IN VIVO STUDIES OF YEAST MITOCHONDRIAL INTRON SPLICING: ECTOPIC BRANCHING AND A SCREEN FOR NUCLEAR ENCODED SPLICING FACTORS.

APPROVED BY SUPERVISORY COMMITTEE

Philip S. Perlman

Wade Winkler

Kristen Lynch

Hui Zou

Michael Roth

FOR TIM AND ISAAC

I would like to thank the members of my Graduate Committee, particularly Phil Perlman for being a wonderful mentor. I have many fond memories of my lab family and cannot think of a better place for me to have done this work. I would like to thank Melanie Cobb for her encouragement and the Division of Biological Sciences in general for providing great support and many friends. I am very grateful to my family for all the help and support they provided, in particular my husband Tim. Finally I would like to thank God, without whose strength I would never have had the courage to finish.

IN VIVO STUDIES OF YEAST MITOCHONDRIAL INTRON SPLICING: ECTOPIC BRANCHING AND A SCREEN FOR NUCLEAR ENCODED SPLICING FACTORS.

by

TARAH MICHELLE NYBERG

DISSERTATION

Presented to the Faculty of the Graduate School of Biomedical Sciences

The University of Texas Southwestern Medical Center at Dallas

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

The University of Texas Southwestern Medical Center at Dallas

Dallas, Texas

August, 2006

IN VIVO STUDIES OF YEAST MITOCHONDRIAL INTRON SPLICING: ECTOPIC BRANCHING AND A SCREEN FOR NUCLEAR ENCODED SPLICING FACTORS.

Publication No. _____

Tarah Michelle Nyberg, Ph.D.

The University of Texas Southwestern Medical Center at Dallas, 2006

Philip S. Perlman, Ph.D.

The splicing mechanism of group II introns is analogous to that of nuclear introns and it is generally thought that both share a common ancestor. This work contains two studies of group II intron splicing in yeast mitochondria.

Previous studies done in collaboration with Dr. Anna Pyle at Yale identified several important determinants for *in vitro* branch-site selection of intron aI5γ: the presence of a bulged A(A880), the 5' flanking GU base pair and the branch location within domain VI. I confirmed the *in vitro* findings *in vivo* and show that displacing the branch adenosine by one nucleotide in either direction can support branching at the shifted bulged A *in vivo*. Returning

the base-pairs flanking the shifted branch-points to GU pairs increased both the efficiency and fidelity of branching at the ectopic branch A. However, for the shifted down ectopic branch A, it is not the presence of the GU pair flanking the branch that restores branching but the presence of a GC pair located two base-pairs above the branch. This finding is consistent with our observations that for the wild-type branch location, the branch environment above and below the branch are distinct. It appears that the short stem below the branch is important for the second splicing step.

The goal of the second project was to identify novel nuclear genes that are involved in mitochondrial intron splicing. Based on the yeast genome project and several recent proteomic studies of yeast mitochondria, we identified 808 nuclear genes coding for potential mitochondrial proteins that can be deleted without lethality. Of these, 476 deletion strains retain a complete copy of the mtDNA (13 introns) and have a respiratory growth defect. Those strains were screened by northern blot analysis for intron splicing defects. I observed the expected splicing defects in strains deleted for *MSS18, CBP2* and *PET54*. I observed a novel splicing pattern in strains deleted for *IMP1, CBS2, PET111, MNE1, AAT1, ATP10* and *PIF1*.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

LIST OF DEFINITIONS

ORF – Open reading frame: coding region for a gene beginning with the start codon and ending with a stop codon.

RNP – Ribonucleoprotein: Particle consisting of both RNA and protein components.

IEP – Intron encoded protein: Protein derived from an ORF located within an intron.

IVS – Intervening sequence, common term for introns in older work. The term "intron" was coined by Walter Gilbert (1978).

E1-IVS – RNA containing the upstream, 5' exon and the intron. Not a normally occurring RNA species.

IVS-E2 – splicing intermediate that contains the intron and the downstream, 3' exon.

Lariat – Term used to describe normal branched intron RNA with a 3' tail.

mtDNA – mitochondrial DNA

Petite – strain that carries a mutation that renders it unable to grow on a non-fermentable carbon source such as glycerol.

Rho minus **(ρ^-)** – strain that contains a mtDNA with a mutation or deletion that renders it dysfunctional, e.g. the yeast strain cannot grow a non-fermentable carbon source such as glycerol.

Rho zero **(ρ^0)** – strain that completely lacks all mtDNA. Note: these strains still have mitochondria.

Rho plus **(ρ^+)** – strain that contains a fully functional mtDNA

CHAPTER ONE
General Introduction

1-1 CATALYTIC RNA

Intervening sequences, or introns, were first observed in the adenovirus genome using electron microscopy. It was observed that DNA and RNA from the same gene hybridized with large loops of DNA remaining unpaired (Berget et al., 2000). Shortly thereafter intervening sequences were also observed in the mouse β-globin gene using the same methods (Tilghman *et al.*, 1978), indicating that the presence of introns may be wide spread. The authors stated that "the advantage to the cell of introducing interruptions in gene segments is not obvious" and proposed that they might serve to generate genetic diversity (Tilghman *et al.*, 1978). The utility of the ubiquitous intron is still a matter of study and debate, though it is now clear that they do benefit cells by generating genetic diversity through alternative splicing and other gene rearrangements. They are also repositories for regulatory sequences. The genomes of most organisms contain introns in some form, and each intron must be removed along with the simultaneous fusion of two exons (Figure 1-1).

The discovery that a *Tetrahymena* group I intron can catalyze its own excision from message RNA in the absence of protein (Kruger *et al.*, 1982b) challenged the prevailing paradigm of DNA – RNA – Protein – catalysis and created the ribozyme concept. RNA has since been shown to be the catalytic component of not just parasitic DNAs like group I and group II introns (Peebles *et al.*, 1986a), but also essential cellular enzymes like the ribosome, spliceosome, and RNase P (Doudna & Cech, 2002). In addition, RNA is a defining

component in many processes, among them the guide RNA in telomere elongation and maintenance (Harrington, 2003), miRNAs in RNA interference (Hammond, 2005), snoRNAs in rRNA maturation (Kiss, 2002), the large non-coding RNAs (ncRNAs) involved at the chromosomal level in dosage compensation and gene silencing (Bernstein & Allis, 2005), and riboswitches that regulate gene expression in bacteria (Winkler & Breaker, 2005). For a good review of the diverse roles of RNA in even more processes see Storz *et al.* (2005). The fact that RNA can both code for genetic information and catalyze important biological reactions led to the RNA world hypothesis that proposes RNA as the primary component of the first self-replicating system (Doudna & Cech, 2002).

Dissecting out the precise role that RNA plays in many biological processes is complicated by the fact that RNA is virtually always found in complex with proteins. Although many RNA reactions can be reconstituted *in vitro* with no or very few proteins, this is not always the case. RNA can play both structural and catalytic roles in RNA-protein (RNP - ribonucleoprotein) complexes making it difficult to prove that RNA is the catalytic component of a RNP complex. RNA catalysis has been relatively poorly understood compared to protein-based catalysis, however recent progress has greatly increased our knowledge of RNA chemistry (Doudna & Lorsch, 2005). How RNA and protein function together is less understood, as are the respective roles of the protein and RNA components. The task understanding how RNA and protein function together is particularly mind-boggling in the case of the ribosome and spliceosome, both of which are massive complexes composed of many different RNAs and proteins. The spliceosome alone contains five RNAs and as many as 300 proteins (Nilsen, 2003). Given this complexity, the study of smaller,

simpler ribozymes like group I and group II introns is a great way to begin to understand the fundamentals of RNA catalysis and how RNA and protein interact in the cell. This work aims to both look at the RNA catalysis of group II intron branch formation and perform a screen for new proteins involved in the relativity self-sufficient yeast mitochondrial intron splicing reactions.

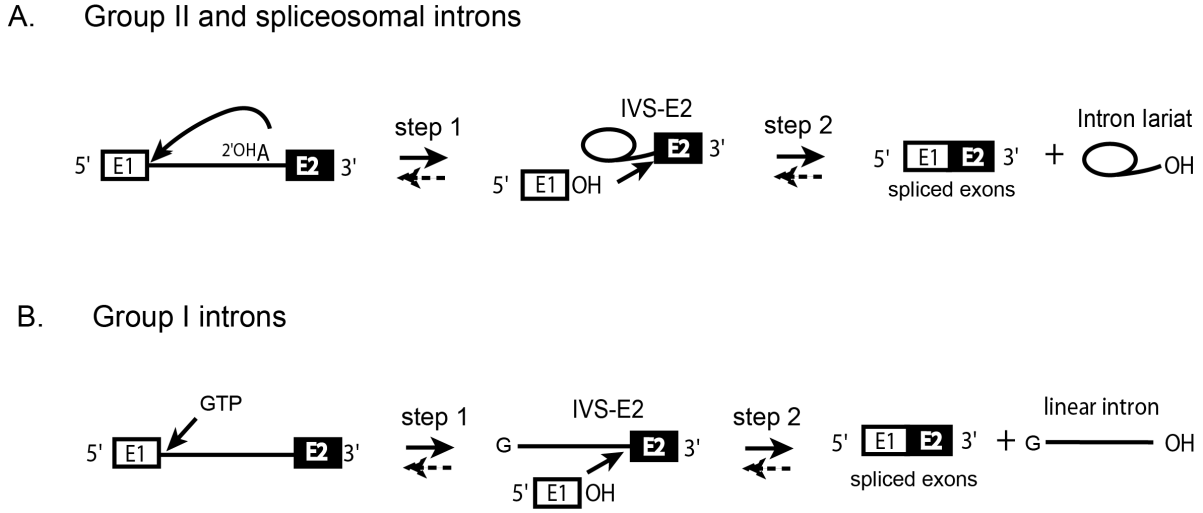


Figure 1-1. Splicing mechanism of group II and group I introns. A. Group II and spliceosomal introns splice using the 2'OH of an internal branch adenosine in the first transesterification step. B. Group I introns use an external guanosine as the nucleophile in the first splicing step.

1-2 GROUP I AND GROUP II INTRONS

Group I and group II introns are ribozymes that consist of a single RNA that is necessary and often sufficient for splicing *in vitro*. *In vivo,* both classes of introns require some form of protein assistance for efficient splicing. Although they use different chemistries and protein co-factors, both can be site-specific transposable elements that catalyze their own

excision from host RNA. RNA splicing enables them to insert into genes with minimal harm to the host organism. Group I introns were the first established case of protein free RNA catalysis and are the best studied ribozymes. Group II introns splice by the same mechanism as nuclear spliceosomal introns and the discovery of their self-splicing was the first evidence supporting the notion that RNA catalyzes the splicing reaction of nuclear introns.

1-2-1 Group I introns

Group I introns are mostly found in the chloroplast genomes of land plants, green and red algae and the mitochondrial genomes of fungi. However they can also be found in the nuclear genomes of a wide variety of protists besides algae, in bacteria, and in bacteriophages. Group I introns have not been found in the nuclei of multicellular eukaryotes (Haugen *et al.*, 2005).

1-2-2a Group I intron structure and splicing mechanism

The structure and splicing mechanism of group I introns has been extensively studied and crystal structures exist for a number of different classes of group I introns. Group I intron RNA, without an open reading frame (ORF), can be quite small, only 250-500 nucleotides long. Most group I introns analyzed in this study contain an ORF or ORF remnant and the combined elements are 737-1622 nucleotides long. Group I introns are classified on the basis of their shared secondary structure which consists of nine paired regions (P1 through P9; see Figure 1-2 adapted from Doherty & Doudna (2001)). The conserved catalytic core consists of pair regions P1, P3-P7 and P9. Peripheral paired regions define the five subclasses

(IA – IE) of group I introns which are classified based on differences in domains P2, P8 and P9 as well as extensions of P5 and P6 (Doherty & Doudna, 2001; Engelhardt *et al.*, 2000). In the tertiary structure, these paired regions stack to generate three major helical domains that form the basis of the intron super-structure. The catalytic core of the intron is found in the center of these three helical domains where the two exons and the guanosine nucleophile are held together (Adams *et al.*, 2004). This structure is completely unrelated to that formed by group II introns or the RNA components of the spliceosome.

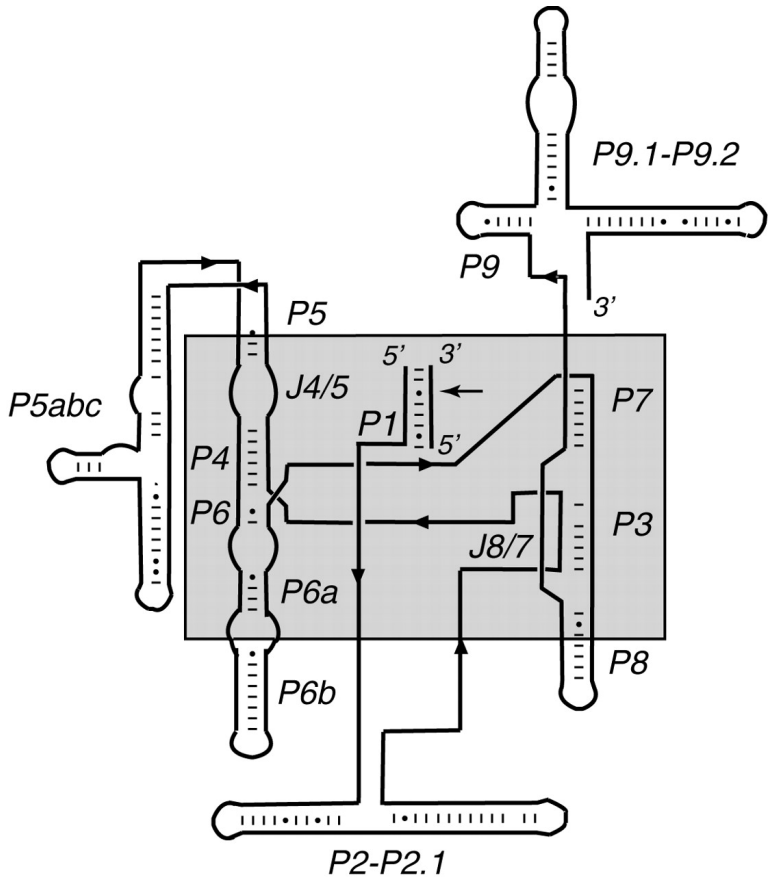


Figure 1-2. Group I intron secondary structure. The nine paired regions, P1-P9 are as labeled. The catalytic core is in the shaded gray box. (Doherty & Doudna, 2001)

As with group II and spliceosomal introns, group I intron splicing proceeds by two consecutive transesterification reactions. However, group I intron splicing is distinguished by the fact that the first step requires an external guanosine to act as the nucleophile. This guanosine attacks the phosphodiester bond between the upstream exon and 5' end of the intron forming a 3'-5' phosphodiester bond with the first nucleotide of the intron. The free 3'-OH at the end of the upstream exon serves as the nucleophile in the second transesterification reaction and attacks the phosphodiester bond between the 3'end of the intron and the downstream exon, and joins the two exons together with a 3'-5' phosphodiester bond. The resultant free intron RNA is linear and has an additional guanosine at the 5' end (Figure 1-1). However, group I introns can form circles (not lariats) *in vitro* by an internal transesterification reaction where the 3'OH attacks a phosphodiester bond (the reaction was discovered with the original Tetrahymena intron (Kruger *et al.*, 1982a) but in yeast are made by omega (Tabak *et al.*, 1984), aI3α (Tabak *et al.*, 1987), and presumably other group I yeast introns. The RNA fold, the guanosine cofactor, and the excision of linear (as opposed to lariat) intron RNA are what distinguishes group I from group II intron splicing. In addition, each intron class has a separate transposition pathway as summarized next.

1-2-2b Group I intron transposition

Many group I introns contain an open reading frame encoding a homing endonuclease. This homing endonuclease can be conceptualized as its own separate parasitic element that takes advantage of the splicing capabilities of group I introns. Four different

types of homing endonucleases are found in group I introns (Chevalier & Stoddard, 2001). The most common class of homing endonuclease**,** contains a characterisitic LAGLIDADG domain, usually in duplicate. In some group I introns, including yeast group I introns aI4α, bI4 (Henke *et al.*, 1995), bI2 (Szczepanek & Lazowska, 1996) and bI3 (Longo *et al.*, 2005), the ORF has evolved an additional separable splicing (maturase) activity and is a potentially a bi-functional protein. The activity of one of the domains has usually been lost, but for bI2 (Szczepanek & Lazowska, 1996) and aI4α (Henke *et al.*, 1995) it can be restored by corrective mutations. All homing endonucleases perform essentially the same function. They make a double strand break in the DNA of an allele that lacks the group I intron. The group I intron (and homing endonuclease) is transferred to the new allele via double strand break repair which uses the intron-containing allele as a template. Homing endonucleases are lost and regained in lineages and some have acquired the additional ability to aid the splicing of group I introns (maturase activity). This has led some to postulate that the acquisition of maturase activity helps the homing endonuclease prevent its loss from the intron (reviewed by Haugen *et al.*(2005)). The co-evolution of group I intron RNA and its encoded protein is an interesting example of how RNA and protein can be inter-dependent on each other yet experience unique evolutionary pressures.

1-2-2 Group II introns

Group II introns are mobile, self-splicing introns. They are found in the mitochondria and chloroplast genomes of lower eukaryotes and land plants. Also, about $1/4^{th}$ of all

currently sequenced bacterial genomes contain one or several group II introns (Lambowitz & Zimmerly, 2004). Recently Group II's have been found in archaebacteria (Dai & Zimmerly, 2003). So far none has been found in a nuclear genome of any organism. Many group II introns code for a multifunctional protein (IEP, or intron encoded protein) that contain reverse-transcriptase, endonuclease and maturase domains. Group II introns are similar to both nuclear spliceosomal introns and non-LTR retro-transposons (NLRs) and are thought to have descended from the common ancestor of both. The study of group II introns has lead to important insights into both nuclear intron splicing and the activity of widespread NLRs such as the LINE elements found in mammalian genomes. In addition group II introns have a large and complex RNA fold the study of which is expanding the knowledge of RNA catalysis gained from studying smaller RNAs like group I introns. For an extensive review see Lambowitz and Zimmerly (2004).

1-2-2a Group II intron structure and splicing mechanism

Group II introns splice via two consecutive transesterification reactions (Figure 1-1). Many group II intron RNAs can self-splice *in vitro* and require only magnesium ions as a cofactor (Lambowitz & Zimmerly, 2004; Peebles *et al.*, 1986b). Unlike group I introns, the first reaction uses the 2'OH of an internal adenosine near the 3' end of the intron to attack the phosphodiester bond located between the upstream exon and the 5'end of the intron. This reaction creates a 2'-5' linkage between that internal adenosine and the first nucleotide of the intron resulting in the characteristic lariat structure which is a loop with a short 3' tail. The second transesterification reaction uses the 3'OH of the upstream exon to attack the

phosphodiester bond between the 3' end of the intron and the downstream exon, thus joining the two exons and liberating the lariat intron RNA. Under certain *in vitro* salt conditions, water can serve as the nucleophile in the first transesterification step, resulting in linear instead of lariat intron RNA (Jarrell *et al.*, 1988b). Later, it was shown that the hydrolysis pathway of splicing can also occur *in vivo* (Podar *et al.*, 1998) and still later it was found that a natural group II intron has lost the ability to splice by transesterification but still splices by the hydrolysis pathway (Vogel & Borner, 2002b).

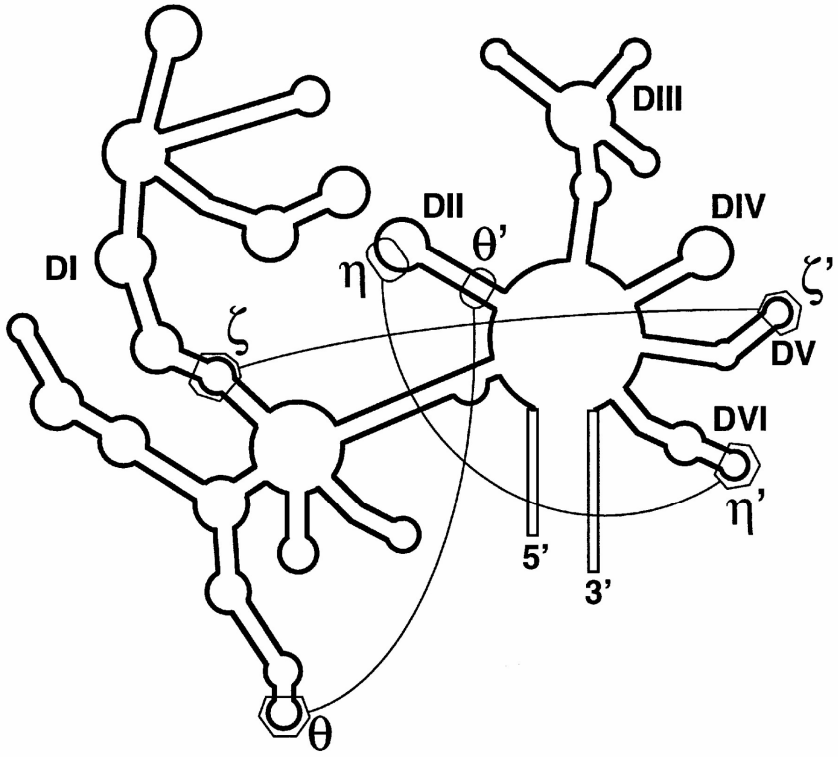


Figure 1-3. Secondary structure of aI5γ. Six helical domains are labeled DI-DVI. (adapted from Chu *et al.*, 2001)

In general, group II introns are larger than group I introns. The most extensively studied group II intron, aI5γ, does not contain an ORF and is one of the smallest group II introns at 887 nucleotides. ORF-containing introns that are active for retro-homing are

usually ~2.5KB long because of the inclusion of the long RT-encoding ORF. Group II intron secondary structure is usually depicted as six conserved helical domains (DI-DVI) connected by short 3-7 nucleotide linkers (Figure 1-3) (for structure review see Fedorova *et al.* (2002)). DI is the largest domain with many branches that vary among group II intron sub-types (for subtypes see, Toor *et al.* (2001)). DI contains the two exon binding sites (EBS1 and EBS2) which base pair with the two intron binding sites (IBS1 and IBS2) in the 5'exon to align the upstream exon on the intron for correct splicing and the mRNA (or DNA) for reverse-splicing (Jacquier & Michel, 1987). Only DI and DV are required for catalysis (Koch *et al.*, 1992). DI is the major scaffolding domain for the ribozyme and makes a number of critical contacts with the catalytic domain, DV (Michels, Jr. & Pyle, 1995). D2 is a relatively small domain that aids in the binding of DIII (Fedorova *et al.*, 2003) and makes contacts with DVI that are important for the second transesterification reaction (Costa *et al.*, 1997). DIII is not required for catalysis of aI5γ *in vitro* but greatly enhances the rate of splicing and is described as a "catalytic effector" (Fedorova *et al.*, 2002; Fedorova *et al.*, 2003; Fedorova & Pyle, 2005). Deleting DIII actually eliminates bI1 self-splicing (Bachl & Schmelzer, 1990). DIV contains the optional reading frame but is not required for splicing (Jarrell *et al.*, 1988a; Koch et al., 1992). DV is considered to be the primary catalytic domain and is absolutely required for catalysis (Peebles *et al*., 1995). DVI is not required for cleavage of the 5' splice site, but it does contain the branch nucleotide that normally acts as the nucleophile in the first transesterification reaction and is required for accurate splicing (Dib-Hajj *et al.*, 1993). Many of the domains in group II introns have similar counterparts in the spliceosome, though most counterparts are in the snRNAs, rather than in the intron sequences (Jacquier, 1990). This is

most clearly demonstrated by the structural and catalytic similarities between the group II intron DV domain and the U6 snRNA (Sashital *et al.*, 2004).

Unlike group II introns, nuclear introns have few conserved nucleotides and are wholly dependent on RNAs in the spliceosome to impose an active configuration for splicing. The spliceosome utilizes five separate RNAs (U1, U2, U4, U5 and U6) to identify and splice nuclear introns (Butcher & Brow, 2005). The best example of the similarity between group II and spliceosomal introns is that the U6 and U2 RNAs base pair with each other and with the branch point in the intron forming two helices that are structurally similar to DV and DVI of group II introns. The catalytic region of the U6 RNA and DV have the same AGC "catalytic triad" and coordinate magnesium ions similarly (Sashital *et al.*, 2004). *In vitro* DV of aI5γ was able to functionally replace a portion of the U6atac RNA, further indication of the similarity between the mechanism of group II and spliceosomal intron splicing (Shukla & Padgett, 2002). It has been proposed that nuclear spliceosomal introns evolved from a group II intron-like precursor that was slowly disassembled by the host organism in order to better control the splicing process (Roy & Gilbert, 2006). Despite the similarities, the spliceosome is huge and the recognition and splicing of nuclear introns presents many unique challenges not encountered by group II introns.

1-2-2b Group II intron transposition

Many group II introns, including aI1 and aI2 in yeast mitochondria, are site-specific non-LTR retrotransposons (Lazowska et al., 1994; Moran et al., 1995). Group II introns retrotranspose via the combined efforts of the catalytic RNA and the intron encoded protein

(IEP). The RNA, with the help of the maturase domain of the IEP, assumes a catalytically active fold and excises itself out of the host RNA. During this process the IEP becomes associated with the RNA making high-affinity contacts with a substructure near the 5' end of DIV (Huang et al., 2003; Singh et al., 2002). This compact RNP particle is readily enriched for use in *in vitro* experiments(Moran et al., 1995). *In vivo*, this RNP particle can insert itself into the intron-less DNA of the host allele in a process called "homing" (Figure 1-4).

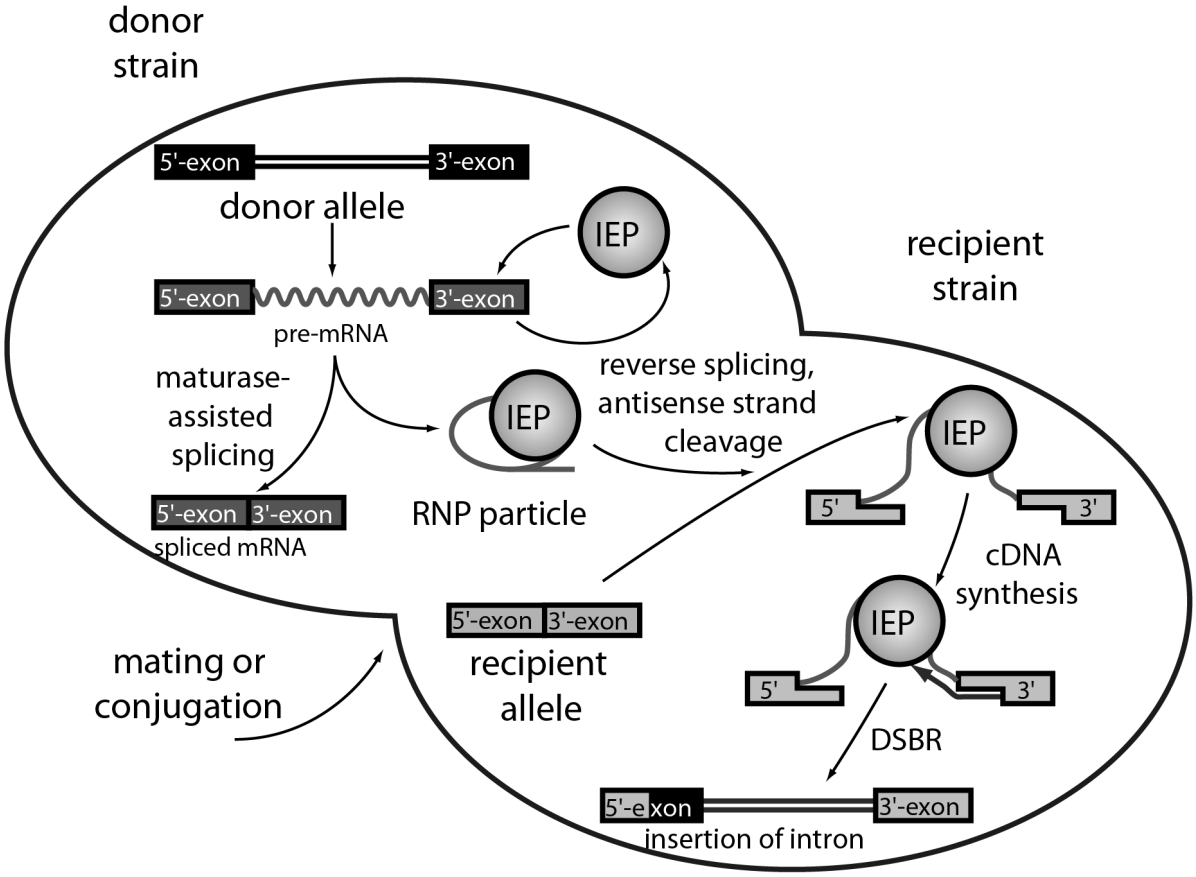


Figure 1-4. yeast group II intron retrotransposition. Intron mobility requires the intron-encoded protein (IEP) that has reverse transcriptase, endonuclease and splicing (maturase) activities.

AI1 and aI2 transpose via the template-primed reverse transcription (TPRT) mechanism which was first worked out using the R2Bm non-LTR element from silkworms

(Luan *et al.*, 1993). The detailed mechanism of group II intron homing was worked out using aI2 (Zimmerly *et al.*, 1995a; Zimmerly *et al.*, 1995b), aI1 (Eskes et al., 1997; Yang et al., 1996; Yang et al., 1998) and the L. lactis intron LtrB (Matsuura *et al.*, 1997). The first step in homing is the recognition of the DNA target sequence. The intron RNA recognizes the sense strand of the DNA target with the IBS-EBS base pairing contacts it uses to splice out of the pre-mRNA. The DNA unwinding function of the endonuclease domain permits double stranded DNA targets to be used and also makes contacts with nucleotides in the 5' and 3' exon (Guo *et al.*, 1997)(Figure 1-5).

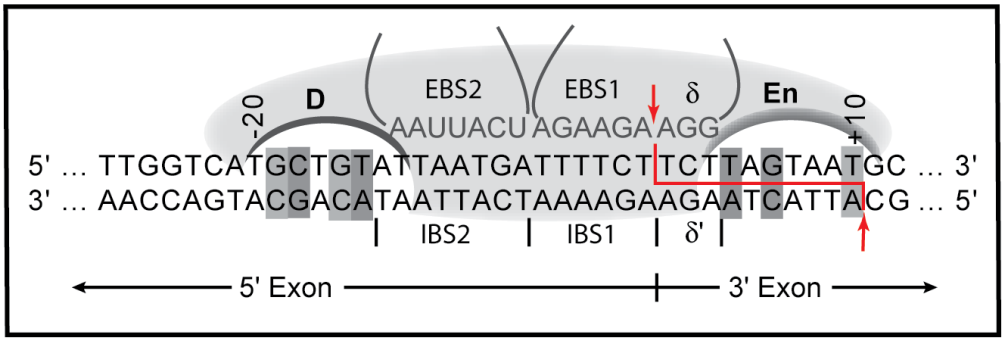


Figure 1-5. Illustration of aI2 RNP contacts with its DNA target site. Shaded pockets indicate the two separate DNA binding and endonuclease domains. Red arrows indicate the sense and anti-sense cleavage sites respectively.

The insertion of the intron RNA into the sense strand of the DNA target occurs by the reverse-splicing of the intron RNA into the DNA target. Next follows anti-sense strand cleavage by the endonuclease domain of the IEP at a site 10 bp downstream from the intron insertion site in the sense strand. The next step of homing is the reverse-transcription of the intron RNA. The RT domain of the IEP uses the 3'OH of the anti-sense strand break as a primer to initiate reverse-transcription. Thus, this step first copies a short sequence of the downstream exon and then progresses into the intron sequence. Ideally the RT proceeds through the entire intron RNA and into the DNA sequence of the upstream exon before

falling off. The final steps of homing, removing the RNA and filling in the gap, have been barely studied, but are assumed to be carried out by host DNA repair processes.

This homing pathway, with complete reverse-splicing and reverse-transcription, occurs *in vitro* under ideal conditions with fully functional RNP particle preparations. The hallmark of such a homing event is that the insertion is clean, occurring without altering any nearby single nucleotide polymorphisms. This mechanism of homing would be required for transposition of a group II intron to a novel, ectopic location. However for homing to proceed in yeast mitochondria (but not in bacteria), it is sufficient for the RNP particle to make a double-strand break and rely on the double-strand break repair mechanism to insert a copy of the intron into the empty allele (i.e., the same strategy employed by group I introns) (Eskes *et al.*, 2000).

The current work is on group II intron splicing and nuclear encoded group I and group II splicing factors and not on homing, but it is important to keep in mind that group I and group II introns are parasitic DNA elements that spread by both linear and horizontal transmission (Zimmerly *et al.*, 2001). This situation explains much of the similarity between introns in divergent organisms but also the lack of conservation of host factors involved in intron splicing.

1-3 YEAST MITOCHONDRIA AS A MODEL SYSTEM FOR THE STUDY OF INTRON SPLICING

Bakers yeast, *Saccharomyces cerevisiae*, is an established model organism that is easy to maintain and genetically manipulate. Specifically, yeast mitochondria are uniquely suited for studying mitochondrial genetics. Strains with nearly any mitochondrial defect can be maintained because yeast can use fermentation as an alternative energy source (Sherman, 1963), and only a handful of nuclear genes for mitochondrial proteins are essential for cell viability. Unlike the situation in higher organisms, genetic exchanges between mitochondria are standard in yeast crosses since both parents of a cross contribute mtDNA to the offspring and mitochondrial genomes recombine with each other (Piskur, 1994). Yeast mitochondria can be genetically altered using the biolistic transformation of plasmid DNA into the mitochondria (Butow *et al.*, 1996). This DNA can be maintained as a mitochondrial plasmid (pMIT) or recombined into natural mtDNAs via crosses to generate deletions and diverse new alleles. The mitochondria of yeast contain up to four group II and nine group I introns which has made it the primary system for the study of these introns *in vivo* (Foury *et al.*, 1998).

1-3-1 The yeast mitochondrial genome

The discovery that chloroplasts have DNA-containing nucleoids similar to those seen in cyanobacteria (RIS & PLAUT, 1962) revived the idea first postulated in the late 19th

century that mitochondria and chloroplasts had bacterial origins. Lynn Margulis (student of Hans Ris in the 60's) formally proposed the endosymbiotic theory in her 1970 book based on this and other observations of bacterial and mitochondrial structure (Margulis, 1970). The presence of separate genomes in mitochondria and chloroplasts along with the marked homology between mitochondrial, chloroplast and bacterial proteins further supports the endo-symbiotic theory of eukaryotic origins (Kurland & Andersson, 2000). The yeast mitochondrial genome is ~85 KB long (Foury *et al.*, 1998) and thought to be circular or consist of linear concatamers (Maleszka *et al.*, 1991), akin to bacterial genomes. Mitochondria still retain bacterial features; however, most of the genes that are predicted to have originally been present on the earliest mitochondrial genome have been transferred to the nucleus or replaced by similar nuclear-encoded proteins (Kurland & Andersson, 2000). Consequently genes for most mitochondrial proteins are transcribed in the nucleus, translated in the cytoplasm, and imported into the mitochondria. Some contain identifiable mitochondrial import sequence tags, but many do not. Only a few genes remain that are encoded, transcribed, and translated in the mitochondria. This number varies among organisms and in yeast only eight necessary protein coding genes and twenty-seven RNA coding genes are located in the mitochondria (Foury *et al.*, 1998). This does not include the introns and intron encoded proteins which are not required for mitochondrial function.

1-3-1a Genes located on mitochondrial DNA

Most of the eight protein-encoding genes located on the mitochondrial genome code for membrane spanning members of respiratory complexes. Three genes code for the

cytochrome oxidase subunits I (*COX1*), II (*COX2*) and III (*COX3*). Three genes code for the ATP synthetase subunits 6 (*ATP6*), 8 (*ATP8*) and 9 (*ATP9*). Apocytochrome b, a subunit of ubiquinol-cytochrome c reductase complex is coded for by the *COB* gene. Finally one gene for a ribosomal protein (*VAR1*) is present. The RNA genes required for the translation of these eight genes are also located on the mitochondiral genome. They include 24 tRNA genes needed to translate the alternative mitochondrial DNA code and the 21S and 15S rRNA genes. Yeast mtDNA also encodes the 9S RNA component of the mitochondrial RNaseP (RPM2) (Poyton & McEwen, 1996).

COX1 of *S. cerevisiae* mtDNA contains up to seven introns (Figure 1-6) and codes for subunit I of cytochrome c oxidase. Cytochrome c oxidase is the terminal enzyme of the mitochondrial inner membrane electron transport chain (complex IV). Cox1p is the core, almost completely membrane-bound member of cytochrome c oxidase, which in addition to the mitochondrial proteins encoded by *COX2* and *COX3* contains nine nuclear encoded proteins: subunit IV (*COX4*), subunit Va (*COX5A*), subunit Vb (*COX5B*), subunit VI (COX6), subunit VIa (*COX13*), subunit VIb (*COX12*), subunit VII (*COX7*), subunit VIIa (*COX9*) and subunit VIII (*COX8*) (Taanman & Capaldi, 1992).

COB contains up to five introns (Figure 1-6) and codes for the apoprotein of cytochrome b, the only mitochondrially encoded subunit of the ubiquinol-cytochrome c reductase (complex III) which also includes Rip1p, Cyt1p, Cor1p, Qcr2p, Qcr6p, Qcr7p, Qcr8p, Qcr9p, and Qcr10p (Hunte *et al.*, 2000; Kreike *et al.*, 1979). This complex is also known as the cytochrome bc_1 complex.

1-3-1b Introns

The *S. cerevisiae* mitochondrial genome contains up to thirteen introns. None of these introns is required for mitochondrial function. Yeast strains with a mitochondrial genome lacking all introns or a genome with 13 introns have identical growth rates on non-fermentable carbon sources (Seraphin *et al.*, 1987). Twelve of the introns are located in the *COX1* or *COB* genes. The thirteenth intron is the group I intron ω and is located in the 21S rRNA gene. Intron ω was the first group I intron to be studied, but it does not figure very largely in this work.

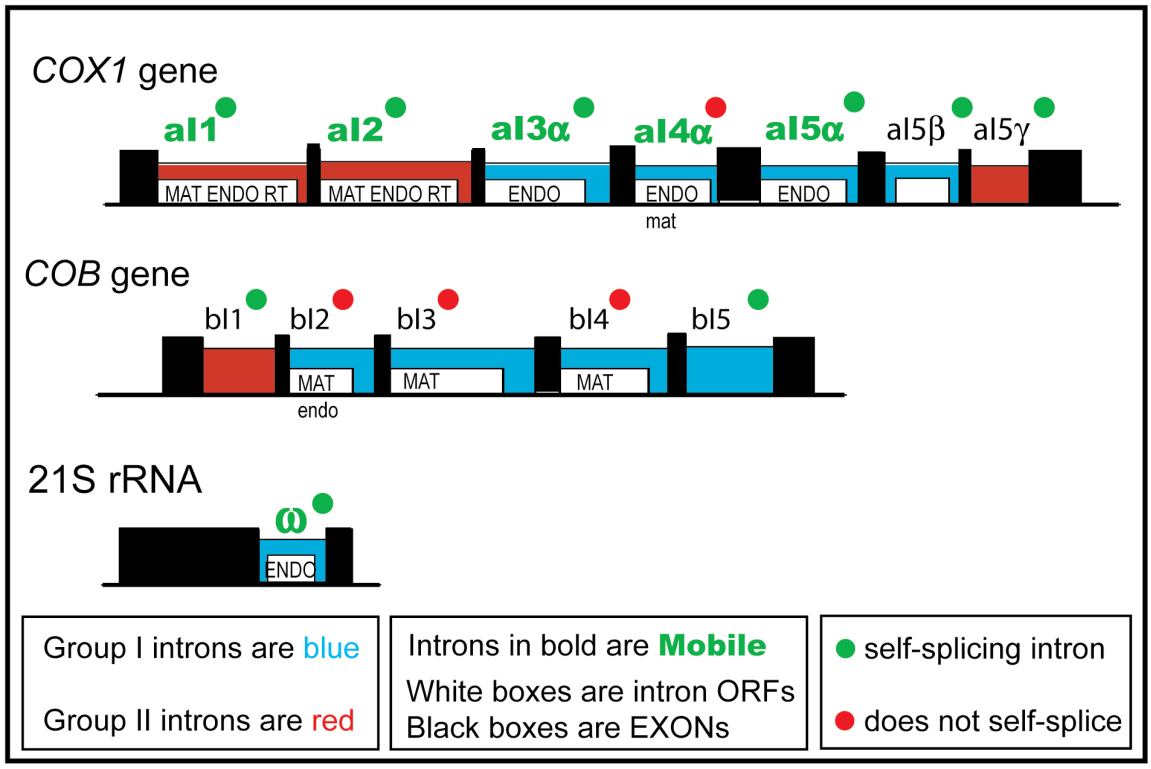


Figure 1-6. Introns found in yeast mitochondria. ENDO = active homing endo nuclease. endo = inactive homing endonuclease. MAT = active maturase, mat = inactive maturase. RT = Reverse transcriptase activity.

All the introns in *COX1* are named aI#, where "a" stands for *COX1*, "I" stands for intron and a number that represents the order of the intron (e.g. aI1 is the first intron in *COX1*). Intron aI5γ, originally thought to be the fifth intron but now known to be the seventh, was so named because the first intron-containing *COX1* allele sequenced did not have introns aI5α and aI5β. The Greek nomenclature was added to accommodate new introns after they were discovered.

The *COB* introns are similarly named with "b" standing for cytochrome b and the *Saccharomyces cerevisiae COB* gene contains five introns (bI1-bI5). Intron bI1 is the only group II intron and it is a small ORF-less intron like aI5γ. Introns bI2, bI3 and bI4 are group I introns containing ORFs with maturase activity. The terminal group I intron, bI5, is ORF-less. In *S. pombe*, the *COB* gene may contain an additional large group II intron that is similar to aI1 and aI2 in COX1 (Lang *et al.*, 1985).

The intron encoded proteins (IEP) of introns aI1, aI2, aI3α, aI4α, aI5α, bI2, bI3 and bI4 are each initially translated as a fusion protein with the upstream *COX1* or *COB* exons respectively. The pre-protein is subsequently processed at a site internal to the intron. The exact process by which this occurs is not well understood and the unprocessed forms of the IEPs are not normally observed. The ORF remnant of intron aI5β is not translated (Johnson C.H. & McEwen J.E., 1997) and a separate message is made for the ω ORF(Jacquier & Dujon, 1985). Introns aI5γ, bI1 and bI5 have no ORF. Despite their dependence on maturases and/or nuclear encoded proteins to splice *in vivo* (Chapter 4), all the *COX1* introns, with the exception of intron aI4α, can self-splice *in vitro*. Even *in vivo*, aI4α has the lowest steady-

state level of splicing with an aI4α intermediate accumulating even in wild-type strains. Of the *COB* introns, only bI1 and bI5 self-splice *in vitro.*

Of the thirteen mitochondrial introns, six are active for transposition. The group II introns aI1 and aI2 and the group I introns ω, aI3α, aI4α, aI5α are mobile. The group I intron bI2 of *S. cerevisiae* possesses debilitating mutations in its homing endonuclease, however a homologous bI2 intron from *S. capensis* is mobile (Lazowska *et al.*, 1992).

1-3-2 The yeast mitochondrial proteome

The mitochondrion performs a great many functions in the cell, including oxidative phosphorylation, the Krebs cycle, biosynthesis of amino acids, lipids and hemes, and many others. Mitochondria also contain many proteins involved in the import and export of proteins and metabolites, protein folding, degradation and synthesis as well as proteins involved in mitochondrial morphology and maintenance. Some of these proteins are exclusive to the mitochondria, but a growing number of them perform dual functions in other cell compartments.

In recent years a great deal of effort has been dedicated to determining which nuclear-encoded proteins are found in the mitochondria and perform mitochondrial functions. A number of different approaches have yielded a combined set of an estimated 700 yeast proteins that either localize to the mitochondria or perform mitochondrial functions (Reviewed in (Reichert & Neupert, 2004). No individual method (deletion screens, localization of tagged proteins, mass spectrometry on highly purified mitochondria) has

identified the complete set of mitochondrial proteins and although there is a good deal of overlap between various studies, a large fraction of proteins are also unique to each analysis. For a complete list of the 20 studies addressing this issue in yeast alone see Reichert & Neupert (2004). For the purposes of this work, which began in earnest while the proteomic studies were first starting to appear, Dr. Hon-Ren Huang, in our laboratory, compiled a liberal list of 867 genes that have been implicated in mitochondrial functioning, although some of them are likely not mitochondrial proteins (see Chapter 4 for more detailed description of the mitochondrial proteome). Our list of 867 potential mitochondrial proteins includes only 44 essential genes. This makes yeast uniquely suited for studying mitochondrial genes, including genes involved in mitochondrial intron splicing.

1-4 AIM OF THIS WORK

The study of group I and group II intron splicing has yielded productive insights into the general principles of RNA structure and catalysis as well as illuminated the specific role that RNA plays in nuclear intron splicing. As the protein involvement in nuclear intron splicing grows increasingly more complex, the simpler mitochondrial intron system may yet hold useful insights into splicing complexes and how they interact with other processes. The yeast, as a system, has proven itself to be particularly useful for *in vivo* work that would be more difficult (if not impossible) in other systems. The current work uses *Saccharomyces cerevisiae* to both expand upon previous *in vitro* studies of group II intron branch site selection and to identify new nuclear proteins involved in mitochondrial intron splicing.

CHAPTER TWO
Materials and Methods

2-1 STRAINS

2-1-1 Yeast strains

Yeast strains used in this thesis are listed in Appendix A, table A-1. All of the aI5γ mtDNA mutations were analyzed in the nuclear background of strain ID41-6/161(*MAT*a *ade*1 *lys*1) containing the mitochondrial genome GII-5γ (in which aI5γ is the only group II intron; the *COX*1 gene has only introns aI3α, aI4α and aI5γ; and the *COB* gene has only bI4 and bI5). Site-directed mutations were transformed into mitochondria of a ρ^o derivative of strain MCC109 (*MAT*α *ade*2-101 *ura*3-52 *kar*1-1)(see section 2-2-4), and the mutant alleles placed into otherwise wild-type GII-5γ mtDNAs by recombination as described in section 2-2-2. The *MRS*2-L232F allele was introduced into the ID41-6/161 nuclear background by one-step gene replacement using PCR amplified DNA from the previously described B-loop-R7 strain (Schmidt *et al.*, 1998). Candidates were selected for their ability to suppress the B-loop mutation of aI5γ, and then were sequenced. The *DBR*1 knockout was introduced into the ID41-6/161 nuclear background by one-step gene replacement using PCR amplified DNA from the Genome Deletion Project *ΔDBR*1::*kan*1-1 strain (strain BY4999 is currently available from Open Biosystems). Candidates were selected for the ability to grow in the

presence of kanamycin (G418) and confirmed by PCR and by probing for the presence of accumulated excised actin gene intron RNA (Chapman & Boeke, 1991).

The strains used in the intron splicing factor screen were ordered from Open Biosystems (www.openbiosystems.com – 6705 Odyssey Drive, Huntsville, AL 35806). The *MAT*a haploid set comprises fifty-three 96-well plates with a total of 4,848 strains. This collection contains most of the viable ORF deletion strains of the yeast genome. Deletions strains of the forty-four essential genes and eleven additional, presumably viable, strains were unavailable as a *MATa* haploid and thus were not included in this study. The deletions were generated by a PCR-based gene deletion strategy that replaces the entire ORF with a KanMX module (Giaever *et al.*, 2002). An overview of the Saccharomyces Genome Deletion Project with detailed descriptions of how the library was generated can be found at: http://sequence-www.stanford.edu/group/yeast_deletion_project/deletions3.html
A sub-library of 867 yeast strains was derived from the full set of 4848 strains by picking only those strains that are deleted for a known or suspected mitochondrial gene. The selection process is described in Chapter 4.

The nuclear background of the *MAT*a haploid deletion strain set is that of strain aBY4741 (*MAT*a his3Δ1 leu2Δ0 met15Δ0 ura3Δ0.)(Brachmann *et al.*, 1998). Strain aBY4741 contains a mitochondrial genome with 13 introns. The *COX*1 gene contains introns aI1, aI2, aI3α, aI4α, aI5α, aI5β and aI5γ. The *COB* gene contains introns bI1, bI2, bI3, bI4 and bI5. The 21S rRNA gene contains intron ω. The sequence of this mitochondrial genome is published (Foury *et al.*, 1998). The presence of all 13 standard yeast mitochondrial group I and group II introns makes this strain ideal for surveying general and specific splicing

defects. Other strains used in test crosses with strain aBY4741 are listed in Appendix A, Table A-1.

2-1-2 Bacterial strains

E. coli strain DH10B (electro competent cells purchased from Invitrogen, 1600 Faraday Avenue, PO Box 6482, Carlsbad, California 92008) were used for cloning. Standard growth conditions and selection methods were used (Sambrook & Russell, 2001).

2-2 YEAST METHODS

2-2-1 Culture conditions

Standard methods for yeast maintenance were employed (Amberg *et al.*, 2005). Yeast strains were generally maintained on YPD-agar petri-plates containing 1% yeast extract, 2% peptone, 2% dextrose and 2% bacto-agar. Dextrose is a fermentable carbon source; cells which were unable to respire still grew on dextrose medium by generating ATP via glycolysis. YPD, a complete medium on which all yeast strains in this study were able to grow, was also used for liquid culture; however, respiratory deficient strains were usually grown in YPD with 10% dextrose instead of 2%. The additional dextrose ensured that these strains were never glucose limited during the growth period.

Raffinose containing medium (YPR – 2% raffinose) was used to grow cells for RNA purification. Raffinose supported growth by fermentation but only slightly repressed expression of genes needed for respiratory growth, unlike dextrose which lowered the level of many mitochondrial proteins by means of catabolite repression. Raffinose medium is routinely used to achieve more consistent results when comparing strains with differing respiratory ability.

Medium containing 2% glycerol (YPG) was used to select for respiratory competent cells; glycerol, a non-fermentable carbon source, requires that cells be able to respire in order to grow. The growth rate of cells on YPG reflected the level of mitochondrial function in various strains, even though cells with as little as 1% of wild-type respiration can still grow detectably on YPG medium

Media were supplemented with 0.2mg/ml G418 when selecting for kanamycin resistance.

Minimal media supplemented with 50mg/L of the appropriate amino acids or nucleotide bases were used to select for auxotrophic markers. Minimal medium contained 0.67% yeast nitrogen base without amino acids, 0.5% ammonium sulfate and 2% of the appropriate carbon source, usually dextrose. Plates used to select only for ura^- mutant strains were supplemented with 1% casamino acids, because casamino acids support much better growth than minimal media alone.

Biolistic transformation required special media to allow yeast cells to recover from the stress of being bombarded. Biolistic transformation plates contained 10% sorbitol, 13%

manitol, 5% dextrose, 0.67% yeast nitrogen base without amino acids, 0.5% ammonium sulfate, 1% casamino acids, 0.01% adenine and 3% agar (Butow *et al.*, 1996).

2-2-2 Yeast crosses and replica-plating of yeast strains.

To understand yeast crosses involving mitochondria it is important to be aware of the special nomenclature used to describe the yeast mitochondrial genome (mtDNA). Consistent with this nomenclature, yeast strains that contain an intact, respiratory competent mitochondrial genome are labeled "ρ^+." Yeast strains that can not grow on non-fermentable carbon sources (e.g. glycerol) are called petites -- a petite phenotype can be caused by mutations of either the nuclear or mitochondrial genomes. A mitochondrial genome that harbors mutations or deletions causing a petite phenotype is labeled "ρ^-." The ρ^- designation includes any mutation or deletion that disrupts mitochondrial function. A yeast strain shown - usually by DAPI staining - to completely lack mtDNA is labeled ρ^o. When it is unknown if a strain is ρ^- or ρ^o, it is labeled ρ^- . For the present work, many strains that probably lacked mtDNA were labeled ρ^- because it could not be directly shown that mtDNA was entirely deleted.

Yeast haploid cells are of two mating types, "α" and "a," corresponding to the pheromones produced by each respective mating type. (Natural isolates of *Saccharomyces* strains are homothallic, i.e., able to switch mating types, but most standard laboratory yeast strains are defective for mating type switching.) Yeast haploid cells generally mate efficiently without special considerations, and it is sufficient to stamp cells from one mating

type onto a lawn of cells of the opposite mating type. This was usually accomplished by replica-plating using sterile velvet squares. When mating yeast strains suspended in liquid in a 96-well plate, a sterilized 48-pronged metal stamp was dipped separately into each half of the 96-well plate and then lightly stamped onto a lawn of the opposite mating type. Diploid cells are larger, and diploid yeast colonies grew somewhat faster than haploid colonies. This was especially true if neither parent strain was able to respire and the cross complemented the respiration defect. For example, the presence of mtDNA was assayed in the yeast library strains that were unable to grow on glycerol (*MAT*a) by stamping the library plates onto a lawn of αW303ρ^o (lacking any mtDNA). Neither parental haploid strain could grow on glycerol, due to either a nuclear or mitochondrial defect. When the library strain was respiratory deficient because of a nuclear defect and contained ρ^+ (wild-type) mtDNA, the diploid offspring of the cross were gly^+. Further stamping of the mating plates onto glycerol medium was sometimes performed to provide confirmation but was usually unnecessary to determine the outcome of each cross; gly^+ diploids grew faster than tester cells and were also easily distinguished from gly^- diploids.

