G+C-rich sequences commonly found in laboratory strains (the common cluster and the "a" element). However, the reading frame of the varl gene in S. capensis is otherwise nearly identical to that of our
standard laboratory strain ID41-6/161 (Hudspeth et al., 1982). The absence of both GC clusters in the *S. capensis* varl allele makes it the most AT-rich reading frame documented (93.1%). Both of the 20 bp A+T-rich sites which flank the common cluster and the "a" insert in other strains are perfectly conserved in *S. capensis* and their sequence in *S. capensis* is shown in Figure 8.

III.A.3. Polymorphisms in the *S. capensis* varl allele

The salient features of the sequence of the varl allele of *S. capensis* are shown in Figure 9. Since the varl gene of *S. capensis* lacks both the "common cluster" and the "a" inserts it is ideally suited as a recipient for testing the capacity of the common cluster to engage in conversion despite its reversed orientation relative to its target sequence in *S. capensis*. The *S. capensis* varl gene is the only strain characterized which lacks the "common cluster" element and is thus an appropriate recipient for such experiments. Furthermore, since *S. capensis* also lacks the "a" insert it is a potential universal recipient of GC cluster conversion in the varl gene.

The varl region of *S. capensis* also provides markers to test for flanking marker coconversion in crosses. A
novel 40 bp G+C-rich sequence ("j") is located 49 bp 5' of the var sequences in the S. capensis varl region. Its sequence is shown in Figure 10. At the site of the new "j" insertion, 11 bp are missing from the S. capensis sequence relative to that of strain ID41-6/161 (see Figure 11).

The "j" GC rich sequence bears no resemblance to those within the varl coding sequence and can be assigned to the "al" family of GC clusters according to de Zamaroczy and Bernardi (1986). The consensus sequence and salient features of this family of GC clusters is shown in Figure 12 along with the sequence of the novel "j" cluster. This family of clusters shares the highest degree of homology among its members. Typical "al" clusters are 44 bp with a GC content of 70%. Deletions are rare among "al" GC clusters and usually involve 1-3 bp; however, there are three known deletions (de Zamaroczy and Bernardi, 1986). Cluster "j" is deleted for the second nucleotide in the 13 bp "m" region (position 19) and 3 nucleotides in the nonanucleotide "b" region (positions 35, 36, and 37) compared to the "al" consensus sequence. Of the four known hypervariable positions (7, 23, 43, and 44), cluster "j" has a T to C substitution at position 7. Despite these deviations from the "al" GC cluster consensus sequence, the GC content of cluster "j" is 70%.
As shown in Figure 13, GC clusters of the "al" family can be folded into hairpin structures (Goursot et al., 1982). Most alterations in the sequence of GC clusters of this family occur at postions that are not crucial for the base pairing of the stem structure. The substitution at position 7 of the "j" cluster would actually strengthen the first portion of the stem structure shown in Figure 13, however the deletion of three nucleotides at postions 35, 36, and 37 would radically alter the latter portion of the stem structure.

The sequence of the varl allele of S. capensis also revealed a G to A transition thirty-six bp immediately before the beginning of the varl reading frame. Together, the "j" insert and the transition upstream of the varl open reading frame can serve as 5' flanking markers to measure co-conversion of sequences. The S. capensis varl gene contains a "blp" element (2 AAT codons) in a string of AAT codons starting at position 489 of the varl coding sequence which can serve as a flanking marker for co-conversion 3' of the common cluster. The sequence of the blp element is shown in Figure 14. An A to T transition is located at position 851 of the varl coding sequence and it can be used as a marker further downstream of the "blp" element. Except for these few sequence differences, the sequence of the varl gene of S. capensis is identical with
that of standard laboratory strain ID41-6/161 throughout the 1288 bp from \textit{S. capensis} examined (compared with Hudspeth et al., 1982).

III.A.4. Preferential Transmission of the Common Cluster

The \textit{var1} gene of \textit{S. capensis}, present on mtDNA in a heterothallic haploid nuclear background, permits the evaluation of a number of features of \textit{var1} conversion not previously testable. Since the new \textit{var1} allele is potentially a universal recipient of the mobile GC clusters, it is able to serve as a recipient in a cross with a strain that contains only the common cluster element (e.g., strain ID41-6/161, \textit{cc+ a- b-}). It should be noted that the \textit{var1} gene of \textit{S. capensis} has two indistinguishable cluster insertion sites but that the donor cluster is inverted relative to the known mobile "a" element (see Figure 9).

Figure 15, lanes 1 and 2, are mitochondrial protein profiles of the two parent strains; these data show that the \textit{var1} alleles used have similar mobility, with the \textit{S. capensis} form migrating somewhat slower than the tester allele. If GC cluster conversion occurs in the cross, then it would be expected that the recipient form (\textit{S. capensis}) would gain 18 amino acids and would thus
migrate more slowly. In such, experiments, mitochondrial profiles of mixed progeny can detect the presence of recombinant varl alleles so long as they represent at least 10% of total progeny (see Strausberg and Butow, 1981). As shown in Figure 15, lane 3, progeny of the cross express both parental varl forms but, in addition, contain a small proportion of cells expressing a larger varl protein. When individual progeny were screened in this way, roughly 19% (four of a set of 21 isolates examined) were found to make a larger varl protein than either parent and no cells making more than one size varl protein were encountered. Figure 16 shows a set of individual progeny screened by examination of their mitochondrial protein profiles. The sets of individual progeny were enriched for recombinants by isolating colonies expressing drug markers from both parents. In this case, recombinants bearing the cap^oli^S markers were isolated from the cross between ID41-6/161 (cap^coli^R) and S. capensis (cap^coli^S).

To test whether these data result from formation of a recombinant form (or forms) of the varl gene, mitochondrial DNA fragments containing the varl gene were analyzed directly. Figure 17A illustrates the strategy for the Southern blot experiment presented in Figure 17B. Mitochondrial DNA was cleaved with HpaII, fractionated on
a 3.0% agarose gel, blotted and hybridized with a probe capable of detecting portions of the var1 gene relevant to this analysis (see Figure 17A). The parental fragments detected are 1.2 and 0.7 kb long (S. capensis and ID41-6/161, Figure 17B, lanes 1 and 2, respectively) and both of those fragments are present in the progeny sample (Figure 17B, lane 3). However, a non-parental fragment 0.3 kb long is also present in lane 3; the fragment is generated by the HpaII site in the S. capensis "j" cluster and the HpaII site in the common cluster. The presence of the recombinant fragment is consistent with the suggestion that the S. capensis allele can acquire the common cluster, resulting in a larger gene and a larger protein product. The proposed structure of the recombinant var1 gene that is consistent with these protein and DNA data is shown in Figure 18.

III.A.5. The Common Cluster is Directed to Only One Target Site

To verify the above inferences and to determine the precise location of common cluster insertion, the sequence of the var1 region of three independent recombinant isolates was determined. The sequencing strategy employed is shown schematically in Figure 19. Although two
identical target sequences are present in the S. capensis varl gene, all three recombinants have the common cluster element inserted at the same site resulting in a region of the recombinant gene form identical to the donor. None has the common cluster element inserted in the "a" site. Since that hypothetical product would yield a HpaII fragment roughly 0.5 kb long and readily detectable in Figure 17B, lane 3, common cluster conversion is most efficient at (or even limited to) its homologous target site.

The results of the sequencing experiments are summarized in Figure 18. A portion of the varl allele of each recombinant was sequenced to permit scoring all sequence differences between the parents noted above. These progeny were chosen because the upstream GC cluster is now linked to the common cluster. However, the other three markers, the upstream transition mutation, the "blp" element and the downstream transversion mutation present in the S. capensis gene are all present in all three recombinants. Since none of these markers was co-converted (or lost, given the parental configurations), these data are the first indications that GC cluster conversion does not necessarily result in flanking marker co-conversion. This finding makes it less likely that the known mechanism of intron mobility in yeast mtDNA can
account for GC cluster mobility. The mobile rRNA gene intron co-converts single base or insertion/deletion flanking markers as far away as 1.0 kb (Jacquier and Dujon, 1986). All of the markers used in this study are much closer to the insertion site than those in the rRNA gene region that are readily co-converted.

III.A.6. Backcrosses Between the Parental and Recombinant Strains

Although the experiments described above indicate that a class of recombinants that has acquired the common cluster has varl protein with a slower mobility than that of the *S. capensis* recipient, it is possible that transmission of the "blp" element to the ID41-6/161 varl allele contributes to the population of recombinants seen in Figure 16. To determine if the "blp" element, or only the common cluster, is mobile, haploid derivatives of the recombinants (having acquired the common cluster) were crossed back to parental strains ID41-6/161 and *S. capensis* as shown schematically in Figure 20A. The mitochondrial proteins of pooled diploid progeny of each cross (Figure 20B, lanes 3 and 6) and the parental strains (Figure 20B, lanes 1, 2, 4, and 5) were labeled with $^{35}\text{S}$ to measure conversion of the inserts.
Directional conversion can be determined by comparing the ratios of the two parental var1 protein bands with drug marker outputs from each cross. The mitochondrial drug markers for strains ID41-6/161, \textit{S. capensis}, and the recombinants are $\text{cap}^\text{R}$, $\text{cap}^\text{S}$, and $\text{cap}^\text{S}$, respectively. The output of the progeny in the cross between the recombinant and ID41-6/161 can be scored for oligomycin resistance while the output of the progeny of the cross between the recombinant and \textit{S. capensis} can be determined by chloramphenicol resistance. In crosses between the recombinant and \textit{S. capensis}, where the output of the chloramphenicol marker is 33% $\text{cap}^\text{S}$ and 67% $\text{cap}^\text{R}$, only recombinant var1 protein types were observed in the pooled progeny at the expense of the \textit{S. capensis} var1 form (Figure 20B, lane 6). In a similar cross between the recombinant and ID41-6/161, the intensity of the parental var1 protein bands reflects the output of the drug markers (63% $\text{oli}^\text{R}$, 37% $\text{oli}^\text{S}$) and there is no obvious depletion of the ID41-6/161 var1 allele (Figure 20B, lane 3). This experiment corroborates the previous conclusion, that the common cluster is efficiently acquired by var1 genes lacking it. To the limits of sensitivity of labeling of mitochondrial translation products, the "blp" element does not appear to be a mobile insert.
III.A.7. Varl GC Clusters can convert separately

The "a" insert is another flanking marker that can be tested for co-conversion so another cross was designed which would permit the analysis of both mobile clusters (the common cluster and the "a" insert) in a single experiment. It is possible that conversion of one insert is either preferred over or else precludes conversion of the other. Strain 5DIT contains both GC clusters and both downstream "AAT" inserts in its varl gene, genotype cc+ a+ bl+ b2+. The cross between 5DIT and the strain containing mtDNA of S. capensis, used above, was carried out and analyzed by protein labeling of mitochondrial translation products (Figure 21B, lanes 1-3). The varl proteins of the pooled diploid progeny of the cross (Figure 21B, lane 3) show a significant fraction of non-parental forms consistent with the presence of only one of the GC-rich clusters. Individual isolates of the progeny were examined by labeling of mitochondrial proteins (Figures 22 and 23) and were then subjected to restriction analysis.

Most of the recombinants were determined by Southern analysis to have either the common cluster or the "a" element (shown schematically in Figure 24, for recombinants R1, R18, R4, and CR1). However, some
recombinants had both GC clusters, and contained the *S. capensis* flanking markers (recombinant CR12). Figure 25 shows the Southern analysis which determined the *var1* gene structure of the recombinants. Mitochondrial DNA from each of the recombinants and control mtDNA from ID41-6/161, *S. capensis*, and 5DIT was digested with *HhaI* and *HaeIII*, fractionated on a 1.0% agarose gel, blotted to a Hybond-N membrane, and hybridized to the probe indicated in Figure 24 (probe 1). All of the strains contain the *HaeIII* site just downstream from the *var1* open reading frame. However, the presence of all three GC clusters in the region ("j", "cc", and "a") can be assessed since each contains either a *HaeIII* or *HhaI* site (see Figure 24). Strains ID41-6/161 and 5DIT have fragments of 0.7 kb extending from an upstream *HaeIII* site to the *HhaI* site in the common cluster which hybridize to the probe. Since *S. capensis* contains only the "j" cluster a fragment of 1.2 kb hybridizes to the probe. Recombinants that have acquired the common cluster (R1, CR1) display a fragment of 0.3 kb, while recombinants that have acquired the "a" insert (R4, R18) have a fragment of 0.5 kb which hybridizes to probe 1 (see Figure 24). Those recombinants that have acquired both the common cluster and the "a" insert (recombinant CR12) have identical fragment patterns to strains that have only the common cluster
using probe 1. Lane 11 of Figure 25 shows mtDNA isolated from the pooled diploid progeny of the cross. Both parental mtDNA fragments of 5DIT and *S. capensis* are apparent (0.7 and 1.2 kb, respectively), however a small proportion of the 0.5 and 0.3 kb fragments indicative of the presence of varl alleles containing the "a" and common cluster inserts is also present.

To confirm the Southern blot data above and to determine which recombinants contained both GC cluster elements, the same filter was stripped and hybridized to the probe indicated in Figure 24 (probe 2). Due to the presence of the common cluster strains ID41-6/161, R1, and CR1 have fragments of 1.0 kb which hybridize to the probe. Strains 5DIT, R4, R18, and CR12 have fragments of about 0.8 kb which hybridize to the probe indicating the presence of the "a" insert. The Southern blot data demonstrate that in a cross between strains 5DIT and *S. capensis* recombinants R1 and CR1 have acquired the common cluster while recombinants R4 and R18 have acquired the "a" insert implying that the two clusters convert separately. Recombinant CR12 is a double convertant and may have arisen from sequential events. Although the Southern blot data for recombinant CR3 suggest that its varl configuration is identical to that of *S. capensis*, the slower mobility of the CR3 varl protein (see Figure
23) compared to that of *S. capensis* suggests that the *varl* gene of CR3 contains elements of the "b" family. The 1.0 kb fragment was found in mtDNA isolated from diploid progeny of the cross (lane 11) signifying the presence of alleles containing the common cluster. CR9 shown in lane 9 of Figures 24 and 26 was originally isolated as a recombinant but is identical to the *varl* configuration of 5DIT.

**III.F. Varl Recombinants Do Not Arise by Homologous Cross Overs**

To rule out the possibility that single crossover events between the clusters produce the recombinants characterized, haploid derivatives of recombinants containing the "j" insert and either the common cluster (CR1) or the "a" element (R18) were mated and their progeny analyzed by examining mitochondrial protein profiles (Figure 27). Simple crossover events could indeed generate the characterized recombinant types; however, the process should also generate a similarly frequent class of recombinants lacking both the common cluster and "a" element. When the mitochondrial proteins of the pooled diploid progeny (Figure 27, lane 3) were analyzed, a class of recombinants with both the common
cluster and "a" inserts is evident. However, the reciprocal recombinant lacking both clusters is not evident. This class of recombinant would co-migrate with the *S. capensis var1* allele used as a standard in lane 4. Its absence implies that these high frequency transmission events do not proceed chiefly by reciprocal crossing over.

III.G. GC-Rich Cluster Conversion is Independent of Mitochondrial Translation

Conversion of the optional introns omega in the 21S rRNA gene and aI4a in the *coxI* gene has a clear dependence on mitochondrial protein synthesis. The site-specific endonucleases which initiate those recombination processes are encoded within the introns themselves (Zinn and Butow, 1985; Macreadie et al., 1985; Wenzlau et al., 1989). To assess whether mitochondrial translation products are required for common cluster conversion, a petite mutant derivative containing the *var1* region of *S. capensis* was isolated by ethidium bromide mutagenesis and crossed to a petite derivative of ID41-6/161 retaining the *var1* region. Mitochondrial DNA from the diploid progeny of the petite cross was analyzed as in the previous cross between the wild type strains. The Southern blot data from this experiment are shown in Figure 28. The appearance of the
same 0.3 kb recombinant HpaII fragment in the mtDNA from pooled diploid progeny of the cross (see section II.C.1.) (as seen in the wild type cross) monitors common cluster transmission in the petite cross. The presence of this recombinant fragment in the mtDNA of Figure 28, lane 3 confirms the transmission of the common cluster in petite crosses and implies that mitochondrial translation plays no role in this process.

III.B. Group I intron conversion

III.B. Some features of mitochondrial introns

The mitochondrial genome of *Saccharomyces cerevisiae* contains another distinct class of non-essential DNA sequences some of which are known to engage in conversion. Typical laboratory strains contain up to 13 introns, though each strain is characterized by a unique set of introns. A novelty of mitochondrial introns is that many contain open reading frames; some of them have been shown to encode intron-specific proteins which facilitate splicing of the introns that encode them (maturases). The view developed in the early 1980's that if an intron has an open reading frame, the product of that open reading frame is likely involved in splicing. However, studies
of conversion of the optional omega intron in the 21S rRNA gene altered this view.

III.B.2. Intron nomenclature in the coxI gene

Recently, the configuration of the coxI gene in mtDNA of many homothallic species of the genus Saccharomyces has been examined (Ralph, 1986; Dib-Hajj et al., in preparation). Table 3 lists these species and summarizes the intron structures of their coxI genes. The seven previously identified introns in the coxI gene (aI1, aI2, aI3, aI4a, aI5a, aI5B and aI5g) as well as two new introns are distributed in various combinations among the species and laboratory strains listed in Table 3. Ralph (1986) defined two new introns; aI3g, which is present in the coxI genes of seven of the organisms surveyed, and aI4B, which is present in the coxI genes of three organisms. The discovery of these two new introns in the coxI gene (aI3g and aI4B) and the identification of a third new intron in S. douglasii (Kotylak et al., 1985) (not listed) mandated a new nomenclature for the coxI introns. The three new introns are interspersed among previously studied introns aI3 and aI4 (which have been renamed aI3a and aI4a) as well as aI5g. The intron identified in S. douglasii is the closest new intron to aI3a and has thus
been termed aI38. Another new intron is positioned between aI38 and aI4a at position 7188 of the published sequence of the coxl gene (Bonitz et al., 1980) and has been termed aI3g. The final new intron has been named aI48 since it is located between introns aI4a and aI5a.

III.B.3. Survey of mitochondrial introns for mobility

To learn whether intron conversion is a special feature of the omega intron in the 21S rRNA gene or perhaps a property of some of the other 15 optional introns in the mitochondrial genome, coxl genes with and without particular introns were required. Table 3 demonstrates that a wide diversity in coxl gene structure exists among the species and laboratory strains of Saccharomyces; each strain has its own unique collection of introns. By designing crosses between strains that have and lack these introns, most coxl introns can be tested for their capacity to engage in conversion.

To utilize these homothallic yeast strains, it was necessary to transfer their mtDNA from the diploid, homothallic nuclear backgrounds in which they were originally found to stable haploid nuclear genomes of standard laboratory strains. The desired constructs required for the various crosses were accomplished by a
combination of protoplast fusion experiments and karl mediated cytoduction (see Materials and Methods).

Nuclear-mitochondrial incompatibility, as was found in similar hybrids between _S. douglasii_ and _S. cerevisiae_ (Kotylak et al., 1985) was not encountered.

A typical experiment to follow conversion of introns can be designed to monitor several introns simultaneously. For instance, a cross between strain D273-10B and _S. norbensis_ is shown schematically in Figure 29. These strains have four introns in the _coxl_ gene in common, aI1, aI2, aI3a, and aI5g. _S. norbensis_ lacks aI4a which is present in strain D273-10B, while strain D273-10B lacks intron aI3g and aI5B which are present in strain _S. norbensis_. Upon conversion of one of these introns, the _coxl_ gene will be restructured. Such conversion events can be followed by restriction site mapping followed by Southern analysis using defined probes. These strains are also distinguished by the omega intron in the 21S rRNA gene which is present in D273-10B but absent from _S. norbensis_. A point mutation in the _olil_ gene conferring resistance to oligomycin in strain D273-10B is located more than 10 kb from the _coxl_ and rRNA genes and further distinguishes it from _S. norbensis_.

III.B.4. _CoxI_ intron aI4a is mobile in crosses
D273-10B and *S. norbensis* cells were mated and diploid progeny were selected and scored for the transmission of the *olil* gene alleles. The cross is essentially unbiased since approximately 60% of the progeny express the *S. norbensis* *olil* allele and 40% express the D273-10B *olil* allele in the nuclear backgrounds of these strains. Thus intron conversion can be easily monitored in such a cross since there is no biased transmission of parental alleles.

Mitochondrial DNA was isolated from both haploid parents and pooled diploid progeny of the cross (see section II.C.1.) and cleaved with various restriction enzymes. Blots of fragment patterns were then hybridized with a defined set of probes to assess the output of specific introns (Figures 30-33). Figure 30A shows the transmission of *HpaII* fragments in the 21S rRNA gene region in strains D273-10B and *S. norbensis*. When hybridized to an oligonucleotide probe complementary to sequences in the downstream exon (see asterisk in Figure 30A), the parental fragments of D273-10B and *S. norbensis* are 1.9 kb and 0.8 kb, respectively. As demonstrated by the blot in Figure 30B, omega conversion proceeds normally in this cross since all of the progeny possess the intron-containing restriction fragment. Although there are many
differences between these two mitochondrial genomes, omega conversion still occurs efficiently.

In Figure 31A, a schematic diagram shows how the output of intron aI5B was determined. A 3.0 kb BclI fragment containing aI5B and aI5g is detected in S. norbensis using a probe containing the last three exons of the coxl gene. The same probe hybridizes to a 1.5 kb BclI fragment from D273-10B (which lacks aI5B). When mtDNA isolated from the pooled diploid progeny of the cross is examined with the same probe it is clear that both parental fragments are transmitted to the progeny. The frequencies of these alleles among the progeny is coordinate with the ratio of olil alleles. That is, the output of olil parental alleles is roughly 60% S. norbensis and 40% D273-10B and the approximate relative porportion of the BclI parental fragments in the coxl gene reflects this ratio. Thus intron aI5B does not appear to be active in biased gene conversion at least in this configuration.