In general, the yeast mitochondrial sub-library generated in this work (Chapter 4) was freshly maintained and tested for the ability of its members to grow under a variety of conditions by replica-plating from either 96-well plates or directly from agar plates onto fresh agar plates of the desired type (YPD or selective medium). Many copies of the same plate could be made in this way and growth extents compared at different temperatures.

2-2-3 Cytoduction of mitochondrial genomes

Haploid yeast strains carrying the *kar*1-1 allele mate well but are defective in nuclear fusion (Conde & Fink, 1976). If one or both of the parental strains in a cross have the *kar*1-1 mutation, fewer than 10% of the mating events result in diploid progeny. Most such mated cells are zygotes that contain both haploid nuclei; buds from such zygotes contain the replication product of one or the other parental haploid nuclei instead of a copy of the fused diploid nucleus. The actual extent of karyogamy varies depending on the strain used; the *kar*1-1 strain most commonly used in this study (αMCC109), produced about 10% diploid cells. This phenotype was extremely useful for mitochondrial genetics because, although the nuclei did not fuse, mitochondria from both parental strains mixed in the zygote so that mtDNA from both parents (or recombinant mtDNAs) could be transmitted to the buds containing one or the other unaltered nuclear genome. If one parent did not contain mtDNA, these crosses provided an easy means of transferring a whole mtDNA from one nuclear background to another. This technique, called cytoduction (Zakharov & Yarovoy, 1977), was used extensively in this work.

Any method of crossing yeast is effective for cytoduction. In order to maximize the fraction of cells of the target nucleus that mated, the parental cells were usually mixed at a ratio of 10 donor cells to 1 target cell. Best results were achieved when cells from log-phase liquid YPD cultures were mixed with some of the pheromone-containing supernatant from each culture, allowed to settle and mate for >10 minutes, and then plated on YPD (without spreading) and left to grow overnight. Cells were suspended in water and plated on

appropriate medium to select for the desired nuclear genotype. The standard *kar*1-1 strain in the lab, αMCC109, also contained the ade2-101 allele which disrupts the adenine biosynthetic pathway at a point that causes the accumulation of a dark red pigment — yielding red colonies on solid YPD medium (Jones & Fink, 1982). It was noted many years ago in the Perlman lab that respiration deficient mutants did not develop the red color, and that visual phenotype allowed for rapid scoring of the respiratory state (and hence presence of mtDNA) of yeast colonies prior to further analysis, and could be used to gauge at a glance the success of a *kar*1-1 mediated cytoduction or the rate of mtDNA loss. This method could not be used exclusively, however, because reversion or suppression was known to lead to the loss or lessening of the red colony color.

2-2-4 Biolistic transformation of yeast mitochondria

Biolistic transformation was used to introduce site-directed mutations into yeast mitochondrial DNA according to published protocols (Butow *et al.*, 1996; Chao, 2001). In brief, plasmid DNAs containing the desired mitochondrial mutation and a separate plasmid with the *URA3* gene were coated onto gold beads and "shot" using a gene gun (Biorad) onto biolistic transformation plates spread with a lawn (~$4x10^8$ cells) of the recipient yeast strain. Colonies that gained the *URA3* containing plasmid grew up on the transformation plates. Several hundred colonies were expected, with greater than 500 Ura+ colonies being ideal. These colonies were replica plated directly onto a lawn of *MAT a* yeast with a gly- deletion or poorly reverting mutation near location of the new change. For example, the D5UA

mutation located in the adjacent DV was used for crossing the aI5γ DVI mutants described in Chapter 3. Because all of the mutations used in this study were expected to permit glycerol growth, selection for restored glycerol growth was used to isolate candidate recombinants containing the transformed mutation. The success rate of biolistic transformation was generally 0-5 transformants per five plates shot (approximately 1 transformant per ~10^9 cells).

2-3 DNA METHODS

Standard molecular cloning techniques were used unless otherwise noted. The reference book set Molecular Cloning: A Laboratory Manual 3rd ed.(2001) by Russell and Sambrook was used extensively. Oligonucleotides used in this thesis have been listed in Appendix A, table A-3.

2-3-1 Mutagenesis and plasmid construction

Plasmids used in this thesis have been listed in Appendix A, table A-2. Plasmid pJD20 (Jarrell *et al.*, 1988a) contained the full length wild-type aI5γ intron plus flanking exon sequences. pJD20 was used as the starting material for generating the aI5γ mutants used in this thesis. Most of these mutants have been previously described (Chu *et al.*, 2001; Liu *et al.*, 1997; Podar *et al.*, 1998). New mutants were generated for this study in the Pyle lab using the Strategene® Quick-change® Site-Directed Mutagenesis Kit.

2-3-2 TENS E. coli plasmid mini-prep

I used a simple, mini-prep method that is a modified version of the TENS mini prep method (Zhou *et al.*, 1990). This method was both inexpensive and rapid, and was particularly useful given the hundreds of mini-preps I prepared to sequence intron branch points. A 1.5ml LB-1mg/ml carbampacillin culture was inoculated with a single *E. coli* colony and grown overnight. This culture was transferred to a 1.5ml eppendorf tube and spun at 13K rpm for one minute to pellet the cells. The supernatant was decanted and the pellet resuspended in 50µl of doubly distilled (dd) H_2O by scraping the tubes across the top of an eppendorf tube rack 3X. (This was much faster than vortexing). Then 300µl of a TENS solution (10mM TrisHCl pH 8.0, 1mM EDTA pH 8.0, 0.1N NaOH, 0.5% SDS) was added and the sample was mixed by inverting repeatedly followed by an incubation at RT for ~2-4 minutes. Then 150µl of 3M NaOAc pH 5.2 was added, the sample mixed by inversion and incubated on ice for 5 -10 minutes. The cell mass was pelleted at 13K rpm for 10 minutes and the slimy pellet removed using the flat end of a toothpick. If done properly, virtually no pellet remained. Then 5µl of a 10mg/ml RNaseH (Roche Applied Science, P.O. Box 50414, 9115 Hague Road Indianapolis, IN 46250-0414) solution was added and the sample was incubated at 37^0C for 15-30 minutes. One ml of pre-chilled 100% Ethanol was added, mixed by inversion, and spun at 13K rpm for 10 minutes to precipitate the DNA. The supernatant was decanted, and the pellet was washed once with 70% Ethanol and allowed to dry. I

resuspended the pellet in 50µl of ddH_2O. The resulting DNA was usually concentrated (~0.5-1µg/µl) and clean enough for DNA sequencing or further cloning.

2-3-3 Yeast whole-cell DNA mini-prep

Yeast whole-cell DNA mini-preps were performed to isolate DNA for further PCR analysis. I used the published glass-bead/phenol method (Amberg *et al.*, 2005). However, if I only needed the DNA for PCR, then the following simplified method was used: A ~50-100µl pellet of cells was scraped with a sterile toothpick off of a fresh YPD plate and suspended in 200µl of Breakage Buffer (2% Triton X, 1% SDS, 100mM NaCl, 10mM TrisHCl pH 8.0, 1mM EDTA pH 8.0). To this ~250 µl of 0.5mm glass beads (measure by placing glass beads in the eppendorf tube up to the 250 µl mark) and 200µl of phenol/chloroform were added. The samples were vortexed for ~2 minutes, placed on ice for ~2 minutes, and then vortexed again for ~2 minutes. The samples were microfuged at maximum speed (13,000rpm) for 5-10 minutes and the aqueous upper layer transferred to a fresh eppendorf tube. This sample was used (~1-5µl) directly for PCR, or was ethanol precipitated and resuspended in 50µl ddH_2O. Though taking somewhat longer than colony PCR, this procedure yielded more consistent PCR results.

2-2-4 Confirming deletion library strains by PCR.

In order to confirm deletions in library strains, I essentially followed the recommended online PCR protocol using primers A and D (See the online table at http://sequence-www.stanford.edu/group/yeast_deletion_project/deletions3.html for a list of recommended primers and the expected wt and deletion sizes). The only exception to the recommended protocol was that I used 2µl of a whole-cell yeast DNA prep -- as described in the previous section -- instead of performing colony PCR with a zymolyase treatment. The PCR reactions were fractionated on agarose gels directly or following cleavage within the kan cassette with *Hind*III . Samples of the correct deletion size that were cut by HindIII were considered confirmed. See also: (http://www-sequence.stanford.edu/group/yeast_deletion_project/verification.html)

2-4 RNA METHODS

RNase contamination can be easily avoided without taking extreme measures. The section on working with RNA in Molecular Cloning: A Laboratory Manual 3rd ed. (2001) by Russell and Sambrook is very informative on this topic. I found that using RNase-zap (ambion) or a similar product to clean glassware, and using untouched plastic ware and weigh boats, adequately protected from RNase contamination. I did not bake my glassware. All reagents to be used for RNA work, including dry-goods, were stored separately, and care was taken not to touch them with potential sources of RNases (e.g., bare hands and/or gloved hands that have possibly touched skin/hair). I used non-autoclaved ultra-pure water collected into RNase-free containers with good results. In general, when I encountered RNA

degradation, the most likely cause was high heat (over-incubation at >60^{o}C or electrophoresis under conditions that significantly heat the gel and buffer) and not RNase contamination.

2-4-1 Yeast whole-cell RNA purification

Yeast were grown in YPR media (YP with 2% raffinose) at 30^{o}C with shaking at 250 rpm unless otherwise specified. Temperature sensitive strains were harvested following growth at the non-permissive temperature: 20^{o}C (c^{s}) or 37^{o}C (t^{s}). Note that such strains are temperature sensitive for respiration but not for growth by fermentation; thus, mutant cells were easily grown in YPR medium at the non-permissive temperature. Cells were harvested before they reach a diluted OD_{600}>1.5. Stationary phase cells were harder to prep because their cell walls are thicker. Ideally, cells were harvested at an OD_{600} of 0.8-1.0. To achieve consistent results, 2 ml YPR starter cultures were inoculated with a single colony and grown at 30^{o}C for ~24hrs. Depending on the amount of RNA desired, a 50 ml YPR culture (standard volume) or a 12 ml YPR culture (the volume used in the library screen making it easier to grow 20 cultures at a time) was started from the 2 ml culture. Cells were allowed to double at least three times prior to harvesting (starting OD_{600} of ~0.05-0.1 or, for the library screen, 2 μl of the 2 ml culture was used). Yeast cells grew more slowly in YPR than in YPD, with a doubling time of ~3.5 hours. The final YPR cultures were usually started in the afternoon, grown overnight, and harvested the next morning.

RNA was isolated from yeast cells using a modified version of the published acid phenol method (Schmitt *et al.*, 1990). A scaled down version of the method was used for

screening the yeast mitochondrial library; those volume have been indicated in parentheses: Cells were harvested from the above-mentioned 50 ml (or 12 ml) cultures, resuspended in 1ml AE buffer (50mM NaOAc pH 5.3, 10mM EDTA) and transferred to a 1.5 ml eppendorf tube. Cells were pelleted, resuspended in 400 µl (200 µl) of AE buffer and 50 µl (25 µl) of 10% SDS and then vortexed for one minute. Added to this suspension was 600 µl (300 µl) acid phenol equilibrated with AE buffer. The samples were vortexed for four minutes and incubated at 65^{o}C for ~2 minutes, vortexed for 2 minutes and then incubated a second time at 65^{o}C for ~2 minutes. Extended incubation (>10 minutes) of the samples at 65^{o}C was found to lead to RNA degradation. The samples were then placed in the -80 ^{o}C freezer for at least 10 minutes. The frozen samples were centrifuged immediately at room temperature at maximum speed in a microcentrifuge for 10 minutes. The cell pellet should have been translucent; an opaque white cell pellet generally indicated poor extraction of mRNA. The aqueous upper phase was carefully removed and transferred to a new eppendorf tube. This was a critical step; it was important to remove the upper layer as rapidly as possible without disturbing the white cell debris found at the interface, and was better to leave some of the aqueous phase behind than to contaminate the sample with RNase-containing cell debris. The aqueous phase was then extracted with 500 µl (250 µl) phenol/chloroform twice followed by a standard ethanol precipitation using RNase-free reagents. The dry pellet was resuspended in 50-100 µl (15 µl) of RNase-free TE. The 50 ml prep yields 150-250 µg RNA. Generally I diluted the RNA to 2-3 µg/µl and used 15 µg of RNA/lane for the initial, unbalanced northern blot analysis. For the library screen blots, 5 µl of the sample was used for northern analysis without quantification of the amount of RNA.

2-4-2 Northern blot analysis using agarose gels

For northern blot analysis, 10-20µg of whole cell RNA were mixed with freshly prepared RNA loading/denaturing buffer (400µl formamide, 160µl formaldehyde (38% aqueous solution), 80µl MOPS, 20µl of a bromophenol blue solution and 1µl ethidium bromide) (Sambrook & Russell, 2001). When indicated, blots were balanced to +/- 20% for content of *COB* mRNA by repeating the northern and adjusting the input sample volumes. I used a 1:1 or greater ratio of loading buffer to RNA (For example 10µl loading buffer and 5µl RNA sample) and incubated the samples at 60^{o}C for 10 minutes. The denatured samples were placed on ice prior to loading on the agarose gel. RNAs were analyzed on 1.2% SeaKem LE (BMA), TBE buffered agarose gels containing 25mM guanidinium thiocyanate (Goda & Minton, 1995; Podar, 1997). The guanidinium thiocyanate (0.36g for a 150ml gel) was mixed in 5ml of TBE and added to the molten agarose before pouring. Samples are loaded, up to a total volume of 30µl, and any empty lanes were loaded with RNA loading buffer to reduce sample smiling. The gel was run at less than 5V/cm in TBE buffer until the bromophenol blue dye was in the last third, usually right at the edge, of a 15cm gel. It was important that the gel did not become too hot to prevent RNA degradation. If the gel felt hot, I would reduce the voltage or use a circulating buffer system. The RNA was transferred onto a charged nylon membrane (S&S Nytran) in 20X SSC using the standard capillary action method (Sambrook & Russell, 2001). After transfer, the RNA was fixed to the membrane by UV-crosslinking in a Stratalinker (Stratagene).

Blots were hybridized with ^{32}P-labeled oligonucleotide probes in RapidHyb buffer (Amersham Biosciences) according to the provided protocol. (The sequences of the oligonucleotides used as probes have been reported in Appendix A, table A-3.) To increase blot sensitivity, mixtures of up to three separate probes were used: for the 3'exon, oligos E2-bottom, E6 9717as, cox1E6as; for aI5γ intron, oligos 5g5E, SnabAS and DIII-lower; for *COB* message, oligo cobE6as. A single probe was used for each of the remaining 12 introns, as listed in Appendix A, table A-3. Blots were washed once in 5X SSC, 0.1% SDS for 15 minutes and twice in 2X SSC, 0.1% SDS for 10 minutes at 42^{o}C. Blots were wrapped in plastic-wrap and placed on a PhosphorImager screen for a minimum of 16 hours prior to scanning.

PhosphorImager scanning was used to quantitate levels of various RNA signals. The digital images obtained in this manner were quantified with the ImageQuant software. To estimate the extent of splicing, the net message signal was represented as a fraction of the total net signal for all bands that hybridize to the exon probe. Specifically, identically-sized boxes were drawn around and in the blank space directly below the four major bands: IVS-E2, message, precursor 1 and precursor 2. The net signal was obtained by subtracting the signal in the blank box from the signal in the band containing box. The fraction of the total RNA that underwent the first splicing reaction was the sum of the signal for message plus the signal for IVS-E2 over the total net signal for all 4 bands. For the library screen this calculation was slightly modified because the library strain contained 13 introns and potentially had many precursor bands. When calculating percent total splicing, I replaced the box for the slowest moving precursor, "precursor 2," with a larger box that encompassed all

the major slow moving precursors and made an identically sized box for the background. As long as the gel lane was "clean" (free of major smudges, spots etc…), I made separate background boxes for each lane.

One of the major advantages of using the guanidinium thiocyanate gel system is that it has improved resolution for linear from branched intron and intron-containing precursor RNAs over formaldehyde based systems (Podar *et al.*, 1998). However, linear and broken lariat RNAs still migrate identically (Jarrell *et al.*, 1988b; Podar *et al.*, 1998).

2-4-3 Northern blot analysis using polyacrylamide gels

Polyacrylamide gel electrophoresis was used to determine whether unbroken lariat was present in the band that migrated as linear intron RNA on agarose gels. I performed the method exactly as published in the Thesis of Mircea Podar, starting on p. 84 (Podar, 1997). Basically, 15 mg of whole cell yeast RNA was fractionated on a standard 40 cm, 4% polyacrylamide gel. The RNA was transferred to a charged nylon membrane by liquid electroblotting in 25mM sodium phosphate buffer pH 7.0. for three hours at 4^{o}C and then crosslinked and probed as described above.

2-4-4 Branch point mapping using RT-PCR

Branch points were mapped as described, with some minor modifications (Vogel *et al.*, 1997): Input RNAs for all semi-quantitative cosmetic gels were first balanced by northern blot analysis. First strand synthesis was carried out using Superscript II RNase H^- RT (Invitrogen, Carlsbad CA) according to the provided protocol. Of the two primers used, 116AS and SnabAS, SnabAS gave cleaner results, but data from both was included. No RNase treatment was necessary prior to PCR. One or two µl of the 20µl RT reaction were used for PCR. I obtained the best PCR results using 2.5U of 5U/µl TAQ polymerase (Roche), the supplied buffer (used as 8X instead of 10X), and 50 pmol of nested primers 5g5E and 657S in a total reaction volume of 50 µl. Preincubation for one minute at 94°C was followed by 36 (unless specified) cycles of 10 seconds at 94°C, 30 seconds at 50°C, 10 seconds at 72°C and a final step of one minute at 72°C on a MJ Research PTC-200 thermocycler. Products were analyzed on 3% Metaphor® (BMA) agarose gels, containing TAE, run at 4°C with circulating buffer, and bands of interest were excised. Cloning was performed using the pGEM-T Easy® vector TA-cloning kit (Promega). Clones were transformed into *E. coli* and initially screened for inserts by color (blue-white) then mini-prepped (TENS protocol, see section 2-3-2) and digested with EcoRI (Roche) to further screen for inserts of the appropriate size and finally sequenced from the T7 primer (IDT) on an ABI machine. Fragments were cloned from a minimum of three separate RNA preps.

2-4-5 *In vitro* RNA-splicing assay

The methods article by Perlman and Podar (1996) was used exclusively to generate this simplified non-quantitative splicing protocol. Plasmid pJD20 containing wt or mutant aI5γ was used to generate a DNA template for *in vitro* RNA transcription by PCR amplification of the intron using the T7 primer and E6-9717as (table A-3). The PCR reaction was ethanol precipitated, washed, and resuspended at a concentration of approximately 1μg/μl. Transcription reactions were carried out using T7 RNA polymerase (Roche Applied Science, for protocol see www.roche-science.com/pack-insert/0881767a.pdf) according to the provided protocol using 1μg of template DNA. Since the goal of these reactions was to map branch points, I did not purify or quantify the RNA products but directly proceeded to self-splicing reactions. The transcription reaction products were ethanol precipitated and resuspended in 40μl of splicing buffer (0.5M $(NH_4)_2SO_4$, 0.1M $MgCl_2$, 40mM Tris-HCl, final pH 7.5) and incubated at 45^{o}C for 90min. A maximum of 1μλ of this reaction, but usually 1μl of a 1/10 or 1/100 dilution, was used to map intron branch points using the RT-PCR protocol described in section 2-4-4.

CHAPTER THREE
The Role of Domain VI in Splicing and Branch Site Selection in the Group II Intron AI5γ

3-1 INTRODUCTION

Group II introns share conserved secondary and tertiary structures and splice by two transesterification reactions that yield excised intron lariat and spliced mRNA as products (Bonen & Vogel, 2001; Michel & Ferat, 1995; Qin & Pyle, 1998) . Splicing of intron aI5γ from yeast mtDNA has been studied intensively and findings obtained from both *in vitro* and *in vivo* experiments are generally in good agreement (*e.g.*, Boulanger et al., 1995; Podar et al., 1998). The branch nucleotide is a nearly invariant adenosine within domain VI (DVI) located 7 nucleotides from the 3' end of the intron (A880 see Figure 3-1). DVI is generally depicted as a largely helical structure in which the branch A is bulged out from the helix as a single unpaired nucleotide. A crystal structure obtained with a model RNA that contains most of the DVI of aI5γ, together with DV, suggests a different configuration in which both the branch A (A880) and the U residue 3' of it (U881) are unpaired (Zhang & Doudna, 2002). This configuration is probably

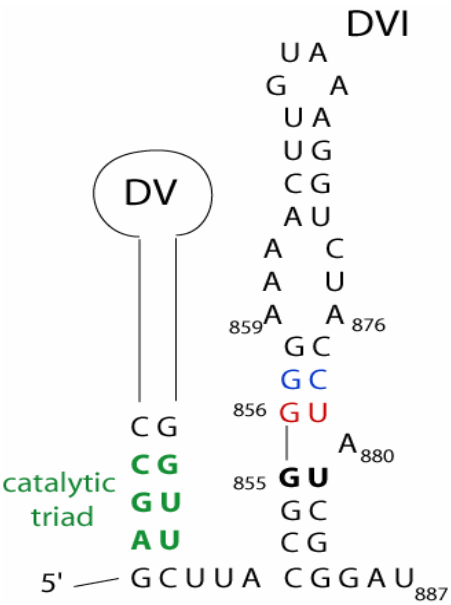


Figure 3-1. Secondary structure of aI5γ DV-DVI. The branch nt, A880, is shown as an unpaired A between a red and a bold GU pair. The DV bases highlighted in green are called the "catalytic triad" and are important, for efficient catalysis. In mutant D5UA the AU pair is flipped to a UA pair.

a crystallization artifact because it is inconsistent with most other published data on DVI structure(Schlatterer *et al.*, 2006; Sigel *et al.*, 2004); however the present study does not depend on one of those models being correct.

The presented work was an *in vivo* study of how intron aI5γ selects the branch nucleotide. This work would not have been possible without the help of a new method for mapping branch points.

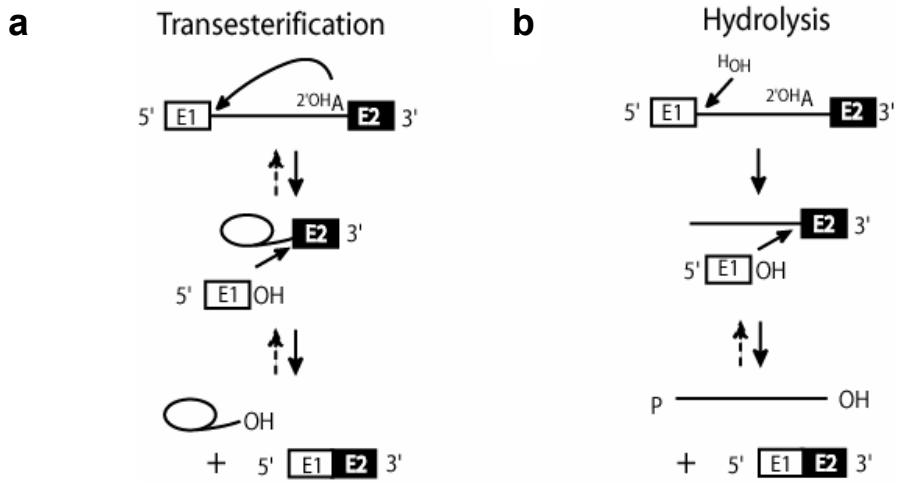


Figure 3-2. Group II intron splicing pathways. a, Splicing by first step transesterification. **b.** Splicing by first step hydrolysis.

Under optimal *in vitro* self-splicing conditions, most precursor RNAs containing aI5γ have spliced with branching (Jarrell *et al.*, 1988b; Peebles *et al.*, 1986b; van der Veen *et al.*, 1986). *In vivo,* the excised aI5γ lariat RNA has been shown to accumulate to detectable levels (Hensgens *et al.*, 1983). Mapping of branch sites by primer extension did not yield a result with *in vivo* RNA from a wild-type strain (Peebles *et al.*, 1993), though solid evidence with wild-type *in vitro* RNAs showed that most or all branching occurred at A880 (Liu *et al.*, 1997; Peebles *et al.*, 1986b). Branch sites have been mapped by primer extension on *in vivo*

RNAs from several yeast mutants that accumulate detectable levels of intron-3' exon (IVS-E2) intermediate, and in those instances branching was at A880 (Peebles *et al.*, 1993).

As it has been performed, primer extension could not demonstrate that all of the splicing was at A880; a low level of splicing at an ectopic site would not give a strong enough signal on a sequencing gel. Similarly, the DNAzyme method for mapping *in vitro* branch points used by Chu *et al.* (2001) could not map very low levels of branching or samples that contained a mixture of branch points. This method uses DNAzymes to digest the lariat into a smaller, 127bp branched molecule that is 5' or 3' labeled and digested by alkaline hydrolysis. Branched molecules result in a gap in the ladder of products produced by alkaline hydrolysis corresponding to the unlabeled 5' or 3' tail, respectively. A change in the branch point would alter the gap in each reaction. Even in the "gap" region, a faint laddering of bands is always observed and consequently a low level of branching would be poorly distinguished from this faint background laddering. An RT-PCR based approach for mapping branch sites has been developed that gets around this problem because it looks at individual branching events and not an average of a population of branching events (Vogel *et al.*, 1997)(see also figure 3-4). This was the method used to map branch points in this work as it allowed me to detect branching at multiple sites *in vivo*.

This method was used to provide strong evidence that the predicted branch A is the site of branching *in vivo* for a set of chloroplast group II introns (Vogel & Borner, 2002a). Early *in vitro* studies with wild-type aI5γ RNA showed that several variations of the self-splicing reaction conditions activate an alternative splicing pathway in which A880 is replaced by water as the nucleophile in the first splicing reaction (*i.e.*, splicing by hydrolysis,

see Figure 3-2) (Jarrell *et al.*, 1988b). Under the original "low salt" reaction conditions, the wild-type intron self-splices relatively slowly and mainly with branching (Jarrell et al., 1988b; Peebles et al., 1986b). In a high salt medium containing 100 mM $MgCl_2$ and >0.5 M $(NH_4)_2SO_4$, the rate constant for splicing with branching is nearly four-times larger and the rate constant for splicing by hydrolysis is not enhanced (Chu *et al.*, 1998). Replacing the $(NH_4)_2SO_4$ with >0.5 M KCl substantially inhibits the branching reaction; however, splicing remains quite active with linear RNA as the main intron product (Jarrell *et al.*, 1988b). Splicing in >0.5 M KCl also activates a "spliced exons reopening" (SER) reaction that generates free exons (Jarrell *et al.*, 1988b). The SER reaction has been described as the reversal of the second splicing reaction, by hydrolysis (Podar *et al.*, 1995).

Splicing by the hydrolysis pathway is observed with various reconstituted intron RNAs under reaction conditions where the wild-type intron branches very efficiently. *In vitro* splicing studies are often performed with the minimum intron RNA necessary for the first splicing step to occur. For example, an E1-IVS substrate RNA containing intron domains I through III (or DI to DIII) is completely inactive, but adding just DV RNA in *trans* reconstitutes an active ribozyme that forms free E1 and linear IVS RNA as products (Jarrell *et al.*, 1988a). Similarly, adding DV RNA to a wild-type pre-mRNA deleted for DV yields free E1 and linear IVS-E2 as the main products (Dib-Hajj *et al.*, 1993). The latter reaction is especially interesting because a wild-type DVI is present in the substrate RNA but fails to participate in most reactions. That reaction indicated that DV plays a key role in positioning DVI for its role in branching but is effective only when DV and DVI are in *cis*. Domain VI is not essential for the first splicing step because a mutant deleted for DVI retains self-

splicing activity in buffers containing $(NH_4)_2SO_4$, yielding E1 and linear IVS-E2 as the main products, with no branching (Koch *et al.*, 1992). ΔDVI RNA can carry out the second splicing reaction at a low level, but selects the 3' intron boundary inaccurately and fails to join E1 and E2 correctly. These findings illustrate that DVI plays a role in the second splicing reaction and influences selection of the second step splice site.

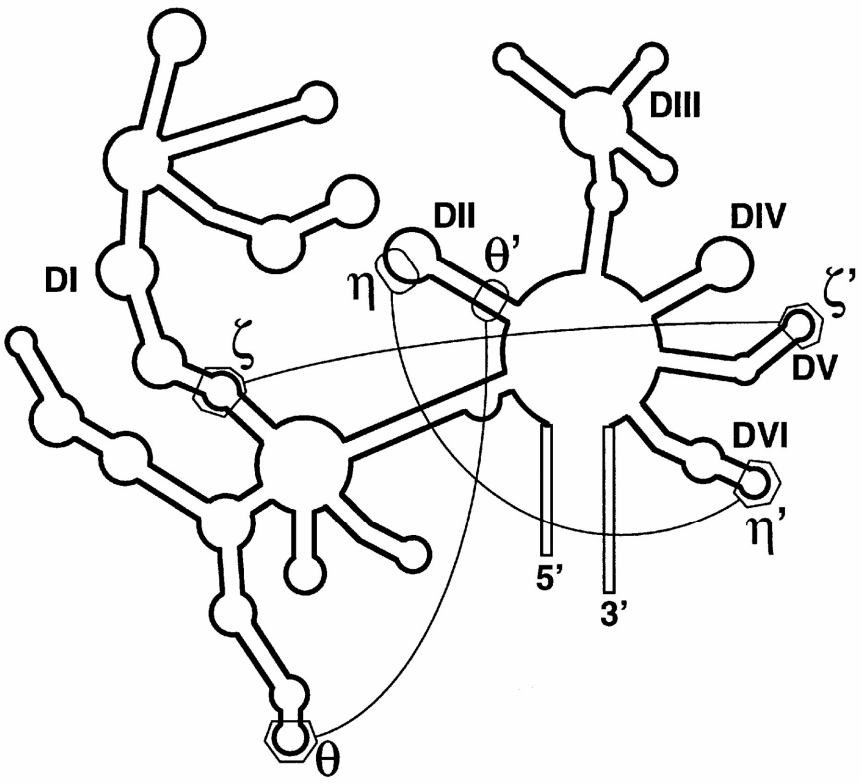


Figure 3-3. Secondary structure of aI5γ. Six helical domains are labeled DI-DVI. The η – η' tertiary interaction is as indicated. (taken from Chu *et al.*, 2001)

In aI5γ RNA, a tertiary interaction between the terminal loop of DVI and a docking site in the distal part of DII, called η – η' (Figure 3-3), is important for efficient second step reactions. This interaction is thought to promote a structural rearrangement that follows the first step reaction (Chanfreau & Jacquier, 1996). Domain VI then, folds in one configuration for the first splicing step and then rearranges prior to the second splicing step.

Many mutations of the branch nucleotide in DVI of aI5γ have been analyzed *in vitro* (Chu *et al.*, 2001; Chu *et al.*, 1998; Liu *et al.*, 1997; van der Veen *et al.*, 1987). Those mutations strongly inhibit or block branching but permit efficient splicing by hydrolysis under salt conditions that support branching by the wild-type intron. Of the four natural nucleotides, only A supports efficient branching; G and U permit a low level of residual branching at nucleotide 880 (G>U) and C supports no detectable branching. Atomic mutations of A880 showed that adenine is mainly recognized by its N6 amino group; interestingly, a ribo-abasic residue at nucleotide 880 supports branching at ~1% the rate constant of control reactions (Liu *et al.*, 1997). Deletion of A880, or pairing A880 with a U, inhibited branching to an undetectable level (more than 1000-fold), though splicing by hydrolysis remained (Chu *et al.*, 1998). It is interesting that the paired-A mutant is more inhibited for branching than is even the abasic mutant; however, changing the UA pair to GA restored nearly wild-type branching. Overall, those experiments establish that the identity of the branch nucleotide base and its arrangement in the DVI substructure are important factors for branching. For those mutants where the branch sites were mapped, the expected branch nucleotide was always employed (Chu *et al.*, 1998).

A critical issue with any *in vitro* study is whether it reflects what really happens inside of a cell. Some of the previously studied aI5γ DVI mutations, when transformed into yeast mitochondria, have had similar effects on branching and splicing *in vivo* (Podar *et al.*, 1998). Most branch nucleotide mutations inhibit splicing ~70% *in vivo*. Branch nucleotide mutants have little if any branching and cause a noticeable defect of the second splicing step, as evidenced by the accumulation of linear (non-branched) IVS-E2 RNA. Even for several

mutants that have a low level of residual branching, little excised intron RNA accumulates, evidently because linear intron RNA is rapidly degraded (Podar *et al.*, 1998). Importantly, those experiments show that aI5γ can splice by hydrolysis *in vivo*; thus, the hydrolysis pathway may be an ancient feature of group II introns rather than just an artifact of non-physiological, *in vitro* reaction conditions.

The current study began as the *in vivo* counterpart to a second set of *in vitro* experiments in our on-going collaboration with the lab of Anna Marie Pyle at Yale University (Chu et al, 2001). The goal of this new set of DVI mutations was to identify other factors that influence splice site choice. Chu *et al* (2001) generated mutants with the branch A moved either one base pair 3' or 5' of its natural location. They found that shifting the branch A one base pair upstream (Br⇑) or downstream (Br⇓) from its usual location, while retaining the ability to remain bulged from the DVI helix, strongly inhibited branching *in vitro* but still permitted robust splicing by hydrolysis. Mutant BR⇑ branched at 1/203rd of the wild type rate, and mutant Br⇓ at 1/2200th of the wild type rate, *in vitro*. It was not possible to map the branch sites in those mutants with the DNAzyme-assisted mapping method. However, they were able to map a similar mutant that combined the BR⇑ mutation with a mutation that increased the length of the single-stranded joiner sequence connecting DV and DVI (see mutant 4B Figure 3-22). This mutation slightly increased the amount of branching and was postulated to have extended the DVI stem by one base pair. DNAzyme-assisted mapping showed that the branching was at an ectopic U located at the fifth base pair from the base of the DVI helix (giving DVI a four base pair stem like the native branch A), rather

than at the shifted A. These data further supported the notion that group II introns are hard-wired to branch at a single location only.

The phylogenetic analysis of DVI sequences and structures presented by Chu *et al.* showed that the base pairs adjacent to the branch A in the DVI helices of group II introns are highly conserved, though not invariant. In wild-type aI5γ, A880 is flanked on both sides by GU pairs. Among a large sample of group II introns, GU is the most common pair above the branch A. The previous *in vitro* analysis of a mutant in which both GU pairs were changed to GC (mutant 2XGC) showed that splicing with branching was significantly inhibited (Chu *et al.*, 1998). *In vivo,* splicing was inhibited 80-90% with predominantly linear IVS-E2 RNA as the main intron product (Podar *et al.*, 1998). The 2XGC mutant further indicated that the base pairs flanking the branch A may be important for branching. It was consequently noted that mutant BR⇑ has a GC pair above the shifted A. Changing that GC to GU improved the efficiency of branching about 40-fold (to $1/5^{th}$ the control rate for branching). DNAzyme-assisted mapping showed that the rescued branching was at the shifted A (Chu *et al.*, 2001). Thus, the finding that this combination of mutations activates branching at the displaced bulged A was the first evidence of efficient ectopic branching. It also indicated that the environment of the unpaired A in DVI is an important factor in defining the branch site.

Taken together, the above studies identified four major structural determinants for branch-point selection in group IIB introns: i) an unpaired adenosine ii) a 4 base pair stem beneath the branch site; iii) a DV-VI linker at least 3 nucleotide long; and iv) a GU pair upstream of the branch nucleotide. Those branching determinants are based on results of *in*

vitro experiments and it remained unknown whether this intron depended on more or less stringent conditions for branching *in vivo*.

Here I analyzed the *in vivo* phenotypes of key mutations from this *in vitro* study with several illuminating additions. The key mutations all supported respiratory growth and spliced at detectable levels. With one exception, no lariat RNAs were detected on RNA blots but an RT-PCR assay allowed me to detect and map branching in most mutants. The RT-PCR assay sampled individual branch points rather than averaging among all branch, which increased sensitivity and revealed that several of the new mutants branch ectopically *in vivo* and also that, given conflicting branch determinants, some mutants respond by branching at alternative locations. The following experiments provide strong evidence that the environment of the branch nucleotide is important for splicing and branch site choice, both in mutant and wild-type contexts.

3-2 RESULTS

3-2-1 Calibration of an RT-PCR assay for branching at nucleotides other than A

The *in vivo* analysis of splicing phenotypes of DVI mutants depends on our ability to identify branched RNAs and map their branch nucleotides. Although primer extension has yielded some insight to *in vivo* branching by group II introns, it is not generally useful because only mutants that accumulate detectable amounts of the splicing intermediate IVS-E2 can be mapped readily. Chemical and DNAzyme-assisted methods for mapping branch-points are impractical for strains that branch inefficiently *in vivo*, but an RT-PCR assay has proven to be very effective (Vogel *et al.*, 1997).

This assay is diagrammed in Figure 3-4. An RT reaction using an antisense primer that is complementary to intron sequence usually yields a strong stop at the 5' end of the intron on both linear and lariat RNAs. On branched molecules, the RT occasionally inserts a nucleotide in response to the RNA branch and then the cDNA is extended through DVI back toward the beginning of the intron (Vogel *et al.*, 1997). PCR amplification of those longer cDNAs using a new pair of primers (as indicated) yields double-stranded cDNAs that are readily cloned and sequenced. Each cDNA clone contains a linear sequence that defines the branch site used in an individual splicing reaction. Sequencing a number of such clones from each intron allele surveys the branch-point location(s) in the strain.

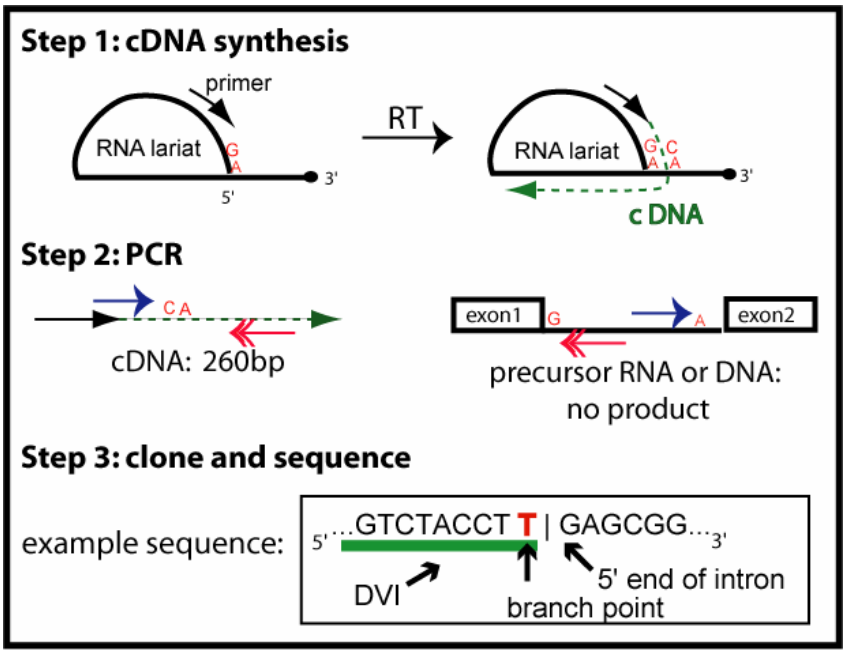


Figure 3-4. RT-PCR assay for mapping branch points. Step 1 – cDNA synthesis from an internal primer. Most extensions terminate after copying the first intron nt, but occasional copying of the branch nt permits formation of longer cDNAs such as indicated in green. Step 2 – PCR using primers that selectively amplify longer cDNAs from branched RNAs. Step 3 –PCR products of roughly the correct size are band isolated, cloned and sequenced. This sizing step permits ectopic branches to be detected without diverting much effort to artifactual RNAs that are present in most reactions where branching is inefficient.

In extensive experiments with lariat RNAs from plant introns it was found that a branch A specifies incorporation of dAMP in the cDNA instead of dTMP and that more than one A is occasionally inserted or that position is skipped (Vogel et al., 1997; Vogel & Borner, 2002b). Figure 3-5 shows an ethidium bromide stained gel on which RT-PCR reactions using *in vivo* yeast RNA templates were fractionated. Lane 1 shows that RNA from the wild type strain yields a strong signal of the expected length. That material was cloned and sequenced and the sense strands obtained have been summarized in Figure 3-6, line 1. The displayed sequence began with part of DVI followed by the sequence of the 5' end of the intron with the location of the branch nucleotide indicated by a letter in bold font. Each of

twenty clones of wild-type aI5γ *in vivo* lariats had a T at the position of the branch site due to incorporation of an A in the cDNA in response to the branch A.

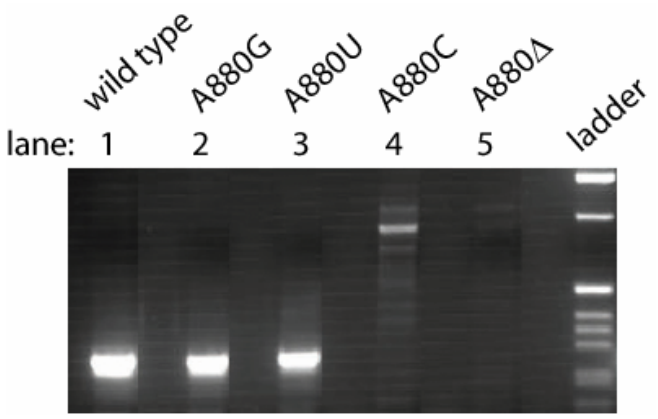


Figure 3-5. Calibrating the RT-PCR branch point mapping assay using branch nucleotide mutants. RT-PCR products run on a stained gel. The major band is ~260bp.

Our lab, in collaboration with the Pyle lab, has reported that the branch nucleotide mutants A880G and A880U (for diagrams, see Figure 3-7) can branch inefficiently *in vitro* at the mutated base (Liu *et al.*, 1997) and *in vivo* evidence indicated that each mutant branches at a low level (Podar *et al.*, 1998). I used these known mutants to determine how an RT responds to branch nucleotides besides A. Nineteen cDNA sequences amplified from *in vivo* RNA from A880G were sequenced and the results demonstrated that the RT incorporates a G when it encounters a branch G (Figure 3-6, line 2). Forty-five cDNA sequences amplified from *in vivo* RNA from A880U were sequenced and the results demonstrated that the RT incorporates an A or rarely a G when it encounters a branch U (Figure 3-6, lines 3 and 6, respectively). It is important to add that, of the eleven events depicted in line 6, only one came from the wild-type background. This was the only significant difference that I observed for all my branch point mapping results between the *MRS2-L232F* suppressor background

and the wild-type (see section 3-2-4 for a description of *MRS2-L232F*). The eight sequences in line 4 either resulted from skipping the branch U or from ectopic branching at U879 with incorporation of an A. The four sequences in line 5 appeared to result from skipping an ectopic branch at U879. I was unable to detect any lariats in the A880C or A880Δ mutant strains (see Figure 3-7) or from self- splicing reactions of those alleles (not shown, see section 3-2-9).

Line	allele	#	3' intron	5' intron
1	wild-type	20	AAG GTC TAC CT **T**	GAG CGG
2	A880G	19	AAG GTC TAC CT **C**	GAG CGG
3	A880U	22	AAG GTC TAC CT **T**	GAG CGG
4	A880U	8	AAG GTC TAC CT **-**	GAG CGG
5	A880U	4	AAG GTC TAC C **- -**	GAG CGG
6	A880U	11	AAG GTC TAC CT **C**	GAG CGG

Figure 3-6. Calibrating the RT-PCR branch point mapping assay using branch nucleotide mutants. Sequencing results. Results shown for mutant A880U are pooled from the WT and *MRS2-L232F* nuclear backgrounds (see section 3-2-4).

In summary, these controls demonstrated that the RT almost always incorporates a purine when it encounters a branched nucleotide; it incorporates an A opposite branch A, a G opposite branch G and an A or sometimes G opposite a branch U. As described in the plant studies of Vogel *et al*, I also observed instances when the RT skipped the branched nucleotide. Although I never observed skipping at the wild-type branch A. This absence of skipping at a wild-type aI5γ branch indicates that branch skipping may be a result of the RT copying a less-than-perfect DVI. This interpretation would indicate that many of the plant

introns studied by Vogel *et al*. have less-than-ideal DVI's. In general, my findings with alternative branch nucleotides aided in the interpretation of later sequences where the branch point was unknown. It appears that the identity of the branch nucleotide influences reverse transcription, but Watson-Crick pairing is not the basis of the nucleotide choice.

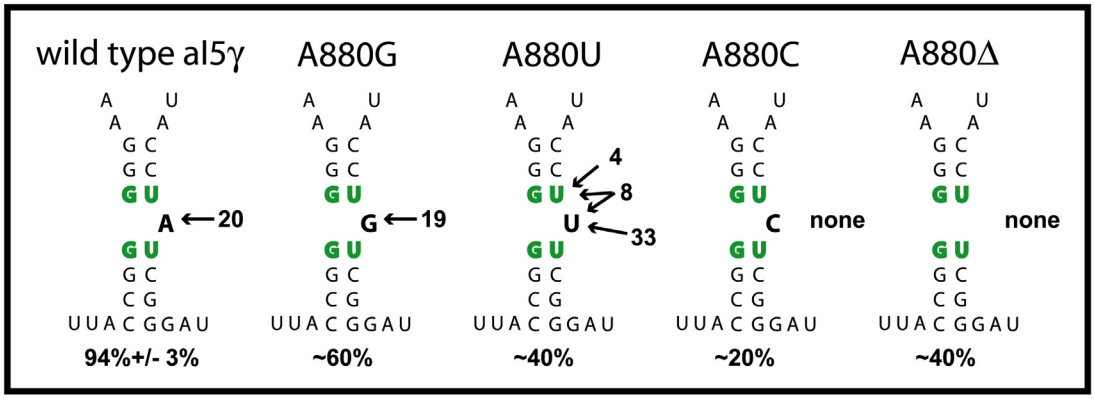


Figure 3-7. Branch nucleotide mutant diagrams. The number of sequences mapped is shown next to the arrow pointing to the branch location. A split arrow means that the sequence data could be interpreted as a splicing event at either site.

3-2-2 *In vivo* branching efficiency and fidelity are affected by the location of the bulged A.

Next I analyzed the role of the location of the unpaired A in DVI to determine the extent and location of branching *in vivo*. Mutants BrA⇑ and BrA⇓ (see Figure 3-8) were transformed into mtDNA and both resulting strains were found to be gly$^+$.

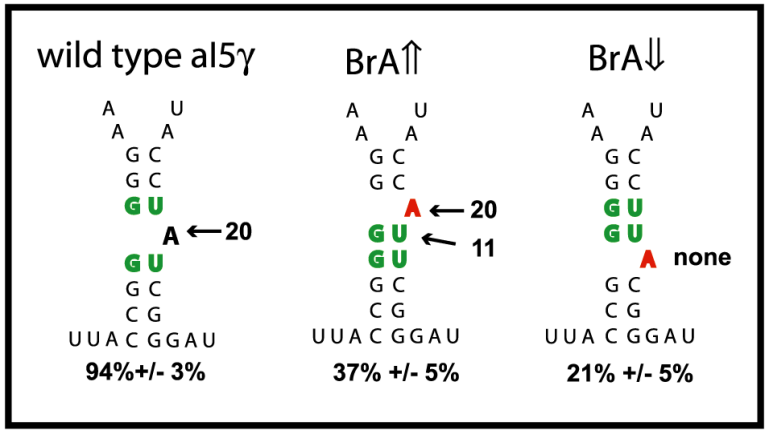


Figure 3-8. BrA⇑ and BrA⇓ Mutant diagrams. Average percent splicing from 3-5 independent blots is shown below each diagram (+/- SD). The number of sequences mapped is shown next to the arrow pointing to the branch location.

From RNA blots, those strains had 37% and 21% spliced mRNA respectively, compared to 94% for the control strain. The most abundant exon-containing transcript in each was pre-mRNA (Figure 3-11, lane 3 and Figure 3-16, lane 4 respectively). The extent of inhibition of splicing in strains BrA⇑ and BrA⇓ was similar to that of the branch nucleotide mutants A880G and A880U, as analyzed *in vivo* previously (Podar *et al.*, 1998), and repeated for this study (Figure 3-7). RNA from strain botGC, which has a reduced level of branching and second-step splicing (see section 3-2-6), was used as a size marker for linear vs. lariat IVS-E2 and IVS RNAs (Figure 3-11 and Figure 3-16). From visual inspection of the RNA blots,

it appears that both mutants had some linear IVS-E2 but no lariat IVS-E2 or excised intron RNA (Figure 3-11, lane 3 and Figure 3-16 lane 4). I concluded that mutants BrA⇑ and BrA⇓ partially inhibit splicing and strongly inhibit branching.

Wild-type nuclear background

Line	allele	#	3' intron	5' intron
1	wild-type	20	AAG GTC TAC CT **T**	GAG CGG
2	BrA⇑	4	AAG GTC TAC CA **T**	GAG CGG
3	BrA⇑	3	AAG GTC TAC C **- -**	GAG CGG
4	BrA⇑	2	AAG GTC TAC C **T**	GAG CGG
5	BrA⇑	2	AAG GTC TAC CA **C**	GAG CGG
6	BrA⇑	1	AAG GTC TAC C **TT**	GAG CGG
7	BrA⇓	none		

Figure 3-9. Mutants BrA⇑ and BrA⇓ RT-PCR branch point mapping results. Sequencing results from RT-PCR branch point mapping. Each sequence begins with the 3'end of the intron switches to the first nucleotide of the intron at the branch point. The nucleotides inserted opposite the branch point are highlighted in bold type. Lane 1 contains the wild-type reference sequence.

I then attempted to map rare branching events for mutants BrA⇑ and BrA⇓, neither of which had been successfully mapped *in vitro* (Chu *et al.*, 2001). RT-PCR reactions with *in vivo* RNA from both strains only yielded a signal for mutant BrA⇑ (BrA⇑ – Figure 3-11, panel E; BrA⇓ – Figure 3-16, panel D). Twelve BrA⇑ clones were obtained and sequenced, yielding five different sequences (Figure 3-9, lines 2-6). Based on the results that calibrate this assay (shown in Figure 3-6), I can interpret most of those sequences: BrA⇑ permits some branching at two sites in DVI, the shifted A and the U 3' of it (summarized in Figure 3-8). Figure 3-9, line 4 reports two events where branching occurred at the shifted A and a single

A was inserted by the RT; line 6 shows one event where two A's were inserted there; and line 3 shows three events where no nucleotide was inserted. The six sequences in lines 2 and 5 resulted from branching at U880 in the mutated DVI (see Figure 3-8) where either A or G was inserted. In sum, six events occurred at the shifted A, and six events occurred at the U 3' of the shifted A (i.e., the natural branch location).

3-2-3 Splicing and branching of BrA⇑ derivative strains

The *in vitro* analysis of BrA⇑ showed that repairing the environment of the shifted A (*i.e.*, the neighboring basepair) greatly improved the extent of branching (Chu *et al.*, 2001). Specifically, the GC above the shifted A was changed to GU (mutant BrA⇑^{C878U}) (see Figure 3-10). *In vitro* that mutation increased k_{br} about 40-fold (to ~20% of the control). Mapping the branch site in self-spliced lariat RNAs using the DNAzyme-assisted method showed that branching occurs at the shifted A. In an attempt to further improve *in vivo* branching at the shifted A, we made mutant BrA⇑$^{C878U}{}_{botGC}$ (see Figure 3-10). This mutation restores the wild type configuration of 5 GC and 2 GU pairs in the branch-containing stem. *In vitro* this mutation (BrA⇑$^{C878U}{}_{botGC}$) branched more poorly than mutant BrA⇑^{C878U}, only increasing k_{br} about 8-fold over that of BrA⇑. The branch site for mutant BrA⇑$^{C878U}{}_{botGC}$ was also mapped to the ectopic A. The *in vitro* analysis demonstrated that the base pair above the branch is important for efficient branching.

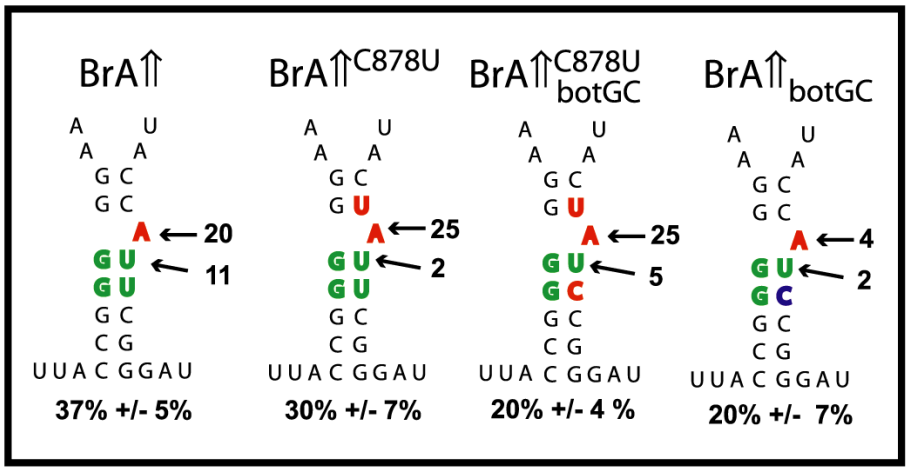


Figure 3-10. BrA⇑ mutant diagrams. Average percent splicing from 3-5 independent blots is shown below each diagram. The number of sequences obtained by RT-PCR branch point mapping is shown next to the arrow pointing to the branch location.