In Figure 32, BamHI/EcoRI digests allow the examination of the output of the other two polymorphic introns in the coxl gene, aI3g and aI4a. Both strains contain a BamHI site in the 3' portion of aI3a and an EcoRI site in the last exon of the coxl gene. When hybridized to an oligonucleotide specific for exon 5 (of
strain ID41-6/161) the parental fragments of D273-10B and *S. norbensis* can be visualized at 5.5 and 3.1 kb, respectively (Figure 32B, lanes 1 and 2). In the mtDNA isolated from the pooled diploid progeny of the cross (Figure 32B, lane 3), about half the progeny contain the D273-10B parental fragment (3.1 kb) while the other half contain a novel fragment of 4.6 kb. Furthermore, there appears to be a complete loss of the *S. norbensis* parental allele (5.5 kb).

In Figure 33A, when BclI digests are performed using the same oligonucleotide probe used in Figure 32 (exon 5 specific), the outputs of introns a13g and a14a can be followed. The parental fragments are 2.2 kb and 1.3 kb for *S. norbensis* and D273-10B, respectively (Figure 33B, lanes 1 and 2). In Figure 33B, lane 3, all of the progeny appear to lack a13g and have a14a. This Southern blot data indicates that a14a of strain D273-10B is transferred to the *coxI* gene of *S. norbensis* with the simultaneous loss of a13g by co-conversion of flanking sequences.

III.B.5. The structure of recombinants which have acquired a14a

The proposed structure of the recombinant *coxI* allele is shown in Figure 34. Several experiments were performed
to confirm this recombinant structure. The conformation of the \textit{coxl} gene of three independent recombinant isolates was subjected to detailed restriction site mapping to confirm the structure shown in Figure 34. All the restriction sites shown have been confirmed. Also the BclI fragment shown in Figure 34 from a recombinant isolate chosen at random was cloned into the pUC-18 vector and then subcloned into M13 mp-18 and mp-19 for sequence analysis. The regions indicated by the arrows in Figure 34 represent the sequence which was determined and compared with that of the known sequence of D273-10B in this region. The sequence data (as shown in Figure 35) confirm that intron aI4a is located at precisely the same position in both the donor (D273-10B) and the recombinant.

As shown in Figure 36, among a set of 26 individual progeny isolates those that contain the \textit{oliI}s allele from the \textit{S. norbensis} parent, most have the novel 4.6 kb BamHI/EcoRI fragment, whereas most of the progeny that have the 3.1 kb BamHI/EcoRI fragment carry the \textit{oliI}f allele of the D273-10B parent. Thus the novel recombinant 4.6 kb fragment appears to be linked to the \textit{oliI}s marker and was generated at the expense of the 5.5 kb BamHI/EcoRI fragment of the \textit{S. norbensis} parent.

III.B.6. Recombinants do not have a selective advantage
There is an alternative explanation for the observation that recombinants containing the aI4a intron predominate in the population of mitochondrial genomes following this cross. Figure 37 shows that a single crossover event in exon 5 (of D273-10B) would result in the recombinant coxI structure shown as "R1" in the figure. The reciprocal recombinant "R2", however, was not detected among the progeny. It is possible that "R1" possesses some type of selective advantage over the other mitochondrial genome types. "R1" may be preferentially replicated such that its proportion in the output exceeds that of "R2" and is similar to that of the D273-10B (donor) parental genome.

To test this possibility mixed growth experiments were conducted with diploids harboring different mitochondrial genomes. The relative fraction of mitochondrial genomes from norbensis, capensis, and the norbensis-derived recombinant (Rec-N) diploids in mixed cultures with D273-10B diploids was measured over 35 generations. At various generation times, an aliquot of each mixed culture was assayed for the ratio of oli (norbensis, capensis, and Rec-N) and oli (D273-10B) cells. Although the ratios (shown in Figure 38) varied slightly over the course of the growth period, there
appears to be no obvious growth advantage conferred to the recombinants which have acquired the aI4a intron.

III.B.7. Flanking marker coconversion accompanies aI4a conversion

One feature of omega intron conversion events is coconversion of sequences adjacent to the converted intron. In the recombinant, intron aI3g has been exactly excised, probably by coconversion 5' of aI4a insertion. The BclI fragment containing aI3g shown in Figure 34 was cloned from S. norbensis and compared to the sequence of D273-10B in the regions indicated by the arrows (Figure 34). The sequences in the region 3' of the intron insertion site differ at the two residues indicated in Figure 39 by the asterisks. These sequence polymorphisms that are present in the donor, are also present in the recombinant implying that co-conversion extends 3' of the intron insertion site as well. The extent of flanking marker co-conversion cannot be efficiently measured in this cross for lack of more distant polymorphic flanking markers. However, intron aI58 has not been removed by coconversion.

III.B.8. Site specific double-strand cleavage of
Several lines of evidence have shown a role for double stranded cuts as recombinogenic structures in yeast. The most appropriate examples for the studies conducted here are the mechanisms of mating type interconversion at the MAT locus in the yeast nucleus and conversion of the omega intron in the 21S rRNA gene (Zinn and Butow, 1985; Nickoloff et al., 1986). To determine whether conversion of the aI4a intron involves a double strand cleavage of recipient molecules, synchronously mated cells of a cross between D273-10B and S. norbensis were analyzed. At various time points mtDNA was extracted from the mated cells and the "intactness" of a fragment spanning the intron insertion site was measured by fractionating HaeIII digests on 6.0% acrylamide gels, blotting the fragments to Hybond-N membranes, and hybridizing them to an oligonucleotide probe. The strategy for this experiment is shown schematically in Figure 40B. In mixed cells (Figure 40B, lane 3) and cells mated for 3 hours (Figure 40B, lane 4) only the parental HaeIII fragment of S. norbensis is present since the oligonucleotide probe is complementary to sequences in intron aI4B which are not present in D273-10B. However, by 6 hours of mating (Figure 40B, lane 5) the expected
0.9 kb product of cleavage of the *S. norbensis* parental fragment near the site of intron insertion is evident. This double strand cut is transient; that is, it persists through 9 hours of mating, but is absent from samples prepared from cells mated for 24 hours (Figure 40B, lanes 6 and 7).

As a control, the same mtDNA samples were digested with HpaII to demonstrate cleavage of the omega- alleles of the 21S rRNA gene. The relevant region of the 21S rRNA gene is shown schematically in Figure 40A. HpaII digests were fractionated on 6.0% polyacrylamide gels, blotted, and probed with an oligonucleotide complementary to sequences in the 3' exon of the 21S rRNA gene. Similar to the experiment described above, at early time points of mating, only the parental fragments of D273-10B (1.9 kb) and *S. norbensis* (0.8 kb) are evident. However, by 6 hours of mating, a new fragment of 0.6 kb is present whose size is consistent with cleavage in the vicinity of intron insertion. Again, the signal persists through 9 hours of mating (Figure 40A, lane 6) but is absent from samples taken from cells mated for 24 hours (Figure 40A, lane 7). Thus, during mating an endonuclease activity is present that is capable of cleaving recipient genomes in the proximity of the intron insertion site.
III.B.9. AI4a conversion requires mitochondrial translation

Studies by Zinn and Butow (1984) demonstrated that conversion of the 21S rRNA intron requires a product of mitochondrial protein synthesis that was later shown to be a product of the intron reading frame (Jacquier and Dujon, 1985; Macreadie et al., 1986). To determine whether a mitochondrial translation product is necessary for conversion of aI4a, crosses were conducted between petite mutants retaining the \textit{coxl} gene of \textit{S. norbensis} (recipient) and the long form of the \textit{coxl} gene of strain ID41-6/161 (donor).

Petite derivatives of \textit{S. norbensis} that conserve the \textit{coxl} region of the mitochondrial genome were generated by ethidium bromide mutagenesis and were mapped to the \textit{coxl} region by standard petite deletion mapping (see Materials and Methods). Figure 41A shows a schematic diagram of the strategy employed for the Southern analysis shown in Figure 41B and 41C. Mitochondrial DNA from parental strains Hp- and Np- and pooled diploid progeny of a cross between these strains was digested with BclI, fractionated on a 1.0% agarose gel, blotted, and probed with a fragment containing aI3g and its flanking exons. The parental fragments of Hp- and Np- are 1.3 and 2.2 kb,
respectively. As shown in Figure 41B, lane 3, there appears to be no loss of the 2.2 kb Np− parental fragment as would be expected upon conversion of intron aI4a. The Np− parental 2.2 kb fragment is more prevalent since the probe has more homology to the Np− mtDNA.

Next, the same mtDNA samples were digested with HaeIII, analyzed by Southern analysis with the same probe, and the results are shown in Figure 41C. The Hp− parental fragments are 3.4 and 0.8 kb (Figure 41C, lane 1) while the Np− parental fragments are 2.3 and 1.5 kb (Figure 41C, lane 2). All four parental fragments are transmitted to the progeny (Figure 41C, lane 3). If aI4a conversion occurs, a novel fragment of 1.9 kb extending from the HaeIII site in aI4a to the HaeIII site in aI4B would be present. Since this fragment is absent, it appears that aI4a conversion does not proceed in petite mutants which are devoid of mitochondrial translation. Furthermore, there is no obvious depletion of the Np− parental fragment which would be the recipient of aI4a conversion. This is consistent with the conclusion that a product of mitochondrial protein synthesis is necessary for aI4a conversion.

III.B.10. The source of the endonuclease required for aI4a conversion
To begin to identify the source of the needed mitochondrially synthesized protein, a number of other strains were surveyed for their ability to serve as donors or recipients of a\textsuperscript{14}a in crosses (summarized in Table 4). For instance, the \textit{fit1} product of the omega intron might direct both its own conversion, and the conversion of another intron (such as a\textit{14}a). In each case described below, mtDNA was isolated from parental strains and pooled diploid progeny of the cross and subjected to Southern analysis to assess whether a\textit{14}a conversion occurs.

Crosses were conducted with \textit{S. capensis} (recipient) and D273-10B (donor) to ensure that a\textit{14}a conversion was not simply a special feature of the cross between D273-10B and \textit{S. norbensis}. The \textit{coxl} gene of \textit{S. capensis} (Table 4, line 2) differs from that of \textit{S. norbensis} (Table 4, line 1) in that it lacks intron a\textit{11} and a\textit{12} and has a second new intron a\textit{14}B (Dib-Hajj et al., in preparation). The strategy for this Southern blot analysis is shown in Figure 42A. Mitochondrial DNA from both parental strains \textit{S. capensis} and D273-10B and pooled diploid progeny was digested with EcoRI and HindIII, separated on a 1.0\% agarose gel, blotted, and probed with the fragment indicated in Figure 42A. The parental fragments of D273-10B and \textit{S. capensis} are 1.6 and 2.0 kb, respectively
Upon conversion of aI4a, a novel fragment of 1.1 kb is generated which is present in the mtDNA prepared from the pooled diploid progeny (Figure 42B, lane 3). Lane 4 (Figure 42B) shows mtDNA from the pooled diploid progeny of a cross between S. capensis and a constructed omega- derivative of D273-10B (see Materials and Methods). In this cross both strains lack omega; however, the expected product of aI4a conversion, the 1.1 kb fragment, is present. Thus, an omega- derivative of the same strain is as good a donor as is the omega+ version. These data also demonstrate that S. capensis is a recipient for aI4a conversion and that intron aI4B (present adjacent to aI3g in the S. capensis coxI configuration) is not lost by co-conversion.

S. ellipsoideus (Table 4, line 3), which has only introns aI1, aI3a and aI3g is also a recipient of aI4a. Figure 43A shows the Southern blot strategy used for a cross between D273-10B and S. ellipsoideus using the Bcoli fragment containing aI3g and its flanking exons as a probe (see Figure 43A). In HaeIII/EcoRI digests the parental fragments of D273-10B and S. ellipsoideus are 2.5 and 1.3 kb (Figure 43B, lanes 1 and 2), respectively. Upon conversion of aI4a, a novel fragment of 1.6 kb is generated at the expense of the S. ellipsoideus 1.3 kb parental fragment. Lane 3 of Figure 43B contains mtDNA
from pooled diploid progeny; the expected product of conversion is evident (1.6 kb fragment) as is the diminished \textit{S. ellipsoideus} parental fragment (1.3 kb).

Strain ID41-6/161 (Table 4, lane 5) contains intron aI5a and aI5B in addition to the set of introns present in strain D273-10B (Table 4, lane 4). In crosses with \textit{S. norbensis}, it is also a donor of aI4a. Figure 44A shows the strategy for the Southern analysis which is shown in Figure 44B. Using an oligonucleotide specific for exon 5 (of ID41-6/161), the BamHI/EcoRI parental fragments are 6.0 and 5.5 kb for ID41-6/161 and \textit{S. norbensis}, respectively (Figure 44, lanes 1 and 2). Upon conversion of aI4a, a novel fragment of 4.6 kb extending from the BamHI site in aI3a to the EcoRI site in the last exon of the \textit{coxl} gene (see Figure 44A) is generated (Figure 44B, lane 3).

A derivative of ID41-6/161 which is cleanly deleted for introns aI1, aI2 and aI3a (ID41-6/ΔI-3, Table 4, line 6) was also tested as a donor for intron conversion in crosses with \textit{S. norbensis}. In EcoRI/HindIII digests the parental fragments of ID41-6/ΔI-3 and \textit{S. norbensis} are 4.6 and 4.0 kb, respectively (Figure 45B, lanes 1 and 2), using the probe indicated in Figure 45A. The Southern blot in Figure 45B shows that in mtDNA isolated from pooled diploid progeny from the cross, the expected 3.3 kb
fragment is present due to conversion of aI4a (Figure 45B, lane 3). Thus the presence of introns aI1, aI2 and aI3a is not necessary for aI4a conversion. Both strain ID41-6/161 and its derivative ID41-6/A1-3 (Table 4, lines 5 and 6) lack the omega intron in the 21S rRNA gene which corroborates earlier results showing that aI4a conversion is independent of the omega intron.

Only introns aI3a, aI4a, and aI5g are present in strain GII-5g (Table 3, line 5) (see Materials and Methods for strain construction). When crossed to S. norbensis, this strain serves as a donor of aI4a. As shown in Figure 46A, BamHI/EcoRI digests of the parental mtDNA yield fragments of 3.1 and 5.5 kb for strains G-II-5g and S. norbensis, respectively, when hybridized to the probe indicated in Figure 46A. Lane 4 of Figure 46 shows that the recombinant which has acquired aI4a contains a BamHI/EcoRI fragment of 4.6 kb. Lane 5 of Figure 46B shows that in mtDNA isolated from pooled diploids from a cross between GII-5 and S. norbensis a 4.6 kb fragment, which co-migrates with that of the recombinant, is evident.

A derivative of GII-5, GII-0, which contains only introns aI3a and aI4a is shown in Table 3, line 6. This strain also serves as a donor of aI4a in crosses with S. norbensis. In BamHI/EcoRI digests, the parental fragment
of GII-0 is 2.2 kb while that of *S. norbensis* is 5.5 kb using the indicated probe (Figure 46A). Again, in mtDNA isolated from pooled diploids of a cross between GII-0 and *S. norbensis* (Figure 46B, lane 6), the recombinant 4.6 kb fragment is present indicating conversion of aI4a. Thus, none of the introns that are absent from strains GII-5g or GII-0 are necessary for aI4a conversion.

A mutant of ID41-6/161, PZ27, which is deleted for the entire cob gene (Perlman et al., 1980) is also a donor of aI4a in crosses with *S. norbensis*. Figure 44B (lanes 1 and 2) shows that in BamHI/EcoRI digests of mtDNA samples prepared from parental strains PZ27 and *S. norbensis*, 6.0 and 5.5 kb fragments hybridize to the indicated probe. In lane 4 of Figure 44B, the recombinant 4.6 kb fragment is present in mtDNA isolated from the pooled diploids of the cross. Thus none of the introns in the cob gene (bI1, bI2, bI3, bI4 and bI5) appear to be necessary for conversion of aI4a.

By the process of elimination and the above survey of 15 introns in the mitochondrial genome (summarized in Table 4), if an intron encodes the needed endonuclease function to initiate the recombination process of aI4a conversion, it is likely aI4a itself. This inference was tested directly in the following section.
III.B.11. Mutations in coxl intron aI4a

The products of some intron open reading frames have been implicated in splicing (see Introduction section I.B.6.). Mutations in intron reading frames fall into two classes. One group, termed trans-recessive mutations, either truncates the intron reading frame or creates missense mutations. Trans-recessive mutants are incapable of splicing since they cannot provide a functional maturase for splicing. Another group of mutations, termed cis-dominant mutations, destroy secondary structures within the intron which are crucial for splicing. These mutations usually leave the open reading frame intact, but render the intron unspliceable. An interesting consequence of these mutations is that the product of the intron open reading frame is overexpressed.

Earlier splicing studies conducted by Anziano (1984) required the generation of a set of mutations in the intron reading frame of aI4a. The genetic screen employed a mutagenesis of a trans-recessive mutation in bI4 which was suppressed by a NAM-2 nuclear mutation and was thus dependent upon the aI4a intron product for splicing. Respiratory deficient mutations were mapped to the aI4a intron and classified as either cis-dominant or trans-recessive mutations by standard petite mapping (Anziano,
For instance, trans-recessive mutants of aI4a could be rescued by either a petite containing the aI4a or bI4 intron, while cis-dominant aI4a mutants could only be rescued by petites containing aI4a. A subset of these mutants was employed in studies of aI4a conversion.

The exact sequence alterations of two cis-dominant and two trans-recessive aI4a mutants was determined by cloning a BclI fragment containing aI4a and flanking exons into pUC-18. These clones were then subcloned into either M13 mp-18 or mp-19 to determine the nature of their mutations using a set of oligonucleotides for dideoxy sequencing that were complementary to sequences in the intron (shown in Figure 47). The entire aI4a intron of each mutant was sequenced and the nature of cis-dominant mutants E and G, and trans-recessive aI4a mutants K and A3 were determined in collaboration with R. J. Saldanha. The positions of the mutants are shown schematically in Figure 48A. Mutant E is a G to A silent change in the internal guide sequence at position 7206 of the D273-10B sequence (numbered as in Bonitz et al., 1980). This sequence change destroys a HaeIII site just 7 bp from the 5' end of the intron. The mutant G is a C to T change at position 7400 in the cis-acting sequence defined as P6 by Burke et al. (1987), which changes a serine to phenylalanine. Mutant K has -1 frameshift mutations at
positions 7596 and 7815 and bears an A to T transversion at position 7446 which changes a lysine to asparagine. Mutant A3 has a G to T mutation at position 8125 that terminates the reading frame prematurely, just eight amino acids from the end of the reading frame.

III.B.12. Mutations that truncate the aI4a protein block conversion

The above mutants were employed as donors in a standard conversion assay to measure their ability to engage in gene conversion. Mitochondrial DNA from parental strains ID41-6/161 and *S. norbensis* and diploid progeny from crosses between *S. norbensis* and ID41-6/161 and its mutant derivatives was digested with BamHI and EcoRI, electrophoresed on a 1.0% agarose gel, transferred to a Hybond-N membrane and hybridized to an exon 5 specific oligonucleotide. Parental fragments of ID41-6/161 and *S. norbensis* are shown in lanes 1 and 2 of Figure 48B, respectively. In the control cross (Figure 48B, lane 3) between ID41-6/161 and *S. norbensis*, conversion of aI4a generates the novel 4.6 kb fragment in place of the recipient 5.5 kb fragment. Crosses between the recipient strain and cis-dominant aI4a mutants E (Figure 48B, lane 5) and G (Figure 48B, lane 6) carry out
intron conversion as measured by the appearance of the 4.6 kb fragment. In crosses with trans-recessive aI4a mutant K (Figure 48B, lane 4) and A3 (Figure 48B, lane 7), conversion is abolished. The data indicate that the former two mutants which do not truncate the intron reading frame do not abolish conversion while the latter two mutants which truncate the intron reading frame also inhibit conversion. Thus aI4a conversion is dependent upon an intact aI4a open reading frame.

In the crosses conducted with ID41-6/161 and its derivatives as donors of aI4a, the outputs of the crosses differ from crosses conducted with D273-10B as a donor. For instance, in crosses conducted with D273-10B aI4a conversion is nearly quantitative, while in crosses with ID41-6/161, many recipient molecules remain in the progeny. Such fluctuations in the bias of the cross are believed to be dependent upon the nuclear background of the strains and the physiological condition of cells being mated.

III.B.13. Sequence specific endonuclease activity in mitochondrial extracts

The endonuclease activity that cleaves omega^- mtDNA flanking the intron insertion site has not been detected
in extracts of yeast mitochondria presumably because very little is synthesized in wild-type strains. Since cis-acting mutants of aI4a naturally overproduce the product of the aI4a reading frame due to a block in aI4a splicing, R. J. Saldanha prepared mitochondrial extracts from these mutants to detect the endonuclease activity in vitro (Wenzlau et al., 1989). Mitochondrial lysates were shown to cleave plasmid DNA containing the intron insertion site (or the recipient allele), but not plasmids containing the donor allele. Overproduction of the endonuclease is crucial since no activity could be detected in wild type strains even with an excess of extract. The in vitro experiments conducted by R. J. Saldanha are in complete agreement with the in vivo data presented above since extracts prepared from trans-recessive mutants which can not convert aI4a have no endonuclease activity, while extracts prepared from cis-dominant mutants which can convert aI4a have a potent endonuclease activity. These data suggest that the endonuclease activity is encoded by the aI4a intron itself. This has been confirmed by experiments conducted by C. Jacq who has demonstrated that the corrected aI4a reading frame engineered in E. coli is expressed as an endonuclease (Delahodde et al., 1989).