I proceeded to the *in vivo* analysis of these complex mutations with the addition of mutant BrA⇑$_{botGC.}$ Mutant BrA⇑$_{botGC}$ was added to determine the contribution of the botGC change to the phenotype of mutant BrA⇑ ^{C878U} $_{botGC.}$ Mutants BrA⇑ ^{C878U}, BrA⇑ ^{C878U} $_{botGC}$ and BrA⇑$_{botGC}$ were transformed into mitochondria and I analyzed the splicing phenotypes of the resulting gly^{+} strains. As shown in Figure 3-11A and graphically in Figure 3-13, these strains spliced about as well as the original BrA⇑ strain. However, mutant BrA⇑ ^{C878U} accumulated a very low level of excised linear intron RNA (see overexposed portion of the blot hybridized with the intron probe shown in Figure 3-11C, lane 5). The RT-PCR branching assay yielded a stronger signal for BrA⇑ ^{C878U} than was obtained with BrA⇑ (Figure 3-11E, compare lanes 3 and 5). Branch site choice was easily mapped for BrA⇑ ^{C878U} with 9 out of 11 events occurring at the shifted branch A (Figure 3-12, line 7).

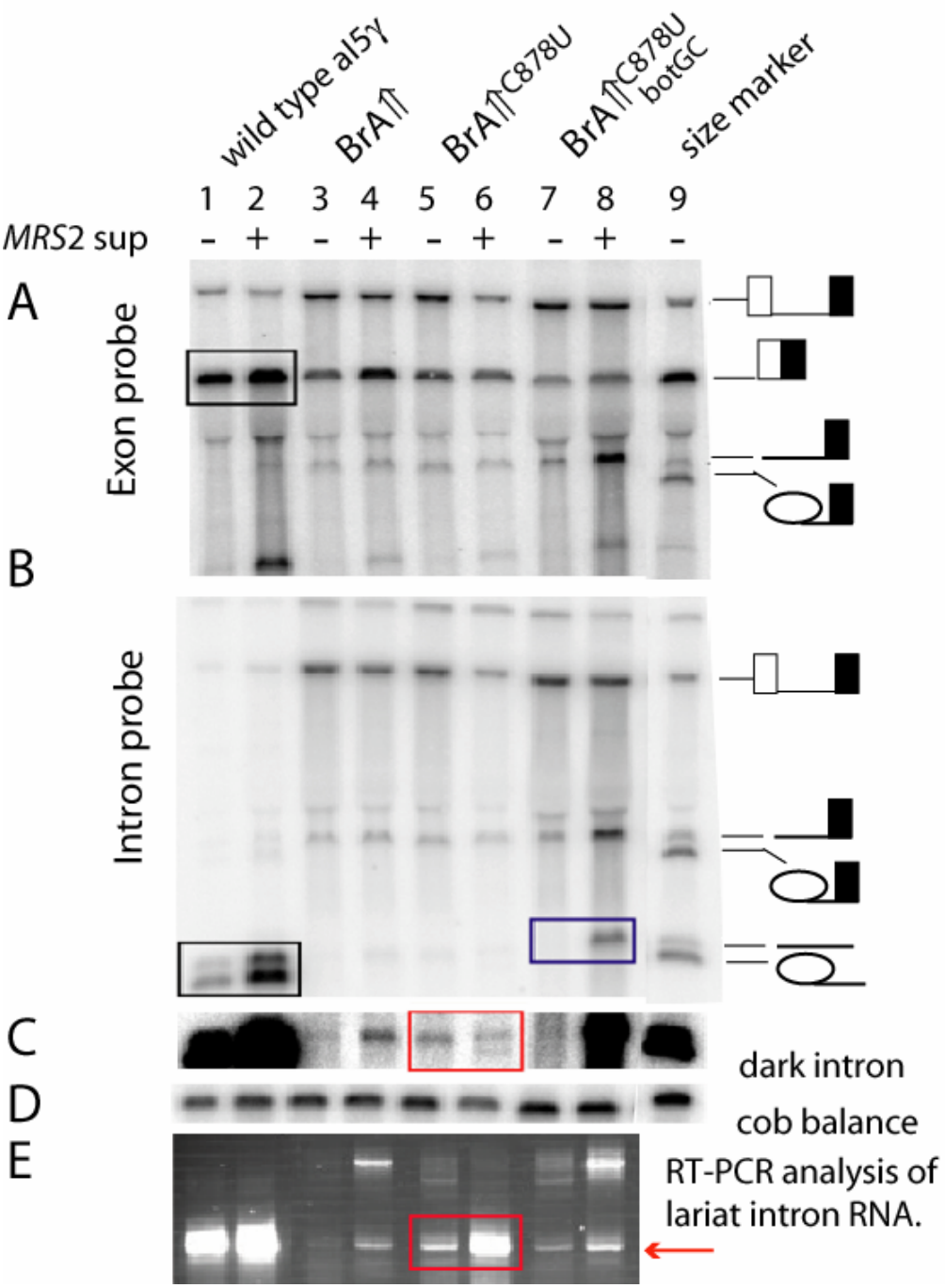


Figure 3-11. Analysis of BrA⇑ mutant strains. These mutations are depicted in Figure 3- 10. Strains were analysed in both the wild-type MRS2 and MRS2-L232F backgrounds. Key bands altered in the MRS2-L232F background are enclosed in a box for easy comparison. (A) RNA blot probed with primers complementary to COXI exon 6. (B) RNA blot probed with primers complementary to intron aI5γ. (C) Darker exposure of the intron RNA signal of the intron-probed blot. (D) RNA blot probed with a primer complementary to the COB gene (be specific, was it exon 4?). This blot is used to estimate the relative amounts of mitochondrial RNA the samples. (E) RT-PCR products used to map branch points fractionated on a stained agarose gel.

The two events shown in Figure 3-12, line 8, appeared to have resulted from a novel sort of ectopic branching or from an RT artifact. This was the only instance I observed of a sequence that would place the branch point in the upstream un-paired loop. RNAs from mutants BrA⇑ $^{C878U}{}_{botGC}$ and BrA⇑$_{botGC}$ yielded little RT-PCR signal although I was able to map them (Figure 3-12, Figure 3-10).

Mutant BrA⇑ $^{C878U}{}_{botGC}$ branched at the shifted A in 6 out of 9 events and at the 3'U in the remaining 3 events. Mutant BrA⇑$_{botGC}$ branched at the shifted A in 4 out of 6 events and at the 3' U in the remaining 2 events. In sum, the repaired mutant BrA⇑ ^{C878U} branched more often and had a higher fraction of branching events at the shifted A relative to strain BrA⇑. The addition of the botGC mutation in the presence of a shifted branch A did not improve splicing or branching relative to the simple BrA⇑ strain. These results demonstrate that a misplaced bulged A at nucleotide 879 can specify branching if it is situated in an environment resembling that of A880 in the wild-type intron.

Wild-type nuclear background

Line	allele	#	3' intron	5' intron
1	wild-type	20	AAG GTC TAC CT **T**	GAG CGG
2	BrA⇑	4	AAG GTC TAC CA **T**	GAG CGG
3	BrA⇑	3	AAG GTC TAC C **-**	GAG CGG
4	BrA⇑	2	AAG GTC TAC C **T**	GAG CGG
5	BrA⇑	2	AAG GTC TAC CA **C**	GAG CGG
6	BrA⇑	1	AAG GTC TAC C **TT**	GAG CGG
7	BrA⇑^{C878U}	9	AAG GTC TAC T **T**	GAG CGG
8	BrA⇑^{C878U}	2	AAG GT	GAG CGG
9	BrA⇑$^{C878U}_{botGC}$	6	AAG GTC TAC T **T**	GAG CGG
10	BrA⇑$^{C878U}_{botGC}$	2	AAG GTC TAC TA **C**	GAG CGG
11	BrA⇑$^{C878U}_{botGC}$	1	AAG GTC TAC TA **T**	GAG CGG
12	BrA⇑$_{botGC}$	2	AAG GTC TAC C **T**	GAG CGG
13	BrA⇑$_{botGC}$	2	AAG GTC TAC C **-**	GAG CGG
14	BrA⇑$_{botGC}$	1	AAG GTC TAC CA **T**	GAG CGG
15	BrA⇑$_{botGC}$	1	AAG GTC TAC CA **C**	GAG CGG

Figure 3-12. BrA⇑ mutant strains RT-PCR branch point mapping results. Sequencing results from RT-PCR branch point mapping. Each sequence begins with the 3'end of the intron switches to the first nucleotide of the intron at the branch point. The nucleotides inserted opposite the branch point are highlighted in bold type. Lane 1 contains the wild-type reference sequence.

3-2-4 An MRS2 nuclear suppressor partially rescues intron RNA in BrA⇑ mutant strains.

Our lab has previously reported three missense mutations of the *MRS*2 gene that suppress the splicing defects of a number of point mutants of aI5γ (Schmidt *et al.*, 1998). The *MRS*2 gene codes for a Mg^{2+} transporter of the mitochondrial inner membrane (Bui *et al.*, 1999) and missense suppressor alleles alter the protein in such a way that the level of the protein increases substantially (Schmidt *et al.*, 1998) and the intra-mitochondrial concentration of Mg^{2+} increases measurably (Gregan *et al.*, 2001).

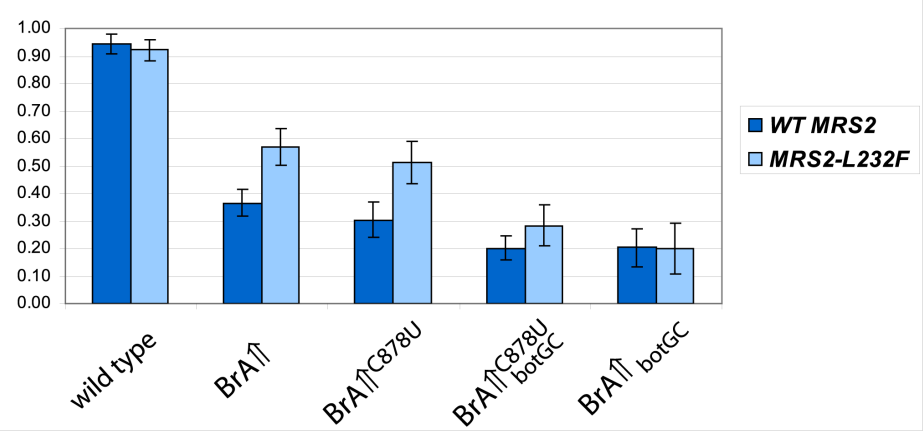


Figure 3-13. Percent splicing of the BrA⇑ mutant strains in the wild-type and *MRS2-L232F* nuclear backgrounds.

I tested whether our strongest suppressor allele, *MRS2-L232F* (Schmidt *et al.*, 1998), influences splicing by the BrA⇑ mutant strains. As shown in Figure 3-11 and Figure 3-13, this suppressor improved the extent of splicing in the BrA⇑ and BrA⇑ ^{C878U} strains from ~30% to ~60% spliced *COXI* mRNA. The *MRS2-L232F* allele had no significant effect on net splicing in the BrA⇑$^{C878U}{}_{botGC}$ and BrA⇑$_{botGC}$ strains. There was some IVS-E2 RNA in

the mutant strains with or without *MRS2-L232F*, but it was linear (Figure 3-11A, lanes 4 and 6, vs. the marker for lariat IVS-E2 in lane 7). Some linear intron RNA was also rescued in all strains. The strongest rescue of linear IVS-E2 and IVS was seen in mutant BrA⇑^{C878U} $_{botGC}$. This rescue was unexpected because linear intron RNA is normally degraded and so is not observed on northern blots. The most interesting finding, however, was that lariat intron RNA is rescued in strain BrA⇑^{C878U} such that a faint lariat band is observed on the RNA blot (Figure 3-11A, lane 6, dark inset). I also placed the wild type intron into the *MRS2-L232F* background and found that both lariat and linear intron levels are increased (Figure 3-11A and Figure 3-16A, lanes 1-2). In wild-type, the "linear" is generally believed to be broken lariat. The *MRS2-L232F* background allows direct confirmation by northern blot of the previous observation that the GU base pair above the branch dramatically improves branching at an ectopic branch A.

The RT-PCR branch point mapping assay yields a stronger signal in the *MRS2-L232F* background than for the same alleles in the wild-type nuclear background (Figure 3-11E). Five different sequences were obtained among the 18 lariat clones analyzed from strain *MRS2-L232F* BrA⇑ (Figure 3-14, lines 2-6), nearly all of which were also obtained from the *MRS2* control strain, albeit in different proportions. The 12 events in lines 2 and 4 reflect branching at the shifted A, and the other six events (lines 3, 5 and 6) reflect branching at U880. From strain *MRS2-L232F* BrA⇑^{C878U}, 16 of the 18 sequences reflect branching at the shifted A (lines 7 and 9); two sequences, because the terminal "A" was not changed to a "T", probably reflect branching at U880 3' of the shifted A (line 8). Out of 21 clones derived from strain *MRS2-L232F* BrA⇑^{C878U} $_{botGC}$ 19 reflect branching at the shifted A and 2 branching at

the 3'U. Though branching in the *MRS2* and *MRS2-L232F* backgrounds was at different levels, the branch sites used were very similar, perhaps with a greater preference for the shifted A in the suppressor background.

MRS2 L232F nuclear background

Line	allele	#	3' intron	5' intron
1	wild-type	7	AAG GTC TAC CT **T**	GAG CGG
2	BrA⇑	3	AAG GTC TAC C A **T**	GAG CGG
3	BrA⇑	2	AAG GTC TAC C **-**	GAG CGG
4	BrA⇑	10	AAG GTC TAC C **T**	GAG CGG
5	BrA⇑	1	AAG GTC TAC C A **C**	GAG CGG
6	BrA⇑	1	AAG GTC TAC C A **-**	GAG CGG
7	BrA⇑^{C878U}	15	AAG GTC TAC T **T**	GAG CGG
8	BrA⇑^{C878U}	2	AAG GTC TAC T **A**	GAG CGG
9	BrA⇑^{C878U}	1	AAG GTC TAC T **-**	GAG CGG
9	BrA⇑$^{C878U}_{botGC}$	18	AAG GTC TAC T **T**	GAG CGG
10	BrA⇑$^{C878U}_{botGC}$	1	AAG GTC TAC T **C**	GAG CGG
11	BrA⇑$^{C878U}_{botGC}$	1	AAG GTC TAC T **A**	GAG CGG
12	BrA⇑$^{C878U}_{botGC}$	1	AAG GTC TAC TA **T**	GAG CGG
13	BrA⇑$_{botGC}$	none		

Figure 3-14. RT-PCR branch point mapping sequencing results for the BrA⇑ mutant strains in the *MRS2-L232F* nuclear background. Sequencing results from RT-PCR branch point mapping. Each sequence begins with the 3'end of the intron switches to the first nucleotide of the intron at the branch point. The nucleotides inserted opposite the branch point are highlighted in bold type. Lane 1 contains the wild-type reference sequence.

3-2-5 Splicing and branching of complex BrA⇓ mutant strains

I next analyzed strain BrA⇓, and a derivative strain in which the environment of BrA⇓ was altered by mutating C882 to U and U879 to C (BrA⇓$^{topGC}_{C882U}$; see Figure 3-15). I hoped to improve branching and map branch points in the derivative mutant, as I had done for BrA⇑. As previously noted, BrA⇓ has a splicing defect typical of other strains with mutant branch points. Surprisingly, mutant BrA⇓$^{topGC}_{C882U}$ barely splices at all (Figure 3-16A, lane 8). To determine the cause of this unexpected finding, I analyzed each mutation separately, namely BrA⇓^{topGC} and BrA⇓$_{C882U}$. Of those three alleles, only BrA⇓^{topGC} spliced at a consistently detectable level, though significantly less than BrA⇓.

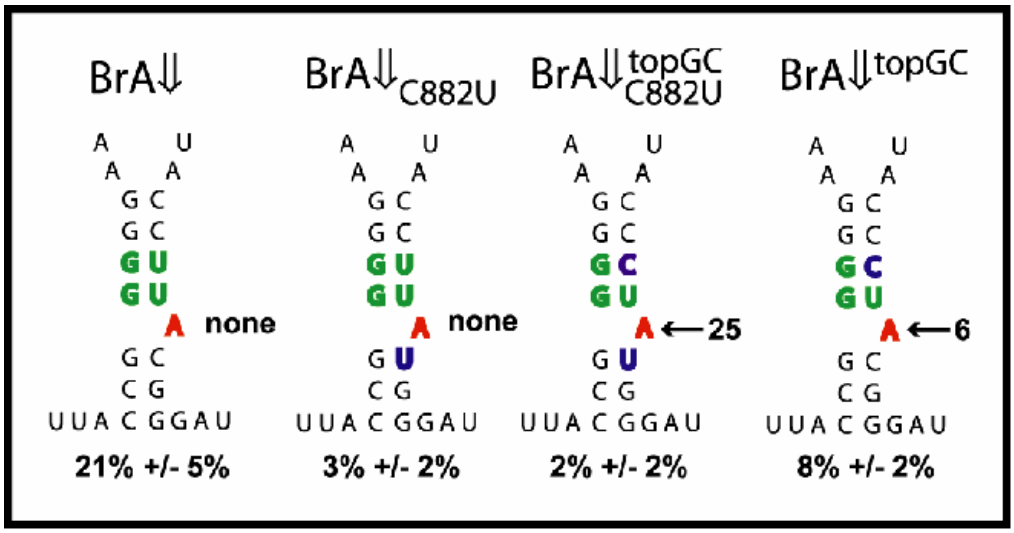


Figure 3-15. BrA⇓ mutant diagrams. Nucleotides highlighted in blue indicate the additional mutations made for each strain. Average percent splicing from 3-5 independent blots is shown below each diagram. The number of sequences obtained by RT-PCR branch point mapping is shown next to the arrow pointing to the branch location.

These data indicated that the C882U mutation is deleterious for splicing. This was not a complete block, because the *MRS2-L232F* background rescued some splicing by both

BrA⇓$^{topGC}_{C882U}$ and BrA⇓$_{C882U}$. Although none of the BrA⇓ mutant alleles had any branched RNA detectable on RNA blots (Figure 3-16A, lanes 4-11), lariat RNAs were cloned by RT-PCR for BrA⇓^{topGC} and for BrA⇓$^{topGC}_{C882U}$, the level of the latter being somewhat improved in the *MRS2-L232F* background (Figure 3-16B, lanes 9-10). As shown in Figure 3-17A, lines 11 and 12, and Figure 3-17B, lines 12-13, all of the sequences reflected branching at the shifted A. In the single event shown in Figure 3-17B, line 13, the RT inserted a C in response to the branched A, which was the only time in this study that a pyrimidine was inserted in a cDNA in response to a branch. It appears that shifting the unpaired A toward the base of the DVI helix was more inhibitory to both splicing and branching than is shifting it upward, but even so, changing the environment of the shifted bulged A so that it is flanked by GU pairs improved branching there, if the destabilizing effect of the flanking GU pairs was compensated by an additional GC pair in DVI.

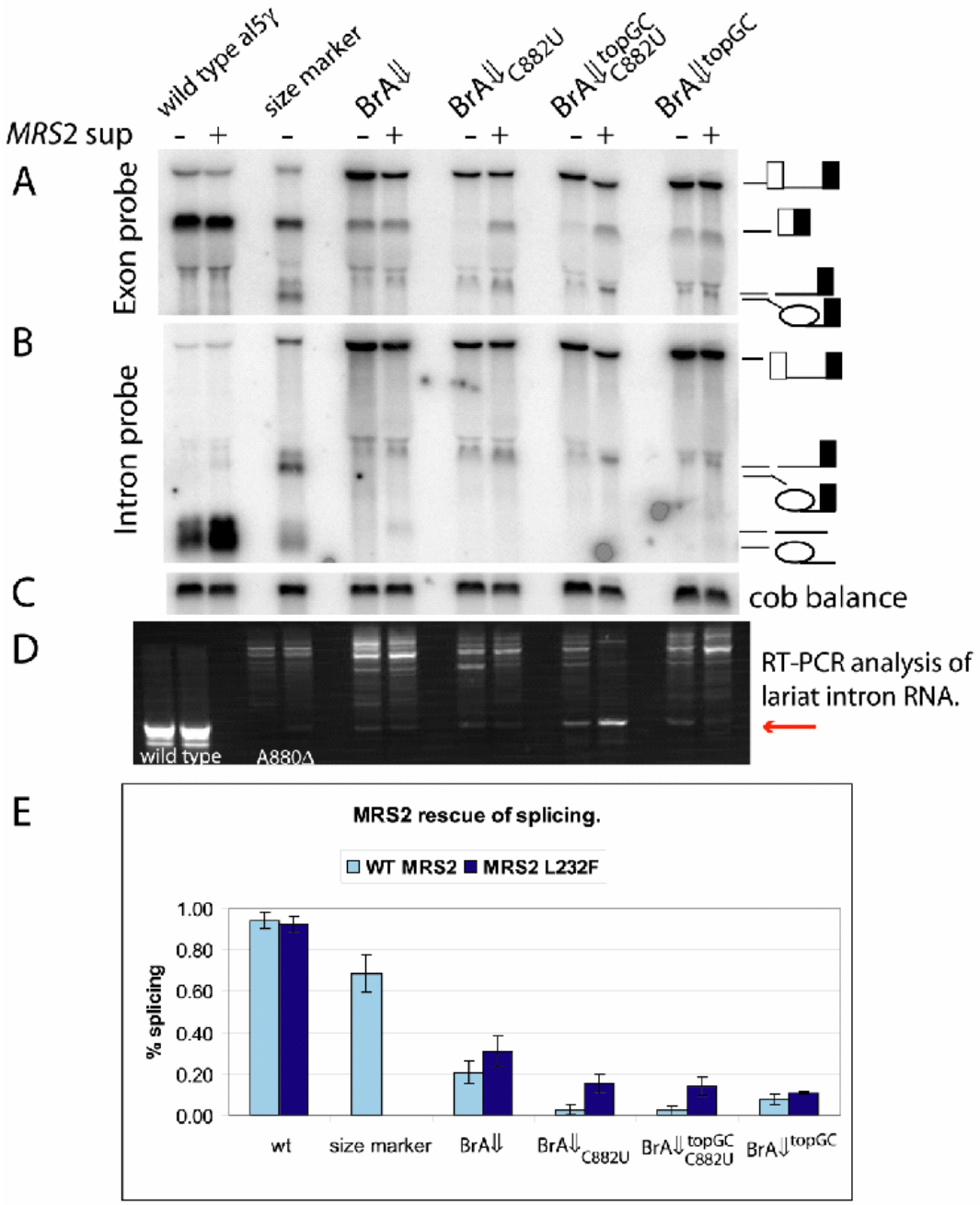


Figure 3-16. Analysis of BrA⇓ mutant strains. These mutations are depicted in Figure 3- 12. Strains were analyzed in both the wild-type *MRS2* and *MRS2-L232F* backgrounds. (A) RNA blot probed with primers complementary to *COXI* exon 6. (B) RNA blot probed with primers complementary to intron aI5γ. (C) RNA blot probed with a primer complementary to the *COB* gene. This probe is used to balance the input material. (D) RT-PCR products used to map branch points fractionated on a stained agarose gel. Analysis of cloned material from gels such as this one showed that the very faint band observed in some lanes is not lariat (see section 3-2-9). (E) Bar graph of the quantification of 3-5 independent blots. % splicing is mRNA over total RNA.

A

Wild-type nuclear background

Line	allele	#	3' intron	5' intron
1	wild-type	20	AAG GTC TAC CT **T**	GAG CGG
2	BrA ⇓	none		
3	BrA $\Downarrow_{C882U}$	none		
4	BrA $\Downarrow^{topGC}_{C882U}$	8	AAG GTC TAC CCT **T**	GAG CGG
5	BrA $\Downarrow^{topGC}$	6	AAG GTC TAC CCT **T**	GAG CGG

B

MRS2 L232F nuclear background

Line	allele	#	3' intron	5' intron
1	wild-type	7	AAG GTC TAC CT **T**	GAG CGG
2	BrA⇓	none		
3	BrA$\Downarrow_{C882U}$	none		
4	BrA$\Downarrow^{topGC}_{C882U}$	16	AAG GTC TAC CCT **T**	GAG CGG
5	BrA$\Downarrow^{topGC}_{C882U}$	1	AAG GTC TAC CCT **G**	GAG CGG
6	BrA ^{topGC}	none		

Figure 3-17. BrA⇓ mutant strains. RT-PCR branch point mapping results. (**A**) Results from the *MRS2* nuclear background. (**B**) Results from the *MRS2-L232F* nuclear background. The nucleotides inserted opposite the branch point are highlighted in bold type.

3-2-6 Branch site environment mutations affect splicing and branching efficiency *in vivo*

These experiments extended to the *in vivo* situation the earlier *in vitro* finding that GU pairs above and below a shifted branch A are very important for branching. Here, I present evidence that both GU pairs flanking the branch A of wild-type aI5γ (A880) are important for efficient splicing and branching *in vivo*. From phylogenetic data assembled for the *in vitro* paper, 11 of 16 possible sequences are present at the base pair above the branch A (nt 856 and 879 in aI5γ) in at least one group II intron (Figure 3-18). The most frequent pair is GU (41%) and the remaining introns are divided roughly equally between Watson-Crick and non-Watson-Crick pairs.

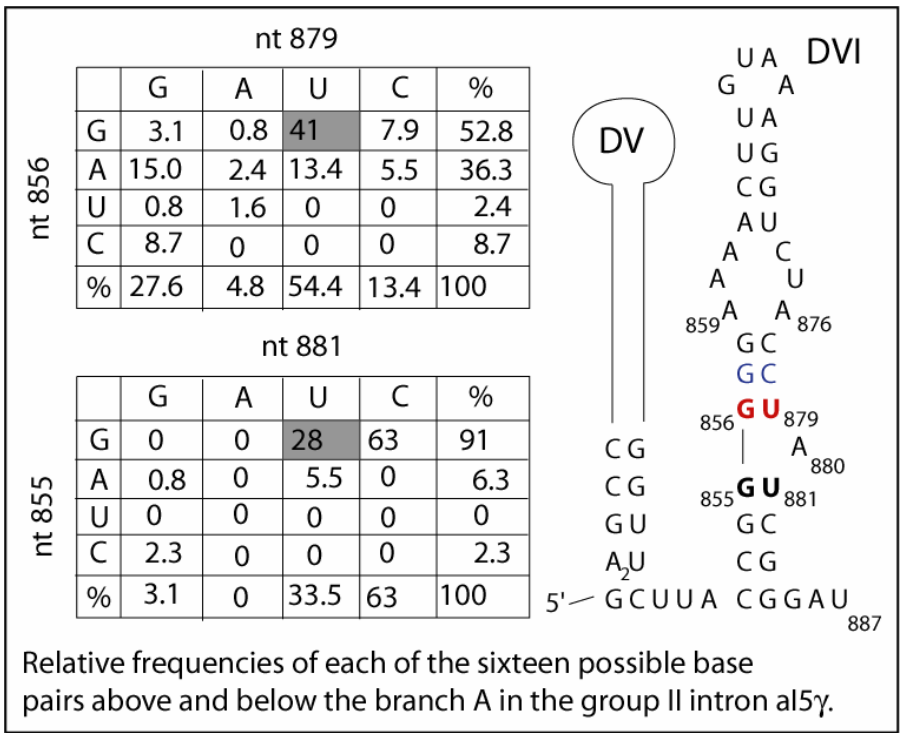


Figure 3-18. Phylogenetic analysis of the branch environment.

The base pair below the branch A (nt 855 and 881 in aI5γ) was much more constrained, with only 5 of the 16 possible pairs present in this sample of introns. GC and GU accounted for 91% of the introns and though GU is present in wild-type aI5γ, GC is by far the most common pair at that position.

Dr. QingShi Zhao created mutants in which the GU pairs that flank A880 in the wild-type intron were changed singly to GC (mutants U879C and U881C, or topGC and botGC, respectively and previously constructed by Dr. Mircea Podar). These mutants were transformed into mtDNA (for structures see, Figure 3-19). I analyzed the resulting strains. They were found to be respiratory competent, and RNA blots show that each mutant spliced aI5γ quite well, though each had a clear splicing defect (Figure 3-20A, B, lanes 2-3 and Figure 3-20D). Compared with the wild-type control, RNA from topGC contained an elevated level of pre-mRNA and had a low but elevated level of IVS-E2 RNA. Despite the high level of spliced mRNA in this mutant strain, it is noteworthy that hardly any excised intron RNA accumulated (Figure 3-20B, compare lanes 1 and 2). This gel system distinguished branched from linear IVS and IVS-E2 RNAs for this intron (Podar *et al.*, 1998), and based on the motilities of those products, the IVS-E2 in topGC is linear but some of the low level of IVS RNA is branched. Those inferences are confirmed by analysis of these RNAs on polyacrylamide gels where lariat RNAs migrate more slowly than pre-mRNA (not shown but see Podar *et al.* 1998). Comparisons of the levels of mRNA and excised intron RNA in the mutant strains relative to their levels in the wild-type strain indicated that the mutant has ~70% of the wild-type level of *COXI* m*RNA* but only ~20% of the expected level of IVS RNA (Figure 3-20D). Thus, the *topGC* mutation partially inhibited splicing,

strongly inhibited branching and destabilized the excised intron RNA. The botGC mutant had less pre-mRNA than topGC but splices comparably at ~70% (Figure 3-20A, lane 3). It had more IVS-E2 RNA than pre-mRNA, showing that it had a slight defect in the first splicing reaction and a greater defect in the second splicing reaction. On the gel shown, the IVS-E2 RNA is mostly branched (Figure 3-20A, B, lane 3) and that observation was confirmed by analyzing this RNA on a polyacrylamide gel (not shown). This mutant accumulated at least four times the level of topGC IVS RNA, and much of it was branched (Figure 3-20B, line 3). BotGC did not appear to alter the branching reaction, and its main effect is inhibition of the second splicing step. Thus, both mutations of the environment of the branch A of aI5γ affected the efficiency of splicing, but in distinguishable ways. A strain in which both GU pairs flanking A880 were changed to GC (U879C U881C or 2XGC) grew well on glycerol medium, but was significantly inhibited for both splicing and branching (Podar *et al.*, 1998).

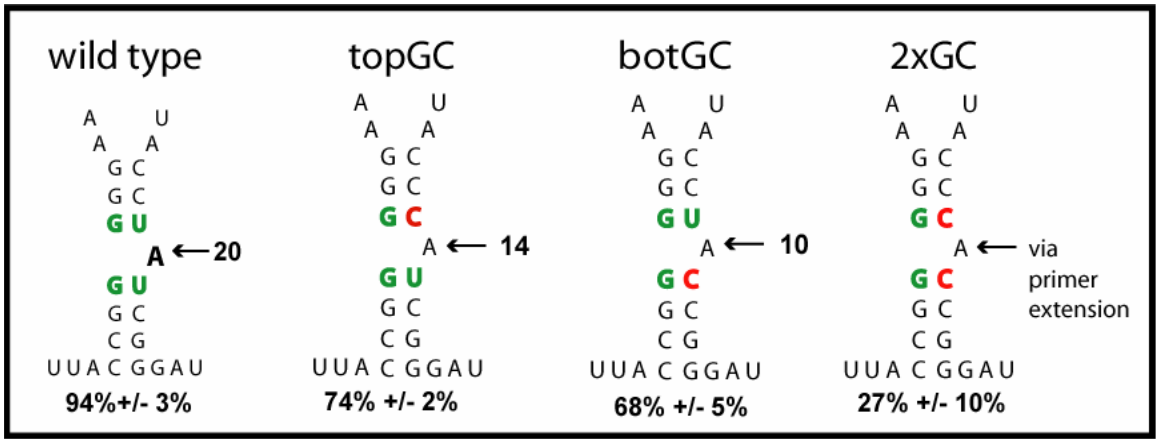


Figure 3-19. Branch environment mutant diagrams. Average percent splicing +/- one standard deviation from 3-5 independent blots is shown below each diagram. The number of sequences obtained by RT-PCR branch point mapping is shown next to the arrow pointing to the branch location.

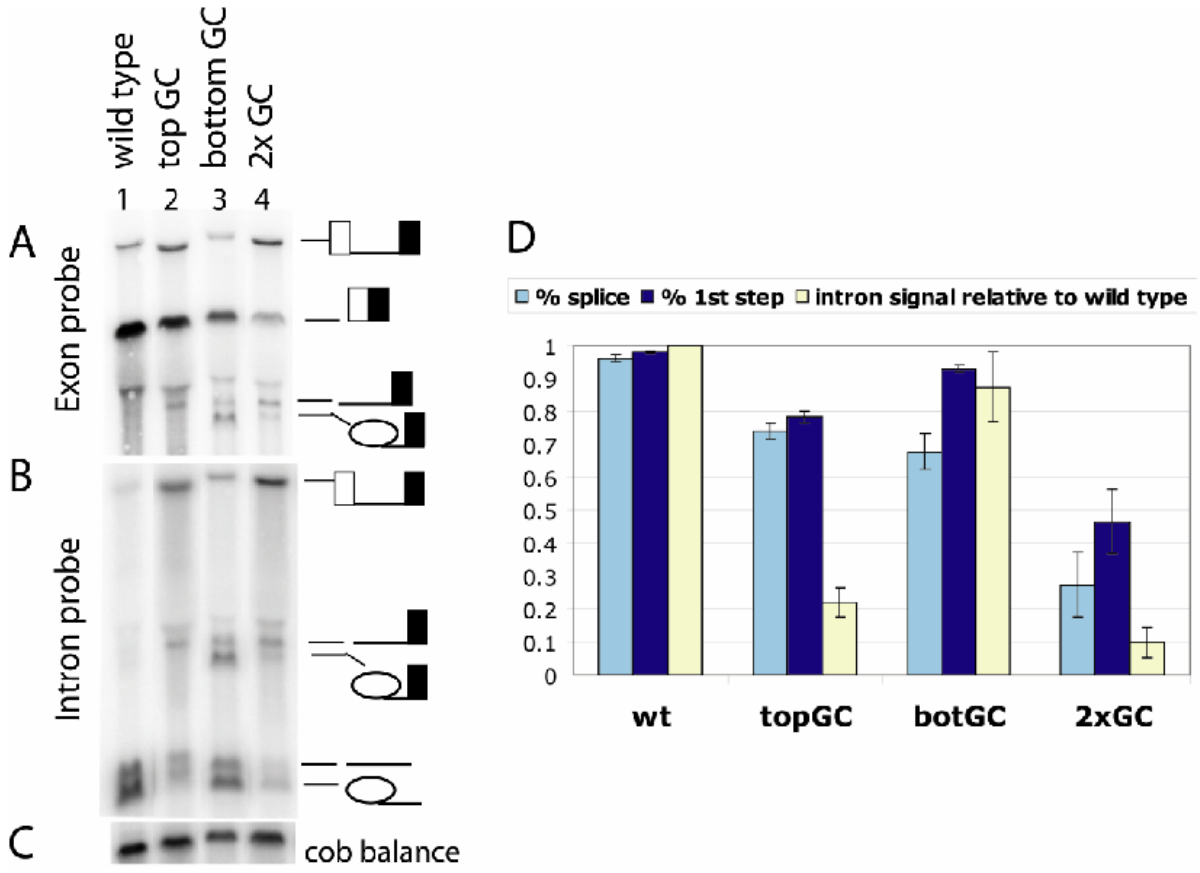


Figure 3-20. Analysis of branch environment mutants. Diagrams of these mutations are shown in Figure 3- 19. (A) RNA blot probed with primers complementary to *COXI* exon 6. (B) RNA blot probed with primers complementary to intron aI5γ. (C) RNA blot probed with a primer complementary to *COB* exon 6. (D) Quantification of 6 independent blots. Percent splicing is calculated as (spliced message) / (total signal) from blots such as shown in panel A. Percent first splicing step is calculated as (spliced message + IVS-E2) / (total signal). The relative intron value is derived from the average of four independent blots such as shown in panel B balanced within a factor of 2. The total signal was normalized using the *COB* mRNA level from panel C or the wild-type level of intron set to equal 1. This is a relative measure of intron abundance and should not be quantitatively compared to the level of spliced message.

As shown in Figure 3-20A, lane 4, mutant 2XGC had about 20% of the wild-type level of mRNA. It had about the same amount of IVS-E2 RNA as spliced mRNA, showing that it had a significant second step defect. It accumulated a very low level of lariat RNA, much less than expected based on the amount of spliced mRNA that is present. Some of the IVS-E2 RNA was shown previously to be branched at A880, but most of the IVS-E2 RNA was linear, demonstrating that the branching and hydrolysis pathways can coexist *in vivo* (Podar *et al.*, 1998). The defects of 2xGC were a combination of those observed for topGC and botGC, further demonstrating that the nucleotides above and below the branch were not equivalent. Even though the alleles that were analyzed here exist in other group II introns, these data show that the environment of the wild-type branch A is important for splicing and branching by this intron. RT-PCR followed by sequencing of a total of 25 cDNA clones indicated that these three mutants (2xGC, topGC, botGC) branch exclusively at A880 (sequences not shown).

3-2-7 The yeast nuclear debranching enzyme (Dbr1p) debranches group II introns.

Our lab has previously characterized other DVI branch mutants of aI5γ that accumulate mainly linear IVS-E2 and have little or no detectable excised intron RNA – the defining *in vivo* syndrome for splicing by hydrolysis (Podar *et al.*, 1998). One of the supporting arguments for the existence of splicing by hydrolysis *in vivo* was the demonstration that the disruption of the gene coding for the nuclear debranching enzyme (*DBR*1) did not rescue lariat RNA in those mutants (Podar *et al.*, 1998). The *Δdbr1* experiment reported in that

study was required by the reviewers, even though there was no prior evidence suggesting that Dbr1p has any function in mitochondria. In order to rule out the possibility that the branch nucleotide mutants produced unstable branches or that perhaps ectopic branches were inherently unstable, I analyzed all the branch point mutants described in this thesis in the *Δdbr*1 background. In additon to confirming the knockout by PCR, I also probed for the nuclear actin intron to confirm that it accumulated in the deletion strain (Chapman & Boeke, 1991). The *Δdbr*1 derivatives of the branch point mutants did not significantly alter their RNA phenotypes. In particular, *Δdbr*1 did not rescue any lariat IVS-E2 or lariat IVS RNA in the branch nucleotide mutants or increase the level of spliced mRNA.

Our lab has encountered a number of mutations of aI5γ outside of DVI that splice well but accumulate little excised intron RNA (Boulanger *et al.*, 1995). One good example is a DV mutant (A2U U33A, here called D5UA) in which the second base pair of the DV helix was mutated from AU to UA (see Figure 3-1). As shown in Figure 3-21, lane 9, D5UA spliced very efficiently in a *DBR*1 strain but accumulated only a little excised, branched, intron RNA. D5UA RNA self-spliced well and had no branching defect, nor was its excised lariat RNA inherently prone to spontaneous debranching *in vitro* (Peebles *et al.*, 1995). In a *Δdbr*1 D5UA strain, the level of splicing was not altered but the level of excised intron lariat increased 5-10-fold, reaching about 50% of the control level (Figure 3-21, lane 10). RT-PCR branch point mapping of mutant D5UA in both the *DBR1* and *Δdbr*1 backgrounds revealed that branching occurred only at the branch A (9 clones each). This finding confirmed that ectopic branches are not intrinsically more unstable than branching at A880.

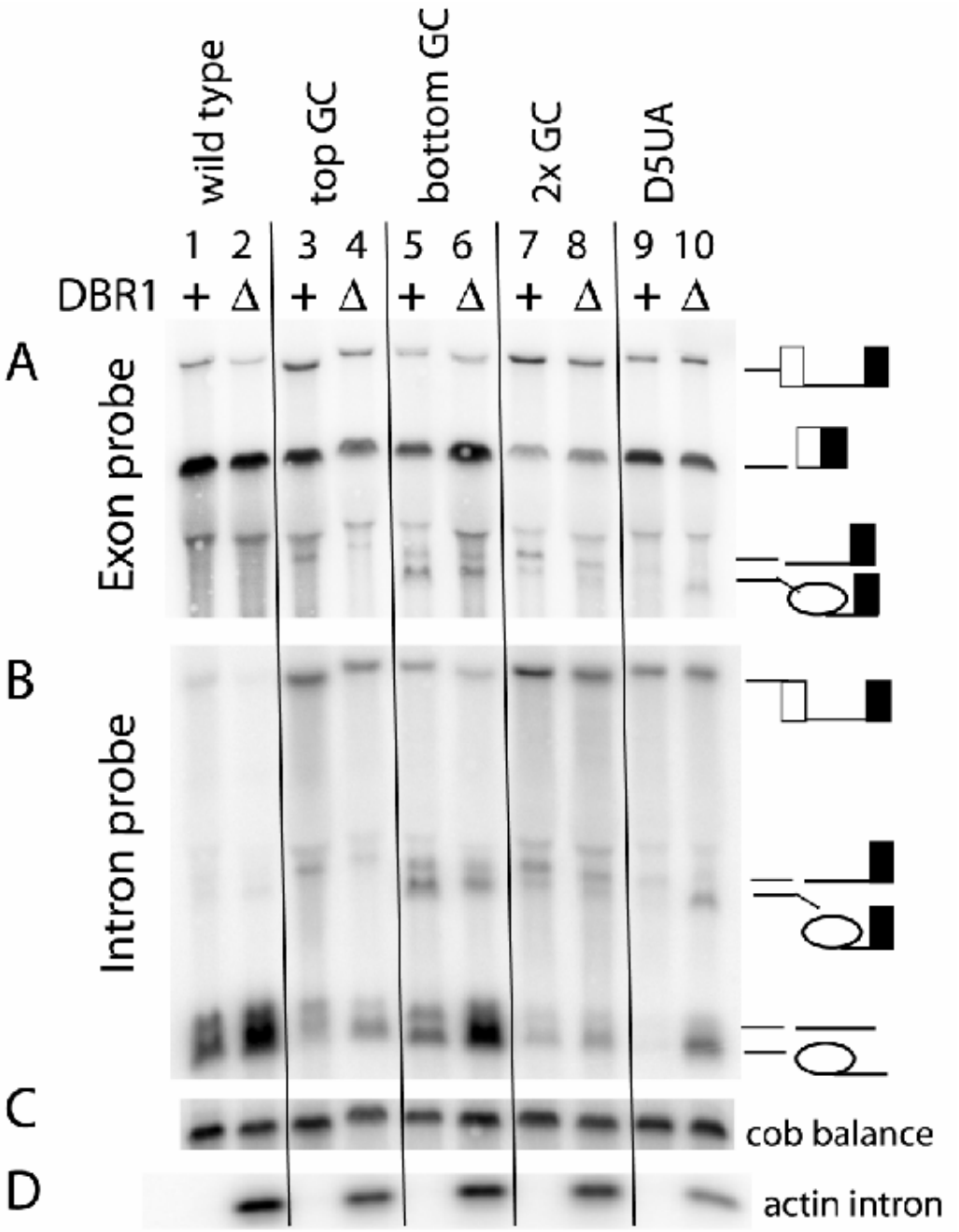


Figure 3-21. Effects of DBR1 alleles on splicing phenotypes of branch environment mutants. The branch environment mutations are depicted in Figure 3- 19. Mutant D5UA is a flip of the terminal base pair in DV- see Figure 3-1. (A) RNA blot probed with primers complementary to *COXI* exon 6. (B) RNA blot probed with primers complementary to intron aI5□. (C) RNA blot probed with a primer complementary to *COB* exon ?. (D) In the absence of Dbr1p, nuclear introns accumulate. shown is an RNA blot of the same samples run on a separate gel probed for actin intron RNA (nuclear spliceosomal intron) to confirm that Dbr1p is absent.

I also found that even wild-type aI5γ is not completely protected from debranching, as was previously thought. In a *Δdbr*1 strain, the level of wild-type intron RNA was consistently twice that of the corresponding *DBR1* strain (Figure 3-21, lanes 1-2). This increase was not due to increased splicing or transcription because the level of spliced message remained the same. Given that wild-type aI5γ intron RNA increases twofold in the absence of Dbr1p, the finding that intron RNA also increases ~twofold in botGC and topGC strains only serves to confirm that intron lariats are normally incompletely degraded in the presence of Dbr1p. Don Yoo, a summer research fellow, showed that a GFP-tagged Dbr1p localizes to the mitochondria. Unfortunately the GFP-*DBR*1 construct was non-functional, such that his finding is not conclusive. Taken together, however, these data suggest that a fraction of Dbr1p is localized to the mitochondrial matrix where it can debranch group II intron RNA lariats. Although we have no direct evidence for such an inference, it follows that the wild-type intron lariat is protected from debranching and that some mutations interfere with that protection leading to increased debranching followed by degradation of the intron RNA.

3-2-8 Some complex DVI mutations support ectopic branching *in vitro* but not *in vivo*

Chu *et al.* (2000) reported a number of other derivatives of the BrA⇑ and BrA⇓ mutations that support low levels of ectopic branching *in vitro* (RNAs 4A through 4E, Figure 3-22 here, also see (Chu *et al.*, 2001)).

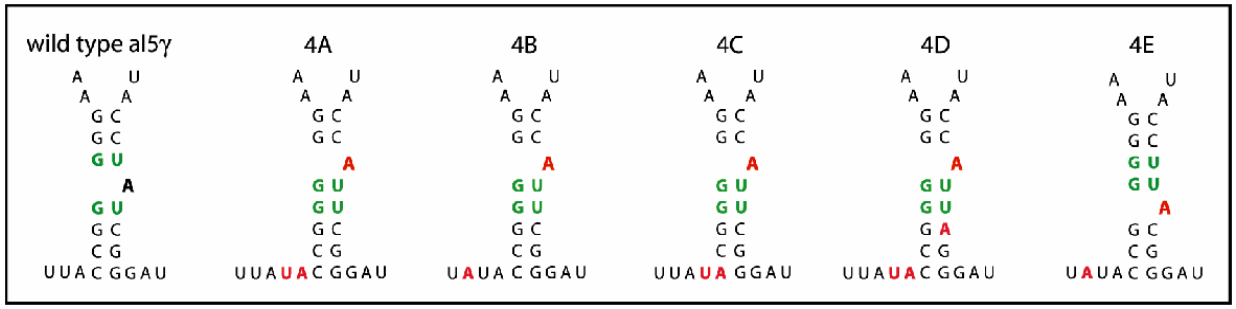


Figure 3-22. Diagrams of mutations from Chu *et al.*. These alleles combine DV-DVI joiner mutations with the BrA⇑ and BrA⇓ mutations. Nucleotide changes are depicted in red. The two GU base pairs that flank the branch A in the wild type intron are shown in green.

Those constructs were transformed into yeast mitochondria, but all of the resulting strains were respiration deficient with no detectable spliced mRNA. Each mutant accumulated a very low level of IVS-E2 RNA that was linear based on its gel mobility relative to standards and RT-PCR analysis (not shown). *Δdbr*1 did not improve splicing or rescue any lariat IVS-E2 RNA, so I concluded that each of those compound mutations completely blocked both branching and the second step reaction *in vivo,* even though some step 1 by hydrolysis remains. *MRS2-L232F* did not restore glycerol growth and also did not alter the splicing phenotype of any of those mutants. It should be noted that three of these constructs (4B, 4C and 4E) combine mutations that individually permitted good splicing and glycerol growth. Revertants were isolated and characterized by Qingshi Zhao, but no new insights into DVI structure or function were obtained; that is, one or another of the contributing mutations simply reverted, restoring some splicing. Mutant 4E combines BrA⇓ with a 4 nucleotide long joiner between DV and DVI, and it branches *in vitro* nearly as well as does mutant BrA⇑$^{C878U}{}_{botGC}$, yet only the latter branches detectably *in vivo*. Even though there is generally good agreement between *in vitro* and *in vivo* results for most mutants analyzed in

this study, these differences demonstrated that requirements for folding and splicing of this intron are somewhat different under the two conditions and the *in vitro* conditions are generally more permissive.

3-2-9 Novel branched intron circle RNA is detected by RT-PCR in both *in vivo* and *in vitro* RNA samples

Vogel *et al.* (2002a) reported a very large set of sequences of RT-PCR products to document the branch site locations of the group II introns of chloroplasts of several plants. In that study, the RT-PCR assay yielded a number of other sequences, most of which could not be explained by misbranching. Some may have reflected potentially interesting, as yet unprecedented, alternative events, though most were probably just PCR artifacts that were detected in cases where branching was minimal. One such event in my data set has been depicted in Figure 3-12, line 8. Inspection of stained or radiolabeled gels of my RT-PCR products (Figure 3-5, Figure 3-11E, Figure 3-16D) revealed that some reactions contained other bands that were substantially larger or smaller than the expected band. I found that changing one of the primers or altering the conditions of the PCR reaction usually increased the amount of the desired product and decreased the amount of alternative products, suggesting that most were PCR artifacts. To avoid documenting events that, because of their length, did not reflect branching within or near DVI, I cloned from gel-purified material that included the wild-type junction plus junctions resultant from mis-branching in the general vicinity of the natural branch site. The sequences compiled in this manner demonstrated that

focusing our attention in this way did not interfere with my ability to detect both ectopic branching and alternative branch sites for several mutant strains.

Murray *et al.* (2001) reported the existence of a novel branching reaction that results in the formation of excised intron circles *in vitro*. I here present evidence that intron circles are formed *in vivo*. A total of 14 sequences were obtained from Δbranch, A880C, A880U, BrA⇑ and BrA⇑$^{C878U}_{botGC}$ (Figure 3-23) showing that those mutant introns can make intron circles *in vivo*. Two different junctions, both of which were the same sequences reported previously by Murray *et al.*, were obtained from the circle intron clones using a similar RT-PCR assay and self-spliced RNAs of a different mutant allele. Most of the sequences shown in Figure 3-23 directly fuse the first nucleotide (G1) and the last nucleotide (U887) of the intron. Some of the sequences shown have an extra A inserted between the 3' and the 5' end of the intron (U887-A-G1). All of these sequences were obtained in the *MRS2-L232F* background except for the one from A880U.

To test whether the wild-type intron can make intron circles at any level, I developed a new PCR primer designed to amplify circle junctions from the cDNA made by the RT but not lariat junctions. From that reaction I obtained and sequenced four circle junction sequences from the *MRS2-L232F* strain carrying the wild-type intron (Figure 3-23, lines 7 and 8). In an attempt to find circular intron clones of the wild type intron without using the circle specific PCR, I performed *E. coli* colony hybridization on transformation plates with greater than 500 insert containing (white) colonies. I designed probes to distinguish between lariat and circular inserts and was unable to find any wild type circle clones either in the *MRS2* or *MRS2-L232F* background.

In vivo evidence for circular introns

Line	source	#	3' intron	5' intron
1	MRS2 / A880U	1	TAC CT **T** TCG GGA **T**	GAG CGG
2	*MRS2-L232F* / A880C	6	TAC CT **C** TCG GGA **T**	GAG CGG
3	*MRS2-L232F* / A880Δ	2	TAC CT **-** TCG GGA **T**	GAG CGG
4	*MRS2-L232F* / BrA⇑	1	TAC C **A** T TCG GGA **TA**	GAG CGG
5	*MRS2-L232F* / BrA⇑ C878U botGC	3	TAC T **A** T TCG GGA **T**	GAG CGG
6	*MRS2-L232F* / BrA⇑ C878U botGC	1	TAC T **A** T TCG GGA **TA**	GAG CGG
7	*MRS2-L232F* /wild-type*	2	TAC CT **A** TCG GGA **T**	GAG CGG
8	*MRS2-L232F* /wild-type*	2	TAC CT **A** TCG GGA **TA**	GAG CGG

* used selective PCR primers to preferentially amplify circle derived cDNAs

Figure 3-23. *In vivo* RT-PCR branch point mapping data identify circular intron products. The natural branch point in DVI is highlighted in green. The terminal nucleotide / presumed "branch point" is highlighted in bold type. The source of the additional A in some sequences is unclear.

I also used the circle and lariat specific probes to screen clones from mutants Δbranch, A880C and A880U. Both Δbranch and A880C colonies hybridized to the circle probe but not the lariat probe, although many of the insert containing colonies did not hybridize to either. Most of the A880U clones hybridized to the lariat probe, however a few circle clones were also isolated**.** This indicated that the faint RT-PCR band observed for Δbranch and A880C in the *MRS2-L232F* background contained mostly circle PCR product. Mutants BrA⇓ and BrA⇓$_{C882U}$ also had a similar faint RT-PCR band (Figure 3-16), possibly resulting from circle intron RNAs; however, I was unable to clone those bands.

In vitro spliced RNAs were used to map branch points of mutants that did not branch at a detectable level *in vivo.* I did not obtain any lariat sequences from self-spliced A880C,

A880U or BrA⇓ RNAs, but obtained two circles from A880C, six from A880U and three from BrA⇓ (Figure 3-24). Because circle sequences contained the branch point, I could confirm the mutant identity of each sequenced clone. Screening of clones from self-spliced wild-type aI5γ yielded one intron circle (Figure 3-24, line 1) out of 10 clones analyzed, but no circles were obtained with the other strains analyzed *in vivo* here. I have now observed what appears to be the excised circular intron RNA *in vivo* as well as *in vitro*. Furthermore, a rotation student in our lab, Ting Ting Li was able to show that intron aI1 also forms mis-spliced intron circles *in vitro*. These data support the existence of a novel intron splicing side-reaction of unknown mechanism or function *in vivo*.

In vitro evidence for circular introns

Line	source	#	3' intron	5' intron
1	wild type self-spliced RNA	1	TAC CT **A** TCG GGA **T**	GAG CGG
2	A880U self-spliced RNA	6	TAC CT **T** TCG GGA **T**	GAG CGG
3	A880C self-spliced RNA	4	TAC CT **C** TCG GGA **T**	GAG CGG
4	A880Δ self-spliced RNA	2	TAC CT - TCG GGA **T**	GAG CGG
5	BrA⇓ self-spliced RNA	3	TAC CT T **A** CG GGA **T**	GAG CGG

Figure 3-24. *In vitro* RT-PCR branch point mapping sequences of circular introns. The natural branch point in DVI is highlighted in green. The terminal nucleotide / presumed "branch point" is highlighted in bold type.