III.B.14. The cleavage site of the aI4a-encoded
endonuclease

In collaboration with R. J. Saldanha, the precise point of cleavage by the aI4a endonuclease was determined. The substrate, containing 70 bp of the 3' terminal portion of aI3g and 158 bp of the 5' flanking exon, was cloned into an M13 vector and using the universal primer we generated an "end-labeled" double stranded substrate. The primer was extended for a short distance with [32P]dATP to label the substrate. This material was then either chased with an excess of unlabeled nucleotides to convert all the substrate molecules to their complete double-stranded form or used in dideoxy-chain termination reactions to generate a sequencing ladder. The substrate was cleaved with the extract, denatured and fractionated alongside of a set of sequencing ladders. Both strands were used as substrate to determine the precise point of cleavage on each. In Figure 49A and 49B, the cleavage on the sense strand and anti-sense strand has been determined, respectively. Figure 49C summarizes the data and shows that the enzyme creates a 4 bp 3' overhang and cleaves downstream of the site of intron insertion (Wenzlau et al., 1989).

R. J. Saldanha used an extension of the above approach to estimate the boundaries of the recognition
sequence. The sequencing ladders used above are in effect a set of nested deletions that differ in one nucleotide steps. When used as a substrate for cleavage, only those ladders that have extended far enough to generate a double stranded recognition sequence will be cleaved. Using ladders generated from both strands as substrates, the recognition sequence was determined to be no longer than 18 bp and is situated asymmetrically about the cleavage and intron insertion sites (see Figure 49C) (Wenzlau et al., 1989).

III.B.15. Dissection of the recognition sequence of the aI4a endonuclease

Although a saturation mutagenesis of the target site for the omega endonuclease has shown that at least one substitution at each position of its 18 bp recognition sequence can abolish cleavage of the substrate (Colleaux et al., 1988), the recognition site for the aI4a endonuclease appears to be much less stringent. Figure 50 shows that the recipient allele used in our studies differs at 3 positions from that used by the group that expressed the aI4a endonuclease in E. coli. However, both substrates are efficiently cleaved by mitochondrial extracts (Wenzlau et al., 1989). These sequence
polymorphisms that appear in the aI4a substrate occur in similar positions to changes in the MAT and omega substrates that render those substrates inactive (see Figure 51). Interestingly, the donor substrate contains a large portion of the recognition sequence; however, it is resistant to cleavage, indicating at least one of the first five bases of the recognition sequence are essential for cleavage to occur (Figure 50, line 3).

In collaboration with R. J. Saldanha, an extensive mutagenesis of the recognition sequence was undertaken. An oligonucleotide was generated that had EcoRI and BamHI ends which were complementary (see Figure 52). The mixes used in the oligonucleotide synthesis were adjusted to yield a 10% mutation rate at each position of the 18 bp recognition site. The oligonucleotides were annealed at their complementary ends and extended with PolI Klenow fragment and dNTPs to yield a "double" recognition site with blunt ends. This double stranded material was then cleaved with EcoRI and cloned into the EcoRI/SmaI sites of a Bluescript vector. After transformation, we screened for clones containing a recognition site with only a few substitutions by colony hybridizations using an oligonucleotide containing the correct recognition sequence as a probe. Each positive clone was then sequenced and tested for cleavage with the mitochondrial
extract. Figure 53 shows a summary of the mutant substrates which have been analyzed and whether they are compatible with cleavage by the mitochondrial extract. Several single internal substitutions are consistent with cleavage (positions 1, 2, 5, 8, 9, 11, 14, 17, and 18). Substitutions at the ends of the recognition sequence (substitutions 1, 2, 17, and 18) may reduce the minimum recognition site to 14 bp.

III.C. Another mechanism for intron mobility

III.C.1. Intron mobility through RNA intermediates

Gene conversion appears to be a common mechanism for transmission of group I introns as intron transmission has been demonstrated in quite diverse organisms. It need not, however, be the only mechanism by which stable intron configurations can change. Other mechanisms for intron propagation, as well as intron loss, may involve RNA intermediates. For instance, Cech has recently demonstrated that splicing can proceed in the reverse direction in vitro (Woodson and Cech, 1989). In fact, they also showed that introns may be introduced into new exonic environments as long as a few simple constraints are met. A reverse transcriptase acting on these products
of reverse splicing could generate cDNAs which, upon recombining with genomic DNA, could result in transposition of the intron to new locations. Similarly, a reverse transcriptase acting on mature message could lead to intron loss.

As discussed in section I.I.1.a., several groups have noted that one mode of reversion of splicing deficient intron mutants is the deletion, or "popping out" of the offending intron (Jacq et al., 1982; Gargouri et al., 1983; Perea and Jacq, 1985; Hill et al., 1985; P. Q. Anziano, unpublished). This is frequently accompanied by the loss of neighboring introns. The simplest interpretation is that a DNA copy of mature message (that lacks all introns) recombines with the genome to delete the introns.

The ORFs of two yeast mitochondrial introns, aIl and aI2 of the coxI gene, as well as one intron in Podospora anserina and one intron in Schizosaccharomyces pombe, and the intron-like Mauriceville plasmid of Neurospora crassa share striking homologies to viral reverse transcriptase (Michel and Lang, 1985). The homology occurs at distinct blocks of amino acid stretches in a portion of the intron reading frames. The order of, and distance between, conserved blocks (1-7) (see Figures 54 and 55) are similar. Figures 54 and 55 show the alignment of similar
segments in mitochondrial intron and plasmid encoded proteins, proteins encoded in mobile genetic elements from fungi, insects, and ciliates, and viral reverse transcriptases taken from Xiong and Eickbush (1988).

Several lines of evidence have indicated the presence of reverse transcriptase activity in fungal mitochondria. A reverse transcriptase activity encoded by Neurospora mitochondrial plasmids has been documented and is believed to function in replication of the plasmids (Kuiper and Lambowitz, 1988). Furthermore, some senescence events in *Podospora anserina* are associated with the appearance of asenDNA, which is a cleanly excised group II intron that encodes a protein similar to reverse transcriptase (Steinhilber and Cummings, 1986).

Reverse transcriptase activity is also present in the yeast nucleus. The yeast transposon Tyl has been shown to encode a reverse transcriptase and transpose via an RNA intermediate. There are approximately 50 copies of Ty elements belonging to six families in a typical yeast genome. These elements are highly expressed and account for 10% of the cellular message. Not all elements are active in transposition since the rate of transposition is relatively low (10^-7 per cell division). The limiting factor appears to be reverse transcriptase since in constructs where an active marked Ty element is driven by
an inducible promotor, between 30 and 70% of the cells have marked transposition events (reviewed by Boeke, 1988). These elements may have been responsible for the streamlining of the yeast genome since the reverse transcriptase could have systematically directed intron removal (Fink, 1987).

III.C.2. A strategy to assay reverse transcriptase activity in yeast mitochondria

To test whether a reverse transcriptase activity is present in yeast mitochondria, and therefore might play a role in reversion of yeast mitochondrial splicing defective mutants by intron excision, I employed mitochondrial genetics. To determine whether CoxI introns aII and aI2 are sources of reverse transcriptase, a "reporter gene" for the putative reverse transcriptase activity was required. A splicing defective mutant in intron aI5g (C2116) had been characterized genetically by R. Dietrich and physically by S. Belcher. The mutant has a G to A substitution near the 5' end of the intron just before domain 1 (at position 8625) and an insertion of an A after nucleotide 9282 in domain 3 of intron aI5g (numbered as in Bonitz et al., 1980) (R. Dietrich and S. Belcher, unpublished). Of the many types of revertants
possible (nuclear suppressors, second site suppressors, true revertants, and "popouts") only two types of revertants have been identified. R. Dietrich characterized several temperature resistant (TR) revertants of C2116 and found they were clean intron deletions of aI5g. Most temperature sensitive revertants (TS) have been shown to be changes within domain 3 of intron aI5g (S. Belcher, unpublished).

III.C.3. C2116 is a suitable reporter intron mutation

To establish whether mutant C2116 would be a suitable "reporter" strain for reverse transcriptase activity, I expanded the initial analysis of revertants of C2116. Figure 56 shows the strategy used to analyze a set of 46 revertants of C2116 using Southern blot analysis. In BamHI/EcoRI digests using the indicated oligonucleotide probe in exon 5, the presence or absence of all coxl introns downstream of aI3a can be assessed. A fragment of 6.0 kb containing introns aI4a, aI5a, aI5B, and aI5g hybridizes to the oligonucleotide probe in the initial mutant strain and its wild-type parent. In strains where aI5g has been excised, the analogous fragment is somewhat smaller, 5.1 kb (Figure 56, line 2). A 3.5 kb fragment (Figure 56, line 3) would be observed if both aI5g and
aI5B were excised. Similarly, if aI5a, aI5B and aI5g were all excised, the BamHI/EcoRI fragment would decrease in size to 2.1 kb (Figure 56, line 4).

Figure 57 shows the Southern blot data of BamHI/EcoRI digests of DNA from 46 revertants of C2116 (which contains all and aI2 coxi introns, and hence a potential source of reverse transcriptase). All of the TS revertants shown in lanes 1-10 contain the 6.0 kb fragment indicating that no introns have been excised. This sample was expanded by M. Henke who examined a set of 149 TS revertants of C2116 by dot blot analysis using an oligonucleotide complementary to sequences in intron aI5g. All TS revertants examined still contained intron aI5g. As a positive control the same dot blot filters were hybridized to an exon probe. Thus all TS revertants appear to retain aI5g.

However when I examined a set of 36 TR revertants from C2116 by Southern blot analysis, a majority had at least one intron excised. In the set of 36 TR revertants shown, 15 contained the 5.1 kb fragment suggesting that intron aI5g had been excised, and 11 contained the 3.5 kb BamHI/EcoRI fragment indicating that both aI5g and aI5B had been excised. Thus, in this sample, 72% of TR revertants are "popouts" and the excision of only one intron, aI5g appears to be slightly favored (58%).
is in contrast to conclusions by Gargouri et al., (1983) who found that multiple intron excisions in the cob gene were more frequent than single excisions.

Those revertants that are designated with an asterisk in Figure 57 were examined by DNA sequence analysis to determine whether the intron excision events were precise. HindIII/EcoRI fragments (shown in Figure 58) from revertants TR1, TR2, TR3, and TR4 containing the fused coxI exons were cloned into Bluescript vectors. Figure 59 shows the sequence analysis from each type of revertant and the relevant sequence of the splice junctions. In revertants TR1 and TR4 intron aI5g has been precisely deleted. Similarly, in revertants TR3 and TR2 both aI5B and aI5g have been cleanly deleted.

It is possible that not all intron deletion events are clean and that imprecise excision of the intron could occur. If so, these rare events will be missed using the Southern blot analysis described above and could only be detected by direct DNA sequencing of each revertant. The possibility that imprecise excisions might be present among the TS revertants was tested and appears remote.

Often revertants of respiratory deficient strains can be obtained easily on glucose solid medium by incubating colonies of mutant strains for about ten days and isolating revertant sectors that overgrow the surface of
the colony as bumps or papillae. This strategy can be enhanced by including an ade-2 auxotrophy in the strain; mutants yield tan colonies while revertants are often observed as red papillae. TS revertants are usually not red so this approach enhances the efficiency of screening for TR revertants.

Revertants of strain W303/C2116 (an ade-2 nuclear background) which arose as red colonies on glucose containing medium were also examined by the Southern blot analysis described above. Figure 60 shows a set of 23 TR revertants; nearly all retain the 6.0 kb fragment indicating no loss of introns. However, two of the TR revertants resulted from intron excision events: one lost aI5g (TR7) and one lost aI5β and aI5g (TR23). Thus, there appears to be a different frequency of reversion by intron deletion on fermentable versus non-fermentable carbon sources. Hence all further analyses were conducted with revertants isolated from YG media.

From the control experiments described above I concluded that examination of TR revertants of mutant C2116 isolated from YG medium is a suitable assay for the capacity of C2116 to revert by intron excision. Using this assay I tested whether reverse transcriptase is responsible for these reversion events, and whether introns aI1 and aI2 are the sources of such a reverse
transcriptase activity.

III.C.4. Intron subtraction of aI1 and aI2 from the mutant C2116

To assess whether coxl introns aI1 and aI2 are necessary for aI5g excision (sources of reverse transcriptase), I designed derivatives of C2116 lacking one or both of these introns. The structure of the desired strains is shown in Figure 61.

The strategy to construct first a derivative of C2116 devoid of aI1 and aI2 is shown in Figure 62. Petite derivatives of an isolate of ID41-6/161 which is deleted for introns aI1 and aI2 (see Table 1, and section I.I.1.b of Introduction), were isolated by ethidium bromide mutagenesis. The precise deletion of introns aI1 and aI2 in this strain was verified by DNA sequence analysis conducted by P. Q. Anziano (unpublished). Petites that retain the upstream portion of the coxl gene were mapped by standard petite deletion mapping (R. Dietrich, unpublished data). Petites that were capable of restoring mit- mutations in aI1 and aI3a, but incapable of restoring a coxl deletion (from the promotor through the 5' half of intron aI5B) were isolated for these constructions. Using karl mediated cytoduction, the desired petite, 159p-, was
moved into a nuclear background carrying resistance to canavanine among its nuclear markers [W303/p°(ade, his, ura, trp, leu)]. The petite, W303/159p- was then mated to the C2116 mitochondrial mutant residing in an α-kar1 nuclear background. Dilutions of the mated cells were spread on minimal medium supplemented with the appropriate amino acids plus canavanine to select for the W303 nucleus. On that medium, since canavanine resistance is a recessive allele, only petites, cytoductants, and the desired recombinants can grow. Large colonies, readily distinguished from petite colonies, were isolated and tested genetically for the presence of the C2116 mutation by petite mapping. C2116 is restored by the coxl deletion described above, but not by petites containing only the 5' half of the coxl gene. The mtDNA of those candidates that contained the C2116 mutation was examined in the coxl region by Southern analysis to distinguish recombinants which have lost aI1 and aI2 by recombination with the 159p- from simple cytoductants. The desired W303/ΔaI1, aI2-C2116 construct was isolated and shown to contain only introns aI3a, aI4a, aI5a, aI5β and a mutant aI5g.

Figure 63A shows the Southern blot strategy used to identify the desired recombinants. In HpaII/BamHI digests, C2116 generates a 3.6 kb fragment when hybridized
to an oligonucleotide probe complementary to a sequence in
*coxI* exon 1. The 159p-, in which introns aI1 and aI2 have
been deleted, generates a 1.6 kb fragment that hybridizes
to the probe. Figure 63B, shows the Southern blot data
for these strains as well as a strain isolated from the
selective medium as a possible recombinant. The strain
in lane 3 was tested genetically and found to contain the
C2116 mutation in aI5g. Also, as shown in Figure 63B,
lane 3, it lacks introns aI1 and aI2 since it contains a
1.6 kb HpaII/BamHI fragment that co-migrates with the
petite fragment. Lane 4 shows a cytoductant of C2116.

W303/ΔaI1, aI2-C2116 and the related strain
(W303/C2116) retaining introns aI1 and aI2 were then
assayed for their ability to revert to growth on glycerol
medium by excision of the mutant intron aI5g. Roughly 10^9
cells were spread on each of 30 YG plates and incubated
for 10 days at 30°C. Both temperature resistant and
sensitive revertants were examined by analyzing crude DNA
preparations from revertants by Southern blot analysis
using the same strategy shown in Figure 56.

In a set of 33 revertants isolated from the W303/Δ
aI1, aI2-C2116, 25 of 25 TR revertants still contained
aI5g. Furthermore, 8 of 8 TS revertants still contained
aI5g. Figure 64 shows BamHI/EcoRI digests used in
Southern analysis with the indicated probe. All the
revertants (both TR and TS) contain the same 6.0 kb fragment indicating that no introns have been excised.

From the above analysis, there appears to be a relationship between the ability of mutants to revert by excision of the mutant intron, and the presence of intron aI1 and aI2 which are potential sources of reverse transcriptase. In strain W303/C2116, 25 out of 25 TR revertants still contained the mutant aI5g (Figure 64), while in strain W303/AaI1,aI2-C2116 only ten out of 36 TR revertants still contained aI5g (Figure 56). In a 2 X 2 Chi Square contingency table, the probability that these sets of revertants are not significantly different is much less than .005.

III.C.5. CoxI intron 1 is sufficient for intron excision

To determine whether aI1, aI2 or both introns are needed for intron excision, construction of more strains was required. Each contains the C2116 mutation and one has aI1 without aI2, and the other has aI2 without aI1. The structures of these coxI gene constructs are shown in Figure 61, line 3. Such constructs were accomplished by an analogous strategy to the one described in the previous section.

The mitochondrial genome of S. diastaticus, which
contains introns aI1, aI3a, aI4a, aI4b, aI5b and aI5g, was moved into the W303 nuclear background by karl mediated cytoduction (see Table 2, line 11). Petites which retain the coxI region were isolated following ethidium bromide mutagenesis of that strain. Petites capable of restoring a large coxI deletion, but not a cob deletion were used for the construction. The strategy described in the previous section utilizing selection on canavanine containing media was used to isolate potential recombinants between a W303/diastaticus petite and the karl/C2116 mitochondrial mutants. Since the petites were capable of restoring the C2116 mutation, four types of colonies were capable of growth on the selective medium: petites, cytoductants, recombinants (with C2116), and recombinants (restored for C2116). Recombinants that were positive for growth on YG medium were eliminated from the set of candidates but served as a positive control that the cross was a suitable one. Candidates were screened genetically for the presence of the C2116 mutation and then physically for the configuration of the coxI gene.

Figure 65A shows the strategy for Southern blot analysis to establish the intron configuration of the W303/ΔaI2-C2116 candidates. In HpaII/BamHI digests, C2116 generates a 3.6 kb fragment when hybridized to an oligonucleotide complementary to sequences in the first
exon of the \textit{coxl} gene. The \textit{S. diastaticus} petite generates a 4.1 kb fragment since it lacks aI2 and hence the HpaII site present in that intron. Figure 65B shows the Southern blot data to confirm the configuration of the constructs. Lanes 1 and 2 contain DNA from C2116 and the diastaticus petite, respectively. Lane 3 contains DNA from a candidate that was found to contain the C2116 mutation in aI5g. This candidate lacks intron aI2 since it co-migrates with the the HpaII/BamHI fragment from the diastaticus petite. Lane 4 contains DNA from a cytoductant.

Recombinants containing aI1, aI3a, aI4a, aI5a, aI5B and a mutant aI5g (termed W303/ΔaI2-C2116) were then plated at a high cell density ($10^9$ cells/plate) on YG solid medium to isolate revertants. After incubation at 30°C for approximately 10 days, many TR revertants were present and some were analyzed by Southern analysis of BamHI/EcoRI fragments using the same oligonucleotide probe as above. Figure 66 shows the Southern blot from a set of 34 TR revertants examined. 16 TR revertants still contained aI5g. Only three TR revertants were deleted for both aI5B and aI5g while 12 TR revertants were deleted for aI5g alone. Thus the presence of aI1 alone is sufficient for reversion by intron excision. This result rules out the hypotheses that aI2 alone is responsible and that both
all and aI2 are required. The possible significance of the proportion of single to double intron popouts in the various intron configurations and the mechanisms for such events will be addressed in the Discussion (Chapter V). Pending experiments aimed at better defining this new experimental system are also described in the Discussion.
CHAPTER IV

DISCUSSION

IV.A. GC cluster conversion in the varl gene

The yeast mitochondrial genome contains an assortment of G+C-rich clusters ranging from 30 to 60 bp. Most are distributed throughout the intergenic regions but a few are located in the coding sequences of genes (de Zamaroczy and Bernardi, 1986). One G+C-rich cluster, the "a" element of the varl gene, has been shown to engage in a specialized recombination process in which alleles lacking that sequence acquire it in crosses by unidirectional gene conversion (Strausberg et al., 1978; Strausberg and Butow, 1981). The present study shows that a G+C-rich cluster (the common cluster of the varl gene) identical in sequence, but inverted in orientation relative to the target sequence, is likewise inserted into the varl alleles that lack it by unidirectional gene conversion.

IV.A.1. GC clusters as recombinogenic elements

117
GC clusters are known to engage in a variety of other recombination phenomena in the mitochondrial genome. Preferential recombination between GC clusters can lead to the integration of a petite genome into a wild type mitochondrial genome (Dieckman and Gandy, 1987) or a mit-mitochondrial genome (A. Morawiec, unpublished data). Furthermore, spontaneous petite genomes are often generated via recombination between G+C-rich clusters (de Zamaroczy et al., 1983). And finally at least one mit-deletion resulted from recombination between two GC clusters (Conrad-Webb et al., in preparation). Studies by Weiller (1987) and de Zamaroczy et al., (1986) have shown that G+C-rich clusters fall into specific classes that share identical flanking sequences, or target sites. Furthermore, possible nucleotide duplications are present at these target sites, a hallmark property of mobile genetic elements.

IV.A.2. Comparison between GC cluster conversion and intron conversion

The only well-characterized asymmetric gene conversion phenomenon in yeast mitochondria besides GC cluster conversion involves two mobile group I introns. G+C-rich cluster and intron conversion have appeared
similar since they both are unidirectional gene conversion events in which a "short" allele acquires sequence from a "long" allele. However, they are independent of each other and it is clear from the experiments described in Chapter III.A. that they have different molecular mechanisms.