3-3 DISCUSSION

The current work is the *in vivo* extension of a set of *in vitro* experiments designed to test if branching of intron aI5γ can be redirected. I performed the *in vivo* analysis of the mutants that were tested *in vitro* (Chu *et al.*, 2001) and transformed into yeast by Dr. Qingshi Zhao. A few additional mutants, not tested *in vitro,* were also analyzed in this work. *In vivo,* branching at greater than a fraction of a percent of the wild type level requires an unpaired adenosine at the correct location. Another study in plants (Vogel & Borner, 2002b) mapped the branch points of many natural group II introns *in vivo*, and they did not observe any instances of ectopic branching. Among the 78 cDNA clones I sequenced (from wild-type and five mutants that have the wild-type branch point: D5UA, topGC, botGC, J4, b-loop), I did not observe a single instance of mis-branching. In addition, 19 cDNA clones were sequenced from A880G (which still has a purine at nucleotide 880), and all of them reflect branching exclusively at nucleotide 880. The ectopic branch mutant that branched most efficiently *in vitro* (BrA⇑^{C878U}) did so at 1/5th of the wild-type level but the same mutant *in vivo* branches at ~1% of wild type (but splices by hydrolysis at about 1/3rd wild type). Apparently, the *in vivo* conditions make ectopic branching even less likely than was observed *in vitro*.

Here I show that ectopic branching in DVI occurs detectably in several mutant strains at locations both 5' and 3' to nucleotide 880. My *in vivo* analysis of the mutants that moved the branch point one nucleotide 5' or 3' of the natural site confirms and extends the *in vitro* finding that branching can be redirected most efficiently if the ectopic branch A is flanked by

a 5' GU wobble pair. In addition I detect branching at a 3' ectopic branch A when the DVI stem is stabilized by an additional GC pair two base pairs 5' of the ectopic branch, a mutant that has not been analyzed *in vitro*.

I also analyzed the effects of mutating the flanking base pairs on branching at the wild-type branch A. I found that the GU pair above the branch is required for efficient branching at the wild-type location and the GU pair below the branch influences the second splicing step, but not branching.

3-3-1 *In vivo* results for BrA⇑ and BrA⇓ compared to the *in vitro* data

The *in vitro* experiments by Chu *et al.*, (2001) demonstrated that having an unpaired adenosine at the wrong location in DVI (specifically at nucleotide 879 and 881) does not support efficient branching. Specifically the A880U, U881A mutant (3A in their work, BrA⇓ here) reduces branching to 1/2200th of wild-type while splicing by hydrolysis proceeds at 1/14th of the total wild-type splicing rate *in vitro*. Chu et al, not surprisingly, could not map these extremely rare branches. I also did not detect branching in the BrA⇓ mutant *in vivo*, even though it splices by hydrolysis at ~1/5th of the wild-type level. The A880U, U879A mutant (3B their work, BrA⇑ here) spliced and branched better than BrA⇓ both *in vitro* and *in vivo*. *In vitro* BrA⇑ branches ten times the rate of BrA⇓, at 1/203rd of the wild-type rate and splices by hydrolysis at 1/7th of the wild type rate. *In vivo* mutant BrA⇑ splices about 40% of the *COXI* transcript, mainly by hydrolysis, somewhat less than half the wild-type level (94%). This level of splicing is similar to the branch nucleotide mutants (Figure 3-7,

they splice in the 20%-60% range). I detected and mapped branching for mutant BrA⇑ by RT-PCR, despite the fact that no branched intron was detected by northern blot. Mutant BrA⇑ must therefore branch at <1% of spliced message, a level below the observable threshold for northern blots. I found that branching by mutant BrA⇑ occurs at two sites, the shifted branch A (20 events) and the U that now occupies the location of the native branch A (11 events). Branching can be redirected at a low level *in vivo* by simply providing an alternate branch point near the native branch point. This indicates that there is some flexibility in branch point selection. However, as evidenced by the low level of branching in mutant BrA⇑ that still occurs at the native site (1/3rd of the time), it is clear that aI5γ is "hard-wired" for branching at position 880, despite the presence of a uracil.

I was able to map the branch point for mutant BrA⇑ even though the very low level of branching by this mutant could not be mapped *in vitro* by the DNAzyme method. It is interesting that Chu *et al.* were able to map mutants 4C and 4D (see Figure 3-22) using the DNAzyme method, both of which branch at less than half the rate of BrA⇑. It is possible that the experiment failed because BrA⇑ also branches at multiple locations *in vitro* – so that the low level of purified material might not contain enough of any one product to yield a signal.

3-3-1a Mutant BrA⇑^{C878U}

One key finding of Chu *et al.* was that an additional C878U mutation activated ectopic branching at the ectopic A879 *in vitro*. Phylogenetic data show that the base pairs flanking the bulged A in DVI of group II introns are rather well-conserved (Figure 3-18, see

Discussion section 3-3-5). That observation led Chu *et al.* to test whether improving the environment of the unpaired A in mutant BrA⇑ would rescue some branching. Remarkably, changing the GC pair above the shifted A to a GU wobble pair (3C their work, BrA⇑^{C878U} here) restored branching to 1/5th the wild type rate. DNAzyme mapping shows that most or all of the branches occur at the shifted A. I analyzed the same mutation *in vivo* and found that mutant BrA⇑^{C878U} splices ~30%, mainly via hydrolysis, not significantly less from BrA⇑. RT-PCR branch point mapping yielded more product for BrA⇑^{C878U} than for BrA⇑, and 25 out of 27 events mapped to the ectopic A879. Furthermore, intron lariat was visible on northern blots of BrA⇑^{C878U} in the *MRS2-L232F* suppressor background (*i.e.*, in the presence of increased intra-mitochondrial Mg^{2+}). The level of branching in the suppressor background can be roughly estimated at ~1-3% of the wild type level. Ectopic branching occurs at a much higher level in mutant BrA⇑^{C878U} than in any of the other mutants I tested *in vivo*, as was also the case in the *in vitro* study. My *in vivo* data confirm the *in vitro* findings for mutant BrA⇑^{C878U}, further supporting the notion that the base pair above the branch site is critical for efficient branching. These data show how important the flanking basepair can be for branching – but do not inform us about the role of those flanking basepairs in a wild-type intron.

3-3-1b Other BrA⇑ mutants do not improve ectopic branching at A879.

Mutant BrA⇑^{C878U} has three GU pairs in the DVI stem, while wild-type only has two. Thinking that perhaps I could further increase ectopic branching at A879 by restoring the

wild-type two GU basepair configuration, I made the BrA⇑ $^{C878U}_{botGC}$ mutant strain. I also made and analyzed mutant BrA⇑ $_{botGC}$ as a control for the effect of U881C (botGC) mutation in this context. Both strains splice by hydrolysis at 20% of wild type, which is significantly less than BrA⇑ (37%). Using semi-quantitative RT-PCR, branching was detected in both strains. For both BrA⇑ $^{C878U}_{botGC}$ and BrA⇑ $_{botGC}$, branching was only marginally increased over the amount observed for BrA⇑, and nowhere near the level of branching observed in BrA⇑ ^{C878U}. Despite a low level of branching, RT-PCR mapping revealed that BrA⇑ $^{C878U}_{botGC}$ branches at the ectopic A879 for 25 out of 30 events; the remaining 5 events occur at U880. In contrast, for mutant BrA⇑$_{botGC}$ I mapped 4 events to A879 and 2 events to U880. Taken together, it appears that the improvement in branch fidelity observed in strain BrA⇑^{C878U} can also be seen in strain BrA⇑ $^{C878U}_{botGC}$, even though the botGC change does not alter the ratio of ectopic branching relative to BrA⇑. In sum, the botGC change did nothing to improve splicing or branching. To the contrary, it reduced splicing by hydrolysis and negated the increased branching caused by the C878U change. These data will be further discussed in section 3-3-4 insofar as they relate to the botGC change in the wild-type context. It suffices to say that the botGC change does not appear to be directly involved in branching but is critical for the second splicing step.

3-3-2 Complex BrA⇓ mutants activate ectopic branching at A881.

Moving the branch nucleotide one base pair downstream of its natural location is more detrimental to overall branching and splicing than is shifting the branch nucleotide one

base pair upstream. Mutant BrA⇓ splices about half as well as mutant BrA⇑ *in vitro* ($1/14^{th}$ vs. $1/7^{th}$ the wild-type rate) and *in vivo* (21% vs. 37% spliced message). Despite many attempts, I was never able to clone any branch events from mutant BrA⇓. That mutant BrA⇓ does not branch at a detectible level is also consistent with my finding that A880U never mis-branched at that location. Unlike mutant BrA⇑, the *MRS2-L232F* suppressor does not significantly increase splicing and no branches were detected in the suppressor background for mutant BrA⇓. This means that the problem with mutant BrA⇓ cannot be easily rectified by providing additional Mg^{2+}. Introducing an unpaired A four nucleotides from the base of the DVI stem partially inhibits the hydrolytic splicing pathway and eliminates branching. This disruption may reflect a failure of the intron to fold or dock DVI properly, or some other structural problem.

3-3-2a Mutant BrA⇓^{topGC} (U879C)

I analyzed a number of derivatives of mutant BrA⇓ attempting to improve the level of splicing and/or branching. Changing the GU base pair two nucleotides above the ectopic branch point (A881) to a GC pair (U879C, mutant BrA⇓^{topGC}) significantly reduced splicing by hydrolysis (from 21% for BrA⇓ to 8% for BrA⇓^{topGC}), however it restored branching to detectable levels. All six branching events mapped by RT-PCR were at the ectopic A881. The nucleotide directly above the branch is a GU pair in BrA⇓ and BrA⇓^{topGC}. I demonstrated that a GU pair is critical for branching at A879 and A880. However, merely having a GU pair above the branch is not sufficient for branching at A881. The change that permits branching at A881, namely topGC, is actually inhibitory for branching at the wild-

type A880. Mutant topGC, in conjunction with the native branch (A880), reduces branching to 1/5th of the wild-type level. This discrepancy may arise because a general *tightening* of the DVI helix is needed to permit a low level of branching at the shifted A (and not at U880) in the context of BrA⇓.

As was observed for mutant BrA⇓, the *MRS2-L232F* suppressor does not improve splicing or branching of the BrA⇓^{topGC} mutant. This finding is in direct contrast to the results obtained with the C882U containing BrA⇓ derivative mutants.

3-3-2b Mutants BrA⇓$_{C882U}$ and BrA⇓$^{topGC}{}_{C882U}$

Just restoring the environment around the shifted branch point (mutant BrA⇓$_{C882U}$) does not restore branching, and in fact reduces total splicing to a level barely detectable by northern blot (from 21% for BrA⇓ to 2% for BrA⇓$_{C882U}$). Specifically, the C882U change blocks the second splicing step and not first step hydrolysis – the first splicing step proceeds in BrA⇓$_{C882U}$ just as well as in BrA⇓ (~20-25% 5' cleavage, Figure 3-25).

I also made mutant BrA⇓ $^{topGC}{}_{C882U}$ which, like wild-type, contains two GU pairs in the DVI stem. Just like mutant BrA⇓ $_{C882U,}$ mutant BrA⇓ $^{topGC}{}_{C882U}$ is strongly inhibited for splicing (reduced to 2%) and also has a clear second step defect. These findings emphasize that DVI plays a key role in the second splicing step. However, the branch environment is not totally unimportant in this context; using RT-PCR signal as a semi-quantitative readout of the branching level, it appears that mutant BrA⇓ $^{topGC}{}_{C882U}$ branches to a greater extent than the other BrA⇓ mutants. The second step defect of the C882U-containing mutants is probably

independent of branching; given the extremely low level of spliced message, most of the branched material detected by RT-PCR is probably in the form of IVS-E2. Although the U879C change is sufficient to activate branching at A881, additionally changing C882U greatly improves the branching rate.

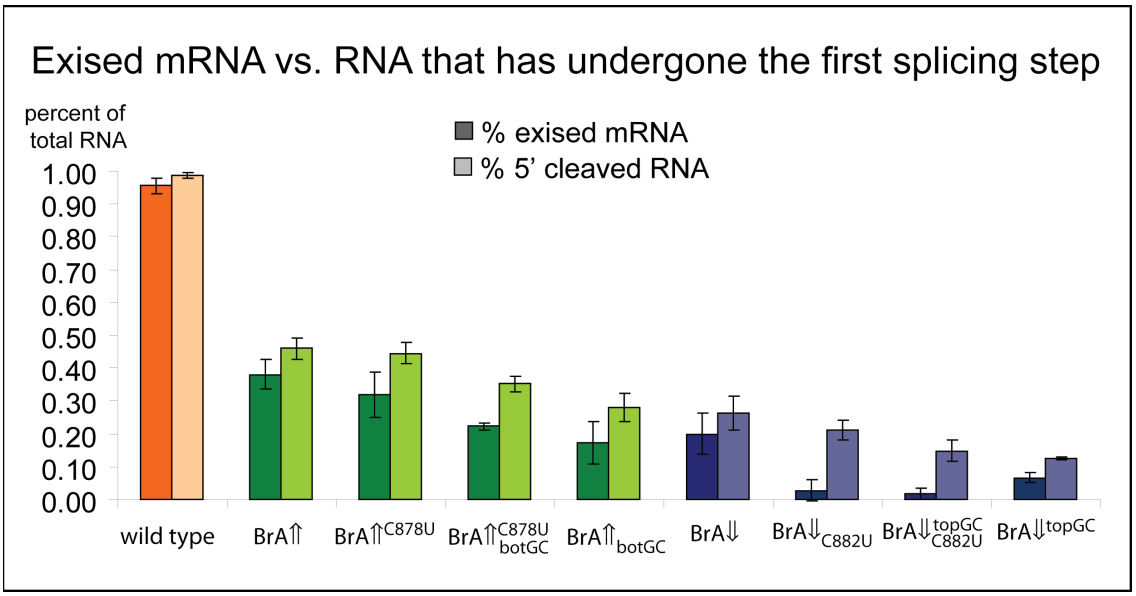


Figure 3-25. Comparison between steady state level of 5' cleaved RNA (1st splicing step only) and completely exised mRNA for the shifted branch point mutants. Wild type is in orange. Br⇑ mutants are in green. Br⇓ mutants are in blue.

3-3-2c The MRS2-L232F suppressor rescues splicing of C882U containing mutants.

The *MRS2-L232F* suppressor background rescues some splicing by the C882U-containing mutants, by improving both splicing steps. The extra Mg^{2+} provided in the *MRS2-L232F* background increases first step hydrolysis, first step transesterification, and second step transesterification. Because the *MRS2-L232F* suppressor is able to dramatically rescue splicing, the C882U change probably causes a mild unfolding of the intron. Magnesium ions are important for group II intron folding and bind at critical locations, particularly in the

catalytic core of the intron (Sigel *et al.*, 2000; Sigel *et al.*, 2004). It is therefore not surprising that *MRS2-L232F* is able to rescue the splicing by generally stabilizing the intron fold. The C882U mutation may decreases the stability of the DVI helix making it more difficult to carry out the second splicing step. I conclude that disrupting the helical structure of the base of DVI by both replacing a GU pair with an unpaired A and a GC pair with a GU wobble pair is sufficient to prevent the second splicing step.

Taken together, my findings with the BrA⇓ derivative alleles suggest that creating an unpaired A at nucleotide A881 destabilizes DVI sufficiently to prevent branching. The general tightening of the DVI stem helix caused by introducing a Watson-Crick pair (U789C) stabilizes the correct structure enough for some branching at A881. Because additional Mg^{2+} (*MRS2-L232F*) reduces the branching in mutant BrA⇓^{topGC}, and the introduction of a GU wobble pair (C882U) adjacent to the branch nucleotide increases branching, I conclude that a more stable DVI, in combination with GU wobble pairs around the branch nucleotide, permits the best ectopic branching at A881.

3-3-3 Mutant A880U also supports ectopic branching *in vivo*.

In order to calibrate the new RT-PCR method of branch point mapping, I performed experiments with all four branch mutants. These studies of the *in vivo* phenotypes of branch nucleotide mutants confirmed that adenosine is the preferred branch nucleotide, with a low level of branching possible at an unpaired guanosine or uracil. The presence of an unpaired cytosine at nt880 of aI5γ, five nucleotides from the base of the DVI stem, has yet to result in

any detectable branching *in vitro* or *in vivo*. For the purposes of this study, the A880U (see figure 3-7) mutant was of particular interest because it activated detectable ectopic branching *in vivo*.

The ectopic branching observed in mutant A880U is understandable because there is no distinction between the branch uracil and the two flanking nucleotides, both of which are uracils. Because the other side of the stem is composed of a string of guanosines, any one of the three uracils could be unpaired. This mutant effectively tests which branch is chosen when the only branch point indicator available is location. Under these circumstances, the wild-type location is still the predominant branch partner, although the uracil 5' to the branch (U879) is also used for branching detectably (4-12 out of 45 mapped branches, depending on the interpretation of the cDNA sequences). The lower uracil is never chosen in this mutant background, perhaps because it is rarely un-paired.

The two positions that support branching in mutant A880U also support branching in the BrA⇑ mutant. The only difference between the two mutants is that BrA⇑ has an unpaired adenosine at nucleotide 879 whereas A880U has a uracil. Mutant BrA⇑ branches at the shifted branch A and at the natural branch location, which is now a U, at a ratio of 2 to 1. In contrast, the ratio of branching at these same two sites in mutant A880U is somewhere between 1 to 3 and 1 to 10 (some A880U cDNA sequences are ambiguous for the branch location). This testifies to the importance of the branch identity because the presence of an adenosine at a location one nucleotide upstream of the natural location makes it the primary instead of the secondary branch site.

For mutant A880U, I never observed branching at U881. This is consistent with my observation that BrA⇑ and its derivative mutations branch at a much greater frequency than do the analogous BrA⇓ mutants. The A880U data lend further support to the notion that the nucleotide below the branch is not available for branching, perhaps because it remains base paired in the proper RNA fold. The A880U mutation does not alter the environment of nucleotide 880 and probably does not alter the basic structure of DVI to the same extent as those mutations that shift the location of the bulged A. Thus, the A880U mutation provides a relatively simple test of the hypothesis that nucleotide 880 is the preferred branch point location. The A880U mutant data indicate that the branch location choice is an intrinsic part of the RNA structure and is not dependant on a branch A. However, in the absence of a favorable branch nucleotide, branching at two locations is possible.

3-3-4 Mutations that alter the environment of a branch A

In vitro and *in vivo* experiments with branches located at nucleotide 879 (BrA⇑) or nucleotide 881 (BrA⇓) demonstrate that the nucleotides flanking the branch influence the efficiency of ectopic branching. These branch mutants alone, however, do not define the role that the branch environment plays for branching at the wild-type A880 position. Phylogenetic studies indicate that the base pairs above and below the putative branch A are partially conserved. Many of the 16 possible pairs exist above the branch. In this data set, GU is the most common pair (41%), followed by AG and AU (15% and 13% respectively) (Figure 3-18). From the phylogeny, one would probably conclude that the base pair above the branch is

not important for branching or splicing. However, a GC base pair is present in ~8% of the introns, and my data for U879C show that a GC there strongly inhibits branching. The nucleotide below the branch is much more highly conserved and 91% are either GU or GC with GC predominating (63%). Either there are intron-specific accommodations for deviations from the most common base pairs and all of the found sequences allow for good splicing and branching, or some natural introns do not splice or branch optimally. It is important to note that GU wobble pairs, besides having one less hydrogen bond, create a much more electronegative environment in the major groove then does a GC base pair. Given the importance of metal ion binding to group II intron catalysis, the change in charge created by mutating GU pairs may explain some of the deleterious phenotypes .

One relevant mutant allele – 2xGC – has already been partially described (Podar *et al.*, 1998). In 2xGC, both the basepair above and below A880 of aI5γ are mutated from GU to GC pairs. *In vitro*, mutant 2xGC splices and branches well under permissive high salt conditions (Podar *et al.*, 1998). I confirm the *in vivo* results and find that 2xGC contains ~30% spliced message that largely results from splicing by first step hydrolysis. Mutant 2xGC accumulates IVS-E2 and contains only $1/10^{th}$ of the wild-type level of excised intron RNA. Intron RNA is not significantly rescued by deleting *DBR1*. From these data I conclude that mutant 2XGC has an overall splicing defect that includes both branching and second step deficiencies. The phenotype of 2xGC cannot be understood without analyzing the component mutations separately. These mutants (topGC and botGC) had previously been made by Mircea Podar in our lab, so I set out to complete his *in vivo* analysis.

3-3-4a The 5' flanking GU wobble pair is important for branch-site selection.

Changing the base pair above the natural branch A from a GU wobble pair to a GC pair (mutant G879C or topGC) only reduces splicing from 94% to 74% and does not accumulate much IVS-E2. However, topGC only has 20% of the wild type amount of lariat intron RNA. The level of branched intron RNA is not significantly increased in the *Δdbr1* background. The primary defect of the topGC mutant is reduced branching. Consistent with the observation that the GU pair above the branch is primarily involved in branching, I find that fidelity and efficiency of branching at the new, ectopic branch site is greatly increased by changing the base pair above the ectopic branch to a GU wobble pair (mutant BrA$\Uparrow^{C878U}$ see Discussion section 3-3-1). Interestingly, making the G879C (topGC) change in the context of mutant BrA$\Downarrow$ (mutant BrA$\Downarrow^{topGC}$, discussion section 3-3-2) improves branching. In this latter case, the U879C change is no longer adjacent to the branch point at A881. This indicates that it is not the absolute location of that base pair in DVI that is important, but rather it must be adjacent to the branch A. It has been proposed that the branch nucleotide must be able to flip out of the DVI helix and must also be positioned in the active site (near the base of DV) in order to act as the nucleophile in the first splicing step (Liu *et al.*, 1997). The requirement for the GU wobble pair above the branch may imply that that nucleotide must also undergo some conformational change to assume the correct step 1 structure. For example, the GU may facilitate the flipping out of the bulged A or facilitate the binding of metal cofactors that are required for efficient branching.

One way to test this would be to change the base pair above the branch from a GU to an AU. Based on phylogenetic evidence, this base pair should be well tolerated. This mutant

would indicate if it is the presence of a uracil *per se* that is required, or the wobble pair. The observation that most base pairs are tolerated above the branch point – despite the finding that a GU pair is critical for efficient branching in aI5γ – may be explained by the ability of group II introns to splice relatively efficiently *in vivo* by first-step hydrolysis. The topGC mutation only has a modest effect on net splicing, and in the absence of any selective pressure for efficient branching, naturally occurring mutations of the base pair above the branch are likely to be well tolerated. Indeed this would be an efficient method for the host organism to reduce the ability of the group II intron to retrotranspose, thus reducing the spread of group II introns in general.

3-3-4b The second step defect of botGC cannot be explained by a lack of branching.

Mutant botGC has the same amount of spliced mRNA as topGC (~70%), however it accumulates IVS-E2 and contains the same amount of unspliced message as wild type. This indicates that the first splicing step is not at all inhibited and that the drop in splicing from ~95% in wild-type to ~70% in botGC is entirely due to a second step defect. This is in direct contrast to topGC which accumulates very little IVS-E2; that is, most molecules that undergo the 5' cleavage reaction also perform the second splicing step. In mutant botGC, the second step defect cannot be attributed to a branching defect because it accumulates nearly as much intron lariat RNA as wild-type (figure 3-20).

Because published intron mutants that fail to branch *in vivo* always have a second step defect, it was previously inferred that the branching *per se* promotes efficient second step reactions. BotGC is a mutant that carries out the first step by branching almost as well as

wild type, but splices only ~70% of the transcripts. The drop in splicing compared to the wild type is due almost entirely to a second step splicing defect. This result demonstrates the importance of the helix below the branch for the second splicing reaction in the wild type context. The same mutation was also made in the context of BrA⇑. The botGC change is no longer adjacent to the branch A in these mutants, however it still causes a second step defect.

Mutants BrA⇑ and BrA⇑ ^{C878U} have a mild second step defect, and for these two mutants the first and second splicing steps are significantly increased in the *MRS2-L232F* suppressor background. Mutants BrA⇑ $_{botGC}$ and BrA⇑^{C878U} $_{botGC}$ do not splice as well as BrA⇑ and BrA⇑ ^{C878U} and this difference is due to an increase in level of the IVS-E2 intermediate, hence an increased second step defect. The *MRS2-L232F* suppressor does not rescue splicing for mutants BrA⇑ $_{botGC}$ and BrA⇑^{C878U} $_{botGC}$ as it does so nicely for mutants BrA⇑ and BrA⇑^{C878U}. These observations show that the botGC mutation in the context of BrA⇑ also causes a second step defect, although it requires the *MRS2* suppressor to become evident. It is interesting to note that the base pair below the branch is much more evolutionarily constrained than the base pair above the branch, and 99.2% of the available group II intron sequences have a Watson-Crick or GU wobble pair there. Because first step hydrolysis is a viable alternative to branching for aI5γ, mutations that disrupt branching are more likely to be tolerated than mutations that disrupt the second splicing reaction. This would explain why the base pair above the branch tolerates many alternative sequences whereas the base pair below the branch is restricted to a few. Taken together, these data indicate that the botGC mutation primarily causes a second step defect and that this defect is independent of branching.

3-3-5 Naturally occurring introns may splice predominately by hydrolysis.

For aI5γ, alternative flanking nucleotides do not support the efficient wild-type level of splicing *in vivo*. The efficient splicing of other introns with divergent branch-nucleotide environments may depend on accommodations by other changes in DVI or elsewhere in the intron. Alternatively, it is possible that many of the introns in the database may splice inefficiently or with little branching because of their branch nucleotide environment. The literature contains little information that can help sort out those alternatives because the efficiency of splicing *in vivo* is not known for most group II introns and *in vitro* data are also lacking (sometimes unattainable) for most of them. In plants, a natural intron lacking a branch nucleotide was shown to splice *in vivo* by hydrolysis, though the efficiency of splicing was not estimated (Vogel & Borner, 2002a). Therefore, splicing by hydrolysis is probably sufficiently robust in organisms besides yeast for the existence of inefficient, naturally occurring group II introns due to their DVI sequence (or other features). Based on quantitative measures of *in vivo* splicing by aI5γ mutants (examples in this work include: A880U, A880C, A880Δ, BrA⇓, BrA⇑ etc.) that are essentially blocked for branching and that splice by hydrolysis to less than half the wild-type extent, it would be surprising (and very interesting) if the chloroplast intron splices efficiently (~90%). Measuring that parameter may be quite difficult because splicing of chloroplast introns is highly regulated by physiological factors and, when splicing is inhibited, high levels of precursor RNA do not accumulate (as opposed to the more-readily analyzed case of defective mutant splicing in yeast mitochondria). If naturally occurring introns in organisms besides yeast splice by

hydrolysis at a reduced rate, it would be interesting to learn how widespread this might be, and if it is used as a mechanism for altering gene expression.

3-3-6 Our data do not support the DVI structural model proposed by Doudna's group

The published crystal structure model of DV-DVI RNA from aI5γ has both the branch A and the 3'U bulged out of the helix (Zhang & Doudna, 2002). If that conformation is accurate, one would expect that mutant A880U would branch at U880 and U881, instead of at U880 and U879 as I found. Also, if the crystal structure is correct, our mutant botGC, in which it is very unlikely that the base pair below the branch is unpaired, should branch poorly. Again I find the opposite, that botGC essentially branches as well as wild-type, and that mutant topGC has a strong branching defect. If anything, my data support a model that requires the base pair above the branch to be unpaired and possibly flipped out of the helix, or at least in a weaker non-Watson-crick base pair configuration. Indeed a recent solution structure of DVI directly contradicts the crystal structure and supports a model of DVI in which the branch A is extrahelical and the two flanking base-pairs are stacked in the helix (Schlatterer *et al.*, 2006).

3-3-7 Lariat RNA is protected from complete debranching by Dbr1p.

Spliceosomal introns are released as intron lariats that are debranched by the yeast debranching enzyme, Dbr1p. A yeast strain deleted for the *DBR1* gene has no discernable

growth defect, but nuclear intron lariats accumulate to a high level (Chapman & Boeke, 1991). Previously, there was no evidence of Dbr1p entering the mitochondria and debranching mitochondrial group II introns. To the contrary, the fact that intron lariats accumulate in the mitochondria suggested that mitochondria either lack a debranching activity or somehow protect the lariats from its action. Because some group II introns are mobile and their mobility depends on reverse splicing, protecting the branched RNA from degradation must be a prerequisite for intron mobility by this mechanism. The findings presented here, that wild-type aI5γ lariat RNA is at least partially degraded by Dbr1p and that some aI5γ mutants are more susceptible to debranching than others, indicate that Dbr1p has access to group II introns but is unable to completely debranch them as it can for nuclear introns. It would not appear that this is because Dbr1p is limiting in the mitochondria, because the level of branched intron is consistently doubled in the *Δdbr1* background of most mutants regardless of the overall level of branched intron. Mutant D5UA splices well in vivo and in vitro but has very little lariat RNA in vivo. In the *Δdbr1* background D5UA lariat RNA is rescued to a level that is consistent with most splicing in D5UA having occurred with branching. It is likely that some aspect of the intron structure or an intron binding protein protects it from debranching, and that protection is greatly reduced in mutant D5UA.

Because branching is an important aspect of group II intron transposition, branch nucleotide mutants of aI2 permit some splicing but block homing completely. It would be interesting to look at the transposition rate of either the group II intron aI2 or aI1 in the *Δdbr1* background to examine the possibility that deleting *DBR1* would increase homing.

For example, if the null mutant increases the amount of lariat RNA, then more RNP particles should be available for retrotransposition.

3-3-8 The *MRS2* suppressor stabilizes certain mitochondrial RNAs.

MRS2 missense suppressor mutations are thought to act by increasing the intra-mitochondrial Mg^{2+} level (Gregan *et al.*, 2001; Weghuber *et al.*, 2006). The suggestion that the mutated protein may interact directly with group II intron RNAs (Schmidt *et al.*, 1998) has not been tested directly, and is unlikely in light of the convincing correlation between *MRS2* mutations that alter Mg^{2+} levels and intron splicing (Weghuber *et al.*, 2006). My data is consistent with the indirect method of action. It is my observation that, in addition to increasing splicing in some mutants, *MRS2-L232* generally stabilizes RNAs. For example, the wild-type intron in a suppressor context does not show an increase in spliced message, but has twice as much intron lariat, and increased levels of linear IVS-E2 and IVS. A new band also appears in the suppressor background, and probing revealed this band to contain upstream exons. The increased Mg^{2+} provided in the *MRS2-L232F* background may make some RNAs more difficult to degrade. In conjunction with my *Δdbr1* results, it appears that although lariat RNA is easily detectable by northern blot, it is still being rapidly degraded in wild-type cells.

3-3-9 The presence of branched intron circles points to a novel branching pathway that is not likely to result in exon joining.

Group II intron circles were discovered by accident in the lab of Kevin Jarrell as an unexpected "branched" (i.e., a slow moving splicing product that for group II introns usually means branched) intron band in first step self-splicing reactions of the Δbranch mutant (Murray *et al.*, 2001). They carefully characterized this reaction and found that the novel band consists of RNA circles in which the 2'OH of the terminal U of the intron is joined to the 5' phosphate of the first intron nucleotide (a guanosine) via a 2'-5' linkage. These molecules have a 3'OH that renders them resistant to debranching, further indicating that they are circles and not a lariat with the downstream exon as the 3'tail. They used the RT-PCR branch point mapping assay and were able to show that purified linear intron RNA does not yield a circular RT-PCR product. A concern is that, because linear excised intron RNA folds such that the 5' and 3' ends of the intron are in relatively close proximity, this may allow the RT to "jump" from the 3' end to the 5'end and make an artifactual cDNA. Their data indicated that the reaction forming these circular RNAs requires a free 3' end. They proposed a mechanism in which the second splicing step occurs first, liberating E1-IVS, followed by the first step reaction using the terminal nucleotide as the branch point.

They attempted to demonstrate that circles are also formed *in vivo*, but were unable to get the RT-PCR reaction to work for intron aI5γ. They also used the RT-PCR reaction to look for intron circles of introns aI1 and aI2 *in vivo.* However, introns aI1 and aI2 are known to reverse-splice into intron-containing alleles resulting in a transient tandem intron allele

(Dickson et al., 2001). A tandem intron duplication would easily serve as template for the cDNA that defines circular intron RNAs. Our lab possesses mutants that do not reverse splice, however Dr. Jarrell did not request those mutants; consequently, their finding that introns aI1 and aI2 form circles *in vivo* is inconclusive.

My finding that branched intron circles of certain aI5γ mutants exist *in vivo* is surprising because the proposed mechanism for circle intron formation requires free E1-IVS. There are several reasons why E1-IVS is not expected to exist *in vivo.* E1-IVS is generated *in vitro* by adding a great excess of E1 to a splicing reaction containing E1-IVS-E2. The excess E1 reacts in a *trans*-splicing reaction to liberate E1-IVS and E1-E2. It is hard to imagine enough free exon RNA *in vivo* to support that reaction. Even if linear E1 or E1-IVS were generated *in vivo,* they are not expected to accumulate because, just like linear IVS, they lack the RNA stability determinants found on the 3' tail of mitochondrial mRNAs. Furthermore, *in vitro* wild-type E1-IVS is very active for branching with release of free E1.

Further work is needed to determine how intron circles are formed *in vivo*; however, based on the data available, I can propose a possible mechanism. Most of the intron circles I cloned were observed in the *MRS2-L232F* suppressor background. This could be tied to my finding that *MRS2-L232F* stabilizes linear intron RNA. The *MRS2-L232F* background may stabilize E1-IVS, creating an opportunity for the circle-forming reaction to occur. If linear E1 is also rescued, the Jarrell model in which exogenous E1 attacks precursor to yield E1-IVS might hold true. It is still unlikely that sufficient E1 accumulates to drive this reaction. However an alternative exists. The second step reaction could in theory precede the first step reaction by intron mediated second step hydrolysis. Group I introns generate RNA circles by

this mechanism (Cech, 1990). This second step hydrolysis may be enhanced in the *MRS2-L232F* background, as is first step hydrolysis. Although this exact reaction has not been observed *in vitro*, a similar reaction has been – spliced exon reopening (SER) (Jarrell *et al.*, 1988b). SER is thought to be the intron mediated *reversal* of the second splicing step by hydrolysis (Podar et al., 1995).

In contrast with the cDNAs obtained with lariats, the cDNA sequences from intron circles show that a circle branch does not pose the same challenge to the RT because U887 is copied faithfully in all 23 of the events compiled in Fig. 8; this contrasts with our finding that a branch U in a lariat RNA results in incorporation of mostly A with a few G's. The four events in which there is an A inserted in the sense strand cDNA between the nucleotides templated by U887 and G1 are a puzzle (see Figure 3-23; lines 4, 6 and 8). It is interesting to note that the first nucleotide of the downstream exon is an A, and I speculate that those events may reflect the use of that A in the branching reaction – in which case it was used correctly to template a U in the initial cDNA. But that would require release of the rest of E2 so that the A is the last nucleotide of the E1-IVS molecule. Further research on the reactions leading to branched circles will be needed to understand the origin of those sequences.

CHAPTER FOUR
Functional Screen of the Mitochondrial Proteome to Identify Novel Proteins Involved in Group I or Group II Intron Splicing.

4-1 INTRODUCTION

4-1-1 Self-splicing introns require protein cofactors *in vivo*

Nuclear introns are spliced out of pre-mRNA by a large molecular machine – the spliceosome – that contains numerous proteins and small RNAs (Jurica & Moore, 2003). In contrast to the highly conserved secondary and tertiary structures of group I and group II introns, nuclear introns have very few innate structural features and require a large number of proteins and RNAs to recognize key intron sequences, fold the intron properly, and perform the splicing reaction. This complexity also gives control of the splicing reaction to the host organism, allowing for sophisticated control of gene expression - even the expression and regulation of alternative isoforms of the same message (Matlin *et al.*, 2005). Group I and group II introns can be thought of as substantially self-sufficient or parasitic elements, that are encoded by "selfish DNA" (Doolittle & Sapienza, 1980; Edgell *et al.*, 1996) and, relative to nuclear introns, appear independent of the host organism. Many group I and group II introns can self-splice *in vitro* in the complete absence of proteins (Kruger *et al.*, 1982a; van der Veen *et al.*, 1986). Group I and group II introns possess completely different folded structures, yet the RNA of both intron types folds into a large and complex structure that positions the reactive groups, including metal ions, and catalyzes the splicing reaction (see

Chapter 1). The group II introns aI1 and aI2 can carry out almost their entire transposition pathway (splicing, reverse splicing, reverse-transcription) *in vitro* with intron RNA, the intron encoded protein and a 50 base pair piece of double-stranded target DNA (Zimmerly *et al.*, 1995b). When group I ribozymes were discovered by Tom Cech (Kruger *et al.*, 1982a), the existence of protein-free chemistry gave credence to the idea of an ancient "RNA world" that pre-dated protein-based life and from which proteins eventually evolved. Current progress towards supporting this theory is reviewed in (Muller, 2006).

In vivo, the story is a bit different. Modern-day group I and group II introns require varying degrees of protein assistance in order to splice and transpose. This is partly because the *in vivo* environment is not as facilitative of self-splicing as the high-salt and high temperature conditions necessary for optimal *in vitro* reactions (Perlman & Podar, 1996). The protein-dense intra-cellular (bacterial or mitochondrial) environment may also exacerbate the folding difficulties of large (>0.5 KB) RNA structures. Also, some splicing factors may have evolved as a means of controlling otherwise parasitic DNA sequences. Over time, the host may not be able to easily eliminate the parasitic intron DNA because splicing is required for gene expression. However, the host can eliminate its "infectiousness" and control further spreading by providing host proteins to suppress intron mutations (Eickbush, 2000). The ultimate example of this, if we assume that nuclear introns evolved from a parasitic precursor intron (which some would argue – (Roy & Gilbert, 2006)), would be the evolution of the nuclear spliceosome. The host has complete control of what once was a bit of "selfish DNA" and further has evolved to use introns as an integral part of gene expression and control.

Consistent with the idea that host organisms adapted individually to the presence of parasitic introns, the protein co-factors shown to date to be involved in group I and group II intron splicing vary significantly between different introns, even similar introns (yeast vs. mold vs. plants and between different group Is in yeast – see following section). This would indicate that the acquisition of protein co-factors is a dynamic and not easily predicted process.

4-1-2 Overview of known group I and group II intron splicing factors:

Although it is thought that most, if not all, group I and group II introns require some protein assistance to splice *in vivo*, only a handful of proteins are established splicing cofactors. In yeast, it is well established that each of the 13 introns requires at least some splicing assistance from intron-encoded or nuclear-encoded proteins (Huang *et al.*, 2005), although the list of known factors is far from comprehensive. Nuclear encoded splicing factors have also been studied in Maize (Ostheimer *et al.*, 2003; Pereira de *et al.*, 1991) and *Neurospora crassa* (Lazowska *et al.*, 1980; Paukstelis *et al.*, 2005). In general, many group I and group II introns do not self-splice *in vitro*, so it is largely assumed that proteins must play a role in splicing even in organisms for which no splicing factor has been identified.

4-1-2a Maturases

The first splicing factor identified was the "maturase," encoded by the group I intron bI2. The possibility of an intron encoded protein was indicated by the analysis of point mutations.

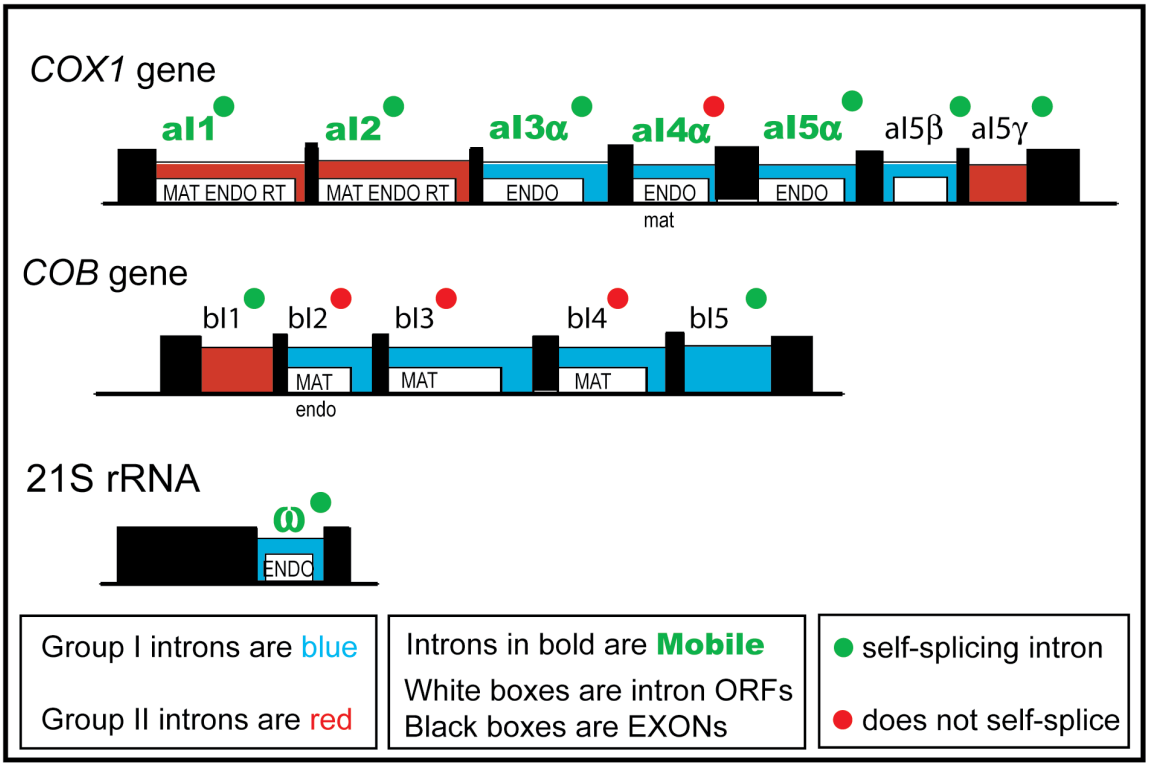


Figure 4-1. Intron encoded proteins. ENDO = active homing endo nuclease. endo = inactive homing endonuclease. MAT = active maturase, mat = inactive maturase. RT = Reverse transcriptase activity.

Point mutations within bI2 were mapped and shown to define intermingled complementation groups (Lazowska et al., 1980). Using a mitochondrial version of the complementation test, it was found that some mutations were *cis*-dominant while others were *trans*-recessive. That finding genetically proved that one segment of the *COB* gene codes for a trans-acting factor that is needed to splice the intron. Direct sequencing of the intron – which appeared in the same study – showed the unexpected finding that bI2 contains a reading frame and most of the *trans*-acting mutations truncated the reading frame without altering any of the *cis*-acting intron sequences later shown to be essential for splicing. Shortly thereafter, the same and other groups also obtained analogous genetic evidence showing that several other group I

introns (and later, two group II introns) also code for a maturase protein (Anziano et al., 1982; Carignani et al., 1983; Lazowska et al., 1980; Weiss-Brummer et al., 1982) (Figure 4-1). At the same time that intron encoded splicing factors were being identified, groups also started looking for nuclear encoded protein co-factors.

4-1-2b NAM2

The first nuclear encoded mitochondrial intron splicing factor was identified as the result of a screen for suppressors of intron mutations (Dujardin *et al.*, 1980; Groudinsky *et al.*, 1981). That screen utilized a collection of 250 mitochondrial mutants of the *COB* gene or its introns (mit-) created by the Tzagoloff group and others in the 1970's (Dujardin *et al.*, 1980; Slonimski & Tzagoloff, 1976). The strategy of the screen was to first mutagenize each strain in the mit- mutant collection individually with EMS. Revertants that could grow on glycerol were selected and analyzed. Revertants that grew exactly as well as wild-type on glycerol, ethanol or lactate at 18^{o}C, 28^{o}C or 36^{o}C were considered to be back-mutations and so were eliminated from further analysis, but strains that had partial growth under any of those conditions were further analyzed. This triage step reduced the number of candidates from thousands to hundreds. Next, they tested if the suppressor mutation was located in the nucleus or mitochondria by performing test crosses back to the original un-mutated strain, either with or without mitochondrial DNA. If the suppressor mutation was located on the mtDNA, then removal of the mtDNA also prevented suppression of the original mutation. Revertants were then grouped using traditional methods (linkage mapping) and placed into linkage groups.

This screen found a dominant suppressor of bI4 maturase mutations that was later demonstrated to be directly involved in the splicing of bI4 and aI4α. This gene was named *NAM2* (nuclear accommodation of mitochondria), and later found to also be the mitochondrial leucyl-tRNA synthetase (Herbert *et al.*, 1988). Normally, the aI4α maturase is inactive because it contains mutations that disrupt its splicing activity; consequently aI4α normally requires the bI4 maturase to splice. However, the NAM2 suppressor protein compensates for the defect in the aI4α maturase and acts with the aI4α protein in the suppressed strain (which entirely lacks bI4 maturase) to splice both bI4 and aI4α (Dujardin *et al.*, 1982).

NAM2 is required along with the bI4 maturase for the splicing of intron bI4 (fourth *COB* intron). *In vitro* work demonstrated that the splicing function of Nam2p is confined to an inserted region that is specific to *leucyl*-tRNA synthetases (Rho *et al.*, 2002). It was shown, using two and three hybrid analysis, that Nam2p and the bI4 maturase bind directly and independently to the bI4 intron to facilitate splicing (Rho & Martinis, 2000). The mechanism by which either Nam2p or the bI4 maturase facilitate splicing is unclear, but they may stabilize the active folded RNA, as is the case for *CYT18* (Paukstelis *et al.*, 2005).

Leucyl-tRNA synthetases from *S. douglasii* (Herbert *et al.*, 1988), *Mycobacterium tuberculosis* and human mitochondria (Houman *et al.*, 2000) are able to functionally substitute for the splicing function of NAM2 in yeast mitochondria. This indicates that the splicing ability of NAM2 is part of its conserved structure and not a diverged specialized function (as it is for another tRNA synthetase, CYT-18, see section 4-1-3).

4-1-2c CBP2

A number of groups generated large sets of PET strains in the late 70's, and these were extensively used to characterize mitochondrial proteins. A sub-set of this strain collection was used in the screen just described above. The Tzagoloff group generated a large percentage of these strains and went on to characterize many mitochondrial proteins (Tzagoloff & Dieckmann, 1990). Because cytochromes can be conveniently detected and differentiated using spectroscopic methods, this PET (for nuclear segregating petites) strain set was sorted into 136 complementation groups and then screened for specific spectral deficiencies. The cause of the cytochrome b (*COB*) deficiency in one of the complementation groups was found to be incomplete processing of the *COB* mRNA (McGraw & Tzagoloff, 1983). Only two *COB* introns, bI4 and bI5, are present in the strain used to make the Tzagoloff mutant collection, making it relatively easy to determine that the defect was the result of the failure to splice intron bI5. The gene responsible was cloned by complementation and named *CBP2* (for Cytochrome B pre-mRNA-processing protein 2) (McGraw & Tzagoloff, 1983). *CBP2* is not a widely conserved protein. It has a close ortholog in *S. douglassi* that is also required for intron splicing (Li *et al.*, 1996). The only other possible homolog present in genebank (NCBI) is a protein with weak homology (24% identical) found in *Candida glabrata* (gi:49525168; (Dujon *et al.*, 2004)).

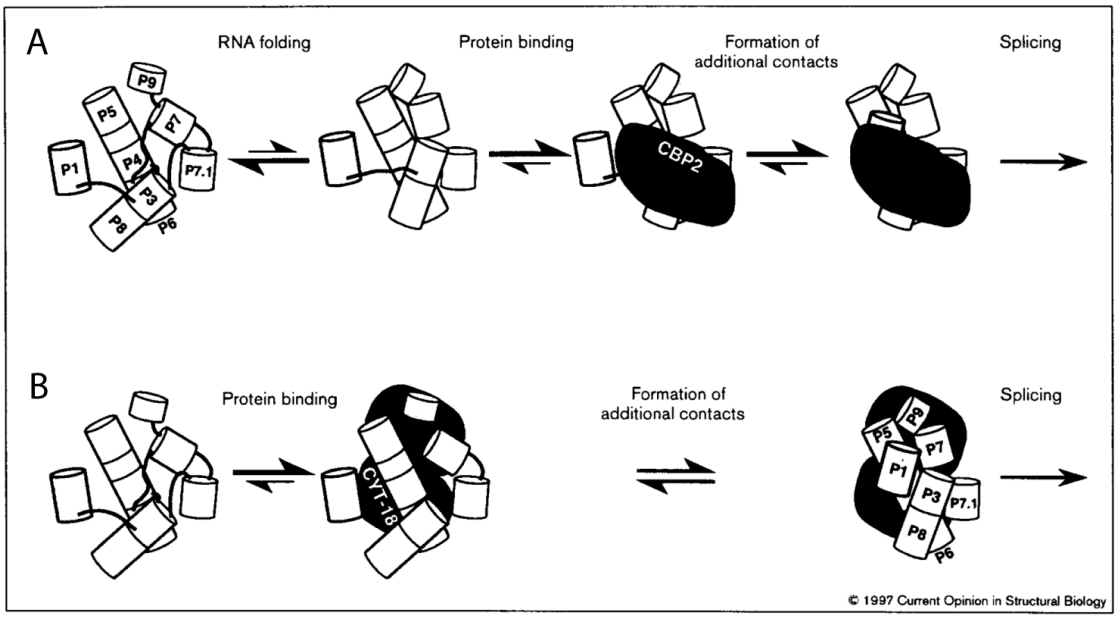


Figure 4-2. Protein assisted group I intron folding. (A) Cbp2p stabilizes the correctly folded intron. (B) CYT-18 binds the intron and then facilitates the formation of the correct RNA fold. Based on studies with intron bI5(Webb et al., 2001). Proteins are depected in black, intron structure is depicted as nine cylinders representing helical domains P1-P9. Figure adapted from (Weeks, 1997).

Initial biochemical experiments in the Tzagoloff laboratory (Gampel *et al.*, 1989) showed that Cbp2p facilitates the splicing of the terminal group I intron of the *COB* gene. Later, more detailed biochemical studies performed by Kevin Weeks demonstrated that Cbp2p acts primarily by binding and stabilizing the correct folded RNA structure (Weeks, 1997; Weeks & Cech, 1995a; Weeks & Cech, 1995b; Weeks & Cech, 1996). Weeks has compared the mechanism by which Cbp2p facilitates bI5 splicing to that by which CYT-18 facilitates splicing also of intron bI5 (although it is not the native substrate for CYT-18). He found that Cbp2p and CYT-18 bind to opposite sides of the intron RNA and that CYT-18, unlike Cbp2p, actually binds the intron RNA in a pre-folded state and facilitates the correct fold (Webb et al., 2001), (see Figure 4-2 adapted from Weeks, 1997). It is important to note that intron bI5 neither contains an open reading frame nor does it use the maturase of another

intron. In addition, intron bI5 is a type group IA intron meaning that it does not possess the stabilizing P5abc domain. It has been postulated that Cbp2p physically substitutes for this missing domain (Tirupati *et al.*, 1999). The upshot of all this is that intron bI5 cannot splice in the absence of *CBP2 in vivo*. This is also true for the *Δcbp2* strain in this study in which the *COB* gene is completely blocked for bI5 splicing.

4-1-2d MRS1

Another screen published in 1983 isolated mutants that were later shown to include the disruption of one genuine splicing factor (*MRS1*) and two translational activators of *COB* mRNA (*CBS1, CBS2)* (Pillar *et al.*, 1983). This screen used the same basic method as the McGraw and Tzagoloff screen that isolated *cbp2* mutants, except that it generated novel PET strains from a yeast strain that contained all twelve *COB* and *COX1* introns and employed sodium nitrite as the mutagen. Pillar *et al.* restricted their analysis to petite mutations that were absolutely unable to grown on glycerol or ethanol even at 30^{o}C but that had retained wild-type mtDNA. They used spectrographic analysis to determine which mutants had specific defects in cytochrome aa_3 (*COX1*) or cytochrome b (*COB*) production. They reported mutants falling into five complementation groups (MK1-MK5) that contain unspliced *COX1* or *COB* message. Apparently they also isolated a number of other mutants (including *cbs2*), but did not discuss them or even say how they narrowed the list to five. Mutant MK3 was shown in a subsequent paper to be required only for the splicing of bI3 (Kreike et al., 1986- they did not look at COX1 splicing). The mutated gene was cloned and renamed *MRS1* (Kreike *et al.*, 1987). A mutant named *PET154* was independently isolated

by the Slonimski group and found to be *MRS1*. In that study they examined both *COB* and *COX1* genes and found that Mrs1p is also required to splice intron aI5β (Bousquet *et al.*, 1990). *MRS1* is homologous to the gene encoding the yeast nuclear-encoded mitochondrial cruciform cutting enzyme (CCE1). CCE1 is a member of the RNase H fold family, but Mrs1p is highly diverged from this fold. Mrs1p is proposed to have evolved in *Saccharomyces* from *CCE1*, presumably because cruciform binding was amenable to adaptation to group I intron binding (Wardleworth *et al.*, 2000). *MRS1*, like *CBP2,* only has homologs in *S. douglasii* and *Candida glabrata* (BLASTP at NCBI).

In vitro, bI3 requires both Mrs1p and the bI3 maturase in order to splice. A single molecule of the maturase binds the RNA, and this binding is independent of the binding of two dimers of Mrs1p (Bassi *et al.*, 2002). This established that Mrs1p, like its homolog Cce1p, functions as a dimer, but Bassi *et al.* show that Mrs1p is unable to cleave DNA junctions. Finally, the authors propose that Mrs1p functions to compensate for deficiencies in the bI3 maturase. They demonstrated that the bI3 maturase has lost its endonuclease activity due to mutations in the conserved LAGLIDADG motif. If so, this is a nice example of the process by which a parasitic DNA element is converted into a host-controlled non-infectious intron.

This process is perhaps even one step further along in the case of intron aI5β, which possesses an un-expressed ORF remnant and is spliced entirely by nuclear-encoded proteins including Mrs1p. The other two proteins that have been convincingly shown to be involved in aI5β splicing are Pet54p and Mss18p. It has been demonstrated that mitochondrial translation not required for aI5β splicing (Johnson C.H. & McEwen J.E., 1997). Partial

inhibition of SUV3, a component of the mitochondrial degradasome, also reduces aI5β splicing (Conrad-Webb et al., 1990). This phenotype is almost certainly a secondary effect caused by the accumulation of group I intron RNAs (due to inhibition of the degradation of the excised RNAs) and consequent sequestering of splicing factors.

4-1-2e PET54

PET54 was first described as a gene required for the accumulation of Cox3p. In a *pet54* null strain, wild-type levels of *COX3* mRNA are present but no Cox3p is made. This places Pet54p in the class of nuclear-encoded mitochondrial proteins that are required for the translation of most mitochondrial mRNAs (Costanzo *et al.*, 1986). Some of these translation activators were originally isolated as intron splicing factors but were later shown to only affect splicing indirectly by way of inhibiting maturase translation. *PET54,* however, was also shown to be required for the splicing of the *COX1* intron aI5β (Valencik *et al.*, 1989). In that study it was shown that a *Δpet54* strain accumulates a single 3.2 KB *COXI* precursor RNA that hybridizes to aI5β probes; also, a strain lacking intron aI5β only contained the 2.2KB *COX1* mRNA band. Those data conclusively show that Pet54p is only required for the splicing of intron aI5β; however, in this and all subsequent *PET54* papers, the exon probe used hybridizes to the 4th exon (prior to introns aI4, aI5α, aI5β and aI5γ), and they could not detect any RNA species that did not contain exon 4 – for example a half-spliced intron-3'exon molecule or a mis-spliced RNA that skipped exon 4.

Pet54p has a number of clear homologs in other yeasts and fungi, none of which has been characterized. A systematic mutation study of Pet54p revealed that the *COX3*

translation and *COX1* splicing functions were separable (Valencik & McEwen, 1991). A BLAST search revealed that Pet54p possesses a RPM1 domain in the C-terminal half of the protein. The RPM1 domain is a RNA-binding motif found in a large number of proteins, particularly RNP proteins (Marchler-Bauer & Bryant, 2004). In the Valencik & McEwen 1991 study, one mutation was located in this RPM1 domain and it disrupted the *COX3* but not *COX1* function of Pet54p. Both of the mutations that disrupt the *COX1* splicing function and several that disrupt the *COX3* translation function are located in the N-terminal half of Pet54p.

Further studies into the translational activator role of Pet54p have shown that it binds directly to sites in the 5'-untranslated region of *COX3* mRNA(Valencik & McEwen, 1991). Also, Pet54p is a peripheral membrane protein that makes direct contacts with a number of integral membrane proteins also involved in translation (Towpik, 2005). It is still not clear if the splicing role of Pet54p is somehow related to its role in organizing translation, perhaps related to the sub-cellular localization of aI5β messages, or if the two functions are completely separate. Because my research identifies even more proteins potentially involved in aI5β splicing, it is reasonable to postulate that sub-mitochondrial localization may be critical to proper splicing.

4-1-2f Screen that isolated the MSS genes.

Several genes named *MSS* (mitochondrial splicing system) were isolated in a screen similar to those already described (Faye & Simon, 1983). Twenty-one pet mutants defective in cytochrome oxidase (generated by the authors and from the Tzagoloff group) were

screened directly by northern blot for mitochondrial intron splicing defects. Two mutants had a strong general splicing block and accumulated many precursors; these were named *mss51* and *mss116*. Another mutant, *mss18*, had a partial splicing defect and accumulated a 4.2KB *COX1* precursor band. This phenotype analysis is very similar to the one presented in this work and the mutants isolated are representative of the phenotypes I observed: i.e., strains with many, often large, precursor bands and a reduced level of spliced message (like *mss116*); or strains with an increased signal for one or at most only a few precursor bands (like *mss18*). The authors first cloned and analyzed *mss51* (Simon & Faye, 1984), which ironically turned out to be a translational regulator of *COX1* that binds newly synthesized COX1p and regulates its translation as part of a larger cytochrome c oxidase assembly quality control mechanism (Barrientos *et al.*, 2004). This is yet another case of a putative splicing factor that turned out to have its primary function in mitochondrial translation.

The same group went on to perform a more involved screen using an intron-less (I^o) mtDNA to search for nuclear proteins that splice mitochondrial introns (Seraphin *et al.*, 1987). First they generated the intron-less mtDNA and created a collection of nuclear petite strains using EMS mutagenesis followed by sorting the strains into complementation groups. In this manner they generated 201 pet mutant strains from different complementation groups. They transferred the intron-less mitochondrial DNA into a ρ^o (no mtDNA) version of each pet mutant strain and scored for mutants that regained glycerol growth when I^o mtDNA was provided. The assumption was that any gene chiefly required for the splicing of introns would *not* be necessary for glycerol growth in a strain lacking introns. In this manner, they re-isolated two mutants, *mss116* and *cbp2,* and showed northern blot data for three additional

mutants having moderate general splicing defects. No further work has been published on the new mutants isolated in that screen. That study generated the intron-less (I^o) mitochondrial genome used in this work.

4-1-2g MSS18

Simon and Faye had better luck isolating a splicing factor with two other mutants (*mss18* and *mss116*) isolated in their first screen. They cloned the *MSS18* gene and showed that it is specifically required for the splicing of the *COX1* intron aI5β (Seraphin *et al.*, 1988). Using S1-nuclease protection assays, they demonstrated that *MSS18* is specifically required for the 5' cleavage reaction, and suppressors of *mss18* mutants were found in the last 13 nucleotides of the upstream exon. They proposed that Mss18p may bind to the upstream exon and somehow promote the proper folding of the exon-intron boundary that is critical for the first splicing step. Mss18p has no apparent homologies to known proteins and no further work has been done with *MSS18*. I also observe that the *Δmss18* strain accumulates an aI5β containing precursor band.

4-1-2h MSS116

MSS116 codes for a DEAD-box RNA-helicase-like protein and was cloned and initially found to be involved in splicing of at least some mitochondrial introns and also to have a translation defect (Seraphin *et al.*, 1989). *MSS116* was also shown to be capable of stimulating the splicing of the *COB* intron bI1 in an *in vitro* system using a crude mitochondrial extract (Niemer *et al.*, 1995). A former graduate student of the Perlman lab,

Dr. Hon-Ren Huang, was the first to do a detailed analysis of the splicing and translation functions of *MSS116* (Huang *et al.*, 2005). He showed that *MSS116* is a general splicing factor involved in the splicing of all group I and group II introns in yeast mitochondria. Because *Δmss116* also has a partial mitochondrial translation defect, it was important to show that the splicing defect in this strain was not an indirect result of the translation defect, as had proved to be the case for many previously isolated "splicing factors". To do this, Huang *et al.* used petite mutants deleted for most of the translation apparatus and looked at the effect of *Δmss116* on the splicing of introns aI5γ, bI1 and omega. All three introns were able to splice efficiently in the control (*MSS116*) strain, although not as well as wild-type. Introducing the petite mtDNA into the *Δmss116* strain background resulted in a further reduction in the splicing of all three introns – down to the level observed in ρ^+ strains. For the two group II introns, aI5γ and bI1, splicing was essentially inhibited, making *MSS116* the first gene to be conclusively shown to be necessary for splicing of group II introns. (All of the previously discussed splicing factors were for group I introns.)

To show that each intron was independently affected in the *Δmss116* strain background, Huang *et al.* generated a set of single-intron-containing mtDNAs and used northern blot analysis to measure the level of splicing in each. The single intron strains permitted him to obtain conclusive evidence that inhibitory effects on a given intron are not caused indirectly by effects on another, perhaps master, intron. Also, this device provided much simpler RNA blots, and the effect of *Δmss116* on splicing of each intron could be quantified for the first time. For several group I introns, which were not of particular interest to the lab at that time, he analyzed strains with two introns in a given gene so that some

interdependence could not be ruled out. He found that all introns are affected, although not to the same extent. The splicing of all four group II introns is reduced to barely detectable levels, whereas the splicing of the group I introns is visibly reduced in all cases, but varies from 20% to 50% spliced message. This same set of individual intron strains will be instrumental for further characterizing any new mutant strain that has an apparent splicing defect isolated in the current screen.

The original *Δmss116* mutant strain (GRF18/B601B) could not grow on glycerol medium, even in the absence of introns, probably because of a translation defect (Seraphin *et al.*, 1989). The *Δmss116* intron-less strain made in our lab is derived from a different parental yeast strain (ID41-6/161) and is cold sensitive for glycerol growth, but can grow on glycerol medium under normal conditions. This illustrates that strain background differences can significantly alter a mutant phenotype. For Dr. Huang's *Δmss116* strain, the more introns that were added back, the slower the strain grew on glycerol, with 3 introns in a given gene being sufficient to eliminate glycerol growth; a strain with two group I introns in *COB* and two others in *COXI* grows on glycerol medium. In general, we find that very little spliced *COX1* or *COB* message (~1%) is necessary for good glycerol growth. The *Δmss116* strain from the commercial library of yeast null mutants (BY4741) could not be included in my analysis because it had lost mtDNA (ρ^o). After completion of the screen, I was able to isolate a gly^- ρ^+ *Δmss116* strain, so any future experiments should include an analysis of this strain.

Huang *et al.*, (2005) also showed that a homologous gene from *N. crassa*, CYT-19, functionally rescues the splicing and translation defects of *Δmss116*, at least partially. Interestingly, Cyt-19p only very weakly rescued aI2 splicing defect of the *Δmss116* strain,

although Huang *et al.* was able to show that translation of the aI2 encoded maturase was restored by the CYT-19 knock-in. Furthermore, an unspliced aI2-maturase complex could be detected in mitochondrial extracts from the *Δmss116*/CYT-19 strain. They proposed that Mss116p binds to the aI2 RNA after its maturase and functions as a RNA chaperone to release aI2 from kinetic traps. The helicase function and generality of Mss116p set it apart from several other described splicing factors, but like Mrs1p, it does not appear to form a higher order complex with an intron encoded maturase.

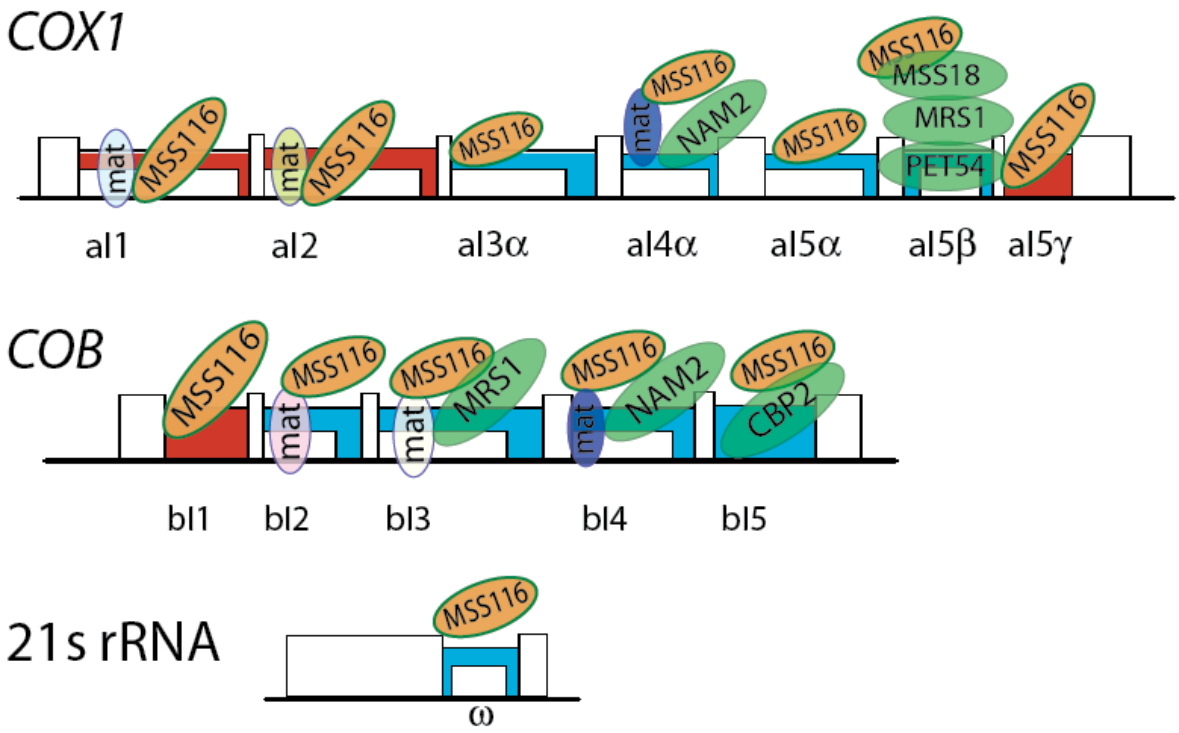


Figure 4-3. Summary of published mitochondrial intron splicing factors. Nuclear encoded proteins are shown in green with the exception of Mss116p which assists the splicing of every intron and is shown in orange. Intron encoded maturases are indicated by "mat".

4-1-3 Biochemical studies of group I intron splicing factors in *N. crassa*.

Direct *in vitro* evidence of a splicing role is ideally required to establish a given protein as a splicing factor, but has been largely lacking for many yeast intron splicing factors. CYT-18 from *N. crassa* was the first bi-functional protein for which it was conclusively shown that it is both the *N. crassa* tyrosyl-tRNA synthetase and yet also directly facilitates group I intron RNA splicing (Majumder *et al.*, 1989).

CYT-18 protein aids the splicing of several *N. crassa* group I introns and is also the *N. crassa* tyrosyl-tRNA synthetase (Akins & Lambowitz, 1987). A subset of *N. crassa* introns require CYT-18 for splicing both *in vivo* and *in vitro*; notably the intron located in the *N. crassa* large ribosomal subunit (LSU) and the intron located in the *NDI* gene (Caprara *et al.*, 1996). CYT-18 has additional domains, not found in most tRNA synthetases, which are required for its intron splicing function. Furthermore, those domains are differentially required for splicing of the LSU and *NDI* introns (Mohr *et al.*, 2001). Despite structural similarities between group I introns and tRNAs, a recent crystal structure of CYT-18 revealed that group I introns and tRNAs are bound by non-overlapping sites on the protein (Paukstelis *et al.*, 2005). CYT-18 appears to have independently evolved intron splicing since the divergence of molds. The yeast protein Nam2p, which is also a tRNA synthetase and intron splicing factor, is not functionally similar to CYT-18. Unlike CYT-18, diverse homologs of Nam2p are able to substitute for the splicing activity of Nam2p (Houman *et al.*, 2000).

CYT-19, like Mss116p, is a member of the DEAD-box family of RNA helicases and has been shown to interact with CYT-18 to facilitate the splicing of the same group I introns that require CYT-18 (Mohr *et al.*, 2002). Interestingly CYT-19 is able to partially substitute for Mss116p in yeast, rescuing both group I and group II intron splicing defects. Both CYT-19 and Mss116p are believed to act as RNA chaperones that release misfolded RNA from kinetic traps (Huang et al., 2005; Mohr et al., 2002; Mohr et al., 2006).

4-1-4 Plant chloroplast group I intron splicing factors

The protein assisted splicing of chloroplast introns is even less well understood than that of yeast mitochondrial introns. Plant group I and group II introns generally contain many mutations and other structural deviations that evidently explain the absence of self-splicing and are believed to rely heavily on protein co-factors for splicing. Plant chloroplast introns may even share maturases (Vogel *et al.*, 1999) and/or use maturases whose genes are no longer located within an intron (Mohr & Lambowitz, 2003). Several nuclear genes affecting chloroplast intron splicing have been isolated in corn (Jenkins *et al.*, 1997). One of these, CRS2, is related to peptidyl-tRNA hydrolase and possesses the characteristic CRM RNA binding domain (Jenkins & Barkan, 2001). The other three plant splicing factors - CRS1, CAF1, and CAF2 - have no homology to known proteins *except* that they also posses CRM domains (Ostheimer *et al.*, 2003). It appears that the only unifying factor between yeast and corn splicing factors is that they both employ RNA binding domains, particularly those that can also bind tRNAs.

4-1-5 Genes that are not splicing factors.

When looking at all known group I intron splicing factors, it appears that they can be single or multi-functional proteins and diverse RNA binding domains have been co-opted to serve as splicing functions. The diversity and multifunctional nature of the known splicing factors makes it difficult to predict potential splicing factors from genomic data bases. Also, it is difficult to establish that a given protein has a direct effect on splicing. Many genes were originally isolated as splicing factors but later shown to alter splicing indirectly by altering RNA transcription or metabolism.

4-1-5a Mitochondrial RNA turnover (SUV3, PET127)

SUV3 is a DEAD-box RNA helicase that is part of the mitochondrial RNA "degradasome" which also includes DSS1 (Dziembowski *et al.*, 2003). This complex functions to unwind, degrade, and separate RNA from protein cofactors in yeast mitochondria. In the absence of this complex, group I intron RNAs accumulate at high levels in mitochondria (Margossian *et al.*, 1996).

SUV3-1 – a dominant partial loss of function mutation – was found as a suppressor of a dodecamer deletion of the VAR1 gene (Conrad, 1987; Conrad-Webb *et al.*, 1990). It was characterized extensively and found to partially inhibit a number of RNA processes in mitochondria (Conrad-Webb *et al.*, 1990). It has been considered to be a splicing factor because the SUV3-1 mutant may have a partial defect of aI5β splicing (Conrad-Webb *et al.*, 1990). This splicing defect may be caused by a failure to recycle splicing co-factors that are

bound tightly to the intron RNA (thus sequestering proteins and leading to an apparent knock-down of splicing factors). SUV3-1 does not, however, affect the splicing of most other introns, all of which were tested (Conrad-Webb *et al.*, 1990). The null mutant strain cannot maintain wild-type mtDNA and the reason is not certain – it may be failure to recycle ribosomal proteins – but is not a defect in splicing (Margossian *et al.*, 1996). *Δsuv3* causes ρ^+ mtDNA loss and was not analyzed by northern blot in this study.

Overexpression of another RNA processing protein, Pet127p, can rescue the phenotypes caused by a *suv3* null mutation (Wegierski *et al.*, 1998). In this study, I used the PET127 deletion as a positive control because it contains many RNA bands. *Δsuv3* causes ρ^+ mtDNA loss and could not be looked at directly by northern blot.

4-1-5b MRS2

MRS2 was identified as a multi-copy suppressor of mutations in intron bI1 that block splicing (Koll *et al.*, 1987), and further publications supported a role for Mrs2p in intron splicing (Schmidt *et al.*, 1998; Wiesenberger *et al.*, 1992). It has since been shown to be a mitochondrial magnesium transporter (Gregan *et al.*, 2001). Mutations of *MRS2* that result in increased Mg^{2+} in the mitochondria are able to suppress a variety of intron mutations (but by no means all mutants) (Gregan *et al.*, 2001; Schmidt *et al.*, 1998). This is not surprising, because magnesium is a critical ion for proper intron folding. *In vitro,* small changes in magnesium concentration can significantly increase the rate of intron splicing (Perlman & Podar, 1996).An *MRS2* allele that increased splicing and raised the steady state levels of intron containing RNAs were used in Chapter 3 as a tool for studying intron point mutations

in vivo, but it should not be considered a true intron splicing factor. This is just one more example of the many ways in which a gene mutation or deletion can indirectly affect intron splicing.

The *Δmrs2* strain is a good example of how the yeast strain background can affect the glycerol growth phenotype. The first disruption of *MRS2* in the aDBY747 background resulted in a strain that was unable to grow on glycerol and accumulated numerous intron-containing precursor RNAs (Wiesenberger *et al.*, 1992). The deletion of *MRS2* in one of our standard yeast backgrounds (DBY947; Luban et al., 2005), and in the library background, αBY4741, has very little effect on glycerol growth. Luban *et al.* did observe a reduced level of aI5γ splicing (50%) in the DBY947 strain background. I included the αBY4741 *Δmrs2* strain on my blots probed for each intron (Appendix E), and although the pattern of RNA bands is not the same as wild-type, it is not consistent with an intron splicing block. It is likely that Mg^{2+} is rate limiting in the original aDBY747 *Δmrs2* strain background but not in the other strains.

4-1-5c Mitochondrial translational regulation.

Many genes that were originally isolated as intron splicing factors were later found to be message-specific translation regulators. The cause for this mix-up is simple: the majority of mitochondrial introns require an intron-encoded maturase to splice. Failure to translate the intron host gene also results in failure to translate the intron-encoded proteins and the strain phenotype appears like a splicing block of all maturase requiring introns. It is only in the absence of introns that the primary translational defect becomes obvious. This is the reason

that many translational regulators have the same three letter gene designator as splicing factors. Just because a protein is involved in translation does not rule out a direct role in splicing. Some proteins, like Pet54p and Mss116p, actually serve dual roles in both translational regulation and intron splicing; reaching that conclusion required much more than genetic and RNA blot evidence. One cannot, therefore, summarily dismiss all translational regulators. Most mitochondrial mRNAs require message specific translational activators (Towpik, 2005). Detailed studies by Tom Fox's group have shown that these proteins act by recognizing and binding to RNA sequences/structures in the 5' leaders of their respective proteins (Fox, 1996; Naithani *et al.*, 2003). This RNA binding ability may make these proteins good candidates for RNA splicing factors. Figure 4-25 shows these genes, what message they bind and how they link the RNA polymerase to the ribosomes and the inner membrane. This topic is addressed in more detail in the discussion.

4-1-6 Screen of the systematic yeast deletion library for nuclear encoded mitochondrial intron splicing factors.

The goal of this project is to find novel proteins involved in the splicing of mitochondrial introns. The yeast genome contains ~5000 non-essential genes, most of which are available as individual knockout strains for this screen. The following strategy was used for the screen: From the ~5000 yeast strains in the library, strains deleted for genes coding for potential or actual mitochondrial proteins were collected into a mitochondrial protein sub-library. The sub-library was further divided into three categories: 1) strains that grew as

robustly as the wild-type strain on glycerol medium, which were not screened further; 2) strains that did not grow as well as the wild-type strain on glycerol medium and also contained mtDNA, which were analyzed in Phase I screening (see below); and 3) strains that did not grow in glycerol medium and lacked mitochondrial DNA, which were reserved for Phase II screening (see below). Phase I of this project – the direct screen of candidates that contain mtDNA – has been completed and a number of strains with interesting splicing defects were found.

4-2 RESULTS

4-2-1 Strains included in the yeast mitochondrial protein knockout sub-library

Genetic and biochemical studies have identified genes that code for mitochondrial proteins (ox phos, TCA cycle, etc) (Grivell *et al.*, 1999; Sherman, 1963). More recently, as genomic approaches have been implemented, several projects have sought to identify all proteins that are found in the yeast mitochondria (i.e., the yeast mitochondrial proteome). The most basic method is to screen all yeast deletion strains for respiratory growth defects (Steinmetz *et al.*, 2002). A more involved approach was to purify intact mitochondria or mitochondrial complexes and then use mass spectrometry to identify proteins found in the purified sample (Gan *et al.*, 2002; Gavin *et al.*, 2002; Prokisch *et al.*, 2004; Sickman *et al.*, 2003). Another approach was to make a library of yeast strains in which each gene had been

tagged (GFP or epitope tag) and then individually determine the localization of each protein (Huh *et al.*, 2003; Kumar *et al.*, 2002). For a thorough review of these and similar studies see (Reichert & Neupert, 2004). Hon-Ren Huang started this project by compiling a list of the mitochondria-localized proteins identified by those published studies. He also included genes that resulted in a glycerol growth defect when deleted but did not show up in any of the mitochondria specific localization studies (Steinmetz *et al.*, 2002). These latter genes may cause a growth defect for a variety of reasons. For example, some may encode a nuclear transcription factor required for the regulation of important respiratory genes; others may represent a very low abundance mitochondrial protein that other studies missed; still others might be non-mitochondrial proteins which, when absent, interfere with mitochondrial function indirectly (by altering protein sorting or membrane composition, for example). Finally, some candidate genes code for "dubious" ORFs that either overlap, or are immediately adjacent to a gene for a mitochondrial protein known to be necessary for glycerol growth. In sum, Dr. Huang's list contains 864 genes. Forty-four of these are essential genes that cannot be analyzed as haploid knockouts. Twelve of them are not present in the commercial haploid deletion library and no steps were taken to analyze them. (See Appendix B, Table B-1 for a list of the genes that were not included in this study.) This leaves a total of 808 unique genes that are present in the haploid deletion strain set and could be included in the screen.

I purchased the yeast haploid *MATa* deletion set which contains 4839 strains spotted onto rectangular YPD plates in a 12 x 8 grid. When the library arrived, the entire lab assisted in picking the desired strains off of the agar plates into a new sub-library grid consisting also

of 96-well plates. This sub-library actually contains 866 strains on 19 plates. There are a number of reasons for the increase in strains over the list of 808. Nine of the desired gene knockouts occur twice in the library; in some instances, only one of the two strains is correct (according to the material provided with the library) so we picked both strains to be conservative. Genes coding for three mitochondrial ribosomal proteins were missing from the original list and were added back from the purchased library plates, which had been sealed and stored at 4 ^{o}C for several months. Nine strains were added because an incorrect strain was accidentally picked off the original library grid. The correct strain was also picked, but it was more efficient to just include these strains in the library than to re-pick the entire library, potentially introducing more errors.

Finally, a number of gene knockouts were present on the purchased library plates as petite or PET mutants (small and white, indicating no mitochondrial function), but had not been reported to be glycerol growth deficient and so were not on our list to be picked. Some of these could be relevant to this screen, but most are probably artifacts due to random or slow loss of wild-type mtDNA in a strain that can maintain mtDNA and is not gly^{-}. I decided to include these strains for a total of 37 additional strains (they later proved to be false negatives).

Altogether, 866 candidate genes were placed in the mitochondrial proteome deletion sub-library. The library was picked into 96-well plates containing 200µl of YPD 20% glycerol and split into two sets of 96-well plates with 100µl each. One of these sets was immediately sealed and stored at -80^{o}C and the other was used to replica-plate onto Petri

dishes containing the desired media, and then stored at -80°C as a backup. The next step was to determine the growth phenotypes of this new library.

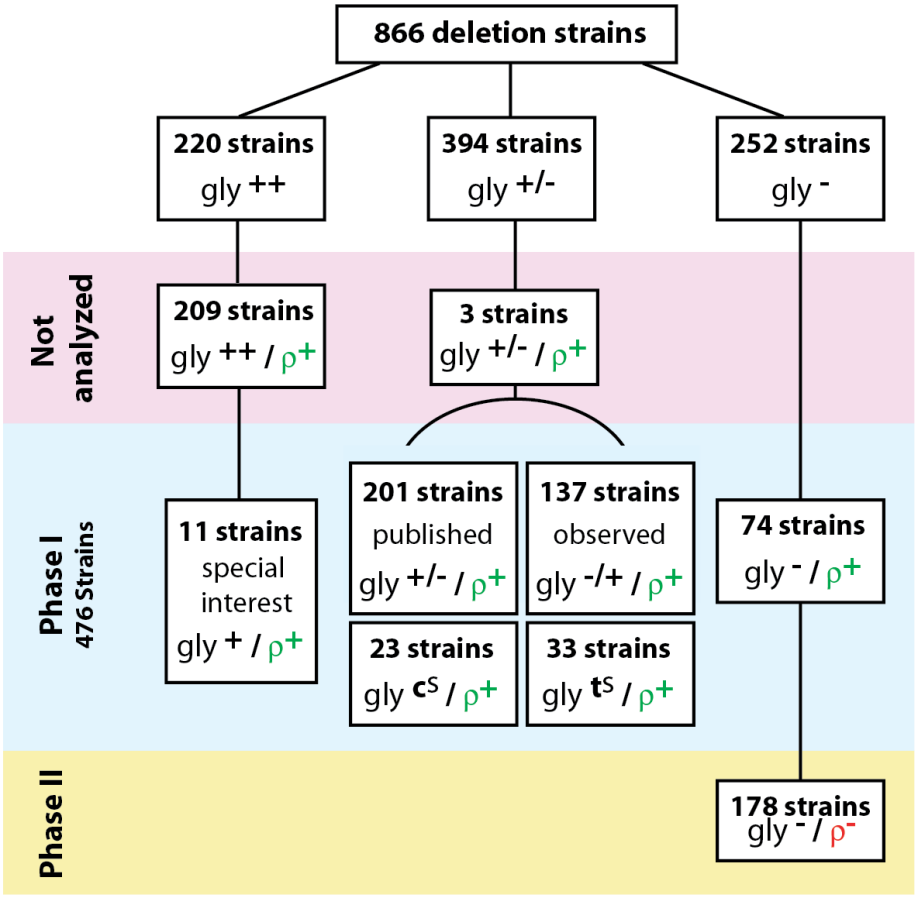


Figure 4-4. Growth phenotypes of the mitochondrial deletion strain library. 866 deletion strains were tested for glycerol growth and the presence of mtDNA. 476 strains with a glycerol growth defect retained mtDNA and were analyzed in phase I of the screen. 178 strains lost mtDNA and cannot be analyzed directly.

4-2-2 Screen of the mitochondrial deletion library for glycerol growth defects and the presence of mtDNA.

Reasonably efficient mitochondrial intron splicing is required for growth on glycerol containing medium, especially in this strain which has 13 introns in its mtDNA. The sub-

library was assayed for growth by plating on YPD and YPG (YP-glycerol) at 20°C, 30°C and 37 °C. Since careful growth curves of the entire library have been previously published (Steinmetz *et al.*, 2002), this assay was intended as a quick screen to confirm the published findings. Steinmetz *et al.* perform careful growth curves and found that 214 strains had a strong reduction in growth on glycerol medium and an additional 288 strains had a mild glycerol growth defect. Figure 4-4 summarizes my growth data. A total of 252 strains did not grow at all on glycerol medium and 196 strains had a detectable partial respiratory growth defect; of the latter strains, 33 grew more slowly at 37 °C (t^s) and 23 grew more slowly at 20 °C (c^s) than at 30 °C. An additional 210 strains have a published glycerol growth defect (Steinmetz *et al.*, 2002) but I did not observe an overt reduction in colony growth on glycerol Petri dishes. The published work generated growth curves from cells grown in liquid culture – technically the correct method of detecting subtle reductions in glycerol growth. I only scored strains with obvious reduction in growth and it is not surprising that a number of strains were missed. For many of these strains I looked up the published growth curves (Steinmetz *et al.*, 2002) and found that they have mild glycerol growth defects. Either way, all strains with published glycerol growth defects were included in this screen.

Each tight gly^- strain was tested for the presence of ρ^+ mtDNA by mating the library strains to the gly^- yeast strain αW303 ρ^0, which is expected to complement any recessive mutation in a library strain. Elizabeth Pickett used DAPI staining to show that the tester strain completely lacks mtDNA. When a library strain contains mitochondrial DNA it will complement the glycerol growth defect of strain αW303 ρ^0; if the library strain has lost ρ^+ mtDNA, diploids resulting from the cross will fail to grow on glycerol medium. Of the 252

haploid gly$^-$ strains, 74 were found to contain intact mtDNA and 178 did not. It is important to note that, of the 37 strains that were unexpectedly petite on the library plates we received, all lacked mtDNA. Restoring mtDNA to these strains also restored the ability to grow on glycerol (See section 4-2-12a). so those strains were not analyzed further. The ρ^+ strains may be analyzed by northern blot analysis without further workup.

An additional 11 strains grow well on glycerol but were included in Phase I of the screen because they were potentially good splicing factor candidates. They coded for uncharacterized mitochondrial proteins or are genes with known RNA or DNA binding properties. Strains deleted for *LSC1*, *NFU1* and *ATP20* should have been included in Phase I of the screen, but were accidentally excluded. The remaining 209 strains that did not have any observable or published glycerol growth defect were excluded from phase I analysis. It is possible that some of them could have a partial mitochondrial splicing defect, but they were not included in the high priority group for initial screening. This leaves 476 strains that were screened by northern blot analysis in Phase I (Figure 4-4). The 178 strains that have lost mtDNA can be screened, along with the strains that lost their mtDNA too rapidly to be screened by northern blot, using a mitochondrial plasmid, or pmit, in Phase II of this project.

4-2-3 Phase I: Screen of 476 strains by northern blot analysis for splicing defects.

The mitochondrial genome of the deletion library strains contains 13 introns (Figure 4-1). Any deletion strain of interest that has retained mtDNA can be screened directly for splicing defects in any of the 13 known introns without having to cytoduct in a different

mitochondrial genome that has fewer introns. As outlined in the previous section, I selected 476 deletion strains to screen for splicing defects by northern blot analysis.

Northern blot analysis is the preferred method of screening for splicing defects because it provides a direct visualization of levels of spliced mRNA, various pre-mRNAs and even the presence of unanticipated novel RNA species. In many cases, an altered band pattern can be interpreted and further analyzed to define the specific defect that is associated with a given gene knockout. Changes in levels of RNA species can only be visualized if the non-mRNA species are stable. Fortunately, yeast mitochondrial mRNAs contain a stability determinant at their 3'ends (Costanzo & Fox, 1990) and any RNA that retains the terminal exon is generally stable, even if it contains one or more unspliced introns. For example, our lab has published evidence that unspliced and mis-spliced RNAs, and even a splicing intermediate containing an intron and the downstream exons (generically called IVS-E2 species), are readily detected on blots of RNA from appropriate mutant strains (Boulanger *et al.*, 1996; Peebles *et al.*, 1986b; Peebles *et al.*, 1993). Low levels of most of those species are also present in RNA from wild-type strains and can often be detected on blots using intron-specific probes or RT-PCR assays. The *COX1* and *COB* probes used in this study are complementary to the terminal exon of each gene because those probes allow the visualization of spliced mRNA, all precursor RNAs and most mis-spliced RNAs. There is a remote possibility that a mis-splice could remove the part of the last exon that is complementary to the probes, in which case the RNA species would be missed in this analysis. In studies conducted by others (Luban *et al.*, 2005; Valencik *et al.*, 1989), probes

complementary to internal exons were used, and in several cases missed relevant phenotypes that I detected.

It would not have been possible to screen 476 strains and perform more than 600 northern analyses in such a short time without the assistance of Stuart Connell, a technician in the lab, who with my guidance performed many of the Northern blots for this portion of the screen. Technical details are described in Chapter 2. In brief, RNA was extracted from late log phase liquid cultures and the RNA fractionated on guanidinium thiocyanate agarose gels. This gel system resolves RNA species well and even separates lariat from linear RNA. All strains on a single blot were usually grown and prepped together but not balanced prior to loading – except to start with the same size culture and using the same volumes in all steps of the preparation. However, the uniformity of the protocol usually resulted in reasonably well balanced blots. The fractionated RNA was transferred onto nylon membranes and probed with ^{32}P-labeled oligonucleotides. Probed blots were analyzed by phosphorimaging which permitted quantification of the findings. Subsequent hybridization with other probes was always carried out by re-probing the same blot.

4-2-3a General description of observed outcomes and how they are interpreted.

The wild type control strain (Figure 4-5, lane 17) and the majority of the deletion strains on each blot have the same pattern of RNA bands for *COX1* and *COB* probes, respectively. This pattern consists of one major band (spliced message) and a number of faint precursor bands (Figure 4-5, most lanes). Most precursor bands are RNAs that contain one or more introns. However, both the *COX1* and *COB* genes are processed from polycistronic

transcripts, and RNA species that have 5' or 3' extensions due to a processing defect are possible (Costanzo & Fox, 1990). Visual inspection remains the best method of detecting genuinely altered RNA band patterns. For example in Figure 4-5 on the *COX1* blot lane 6 the mRNA band (asterisk) is virtually absent and the precursor band at ~3.2KB is increased relative to message.

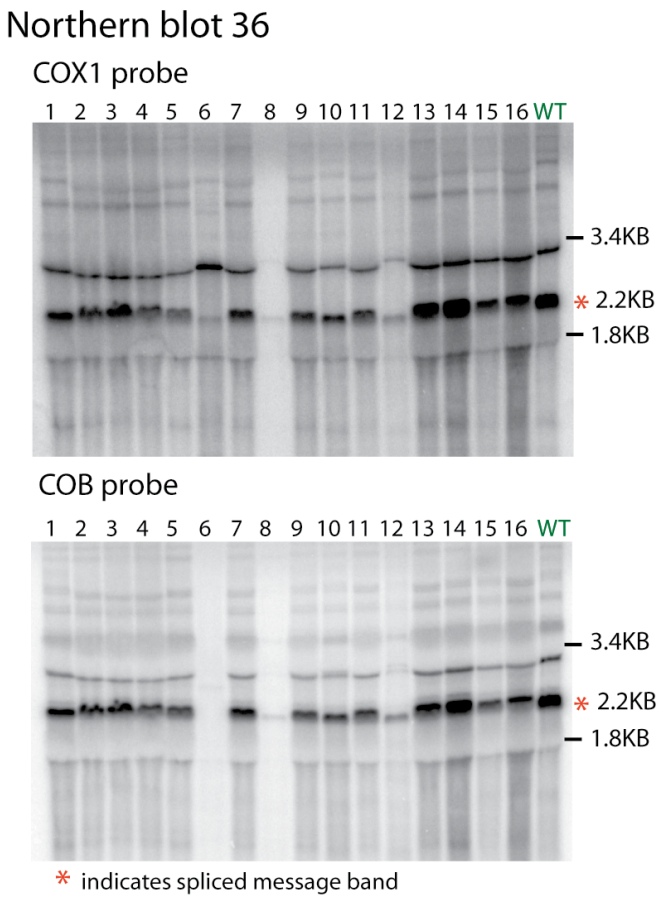


Figure 4-5. Example northern blot. Blots were probed with oligonucleotide probes complementary to the terminal exon. The major band runs at 2.2KB and is spliced message (COX1 mRNA above, COB below) The washed out regions above and below the message band are the cytosolic rRNAs (25S rRNA is 3.4KB and the18S rRNA is 1.8KB) that locally saturate the blot. Lane 17 contains the wild-type control. Lane 6 contains RNA from a strain that fails to splice COX1 and looses COB message.

The sizes of RNA bands can be estimated using the known locations of the cytoplasmic rRNA subunits (25S rRNA is 3.4KB and the18S rRNA is 1.8KB, Figure 4-5) and the *COX1*and *COB* messages (both are 2.2KB, although *COB* mRNA runs slightly slower). The rRNAs are prominent on the ethidium bromide stained gels that were routinely analyzed to assess the quality of each RNA prep. The high level of the rRNAs locally saturates the nylon membrane resulting in two washed-out regions in most northern blot lanes. The single-intron precursors of αI3α, aI5α and aI5β are expected to co-migrate with the 25S rRNA. These and any other RNAs that co-migrate with a rRNA will be detected with reduced sensitivity at the leading edge of the washed out region. The leading edge of each washed out region yields a signal on most blots that can be used as an approximate size marker (Figure 4-5). It is important to note that the leading edge of the slower 3.4KB washed out region is much darker than the leading edge of the 1.8KB washed out region (Figure 4-5). This is because a 3.2KB *COX1* precursor that runs near the 25S rRNA is present in most wild-type strains. In the wild type background, this precursor band contains intron aI4α. Several of the mitochondrial introns are similarly sized (see Table 4-1) and a band at this location may also contain pre-mRNAs with αI3α, aI5α or aI5γ. Intron specific probing is needed to resolve exactly which intron (often several) is present at that location; however aI4α and aI3α are most frequently located in the band.

A change in the splicing pattern can be quantified as a decrease in percent splicing, where percent splicing is calculated as the fraction of spliced mRNA relative to the total RNA signal in that lane using ImageQuant (Amersham). (See Chapter 2 for a detailed description including how background values were calculated individually for each band.)

This simple method of calculating percent splicing does not depend on the quality of any other lane on the blot; however, the lower the total signal in any given lane, the greater the error in the measurement. For this reason, total RNA was roughly quantified for all lanes by staining the RNA with ethidium bromide. Occasionally I found strains that had plenty of good quality RNA loaded in the lane but none of it hybridized to *COB* or *COX1* probes. One example of this is seen in Figure 4-5, lane 8. The most obvious cause of this result is loss of mtDNA. These strains were transferred to phase II of this screen and tested for the presence of mtDNA directly (see section 4-2-12; Table 4-4 for a list).

The deletion library parental strain, aBY4741, was independently prepped for most mutant batches. This enabled me to calculate the average and standard deviation of percent splicing in a single strain background. For wild-type, splicing of *COX1* averages 80% +/- 10% (n = 29) and *COB* splicing averages 86% +/- 6% (n=31). In general, the *COB* data are cleaner than the *COX1* data because the probe used for *COB* has a better signal to noise ratio. To better determine the level at which percent splicing values were significantly reduced, I decided to analyze the entire data set for variance. I calculated the percent splicing range for both *COX1* and *COB* and used a single-tailed t-test to determine the value below which splicing is likely to be significantly reduced. For *COX1,* the average value is 75% +/- 15% and percent splicing at less than 51% ($p < 0.05$) is statistically significant. For *COB* the average value is 78% +/- 14% and percent splicing at less than 55% ($p < 0.05$) is statistically significant. Figures 4-6 and 4-7 show bar graphs of these results for the *COX1* and *COB* probed blots respectively. As can be seen in the graphs, there is a wide range of splicing values among strains. This is why, although percent splicing is helpful, looking at the actual

blots for an increase in specific precursor bands or the presence of new bands is the best method for finding a splicing defect.

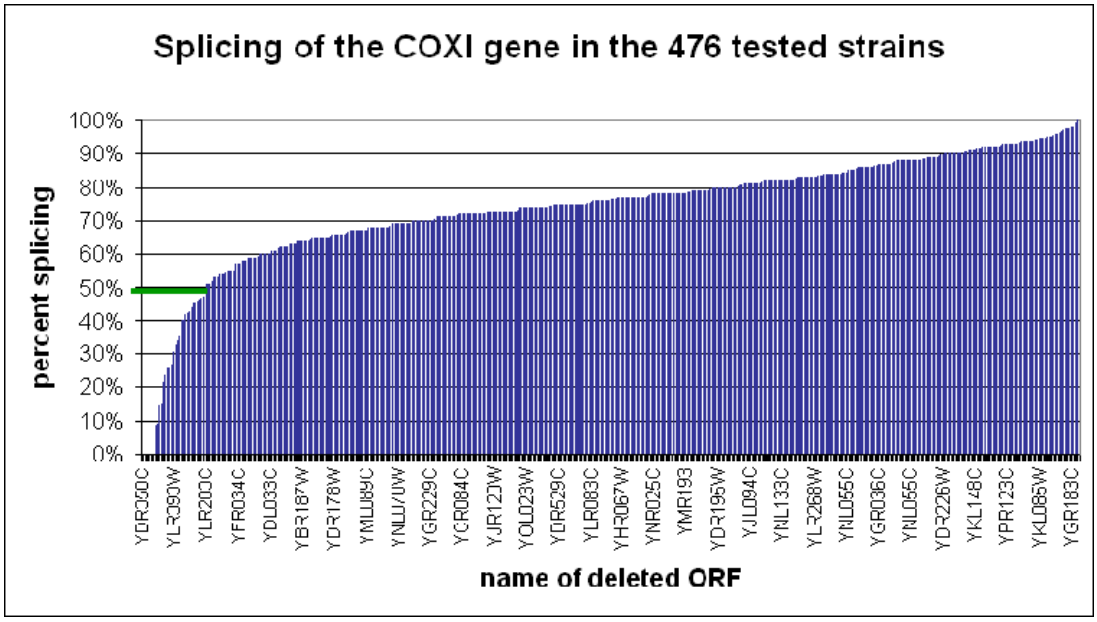


Figure 4-6. Bar graph of *COXI* percent splicing.. Average = 75%; Standard deviation = 15%, p=0.05 at 51% (one tailed t-test). Green line indicates the cutoff.

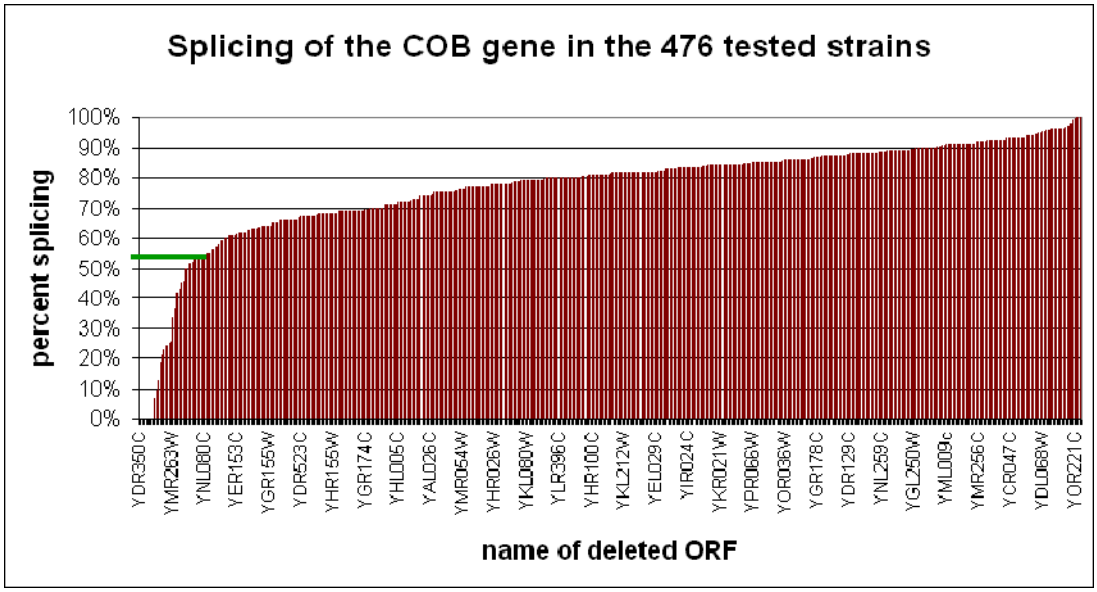


Figure 4-7. Bar graph of *COB* percent splicing.. Average = 78%; Standard deviation = 14%, p=0.05 at 55% (one tailed t-test) Green line indicates the cutoff.

Any strain assayed in this screen can be referenced in Appendix C by its systematic ORF name. Appendix C lists the blot(s) on which the data may be found, the percent splicing for each blot, and how well the strain grows on glycerol as well as other pertinent information. Because the actual appearance of each blot lane is critical for assessing the presence of a splicing defect, I have included all the blot images and tables with the data from each blot in Appendix D.

A **Intron information** | **B** **Published intron splicing factors**

	Intron	Group	intron (KB)	precursor (KB)	stable intron	maturase utilized	MSS116	NAM2	MRS1	PET54	MSS18	CBP2
COXI	**aI1**	II	2.5	4.7	yes	self	*X*					
	aI2	II	2.5	4.7	yes	self	*X*					
	aI3α	I	1.5	3.7	no	endo only	*x*					
	aI4α	I	1.0	3.2	no	bI4	*x*	*X*				
	aI5α	I	1.4	3.7	yes	endo only	*x*					
	aI5β	I	1.6	3.8	no	none	*x*		X	X	X	
	aI5γ	II	0.9	3.1	yes	none	*X*					
COB	**bI1**	II	0.8	3	yes	none	*X*					
	bI2	I	1.4	3.6	no	self	*x*					
	bI3	I	1.6	3.8	no	self	*x*		X			
	bI4	I	1.4	3.6	no	self / aI4α	*x*	*X*				
	bI5	I	0.7	2.9	no	none	*x*					X
	omega	I	1.14	4.4	yes	endo only	*x*					*x*

C

Observed deletion phenotype of published splicing factors							
	COXI	petite	petite	wt / wrong	15%	50%	55%
	COB				75%	80%	<10%
	Glycerol	**-**	**-**	**+**	**-**	**+/-**	**-**

Table 4-1. Individual intron descriptions and phenotypes of known splicing factors. (**A**) mitochondrial introns. Introns are listed in the order they are found in their respective genes. The spliced intron size is given in KB. The size of a single intron precursor RNA is given in KB. The stability of the spliced intron in the wild-type background is indicated. (**B**) Published intron splicing factors are listed along with the intron they assist to splice (large X indicates a strong published effect on splicing, a small x indicates a mild splicing phenotype). (**C**) Observed phenotype from this screen for published splicing factors. Petite strains have lost mtDNA and could not be analyzed by northern blot. A (+) indicates wild-type glycerol growth; (+/-) a partial glycerol growth defect and (-) absolutely no growth on glycerol. Note that the MRS1 deletion strain appears to be incorrect.

By carefully inspecting and quantifying the percent splicing of each blot lane, I derived a list of genes that appear to alter RNA splicing or processing. Three of the known splicing factors gave the expected result (Table 4-1) and an additional 23 strains had an altered pattern of RNA bands (Table 4-2). Appendix F lists all genes of potential interest, a brief description of gene function and the location on each blot that strain can be observed. Most candidate strains have a clear change in RNA pattern. The RNA pattern of deletion strains can deviate from the control pattern in several different ways. The simplest phenotype to interpret is when the level of mature mRNA is reduced and there is an elevated level of a single pre-mRNA band. Good examples of this pattern change are *Δpet54* (section 4-2-4c, figure 4-11), *Δcbs2, Δmne1, Δimp1, Δpet111* (section 4-2-6, figure 4-15) and *Δpif1* (section 4-2-9, figure 4-20). At the other extreme, the pattern of bands does not change, but the relative levels of all or most of the bands is altered. Such a phenotype would be expected of a general splicing factor that reduces but does not block the splicing of a number of introns. Examples of strains with a general increase in RNA precursors are *Δaat1*(section 4-2-7, figure 4-17)*, Δatp10* (section 4-2-8, figure 4-18) and *Δrim9* (section 4-2-8a, figure 4-19).

To test further for a possible splicing defect, strains were probed separately for each of the 13 mitochondrial introns. The expected wild-type band pattern varies because, even in the wild-type background, some intron-containing precursors accumulate and not every intron is stable. Generally, group II intron RNAs are stable and group I introns are not. The wild-type band pattern for each intron is shown in figure 4-8. All four group II introns (aI1, aI2, aI5γ and bI1) have a strong band that runs at the correct size for spliced intron RNA. In addition, the group I introns aI5α and ω have a fainter band that runs at the correct size for

intron RNA. In the wild-type background, substantial precursor RNA accumulates for introns aI3α, aI4α, bI5 and ω. Faint, larger RNA bands are also observable for a number of introns. The goal in examining the intron-specific blots is to determine which intron or introns is contained in the deletion strain RNA bands that differed from wild-type.

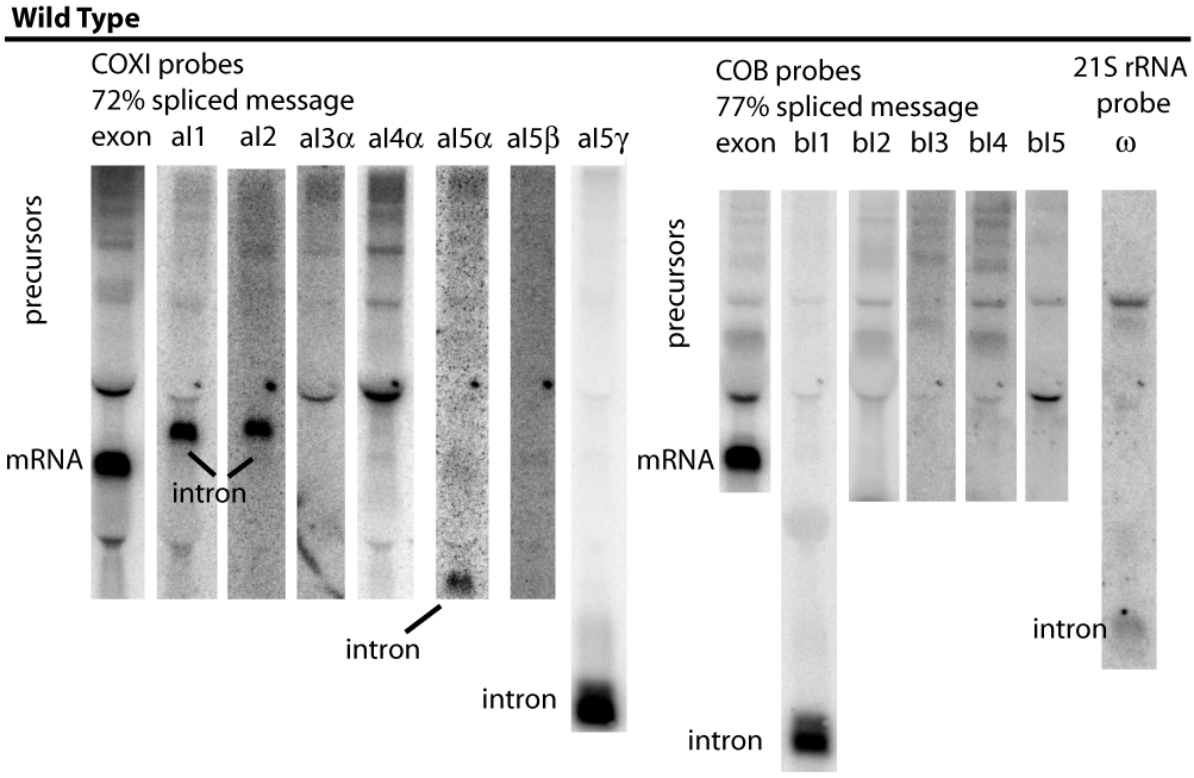


Figure 4-8. Wild-type RNA probed for message and each of the 13 mitochondiral introns.. The value for percent splicing of *COX1* and *COB* above the gel slices was derived from the shown blot. Each gel slice is labled with the intron for which it was probed.