The process of conversion of the optional introns omega (in the 21S rRNA gene) and aI4a (in the coxI gene) has been closely examined and resembles mating type interconversion in the yeast nucleus (Strathern et al., 1982). Intron conversion events have a clear dependence on mitochondrial translation: a site-specific, intron-encoded endonuclease mediates conversion of the optional introns by catalyzing a double-strand break in recipient molecules near the site of intron insertion (Zinn and Butow, 1985; Macreadie et al., 1985; Wenzlau et al., 1989). However, conversion of GC clusters occurs roughly as frequently in crosses between petite strains, which are devoid of mitochondrial translation, as in wild type crosses (Zinn et al., 1988; and this thesis). Thus any site-specific endonuclease or other protein required for this recombination process is encoded in the nucleus. GC cluster conversion also differs from intron transmission quantitatively in the frequency with which the recombinants appear among the diploid progeny. In
crosses, the optional "a" GC cluster in the varl gene inserts into only about 30% of the recipient alleles whereas intron conversion is nearly quantitative (reviewed by Butow et al., 1985).

IV.A.3. Co-conversion does not accompany common cluster conversion

Replicative repair of the double-strand breaks in recipient genomes likely results in another hallmark of intron conversion, flanking marker co-conversion. For both omega and aI4a, markers on either side of the transmitted intron are co-converted to the recipient, intronless, alleles. Flanking sequences show a gradient of co-conversion such that the frequency of co-conversion of markers is inversely related to their distance from the acquired intron (Dujon et al., 1986; Zinn and Butow, 1985). Markers as distant as 700-1000 bp from the omega intron are co-converted at a high frequency (Jacquier and Dujon, 1985).

The structure of the recipient varl allele used in this study, the varl gene of S. capensis, permitted me to test whether conversion of G+C-rich clusters is accompanied by co-conversion of sequences flanking the transmitted GC cluster. No flanking marker co-conversion
accompanying conversion of the common cluster was
detected. Markers as close as 219 bp 5' and 209 bp 3' of
the "a" and common clusters are not efficiently co-
converted in crosses, at least to the limits of detection
by Southern analysis. Furthermore, in the cross between
the universal recipient S. capensis allele ("cc-a-") and
the double donor, 5DIT ("cc+a+") allele, each of the GC
clusters can be considered a flanking marker for the
other. Although they are separated by only 157 bp, the
clusters clearly insert into the recipient allele
independently. Several recombinant isolates recovered
from the cross contained only one or the other GC
cluster. Although some double convertants which acquired
both the "a" element and the common cluster were
identified, this class of recombinants is clearly the
minority. Furthermore, such double recombinants retain
the S. capensis flanking markers, indicating that no (or
minimal) flanking marker co-conversion accompanies
conversion of the GC clusters. These recombinants likely
arise by sequential GC cluster insertion events.

Thus, these new studies of GC cluster conversion in
the var1 gene provide important new data showing that the
mechanism of GC cluster conversion is almost certainly
different from that of intron conversion. Also, I found
that extensive sequence homology appears to play a role in
conversion site selection. And finally, both orientations of the GC cluster relative to its A+T-rich "target" sequence are compatible with conversion activity.

IV.A.4. The role of double strand breaks in GC cluster conversion

Since double strand breaks are known to enhance the frequency of homologous recombination in yeast, Zinn et al. (1988) examined the region of the varl gene for the presence of double strand breaks that might initiate GC cluster conversion. Double strand breaks were detected at the boundaries of the common cluster and "a" G+C-rich elements in vegetatively growing cells. However, no double strand cleavages were observed in "a-" recipient cells. Since G+C-rich cluster conversion occurs in petite mutants and is independent of mitochondrial translation, nuclear encoded endonucleases must be responsible for those double strand cleavages. However, those double-strand breaks have not been shown to play a direct role in the transmission of the GC clusters but their discovery provides some insight to possible mechanisms.

The presence of double stranded breaks in the donor molecules are consistent with "cut and paste" models for
GC cluster conversion. First proposed for transposable elements such as Tn10 or Tn7, these models suggest that the element is excised from the donor and conservatively donated to the recipient (reviewed by Grindley and Reed, 1985). This process is non-replicative and any replication that occurs is to fill small gaps in the recipient. Should subsequent studies confirm that GC cluster conversion resembles conservative transposition, then this will be the first case where host specific rather than element specific enzymes direct a transposition process. However, it is important to note that some transposases can function in trans on transposons that have lost the ability to encode the cutting function.

This analogy to "cut-and-paste" transposition mechanisms forces a reassessment of some of the published data on var1 gene GC cluster conversion. For example, the original reports stress that recipient var1 alleles are "lost", being converted to "a+" alleles while donor alleles are conserved, relative to flanking drug-resistance markers. It is unlikely that the estimates of ratios of parental and recombinant var1 genes are sufficiently accurate to resolve this problem.

If GC clusters are, indeed, excised from donor var1 genes, what then happens to the donor parental genomes?
If they were destroyed, then the loss of donor genomes equal in number to recombinant genomes, would have been obvious (assuming 30% of the progeny are recombinant). One could explain the retention of flanking markers by marker rescue, but the persistence of donor var1 alleles (linked to appropriate drug markers) still requires explanation. Perhaps genomes that have donated a GC cluster are destroyed, but are subsequently replaced by extra replication of unreacted donor genomes. It is generally believed that the amount of mtDNA per cell is fairly rigidly controlled and it is well established that yeast mtDNA replicates throughout the cell cycle. It is not clear, however, how donor genomes might be selected preferentially for extra replication cycles.

Another problem with this analogy to conservative transposition mechanisms is my finding that the orientation of the cluster relative to its target sequence is retained. If the cluster is first excised and then inserted into a suitable target site in the recipient genome, then one might expect that insertion would be in both orientations. It should be noted that Tn7 transposes in a highly site and orientation specific manner (N. Craig, personal communication). It is well-established that most conservative transpositions are concerted events involving interaction between both donor and recipient
prior to the excision event. In the varl gene it appears that extensive flanking homology is important for conversion since in double donor experiments only the target site that is at the same site as the donor element is used. So, interaction between donor and recipient varl genes may predispose this process to be site and orientation specific.

IV.A.5. A candidate enzyme involved in GC cluster conversion

While the nuclear-encoded enzyme responsible for the endonucleolytic cleavages flanking the GC clusters in the varl gene is unknown, a potential candidate is topoisomerase II. Interestingly, the 3' boundaries of the "a" element and common cluster and a portion of their flanking sequences closely resembles the topoisomerase II cleavage consensus sequence for chicken (Zinn et al., 1988; Spitzner and Muller, 1988). Since a mitochondrial topoisomerase has not been identified, it is possible that the nuclear topoisomerase II functions in the mitochondria to allow various mtDNA manipulations. Type II topoisomerases have been well characterized and are known to act by generating a double-strand break in the DNA (Brown and Cozarelli, 1979). Several temperature
sensitive nuclear mutations defective in this enzyme (top2 mutations) have been characterized in yeast (DiNardo et al., 1984; Holm et al., 1985). Also, recently a top3 mutant has been identified in yeast which appears to encode another topoisomerase which is analogous to the gyrase function of E. coli (Wallis et al., 1989). To test whether the nuclear topoisomerase II enzyme is responsible for generating the cleavages flanking GC clusters in the var1 gene, appropriate strains could be constructed to assay for the presence of these double strand cleavages flanking GC clusters in the var1 gene in the top2 nuclear background at the restrictive temperature.

IV.A.6. New strategies to study GC cluster conversion

The development of methods for mitochondrial transformation allows the design of experiments to dissect the mechanism of GC cluster conversion (Johnston et al., 1988). The precise sequence requirements for the target site for GC cluster conversion may be closely examined by distorting the site via mitochondrial transformation. Subtle alterations in the coding region of the var1 gene will likely be compatible with respiration since it can tolerate such a wide variety of insertions in its reading frame. If not, the identical target sequence between the
tRNA-ser gene and the olil gene may be exploited for dissection of the target sequence without phenotypic consequence. A GC cluster identical to those in the varl gene resides within this sequence in some strains. Furthermore, site-directed mutations may be introduced into GC cluster donor strains using mitochondrial transformation. GC cluster conversion upon mating with recipient strains and the presence of double stranded cuts flanking GC clusters could then be assayed in such strains. Mitochondrial transformation also allows the design of flanking markers much closer to the mobile GC clusters to accurately measure the degree (if any) of flanking marker co-conversion that accompanies GC cluster conversion.

The enzyme that introduces the double stranded cuts flanking GC clusters may be identified using extracts of yeast mitochondria. End-labeled, cloned substrates containing the GC clusters and flanking sequences could be incubated with mitochondrial extracts, and the position of the double stranded cuts could be measured by electrophoresing the products of the reaction alongside appropriate sequencing ladders.

Since nuclear-encoded enzymes are responsible for GC cluster conversion it is possible that the process may be active in nuclear DNA. Although this may be unlikely, it
can be tested rather easily; since a positive result would likely yield the preferred system for further study of the process, these experiments are worth doing. Using reverse genetics it is quite easy to incorporate a recipient var1 gene within a yeast chromosome. Then, a multicopy plasmid containing a donor var1 allele would be provided and chromosomal alleles scored using DNA blots after some generations of selection for the plasmid. If the process occurs, then it is obvious how this new system would be exploited to define essential sequences for conversion, to evaluate the requirement for flanking homology, etc.

Most important, however, would be studies aimed at obtaining mutants that have modified in the ability to convert GC clusters in nuclear DNA. For example, if an active target site can be incorporated within a reporter gene (e.g., ura3+) on a chromosome and a ura3− allele formed by inclusion of the cluster on a plasmid then a system that allows selection for increased or decreased incidence of cluster conversion would result. That is, such haploid cells would be ura+ but they would yield ura− progeny by conversion at some frequency. Such convertants would be FOA resistant and colonies could be screened for increased or decreased frequencies of FOAR cells. Alternatively, a system could be established in which a frameshift in a chromosomal lacZ gene is cured by
GC cluster conversion from a donor plasmid. In this case, parental strains would yield some incidence of blue spots on colonies and derivatives with fewer or more spots screened for quite directly.

IV.B. Mobile introns

IV.B.1. The aI4a intron-encoded protein promotes intron transmission

An unusual feature of yeast mitochondrial introns is that many contain open reading frames. The protein products of the intron reading frames that have been examined either promote splicing, or intron propagation. In the studies presented here, introns in the coxI gene were surveyed for their ability to engage in gene conversion. Intron aI4a is efficiently transmitted to genes lacking it in crosses. AI4a conversion is dependent upon the integrity of the open reading frame within the intron. Mutations that truncate the intron encoded protein abolish intron transmission. The experiments in Chapter III.B. support the view that the intron encodes a site-specific endonuclease which recognizes defined exon sequences and cleaves them, thus initiating the process of gene conversion. The aI4a
endonuclease activity has been detected in extracts of mitochondria prepared from mutant strains that overproduce the intron-encoded protein (Wenzlau et al., 1989). Studies conducted by Delahodde et al. (1989) indicate that the aI4a intron reading frame expressed in *E. coli* is sufficient for the endonuclease activity seen *in vivo* and *in vitro* in the experiments discussed here.

IV.B.2. Intron conversion is not limited to yeast mitochondria

IV.B.2.a. Comparison of features of intron conversion in different systems

Conversion of group I introns by site-specific recombination appears to be a common mechanism for the propagation of group I introns (reviewed by Lambowitz, 1989). While my studies of aI4a conversion were being prepared for publication, we learned of research using other experimental systems which suggested that intron proliferation may be an intrinsic property of many group I introns. The recent demonstration of intron propagation via intron-encoded endonucleases in the *sunY* and *thymidylate synthase* (td) genes of phages T2 and T4, and I3 of the nuclear rRNA gene in Physarum supports this
notion. The extent of characterization for each system varies, but the salient features of the phenomenon are consistent among the various mobile introns and all are similar to the "prototype" convertible intron, the omega intron in the 21S rRNA gene of yeast mitochondria. For five of the introns which have been shown to be mobile genetic elements, intron movement involves unidirectional gene conversion events in which intronless genes acquire the intron in crosses (reviewed by Lambowitz, 1989).

Conversion in each case is associated with a double strand cleavage of recipient DNA molecules near the site of intron insertion. These double strand cleavages have been monitored both in vivo and in vitro. The aI4a endonuclease has been characterized from mitochondrial lysates prepared from mutants which naturally overproduce the aI4a endonuclease (see section III.B.13.). Like the omega intron reading frame, the aI4a intron reading frame has also been engineered in E. coli and has been shown to encode a site-specific endonuclease (Delahodde et al., 1989). The endonuclease activities of the large rRNA intron reading frame in Physarum and the td intron reading frame from T4 were demonstrated utilizing in vitro translation systems.

Exonucleolytic degradation of the cleaved DNA has been suggested to result in the observed co-conversion of
flanking DNA sequences. In the double strand break repair model for recombination, the 3' overhangs generated from the double strand cuts serve as primers for DNA synthesis using the intron containing allele as a template. Depending upon the extent of gapping (or exonucleolytic degradation), flanking sequences on the (donor) template show a gradient of co-conversion which decreases with increasing distance from the intron. For the $w^+$ intron in the 21S rRNA gene, sequences extending 700-1000 bp on either side of the intron are co-converted. Although co-conversion clearly accompanies conversion of intron aI4a in the CoxI gene, the limits of co-conversion have yet to be determined. In the new systems where intron transmission has been documented, co-conversion is bidirectional and its frequency decreases with increasing distance from the site of intron insertion.

IV.B.2.b. Intron mobility in T-even phage

Intron mobility has been demonstrated in the thymidylate synthase (td) and sunY genes in bacteriophages T2 and T4. High frequency intron transfer (HFIT) occurs when transformants harboring cloned intron-containing td or sunY genes are infected with intronless phage. Both genes encode ORFs: in crosses between intronless phage
and transformants containing \textit{sunY} and \textit{td} genes in which the ORF has been deleted, HFIT is nearly abolished.

The ORF product of the \textit{td} intron has been expressed in an \textit{in vitro} translation system and used in endonuclease assays. An endonuclease activity induces the same double strand break in a plasmid containing the "homing site" for the intron as that observed \textit{in vivo}. However, this enzyme appears to induce a cleavage 15 to 25 bp from the recognition sequence similar to type I restriction endonucleases (Bell-Pedersen et al., in press). The specificity of the ORF endonuclease has not yet been determined although a 21 bp deletion of the \textit{td} gene "homing site" abolishes HFIT.

Both the \textit{td} and \textit{sunY} introns are inserted precisely at their cognate sites. Extending the exon homology flanking the intron has been shown to increase the frequency of intron transfer. This implies a mechanism similar to that of conversion of the omega intron in the 21S rRNA gene in yeast mitochondria which is thought to proceed by double strand break repair. A requirement of the model is flanking marker co-conversion which has been demonstrated for intron conversion in the T-even phages (Quirk et al., 1989). A set of exon point mutations was generated by hydroxylamine mutagenesis to create flanking markers. High frequency conversion of the intron along
with co-conversion of flanking sequence restores the mutant phage to wildtype (thy+) depending on the degree of co-conversion. Within 50 bp of the homing site, 82-98% of the progeny co-converted exon sequences both 5' and 3' of the intron insertion site. Mutations more distant from the target site (between 56 and 441 bp), were only restored by roughly 50% of the progeny (Bell-Pedersen et al., in press).

IV.B.2.c. A mobile intron in *Physarum polycephalum*

Intron 3, in the rDNA of *Physarum Polycephalum* has also been demonstrated to be a mobile group I intron in matings between amoebae containing and lacking introns. Using restriction polymorphisms flanking intron 3, co-conversion of flanking sequences was demonstrated. A polymorphism 500 bp downstream of the convertible intron accompanies transmission of the intron in crosses; however, a polymorphism 4 kb away does not (Muscarella and Vogt, 1989). This intron also contains an ORF. Intron 3 containing RNA synthesized from a phage T7 promoter is capable of self-splicing *in vitro*. When the excised intron was employed in *in vitro* translations with reticulocyte lysates, a 15 kilodalton (kd) peptide was synthesized. Upon incubation of the 15 kd protein with a
fragment containing the intron 3 insertion site, the substrate was cleaved. Furthermore, a double-strand break near the intron 3 insertion site of recipient DNA molecules has also been demonstrated in vivo (Muscarella and Vogt, 1989).

IV.B.2.d. Gene conversion in other systems

Other examples of unidirectional gene conversion involving group I introns have been described in the 
Chlamydomonas eugametos and moewusii chloroplast DNA genomes (Turmel et al., 1987; Lemieux and Lee, 1987). Studies of polymorphic loci in the 23S rDNA region of the inverted repeat sequence of the chloroplast DNA revealed a polymorphic locus which is inherited consistently from one parent among recombinants examined from interspecific hybrids between C. eugametos and C. moewusii (Lemieux and Lee, 1987). This polymorphic locus resides within group I intron 5 and contains a long open reading frame.

The mitochondrial genome of Aspergillus nidulans var. echinulatus contains six insertions relative to that of Aspergillus nidulans. Employing polyethylene glycol induced protoplast fusions, recombination of mtDNA is possible between these subspecies of A. nidulans. The resulting hybrid genomes invariably contained inserts 1, 2
and 3, presumably as a consequence of high frequency asymmetric gene conversion (Earl et al., 1981). Also, in crosses between *Chlamydomonas smithii* and *C. reinhardtii*, a 1.0 kb insert which is thought to be an intron in the mitochondrial cytochrome b gene is transferred unidirectionally to the "intron absent" *C. reinhardtii* cytochrome b allele. The 1.0 kb insert contains an open reading frame capable of encoding a protein very similar to the *fit1* endonuclease of the omega intron (summarized by Dujon and Belcour, 1989).

Two mitochondrial genomes of *Coprinus cinereus* contain alternative 1.2 kb insertions ("H" and "J") in the cytochrome oxidase subunit I gene. Both the "H" and "J" insertions show biased transmission in crosses between two isolates of *Coprinus cinereus* having alternative forms of the mitochondrial genome. It is likely that these two inserts promote conversion events which lead to their own insertion into "H" and "J" "homing sites" (Economou et al., 1987).

In studies of the interaction between mtDNAs of *Neurospora crassa*, two insertions were determined to be sites of high frequency unidirectional gene conversion. These insertions (1200 bp and 50 bp) are spread through mtDNA populations in heteroplasmsons (Manella and Lambowitz, 1979). Thus high-frequency unidirectional gene
conversion appears to occur in several distinct systems ranging from bacteriophage; organelle genomes of yeast, Chlamydomonas, Aspergillus, Coprinus, and Neurospora; and the nuclear genes of Physarum. The convertible sequences described above are probably all group I introns.

IV.B.3. Cleavage by the intron-encoded enzymes

The ability to study the endonuclease activity of the yeast mitochondrial, Physarum, and T-even phage intron encoded proteins in vitro and/or in E. coli has revealed several features of these enzymes. For those cleavages that have been precisely mapped, they appear to be 4 bp 3' hydroxyl overhangs (see Figure 67). The enzymes are similar to a nuclear encoded endonuclease of yeast, HO, the initiator of mating type interconversion and Scel, a yeast nuclear enzyme whose function is unknown (Kostricken et al., 1985). Figure 67 illustrates that the HO and the aI4a endonucleases cleave downstream of the site of insertion, while the fitl endonuclease cleaves symmetrically about the site of insertion. The td and sunY endonucleases are different since they are believed to yield blunt ends after cleavage and cleave about 20 bp upstream and downstream, respectively, of the intron insertion site.
The reading frames for the intron-encoded endonucleases appear to share only short amino acid sequences in common. The most prominent common sequence element among most of these intron open reading frames is the LAGLI-DADG motif, first identified in group I intron reading frames by Hensgens et al. (1983). The mobile group I introns of phage T4 do not contain the LAGLI-DADG motif while the open reading frame of intron four in the yeast mitochondrial cob gene contains the LAGLI-DADG motif but functions only as a maturase. Although the HO endonuclease is not encoded by an intron ORF, it does contain the peptide LAGLIDSDG, closely resembling the LAGLI-DADG motif. Thus, it is not clear whether this sequence element specifies a functionally important domain for these endonucleases.

IV.B.4. The recognition sequences for intron-encoded endonucleases

Figure 67 illustrates that each of these convertible introns has a distinct intron insertion site, and probably distinct endonuclease recognition sequences; they share no core consensus sequence. However, a portion of the w+ substrate resembles that of the HO substrate. The recognition sequences which have been determined for some
of these endonucleases appear to be quite large: 16 bp for the HO endonuclease and 18 bp for the w⁺ and aI4α endonucleases. Only one recognition site for the aI4α and w⁺ endonucleases is present in the mitochondrial genome. Interestingly, there are at least three copies of the cleavage site for the HO endonuclease, but only that at the MAT locus is accessible to the HO enzyme in wild type yeast strains.

Extensive mutagenesis of the fitl recognition sequence has demonstrated that at least one substitution at every position of the 18 bp recognition sequence influences cleavage of the substrate (Colleaux et al., 1988). Thus, the fitl endonuclease can tolerate little alteration of its recognition sequence. The constraints on the recognition sequence for the aI4α encoded endonuclease may be less stringent. Although originally the recognition site was broadly estimated to be 18 bp, mutagenesis studies indicate that no more than 14 bp are necessary for cleavage (see Figure 53). Furthermore, several internal substitutions can be tolerated at each of several positions. Substitutions at a minority of positions (those indicated by the boxes in Figure 53) interfere with cleavage of the substrate implying that the enzyme may only recognize bases at certain positions. The presence of either purines or pyrimidines at other
positions may be required for efficient cleavage of the substrate. Thus, the sequence constraints for cleavage and/or recognition by the $aI4a$ endonuclease may not be entirely rigid.

Deletion and mutagenesis studies have determined that the minimal recognition site for the HO endonuclease is 16 bp. However, several shorter substrates and mutant substrates are cleaved at low levels \textit{in vitro} with purified HO enzyme. Also, a 24 bp recognition sequence is capable of initiating homologous recombination when inserted into new locations in the yeast genome while a 20 bp sequence does not (Nickoloff et al., 1986). The discrepancy between the HO enzyme specificity \textit{in vitro} and \textit{in vivo} may be due to the lack of cis and/or trans-acting factors \textit{in vitro} which lower the enzyme's specificity. Alternatively, the partially purified HO enzyme may be unstable and the breakdown product may have a lower specificity than the native protein. Such inconsistencies are partially alleviated in studies with crude mitochondrial extracts since any additional trans-acting factors required for cleavage are likely to be present. Furthermore, analogous experiments to those conducted with HO may now be pursued since mitochondrial transformation is possible: the consequence of mutations in the $aI4a$ endonuclease recognition sequence which interfere with
cleavage *in vitro* may be studied with crosses *in vivo*.

Interestingly, all three of these recognition sites are situated asymmetrically relative to the site of sequence insertion (see Figure 67). In each case, the insertion site is located closer to the left boundary of the recognition sequence than to the right boundary. In fact, the donor substrate contains a large portion of the recognition sequence since the majority of the endonuclease recognition site is present in the 3' exon. The *in vitro* mutagenesis studies indicate that, at most, only 3 bp of the 5' exon are essential for cleavage.

IV.B.5. Intron encoded endonucleases as restriction enzymes

Although these enzymes appear broadly similar, given the different recognition sites and sequence constraints of these recognition sites, each has a unique specificity. Because the sequence constraints for the *fit1* endonuclease are quite rigid, it is unlikely that this recognition sequence appears frequently in a given genome. Perhaps this is why the enzyme, when expressed in *E. coli*, is not toxic to cells. However, the sequence constraints for the aI4a endonuclease appear more lenient. As would be expected, the aI4a enzyme, when
expressed in *E. coli*, is toxic.

Both the fit1 and aI4a endonucleases have different optimal in vitro reaction conditions as well. Figure 68 illustrates a comparison of the optimized in vitro reaction conditions for the fit1 and aI4a endonucleases. Both endonucleases require divalent cations; however, only the aI4a endonuclease is stimulated by monovalent ions depending upon the salt concentration. The aI4a endonuclease shows a sharp pH optimum above pH 7.0 whereas the fit1 endonuclease operates in a wide pH range. The optimized in vitro reaction conditions for the HO endonuclease are 100 mM NaCl, 5 mM DTT, and 10 mM MgCl2; however, it has not been reported whether these conditions are stringent. Thus, although these enzymes share general features in common, they have idiosyncracies similar to those seen with type II restriction endonucleases.

Given the large and distinct recognition sequences of these endonucleases they may eventually become useful as genomic mapping reagents. Very few restriction enzymes with relatively long recognition sites are currently available. NotI, which has an 8 bp recognition site, has been used extensively. However, because its recognition site contains only G's and C's, NotI sites tend to be concentrated in HET islands in the human genome (reviewed by Bird, 1986). It is also sensitive to methylation fo
that DNAs from different individuals or even different tissues from the same individual can yield different patterns. SfiI, which has a 13 bp recognition sequence in which the 5 central bases are ambiguous, is also effective for mapping human DNA since it recognizes very infrequent sequences. However, because mutagenesis studies of the recognition sequences for the HO and intron encoded endonucleases have indicated that several substitutions show reduced, but appreciable digestion by the enzymes, these endonucleases may partially cut some sites. It is possible that even if the aI4a, fitl, and HO endonucleases have cleavage sites that are present too rarely in the human genome to be used for mapping, these sites could be integrated into the genome to function as landmark sites for mapping with other enzymes.

IV.B.6. Expression of the intron-encoded endonucleases

These intron encoded proteins have evolved very distinct modes of expression. Figure 69 shows schematically how several of these intron encoded endonucleases are expressed. The aI4 endonuclease in the coxI gene is translated from an unspliced pre-mRNA as a fusion protein with the upstream exons (see Figures 69 and 70). The evidence for this comes from several facts; in
cis-dominant mutants defective in splicing, many downstream introns are not spliced suggesting a fusion protein product. For a14a, the fusion protein has been identified by Hanson et al. (1982) and was found to be 59 kd. The fusion protein is then believed to be proteolytically processed to yield a set of p29 protein species. One of the p29 species is derived from the fused upstream exons, while another may function as the endonuclease (Sass, 1984). Figure 69A shows the model for expression of the fitl protein of the 21S rRNA gene intron. The fitl protein is believed to be translated from a separate mRNA species which is probably a processed form of the excised intron. In Figure 69C, the bacteriophage intron proteins are shown to be expressed independently from their own late promoters which are located within the intron. These late promoters are dependent upon the gene 55 product for transcription initiation. This temporal regulation of the expression of the intron open reading frame guarantees that translation of the intron product does not interfere with splicing since the 3' end of the reading frame overlaps with the sequences required to form the secondary structures crucial for splicing (Quirk et al., 1989). The fact that such distinct modes of expression exist for these intron-encoded endonucleases implies that the reading frames
within the introns may have been acquired separately (see section IV.C.2.).

IV.B.7. The aI4a-ORF encodes two functions

An endonuclease activity is not the first function to be assigned to the product of an open reading frame of a group I intron. Three group I introns, cob I2, cob I3, and cob I4, have been shown by extensive genetic studies to encode maturases which facilitate intron splicing (reviewed by Perlman et al., 1989; Lazowska et al., 1989). The maturase from cob I4 is required for the splicing of two introns, aI4a and bI4, while the other maturases only facilitate the splicing of the intron which encodes them. The product of the aI4a intron reading frame encodes a protein that is closely related to the cob I4 maturase; they are 70% homologous (see Anziano et al., 1982). However, the product of the aI4a intron is dispensable in wild-type cells.

The product of the aI4a intron can be activated in vivo to substitute for the cob I4 maturase by two modes. A single base change in the aI4a intron open reading frame (the mim-2 mutation) can suppress trans-acting mutations in cob I4 (Dujardin et al., 1982). Also, dominant nuclear suppressors of cob I4 trans-acting mutations have been
characterized which are dependent upon the integrity of the aI4a reading frame (the NAM2 mutants) (Dujardin et al., 1983). Thus, the aI4a intron encoded protein is the first example of an intron encoded protein that can interact productively in reactions with both DNA and RNA. By virtue of their selection scheme, the aI4a mutants which were found to abolish intron conversion (see section III.B.12.) also have a defective maturase since they are in a NAM-2 nuclear background. Thus it is not clear whether the endonuclease and maturase activities are located in a single protein domain or distinct domains.

IV.B.8. Introns as mobile genetic elements

There are many observations that suggest that introns are or have been mobile genetic elements. The highly conserved sequences and structures of group I and group II introns imply that each group arose from a common ancestor. The fact that introns are present in vastly divergent organisms suggests that they were acquired by independent insertion events. For instance, the mobile group I intron in the rRNA gene of Physarum is more than 70% homologous (particularly in the region comprising the structural elements of the intron) to the self-splicing intron in the large rRNA gene in Tetrahymena. Although
the Physarum intron is twice the size of the Tetrahymena intron, it is the closest relative to the Tetrahymena intron and they are inserted at precisely the same location in the large rRNA gene. Since the introns in both Physarum and Tetrahymena are 70% homologous and because these two organisms diverged quite early in evolution, these introns were likely acquired relatively recently in evolution, perhaps by horizontal gene transfer (Muscarella and Vogt, 1989).

Another illustration of the same intron present in diverse organisms is that of the second group I intron in the coxI gene of S. pombe which is highly homologous to group I intron 3 in the coxI gene of A. nidulans. These introns contain 70% amino acid homology between intron open reading frames. However, the A. nidulans intron contains a 37 bp GC insert flanked by 5 bp direct repeats. The fact that these introns are inserted at exactly the same location in the coxI gene and that exon sequences extending 60 bp on each side immediately adjacent to the introns are highly homologous between these two species implies lateral (or horizontal) intron transfer (Waring et al., 1984; Lang, 1984). If intron conversion were to occur in the course of lateral transfer, flanking exon sequences would be co-converted with the intron creating the observed homology between the two genes of these
species, where the rest of the coxI sequences are fairly divergent.

Also, two group I introns of the coxI gene of S. pombe are found at precisely the same locations as introns aI4a and aI5β in S. cerevisiae. Although the S. pombe introns lack intron open reading frames, they share distinct homology with the introns in the coxI gene of S. cerevisiae. Additionally, the S. pombe intron that is homologous to aI4a also shares a remarkable sequence homology to cob I4 in S. cerevisiae (Trinkl and Wolf, 1986). These highly homologous introns, cob (bI4) and coxI (aI4a), of S. cerevisiae provide an example of a similar intron at different sites in the same organism.

Moreover, there are two instances of highly homologous introns that are located in different genes in two organisms. That is, introns may insert themselves into new chromosomal locations. An intron closely related to the w+ intron in the S. cerevisiae 21S rRNA gene is present in the ATPase 9 gene of the mitochondrial DNA of Kluveromyces fragilis. The ATPase 9 gene contains the identical recognition site found in w− alleles of the 21S rRNA gene in S. cerevisiae. However, the intron present in the ATPase 9 gene in K. fragilis lacks the fit1 reading frame that is present in the w+ intron of the 21S rRNA gene in S. cerevisiae (Dujon et al., 1986).
Lambowitz (1989) has proposed one possible mechanism for horizontal intron transfer between species which involves the autonomous forms of introns such as the Podospora -senDNA (an excised group II intron) and the Neurospora mitochondrial plasmids (which share similarities to group I introns). Both -sen DNA and the Neurospora plasmids are autonomous elements that have been shown to be associated with a reverse transcriptase activity. Presumably the intron exists independently as a DNA species and then inserts itself into a compatible target site possibly by a mechanism resembling retroviral provirus integration. Site-specific integrations have been documented for the Mauriceville and Varkud plasmids (Akins et al., 1986).

Since many group I introns are capable of self-splicing, their presence would not necessarily be detrimental to expression of the gene in which they reside. The high frequency transmission of these introns reflects the exploitation of the cellular recombination apparatus to spread the intron through the population. Intron transmission is not a transposition event directed by intron-encoded products. The intron initiates recombination between sequences that share extensive flanking sequence homology and therefore appears to be restricted to exchanges between alleles which have and
lack the mobile intron.

IV.C. Transmission of group I introns by reverse splicing

There are several facts which indicate that other mechanisms for intron transmission may exist. Because of the requirement for considerable flanking homology for endonuclease mediated intron conversion, insertion into new locations should be extremely uncommon. Furthermore, not all group I introns contain potential endonuclease encoding intron reading frames. In fact, the self-splicing Tetrahymena intron in the large rRNA gene lacks an intron reading frame, yet the analogous intron that resides in the Physarum rRNA gene and does contain an endonuclease encoding reading frame is mobile. Other mechanisms for stable intron transmission may involve RNA intermediates.

Recently, reverse self-splicing of the Tetrahymena group I intron in the large rRNA gene has been demonstrated (Woodson and Cech, 1989). The reverse reaction follows an inverted pathway of the forward reaction: splicing reversal begins by base pairing interactions between the intron and exon sequences at the splice junction using the same sequence associations that facilitate splicing. Just as the 5' exon sequences are
essential for determining the 5' splice site, they are important in determining the site of intron integration. However, the 3' exon sequences are not required for the reverse reaction. The transposition of introns to new sites via reverse splicing is less restrictive than intron-endonuclease mediated conversion since fewer bases are required. Moreover, the contribution of the 5' and 3' exon sequences in these two processes differs. For instance, for the aI4a intron-endonuclease the 3' exon constitutes most of the recognition site (see Figure 51) while in the only case where reverse splicing has been demonstrated, the 5' exon has the dominant role.

Excised introns released from an mRNA can also integrate into another mRNA species by reverse splicing. Woodson and Cech (1989) have demonstrated that the intron of the Tetrahymena rRNA gene can integrate into a β-globin transcript so long as it contains the sequences that interact with the internal guide sequence of the intron. Reverse splicing is inefficient in vitro and probably in vivo, as well. Intron insertion by reverse splicing also has target sequence limitations; however, they are much more lenient than those for endonuclease-mediated intron insertion. In the case of the Tetrahymena large rRNA gene intron, the sequence requirement is only "CUCU". Thus there are many plausible sites for intron integration for
this intron. Insertion of introns by splicing reversal would ensure that the location of intron insertion is compatible with splicing so that the presence of the intron is not deleterious.

The next step for intron integration would involve reverse transcription of an mRNA into which an intron has inserted to generate a cDNA. The new intron would then be incorporated into the genome by homologous recombination. Figure 71 shows schematically the process of reverse splicing and intron integration. The feasibility of this mechanism depends upon a source of reverse transcriptase. There are at least two potential sources of reverse transcriptase activity in yeast; nuclear transposable elements (i.e. Ty elements), and the products of mitochondrial group II intron reading frames.

The open reading frames of group II introns are typically larger than those of group I introns. They also share a significant region of homology to reverse transcriptase which extends for roughly 250 amino acids in the central region of each group II reading frame. Figure 72 shows the conserved blocks of reverse transcriptase homology for coxI all. Several groups have demonstrated that mutants defective in splicing one or another mitochondrial intron are capable of reverting by clean excision of the impaired intron (Jacq et al., 1982;
Gargouri et al., 1983; Perea and Jacq, 1985; Hill et al., 1985; P. Q. Anziano, unpublished). These deletion events often involve coincident loss of one or several adjacent introns simultaneously, strongly implying recombination with a cDNA copy of mature message.

Experiments described in Chapter III.C. have tested whether the products of group II introns, aI1 and aI2, of the coxl gene are required for intron excision. The majority of TR revertants of the mutant, C2116, were shown to be excisions of either aI5g or aI5B and aI5g. This mutant intron serves as a reporter for the capacity to revert by "popping out". By constructing strains devoid of these two introns, it has been demonstrated that the mutant reporter intron is incapable of reverting by "popping out" when these particular group II introns are absent. This implicates the product of these intron reading frames in the excision process. The product of intron 1 alone is capable of supporting the excision of the reporter intron.

Analysis of the revertants that have deleted the reporter intron reveal that adjacent introns are simultaneously deleted in about 31% of the cases. In such cases the deletions are restricted to introns aI5B and aI5g. If intron deletion via recombination with a cDNA is the process responsible for intron excision, the above
products would imply a very short cDNA product, a hot spot for recombination in exons 6 and 7, or a mechanism that actively suppresses or masks recombination upstream of these exons (see below for more detail).

IV.C.1. Models for intron excision

Conceptually, there are at least two modes by which introns may be deleted from a genome: by precise excision on the DNA level or via an RNA intermediate. Recombination at the DNA level would require enzyme activities that are capable of recognizing and cleaving intron/exon junctions, and ligating them to yield fused exons. There are numerous examples of introns excising from the mitochondrial genome. The fact that more than one intron can be "popped out" in a single event would necessitate either a series of enzymes to excise the sets of introns that are commonly deleted together or a highly plastic enzyme since the intron/exon junctions differ.

A more attractive mechanism is to postulate an RNA intermediate. One can imagine a cDNA being synthesized from mature message which can recombine with the genome based on its homology to exon sequences. The exons that participate in such recombination events will dictate which introns are excised. For example, revertants of
C2116 (see Chapter III.C.) must have a crossover in the last exon of *coxI* and anywhere in the exons upstream of aI5g. Among revertants of C2116, 11 double excisions and 15 single excisions are observed. This suggests that crossovers predominate in exons 6 and 7 (of strain ID41-6/161). These exons are only 134 and 25 bp, respectively, while exons upstream are as large as 479 bp (exon 4). Despite this homology, no crossovers in the upstream exons are obtained. Several possibilities could explain these results. Exon 6 and 7 may be hot spots for recombination. Alternatively, the lack of "pop-outs" spanning aI5a and aI4a could be explained by proposing that cDNA synthesis rarely proceeds past exon 6. In that case, there would be only two regions for the required second crossover and the ratio of single and double "pop-outs" obtained would reflect the array of cDNA lengths. Finally, the absence of recombinants with introns 5' of intron aI5B excised may be readily explained by the presence of the convertible intron aI4a that is present in all of the strains used here to study intron excision.

The presence of convertible introns may be expected to alter the spectrum of products of such "reversion by intron excision" events in several ways. One, if the initial product of recombination has deleted the convertible intron it will be rapidly re-introduced. That
is, if full-length cDNAs are made so that initial recombinant would lack more than two introns, so long as intron aI4a was removed, the recombinant coxI gene would be a substrate for intron conversion to replace some or all of the lost introns. In our studies, it is clear that we selected for cells that lost at least one intron (aI5g); but if excision and conversion actually interact as suggested here, it is possible that only a minority of primary intron "pop-outs" are recovered routinely.

The other way in which intron conversion could limit the types of "popouts" obtained involves the DNA intermediates (see Figure 73). There is no obvious reason why a full-length, or nearly full-length, cDNA could not be a substrate for cleavage/conversion at its cognate site even before recombining with mtDNA. This would divide the cDNA into two half molecules. If conversion does not employ the resulting linear fragments of the cDNA, the two half molecules might participate in recombination. Although both half molecules might be equally recombinogenic, selection in the studies described here is applied for reversion of only intron aI5g and hence only for recombination with the downstream half molecule containing exons 5-8. Gap formation, proposed to play a role in conversion events, would reduce the size of exon 5 perhaps to a point where no crossovers in exon 5 would
occur. In that case, only single and double "pop-outs" would be possible. Or, if cDNA fragments could engage in conversion then they would re-acquire some introns before recombining with mtDNA.

A special strain can be constructed to test whether intron excision and mobility interact. A mutation, A3, shown in this thesis to block conversion but not splicing of aI4a, will be introduced into the aI4a intron of a strain containing the aI5g mutation C2116. The popouts obtained will include excisions of only introns aI5B and aI5g if the aI4a endonuclease does not have a role in the configuration of revertants obtained from the reporter mutant. However, if the aI4a endonuclease participates in either way described above, upstream introns will also be deleted. The "pop-outs" generated in this context will reflect the length of message that was converted to cDNA.

The intron configuration of mitochondrial genes may represent a dynamic equilibrium between two opposing intron encoded functions, endonuclease, and reverse transcriptase. The ability of introns to delete via a reverse transcriptase mediated mechanism may explain why intronless alleles of specific genes still remain. Endonuclease mediated intron transmission is a high frequency event which spreads the intron containing allele throughout a population of intron lacking alleles. If
group II introns are present in the same genome acquiring the new introns, they may provide an active mode for removing introns once they are inserted into a gene.

Even though this genetic assay, by virtue of its design, measures the loss of introns, it is likely that if a reverse transcriptase activity is present in yeast mitochondria it could promote other rearrangements including recombination with cDNAs that still contain introns. Thus, intron insertion and intron excision mediated by an intron encoded reverse transcriptase may interact.

IV.C.2. Evolution of mitochondrial introns

Current theories of mitochondrial intron evolution speculate that ancestral introns were self-splicing and subsequently gained intron open reading frames, possibly by transposition events (Cech, 1985). Several examples of analogous introns which have and lack intron reading frames have been documented. For instance, the Tetrahymena rRNA intron and the mobile intron in the rRNA gene of Physarum are closely related, but the intron in Tetrahymena lacks the intron reading frame. Two introns in S. pombe in the coxI gene are analogous to introns aI4a and aI5B in the coxI gene of S. cerevisiae. However, the
intron in *S. pombe* analogous to aI4a is missing the 948 bp intron open reading frame making it the shortest group I intron in lower eukaryotes. Similarly, the *S. pombe* cognate intron of aI5B lacks the 1536 bp intron reading frame found in *S. cerevisiae* (Trinkl and Wolf, 1986). Probably the best evidence for subsequent insertion of open reading frames into introns comes from a comparison of the intron reading frames in the mitochondrial ND1 gene in two species of Neurospora. Although the introns in the two species are identical for over 200 bp including the essential sequences that form secondary structures, the introns contain entirely different open reading frames located at different positions. In one strain the conserved elements occur near the 5' end of the intron and a free standing intron reading frame is located 183 bp downstream. A completely different intron reading frame is found in the other strain and is in frame with the preceding exon. This intron reading frame is flanked by direct repeats which may have resulted from a transpositional insertion event (Mota and Collins, 1988).

In yeast, Neurospora, and Aspergillus mitochondrial rRNA introns, the intron open reading frames are separated from the conserved sequence elements required for splicing. For the rRNA gene introns, the intron open
reading frames are inserted between the conserved box 9 and box 2 sequences. In contrast, in intron 3 of the coxI gene in *S. cerevisiae*, all of the conserved sequence elements are situated downstream of the intron open reading frame. In group II introns that have intron reading frames, the bulk of the intron open reading frames are sequestered in domain 4, which can be deleted without consequence to *in vitro* splicing of the intron. In each of the above cases it is possible that the reading frame was acquired by an independent insertion event.

For many other mitochondrial introns (such as aI1, aI2, aI4a, and bI4), some of the conserved secondary structure elements crucial for splicing are within the intron reading frame. In many of these introns all of the sequence elements are located within the intron reading frame except box 2 which is downstream of the intron open reading frame near the 3' boundary of the intron. In these cases, where the open reading frames overlap the secondary structures, the intron open reading frame may have been inserted separately like those above but was followed by subsequent elimination of the stop codons between the inserted open reading frame and the upstream exon (Cech, 1985). In the case of the aI4a and bI4 highly homologous introns, they are most similar in the 3' portion of their intron open reading frames, but diverge
significantly in the 5' portion where the stop codons may have been eliminated over time (Anziano et al., 1982). Further evidence for the independent insertion of open reading frames into introns are the proteins encoded by the phage T4 introns. Each intron product is expressed from its own late promoter located within the intron itself.