Strains with convincing phenotypes and individual intron data

	A	B	C	D	E	F	G
	Gene		**Gly**	**COXI**	**COB**	introns affected/ possible cause	**Section**
1	***AAT1***	YKL106W	l+/-	70%	20%	aI4α, bI2, bI3, bI4, bI5	4-2-7
2	***ATP10***	YLR393W	-/+	30%	25%	ALL; aI5α,	4-2-8
3	***CBS1***	YDL069C	-	30%	65%	bI2,bI3,bI4 -- translation defect	4-2-5b
4	*CEM1*	YER061C	-/+	20%	45%	bI2,bI3,bI4 -- translation defect	4-2-5b
5	*MIR1*	YJR077C	-	75%	65%	bI2,bI3,bI5- mild translation defect	
6	***CBP2***	YHL038C	-	25%	15%	bI5, possibly aI5β	4-2-4b
7	***CBS2***	YDR197W	-	<10%	<10%	aI5β, *COB* translation block	4-2-5a
8	*IMP1*	YMR150C	-	25%	90%	aI5β	4-2-6a
9	***MNE1***	YOR350C	-/+	40%	80%	aI5β	4-2-6b
10	***PET111***	YMR257C	-	55%	90%	aI5β	4-2-6a
11	***NIP100***	YPL174C	+/-	75%	90%	ω	4-2-10
12	***PET127***	YOR017W	+	35%	20%	general mito RNA stability	4-2-3b
13	***PIF1***	YML061C	t^s	45%	20%	bI2, aI3α	4-2-9
14	***(MRP4)***	YHL005C	c^s	5%	25%	no intron defect	4-2-10
15	***MRPL39***	YML009c	c^s	40%	50%	no intron defect	4-2-10
16	***OYE2***	YHR179W	c^s	10%	50%	no intron defect	4-2-10
Strains with a convincing phenotype without individual intron data							
17	*CBP1*	YJL209W	-	10%	n/a	Complete block of single COXI intron (aI4α?), no COB RNA.	4-2-5c
18	*CBT1*	YKL208W	-	70%	25%	novel band like CBS2 aI5β band	4-2-5d
19	*COQ3*	YOL096C	-	80%	85%	possible, faint aI5β band	4-2-6a
20	*COQ4*	YDR204W	-	75%	75%	possible, faint aI5β band	4-2-6a
21	*COX9*	YDL067C	-	50%	85%	possible, faint aI5β & strong aI5γ band	4-2-6a
22	*PET117*	YER058W	-	90%	70%	possible, faint aI5β band	4-2-6a
23	*RIM9*	YMR063W	-	35%	40%	definite effect, many precursors	4-2-8a

Table 4-2. Strains with an altered RNA band pattern.
Column **A** lists the gene name – strains shown in bold were confirmed by PCR; column **B** the systematic ORF name (useful for looking up these strains in the appendices). Column **C** lists the glycerol growth phenotype where (+) indicates wild-type glycerol growth; (l+/-) a published, but not observed, growth defect; (+/-) a moderate growth defect; (-/+) a strong growth defect; (t^s) a growth defect at 37°C; (c^s) a growth defect at 20°C and (-) absolutely no growth on glycerol. Column **D** lists COX1 percent splicing; Column **E** lists COB percent splicing; column **F** lists the effected introns and a possible cause for the observations; column **G** lists the section in which the results are presented in more detail.

Strains shown with a green or blue background were probed individually for each of the 13 introns and the introns listed in column F reflect those findings. Strains shown with a yellow background were identified later and have not been repeated and probed for intron defects. In the latter strains, educated guesses as to which intron is altered are based on band sizes and similar patterns in other strains.

RNA preps and Northern blots were repeated for most potential splicing candidates to confirm the splicing phenotype. The small scale of the RNA preps did not allow for repeat runs of the same RNA, so each gel lane was derived from a separate RNA prep. (see details in appendix C where data on each individual prep/gel lane (~600) is listed). The deletion genotype was confirmed by PCR for all the strains listed in Table 4-2 in bold. Primers that flank the ORF were used so that both deletion and wild-type alleles would amplify. PCR products were digested with HindIII which only cuts the deletion allele. (For details see: http://www-sequence.stanford.edu/group/yeast_deletion_project/verification.html.) The strains in Table 4-2 shown in blue or green were individually probed for each of the 13 mitochondrial introns to determine which, if any, introns were affected. This also allowed me to eliminate a number of strains from my initial list of candidate genes because they did not hybridize with probes for any of the 13 introns, or the initial splicing defect was not reproducible. These strains are listed in Appendix F, but otherwise are not discussed further.

Several potentially interesting genes were not screened for individual intron defects. They are listed with a yellow background in Table 4-2 . Strain Δcbp1 was in the last set of strains that were screened for splicing defects after the individual intron probing was completed. Strains *Δcoq3*, *Δcoq4*, *Δpet117* and *Δcox9* all contain a faint novel band that I initially disregarded; however, individual intron probing of other strains with that band revealed that it contains aI5β (see section 4-2-6). Strain *Δrim9* was simply overlooked (see section 4-2-8a). With the exception of *Δpet117,* these latter strains have not been repeated but should be analyzed and included in future experiments. A data summary, the blot location and freezer location of every strain tested by northern blot in this screen is listed in Appendix

C by its systematic ORF name. The actual gel images can be found in Appendix D along with a repeat of the information in Appendix C organized by blot location. This allows the reader to look up a gene of interest and/or peruse the blots and then look up lanes of interest. Blots 1-3 are of a poor quality and most strains found on them were repeated on later blots, but they are included for the sake of completeness.

4-2-3b A deletion strain that generally stabilizes RNA: PET127

Sometimes a pattern change results from blocking a process other than splicing. *PET127* is a good example of this situation. It codes for a protein involved in degrading mtRNAs. RNAs that are efficiently degraded in wild-type strains are stabilized by *Δpet127* mutation (Wegierski *et al.*, 1998; Wiesenberger & Fox, 1997). This is reflected as a net reduction in apparent splicing – the *Δpet127* background has 27% spliced *COX1* and 24% spliced *COB* message. The wild-type sample from the same blot has 72% spliced *COX1* and 77% spliced *COB* message. The reduction in percent splicing for *Δpet127* is probably due to the stabilization of non-message RNA and not a reduction in net splicing. The *Δpet127* strain has many precursors which contain many introns (Figure 4-9). Because I use the steady-state levels of RNA species and calculate splicing as message RNA over total RNA, any strain that increases the steady state level of non-message species will mathematically result in reduced percent splicing. The *Δpet127* result may indicate that in the wild-type background, intron containing precursors are degraded more efficiently than message. (This same observation was made regarding the *MRS1-L232F* and *Δdbr1* strains and reported in Chapter 3). The

Δpet127 strain was included along with splicing factor candidates as a general positive control and size marker for normally present RNA bands.

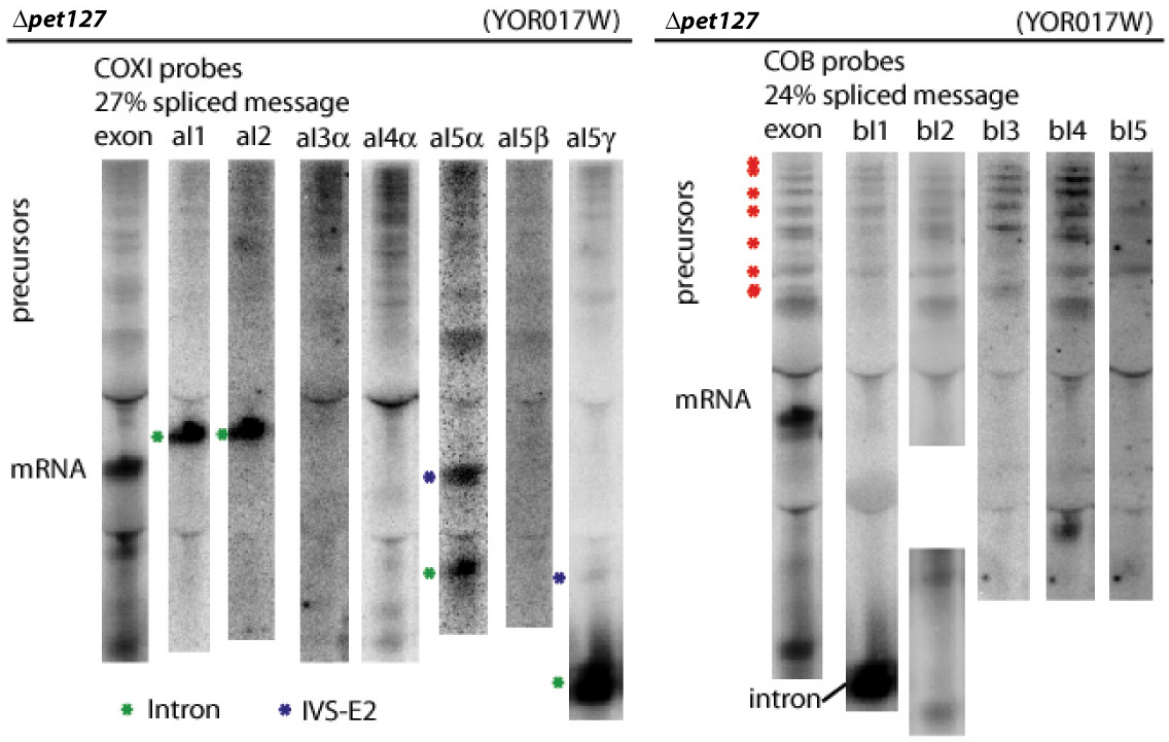


Figure 4-9. *Δpet127* RNA probed for message and each of the 13 mitochondrial introns.. Strain *Δpet127* is a useful size marker because it has many RNA bands including intron RNA (green dots), intron splicing intermediates (blue dots) and precursor RNAs (red dots). Each gel slice is labled with the intron for which it was probed.

4-2-4 Analysis of strains deleted for a previously reported splicing factor

Most known mitochondrial intron splicing factors were included in this screen. Table 4-1 lists the known yeast splicing factors (*MSS116, NAM2, MRS1, PET54, MSS18* and *CBP2)* and the introns they assist. Listed at the bottom of the table are my findings for each strain. The isolates of Δ*mss116* and Δ*nam2* lost mtDNA and were included in the petite sub-

library for Phase II. It was surprising that the Δ*mss116* lost mtDNA, and I was able to later restore wild-type mtDNA to this strain, however not in time to analyze it by northern blot. I tested Δ*mrs1*, and this isolate has no glycerol growth defect and splices as well as wild-type. The published strain (see supplement to Steinmetz *et al.* (2002) online at http://www-deletion.stanford.edu/YDPM/YDPM_index.html), has a glycerol growth defect, though my isolate does not. I am suspicious that this strain is wrong in the deletion library or was contaminated by a neighboring strain.

4-2-4a CBP2

My data confirm that *CBP2* is needed for splicing bI5 (Figure 4-10 panel A, lane bI5). The Δ*cbp2* strain cannot grow on glycerol and has very little *COB* message (2% and 24% from two separate preps). The *COB* gel lane contains two distinct precursor RNA bands (Figure 4-10 panel A, exon lane: two red asterisks). Probing for each *COB* intron, I find that bI5 is present in both bands, with bI2 and bI4 also present in the larger precursor band (Figure 4-10, bI2 and bI4 lane: red asterisk). The upper band could contain all three intron RNAs (the three intron precursor is expected to be ~5.7 KB) or it may contains two co-migrating bands, one band containing bI5 and bI2 (~4.3 KB) and the other containing bI5 and bI4 (~4.3 KB).

COX1 message is reduced relative to wild-type (average of two preps is 55% splicing). This reduction in percent splicing is primarily caused by the appearance of a single ~3.8KB precursor band that runs with precursor bands found in strains Δ*mss118,* Δ*pet54,* Δ*mne1* and Δ*pet127*. For these latter four strains, the band hybridized to aI5β. Unfortunately,

insufficient RNA was loading in the Δ*cbp2* lane to expect a clear aI5β signal, though a faint signal is visible (Figure 4-7, panel B, over exposed lane: red asterix). Strains with aI5β splicing defects will be discussed in greater detail later (section 4-2-6). Neither of the two preps and blots for strain Δ*cbp2* (Appendix D; blot 4, lane 5 and blot c1, lane 5) gave strong signals, but these data agree with expectations and reveal no indications of novel phenotypes in this strain background.

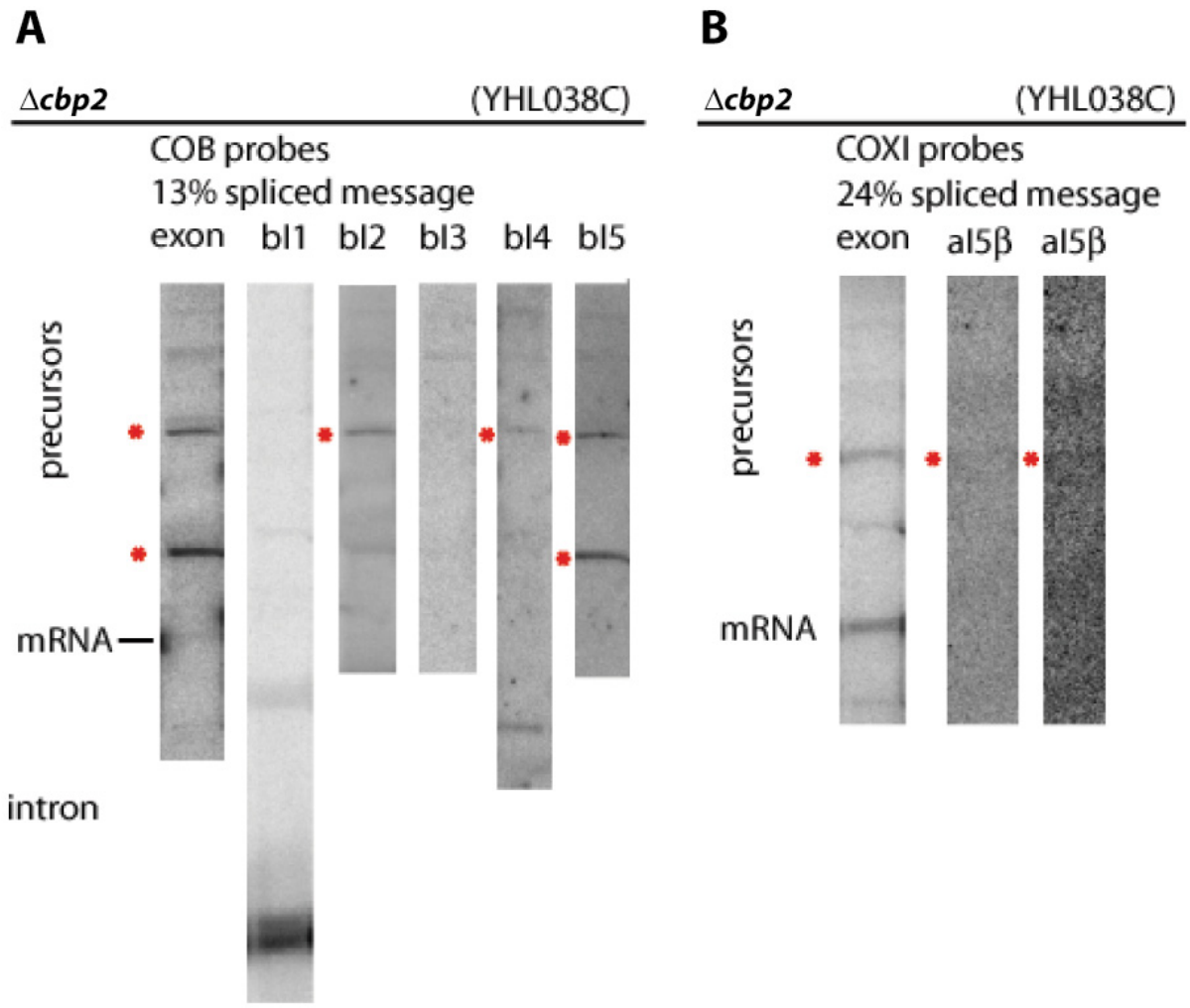


Figure 4-10. *Δcbp2* RNA probed for message introns. (A) COB probes. As expected, the *Δcbp2* strain fails to splice bI5 (B) COX1 probes. *Δcbp2* may have an aI5β containing precursor band.

4-2-4b MSS18

The Δ*mss18* strain has a mild glycerol growth defect, and *COB* splicing that is comparable to wild type (~80%). As expected, *COX1* splicing is reduced to ~50% (wild type is 75% +/- 15%) and an ~3.8KB precursor RNA, containing aI5β, accumulates (Figure 4-11, *Δmss18* lanes). This band is almost certainly *COX1* mRNA with unspliced aI5β, which is predicted to be 3.775KB. Although strain *Δmss18* has the correct phenotype and band pattern, PCR confirmation of that strain gave the wild-type band. Of the 19 tested strains, this was the only one that turned out to have a wild-type copy of the gene. The northern blot data would strongly indicate that the cells used for RNA preps were correct.

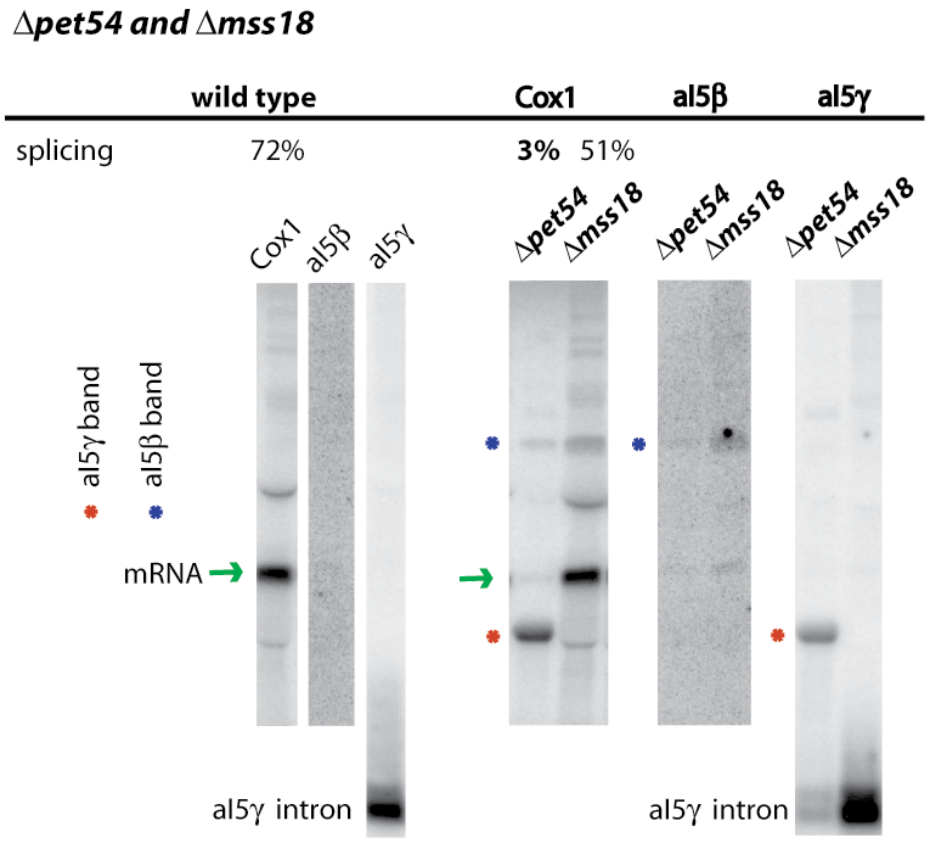


Figure 4-11. *Δpet54* and Δmss18 RNA probed for COX1 exon, aI5β and aI5γ. Both have the expected aI5β band (blue asterisk). Strain *Δpet54* also has an aI5γ containing band (red asterisk)

4-2-4c PET54 and COX9

Pet54p has been shown to be involved in the splicing of intron aI5β (Valencik *et al.*, 1989); see Introduction). I found that strain Δ*pet54* cannot grow on glycerol and is strongly inhibited for *COX1* splicing, with only ~15% of mature mRNA. Consistent with earlier reports, it accumulates an aI5β precursor RNA (figure 4-11, blue asterisk). The surprising finding was that the major transcript is not mature mRNA but is a novel species somewhat smaller than the mRNA. The full set of intron probes showed that the novel band does not contain aI5β but instead contains aI5γ (Figure 4-11). From my other studies of aI5γ splicing, I know where the splicing intermediate of that intron (IVS-E2) runs on this gel system: notably, the observed band is clearly larger than the aI5γ-E6 intermediate. The novel band probably contains one or more upstream exons in addition to aI5γ and exon 6. It could not contain all *COX1* exons, however, because then it would have been larger than the *COX1* mRNA. Probing for individual exons should test that hypothesis and reveal which are still present. Another possibility is that an upstream intron (perhaps aI5β) mis-splices, joining one or more upstream exons to a site within aI5γ, thus blocking the splicing of the remnant of aI5γ. There is precedent for this notion from published studies of an aI3 mutant (Guo *et al.*, 1995). If there is no effect on splicing when aI5γ is the only intron, then mis-splicing is likely. This novel species can be cloned as a cDNA and analyzed by sequencing.

Strain *Δcox9* cannot grow on glycerol medium and has a mild *COX1* splicing defect (50%) but, interestingly, has the same aI5γ-containing band as *Δpet54*. It also has the ~2KB band, containing aI5β, that runs faster than the 2.2KB *COX1* mRNA (Figure 4-16, panel D)

observed in strains *Δimp1* and *Δpet111* (Figure 4-15). Unfortunately, *Δcox9* was not included in the further analysis of candidate genes and it has not been confirmed that those bands actually include said introns. Both RNAs are distinctive and are unlikely to arise for other reasons. Given on-going interest in the splicing of both aI5β and aI5γ, *Δcox9* should be investigated further. *COX9* encodes subunit VIIa of cytochrome c oxidase (Wright *et al.*, 1986). Cox9p is one of the nine proteins that are present in the final, active form of cytochrome c oxidase (Poyton & McEwen, 1996). It is not the catalytic subunit and may play a role in either assembly or the structural support of the holoenzyme. In the absence of Cox9p, levels of other subunits of cytochrome c oxidase are depressed. There is no apparent reason for Cox9p to be involved in splicing, especially splicing of non-maturase requiring introns.

The finding that strain *Δpet54* (and possibly strain *Δcox9*) appears to have a strong defect in aI5γ splicing is potentially interesting because currently Mss116p is the only protein that plays a direct role in its splicing.

4-2-5 Deletion strains that alter mitochondrial translation.

Maturases are intron-encoded, usually intron-specific, splicing factors. An important caveat to note when looking for yeast mitochondrial splicing factors is that deletion of any gene that disrupts the translation of *COX1* or *COB* mRNA should result in reduced splicing of maturase-requiring introns. (Table 4-1 lists which introns require maturases.) Strains that only inhibit the splicing of maturase-requiring introns probably have a primary translation

problem. This matter is complicated somewhat because most maturases cannot be translated until the upstream introns have been spliced; therefore, a primary block in the splicing of an upstream intron is expected to create a secondary block in the splicing of all downstream maturase-requiring introns. It is more difficult to sort this out for cases where the upstream intron is only partially inhibited. One possible example of this effect is *Δpif1*, which appears to have a primary bI2 defect and secondary bI3, bI4 and aI4α defects (see section 4-2-9). Such cases can be resolved by analyzing derivatives of these knockouts that have fewer introns in their mtDNAs. Here I present the results from all strains with splicing defects that probably result from a primary translation defect.

4-2-5a CBS2

A good example of a strain that blocks the splicing of *COB* introns as a result of a primary translation block is the *cbs2* deletion strain. *CBS2* is a *COB*-specific translation factor. The *Δcbs2* strain contains no *COB* mRNA and is totally blocked for the splicing of maturase-requiring bI2, bI3 and bI4 but not that of maturase independent bI1 and bI5 (Figure 4-12, panel A). All three unspliced introns accumulate predominantly as a single, very large, RNA precursor. A slightly larger faint precursor containing unspliced bI5 can also be seen (Figure 4-12, panel A). This level of bI5 precursor in the *Δcbs2* background is comparable to basal level of bI5 precursor seen in the wild type background (Figure 4-8, panel B), so I conclude that deleting *CBS2* has no effect on bI5 splicing. Intron bI1 has no apparent splicing block. As expected, when *COB* translation is inhibited, *Δcbs2* also blocks aI4α splicing because intron aI4α splicing depends on the bI4 maturase (Figure 4-12, panel B).

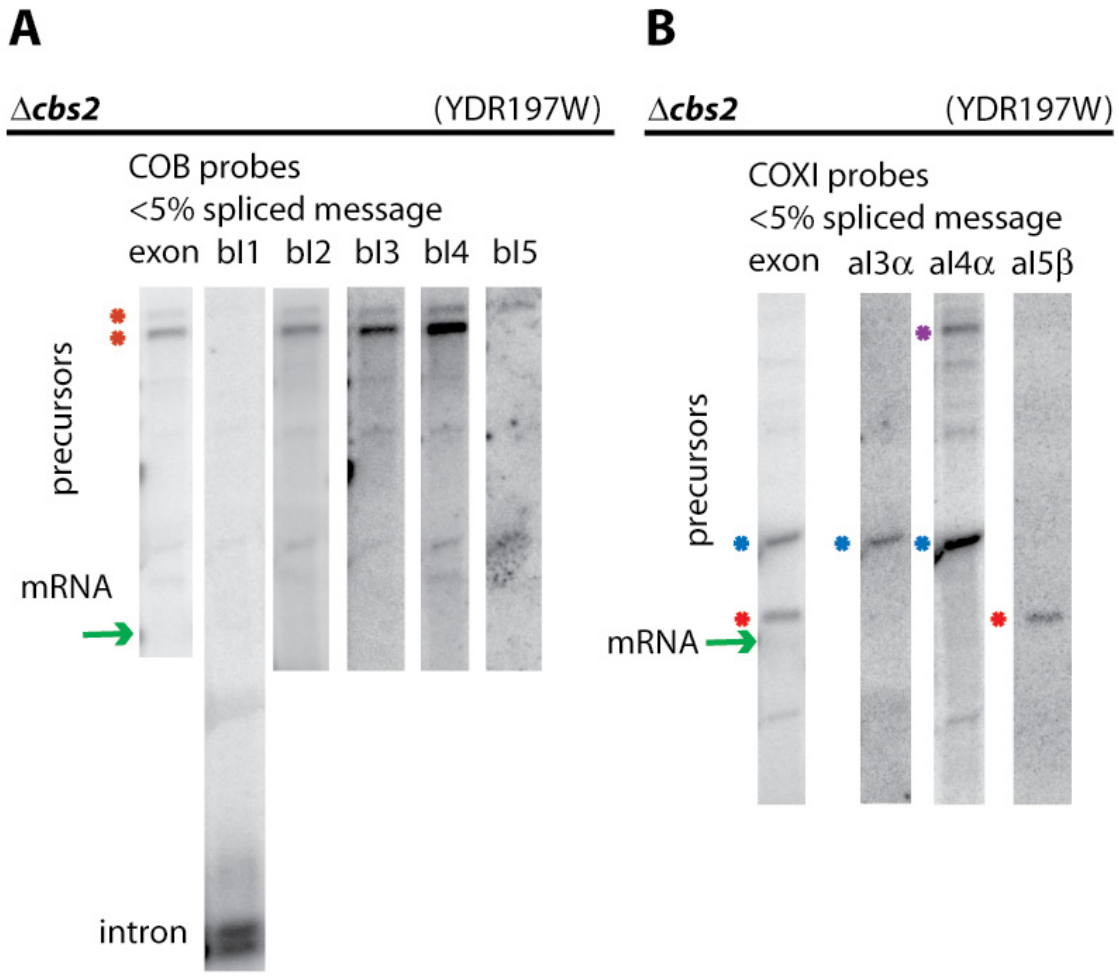


Figure 4-12. *Δcbs2* has a *COB* translation defect and accumulates aI4α, aI3α and aI5β precursor RNA. (A) *COB* probes. (B) *COX1* probes. A purple asterisk is adjacent to large aI4α band, a blue asterisk adjacent to the single intron aI3α and aI4α precursor bands and a red asterisk is adjacent to the novel aI5β containing RNA band.

However, one study of a *cbs2* deletion lacking bI2 and bI3 found that some correctly processed (27kd) bI4 maturase was still generated; however, the majority of the protein was in the unprocessed form (55kd) (Muroff & Tzagoloff, 1990Note: CBP7 is synonymous with CBS2). In the absence of all introns no *COB* protein was made in the *cbs2* strain. The authors conclude from this that *CBS2* may also play a direct role in splicing but is primarily involved in translation. I also find that *CBS2* may play a direct role in splicing because the *Δcbs2*

strain in this study has an aI5β splicing block (Figure 4-12, panel B, red asterisk). The aI5β block cannot be explained as a secondary effect of a translation problem (intron aI5β does not require a maturase) and will be addressed in section 4-2-6 along with other strains that have aI5β splicing defects. In addition aI3α and aI4α precursors accumulate in the *Δcbs2* strain (Figure 4-12, panel B, blue and purple asterisk).

4-2-5b CBS1 and CEM1

The *CBS1* gene also affects the translation of *COB* mRNA, and as anticipated, the deletion shows reduced splicing of introns bI2, bI3 and bI4. Unlike *Δcbs2,* some splicing of all three maturase-requiring introns still occurs leading to a complex pattern of precursor bands (Figure 4-13, panel A). The translation block in the *Δcbs1* strain is probably not as strong as is the block in the *Δcbs2* strain, and that probably accounts for the reduced level of splicing inhibition. It is interesting to note that the *Δcem1* strain inhibits the splicing of the same introns and has almost the identical pattern of precursor bands as does the *Δcbs1*strain (Figure 4-13, compare panels A and B). The only noticeable difference is that *Δcem1* accumulates what appears to be an aI4α intermediate band (Figure 4-13, panel B, lower blue asterisk). *CEM1* is homologous to beta-ketyl-acyl synthetase genes and is not known to be involved in *COB* translation (Harrington, 2003). Despite this, given the similar result to *Δcbs1,* I view the splicing defect in the *Δcem1* strain as a *COB* translation defect.

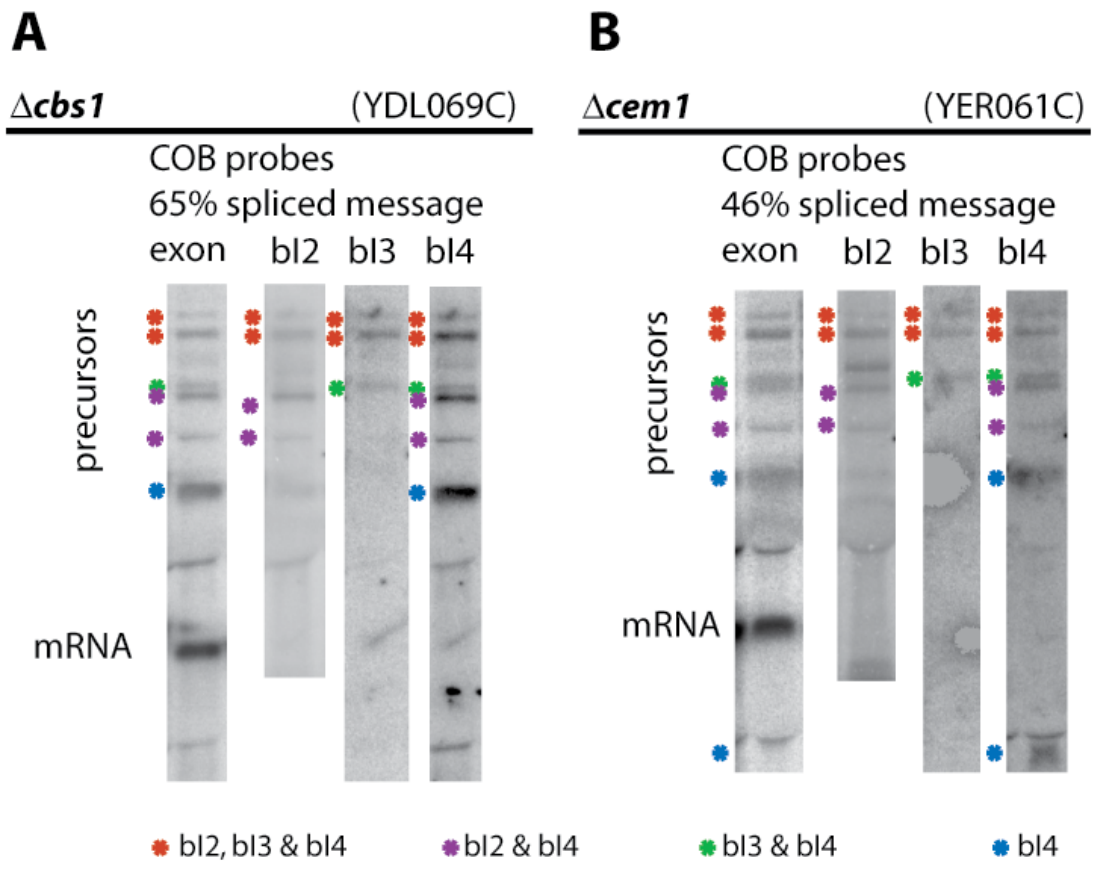


Figure 4-13. *Δcbs1* and *Δcem1* have a *COB* translation defect (A) *Δcbs1 COB* blot. (B) *Δcem1 COB* blot.

4-2-5c CBP1

The *CBP1* gene has been extensively studied and is known to interact with the 5' end of the *COB* transcript and to be essential for the accumulation of detectable *COB* transcript (Dieckmann *et al.*, 1982; Ellis *et al.*, 2005; Mittelmeier & Dieckmann, 1993). My isolate of that disruption confirms that expectation (Figure 4-14, panel A: no *COB* bands). Previous research has shown that the gene is not needed for bI1 splicing and that it is unlikely to play a

direct role in splicing any *COB* introns (Mittelmeier & Dieckmann, 1993). Strain *Δcbp1* contains *COX1* RNA and ~10% of the RNA is spliced message (Figure 4-14; panel A). The bulk of the unspliced RNA is present as a precursor that contains just one intron, probably aI4α, due to the absence of bI4 maturase synthesis. Unfortunately this strain was not probed for individual introns, so that interpretation requires confirmation.

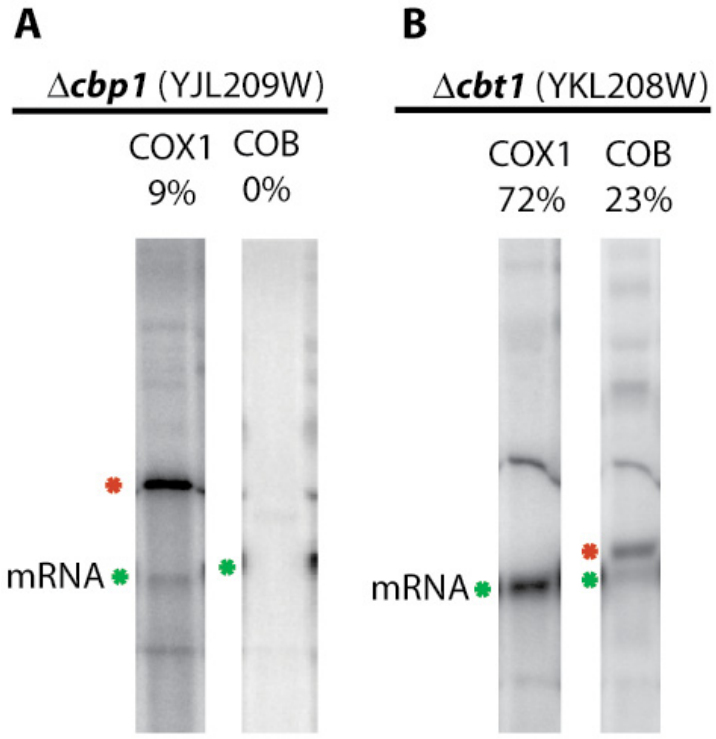


Figure 4-14. *Δcbp1* and *Δcbt1* exon probed blots. Green asterisk is shown next to spliced mRNA; red asterisk is shown next to precursor RNA. (A) *Δcbp1*. (B) *Δcbt1*.

4-2-5d CBT1

Strain *Δcbt1* contains a novel *COB* RNA band slightly larger than mature *COB* mRNA (Figure 4-14; panel B – red asterisk). Cbp1p is known to be involved in the 3' end processing of *COB* mRNA and the deletion of *CBT1* results in mRNAs that are ~200bp

longer than wild-type (Rieger *et al.*, 1997). This is consistent with my observation; the novel band is probably unprocessed message. Assuming that this band is in fact message, *COB* is still only ~23% spliced. The decrease in message level is most likely because the extra 200 base pairs destabilizes the spliced message. *COX1* splicing appears to be unaffected by the *CBT1* deletion.

4-2-6 Up to twelve genes may be involved in the splicing of intron aI5β.

Deletions of three genes known to be involved in aI5β splicing were included in this library screen (*PET54*, *MSS18*, and *MRS1*). This deletion of *MRS1* does not have a splicing phenotype; it is possible that the strain is incorrect or contaminated, or that this gene may yield a different phenotype in different strain backgrounds. Strains *Δ*pet*54* and *Δ*mss*18* have splicing defects, and intron blots revealed that they are inhibited for aI5β splicing (Figure 4-11, aI5β lanes).

At least ten strains were identified that are likely to have a splicing defect involving intron aI5β. Six of these strains (*Δ*pet*54*, *Δ*mss*18* and four new strains– *Δcbs2, Δimp1, Δmne1* and *Δpet111*), contain bands that hybridize to the intron aI5β probe. Four strains (*Δcoq3, Δcoq4, Δcox9* and *Δpet117*) contain bands that are in the same location as bands that hybridize to aI5β but have not been directly shown to have an aI5β splicing defect (see Table 4-2 for a summary). Strain *Δatp10* also has an aI5β splicing block, along with a general splicing block, and is discussed in section 4-2-8.

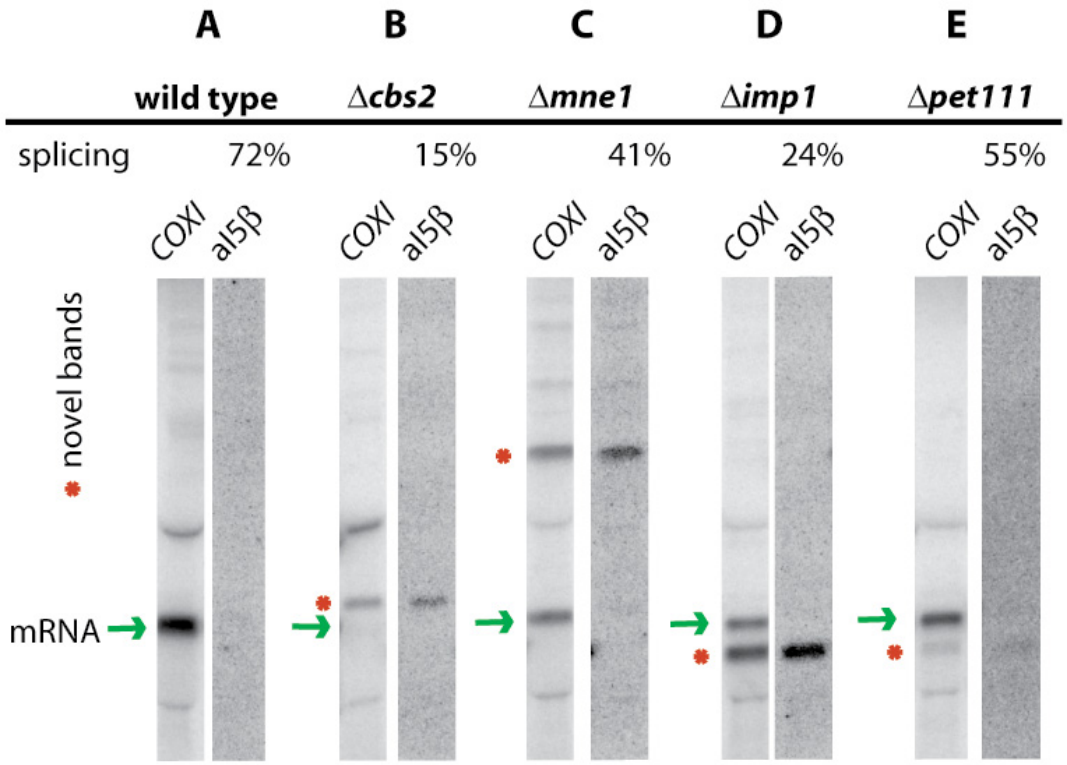


Figure 4-15. Deletion strains with novel RNA bands that hybridize to aI5β. A green arrow is shown next to spliced mRNA; red asterisk is shown next to novel RNA bands. (A) wild type RNA (B) *Δcbs2*. (C) *Δmne1*. (D) *Δimp1*. (E) *Δpet111*.

4-2-6a IMP1, PET111, COX9, PET117, COQ3 and COQ4

I observed a novel band that is slightly smaller than the *COX1* message in six strains. Of those strains, only *Δimp1* and *Δpet111* were included in the intron specific blots and for both of those strains the band contains aI5β (Figure 4-15; panels D and E; red asterisk). The band is about 2 KB long and is clearly larger than excised aI5β RNA (1.57 KB). A band of this size may result from partial splicing and so could be the IVS-exon RNA containing aI5β (1.5KB) and the last two exons (~0.5KB) for a total of ~2 KB. The remaining four strains

that also contain a similar band – *Δcox9*, *Δpet117*, *Δcoq3* and *Δcoq4* (Figure 4-16, red asterisk) were not initially included because the band is faint. With the exception of *Δcox9*, none of these strains has a significant reduction in *COX1* splicing. The band is most prominent in strain *Δimp1*. Strain *Δimp1* only contains ~25% spliced *COX1* mRNA and the darkest band is this novel aI5β–containing band. In the remaining five strains this band is <10% of the total RNA signal. Clearly, future work should focus on *IMP1*.

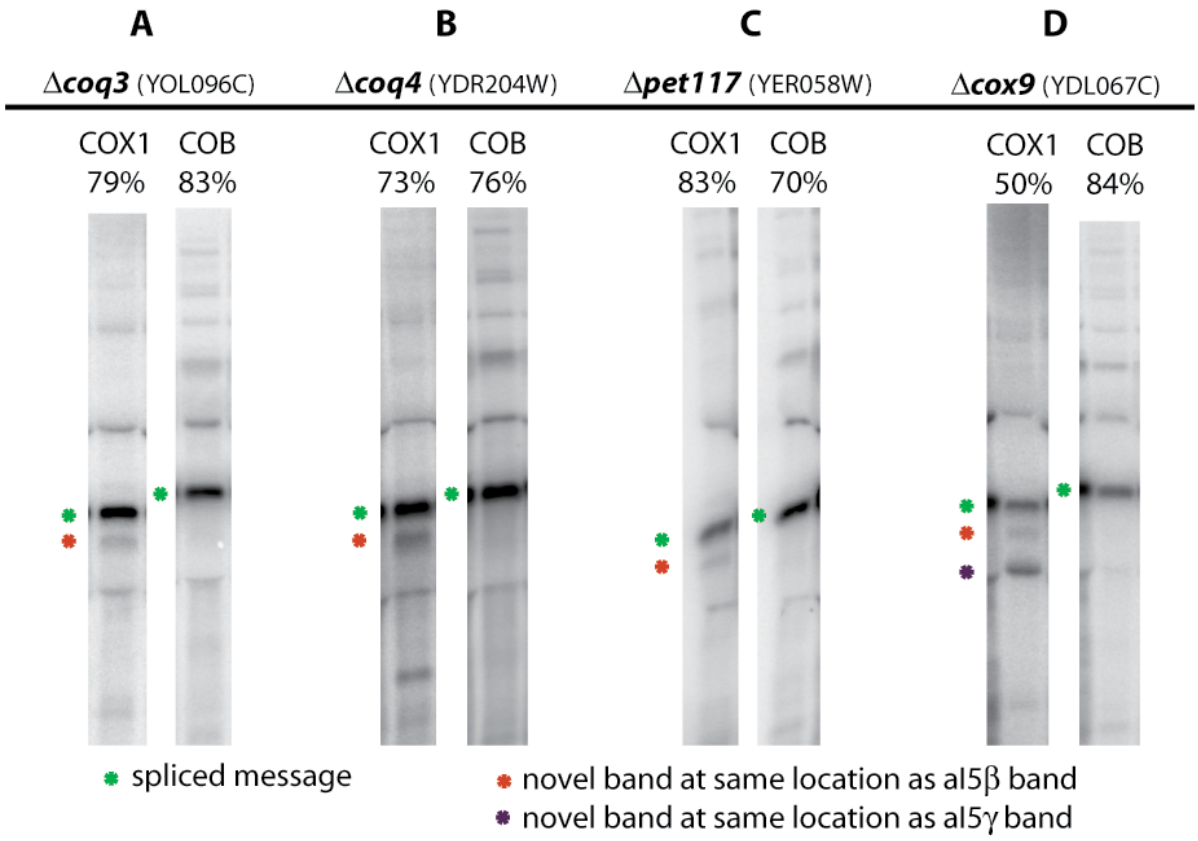


Figure 4-16. Deletion strains with novel RNA bands that may contain aI5β. A green arrow is shown next to spliced mRNA; red asterisk is shown next to novel RNA bands. (A) *Δcoq3.* (B) *Δcoq4*. (C) *Δpet117*. (D)*cox9*.

4-2-6b MNE1, PET54, MSS18 and CBP2

Strains *Δmne1*, *Δpet54*, *Δmss18* (and possibly *Δcbp2* see section 4-2-5) each contain a ~4 KB precursor band in the same location, and each band hybridizes to intron aI5β (Figure 4-15, panel C, red asterisk; Figure 4-11, *Δpet54* and *Δmss18* – aI5β lanes and Figure 4-10, *Δcbp2* – panel B). The expected size of the mature *COX1* mRNA plus intron aI5β is 2.2 KB plus 1.57 KB or ~3.8KB, which is consistent with the observed size of this precursor band.

4-2-6c CBS2

Strain *Δcbs2* has a clear aI5β-containing band at a novel location. It was also discussed in section 4-2-5a because it has a strong *COB* translation defect (Figure 4-12). The band that hybridizes to aI5β is larger than mature *COX1* message (Figure 4-15, panel B, red asterisk) and is probably around 3 KB. There is no obvious single explanation for a band of this size. Because it hybridizes to aI5β and the terminal exon of *COX1*, it should theoretically contain intron aI5β (1.5KB) and the last two exons of *COX1* (~0.5KB), but these do not add up to the observed 3 KB. Some other sequence must also be included in this band. No other intron hybridizes to this band. This band only occurs in the *Δcbs2* background. As was proposed for the *PET54* deletion, this may also be the result of an upstream mis-splice.

These findings bring the possible total of genes involved in the splicing of intron aI5β to twelve: *Δ*pet*54*, *Δmss18, Δmrs1, Δcbs2, Δcbp2, Δimp1, Δmne1, Δpet111, Δcoq3, Δcoq4, Δcox9* and *Δpet117***.** It is likely that a number of these proteins act indirectly to alter aI5β

splicing. Even so, aI5β is distinctive in that it appears to require more proteins for its splicing than any other intron. Intron aI5β may require additional help because it contains a large, untranslatable ORF. This ORF was probably once translated in the usual way but has acquired many mutations and small deletions that render it unquestionably inactive. This significant extra RNA sequence, relative to the ORF-less and Maturase-less introns, may cause folding problems that need to be overcome by various accessory proteins. It will be interesting to learn which of these proteins play direct roles in aI5β splicing and how those proteins interact or function sequentially.

4-2-7 AAT1 deletion alters the splicing of introns bI3, bI4 and aI4α.

The *AAT1* deletion strain has a very mild glycerol growth defect, and contains ~20% spliced *COB* mRNA and ~70% spliced *COX1* mRNA. The *COB* exon blot contains one striking new precursor band and a number of fainter precursors. All *COB* precursor bands hybridize to the bI3 probe (Figure 4-17, compare the *COB* exon lane to the bI3 lane). The major new precursor band hybridizes exclusively to the bI3 probe and is about 4KB in size, as expected of pre-mRNA containing only bI3. A strong bI3 splicing block is expected to reduce the level of bI4 maturase, which may explain the observed reduction in bI4 and aI4α splicing (as they both require the bI4 maturase). The block to bI3 splicing in *Δaat1* is not complete enough to also block both bI4 and aI4α splicing, because some spliced message is observed. Most of the fainter precursor bands in the *COB* exon lane hybridize to bI4 (Figure 4-17, compare the *COB* exon lane to the bI4 lane). The increased precursor bands observed

for *COX1* almost exclusively contain aI4α (Figure 4-17, compare the *COX1* exon lane to the aI4α lane). The deletion of *AAT1* appears to reduce the splicing of introns bI2, bI5 and aI3α to a lesser extent (Figure 4-17). However it is clear that bI3 splicing is the most strongly influenced by the deletion of *AAT1*.

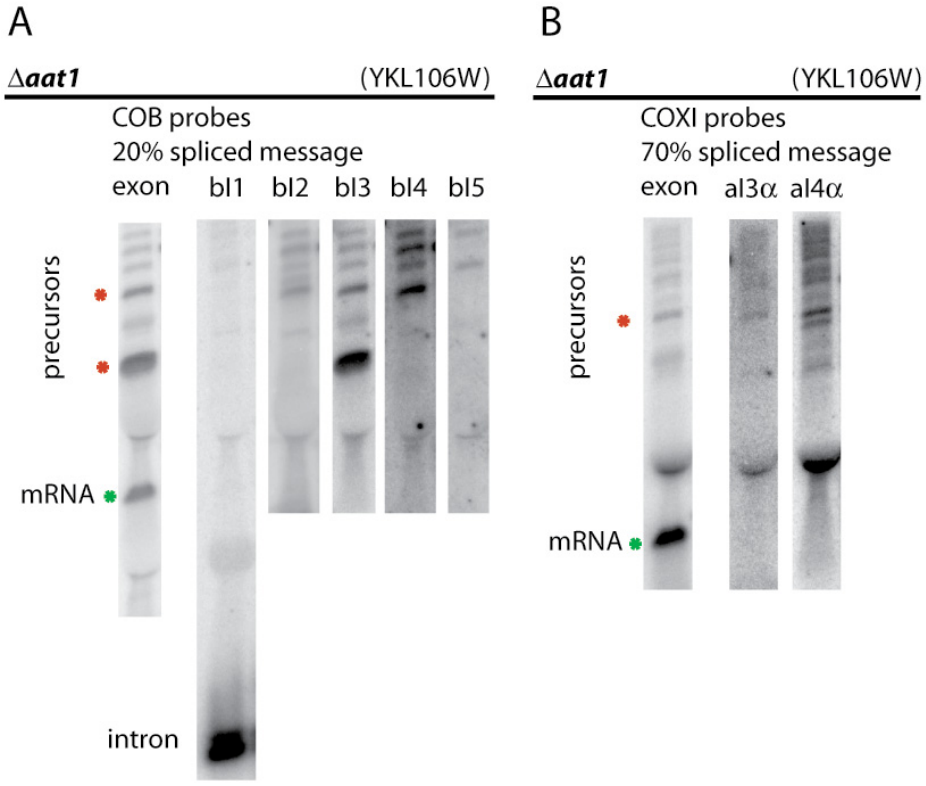


Figure 4-17. *Δaat1* has reduced bI3 splicing. A green asterisk indicates spliced mRNA; red asterisk indicates precursor RNA bands of interest. (A) *COB* lanes. (B) *COXI* lanes.

4-2-8 *ATP10* deletion alters the splicing of almost every intron.

The *Δatp10* strain can grow very slowly on glycerol, has ~30% spliced *COX1* and ~25% spliced *COB* message. Both the *COX1* and *COB* exon probed gel lanes contain many precursor bands. Hybridization with the single intron probes revealed that the splicing of almost every intron (with the possible exceptions of aI1, aI2 and omega) is reduced (Figure

4-18). Even bI1 and aI5γ, the two maturase independent group II introns that rarely have any trace of precursor RNA, show faint precursor bands. The group I introns (aI3α, aI4α, aI5α, aI5β, bI2, bI3, bI4, bI5) are the most strongly affected. In every case, the amount of intron precursor RNAs is significantly greater than that observed for the wild-type strain.

The aI5α splicing defect is striking because *Δatp10* is the only deletion strain that has a strong aI5α splicing block. This is evidenced by an accumulation of large MW aI5α precursors and a reduction in spliced intron RNA (Figure 4-18; panel B, aI5α lane): the deletion library strain is unusual in that aI5α intron RNA accumulates in the wild type). Introns aI1 and aI2 have a wild-type amount of precursor and excised intron RNA and so are unaffected by this mutation.