IV.C.3. Evolution of intron encoded functions

Acquired intron open reading frames may have evolved to facilitate intron splicing and/or intron propagation. It is difficult to assess which function originated first. Both functions may have simultaneously evolved since they both (RNA splicing and DNA cleavage) require the protein to specifically recognize exon sequences. At least one intron encoded protein, that of the aI4a intron, has both maturase activity and endonuclease activity; however, other intron encoded proteins are believed to have either one or the other function. For instance, although bI4 is 70% homologous to aI4a, that intron product is only capable of maturase activity. Similarly, the fitl protein is a potent endonuclease, but the omega intron is known to splice in petites which are devoid of mitochondrial translation.
It is intriguing to consider that the recognition sequences for splicing and for DNA cleavage are similar. For the aI4a intron, the exon sequence that defines the 5' splice site of the mRNA is the left portion of the target sequence required for DNA cleavage by the endonuclease activity. The lack of an in vitro splicing system makes it difficult to assess how the protein interacts with RNA, however studies to determine how the protein interacts with DNA continue.

It is not clear whether these various activities originate from a single protein domain or whether they are separate functions. For the intron product of aI4a, because the NAM-2 nuclear suppressor activates only about 30% of wild-type splicing function it will be difficult to assess whether the maturase and endonuclease functions are mutually exclusive in vivo. That is, whether the maturase activity precludes the endonuclease activity. However, mutagenesis of the aI4a intron reading frame followed by mitochondrial transformation should allow dissection of the functional domains of these proteins.

Many of the proteins encoded by group I and group II introns are likely bifunctional proteins. Such dual function proteins lend themselves naturally to the spread of introns since they encode both the machinery to insert themselves into new sites via the endonuclease activity or
the potential reverse transcriptase activity, and a factor for their splicing via the maturase function to guarantee that they are not deleterious.
Table 1. Yeast strains used for genetic and physical analysis.

The yeast strains used in this dissertation are listed according to their nuclear and mitochondrial genomes. The nuclear genomes are listed according to the strain and its nuclear markers. Wild type genomes and recombinant genomes employed are also listed. Mit- genomes are listed with their location in the genome. Petites employed in these studies are listed along with the strain from which they were derived.
Table 1.

**Yeast Strains Used for Genetic and Physical Analysis**

<table>
<thead>
<tr>
<th>Nuclear Genomes</th>
<th>Wild type Genomes</th>
<th>Recombinant Genomes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Strain</strong></td>
<td><strong>Markers</strong></td>
<td></td>
</tr>
<tr>
<td>ID41-6/161</td>
<td>ade lys</td>
<td></td>
</tr>
<tr>
<td>Karl</td>
<td>leu karl</td>
<td></td>
</tr>
<tr>
<td>GRF</td>
<td>leu his</td>
<td></td>
</tr>
<tr>
<td>COP19</td>
<td>ade lys</td>
<td></td>
</tr>
<tr>
<td>AH22</td>
<td>ade ura canR</td>
<td></td>
</tr>
<tr>
<td>U303</td>
<td>ade ura his trp leu canR</td>
<td></td>
</tr>
<tr>
<td>5055</td>
<td>ade ura</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mitogenomes</th>
<th>Location</th>
<th>Petites</th>
<th>Derived from</th>
</tr>
</thead>
<tbody>
<tr>
<td>PZ27</td>
<td>cob deletion</td>
<td>Hp-</td>
<td>161</td>
</tr>
<tr>
<td>C245</td>
<td>coxl deletion</td>
<td>17-17</td>
<td>161</td>
</tr>
<tr>
<td>C103b</td>
<td>coxl all</td>
<td>159p-</td>
<td>Δall.12-161</td>
</tr>
<tr>
<td>C2126</td>
<td>coxl all</td>
<td>Np-</td>
<td>Norbensis</td>
</tr>
<tr>
<td>H44</td>
<td>coxl A13</td>
<td>Diap-</td>
<td>Diastaticus</td>
</tr>
<tr>
<td>C2116</td>
<td>coxl A15γ</td>
<td>Cp-</td>
<td>Capensis</td>
</tr>
<tr>
<td>PZ16/A3</td>
<td>coxl A1α</td>
<td>ΔVp-</td>
<td>Capensis</td>
</tr>
<tr>
<td>PZ16/E</td>
<td>coxl A1α</td>
<td>206p-</td>
<td>PZ206</td>
</tr>
<tr>
<td>PZ16/K</td>
<td>coxl A1α</td>
<td>A12/4</td>
<td>161</td>
</tr>
<tr>
<td>PZ16/G</td>
<td>coxl A1α</td>
<td>A4/1</td>
<td>161</td>
</tr>
<tr>
<td>TS12 PZ200</td>
<td>varl region</td>
<td>41/36/7</td>
<td>161</td>
</tr>
<tr>
<td>PZ206</td>
<td>varl region</td>
<td>Δlp-</td>
<td>Δall-161</td>
</tr>
<tr>
<td>5B</td>
<td>all/all2 fusion</td>
<td>401</td>
<td>D273-10B</td>
</tr>
</tbody>
</table>
Table 2. Clones used for physical analysis.

The clones used for various experiments in this dissertation are listed according to their name, the vector in which they are cloned, a description of the DNA insert, the yeast strain from which the fragment was cloned, and the researcher who generated the clone.
Table 2.

**Clones Used for Physical Analysis**

<table>
<thead>
<tr>
<th>Clone</th>
<th>Vector</th>
<th>Insert Description</th>
<th>Strain Origin</th>
<th>Provider</th>
</tr>
</thead>
<tbody>
<tr>
<td>varl-ORF</td>
<td>pBR322</td>
<td>HpaII/HpaII of varl ORF</td>
<td>161</td>
<td>P. Zassenhaus</td>
</tr>
<tr>
<td>varl-leader</td>
<td>Bluescribe</td>
<td>HpaII/HpaII of varl leader</td>
<td>161</td>
<td>P. Huffman</td>
</tr>
<tr>
<td>varl-Cap1</td>
<td>pUC/M13</td>
<td>NsiI/BclI varl 5' half</td>
<td>Capensis</td>
<td>J. Wenzlau</td>
</tr>
<tr>
<td>varl-Cap2</td>
<td>pUC/M13</td>
<td>NsiI/BclI varl 3' of varl</td>
<td>Capensis</td>
<td>J. Wenzlau</td>
</tr>
<tr>
<td>varl-CapR1(1-3)</td>
<td>pUC/M13</td>
<td>NsiI/BclI varl 5' half</td>
<td>Capensis Recombinant</td>
<td>J. Wenzlau</td>
</tr>
<tr>
<td>varl-CapR2(1-3)</td>
<td>pUC/M13</td>
<td>NsiI(ccl1/HpaII3' of varl)</td>
<td>Capensis Recombinant</td>
<td>J. Wenzlau</td>
</tr>
<tr>
<td>aI3γ (pDRI)</td>
<td>pEMBL</td>
<td>BclII/BclI (aI3γ &amp; flanking exons)</td>
<td>Capensis</td>
<td>D. Ralph</td>
</tr>
<tr>
<td>aI5γ</td>
<td>pSP</td>
<td>BclII/BclI (aI5γ &amp; flanking exons)</td>
<td>D273-10B</td>
<td>K. Mecklenburg</td>
</tr>
<tr>
<td>coxl-a11</td>
<td>pBR322</td>
<td>MboI(-1600)/MboI(all)</td>
<td>161</td>
<td>D.R/K.M.</td>
</tr>
<tr>
<td>coxl-a12</td>
<td>pBR322</td>
<td>MboI(all)/MboI(aI3)</td>
<td>161</td>
<td>D.R/K.M.</td>
</tr>
<tr>
<td>coxl-a13</td>
<td>pBR322</td>
<td>MboI(aI4)/MboI(exon8)</td>
<td>161</td>
<td>D.R/K.M.</td>
</tr>
<tr>
<td>TRI</td>
<td>Bluescript</td>
<td>HindIII(aI4)/EcoRI(exon8)</td>
<td>C2116 Revertant</td>
<td>J. Wenzlau</td>
</tr>
<tr>
<td>TR2</td>
<td>Bluescript</td>
<td>HindIII(aI4)/EcoRI(exon8)</td>
<td>C2116 Revertant</td>
<td>J. Wenzlau</td>
</tr>
<tr>
<td>TR3</td>
<td>Bluescript</td>
<td>HindIII(aI4)/EcoRI(exon8)</td>
<td>C2116 Revertant</td>
<td>J. Wenzlau</td>
</tr>
<tr>
<td>TR4</td>
<td>Bluescript</td>
<td>HindIII(aI4)/EcoRI(exon8)</td>
<td>C2116 Revertant</td>
<td>J. Wenzlau</td>
</tr>
<tr>
<td>aI4 Substrate</td>
<td>pUC</td>
<td>BclII/BclI (aI4 &amp; flanking exons)</td>
<td>aI4 mutant K</td>
<td>J. Wenzlau</td>
</tr>
<tr>
<td>aI4-K</td>
<td>Bluescript</td>
<td>BclII/BclI (aI4 &amp; flanking exons)</td>
<td>aI4 mutant E</td>
<td>J. Wenzlau</td>
</tr>
<tr>
<td>aI4-E</td>
<td>Bluescript</td>
<td>BclII/BclI (aI4 &amp; flanking exons)</td>
<td>aI4 mutant G</td>
<td>J. Wenzlau</td>
</tr>
<tr>
<td>aI4-G</td>
<td>Bluescript</td>
<td>BclII/BclI (aI4 &amp; flanking exons)</td>
<td>aI4 mutant A3</td>
<td>J. Wenzlau</td>
</tr>
<tr>
<td>pRS1</td>
<td>Bluescribe/M13</td>
<td>Rsal(aI3γ)/BclI(3'exon)</td>
<td>Capensis</td>
<td>R. Saldanha</td>
</tr>
<tr>
<td>ND6-Rec</td>
<td>Bluescript</td>
<td>BclII/BclI (aI4 &amp; flanking exons)</td>
<td>Norbensis Recombinant</td>
<td>J. Wenzlau</td>
</tr>
</tbody>
</table>
Table 3. Coxl intron configuration of yeast strains.

The intron configurations of the coxl genes of several laboratory and industrial yeast species is illustrated. The yeast strains are listed in column 1, while the presence or absence of the various introns in the coxl gene are indicated by a + or - in the columns under the designated introns.
Table 3.

<table>
<thead>
<tr>
<th>Strains</th>
<th>1</th>
<th>2</th>
<th>3α</th>
<th>3γ</th>
<th>4α</th>
<th>4β</th>
<th>5α</th>
<th>5β</th>
<th>5γ</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID41-6/161</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>ID41-6/161Δ1-3</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>ID41-6/161Δ1,2</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>D273-10B</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>G-II-5γ</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>G-II-0</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>S. ellipsoideus</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>S. norbensis</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>S. capensis</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>S. aceti</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>S. diastaticus</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>S. uvarum</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>S. hienipiensis</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>S. coreanus</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>S. oleaeus</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>S. cerevisiae</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>
Table 4. Intron configuration of strains assayed for conversion of aI4a.

Strains differing in their intron configuration of three mosaic genes (*coxI*, *cob*, and the 21S rRNA gene) of mtDNA are listed. Each strain was assayed for its ability to act as donor or recipient of aI4a. All strains are listed in Table 3, and construction of strains is described in Materials and Methods (strain 7). Strain 6 was isolated as a spontaneous revertant of an aI1 mutant of strain 5 (P. Q. Anziano, unpublished data). Strain 8 was generated by MnCl₂ mutagenesis of strain 5 (Perlman et al., 1980).
Table 4.

<table>
<thead>
<tr>
<th>STRAIN</th>
<th>CoxI gene Introns</th>
<th>21S rRNA intron</th>
<th>cob gene</th>
<th>Recipient/Donor</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. S. norbensis</td>
<td>+ + + + - - - + +</td>
<td>-</td>
<td>+</td>
<td>Recipient</td>
</tr>
<tr>
<td>2. S. capensis</td>
<td>- - + + - + + +</td>
<td>-</td>
<td>+</td>
<td>Recipient</td>
</tr>
<tr>
<td>3. S. ellipsoideus</td>
<td>+ - + + - - - -</td>
<td>-</td>
<td>+</td>
<td>Recipient</td>
</tr>
<tr>
<td>4. D273-10B</td>
<td>+ + + - + - - - +</td>
<td>+</td>
<td>+</td>
<td>Donor</td>
</tr>
<tr>
<td>5. ID41-6/161</td>
<td>+ + + - + - + +</td>
<td>-</td>
<td>+</td>
<td>Donor</td>
</tr>
<tr>
<td>6. ID41-6/161Δ1-3</td>
<td>- - - - + + + +</td>
<td>-</td>
<td>+</td>
<td>Donor</td>
</tr>
<tr>
<td>7. GII-0</td>
<td>- - + - + - - -</td>
<td>-</td>
<td>+</td>
<td>Donor</td>
</tr>
<tr>
<td>8. PZ27</td>
<td>+ + + - + - + + +</td>
<td>-</td>
<td>-</td>
<td>Donor</td>
</tr>
</tbody>
</table>
Figure 1. The yeast mitochondrial genome.

Genes are represented by boxes: solid boxes indicate exons while open boxes indicate introns. tRNA genes are represented by circles. Arrows signify origins of replication. Intergenic regions are shown as a line. Drug-resistance markers are shown as cap, ery, oli and par.
Figure 2. GC cluster families.

The consensus sequences for the a1, a2, a3 and a4 GC cluster families are shown in lines 1-4. Representatives of families v and c are shown in lines 5 and 6. "i" indicates the initial octanucleotide sequence CTCCTTTC (for a1) and CTCCTT-C (for a2 and a3). "a" represents cluster A of a typical ori sequence: GGGTTCC (for a1), GGGGT-CC (for a2), and GGGGTCCC (for a3). The a3 family has an a-c segment which is a combination between the "a" segments found in families a1 and a2 and "c", a 9 bp sequence identical to the C sequence found in a typical ori. The "q" segment found in the a3 family can be repeated up to five times. "m" is a 13 bp stretch which is similar in families a1 and a2. "b" represents a segment similar to the B region of the ori region and is GGGCCCCGG (for a1) and GGG--CCGG (for a2). The terminal hexanucleotide sequence "t" is AACTAT (for a1) and -ACTAT (for a2). Several members of the a4 family contain 10-22 bp of the a1 consensus sequence. The a4 clusters differ from a1-a3 in that their central portion is variable in length and sequence, and they are not palindromes. The v family is very GC-rich and often flanked by 2-14 bp AT-rich inverted repeats, or a common
AT-rich "target" site. The GGTCCCGCGGGG sequence found in the $v$ family members corresponds to a fusion of "a" and "m" segments. Three clusters of this family are present inside, 5' and 3' of the $varl$ gene. Family $c$ members are similar to the $C$ region of the ori sequence and cannot form a hairpin structure. Family $n$ (not shown) consists of very short clusters that cannot be correlated with any of the other families.
Figure 2.
Figure 3. Introns in yeast mitochondrial genes.

Typical laboratory strains of yeast contain up to thirteen introns: seven in the oxi3 gene, five in the cob gene, and one in the 21S rRNA gene. Group I introns are represented by the striped boxes, and group II introns are represented by the solid boxes. The open boxes indicate intron open reading frames, some of which are in-frame with the preceeding exon.
INTRONS IN YEAST MITOCHONDRIAL GENES

\[ \text{OXI 3} \]

\[ \begin{array}{cccccc}
\text{aI1} & \text{aI2} & \text{aI3} & \text{aI4} & \text{aI5a} & \text{aI5}\beta & \text{aI5}\gamma \\
\hline
\end{array} \]

\[ \text{COB} \]

\[ \begin{array}{cccc}
\text{bI1} & \text{bI2} & \text{bI3} & \text{bI4} & \text{bI5} \\
\hline
\end{array} \]

\[ \text{21S rRNA} \]

\[ \omega^+ \]

\[ 1 \text{ kb} \]

- \( \square \) = Exon/Intron ORF
- \( \text{Group I Intron} \)
- \( \text{Group II Intron} \)

Figure 3.
Figure 4. Structure of the var1 gene.

The large striped box represents the coding region of the var1 gene. The two dotted boxes indicate optional GC clusters: the common cluster (black box with white dots) is present in all strains of *S. cerevisiae*, the "a" element (white box with black dots) is an optional GC cluster identical in sequence, but inverted in orientation to the common cluster. The locations of the "b" elements (AAT codons) are indicated by vertical arrows. The orientation and position of the AT-rich acceptor sites are shown as horizontal arrows.
Figure 4.

var1

common cluster a b1(b1p) b2

A+T-rich acceptor sites
Figure 5. Salient features of the double strand gap repair model shown at the MAT locus.

Conversion of Yα information is shown schematically. Ya and Yd alleles pair and the recipient MATa duplex is cleaved near the Y/Z junction (A). The cleavage is enlarged by exonucleolytic degradation of Y to create a gap (B). The gapped duplex is repaired via 3' extension of the Z and X ends of MAT using the donor duplex as a template (C and D). E shows resolution of the intermediate (Strathern, 1988).
Figure 5.
Mitochondrial protein profiles of several laboratory strains and species show a wide diversity in \textit{varl} alleles. Strains 161, D273, and 5DIT are known to contain the common cluster as do all other strains of \textit{S. cerevisiae}. In addition 5DIT contains the "a" and "b" inserts which are reflected in the slower migrating \textit{varl} protein (44kd). D273 contains a "blp" element which yields a 42 kd \textit{varl} protein. The other protein profiles shown of various species have intermediate size \textit{varl} proteins. \textit{S. aceti} and \textit{S. capensis} are believed to contain "blp" elements. The \textit{varl} structure of the other species shown has not yet been determined.
Figure 7. Sequencing strategy of the *S. capensis* varl allele.

Three clones of the varl gene of *S. capensis* were employed in sequence analysis using the restriction sites shown: NsiI/BclI, BclI/HpaII, and NsiI/HpaII. The dideoxy sequencing strategy used is indicated by the horizontal arrows.
Figure 7.
Figure 8. The *S. capensis varl* AT-acceptor sites.

The sequence of the 20 bp AT-acceptor site for the common cluster in the *S. capensis varl* gene is shown. The identical target sequence is found 157 bp downstream where the "a" cluster is known to insert in appropriate crosses.
Figure 9. Schematic diagram of the *S. capensis varl* allele.

A composite of the sequence analysis performed on the *S. capensis varl* gene is shown. The coding region of the *varl* gene is shown by the striped box. The novel "j" cluster sequence is shown in a black box. The open boxes represent the two AT-target sequences found in the *S. capensis varl* allele. Two single base pair substitutions were identified and are shown by the asterisks. The "blp" element is shown by the triangle.
Figure 9.
Figure 10. Sequence of the *S. capensis* "j" cluster.

The sequence of the novel "j" cluster found 49 bp 5' of the *varl* open reading frame in *S. capensis* is shown. The cluster is a member of the al family of GC clusters and is 40 bp in length.
Figure 10.

ATTTCTCTTTCCGGGGTTCCGCTCCCCGTGGCGGGCGGCGAAGCTATTAA
Figure 11. Schematic diagram of the 5' region of the *S. capensis* var1 gene.

At the location of the novel "j" GC cluster in *S. capensis* (line 2), 11 nucleotides are deleted relative to the sequence of ID41-6/161 (line 1), a standard laboratory strain. The single base pair substitution in the 5' var1 region is designated by an asterisk.
Figure 11.
Figure 12. Comparison of the "j" cluster to the al consensus sequence.

Cluster "j" belongs to the al family of GC clusters whose consensus sequence is shown in line 1. The "j" cluster is deleted for nucleotide 18 in the "m" region and three nucleotides in the "b" region (35-37). Cluster "j" also has a T to C substitution at position 7 (shown by a black box), one of the known hypervariable postions.
Figure 12.
Figure 13. Proposed secondary structure of the al GC clusters.

The putative secondary structure for the al GC clusters is shown. Two of the hypervariable positions are found in the stem portions; while seven are found within the loop and bulge. The a2, a3, and v families of clusters can also be folded into hairpin structures.
Figure 13.
Figure 14. Sequence of the *S. capensis* "blp" element.

The sequence of the "blp" element found in the *S. capensis varl* allele is shown. The "blp" element consists of two additional asparagine residues inserted into a string of asparagine residues between positions 480 and 511 of the published sequence of the *varl* gene (Hudspeth et al., 1982).
Figure 14.

GATATTAATAATAATAATAATAATAATAATAATTATTA

$[\text{AAT}]_8$

ACGT
Figure 15. Transmission of the common cluster.

A schematic diagram of a cross between strain *S. capensis* and 161 is shown in panel A. Mitochondrial protein profiles of *S. capensis* (lane 1), 161 (lane 2), and the mixed progeny of a cross between these strains (lane 3) are shown in panel B. Recombinant progeny can be observed by a small proportion of a higher molecular weight species of var1 protein in addition to the parental forms in the lane containing proteins labeled from the total progeny of the cross (lane 3).
Individual progeny of a cross between *S. capensis* and 161 were screened by labeling of mitochondrial protein products and evaluating their migration on 10-15% gradient, SDS, polyacrylamide gels. Lanes 1 and 2 contain mitochondrial proteins from parental strains 161 and *S. capensis*. The arrows indicate recombinant isolates whose varl migration is slower than that of either parent. Three such isolates were chosen for further analysis.
Figure 16.
Figure 17. Southern analysis of transmission of the common cluster.

Panel A illustrates the Southern blot strategy to demonstrate conversion of the common cluster. The probe used to hybridize with mtDNA HpaII digests of the parental strains *S. capensis*, 161, and the total progeny of the cross is represented by the black box. Panel B shows the Southern blot analysis of HpaII digests fractionated on a 3.0% agarose gel, blotted, and hybridized with the probe shown in Panel A: the *S. capensis* parental fragment is 1.2 kb (lane 2), while that of 161 is 0.7 kb (lane 1). The novel recombinant fragment of 0.3 kb in lane 3 indicates transmission of the common cluster and extends from the HpaII site in the "j" cluster to that in the common cluster.
Figure 17.

Recombinant

A Capensis

B

ID41-6/161

Recombinant
Figure 18. Schematic diagram of the var1 structure of capensis recombinants.

The striped box represents the var1 coding region of the S. capensis recombinants. The "j" cluster sequence is shown in the black box. The common cluster is indicated by the dotted box while the AT-rich acceptor sites are shown by the open boxes. The two asterisks denote the single base pair substitutions found in each of the recombinants sequenced. The "blp" element is shown by the triangle.
Figure 18.
Figure 19. Strategy used for sequence analysis of the capensis recombinants.

Two clones were employed in sequence analysis of the recombinant isolates containing the common cluster using the restriction sites shown: NsiI/BclI, and HpaII(common cluster)/HpaII(3' of the var1 ORF). Horizontal arrows indicate how the sequence analysis was conducted. The coding region of the var1 gene is shown by a striped box; the "j" cluster and the common cluster are shown by triangles and the "a" AT-rich acceptor site is indicated; the "blp" element is shown by a triangle; and single base pair substitutions are shown by asterisks.
Figure 19.
Figure 20. Backcrosses demonstrate directionality of the mobile cluster.

Panel A shows a schematic diagram of the backcrosses between the recombinant (line 1) and parental strains 161 and \textit{S. capensis} (line 2). The asterisks denote the locations for insertion of each of the elements (common cluster and "blp"). Panel B shows mitochondrial protein profiles of each of the parental strains in the crosses and the mitochondrial proteins of the total progeny of the crosses run on 10-15\% gradient SDS polyacrylamide gels. In unbiased crosses between the recombinant (lane 4) and \textit{S. capensis} (lane 5), only recombinant var1 protein types are seen in the total progeny (lane 6). In unbiased crosses between 161 (lane 1) and the recombinant (lane 2), both of the parental var1 proteins are transmitted to the progeny (lane 3).
Figure 20.
Figure 21. Varl GC clusters are converted independently.

Panel A shows a schematic diagram of a cross between strains 5DIT (cc+a+bl+b2+) and *S. capensis* (blp+). In panel B the cross was analyzed by labeling of mitochondrial proteins of both parental strains (lanes 1 and 2) and the total progeny of the cross (lane 3). A significant fraction of the varl proteins in the total progeny (lane 3) have non-parental forms whose size is consistent with the presence of either the common cluster or the "a" insert.
Figure 21.
Individual progeny of a cross between strains 5DIT and *S. capensis* were screened by labeling of mitochondrial proteins and analyzing their migration on 10-15% gradient SDS, polyacrylamide gels. In each set, lanes 1, 2, and 3 contain parental strains *S. capensis*, 5DIT, and the total progeny of the cross, respectively. The asterisks denote those recombinants used in further analysis.
Figure 23. Co-conversion recombinant isolates.

A composite of the mitochondrial proteins of recombinant isolates from the screen for recombinant var1 progeny from a cross between strains S. capensis and 5DIT is shown. Lanes 1, 2, and 3 contain mitochondrial proteins from parental strains S. capensis, 5DIT, and the total progeny of the cross, respectively.
Figure 23:

Capensis
SDIT
Total Progeny
CR19
CR12
CR1
CR3
R1
R4
R18

cyt b
cox I
cox II
vWA
Figure 24. Strategy for Southern blot analysis of coconversion recombinant varl isolates.

The strategy employed in Southern blot analysis for the varl recombinants isolated from a cross between strains S. capensis and 5DIT is shown schematically. The location of relevant HaeIII and HhaI restriction sites is shown: both the common cluster and the "a" insert contain a HhaI site; HaeIII sites are located about 400 bp 5' of the varl gene, in the S. capensis "j" cluster, and immediately 3' of the varl gene. The probes used in the Southern blot analysis are indicated as probe 1 and probe 2. Probe 1 extends from a HpaII site about 400 bp 5' of the varl ORF to the HpaII site in the common cluster. Probe 2 extends from the HpaII site in the common cluster to a HpaII site immediately 3' of the varl gene. The varl coding region is represented by the dotted box. The common cluster and "a" inserts are indicated by the striped boxes. Horizontal arrows show the orientation of each of the GC clusters. The "j" and "b" elements are represented by triangles. The open boxes in the varl gene signify the AT-rich acceptor sites for the GC clusters. The predicted sizes of the HaeIII/HhaI restriction fragments are shown below each varl gene diagram.
Figure 24.
Figure 25. Southern blot analysis of conconversion recombinants using the 5' var1 probe.

HaeIII/HhaI restriction digests of mtDNA from strains 161, 5DIT, *S. capensis*, and recombinants R1, R4, R18, CR1, CR3, CR9, CR12 and the progeny of a cross between strains *S. capensis* and 5DIT were electrophoresed on 1.0% agarose gels, blotted to Hybond-N membranes, and probed with the 5' var1 region (probe 1 in Figure 24).
Figure 26. Southern blot analysis of coconversion recombinants using the var1 ORF probe.

HaeIII/HhaI restriction digests of mtDNA from strains 161, 5DIT, S. capensis, and recombinants R1, R4, R18, CR1, CR3, CR9, CR12, and the progeny of a cross between strains S. capensis and 5DIT were electrophoresed on 1.0% agarose gels, blotted to Hybond-N membranes, and probed with the var1 ORF probe (probe 2 in Figure 24).
Figure 26.

161
5DIT
Capensis
R1
R4
R18
CR1
CR3
CR9
CR12
Cap X 5DIT
Figure 27. *Varl* recombinants do not arise by homologous cross overs.

Panel A shows a schematic diagram of a cross between haploid derivatives of *varl* recombinants CR1 (cc+a-) and R18 (cc-a+) and the possible recombinant progeny that may be generated by the cross. Panel B shows mitochondrial protein profiles of the parental strains CR1 and R18 (lanes 1 and 2, respectively), and the total progeny of the cross (lane 3). Also shown are mitochondrial proteins from standard strains *S. capensis* and 5DIT in lanes 4 and 5, respectively.
Figure 28. Common cluster conversion in independent of mitochondrial translation.

Panel A shows a schematic diagram of a cross between petite derivatives of *S. capensis* and 161 which contain the varl region. The varl coding region is shown by a striped box; the common cluster is shown by a dotted box, and the "j" and "blp" elements are represented by triangles. The location of relevant HpaII sites and the predicted sizes of the HpaII fragments used in Southern blot analysis are indicated. The probe used in Southern analysis is shown by a black box. Panel B shows the Southern blot analysis for the petite cross. Lanes 1 and 2 contain mtDNA fragments from parental petite strains of *S. capensis* and 161. Lane 3 contains mtDNA isolated from the total progeny of the cross. The expected recombinant 0.3 kb fragment is indicated by an arrow. The 1.2 kb *S. capensis* fragment has been cut off the gel.
Figure 28.
The diagram illustrates the intron configuration of a portion of the \textit{coxI} gene of strains D273-10B and \textit{S. norbensis}. The strains have four introns in common; aI1, aI2 (which are not shown), aI3a, and aI5g. \textit{S. norbensis} lacks aI4a (present in D273-10B); while D273-10B lacks introns aI3g and aI5B which are present in \textit{S. norbensis}. Exons are indicated by solid boxes, while introns are indicated by open boxes. Intron aI4a is shown as a striped box. The relative positions the polymorphic introns would occupy in the strains where they are absent are shown by either broken (aI3g and aI5B) or solid lines (aI4a).
Figure 29.
Figure 30. Transmission of the omega intron in the 21S rRNA gene.

The transmission of the omega intron in a cross between D273-10B and *S. norbensis* was followed by Southern blot analysis using an oligonucleotide probe homologous to a site in the second exon of the 21S rRNA gene (see asterisk). Panel A shows the locations of *HpaII* sites in each strain, the predicted sizes of the *HpaII* fragments, as well as the position of the oligonucleotide probe. Striped boxes represent exons, while open boxes represent introns. Panel B shows the Southern blot analysis. Lanes 1 and 2 contain mtDNA from parental strains D273-10B (*w*+) and *S. norbensis* (*w*−), respectively. Lane 3 contains mtDNA isolated from the total progeny of the cross.
Figure 30.
Figure 31. Transmission of intron aI5B.

The transmission of intron aI5B in a cross between D273-10B and *S. norbensis* was followed by Southern blot analysis using a probe containing the last three exons of the *coxI* gene fused. This probe was derived from *S. ellipsoideus* whose intron configuration is shown in Table 4, line 3. Panel A shows the locations of the BclI sites in each strain and the predicted sizes of the BclI fragments. Striped boxes represent exons, while open boxes represent introns. Panel B shows the Southern blot analysis. Lanes 1 and 2 contain mtDNA from parental strains D273-10B (aI5B⁻) and *S. norbensis* (aI5B⁺), respectively. Lane 3 contains mtDNA isolated from the total progeny of the cross.
Figure 31.
Figure 32. Transmission of the 3' portion of the coxl gene.

The transmission of the 3' portion of the coxl gene in a cross between D273-10B and S. norbensis was followed by Southern blot analysis using an oligonucleotide probe which hybridizes to exon 5 (see asterisk in Panel A). Panel A shows the locations of the BamHI and EcoRI sites in each parental strain (and aI4a recombinants) and the predicted sizes of these fragments. Striped boxes represent exons, while open boxes represent introns. Panel B shows the Southern blot analysis. Lanes 1 and 2 contain mtDNA from parental strains D273-10B and S. norbensis, respectively. Lane 3 contains mtDNA isolated from the total progeny of the cross.
Figure 32.
Figure 33. Transmission of introns aI3g and aI4a.

The transmission of introns aI3g and aI4a in the coxI gene in a cross between D273-10B and S. norbensis was followed by Southern blot analysis using an oligonucleotide probe which hybridizes to exon 5 (see asterisk in Panel A). Panel A shows the locations of the BclI sites in each parental strain and the predicted sizes of these fragments. Striped boxes represent exons, while open boxes represent introns. Panel B shows the Southern blot analysis. Lanes 1 and 2 contain mtDNA from parental strains D273-10B (aI4a) and S. norbensis (aI3g), respectively. Lane 3 contains mtDNA isolated from the total progeny of the cross.
Figure 34. The \textit{coxi} gene structure of the recombinant.

The \textit{coxi} gene structure of a recombinant resulting from conversion of intron aI4a is shown. Exons are shown as solid boxes, while introns are shown as open boxes. Intron aI4a is shown as a striped box. The arrows indicate regions sequenced from recombinant isolates which were cloned using the BclI sites flanking intron aI4a.
Figure 35. Location of intron aI4a in the recombinant

The sequence of regions flanking intron aI4a in a recombinant is shown in the figure. The lanes contain ddG, A, T, and C, reactions from left to right. The arrows indicate the intron/exon junction at both the 5' and 3' ends of the intron.
Figure 36. Oligomycin sensitive recombinant isolates.

Individual progeny from a cross between D273-10B and
_S. norbensis_ were analyzed by Southern analysis of crude
DNA preparations using the strategy outlined in Figure
32A. 20 of 26 isolates acquired intron aI4a via
conversion. Those recombinants (denoted with an asterisk)
were also found to be linked to the oli^S^ allele. The
first two lanes contain parental DNAs of _S. norbensis_ and
D273-10B.
Figure 36.

Norbensis D273

**...**

Q1 U3 CO5 S.

**I**

**W** M1

- oxn

Oligomycin Sensitive Isolates
Figure 37. Reciprocal exchange following crossing over in exon 5.

The figure illustrates the \textit{coxl} structure of recombinants following a reciprocal exchange event in exon 5. Recombinant R1 is the observed form resulting from a cross between D273-10B and \textit{S. norbensis}. The reciprocal recombinant, R2, has never been detected.
Figure 37.
Figure 38. Mixed growth tests of diploids harboring different mitochondrial genomes.

The figure illustrates the data resulting from mixed growth tests conducted between diploids harboring the mitochondrial genomes of D273-10B, and either *S. norbensis*, *S. capensis*, or a recombinant which has acquired intron aI4a (N-Rec). The fraction of mitochondrial genomes from each diploid type was assessed over 35 generations by the proportion of olis/olir genomes to ensure that the recombinant mitochondrial genome did not have some type of selective advantage over other mitochondrial genome types. At each generation point (left column) roughly 600 diploid colonies were scored. Drug markers of each genome type are indicated in the black circles; the number of generations is shown in the left column, and the percent of a given mitochondrial genome type is shown in the other three columns; sample sizes are shown in parentheses.
<table>
<thead>
<tr>
<th>Generation</th>
<th>D273</th>
<th>Norbensis</th>
<th>Capensis</th>
<th>Rec N Qs</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 gen</td>
<td></td>
<td>66 [511]</td>
<td>48 [493]</td>
<td>55 [680]</td>
</tr>
<tr>
<td>5 gen</td>
<td></td>
<td>47 [1079]</td>
<td>52 [682]</td>
<td>52 [875]</td>
</tr>
<tr>
<td>10 gen</td>
<td></td>
<td>58 [813]</td>
<td>52 [725]</td>
<td>50 [657]</td>
</tr>
<tr>
<td>15 gen</td>
<td></td>
<td>58 [493]</td>
<td>60 [577]</td>
<td>48 [551]</td>
</tr>
<tr>
<td>20 gen</td>
<td></td>
<td>58 [521]</td>
<td>58 [688]</td>
<td>49 [618]</td>
</tr>
<tr>
<td>25 gen</td>
<td></td>
<td>64 [617]</td>
<td>58 [619]</td>
<td>50 [745]</td>
</tr>
<tr>
<td>35 gen</td>
<td></td>
<td>69 [633]</td>
<td>69 [711]</td>
<td>54 [560]</td>
</tr>
</tbody>
</table>

Figure 38.
Figure 39. Norbensis sequence polymorphisms in the 3' exon flanking aI3g

The exon sequence 3' of aI3g from Norbensis is shown. The solid arrow indicates the intron/exon junction. The hollow arrow shows the site of aI4a intron insertion. The two sequence polymorphisms are shown in D273-10B and the recombinant adjacent to the asterisks.
A restriction fragment spanning the intron insertion site was examined for double-stranded breaks in mated cells. The strains 5DSS/D273-10B and COP19/norbensis were mated, and mtDNA was isolated at the beginning of mating and 3, 6, 9, and 24 hours after mating began. The mtDNA was digested with either HpaII (Panel A) or HaeIII (Panel B) and analyzed to detect double-stranded cuts near the sites of insertion for omega and αI4α, respectively. Mitochondrial DNA was fractionated on 6% polyacrylamide gels, electrophoretically transferred to Hybond-N membranes, and hybridized to end-labeled oligonucleotides specific to 21S rRNA gene exon sequences (Panel A) or to αI5β sequences (Panel B). A 0.6 kb fragment (arrow) corresponding to an in vivo double stranded cut near the site of omega insertion is seen from 6-9 hours after initiation of mating in the HpaII digest (Panel A). Similarly, a 0.9 kb fragment (arrow) consistent with a double-stranded cut near the site of αI4α insertion is present from 6-9 hours after mating in HaeIII digests (Panel B).
A.

21S rRNA gene

Hpa II

DOUBLE STRANDED CUT

0.8 kb

0.6 kb

B.

CoxI gene

aI3γ aI5β

Hae III

DOUBLE STRANDED CUT

1.5 kb

0.9 kb

Figure 40.
Figure 41. AI4a conversion requires mitochondrial translation.

Panel A shows a schematic diagram of the Southern blot strategy employed in crosses with petite strains to determine whether AI4a conversion occurs without mitochondrial translation. The positions of the relevant HaeIII and BclI sites in each strain are indicated, with the predicted sizes of these fragments. The probe is represented by a black box. Exons are shown as solid boxes while introns are shown as striped boxes. Panels B and C show blots of HaeIII and BclI digests which have been separated on 1.0% agarose gels. In Panels B and C, lanes 1 and 2 contain mtDNA from parental strains Hp- and Np-, respectively. Lane 3 in Panels B and C contains mtDNA isolated from the total progeny of the cross.
Figure 41.
Figure 42. aI4a conversion in a cross between D273-10B and S. capensis.

The strategy for Southern blot analysis of mtDNA isolated from parental strains D273-10B and S. capensis and the total progeny of the cross is shown in Panel A. Exons are represented by solid boxes while introns are shown by striped boxes. The positions of the relevant HindIII and EcoRI sites are indicated, as well as the expected size of the fragments generated by such digests. The probe which is used is represented by a black box.

In Panel B, lanes 1 and 2 contain mtDNA of parental strains D273-10B and S. capensis. Lane 3 contains mtDNA isolated from the total progeny of a cross between D273-10B and S. capensis. Lane 4 contains mtDNA isolated from a cross between an omega- derivative of D273-10B and S. capensis.
Figure 42.
Figure 43. aI4a conversion in a cross between D273-10B and *S. ellipsoideus*.

The strategy for Southern blot analysis of mtDNA isolated from parental strains D273-10B and *S. ellipsoideus* and the total progeny of the cross is shown in Panel A. Exons are represented by solid boxes while introns are shown by striped boxes. The positions of the relevant HaeIII and EcoRI sites are indicated, as well as the expected size of the fragments generated by such digests. The probe which is used is represented by a black box. In Panel B, lanes 1 and 2 contain mtDNA of parental strains D273-10B and *S. ellipsoideus*. Lane 3 contains mtDNA isolated from the total progeny of a cross between D273-10B and *S. ellipsoideus*. 
Figure 43.
Figure 44. aI4a conversion in a cross between *S. norbensis* and ID41-6/161.

The strategy for Southern blot analysis of mtDNA isolated from parental strains *S. norbensis* and ID41-6/161 and the total progeny of the cross is shown in Panel A. Exons are represented by solid boxes while introns are shown by striped boxes. The positions of the relevant BamHI and EcoRI sites are indicated, as well as the expected size of the fragments generated by such digests. The oligonucleotide probe which is used is represented by a bar and asterisk. In Panel B, lanes 1 and 2 contain mtDNA of parental strains *S. norbensis* and ID41-6/161. Lane 3 contains mtDNA isolated from the total progeny of a cross between *S. norbensis* and ID41-6/161. Lane 4 contains mtDNA isolated from the total progeny of a cross between *S. norbensis* and a derivative of ID41-6/161 (PZ27) which is deleted for the *cob* gene.
The strategy for Southern blot analysis of mtDNA isolated from parental strains D273-10B and ID41-6/Δ1-3 and the total progeny of the cross is shown in Panel A. Exons are represented by solid boxes while introns are shown by striped boxes. The positions of the relevant HindIII and EcoRI sites are indicated, as well as the expected size of the fragments generated by such digests. The oligonucleotide probe which is used is represented by a bar and asterisk. In Panel B, lanes 1 and 2 contain mtDNA of parental strains D273-10B and ID41-6/Δ1-3. Lane 3 contains mtDNA isolated from the total progeny of a cross between D273-10B and ID41-6/Δ1-3.
Figure 45.
The strategy for Southern blot analysis of mtDNA isolated from parental strains *S. norbensis*, G-II-5g and G-II-0 and the total progeny of the crosses is shown in Panel A. Exons are represented by solid boxes while introns are shown by striped boxes. The positions of the relevant BamHI and EcoRI sites are indicated, as well as the expected size of the fragments generated by such digests. The oligonucleotide probe used is represented by a bar and asterisk. In Panel B, lanes 1, 2, and 3 contain parental mtDNAs of *S. norbensis*, G-II-5g, and G-II-0, respectively. Lane 4 contains mtDNA isolated from a recombinant isolate which has acquired intron aI4a. Lane 5 contains mtDNA isolated from the total progeny of a cross between *S. norbensis* and G-II-5g. Lane 6 contains mtDNA isolated from the total progeny of a cross between *S. norbensis* and G-II-0.
Figure 46.
Figure 47. Mutations of the aI4a reading frame

The sequence alterations of two cis-dominant mutants, E and G, and two trans-recessive mutants, K and A3, are shown. The mutant E has a G to A silent change in the internal guide sequence; the mutant G has a C to T change in the cis-acting sequence defined as P6 by Burke (1987), which changes a serine to phenylalanine. Mutant K has two frameshift mutations and bears an A to T transversion which changes a lysine to asparagine, and mutant A3 has a G to T mutation that terminates the reading frame prematurely. The relevent regions of sequence are shown. Each mutation is indicated by bold type.
Figure 48. Truncations of the aI4a reading frame block conversion

In Panel A the positions of *cis*-dominant aI4a mutants E and G, and *trans*-recessive aI4a mutants K and A3 are shown schematically. Exon sequences are shown by striped boxes. Intron sequence is shown as a black line, and the aI4a reading frame is shown as a black box. In Panel B, mutants in the aI4a intron were assayed for their ability to engage in gene conversion. mtDNA from parental strains ID41-6/161 and GRF/norbensis and diploid progeny from crosses between GRF/norbensis and ID41-6/161 and its mutant derivatives was digested with BamHI plus EcoRI, electrophoresed on a 1% agarose gel, transferred to Hybond-N membranes, and hybridized to an exon 5-specific oligonucleotide probe. Parental fragments of donor (ID41-6/161) and recipient (GRF/norbensis) are shown in lanes 1 and 2, respectively. In the control cross (lane 3) between the strains shown in lanes 1 and 2, conversion of aI4a generates the novel 4.6 kb fragment in place of the recipient 5.5 kb fragment. Crosses between the recipient strain and *cis*-dominant aI4a mutants E (lane 5) and G (lane 6) carry out intron conversion as measured by the appearance of the 4.6 kb fragment. In crosses with *trans-*
recessive aI4a mutants K (lane 4) and A3 (lane 7), conversion is abolished.
Figure 48.
Figure 49. The cleavage site of the aI4a-encoded endonuclease.