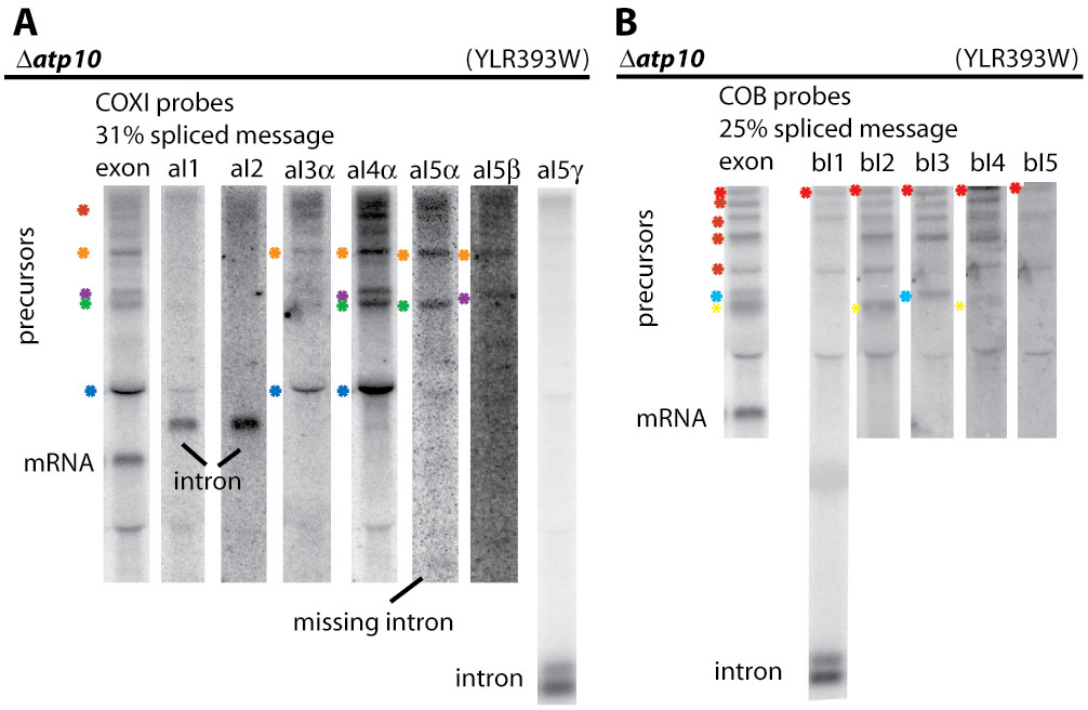


Figure 4-18. *Δatp1* accumulates many precursor RNAs. (A) *COB* probed lanes. (B) *COXI* probed lanes.

4-2-8a The RIM9 deletion accumulates many precursors.

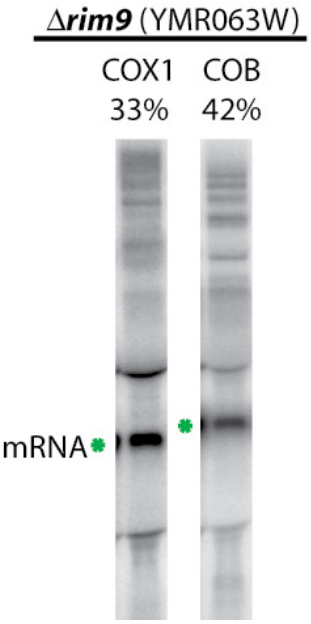


Figure 4-19. *Δrim9* also accumulates many precursor RNAs.

Strain *Δrim9* cannot grow on glycerol and appears to have a general splicing defect. Both *COX1* and *COB* have a significant reduction in percent splicing (33% and 42%, respectively from a single very good blot) and a general increase in the level of precursor RNAs (Figure 4-19). Rim9p is a putative trans-membrane protein that localizes to the mitochondria (Gavin *et al.*, 2002) and is involved in regulating the pH-sensitive cleavage of Rim101p (Li & Mitchell, 1997). Rim101p is a transcription factor responsible for gene expression under alkaline conditions (Boysen & Mitchell, 2006). Unfortunately, *Δrim9* was not included in the screen for individual intron defects.

4-2-9 *PIF1* deletion significantly reduces the splicing of intron bI2 at 37°C.

Pif1p is a mitochondrial DNA helicase that also interacts with telomeres in the nucleus. *Δpif1* is known to be temperature sensitive, although glycerol growth was not specifically examined (Foury *et al.*, 1998; Van *et al.*, 1992). I confirm the t^s phenotype of *Δpif1* and also find that glycerol growth is wild type at 30°C and severely impaired at 37°C. Visually, the reduction in glycerol growth appeared greater than the reduction in growth on dextrose at the non-permissive temperature. At 30°C, *COX1* message is spliced ~85% and

COB message is spliced ~55%. The strain can grow, albeit very slowly, at 37°C. The strain grew well enough to prepare RNA from cells grown at the restrictive temperature for at least

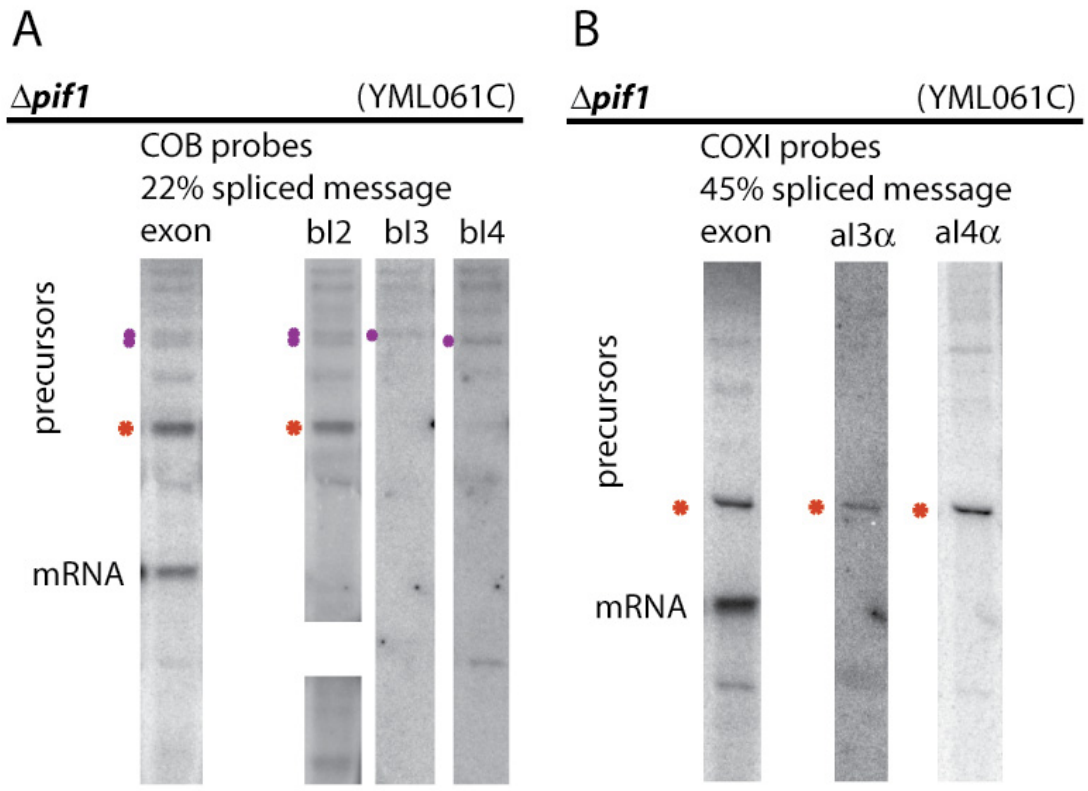


Figure 4-20. *Δpif1* has reduced bI2 splicing at 37°C. A red asterisk indicates precursor RNA bands of interest. (A) *COB* lanes. (B) *COXI* lanes.

two doublings (protocol is three). Growing the cells at the non-permissive temperature (37°C) on raffinose medium reduces *COX1* splicing to ~45% and *COB* to ~20%. The northern blot of *Δpif1* probed for *COB* exon contains a novel band of roughly 3.5-3.8KB and a general increase in slow running precursor bands. Intron specific probing of *Δpif1* revealed that the novel band and all visible precursor bands hybridize to intron bI2 (Figure 4-20, panel A; bI2 lane). This prominent precursor is indicated by a red asterisk and only bI2 and *COB* exon probe hybridize to this band. The expected size of the precursor containing bI2 only is

~3.6KB (1.4KB plus 2.2KB), which is the size of the observed band. A slower running doublet is indicated by two purple asterisks; this doublet hybridizes to *COB* exon and bI2, with the upper band additionally hybridizing to bI3 and the lower band hybridizing to bI4. The reduction in *COX1* splicing appears to be caused by an aI4α block (Figure 4-20, panel B), which is an expected secondary effect of a *COB* splicing problem. Overall, the *Δpif1* strain has what appears to be a very clean splicing block of a single intron.

4-2-10 Several deletion strains accumulate a ~3KB omega containing band.

Strain *Δnip100* does not affect the splicing of *COX1* (~95% splicing) or *COB* (~90% splicing) message. The omega intron probed lane contains a ~3KB band (Figure 4-21, panel A, red asterisk) that is also very faintly visible in the wild type lane. The *Δnip100* lane also contains excised omega intron RNA (1.1KB) and the full-length precursor RNA (4.4KB). The *Δnip100* strain was included in this screen because it has a glycerol growth defect, even though the protein has not been shown to localize to the mitochondria; moreover, it has been convincingly shown to locate to spindle poles. Nip100p is part of the dynactin complex and is required to translocate the mitotic spindle through the bud neck (Kahana *et al.*, 1998). The double *Δnip100, Δatp10* (see section 4-2-8) deletion has a synthetic slow growth phenotype (Tong *et al.*, 2004), which may serve as a link to the mitochondria. Both *Δnip100* and *Δatp10* also have a aI5α splicing defect (sections 4-2-11 and 4-2-8, respectively). Pursuit of the *Δatp10* phenotype may yield insights into the *Δnip100* phenotype and is recommended over further analysis of *Δnip100* which probably alters splicing indirectly.

Three other strains - *Δyhl005c* (partial deletion of *MRP4*), *Δmrpl39* and *Δoye2* - also contain the same novel band on the omega blot. These three strains are all cold sensitive and have reduced *COX1* and *COB* splicing at the non-permissive temperature (20°C) (Table 4-2, lines 14-16). These same three strains have other slight irregularities when hybridized with probes other than omega; however, I do not have a wild-type control for growth at 20°C on the blots probed for individual introns. The nearly identical phenotype, with multiple probes, of all three strains may be due to growth at 20°C. Any further analysis of these strains needs to begin by repeating these findings with the appropriate wild-type control.

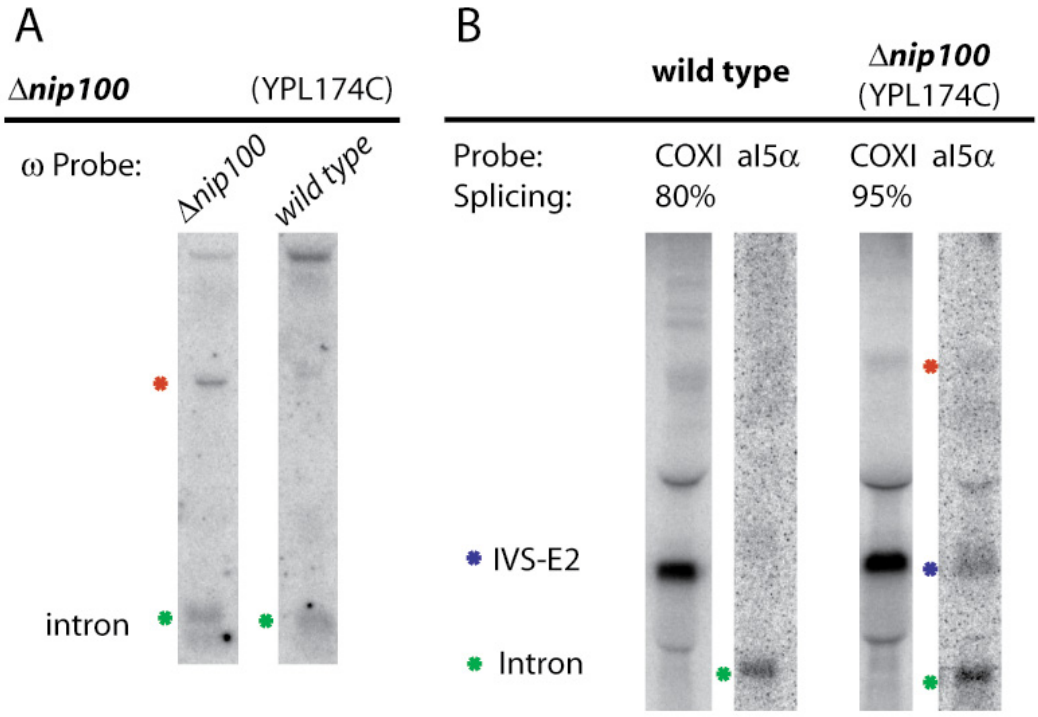


Figure 4-21. *Δnip100* accumulates an ω containing band and al5α containing bands.
(A) ω probed lanes. (B) al5α probed lanes.

4-2-11 A number of deletion strains accumulate intron aI5α intermediate.

Wild type RNA probed for aI5α contains a single band that runs at roughly 1.5KB (Figure 4-21, panel B, green asterisk). Intron aI5α is 1.36KB long, so the presence of such a band would indicate that excised aI5α RNA is relatively stable in this strain. Most mutant strains of this set probed for aI5α contain the spliced intron band (Appendix E, aI5α blot) and most lanes that lack this band are under-loaded with total RNA. As discussed in section 4-2-8, *Δatp10* appears to have an aI5α splicing block and fails to accumulate this aI5α containing band. Other strains with sufficient total RNA that lack the aI5α intron band (but do not have precursor bands like *Δatp10)* are Δ*cbs2, Δimp1, Δpet54, Δyhl005c* (partial deletion of *MRP4*) and *Δyml090w* (partial deletion of *RPM2*).

A number of aI5α probed lanes contain an additional ~2.2 KB band at roughly the same location as *COX1* message. This band cannot be *COX1* message because it is not present in strains that have a strong *COX1* mRNA band as does wild-type. This band is most likely a partially spliced aI5α intermediate containing aI5α and the downstream exons. The presence of downstream exons in this band cannot be easily confirmed because it would be masked by the presence of *COX1* message on the exon probe. A probe that straddled the intron/exon boundary could be used if conditions were first worked out under which the probe did not hybridize to *COX1* message but did to a control aI5α-Exon RNA. The most pronounced examples of strains containing this putative aI5α IVS-E2 RNA are strains *Δpet127* (Figure 4-9, lane aI5α, blue asterisk) and *Δnip100* (Figure 4-21 panel B, blue asterisk). This is expected for the *Δpet127* strain because RNAs are stabilized in the *Δpet127*

background (see section 4-2-3b). For *Δnip100,* I calculate that 95% of the *COX1* transcript is spliced message. This is an unusually high value (i.e., higher than in the wild-type control strain) and could reflect the presence of both *COX1* message and aI5α IVS-E2 signal at the same location. In appendix E on the aI5α candidate blot1 there are a number of other strains that also have this band: *Δslm3*(lanes 8 and 9)*, Δpet111* (lane 10)*, Δmss51* (lane 12) and *Δmrs2* (lane 13). In these strains, the level of the aI5α IVS-E2 band is greater than or equal to that of the aI5α spliced intron band.

4-2-12 Phase II: Initiation of screening the strains without mtDNA for splicing defects

Strains that lack mtDNA or lost their mtDNA during the course of the phase I screen were collected into a petite strain sub-library. Table 4-3 lists those strains that were transferred from phase I of the screen to the petite sub-library. They were all sub-cloned and retested for glycerol growth and the presence of mtDNA. Some of these strains have not completely lost mtDNA. For those strains that are still ρ^+, it is possible that the colony selected for northern analysis was petite, but that it was still possible to repeat the northern analysis with a ρ^+ colony. These strains are currently stored in the petite sub-library, however they should not be analyzed as part of phase II. They should be returned to phase I and be tested directly for splicing defects, with mtDNA cytoducted back in, as needed. Seven out of the eighteen strains can probably be returned to phase I after selecting for sub-clones with intact mtDNA. The remaining eleven strains need to be analyzed as part of phase II.

	A	B	C	D	E	F
	Gene name	Systematic ORF name	General protein function	sub-clone phenotype	strain location petite grid	strain location library grid
1	*FZO1*	YBR179C	m DNA maint./repair	gly$^-$ / ρ^-	3.2.F	3.5.C
2	*MIP1*	YOR330C	m DNA maint./repair	gly$^-$ / ρ^-	3.3.C	2.5.G
3	*MSH1*	YHR120W	m DNA maint./repair	**gly$^+$ / ρ^+**	3.3.E	5.3.C
4	*UGO1*	YDR470C	m fusion	gly$^-$ / ρ^-	3.6.A	19.4.F
5	*RMD9*	YGL107C	m meiosis	**gly$^+$ / ρ^+**	3.5.H	16.2.B
6	*ATP11*	YNL315C	m respir./fermt.	gly$^-$ **/ ρ^+**	3.4.D	15.2.D
7	*ATP22*	YDR350C	m respir./fermt.	gly$^-$ / ρ^-	3.4.G	4.2.D
8	*PPA2*	YMR267W	m respir./fermt.	**gly$^+$ / ρ^+**	3.5.D	2.2.D
9	*QCR7*	YDR529C	m respir./fermt.	gly$^-$ / ρ^-	3.5.E	16.1.H
10	*MRPL13*	YKR006C	m ribosome	gly$^-$ / ρ^-	1.2.E	17.2.E
11	*MRPL4*	YLR439W	m ribosome	gly$^-$ / ρ^-	1.4.D	15.3.D
12	*MRPL9*	YGR220C	m ribosome	gly$^-$ / ρ^-	1.5.C	9.3.E
13	*SRF1 (MDM34)*	YGL218W	m stress resistance (mDNA maint./repair)	**gly$^+$ / ρ^+**	3.3.G	17.4.E
14	*HAP2*	YGL237C	non-mito known	gly$^-$ **/ ρ^+**	4.2.D	8.4.B
15	*HAP3*	YBL021C	non-mito known	gly$^-$ **/ ρ^+**	4.2.E	9.4.F
16	*VPS45*	YGL095C	non-mito known	gly$^-$ / ρ^-	4.3.F	17.2.D
17	*RPB4*	YJL140W	non-mito RNA	gly$^-$ / ρ^-	4.2.A	19.2.D

Table 4-3. Strains transferred to the petite sub-library because they lost mtDNA in phase I of the screen. (A) Common gene name. (B) Systematic ORF name. (C) Protein function. An "m" means mitochondrial. (D) Growth (gly+ = any growth on glycerol, gly- = no glycerol growth) and presence of mtDNA (ρ- = no mtDNA and ρ+ = wt mtDNA). Although the subclone screened in phase I of the screen lost mtDNA, for some strains, a new sub-clone still retained mtDNA. (E) "petite grid" is the new strain location in the 96 well plates for phase II of the screen. (F) "library grid" is the strain location in the 96-well plates that contain all 866 potential mitochondrial gene deletion strains.
The strain in lane 12 is a deletion of *SRF1* and a partial deletion of *MDM34*.

4-2-12a Some ρ^- strains are able to maintain intronless mtDNA.

To determine if any of these strains are able to maintain mtDNA, but had spontaneously lost their mtDNA in the systematic deletion library, I attempted to transfer an intact mitochondrial genome into each petite strain. I crossed each ρ^- strain to αMCC109/I^o, a *kar*1 strain that has ρ^+ mtDNA that lacks all 13 introns. Because that tester strain has the

*kar*1 mutation, most zygotes to fail to undergo nuclear fusion. Consequently, this cross can yield haploid progeny with the nuclear genome of the petite library strain and an intact mtDNA from the tester strain (see Chapter 2 for more details about these *kar*1 x *KAR*1 crosses). An important feature of this cross is that the auxotrophic markers in these strains overlap such that occasional diploids formed cannot be selected for; thus, a negative result (i.e., no cells with mtDNA and the library strain nuclear genome) could result from failure to mate – a false negative result. Each cross was plated for single colonies, and colonies with the aBG4741 background were picked and assayed for glycerol growth plated on medium lacking uracil to determine if they were diploids. The colonies were also tested for the presence of mtDNA by crossing to *a*W303ρ^0. That tester strain is wild-type and will complement the deleted gene in the library strain, and the resulting diploid will grow on glycerol medium provided that the library strain acquired ρ^+ mtDNA in the initial cross.

I originally included 37 strains from the deletion library that did not grow on glycerol medium, even though they did not have a published glycerol growth defect. Thirty-two of those strains proved able to maintain the intact mtDNA that was transferred into them by this cross and they were then removed from the petite sub library. They do not need to be returned to the phase I library and assayed for splicing defects because there is no evidence that they have a mitochondrial location or influence glycerol growth. The remaining five resulted in an inconclusive cross and were retained in the petite sub-library

Ten strains acquired ρ^+ mtDNA from the αMCC109/I^o cross but remained gly^-; those strains are listed in Table 4-4 and in Appendix G. An additional 23 strains with known mitochondrial function or glycerol growth defect were able to maintain I^o mtDNA and the

presence of I^o mtDNA restored at least some glycerol growth (Table 4-5 and Appendix G). These strains need to be crossed to an intron-containing strain to determine if it is the absence of introns *per se* that restored glycerol growth. Any strain that cannot maintain a mitochondrial genome with introns but is able to maintain an intronless genome should be further tested with single intron genomes. For example, failure to splice the group I intron omega is expected (but has never been shown) to result in mtDNA loss because a very low level of mitochondrial translation for most any reason causes loss of mtDNA.

	A	B	C	D	E	F		
	Systematic ORF name	Gene name	General protein function	re-tested phenotype	Result of I^o cross	petite sub-grid plate	col	row
1	YNL184C	*(MRPL19)*	m ribosome	gly^- / ρ^-	gly^- / ρ^+	2	1	H
2	YOR199W	*(PET56)*	m transc./ othr RNA	gly^- / ρ^-	gly^- / ρ^+	2	4	E
3	YOR200W	*(PET56)*	m transc./ othr RNA	gly^- / ρ^-	gly^- / ρ^+	2	4	F
4	YDL044C	*NAM1*	m transc./ othr RNA	gly^- / ρ^-	gly^- / ρ^+	2	4	H
5	YMR064W	*AEP1*	m translation	gly^- / ρ^-	gly^- / ρ^+	2	5	E
6	YLR067C	*PET309*	m translation	gly^- / ρ^-	gly^- / ρ^+	2	6	C
7	YMR089C	*YTA12*	m protein fold/proc.	gly^- / ρ^-	gly^- / ρ^+	3	2	C
8	YPL059W	*GRX5*	m respir./fermt.	gly^- / ρ^-	gly^- / ρ^+	3	4	H
9	YLR091W	*YLR091W*	m respir./fermt.	gly^- / ρ^-	gly^- / ρ^+	3	5	F
10	YGR102C	*YGR102C*	m unknown fx	gly^- / ρ^-	gly^- / ρ^+	3	6	H

Table 4-4. Strains that can maintain (I^o) mtDNA but remain gly^-. (A) Common gene name. (B) Systematic ORF name. (C) Protein function. An "m" means mitochondrial. (D) Growth (gly+ = any growth on glycerol, gly- = no glycerol growth) and presence of mtDNA (ρ- = no mtDNA and ρ+ = wt mtDNA). (E) result of cross. (F) "petite sub-grid" is the new strain location in the 96 well plates for phase II of the screen.

	A	B	C	D	E	F		
	Systematic ORF name	Gene name	General protein function	re-tested phenotype	Result of I^o cross	petite sub-grid plate	col	row
1	YBR122C	*MRPL36*	m ribosome	gly^- / ρ^-	gly^+	1	3	H
2	YDR042C	*(RSM10)*	m ribosome	gly^- / ρ^-	gly^+	2	1	C
3	YLR270W	*DCS1*	m transc./ othr RNA	gly^- / ρ^-	gly^+	2	5	A
4	YER154W	*OXA1*	m protein fold/proc.	gly^- / ρ^-	gly^+	3	1	E
5	YDL104C	*QRI7*	m protein fold/proc.	gly^- / ρ^-	gly^+	3	2	A
6	YLR369W	*SSQ1*	m protein fold/proc.	gly^- / ρ^-	gly^+	3	2	B
7	YMR072W	*ABF2*	m DNA maint./repair	gly^- / ρ^-	gly^+	3	2	D
8	YOR241W	*MET7*	m DNA maint./repair	gly^- / ρ^-	gly^+	3	2	G
9	YLL006W	*MMM1*	m DNA maint./repair	gly^- / ρ^-	gly^+	3	3	D
10	YDL198C	*YHM1*	m DNA maint./repair	gly^- / ρ^-	gly^+	3	3	H
11	YMR083W	*ADH3*	m respir./fermt.	gly^- / ρ^-	gly^+	3	4	A
12	YDR298C	*ATP5*	m respir./fermt.	gly^- / ρ^-	gly^+	3	4	C
13	YGL240W	*DOC1*	m unknown fx	gly^- / ρ^-	gly^+	3	6	C
14	YDR507C	*GIN4*	non-mito known	gly^- / ρ^-	gly^+	4	2	B
15	YER122C	*GLO3*	non-mito known	gly^- / ρ^-	gly^+	4	2	C
16	YER110C	*KAP123*	non-mito known	gly^- / ρ^-	gly^+	4	2	F
17	YLR260W	*LCB5*	non-mito known	gly^- / ρ^-	gly^+	4	2	G
18	YOL100W	*PKH2*	non-mito known	gly^- / ρ^-	gly^+	4	3	A
19	YER070W	*RNR1*	non-mito known	gly^- / ρ^-	gly^+	4	3	C
20	YGR180C	*RNR4*	non-mito known	gly^- / ρ^-	gly^+	4	3	D
21	YAL048C	*GON1*	non-mito unknown	gly^- / ρ^-	gly^+	4	3	G
22	YDL167C	*NRP1*	non-mito unknown	gly^- / ρ^-	gly^+	4	4	A
23	YNL170W	*YNL170W*	non-mito unknown	gly^- / ρ^-	gly^+	4	4	B

Table 4-5. Strains for which the intronless (I^o) mtDNA restores some glycerol growth. (A) Common gene name. (B) Systematic ORF name. (C) Protein function. An "m" means mitochondrial. (D) Growth (gly+ = any growth on glycerol, gly- = no glycerol growth) and presence of mtDNA (ρ- = no mtDNA and ρ+ = wt mtDNA). (E) result of cross. (F) "petite sub-grid" is the new strain location in the 96 well plates for phase II of the screen.

4-2-12b Organization of the petite sub-library by gene function.

The final petite sub-library contains 163 strains and is stored in two 96-well plates organized by gene function. Appendix G contains the complete strain list including gene functions and the grid location of each strain in this sub-library. As described in the previous section, fifty-three strains located in the petite sub-library were not analyzed in phase I of this screen although they can maintain mtDNA, at least briefly. Many of these strains are highly interesting: included are deletions of 8 mitochondrial ribosomal genes, 6 mitochondrial genes involved in transcription or RNA binding and 2 genes involved in mitochondrial translation. There are an additional 25 genes coding for known mitochondrial proteins and 12 genes coding for non-mitochondrial proteins. With the possible exception of the latter 12 deletion strains, this set of strains has a high potential for yielding splicing related proteins.

The 110 unavoidably ρ^o strains in the petite sub-library include 48 mitochondrial ribosomal proteins, 13 mitochondrial tRNA synthetases, 7 mitochondrial proteins involved in transcription or RNA binding, 7 proteins involved in mitochondrial translation, 6 proteins involved in mitochondrial protein folding, processing or transport, 5 mitochondrial DNA maintenance and repair proteins, 7 proteins involved in respiration or fermentation, 2 additional mitochondrial proteins with known functions and 7 mitochondrial proteins with unknown functions. There are also 8 ρ^o knockouts of genes for non-mitochondrial proteins, three of which are involved in RNA-related processes, four have other known functions and one protein remains uncharacterized (Figure 4-22).

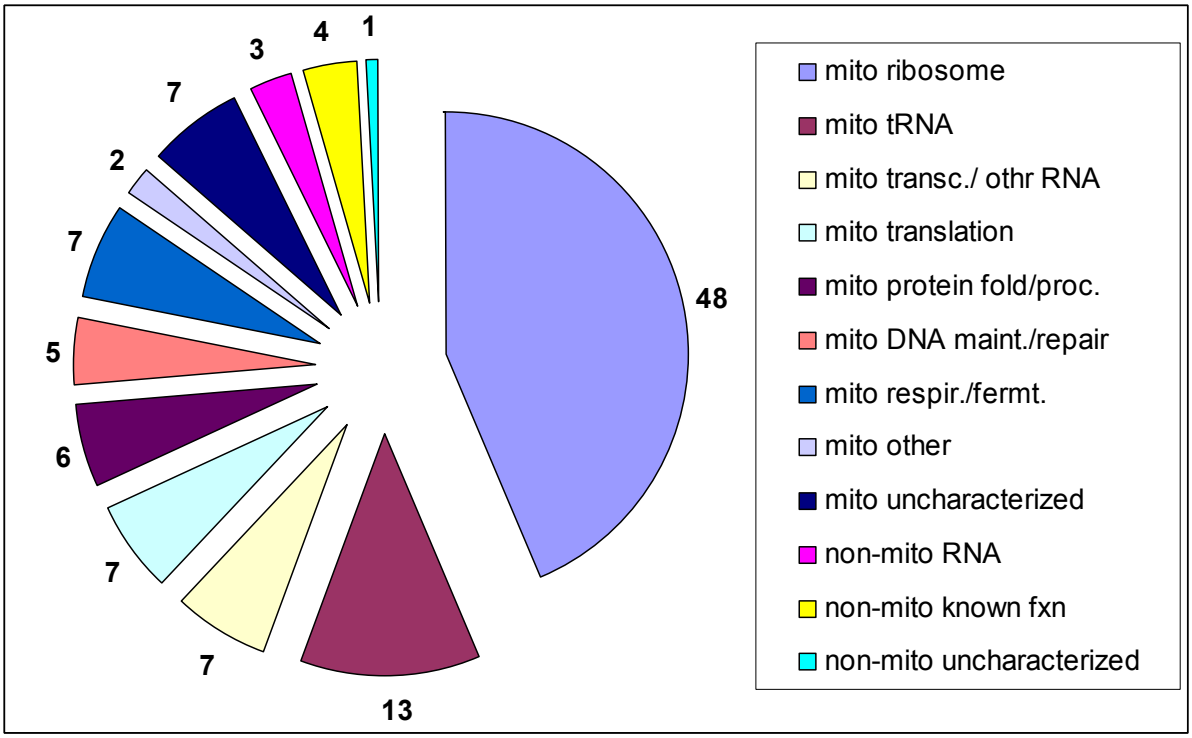


Figure 4-22.Petite strains by protein function. 110 petite strains that did not reacquire mtDNA grouped by protein function. 102 of the 110 are mitochondrial proteins.

4-2-12c Strategy for screening ρ^- strains for splicing defects.

All together fifty-three strains in the petite sub-library can maintain mtDNA and should be analyzed directly by northern blot. This leaves 110 strains that cannot maintain mtDNA (Figure 4-22). Deletion strains that are unable to maintain mtDNA can be screened for mitochondrial intron splicing using specially constructed petite genomes (Fox et al., 1988; Johnson C.H. & McEwen J.E., 1997). Introns omega, aI3α, aI5β, aI5γ, bI1 and bI5 have been shown to splice in the absence of mitochondrial translation (Guo *et al.*, 1995; Huang *et al.*, 2005; Johnson C.H. & McEwen J.E., 1997). A *COX1* derived tester gene will be constructed in a standard plasmid background that contains three of the aforementioned

introns (Figure 4-23, pPetSplice with aI3α, aI5γ and bI1). This petite consists of the *COXI* gene, with bI1 and its flanking exons inserted into a unique site in intron aI4α.

This plasmid will be placed into the mitochondria of strain αMCC109/ρ^o using biolistic transformation to create strain αMCC109/pPetSplice. MATa strains from the petite sub-library will be crossed to αMCC109/pPetSplice and the pPetSplice petite genome transferred into the library background. Colony hybridization will be used to identify *MATa* colonies that contain the petite genome. The resulting derivatives of the library strains that contain this petite genome will then be screen by standard northern blots to determine if the splicing of any of the three introns is inhibited. In addition, omega splicing defects could be tested in a similar manor by using the existing petite F11 used by Dr. Huang in our lab (Huang *et al.*, 2005).

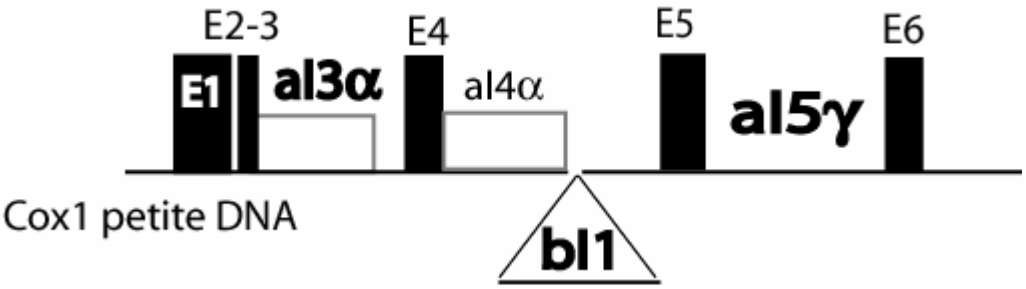


Figure 4-23. pPetSplice: Proposed pMIT to screen petite strains for splicing defects. This is essentially the GII-5γ *COXI* allele with bI1 inserted into intron aI4α.

4-2-13 Not all mitochondrial ribosomal proteins cause mtDNA loss.

It is well established that many mutations that disrupt mitochondrial translation result in the loss of mitochondrial DNA (reviewed byContamine & Picard, 2000). The current work is probably the first time that the systematic deletion library has been screened for the

presence of mitochondrial DNA. In the course of this screen I observed that strains lacking certain mitochondrial ribosomal proteins are actually able to maintain mtDNA. I was not able to test every ribosomal gene - eight of the seventy-seven genes coding for known mitochondrial ribosomal proteins are missing from this strain set. Five of the eight are essential genes so that viable haploid null mutants cannot be obtained, the remaining three knockouts were not available in the deletion library (Δ*mrp4,* Δ*rsm26* and Δ*mrpl19*). Of the remaining sixty-nine strains I found that forty-five lack mtDNA and did not re-acquire mtDNA in the test cross noted above. Surprisingly, twenty-four of those null mutants can maintain an intact mitochondrial genome (Table 4-6). Of those twenty-four, one strain *(Δmrpl39*) had reduced glycerol growth at 20°C. Appendix H lists my findings for the seventy-seven mitochondrial ribosomal proteins, their E. coli and human homologs and splicing data if available. This surprising finding appears to indicate that many ribosomal proteins, including seven proteins with homologs in E coli and mammals (*MRP02, MRPS17, MRPS28, MRPL36, MRPL01, MRPL24, MRPL39*), are not absolutely essential for translation.

77 yeast mitochondrial ribosomal genes
45 are ρ^-
24 can maintain ρ^+ mtDNA
5 are essential genes
3 are unavailable

Table 4-6. mtDNA content of the 77 yeast mitochondrial ribosomal proteins.

4-3 DISCUSSION

To increase our knowledge of protein cofactors involved in mitochondrial intron splicing, I undertook to screen 808 strains deleted for known or suspected yeast mitochondrial proteins for intron splicing defects.

4-3-1 Phase I: Screen of strains that have mtDNA for intron splicing defects.

Of the 808 strains, 476 were directly screened by northern blot analysis because they retain mtDNA and have a mild to severe growth impairment on glycerol. I isolated 14 strains with novel, intron-related, RNA processing defects and at least an additional 10 strains with probable splicing defects (Figure 4-24).

I did not screen 209 strains, although they were deleted for mitochondrial genes and contained intact mtDNA, because did not show any reduction in the ability to grow on glycerol as compared to wild-type. It is possible for a strain to have no glycerol growth defect and still have a mild splicing defect. Given the productive results from the first phase of this screen, a possible phase III would be to screen these remaining 209 strains.

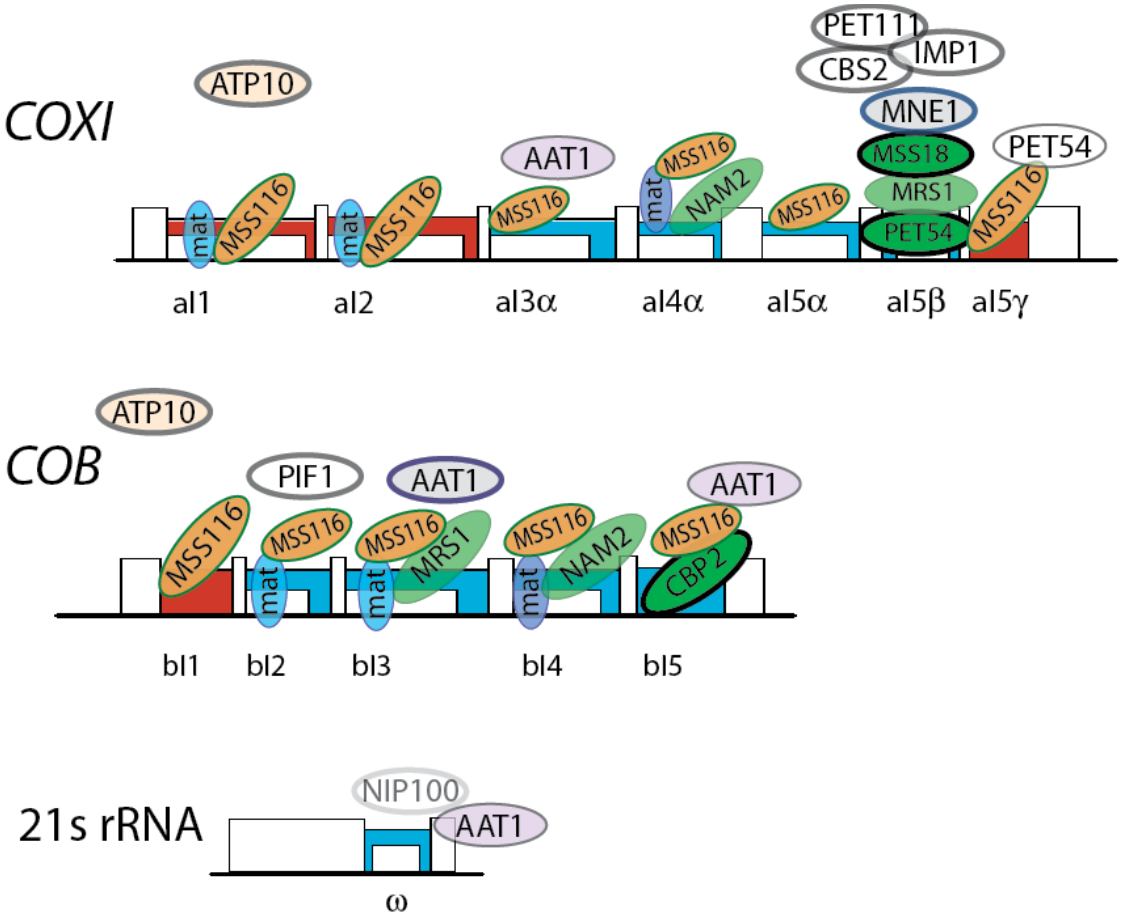


Figure 4-24. Summary of findings including known splicing factors. Published splicing factors include: Mss18, Mss116, Mrs1, Pet54 and Cbp2 and intron encoded maturases (mat). Several potential new mitochondrial intron splicing factors were identified including: Atp10, which may be a general splicing factor; Aat1;Mne1; Pif1; Pet111; Cbs2 and Imp1.

4-3-2 Comparison of my results to those of a similar screen by Udo Schmidt.

Since the 1980s, no new systematic screens for yeast mitochondrial splicing factors have been published. However, at the same time as I was working on the present screen, Udo Schmidt (a former member of our lab) was also screening the yeast deletion library for genes specifically involved in the splicing of the group II intron aI5γ (Luban *et al.*, 2005). In brief,

they acquired, from the European source, the same library that I used. They also initially screened the entire library of 4878 strains for the ability to grow on glycerol and isolated 355 strains that had a strong glycerol growth defect (compare to my 389 strains with a mild to severe growth defect). Unfortunately, they assumed that all 355 strains contained wild-type mitochondrial DNA and did not perform control crosses to check for the status of mtDNA. After treating the 355 strains to force mtDNA loss, they cytoducted in a new mitochondrial genome with only four group I introns (GII-0 mtDNA, containing aI3α, aI4α, bI4, bI5). For 109 out of the 355 strains, the four intron mtDNA restores glycerol growth. They concluded that the initial inability of these strains to grow on glycerol was due to the presence of 9 more introns, and that they had isolated 109 potential splicing factors. They acknowledged that only 44 of the 109 strains are deleted for proteins known to localize to the mitochondria. It is more likely that many of the original library strains had lost their mtDNA, and that the restoration of ANY functional mitochondrial genome would restore glycerol growth.

I compared my data for those 109 strains. In my strain collection, 34 out of the 109 strains are petite (they probably have lost mtDNA, or lack a functional mitochondrial genome); 27 strains were wild type for glycerol growth and thus were not included in my analysis (of these, 25 were non-mitochondrial genes); 35 strains had a glycerol growth defect but a wild-type splicing pattern for *COX1* and *COB*; and finally 13 strains had a northern blot pattern that deviated from wild-type. Of these 13 strains, 7 are discussed in my results as potential splicing factors (*Δpet54, Δcbs2, Δimp1, Δcem1, Δcox9, Δcoq4, Δpet117*). The other six strains either splice just barely worse than wild-type or I was unable to reproduce the initial result (*Δpet122, Δcoq2, Δmdm38, Δgon5, Δsap30, Δyjl046w*). The finding by Luban *et*

al. (2005) that *Δpet54, Δcbs2, Δimp1, Δcox9, Δcoq4* and *Δpet117* strains have improved glycerol growth in the absence of intron aI5β (among other introns) supports my observation that these six strains are blocked for aI5β splicing.

Mitochondrial DNA is spontaneously lost from some strains. In my case I find at least 37 strains for which this appears to be the case. Of the 109 strains listed by Luban *et al.*, it is probable that many of them lost mtDNA. At least, the 34 strains that I found to be petite due to mtDNA loss and the 27 strains that I find have wild-type glycerol growth are likely to lack mtDNA in the yeast library obtained by Dr. Schmidt's lab (summing to 61 strains largely deleted for genes coding for non-mitochondrial proteins).

The screen by Luban et al suggested that 35 strains, which in my hands do not have a splicing defect, have improved glycerol growth with fewer introns. If their hypothesis is correct, then I should have seen a splicing defect for these strains. I did not. Some of these strains may also have lost mtDNA in the parental library, as many of them are able to grow somewhat on glycerol in my hands. It is also possible that the 13 intron genome differs significantly from the 4 intron genome in other ways besides the number of introns. A reduced number of introns may indirectly alleviate a glycerol growth block by generally simplifying mitochondrial biogenesis.

4-3-3 Mitochondrial translational activators and splicing

Transcription, translation and membrane insertion of all three mitochondrially expressed *COX* genes appears to be temporally linked in the mitochondria (Brown *et al.*, 1994; Naithani *et al.*, 2003; Rodeheffer *et al.*, 2001; Towpik, 2005). During or immediately following transcription, translational activators bind RNA structures in the 5' untranslated region (in the case of *COX2*, this is also the RNA that codes for the first 14 amino acids of the preprotein) and localize their target message to sites near ribosomes on the inner membrane surface where the proteins are translated and co-translationally inserted into the inner membrane (Towpik, 2005). Assembly of the active cytochrome c oxidase and bc_1 complexes most likely also occur at these highly protein-dense "nodes" (Herrmann & Funes, 2005; Krause *et al.*, 2004; Towpik, 2005).

The intron-encoded maturases are expressed as protein fusions with the upstream exon sequences and then subsequently proteolytically processed to produce the final, active maturase protein. Our lab has mutated the proteolytic processing site for the aI2 maturase and shown that larger protein precursors do accumulate (Jiang, unpublished observations); it is interesting that most such mutants accumulate multiple larger precursors, suggesting that either the precursor is unstable and is subject to some proteolysis *in vivo* or that the full-length precursor is processed in steps, some of which are observed in those mutant strains.. Those precursor proteins do not appear to be active for intron splicing – because the strains do not splice. We do not know if they could transpose, though it is unlikely. Splicing is likely to occur within or near these highly protein-dense nodes of transcription-translation-

membrane insertion-complex assembly. This is further complicated by the fact that most of the translational activator proteins are RNA binding proteins and good candidates for a direct role in intron splicing. It is already established that Pet54p is both a translational activator for *COX3* and a splicing factor for *COX1*. Pet54p, and indeed most of the message-specific translational activators, are not located in separate, message-specific complexes, but are present in complexes with each other (Towpik, 2005) and (Figure 4-25). Pet54p is therefore available to splice introns out of the proximally located *COX1* message. My finding that many proteins known to localize to these nodes of cytochrome assembly (*PET54, CBS1, CBS2, PET111, PET117, IMP1, CBP1, CBP6, COX*9) also have splicing defects is on one level not surprising, but should also not be brushed aside as a guaranteed secondary effect. A few may be directly involved in intron splicing and merit further consideration. For the sake of this discussion I have re-drawn a figure taken from the review by Topik (2005) that diagrams the physical interactions between the RNA polymerase, message specific translational activator proteins, and the ribosome (Figure 4-25). The small P or I next to each protein indicates whether it is a peripheral or integral matrix facing membrane bound protein. As is illustrated, the translational activators Cbp1p (*COB*), Cbs2p (*COB*) and Pet309p (*COX1*) form a tight complex that interacts with the ribosome and Nam1p (which is an RNA polymerase associated protein) via Cbs2p. The *COX3* specific translational activators (Pet54p, Pet122p and Pet494p) interact with each other, the ribosome, Nam1p and the Pet309p-containing complex. The *COX2* translational activator, Pet111p interacts with Nam1p and the *COX3* translational activators. Cbs1p (*COB*) also interacts directly with the ribosome and Nam1p, but was not shown to have any direct contacts with the other

translation activators. From the studies leading up to these results, it seems probable that the translational activators form part of dynamic complexes and create a network for coordinating mitochondrial transcription (probably including splicing), translation and co-translational membrane insertion.

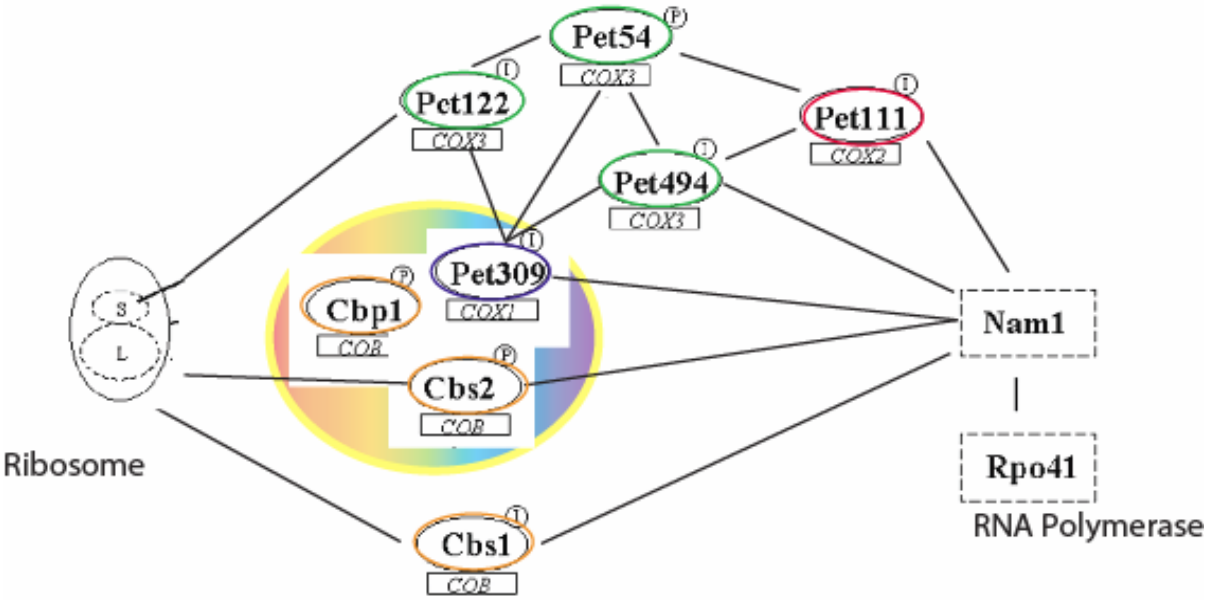


Figure 4-25. Summary of interactions between mitochondrial message specific translational activators. Based on two-hybrid and co-precipitation studies. Adapted from review by (Towpik, 2005).

4-3-3a PET309

Pet309p is the only translational activator that is essential for *COX1* translation. Pet309p binds to RNA sequences in the 5'UTR and localizes *COX1* mRNA to ribosomes. Pet309p is an integral, inner membrane protein and forms an isolatable complex with Cbp1p and Cbs2p (Manthey *et al.*, 1998; McEwen *et al.*, 1986). The original *pet309* mutant lacked any *COX1* precursor or spliced mRNA, but had wild-type levels of *COB* mRNA (McEwen *et*

al., 1986). Reducing the number of introns progressively restores *COX1* RNA, with the intron-less *COX1* gene supporting wild-type levels of correct length mRNA. Removing introns, however, did not restore glycerol growth or the production of Cox1p (Manthey *et al.*, 1998). These data show that, in addition to its well characterized role in translation, Pet309p is involved in stabilizing intron-containing *COX1* transcripts. Consistent with this, I find that strain *Δpet309* has lost the thirteen intron mitochondrial genome but is able to maintain intron-less mtDNA (section 4-2-12a; Table 4-4). In the study by Manthey *et al.* (1995), the destabilizing effect of *Δpet309* was not specific to any particular intron but rather the total number of introns (they looked at 8 different intron combinations). Consequently, the authors proposed that Pet309p may act by suppressing aberrant splicing reactions that are normally targeted for degradation. However, aberrant splicing does not usually cause degradation of mitochondrial RNAs, so it is unclear what role Pet309p plays in stabilizing mRNAs. The unusual sizes of the three novel aI5β-containing bands that I observe in strains *Δpet54, Δcbs2, Δimp1* and *Δpet111* and the similar bands observed in strains *Δcoq3*, *Δcoq4, Δcox9* and *Δpet117* give credence to the idea that alternative splicing reactions may be increased in strains with disrupted translational activators.

4-3-3b PET54

In the current study, I detected an aI5γ–containing band in the *Δpet54* strain. It was surprising that a well-characterized gene like *PET54* would reveal a new phenotype. The choice of exon 4 probe in all of the original papers, which would not hybridize with the splicing intermediate, may explain their failure to observe this novel band. Using an exon 4

probe, it is possible for a *Δpet54* strain that lacks aI5β but contains aI5γ to appear to have no *COX1* splicing defect, but the total signal for mRNA is expected to be reduced. Looking at the figures from older publications, the mutant strains do appear to have less overall spliced *COX1* message than wild-type. My observations may also be strain-specific.

Luban et al (2005) worked with the same library *Δpet54* strain as used in this study, and found that it is gly- but becomes gly+ when the number of introns is reduced to four (aI3α, aI4α, bI4, bI5). This is the expected outcome because intron aI5β was removed. They also cytoducted in a five intron strain that additionally contains aI5γ and performed northern analysis of this strain. If the aI5γ containing band I observe is due to an aI5γ and not aI5β, then they should have observed an aI5γ splicing block in this strain. They also used the same exon 4 probe that is unable to detect the aI5γ-containing band I observe in the *Δpet54* strain, however they also used an intron probe that would have been able to detect band I observe. They claim that the *Δpet54* strain containing introns aI3α, aI4α, bI4, bI5 and aI5γ splices like wild-type and has a wild-type level of excised intron RNA. Therefore, I conclude that the aI5γ splicing block observed in the library strain with thirteen introns is likely an indirect effect of the aI5β splicing block and not due to an inhibition of aI5γ. It is possible that in the *Δpet54* strain that the normal RNA processing is disrupted to such an extent that normally suppressed alternative splicing reactions occur. Such alternative reactions are worth studying and may lead to new insights into how mitochondrial introns are processes as part of the larger transcription-translation machinery.

4-3-3c COX9

Strain *Δcox9* has the same aI5γ-containing band as *Δpet54* and the same aI5β containing band that is observed in strains *Δimp1* and *Δpet111*. Given that for strain *Δpet54* the aI5γ containing band is likely a mis-splicing reaction caused by the aI5β splicing block, strain *Δcox9* probably primarily disrupts aI5β splicing. *COX9* encodes the cytochrome c oxidase subunit VIIa, a widely conserved member of the active cytochrome c oxidase holoenzyme. It is thought to serve a largely structural role and is not a part of the catalytic site (Wright *et al.*, 1986). Still, it is hard to imagine that this protein would be directly involved in intron splicing. It is more likely that the absence of Cox9p disrupts a complex or process needed for aI5β splicing.