Single-stranded DNA from phages containing the fused exons which flank aI4a was labeled using the Sequenase protocol and the universal sequencing primer. This material was then extended either using dideoxynucleotides (to generate a sequencing ladder) or with unlabeled nucleotides (to generate a substrate). Substrate samples were cleaved with a mitochondrial extract. Panels A and B show the sequencing gels generated by this protocol using each strand as a template. In each Panel, lanes 1-4 represent chain termination reactions with dideoxy G, A, T, and C, respectively; lane 5 is the product of a cleavage reaction on the "end labeled" substrate; lane 6 shows this region of the gel using uncut substrate to demonstrate no premature primer extension products in the vicinity of the cleavage site. Panel C shows the sequence surrounding the cleavage site. Intron sequences are shown in lower case letters while exon sequences are in upper case letters. The large arrow indicates the point of intron insertion. The sites of cleavage on each strand are shown by small arrows. The exon sequences of donor and recipient are nearly identical; two differences in the
exon between donor and recipient are shown by asterisks. The bracket sets an upper limit on the boundaries of the recognition site.
Figure 49.

RECIPIENT

MAXIMUM BOUNDARIES OF THE RECOGNITION SEQUENCE

5' AGTCTTTGCTCCAGAAGTATATAT
3' aaccctaagaaaccaggtggcttcatatata

SITE OF INTRON INSERTION
Figure 50. Substrate differences between the recipient, donor and artificial substrate.

The recipient allele in these studies (line 1) differs at the three positions shown in bold print from the substrate used by a group that studied the endonuclease activity in *E. coli* (line 2). Both substrates are efficiently cut by the endonuclease. The donor substrate (line 3) contains the majority of the aI4a substrate, however it is not cleaved by the aI4a-encoded endonuclease.
RECIPIENT:
```
tttgGATTTCTTTGGTCATCCAGAAATAT
aaccCTAAGAAAACCAGTAGGTCTTCATA
```

ARTIFICIAL SUBSTRATE:
```
TTCTGATTCTTTGGTCACCTGAARGTAG
AAGACTAAGAAAACCAGTGGGACTTCATC
```

DONOR:
```
atatatataacagCACCCTGAAATAT
atatatatattgttctGGGACTTCATA
```

Figure 50.
Figure 51. Comparison of sequence polymorphisms in the aI4a and fitl substrates.

The sequence polymorphisms in the aI4a substrate are shown in line 2. Both are consistent with cleavage of those substrates (as indicated in the "cleavage" column). Substitutions at similar positions in the fitl substrate (line 1) are resistant to cleavage (as indicated in the "cleavage" column).
Figure 51.
Figure 52. Strategy for dissection of the recognition sequence of the aI4a-encoded endonuclease.

Mutagenic oligonucleotides were synthesized to yield a 10% mutation rate at each position of the 18 bp recognition site and BamHI and EcoRI ends. The oligonucleotides were annealed at their complementary EcoRI ends and extended with Klenow and dNTPs to yield a "double" recognition site with blunt ends. This material was then digested with EcoRI and cloned into a pUC vector. Competent DH5α cells were transformed with the ligation mix and a set of clones containing mutant recognition sites was obtained by colony hybridization using the wild type substrate as a probe. All mutant substrates were sequenced and tested for cleavage with the mitochondrial extract.
Anneal mutagenic oligonucleotides

\[
\text{CGGGATCCTTTGGTCATCCAGAAGTATACGAATTTCG}
\]

\[
\text{GCTTAAGCATATGAAGACCTACTGGTTTCCTAGG}
\]

Extend with dNTP's and Klenow Fragment

EcoRI digest

ligate into EcoRI/SmaI cut vector

\[
\text{EcoRI} \quad \text{SmaI}
\]

\[
pUC
\]

Figure 52.
Figure 53. aI4a mutant substrates.

The figure shows a summary of the mutant substrates and whether or not they are cleaved by the mitochondrial extract. The wild type substrate is shown in the middle. Substrate mutants above the wild type substrate sequence are consistent with cleavage while substrate mutants below the wild type sequence are resistant to cleavage. Those positions that are essential for cleavage are contained within a box. The dotted line designates the cleavage made by the endonuclease. The open bar shows the intron insertion site.
Figure 53.
Figures 54 and 55. Amino acid sequence alignment of reverse transcriptase related sequences.

The figure illustrates the alignment of amino acid sequences of reverse transcriptase-related sequences taken from Xiong and Eickbush (1988). The abbreviations for each element are as follows: Bm, *bombyx mori*; LImd, LINE1 from *Mus domesticus*; ingi, from *Trypanosoma brucei*; Sc, *Saccharomyces cerevisiae*; Pa, *Podospora anserina*; Sp, *Schizosaccharomyces pombe*; Nc-p, *Neurospora crassa* plasmids; Ty, transposable element of *S. cerevisiae*; 17.6 (copia), 297 (gypsy), 412, I, F, and G are mobile elements from *Drosophila melanogaster*; CaMV, cauliflower mosaic virus; DIRS-1, *Dictyostelium discoideum*; HBV, human hepatitis B virus; DHBV, duch hepatitis B virus; MuLV, Moloney murine leukemia virus; HERV, human endogenous retroviral DNA; RSV, Rous sarcoma virus; HSRV, human spumaretrovirus; MMTV, mouse mammary tumor virus; HTLV-I, human adult T-cell leukemia virus; HTLV-II, human T-cell leukemia virus type II; BLV, bovine leukemia virus; HIV-I, human immunodeficiency virus; HIV-II, human immunodeficiency virus type 2. The letters and (+) symbols on the top of the alignment indicate the largely unvaried and chemically similar amino acids found in at
least 11 of the 13 mitochondrial sequences and non-LTR retrotransposons. The letters and (+) symbols at the bottom of the alignment indicate the largely unvaried and chemically similar residues found in at least 18 of the 22 viruses and LTR-containing retrotransposons. Asterisks indicate those residues conserved in all four major groups of elements.
<table>
<thead>
<tr>
<th>Column 1</th>
<th>Column 2</th>
<th>Column 3</th>
<th>Column 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value A</td>
<td>Value B</td>
<td>Value C</td>
<td>Value D</td>
</tr>
<tr>
<td>Value E</td>
<td>Value F</td>
<td>Value G</td>
<td>Value H</td>
</tr>
<tr>
<td>Value I</td>
<td>Value J</td>
<td>Value K</td>
<td>Value L</td>
</tr>
<tr>
<td>Value M</td>
<td>Value N</td>
<td>Value O</td>
<td>Value P</td>
</tr>
<tr>
<td>Value Q</td>
<td>Value R</td>
<td>Value S</td>
<td>Value T</td>
</tr>
<tr>
<td>Value U</td>
<td>Value V</td>
<td>Value W</td>
<td>Value X</td>
</tr>
<tr>
<td>Value Y</td>
<td>Value Z</td>
<td>Value AA</td>
<td>Value AB</td>
</tr>
<tr>
<td>Value AC</td>
<td>Value AD</td>
<td>Value AE</td>
<td>Value AF</td>
</tr>
<tr>
<td>Value AG</td>
<td>Value AH</td>
<td>Value AI</td>
<td>Value AJ</td>
</tr>
<tr>
<td>Value AK</td>
<td>Value AL</td>
<td>Value AM</td>
<td>Value AN</td>
</tr>
<tr>
<td>Value AO</td>
<td>Value AP</td>
<td>Value AQ</td>
<td>Value AR</td>
</tr>
<tr>
<td>Value AS</td>
<td>Value AT</td>
<td>Value AU</td>
<td>Value AV</td>
</tr>
<tr>
<td>Value AW</td>
<td>Value AX</td>
<td>Value AY</td>
<td>Value AZ</td>
</tr>
</tbody>
</table>

Figure 54.
Figure 55.
Figure 56. \textit{CoxI} structure of "popout" revertants.

Southern analysis of BamHI/EcoRI digests of revertant candidates using an oligonucleotide probe in exon 5 (asterisk) allows the determination of which introns have been excised. Exons are shown by black boxes, while introns are indicated by striped boxes. The mutant intron with the C2I16 mutation is represented by an open box. Line 1 shows the intron configuration of the C2I16 mutant. Line 2 shows the intron configuration of an aI5g popout revertant. An aI5B, aI5g popout revertant is shown in line 3. Revertants which have excised introns aI5a, aI5B, and aI5g have the intron configuration shown in line 4.
Figure 56.
BamHI/EcoRI digests of miniprepped DNA isolated from 46 revertants of strain C2116 was examined by Southern blot analysis after separation on 1.0% agarose gels and hybridization to an exon 5-specific oligonucleotide probe. Lanes 1-36 contain DNA from TR revertants: 15 TR revertants have aI5g excised, 11 TR revertants have aI5B and aI5g excised, and 11 TR revertants still contain aI5g. A set of 10 TS revertants still contain aI5g. TR revertants designated with an asterisk (TR1, TR2, TR3, and TR4) were cloned and sequenced to ensure precise intron excision has occurred.

Figure 57. C2116 revertants.
Figure 57.
Figure 58. Sequencing strategy of "popout" revertants.

Two aI5g excision TR revertants and two aI5g/aI5g excision TR revertants were cloned from the HindIII site in aI4a to the EcoRI site in the last exon of coxI into Bluescript vectors for sequence analysis. The horizontal arrows indicate the approach of the sequence determination. Exons are shown by dark boxes, while introns are shown by striped boxes.
Figure 58.
Figure 59. Sequence structure of TR revertants TR1 and TR2

The region of sequence where introns aI5g (TR1) and aI5B and aI5g (TR2) have been excised is shown. The locations of the exon junctions are indicated by arrows.
Figure 59.
BamHI/EcoRI digests of miniprepped DNA isolated from 23 revertants of strain ΔaIl, aI2-C2116 was examined by Southern blot analysis after separation on a 1.0% agarose gels and hybridization to an exon 5-specific oligonucleotide probe. The set of 23 TR revertants was isolated from YPD media. Lanes 1-23 contain DNA from TR revertants: 1 TR revertant was deleted for aI5g, and 1 TR revertant was deleted for aI5B and aI5g.
Figure 60.

Δal1,al2-C2116

GLUCOSE TR REVERTANTS

6.0 5.1 Δal5y

3.5 Δal5β, al5y
Figure 61. CoxI intron structure of C2116 derivatives

The CoxI intron configuration of strains C2116, Δ1,2-C2116, and Δ2-C2116 is shown schematically. Introns are represented by striped boxes, and exons are shown by solid boxes. The C2116 mutation in a15g is indicated by an open box and asterisk. The reverse transcriptase (RT) domains are indicated above each intron by a solid box.
Figure 61.
Figure 62. Strategy of intron subtraction for construction of strain ΔaI1,aI2-C2116.

A petite deleted for introns aI1 and aI2 of the coxI gene was cytoducted into a nuclear background containing a canR marker (W303) (line 1). When crossed via a patch mating to strain akar/C2116, a mutant in aI5g (line 1), only the cytoductants, petites, and recombinants are capable of growth on minimal media supplemented with growth factors to select for the W303 nuclear background. On such media, cytoductants and recombinants are medium sized colonies which are first tested genetically for the presence of the C2116 mutation in aI5g and then physically to ensure the absence of introns aI1 and aI2 in the coxI gene.
Figure 62.

\[
a_{\text{W303/159}^-} \times \alpha_{\text{kar1/C2116}}
\]

\[
\downarrow
\]

Minimal+ (ade, leu, his, trp, ura, can)

\[
a_{\text{W303/159}^-}
\]

\[
a_{\text{W303/C2116}} \text{ (cytoductants)}
\]

\[
a_{\text{W303/159}^- \text{--C2116}} \text{ (recombinants)}
\]

- genetic analysis
- physical analysis
Figure 63. Southern analysis to confirm the structure of ΔaII, aI2-C2116.

Panel A shows the Southern blot strategy to determine the intron configuration in the 5' region of the coxI gene of recombinant candidates. In HpaII/BamHI digests, mtDNA isolated from C2116 generates a 3.6 kb fragment when hybridized to a probe in exon 1 while mtDNA from 159p- generates a 1.6 kb fragment. The desired ΔaII, aI2-C2116 strain would contain a 1.6 kb fragment that co-migrates with that of the 159p-. Exons are represented by solid boxes while introns are represented by striped boxes. The probe is shown by an asterisk. Panel B shows the Southern blot data from HpaII/BamHI digests of mtDNA isolated from C2116, 159p-, and two recombinant candidates. As shown in lane 3, the ΔaII, aI2-C2116 construct contains the 1.6 kb HpaII/BamHI fragment, indicating it is deleted for introns 1 and 2. Lane 4 contains a cytoductant since the 3.6 kb fragment is present indicating the presence of introns 1 and 2.
Figure 63.
Figure 64. ΔaIl, aIl2-C2116 revertants

BamHI/EcoRI digests of miniprepped DNA isolated from 33 revertants of strain ΔaIl, aIl2-C2116 was examined by Southern blot analysis after separation on 1.0% agarose gels and hybridization to an exon 5-specific oligonucleotide probe. Lanes 1-25 contain DNA from TR revertants, all of which have a 6.0 kb BamHI/EcoRI fragment. A set of 9 TS revertants are shown (lane 1-9, second set) which also contain a 6.0 kb BamHI/EcoRI fragment.
Figure 64.
Panel A shows the Southern blot strategy to determine the intron configuration in the 5' region of the coxI gene of recombinant candidates. In HpaII/BamHI digests, mtDNA isolated from C2116 generates a 3.6 kb fragment when hybridized to a probe in exon 1 while mtDNA from the Diastaticus petite generates a 4.6 kb fragment. The desired Δai2-C2116 strain would contain a 4.1 kb fragment that co-migrates with that of the Diastaticus petite. Exons are represented by solid boxes while introns are represented by striped boxes. The probe is shown by an asterisk. Panel B shows the Southern blot data from HpaII/BamHI digests of mtDNA isolated from C2116, Diap⁻, and two recombinant candidates. As shown in lane 3, the Δai2-C2116 construct contains the 4.1 kb HpaII/BamHI fragment indicating it is deleted for intron 2. Lane 4 contains a cytoducctant since the 3.6 kb fragment is present indicating the presence of introns 1 and 2.
Figure 65.
BamHI/EcoRI digests of crude DNA preparations isolated from 34 TR revertants of strain ΔaI2-C2116 were examined by Southern blot analysis after separation on a 1.0% agarose gel and hybridization to an exon 5-specific oligonucleotide probe. 12 TR revertants were deleted for aI5g, while 3 TR revertants were deleted for aI5B and aI5g.
Figure 66.
Figure 67. Gene conversion sites recognized by recombinogenic endonucleases.

The figure illustrates the sequences recognized by various recombinogenic endonucleases. Lines 1 and 2 show the recognition sequences for the *Saccharomyces cerevisiae* MATα and MATδ sites cleaved by the HO endonuclease (Sc HO MATα, Sc HO MATδ). Lines 3, 4, and 5 list the recognition sequences for the intron-encoded endonucleases in the mitochondrial genome of *Saccharomyces cerevisiae* (Sc mt w, Sc mt aI4a, Sc mt aI4a). Lines 6, 7, and 8 represent the sequences recognized by the enzymes encoded within the thymidylate synthase intron of phages T4 and T2 (T4 td, T2L td), and the endonuclease encoded within the sunY intron of phage T4 (T4 sunY), respectively. Line 9 shows the sequence recognized by the endonuclease encoded within intron 3 of the *Physarum polycephalum* rRNA gene (Pp rRNAI3).
Gene Conversion Sites Recognized by Recombinogenic Endonucleases

| Species | MAT Type | Sequence
|---------|----------|--------------------------------------------------|
| Sc HO MAT \(\alpha\) | TTTGGGACTACTTCGCACACAGTATAATTAAACCCTGATGAAGCGCGTGGTCAATTTA | ![Sequence Diagram](image)
| Sc HO MAT \(\alpha\) | TTAGTTCAGCTTTCCGCAAAGTATAATTAAATCAAGTGCAGGCGTGGTCAATTTA | ![Sequence Diagram](image)
| Sc mt \(\omega\) | AGTTACGCTAGGGATAAACAGGTAAATATGAACCAGGTCTCCATTATATC | ![Sequence Diagram](image)
| Sc mt \(\alpha I4\) | tttgGATTTCTTTGGTCATCCAGAAGTAGTATA | ![Sequence Diagram](image)
| Sc mt \(\alpha I4\) | TTCTGATTTCTTTGGTCACCTGGAAGTAGAAGAAGACTAAAGAACAGTGGGACCTCATTCTT | ![Sequence Diagram](image)
| T4 td | GATGTTTTTCCTGGGTCCTACCCTTTAATT | ![Sequence Diagram](image)
| T2L td | GATGTTTTTCCTGGGTCCTCCATTATTTAATTCTACAAGAAGAACCAGAGTGGCCAAATTATAA | ![Sequence Diagram](image)
| T4 \(\text{sun} Y\) | TCCAAGCTTATGAGTATGAAGTGAAACACGTAGGTTCGAATACTCATTTCACTTCTTGTCGA | ![Sequence Diagram](image)
| Pp rRNA \(\text{I3}\) | TAATATGACTCTCTTTAAAGTGAAGTGCCAAAAGATTGATACAGAGAAATCTCGGGTTTTC | ![Sequence Diagram](image)

Figure 67.
Figure 6.8. Comparison of the \textit{in vitro} reaction conditions for fitl and the aI4a endonuclease.

The diagram shows a comparison of the optimized \textit{in vitro} reaction conditions for fitl and aI4a endonucleases. The conditions compared are shown in column 1. The omega encoded endonuclease requirements are listed in column 2 and the aI4a encoded endonuclease requirements are listed in column 3.
<table>
<thead>
<tr>
<th></th>
<th>Omega</th>
<th>ai4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monovalent ions</td>
<td>not needed</td>
<td>stimulates</td>
</tr>
<tr>
<td>Divalent ions</td>
<td>required 2 - 15 mM</td>
<td>required &gt;1mM (200mM NH₄Cl)</td>
</tr>
<tr>
<td></td>
<td>&gt;10mM (-NH₄Cl)</td>
<td></td>
</tr>
<tr>
<td>pH</td>
<td>3 - 9</td>
<td>&gt; 7</td>
</tr>
</tbody>
</table>

Figure 68.
Figure 69. Modes of expression of intron-encoded endonucleases

Each intron encoded protein has a distinct mode of expression. In Panel A, the fit1 protein is translated from a processed form of the excised intron. Panel B shows that the aI4a endonuclease is translated as a fusion protein with the upstream exons. In Panel C the bacteriophage intron encoded proteins are expressed from promoters located within the intron. In each Panel, exons are shown by open boxes, and intron ORFs are shown by striped boxes. The asterisks and vertical arrows in Panel A represent processing sites. Horizontal arrows indicate the direction of transcription.
Figure 69.
Figure 70. The fused exon/intron aI4a reading frame.

The illustration shows that the aI4a encoded endonuclease is translated as a fusion protein with the four upstream exons to generate a p56. The p56 protein is then processed to yield a set of p29 protein species. Exons are represented by open boxes, the intron is shown as a line. The intron ORF is shown as a striped box. The asterisk denotes the protein processing signal.
PROTEINS ENCODED BY FUSED EXON/INTRON READING FRAME

Figure 70.
Figure 71. Mode of intron transposition via an RNA intermediate

The figure demonstrates a hypothetical mechanism for RNA-mediated intron transposition (taken from Woodson and Cech, 1989). A gene containing a self-splicing intron [G1(I+)], generates excised intron RNA after transcription and self-splicing. The RNA can then integrate into the message of another gene [G2(I-)] via reverse splicing. Following reverse transcription and recombination, the intron is incorporated into the genomic copy of G2. Intron DNA is shown by shaded boxes, exon DNA is shown by hollow boxes, and RNA is indicated by a wavy line.
Figure 71.
Figure 72. Conserved reverse transcriptase domains in coxI all

The seven conserved blocks of amino acids in the reading frame of all having homology to reverse transcriptase are shown. The domains are numbered in roman numerals. Exons are shown by striped boxes, the intron is shown by a line, the intron reading frame is indicated by an open box, the region of the reading frame containing the zinc fingers is shown by a dark dotted box, the region containing homology to reverse transcriptase is represented by a light dotted box.
Figure 72.
Figure 73. Model for interaction of intron conversion and intron excision.

The figure illustrates the possible interaction of the aI4a intron-encoded endonuclease with intermediates in the intron excision pathway. In the mutant C2116, intron aI5g rarely splices, but once spliced the mRNA could be reverse transcribed to generate a cDNA. The cDNA may be a substrate for cleavage by the aI4a endonuclease which would generate two half molecules. Gapping may then reduce the size of exon 5 such that crossover events are restricted to exons 6, 7 and 8 to yield the intron configurations of the revertants.
Figure 73.


mitochondrial DNA and preliminary investigations of the mechanism. Genetics, 89, 615-651.


Lang, F. (1984). The mitochondrial genome of the fission yeast Schizosaccharomyces pombe: highly homologous introns are inserted at the same position of the otherwise less conserved coxI genes in Schizosaccharomyces pombe and Aspergillus nidulans. EMBO J. 3, 2129-2136.


mitochondria requires a protein encoded by that intron. Cell 41, 395-402.


Nargang, F. E., Bell, J. B., Stohl, L. L., and Lambowitz, A. M. (1983). A family of repetitive palindromic sequences found in Neurospora mitochondrial DNA is also found in a mitochondrial plasmid DNA. J. Biol. Chem. 258, 4257-4260.


cerevisiae D273-10B. J. Biol. Chem. 255, 9828-9837.


Ph.D. Thesis, The Ohio State University, Columbus, Ohio.


for reconstructing a mammalian gene. Gene 69, 81-89.