4-3-3d PET111 and *IMP1*

Imp1p is an inner membrane protease that forms a complex with Imp2 and faces the matrix (Schneider *et al.*, 1991). Both Imp1p and Imp2p have protease activity and are involved in processing mitochondrial proteins, but each subunit cleaves different substrates (Nunnari *et al.*, 1993). Among other reactions, Imp1p cleaves the leader peptide of Cox2p. This leader peptide is translated from a hairpin-containing RNA sequence that is required for *COX2* mRNA translation. Destabilizing mutations in this RNA hairpin are suppressed by over-expression of Pet111p (or Mrpl36p) (Bonnefoy *et al.*, 2001). Pet111p is the only *COX2* message-specific translational activator, is also tightly associated with the inner membrane, and is rate limiting for Cox2p expression (Green-Willms *et al.*, 2001). I find that the deletion of either *IMP1* or *PET111* causes a new aI5β-containing RNA band to appear. Although it is

the same band in both strains, the band is much stronger in *Δimp1*. I suspect that both *Δimp1* and *Δpet111* similarly disrupt the assembly complex for Cox2p and that this translation/assembly complex (because it is temporally linked with the translation and membrane insertion of other COX messages) is somehow required for the correct splicing of aI5β.

4-3-3e COQ3 and *COQ4*

I find that both *Δcoq3* and *Δcoq4* do not have a significant reduction in spliced *COX1* or *COB* message but northern blots contain a minor band that runs slightly slower than spliced *COX1* mRNA (~2KB). This band runs at the same position as the intron aI5β-containing band observed in *Δimp1* and *Δpet111*.

Coq3p and Coq4p are the central peptides of a multi-protein coenzyme Q biosynthetic complex. Coq3p is one of the primary catalytic components and is also required for the formation of the polypeptide complex (Marbois *et al.*, 2005). Interestingly, the coenzyme Q biosynthetic complex co-purifies with the bc_1 complex - Cobp is the mitochondrially expressed component of this complex - and the bc_1 complex fails to form in a *Δcoq3* strain (Marbois *et al.*, 2005). This failure to form large MW complexes in the *Δcoq3* or *Δcoq4* backgrounds may be linked to the new band observed in these strains.

4-3-3f PET117

PET117 was isolated as a gene required for the assembly of active cytochrome c oxidase (McEwen *et al.*, 1993). It has membrane spanning domains and localizes to the

mitochondrial inner membrane. Pet117p is an inner membrane protein and its membrane spanning domain is in the N-terminal half of the protein. It also contains a tetratricopeptide repeat (TPR)-like motif, which for Tom70p is thought to mediate protein-protein interactions as part of the import process. I find that the *PET117* deletion also contains the same band that is found in *Δimp1* and *Δpet111*; however, I was unable to confirm that this band contains intron aI5β. It is quite likely, given the little that is known about Pet117p, that it is included in the same complexes or "nodes" as Pet54p, Imp1p, Pet111p and/or Coq3p and Cox9p that are involved in the translation and assembly of the cytochrome complexes and that the defect in *Δpet117* is related to that of the other discussed proteins.

4-3-3g CBS2

Cbs2p is a *COB*-specific translational activator that forms a tight complex with Cbp1p and Pet309p (Krause *et al.*, 2004). This complex interacts physically with both the RNA polymerase associated protein Nam1p and the ribosome (Krause-Buchholz *et al.*, 2005). As expected, *Δcbs2* has splicing defects for all *COB* encoded maturase-requiring introns (bI2, bI3, bI4 and aI4α; section 4-2-5a – Figure 4-12) . Less easily to explain is the fact that *Δcbs2* accumulates an aI5β-containing RNA band. Unlike strains *Δimp1* and *Δpet111*, this band runs faster, not slower than *COX1* message RNA. The band runs at ~3KB, a size that does not neatly match any combination of exons and is clearly smaller than the full *COX1* message containing only aI5β (3.8 KB). Further northern blots with fewer introns and/or probes for upstream COX1 exons could be used to further characterize this novel RNA.

4-3-3h A number of strains that have aI5β splicing defects also have an altered aI5α. band pattern

Intron aI5α self-splices *in vitro* (Winter *et al.*, 1988) and can transpose to empty target sites *in vivo* using the standard group I intron transposition pathway (Moran *et al.*, 1992; Seraphin *et al.*, 1992). I find that, unlike most other group I introns, excised aI5α intron accumulates in this wild-type background (aBY4741). *In vitro* self splicing reactions of intron aI5α accumulate excised circular intron RNA (Winter *et al.*, 1988), as do many other group I introns (reviewed in Nielsen *et al.*, 2003). However circular group I introns are usually shorter than the full length intron because the reaction releases a shore segment of the intron. It is possible that the excised aI5α intron is stabile because much of it is circular (perhaps more so than group I introns generally) and not linear, or that circular aI5α is more stable than most group I intron circles. This hypothesis could easily be tested using RT-PCR to sequence the aI5α intron RNA.. It is interesting that strains Δ*cbs2, Δimp1, Δpet54*, which I show to have an aI5β splicing defect, fail to accumulate excised aI5α RNA. A complete analysis of the splicing defects in these strains should also include a careful look at aI5α splicing.

4-3-3i Summary

In summary, I show that there exists an as yet to be characterized link between large mitochondrial complexes that localize to the mitochondrial inner membrane and the splicing of intron aI5β. It is possible that the intron splicing occurs in this large multi-functional

complex or "node," and that simply disrupting the higher order complex (as has been shown to occur in *Δcoq3* strains) reduces the splicing of aI5β. Alternatively, it is possible that *accurate* splicing requires the proper localization of mRNA (and newly transcribed maturases) to the mitochondrial inner membrane, and that un-localized message activates alternative splicing reactions. Further characterization of the novel RNA bands and those introns that are required for their formation may begin to explain the exact splicing problem in these strains.

4-3-4 MNE1

Of the genes identified in this screen, *MNE1* is one of the best splicing factor candidates. The *Δmne1* strain grows very slowly on glycerol medium and splices *COX1* ~40%. *COB* splicing is essentially wild type at ~80%. Intron aI5β-containing precursors accumulate in the *Δmne1* strain, and the precursor sizes are the same as that seen for the known aI5β splicing factors *Δpet54* and *Δmss18*. It does not contain any of the novel aI5β-containing bands that I observe for the translational activators and associated proteins. This "classic" precursor pattern may indicate that the splicing problem for *Δmne1* is similar to that found in *Δmss18*, which is directly involved in stabilizing the correct intron RNA fold.

MNE1 codes for an uncharacterized protein that localizes to the mitochondria (Huh *et al.*, 2003). It was cloned in 1996 along with 18 other genes on the right arm of chromosome XV, and has already been named (Purnelle & Goffeau, 1996). A BLAST search reveals only one *Candida glabrata* gene with homology to *MNE1*.

Future studies of *MNE1* should start by confirming the phenotype with a single intron strain. Intron aI5β can splice in a petite background, and a mitochondrial petite already exists for looking at aI5β splicing in the absence of translation (Johnson & McEwen, 1997). Splicing of aI5β from this petite mtDNA should be measured in the *Δmne1* background. If splicing of aI5β is still impaired, an even stronger case can be made that *MNE1* plays a direct splicing role, and *in vitro* experiments would be the next logical step. If a functional, tagged Mne1p can be made, *in vitro* experiments could be performed with Mne1p and aI5β to establish if Mne1p is directly involved in assisting aI5β splicing.

4-3-5 AAT1

I show that strain *Δaat1* can grow on glycerol and contains ~20% spliced *COB* mRNA and ~70% spliced *COX1* mRNA. This reduction in spliced message is due to a strong bI3 splicing block and reduced splicing of introns bI4, aI4α, bI2, bI5 and aI3α (section 4-2-7 – Figure 4-17). *AAT1* was first cloned and sequenced serendipitously as part of another screen (Morin *et al.*, 1992). The sequence revealed that it had a high homology to aspartate aminotransferases (AspAT, which catalyzes the reversible reaction that, along with glutamate dehydrogenase, generates NH_4^+ from aspartate) but was not the previously described yeast cytoplasmic AspAT (Morin *et al.*, 1992). The amino acid sequence of AAT1 is highly homologous to TyrB (Expect = 1e-132), the bacterial AspAT family, an exception being that AAT1 possesses four insertions not found in other AspATs, including closely related fungal AspATs (BLAST search). Morin *et al.* (1992) proposed that AAT1 is the mitochondrial AspAT, consistent with its homology and the presence of a likely mitochondrial targeting

sequence. Furthermore, AAT1 was shown to localize to the mitochondria (Sickman *et al.*, 2003). No biochemical studies have shown that AAT1 is indeed an AspAT, but it is highly likely.

Mutations of *AAT1* (named *bmp3*) were previously isolated as suppressors of a recombination event between short repeated sequences in bI4 and aI4α (P. Perlman, personal communication of unpublished data by D. Gerber and J. V. Moran). The recombination event that is inhibited in the *aat1* mutant creates a functional hybrid maturase (mostly encoded by the aI4α maturase). This hybrid maturase is able to splice both aI4α and bI4 in *trans*, and therefore suppresses the gly^- phenotype of the parental strain (which has a stop codon in the bI4 ORF). Moran showed that the defect in the *bmp3* mutant strain affected the recombination event and not the expression of the hybrid maturase. When the petite genome encoding the hybrid maturase was introduced to the *bmp3* bI4 mutant strain, glycerol growth was observed. It is not clear how this relates to my observation that *AAT1* is involved in bI3 splicing, but it is evidence that AAT1 is likely a bi-functional mitochondrial protein.

AAT1 is a strong candidate for a bI3 splicing factor and should be pursued further. Future experiments should begin by looking at splicing using mitochondrial genomes containing a single intron. These experiments are needed to determine the splicing defects that are directly caused by this deletion and which defects, if any, are secondary. It is possible that *AAT1* is required for efficient mitochondrial recombination and that the splicing phenotype is somehow related to reduced recombination.

4-3-6 *ATP10*

The *Δatp10* strain grows slowly on glycerol and appears to have a general splicing block. This is evidenced by a reduction in spliced message for both COX1 (from ~80% to ~30% of total RNA) and *COB* (from ~ 86% to ~25% of total RNA) and an increase in precursor RNAs representing at least 10 different introns. Only the group II introns aI1 and aI2, and the group I intron omega, appear unaffected. The group I introns aI3α, aI4α, aI5α, aI5β, bI2, bI3, bI4, bI5 are more inhibited than the group II introns aI5γ and bI1. However the effect on aI5γ and bI1 is significant because only *MSS116* is known to alter their splicing.

The deletion of *ATP10* does not appear to stabilize otherwise degraded RNAs (as occurs in *Δpet127*), because the level of excised aI5γ and bI1 intron RNA is reduced relative to wild-type and because excised aI5α is also missing. My observations (in addition to the published data, see below) do not support a role for Atp10p in translation because several introns known to splice in the absence of translation - notably aI3α, aI5β, aI5γ and bI1 - are inhibited, whereas several introns that absolutely require translation - aI1 and aI2 - are not.

Atp10p is a protein chaperone required for the stabilization of mitochondrially expressed Atp6p and for the incorporation of Atp6p into the ATPase holoenzyme (Tzagoloff *et al.*, 2004). Atp10p is an integral inner membrane protein that crosslinks to the highly hydrophobic Atp6p shortly after or during its translation. Deletion of *ATP10* does not alter mitochondrial translation (Tzagoloff *et al.*, 2004). The glycerol growth defect of *Δatp10* occurs because very little functional ATPase is made in this strain (Ackerman & Tzagoloff, 1990). The level of splicing I observe is sufficient for the production of wild-type levels of Cox1p and Cobp; studies of *Δatp10* show a small reduction in Cox1p and no reduction in

COB protein (Tzagoloff *et al.*, 2004). It is possible that Atp10p is also required as a chaperone for other proteins, one or more of which may be needed for efficient splicing.

Further characterization of the splicing defect in *Δatp10* should begin by looking at splicing in single intron mitochondria. An *ATP6* mutation (alanine 249 to valine; called *Δatp10/R3*) improves glycerol growth in the *Δatp10* strain (Paul *et al.*, 2000). Splicing should be examined in the *Δatp10/R3* strain background to determine if the splicing defect is somehow linked to reduced ATPase function or the presence of unassembled ATPase components. If splicing is improved in the *Δatp10/R3* strain, than the splicing defect in *Δatp10* is probably an indirect effect of the known function of Atp10p.

4-3-7 *PIF1*

In this study, the *Δpif1* strain is temperature sensitive. Shifting the cells from 30^{o}C to 37^{o}C reduced *COX1* splicing from ~85% to ~45%, and *COB* splicing from ~55% to ~20%. The primary cause of the reduced splicing appears to be a bI2 splicing defect (section 4-2-9; Figure 4-20).

Pif1p is a 5' to 3' DNA helicase that is found in the mitochondria and the nucleus. Pif1p has been most extensively studied in the nucleus, where it regulates telomere formation and elongation by removing telomerase from DNA ends (Boule *et al.*, 2005). Boule *et al.* show that purified Pif1p can unwind DNA:DNA and DNA:RNA duplexes as long as a 20 nucleotide 5' DNA overhang is present. They find that Pif1 cannot unwind RNA:RNA duplexes and does not bind to single stranded RNA. Because the catalytic activity of Pif1p is required, and binding alone is insufficient, they propose that Pif1p removes telomerase from

DNA ends by unwinding the DNA:RNA duplex formed between the telomerase RNA and the DNA substrate.

In the mitochondria, *PIF1* is involved in oxidative damage resistance and mtDNA recombination and/or replication. The t^s phenotype I observe has already been described (Foury & Kolodynski, 1983), and is probably due to the combined effects of *PIF1* deletion. A *PIF1* deletion strain exhibits increased mitochondrial point mutations and increased petite formation due to gross mtDNA rearrangements (Foury & Kolodynski, 1983). It is not known exactly how *PIF1* is involved in these processes. Pif1p may play a direct role, or it may act as a repressor of mtDNA replication; its absence could, by increasing the mtDNA replication rate, indirectly cause an increase in the mutation rate (O'Rourke *et al.*, 2002). From this and numerous other studies, it appears that Pif1p is a multifunctional protein that is involved in a variety of processes both in the nucleus and mitochondria.

My finding that the *Δpif1* strain has a specific bI2 splicing defect at increased temperatures is not easily explained by the known properties of Pif1p. The next logical step is to look at bI2 splicing in a reduced intron strain. If the splicing of bI2 is still strongly reduced at 37^{o}C, further studies are justified. Given that Pif1p unwinds DNA:RNA duplexes but not RNA:RNA duplexes it is more probable that the role of Pif1p in the splicing of intron bI2 is not to aid in RNA folding but to release unproductive DNA:RNA duplexes that form with bI2 precursor RNA. Intron bI2 is has some very AU rich regions, as do all other mitochondrial RNAs, and it is possible that Pif1p relieves the “tangling” that results from short, nearly identical sequences.

4-3-8 Intron omega

Wild-type RNA probed for intron omega contains two major bands and a faint third band. The 1.1 KB excised omega intron RNA is readily detected on RNA blots.. A slower running band is the 4.4KB intron-containing 21S rRNA precursor RNA. A faint ~3.0 KB band is most likely the 2.7KB RNA that is formed by dodecamer cleavage of the intron-containing I-SceI transcript (Zhu *et al.*, 1989). This 2.7KB RNA was proposed to be the mRNA for the omega-encoded endonuclease I-SceI and it appears to be required for the expression of the intron-encoded protein.

The only deviation from the wild-type phenotype that I observe for intron omega is an increase in amount the 2.7KB cleavage product. This band is observed in strains *Δnip100* , *Δyhl005c* (partial deletion of *MRP4*), *Δmrpl39* and *Δoye1*. Those genes do not appear to be needed for omega intron splicing, and it is possible that they are involved in 3' end cleavage of mitochondrial RNAs at dodecamer sequences.

4-3-9 Phase II: screening strains that lack mtDNA for intron splicing defects.

Of the 808 strains deleted for a gene coding for a mitochondrial protein, 178 strains had lost mtDNA and could not be immediately screened by northern blot. I performed mtDNA transfer crosses (kar crosses, called cytoduction), and for 33 of the 178 mitochondrial strains I easily isolated derivatives containing intact mtDNA. A more through search is expected to reveal some more strains that can maintain mtDNA. These strains should be analyzed by RNA blotting after replacing wild-type intron-containing mtDNAs. In

addition, twelve strains in the petite sub-library have retained mtDNA. In sum, 53 strains remain that can be analyzed directly by northern blot.

As many as 110 strains cannot maintain mtDNA and will need to be analyzed with an intron-containing pMIT DNA. Jill Kolesar looked at this set of strains and found that virtually all of them can maintain a pMIT derived petite genome, though she didn't analyze the ribosomal genes. Many proteins in this set are excellent candidates for splicing factors, including tRNA synthetases (13 genes), mitochondrial ribosomal proteins (48 genes), and proteins involved in transcription or with RNA binding motifs (7 genes).

For the majority of these strains that truly cannot maintain ρ^+ mtDNA, a special mitochondrial petite plasmid (pMIT) can be constructed containing three introns known to splice in "natural" petites. Making new mitochondrial constructs can be difficult because mitochondrial DNA has very few useful restrictions sites and it is important not to alter the base sequence by introducing new sites. To enable this project to progress, it may be easier to use a petite genome that already exists (for example a petite that contains only *COX1* with key introns), looking at splicing defects in introns aI3α, aI5β and aI5γ rather than constructing a new pMIT. It will be particularly important to repeat the wild-type several times to determine the basal level of splicing in pMITs. Because there is no translation in a ρ^- petite strain, genes that are directly involved in splicing are more likely to be isolated. Once the system is set up in a petite background, it would be informative to also look at some of the genes isolated in phase I. This would eliminate those genes that only indirectly alter splicing. This technique was used to confirm that MSS116 is directly involved in the splicing of a number of introns (Huang *et al.*, 2005).

Any new genes that are found to alter splicing of introns from a petite genome could be further analyzed using partial mutants. Strategic small deletions were used to determine which domains of NAM2 were important for its intron splicing activity (Rho *et al.*, 2002). Ideally, the translation and splicing functions of a potential splicing factor would be separable, confirming a direct splicing role.

4-3-10 Mitochondrial ribosomal proteins

Mitochondrial translation is required for mtDNA maintenance, though the molecular basis of that phenotype remains a mystery (Contamine & Picard, 2000). For this reason it has been assumed that the deletion of any mitochondrial ribosomal proteins will result in mtDNA loss. However, I find that the deletion of twenty-four of the seventy-seven genes coding for proteins that co-purify with ribosomes can in fact maintain mtDNA. Twenty-one of these are able to grow, at least slowly, on glycerol. One cannot argue that these proteins represent contaminants in the ribosomal preparations because at least fifteen of them are homologous to ribosomal proteins in *E. coli* or humans (Appendix H with homology data from Gan *et al.*, 2002 and SGD). The most likely explanation is that some ribosomal proteins are not essential for translation. For example, the removal of a single peripheral structural protein may not be sufficient to make the ribosome dysfunctional. One way to test this would be to delete several of the ribosomal proteins that maintain mtDNA and are able to grow on glycerol. Such a double or triple mutation may result in mtDNA loss.

APPENDIX A
Strains and oligonucleotides used in this dissertation

Table A-1: Strains.

Column "#": Storage record number in freezer box.

Column "nuclear background": Strain name of the nuclear background, see Chapter 2.

Column "mitochondria": name of mitochondrial genome.

#	nuclear background	mitochondria
93	a161	ρ°
94	a161 *Δdbr1*	ρ°
95	a161*MRS2 L232F*	ρ°
116	aBY4741	ρ°
115	BY4999 *(Δdbr1)*	ρ°
25	αMCC109	ρ°
272	αW303 (epi-4d)	ρ°
268	αMCC109	ρ+ (Foury; 13 intron)
224	αW303	ρ+ (Foury; 13 intron)
267	αMCC109	I°
165	a161	I°
166	a161 *Δdbr1*	I°
167	a161*MRS2 L232F*	I°
164	αMCC109	I°
16	a161	GII-5γ
100	a161 *Δdbr1*	GII-5γ
101	a161*MRS2 L232F*	GII-5γ
96	αMCC109	GII-5γ
102	a161	GII-5γ topGC
103	a161 *Δdbr1*	GII-5γ topGC
104	a161*MRS2 L232F*	GII-5γ topGC
97	αMCC109	GII-5γ topGC
18	a161	GII-5γ Δbranch
222	a161 *Δdbr1*	GII-5γ Δbranch
129	a161*MRS2 L232F*	GII-5γ Δbranch
177	αMCC109	GII-5γ Δbranch
134	a161	GII-5γ D5UA
135	a161 *Δdbr1*	GII-5γ D5UA
114	a161*MRS2 L232F*	GII-5γ D5UA
111	αMCC109	GII-5γ D5UA
105	a161	GII-5γ botGC
106	a161 *Δdbr1*	GII-5γ botGC
107	a161*MRS2 L232F*	GII-5γ botGC
98	αMCC109	GII-5γ botGC

#	nuclear background	mitochondria
7	a161	GII-5γ A880U
216	a161 *Δdbr1*	GII-5γ A880U
128	a161*MRS2 L232F*	GII-5γ A880U
175	αMCC109	GII-5γ A880U
5	a161	GII-5γ A880G
213	a161 *Δdbr1*	GII-5γ A880G
127	a161*MRS2 L232F*	GII-5γ A880G
176	αMCC109	GII-5γ A880G
6	a161	GII-5γ A880C
219	a161 *Δdbr1*	GII-5γ A880C
220	a161*MRS2 L232F*	GII-5γ A880C
174	αMCC109	GII-5γ A880C
8	a161	GII-5γ 2xGC
109	a161 *Δdbr1*	GII-5γ 2xGC
110	a161*MRS2 L232F*	GII-5γ 2xGC
99	αMCC109	GII-5γ 2xGC
234	αMCC109	$BrA\Uparrow_{botGC}$ (pMIT)
242	a161	$BrA\Uparrow_{botGC}$
44	a161	$BrA\Uparrow^{C878U}_{botGC}$
84	a161*MRS2 L232F*	$BrA\Uparrow^{C878U}_{botGC}$
54	MCC109	$BrA\Uparrow^{C878U}_{botGC}$
45	a161	$BrA\Uparrow^{C878U}$
85	a161*MRS2 L232F*	$BrA\Uparrow^{C878U}$
55	MCC109	$BrA\Uparrow^{C878U}$
42	a161	$BrA\Uparrow$
82	a161*MRS2 L232F*	$BrA\Uparrow$
52	MCC109	$BrA\Uparrow$
236	αMCC109	$BrA\Downarrow^{topGC}_{C882U}$ (pMIT)
239	a161	$BrA\Downarrow^{topGC}_{C882U}$
113	a161*MRS2 L232F*	$BrA\Downarrow^{topGC}_{C882U}$
112	αMCC109	$BrA\Downarrow^{topGC}_{C882U}$
235	αMCC109	$BrA\Downarrow^{topGC}$ (pMIT)
241	a161	$BrA\Downarrow^{topGC}$
238	αMCC109	$BrA\Downarrow_{C882U}$ (pMIT)
240	a161	$BrA\Downarrow_{C882U}$
43	a161	$BrA\Downarrow$
83	a161*MRS2 L232F*	$BrA\Downarrow$
53	MCC109	$BrA\Downarrow$
75	a161	bloop
76	a161*MRS2 L232F*	bloop
65	α947	bloop

#	nuclear background	mitochondria
50	a161	4E
77	a161*MRS2 L232F*	4E
60	MCC109	4E
49	a161	4D
78	a161*MRS2 L232F*	4D
59	MCC109	4D
48	a161	4C
79	a161*MRS2 L232F*	4C
58	MCC109	4C
47	a161	4B
81	a161*MRS2 L232F*	4B
57	MCC109	4B
46	a161	4A
80	a161*MRS2 L232F*	4A
56	MCC109	4A

Table A-2: Plasmids.

#	Plasmid name	Plasmid Description
	pJD20	parent plasmid with wild-type aI5γ
	pJD20-A880U	Podar et al. 1998
	pJD20-A880C	Podar et al. 1998
	pVC02	pJD20-ΔA_{880}; Chu et al. 1998
	pVC11	pJD20-BrA$\Downarrow$; Chu et al. 2001
	pEL04	$BrA\Downarrow^{topGC}{}_{C882U}$
231	pEL05	$BrA\Downarrow^{topGC}$
232	pEL06	$BrA\Downarrow_{C882U}$
233	pEL07	$BrA\Uparrow_{botGC}$
265	pJD20-3GU-UP	pJD20-C882U (not in yeast)
266	pJD20-3GU-DN	pJD20-C878U (not in yeast)
	pGEM-T easy	used for TA cloning (Promega kit)

Table A-3: Oligonucleotides.

Name	Sequence	Location
aI1-851as	GTTTCTAATGTTGTACCTGGAG	nt 831-851 of COX1 (aI1)
aI2 2720as	GAATAAACAGAGATATGTTTATC	nt 2697-2720 of COX1 (aI2)
aI3α–AS	GCTAAATAAGGTCCCAACTTATC	nt 5282-5304 of COX1 (aI3α)
aI4 8216 as		?8126
aI5α	TCTATTTGATCTTGGATAAATATC	nt 148-171 of aI5α
aI5β	TAATTATAAGAGAAACCCCGT	nt 1992-2012 of Hensgens 1983
5g5E	GGTAAATATTATTTATGATAACTTTC	nt 8629-8654 of COX1 (aI5γ)
SnabAS	GGGTTTATTCTGTTTTATC	nt 9024-9042 of COX1 (aI5γ)
D3-lower	ATCACCTATAGTATAAGTTA	nt 9279-9298 of COX1 (aI5γ)
116AS	GGATACTTTTGTCATTACAGC	nt 8735-8755 of COX1 (aI5γ)
657S	CTGCTAACTTATACTATAGG	nt 9275-9294 of COX1 (aI5γ)
E2-bottom	GACTTCAATAGTAGTATCCTG	nt 9546-9566 of COX1 (E8)
E6 9717as	GCGAAAGCATCAGGATAATCAGG	nt 9695-9717 of COX1 (E8)
cox1E6as	GAATAATGATAATAGTGCAATGAATGAACC	nt 9743-9772 of COX1 (E8)
cobE6as	AAGGTACTTCTACATGGCATGCT	nt 43509-43531 of COB (E6)
COB bI1AS	NA	From Hon Ren Huang
bI2	CCTTTTCTTTTTTCAGCATGACCATCTCC	nt 1462-1490 of COB (bI2)
bI3	GAGTATTAAATGGTTTAATAGGTCG	nt 3322-3346 of COB (bI3)
bI4	NA	From Hon Ren Huang
bI5	CCTTATGGGAGTTCCCACAAAGCGG	nt 6264-6288 of COB (bI5)
omega	CACGACTACGAATATAAGCATCAC	inside SCE1

APPENDIX B
Yeast mitochondrial proteome deletion strains not included in phase I or phase II of the splicing factor screen.

Table B-1: Strains not available in the mat a haploid deletion strain background.

Listed below are known or potential mitochondrial proteins that were not included in the current screen because they were not available in the mat a deletion background. Most of these are essential proteins that can not be looked at in a haploid background..

Mitochondrial proteome ORFs not availible as a *mat a* haploid deletion strain

ORF designation	Gene name	availability of deletion strain	viability of haploid strain
YNL185C	MRPL19	mat alpha haploid	viable
YDR322C-A	TIM11	not availible	viable
YGL119W	ABC1	not availible	viable
YGL187C	COX4	not availible	viable
YOL077W-A	ATP19	not availible	viable
YOL140W	ARG8	not availible	viable
YDR379C-A		not availible	not characterized
YHL018W		not availible	not characterized
YJL062W-A		not availible	not characterized
YJR101W	RSM26	not availible	not characterized
YML007C-A		not availible	not characterized
YBR039W	ATP3	heterozygous diploid	inviable
YDL120W	YFH1	heterozygous diploid	inviable
YDL164C	CDC9	heterozygous diploid	inviable
YDR041W	RSM10	heterozygous diploid	inviable
YDR232W	HEM1	heterozygous diploid	inviable
YDR376W	ARH1	heterozygous diploid	inviable
YER048W-A		heterozygous diploid	inviable
YGL018C	JAC1	heterozygous diploid	inviable
YGL068W		heterozygous diploid	inviable
YGR046W		heterozygous diploid	inviable

Mitochondrial proteome ORFs not availible as a *mat a* haploid deletion strain

ORF designation	Gene name	availability of deletion strain	viability of haploid strain
YGR082W	TOM20	heterozygous diploid	inviable
YGR147C	NAT2	heterozygous diploid	inviable
YHR024C	MAS2	heterozygous diploid	inviable
YHR083W		heterozygous diploid	inviable
YIL022W	TIM44	heterozygous diploid	inviable
YIL031W	ULP2	heterozygous diploid	inviable
YIL051C	MMF1	heterozygous diploid	inviable
YJL054W	TIM54	heterozygous diploid	inviable
YJL104W	MIA1	heterozygous diploid	inviable
YJR016C	ILV3	heterozygous diploid	inviable
YKL192C	ACP1	heterozygous diploid	inviable
YKL195W		heterozygous diploid	inviable
YLR105C	SEN2	heterozygous diploid	inviable
YLR132C		heterozygous diploid	inviable
YLR163C	MAS1	heterozygous diploid	inviable
YLR259C	HSP60	heterozygous diploid	inviable
YML025C		heterozygous diploid	inviable
YML091C	RPM2	heterozygous diploid	inviable
YMR059W	SEN15	heterozygous diploid	inviable
YMR108W	ILV2	heterozygous diploid	inviable
YMR301C	ATM1	heterozygous diploid	inviable
YNL137C	NAM9	heterozygous diploid	inviable
YNL256W	FOL1	heterozygous diploid	inviable
YNL306W	MRPS18	heterozygous diploid	inviable
YNL310C		heterozygous diploid	inviable
YNR017W	MAS6	heterozygous diploid	inviable
YOR004W		heterozygous diploid	inviable
YOR176W	HEM15	heterozygous diploid	inviable
YOR236W	DFR1	heterozygous diploid	inviable
YPL063W	TIM50	heterozygous diploid	inviable
YPL083C	SEN54	heterozygous diploid	inviable
YPL252C	YAH1	heterozygous diploid	inviable
YCL017C	NFS1	heterozygous diploid	inviable
YMR033W	ARP9	not availible	inviable

Table B-2: Gly+, ρ+ strains eliminated from the splicing factor screen.

The strains listed below all grow as well as wild type on glycerol and contain wt mtDNA. The data is presented in two columns and ordered by ORF name. The plate, column, row designation refers to the stored location of each strain in the sub-library grid, not the location of the strain in the commercial library stocks. No further analysis was performed with these strains, however they have essentially all been shown to localize to the mitochondria.

ORF name	*Gene*	Plate	Column	Row
YAL056W	*GPB2*	1	1	B
YBL013W	*FMT1*	9	4	C
YBL015W	*ACH1*	9	4	D
YBL057C	*PTH2*	9	5	B
YBL059W		9	5	D
YBL064C	*PRX1*	9	5	E
YBL095W		13	6	D
YBL098W	*BNA4*	13	6	E
YBR001C	*NTH2*	13	6	H
YBR047W		14	1	E
YBR085W	*AAC3*	15	3	H
YBR104W	*YMC2*	15	4	A
YBR176W	*ECM31*	3	5	B
YBR238C		3	6	A
YBR262C		3	6	C
YBR263W	*SHM1*	3	6	D
YBR291C	*CTP1*	13	1	C
YCL009C	*ILV6*	5	4	C
YCL044C		5	4	D
YCL064C	*CHA1*	5	4	E
YCR079W		13	1	G
YCR083W	*TRX3*	14	6	H
YDL027C		15	4	F
YDL048C	*STP4*	15	5	C
YJL131C		7	3	G
YJL147C		7	3	E
YJL161W		7	3	C
YJL200C		7	2	F
YJL208C	*NUC1*	7	2	E
YJR003C		15	2	A
YJR051W	*OSM1*	13	2	H
YJR062C	*NTA1*	13	3	A
YJR080C		12	5	E
YJR111C		12	4	A
YKL011C	*CCE1*	5	6	F
YKL027W		5	6	H
YKL029C	*MAE1*	6	1	A
YKL114C	*APN1*	6	2	E
YKL120W	*OAC1*	6	2	H
YKL150W	*MCR1*	6	3	E
YKL157W	*APE2*	6	3	F
YKL162C		6	3	G
YKL188C		6	4	C
YKR042W	*UTH1*	10	3	C
YKR052C	*MRS4*	10	3	E
YKR065C		10	3	F
YKR066C	*CCP1*	11	1	G
YKR070W	*MRPL20*	11	1	H

ORF name	Gene	Plate	Column	Row
YDL066W	*IDP1*	15	5	E
YDL130W-A	*STF1*	9	2	G
YDL157C		13	4	A
YDL174C	*DLD1*	13	4	B
YDL178W	*DLD2*	13	4	C
YDR019C	*GCV1*	12	6	E
YDR036C	*EHD3*	12	6	G
YDR061W		3	6	F
YDR070C		4	1	A
YDR125C	*ECM18*	4	1	H
YDR185C		8	2	A
YDR219C		8	2	F
YDR234W	*LYS4*	8	3	C
YDR258C	*HSP78*	10	3	G
YDR305C	*HNT2*	10	4	F
YDR430C	*CYM1*	4	3	B
YDR474C		15	6	D
YDR514C		16	1	D
YDR516C		16	1	E
YDR538W	*PAD1*	9	2	H
YEL006W		4	3	C
YEL052W	*AFG1*	4	3	H
YER053C		4	4	D
YER069W	*ARG5,6*	4	4	H
YER073W	*ALD5*	4	5	A
YER078C		11	3	D
YER080W		4	5	B
YER086W	*ILV1*	4	5	D
YER170W	*ADK2*	16	3	G
YER182W		16	4	A
YFL027C	*GYP8*	10	6	G
YFL046W		11	1	A
YFR011C		14	3	H
YFR024C-A	*LSB3*	11	1	B
YFR044C		9	3	C
YFR049W	*YMR31*	9	3	D
YGL059W		9	6	E
YGL080W		9	6	H
YGL085W		10	1	B
YGL226W		8	3	H
YLL001W	*DNM1*	1	1	H
YLL001W	*DNM1*	19	1	B
YLL009C	*COX17*	19	1	D
YLR072W		1	3	D
YLR121C	*YPS3*	1	4	B
YLR142W	*PUT1*	5	5	A
YLR188W	*MDL1*	5	5	E
YLR193C		5	5	F
YLR253W		7	5	E
YLR281C		7	6	B
YLR283W		7	6	C
YLR289W	*GUF1*	7	6	D
YLR290C		7	6	E
YLR348C	*DIC1*	7	3	H
YLR351C	*NIT3*	7	4	A
YLR356W		7	4	B
YLR390W	*ECM19*	7	4	E
YML013C-A		1	4	H
YML078W	*CPR3*	1	4	E
YML086C	*ALO1*	1	4	D
YMR003W		1	5	B
YMR023C	*MSS1*	1	5	D
YMR056C	*AAC1*	16	4	C
YMR062C	*ECM40*	11	2	A
YMR115W		11	5	B
YMR145C	*NDE1*	1	5	H
YMR157C		1	6	C
YMR166C		1	6	D
YMR177W	*MMT1*	1	6	F
YMR189W	*GCV2*	1	6	H
YMR225C	*MRPL44*	2	1	E
YMR241W	*YHM2*	2	1	G
YMR252C		2	2	A
YMR302C	*PRP12*	2	3	D
YNL063W		14	2	E
YNL083W		14	4	G
YNL084C		17	5	G
YNL100W		14	5	A
YNL104C	*LEU4*	14	5	B
YNL121C	*TOM70*	14	5	D

ORF name	Gene	Plate	Column	Row
YGL256W	*ADH4*	8	4	D
YGR012W		8	4	E
YGR015C		8	4	F
YGR021W		8	4	H
YGR028W	*MSP1*	14	5	G
YGR031W		6	4	D
YGR033C		6	4	E
YGR049W	*SCM4*	6	4	G
YGR096W	*TPC1*	6	5	C
YGR132C	*PHB1*	4	5	F
YGR193C	*PDX1*	4	6	F
YGR207C		4	6	G
YGR231C	*PHB2*	9	3	F
YGR235C		9	3	G
YGR243W		9	3	H
YGR244C	*LSC2*	9	4	A
YGR286C	*BIO2*	11	1	C
YHL002W	*HSE1*	12	2	C
YHL014C	*YLF2*	5	1	D
YHL021C		5	1	C
YHL032C	*GUT1*	5	1	B
YHR001W-A	*QCR10*	5	1	G
YHR003C		12	2	E
YHR106W	*TRR2*	5	2	H
YHR162W		11	3	G
YHR198C		5	3	H
YHR199C		5	4	A
YIL042C		10	6	C
YIL077C		10	5	G
YIL087C		10	5	H
YIL094C	*LYS12*	16	4	H
YIL111W	*COX5B*	16	4	F
YIL113W	*SDP1*	14	2	G
YIL114C	*POR2*	14	2	H
YIL136W	*OM45*	16	4	G
YJL043W		12	2	A
YJL052W		12	1	G
YJL060W	*BNA3*	12	1	F
YJL066C	*MPM1*	12	1	D
YJL071W	*ARG2*	12	1	C
YJL112W	*MDV1*	11	6	H
YNL122C		14	2	A
YNL168C		10	2	E
YNL198C		10	1	G
YNL211C		10	1	F
YNL239W	*LAP3*	10	1	C
YNR022C	*MRPL50*	13	5	D
YNR040W		13	5	H
YOL004W	*SIN3*	3	1	B
YOL042W	*NGL1*	3	2	A
YOL043C	*NTG2*	3	2	B
YOL150C		12	3	F
YOR040W	*GLO4*	2	4	H
YOR045W	*TOM6*	2	5	A
YOR108W	*LEU9*	6	5	H
YOR130C	*ORT1*	6	6	B
YOR215C		7	1	E
YOR226C	*ISU2*	7	1	H
YOR228C		7	2	A
YOR266W	*PNT1*	15	2	E
YOR286W		7	2	C
YOR297C	*TIM18*	2	5	E
YOR356W		2	6	E
YOR374W	*ALD4*	2	6	H
YPL069C	*BTS1*	8	5	E
YPL072W	*UBP16*	8	5	D
YPL091W	*GLR1*	8	5	A
YPL107W		3	4	F
YPL109C		3	4	E
YPL134C	*ODC1*	11	4	F
YPL168W		3	4	B
YPR001W	*CIT3*	8	6	G
YPR002W	*PDH1*	8	6	H
YPR004C		9	1	A
YPR006C	*ICL2*	11	5	D
YPR011C		4	4	A
YPR061C		11	5	H
YPR098C		11	6	B
YPR125W		9	1	E
YPR140W		9	1	G
YPR155C	*NCA2*	9	2	B

APPENDIX C
Phase I percent splicing data organized by systematic ORF name

See chapter 4 section 2-3 for a description of how the data was derived.

The chart is sorted in alphabetically order by systematic ORF name. Multiple entries for the same strain are independent preps of the same strain or the same strain grown at the permissive and non-permissive temperature, as indicated. The RNA blot lane for each of the listed samples can be looked up using the blot number and lane number in appendix D.

Chart Legend:

ORF name: Systematic ORF name of the strain.

blot: Blot number of the sample. There are a total of 37 blots. The RNA blot images can be found in appendix C.

Blots labeled canDI and cand2 refer to the final candidate splicing factor blots that were probed for each individual intron. (see appendix E)

lane: blot lane of the sample. Use to look up the sample in appendix D.

Gene name: Gene name if available.

COX % splicing: *COXI* percent splicing calculated as described in section 2-4-2.

***COB* % splicing:** *COB* percent splicing calculated as described in section 2-4-2.

rel total RNA : Relative measure of the amount of total RNA loaded into the lane based on the visual inspection of the large and small rRNA bands on the Ethidium bromide stained gel:

(-) no RNA or totally degraded RNA.

(+/-) barely visible RNA, quantification is not reliable.

(+) low but clearly visible and quantifiable level or RNA.

(++), (+++), (++++) good to excessive level of RNA, all give reliable quantification. Too much RNA tends to underestimate % splicing.

temp: Temperature at which the sample was grown. Most but not all temperature sensitive strains were also grown at 30^{o}C.

glycerol growth: Ability of the strain to grow on glycerol/respire.

(-) unable to grow on glycerol.

(+/-) slow growth on glycerol (observed in this study).

(l+/-) reduced growth on glycerol (as reported on the SGD website).

(+) able to grow on glycerol. These strains were included in the screen because the knock out gene was potentially interesting or uncharacterized.

library grid location: location of the sample in the sub-library 96-well grid. The format is (Plate.Column.Row) A portion of the YPR culture assayed by northern blot is also stored at -80^{o}C in 2ml vials. These 2ml freezer stocks labeled with this format: (Plate. Column. Row; for example 16.5.B is the label on the *Δcox9* tube) These tubes are the most reliable method of obtaining a strain from our collection.

Each blot also has a wild-type control lane. The average of the wild-type lanes are summarized below:

Summary of wt samples:	COX % splicing	COB % splicing
WT Average	0.80	0.86
WT Standard Deviation	+/- 0.10	+/- 0.06
n=	29	31
5th - 95th percentile range	0.61 - 0.99	0.73-0.98

ORF name	blot	lane	Gene name	COX % splicing	COB % splicing	rel. total RNA	temp (°C)	Glycerol growth	library grid location
YAL008W	6	5	*FUN14*	**0.90**	**0.88**	+++	30	c^s**+/-**	1.1.g
YAL008W	7	10	*FUN14*	**0.71**	**0.67**	+++	***20***	c^s**+/-**	1.1.g
YAL010C	11	3	*MDM10*	**0.83**	**0.99**	++	30	**+/-**	1.1.f
YAL010C	20	1	*MDM10*	**0.68**	**0.90**	++	30	**+/-**	1.1.F
YAL012W	26	17	*CYS3*	**0.61**	**0.68**	+	30	**+/-**	12.5.B
YAL016W	33	12	*TPD3*	**0.70**	**0.75**	+++	30	**+/-**	17.4.D
YAL026C	9	1	*DRS2*	**0.82**	**0.74**	+++	***20***	l**+/-**	1.1.e
YAL039C	11	2	*CYC3*	**0.66**	**0.83**	++	30	**-**	1.1.d
YAL044C	11	1	*GCV3*	**0.68**	**0.84**	++	30	**+/-**	1.1.c
YBL007C	8	9	*SLA1*	**0.91**	**0.87**	+++	***37***	t^s**-**	9.4.b
YBL019W	22	7	*APN2*	**0.96**	**0.97**	++	30	l**+/-**	9.4.E
YBL021C	22	8	*HAP3*	**1.00**	**1.00**	+	30	**+/-**	9.4.F
YBL033C	33	3	*RIB1*	**0.78**	**0.79**	+++	30	l**+/-**	17.1.H
YBL045C	22	9	*COR1*	**0.99**	**0.96**	+	30	**-**	9.5.A
YBL058W	8	10	*SHP1*	**0.94**	**0.90**	++	***37***	t^s**-**	9.5.c
YBL089W	36	16		**0.62**	**0.53**	+++	30	l**+/-**	9.5.G
YBL099W	28	17	*ATP1*	**0.74**	**0.80**	++	30	**-**	13.6.F
YBL100C	28	18		**0.71**	**0.59**	++	30	**-**	13.6.G
YBR003W	29	1	*COQ1*	**0.74**	**0.81**	+	30	**-**	14.1.A
YBR026C	29	2	*ETR1*	**0.63**	**0.74**	++	30	**+/-**	14.1.B
YBR035C	34	11	*PDX3*	**0.79**	**0.82**	+++	30	l**+/-**	17.6.G
YBR037C	1	20	*SCO1*	**0.96**	**0.15**	+	30	**-**	14.1.c
YBR037C	2	17	*SCO1*	**0.66**	**0.74**	+	30	**-**	14.1.c
YBR037C	4	7	*SCO1*	**0.84**	**0.55**	+	30	**-**	14.1.c
YBR044C	37	2	*TCM62*	**0.65**	**0.65**	++++	30	**+/-**	14.1.D
YBR084W	30	7	*MIS1*	**0.81**	**0.85**	++	30	**+/-**	15.3.G
YBR112C	35	8	*SSN6*	**0.73**	**0.72**	+++	30	**-**	18.2.D
YBR120C	3	7	*CBP6*	**0.00**	**0.55**	+	30	**-**	15.4.b
YBR127C	30	8	*VMA2*	**0.94**	**0.93**	++	30	**+/-**	15.4.C
YBR146W	27	7	MRPS9	**0.68**	**0.78**		30	**+**	MRPS9
YBR185C	14	13	*MBA1*	**0.46**	**0.95**	++	30	l**+/-**	3.5.d
YBR187W	36	2		**0.64**	**0.65**	++++	30	l**+/-**	3.5.e
YBR191W	6	4	*RPL21A*	**0.83**	**0.90**	+++	30	t^s**-**	15.1.b
YBR191W	7	6	*RPL21A*	**0.84**	**0.66**	+++	***37***	t^s**-**	15.1.b
YBR221C	14	14	*PDB1*	**0.73**	**0.92**	++	30	l**+/-**	3.5.f
YBR227C	8	7	*MCX1*	**0.94**	**0.91**	++	***37***	t^s**+/-**	3.5.g
YBR230C	5	13	*YBR230C*	**0.91**	**0.77**	+++	30	t^s**+/-**	3.5.h
YBR230C	7	7	*YBR230C*	**0.83**	**0.69**	++	***37***	t^s**+/-**	3.5.h

ORF name	blot	lane	Gene name	COX % splicing	COB % splicing	rel. total RNA	temp (°C)	Glycerol growth	library grid location
YBR269C	28	1		0.75	0.87	+	30	l+/-	13.1.A
YBR289W	29	16	*SNF5*	0.79	0.92	++++	30	+/-	14.6.C
YCL007C	35	4	*CWH36*	0.68	0.78	++++	30	l+/-	18.1.F
YCL008C	8	5	*STP22*	0.88	0.89	++	***37***	t^s+/-	5.4.b
YCR009C	16	5	*RVS161*	0.93	0.83	++++	30	+/-	5.4.g
YCR044C	34	2	*PER1*	0.70	0.85	+++	30	l+/-	17.4.H
YCR047C	29	17	*BUD23*	0.67	0.93	+++	30	+/-	14.6.F
YCR053W	29	18	*THR4*	0.77	0.93	+++	30	l+/-	14.6.G
YCR065W	28	2	*HCM1*	0.84	0.86	+++	30	l+/-	13.1.E
YCR081W	6	2	*SRB8*	0.82	0.78	+	30	t^s-	13.1.h
YCR081W	7	4	*SRB8*		0.65	+/-	***37***	t^s-	13.1.h
YCR081W	10	6	*SRB8*	0.65	0.69	++	***37***	t^s-	13.1.h
YCR084C	30	1	*TUP1*	0.72	0.84	++	30	-	15.1.A
YDL006W	30	9	*PTC1*	0.67	0.85	+++	30	+/-	15.4.E
YDL032W	3	8	*YDL032W*	0.43	0.57	++	30	+/-	15.4.g
YDL032W	cand 1	8	*YDL032W*	0.40	0.89	++	30	+/-	15.4.G
YDL033C	3	9	*MTO2*	0.51	0.73	+/-	30	+/-	15.4.h
YDL033C	4	6	*MTO2*	0.54	0.38	+	30	+/-	15.4.h
YDL033C	cand 1	9	*MTO2*	0.66	0.86	+++	30	+/-	15.4.H
YDL047W	32	1	*SIT4,*	0.78	0.82	++	30	+/-	16.5.A
YDL056W	31	1	*MBP1*	0.87	0.91	++	30	l+/-	15.5.D
YDL067C	32	2	*COX9*	0.50	0.84	+	30	-	16.5.B
YDL068W	30	10		0.90	0.95	++	30	-	15.5.F
YDL069C	3	10	*CBS1*	0.47	0.70	+	30	-	15.5.g
YDL069C	cand 1	7	*CBS1*	0.15	0.60	++	30	-	15.5.G
YDL106C	28	7	*PHO2*	0.71	0.87	+	30	l+/-	13.3.C
YDL107W	1	19	*MSS2*	0.78	0.69	++	30	-	13.3.d
YDL117W	28	8	*CYK3,*	0.77	0.85	+++	30	l+/-	13.3.E
YDL118W	28	9		0.74	0.81	+++	30	l+/-	13.3.F
YDL119C	28	10		0.73	0.82	++	30	l+/-	13.3.G
YDL142C	10	8	*CRD1*	0.71	0.81	+	***37***	t^s+/-	13.3.h
YDL185W	34	6	*TFP1*	0.80	0.79	+++	30	+/-	17.5.D
YDR027C	27	2	*LUV1*	0.80	0.78		30	l+/-	12.6.F
YDR069C	8	2	*DOA4*	0.90	0.92	+++	***37***	t^s+/-	3.6.h
YDR116C	4	3	*MRPL1*	0.88	0.63	+	30	+/-	4.1.f
YDR120C	15	1	*TRM1*	0.94	0.87	++++	30	l+/-	4.1.g
YDR129C	15	2	*SAC6*	0.96	0.88	++++	30	l+/-	4.2.a
YDR136C	36	3		0.62	0.69	++++	30	l+/-	4.2.b

ORF name	blot	lane	Gene name	COX % splicing	COB % splicing	rel. total RNA	temp (°C)	Glycerol growth	library grid location
YDR148C	19	1	*KGD2*	**0.90**	**0.78**	++	30	**+/-**	8.1.B
YDR150W	9	13	*NUM1*	**0.66**	**0.69**	+++	***20***	c^s**+/-**	8.1.c
YDR162C	36	11	*NBP2*	**0.61**	**0.78**	+++	30	l**+/-**	8.1.D
YDR176W	36	12	*NGG1*	**0.52**	**0.88**	+	30	l**+/-**	8.1.F
YDR178W	36	13	*SDH4*	**0.66**	**0.66**	+++	30	**+/-**	8.1.G
YDR195W	4	15	*REF2*	**0.80**	**0.58**	++	30	**+/-**	8.2.c
YDR197W	2	3	*CBS2*	**0.17**	**0.00**	+	30	**-**	8.2.d
YDR197W	21	10	*CBS2*	**0.10**	**0.09**	+	30	**-**	8.2.D
YDR197W	cand 1	2	*CBS2*	**0.19**	**0.05**	++	30	**-**	8.2.D
YDR202C	29	13	*RAV2,*	**0.74**	**0.85**	++++	30	l**+/-**	14.5.E
YDR204W	19	2	*COQ4*	**0.73**	**0.76**	+	30	**-**	8.2.E
YDR225W	19	3	*HTA1*	**0.89**	**0.91**	++	30	l**+/-**	8.2.G
YDR226W	19	4	*ADK1*	**0.90**	**0.94**	++	30	**+/-**	8.2.H
YDR230W	2	4		**0.75**	**0.86**	+/-	30	**-**	8.3.a
YDR231C	2	5	*COX20*	**0.85**	**0.82**	++	30	**-**	8.3.b
YDR269C	23	8		**0.75**	**0.77**	+++	30	l**+/-**	10.3.H
YDR270W	8	11	*CCC2*	**0.93**	**0.93**	++	***37***	t^s**+/-**	10.4.a
YDR271C	2	8	*YDR271C*	**0.69**	**0.52**	+++	30	**+/-**	10.4.b
YDR276C	23	9	*PMP3*	**0.85**	**0.84**	+	30	**+/-**	10.4.C
YDR293C	23	10	*SSD1*	**0.83**	**0.78**	+	30	l**+/-**	10.4.D
YDR316W	23	11		**0.83**	**0.83**	+	30	**+/-**	10.4.G
YDR323C	24	1	*PEP7*	**0.71**	**0.86**	++	30	l**+/-**	10.5.A
YDR332W	24	2		**0.73**	**0.88**	+	30	l**+/-**	10.5.B
YDR335W	24	3	*MSN5*	**0.73**	**0.89**	+++	30	l**+/-**	10.5.C
YDR336W	24	4		**0.74**	**0.88**	+++	30	l**+/-**	10.5.D
YDR337W	24	5	*MRPS28*	**0.70**	**0.85**	+++	30	l**+/-**	10.5.E
YDR350C	15	3	*TCM10*	**0.59**	**0.99**	++	30	**-**	4.2.d
YDR350C	20	8	*TCM10*	**0.00**	**0.00**	++	30	**-**	4.2.d
YDR375C	15	4	*BCS1*	**0.89**	**0.88**	++	30	**-**	4.2.e
YDR388W	15	5	*RVS167*	**0.96**	**0.82**	++++	30	l**+/-**	4.2.g
YDR393W	15	6	*SHE9*	**0.95**	**0.82**	+++	30	l**+/-**	4.2.h
YDR448W	30	11	*ADA2*	**0.78**	**0.82**	++	30	l**+/-**	15.5.H
YDR455C	30	12		**0.80**	**0.82**	++	30	l**+/-**	15.6.A
YDR456W	30	13	*NHX1*	**0.90**	**0.85**	++	30	l**+/-**	15.6.B
YDR462W	30	14	*MRPL28*	**0.89**	**0.91**	++	30	l**+/-**	15.6.C
YDR470C	37	15	*UGO1*	**0.73**	**0.91**	++	30	**-**	19.4.F
YDR477W	32	3	*SNF1*	**0.74**	**0.87**	+++	30	**+/-**	16.5.C
YDR484W	30	15	*SAC2*	**0.77**	**0.82**	++	30	l**+/-**	15.6.E

ORF name	blot	lane	Gene name	COX % splicing	COB % splicing	rel. total RNA	temp (°C)	Glycerol growth	library grid location
YDR493W	3	13	*FMP36*	0.77	0.69	++	30	+/-	16.5.f
YDR494W	31	2		0.80	0.82	++	30	l+/-	15.6.F
YDR495C	30	16	*VPS3*	0.79	0.90	+++	30	l+/-	15.6.G
YDR509W	31	3		0.83	0.86	++	30	l+/-	16.1.A
YDR511W	31	4	*ACN9*	0.84	0.86	++	30	l+/-	16.1.B
YDR512C	31	5	*EMI1*	0.72	0.75	++	30	+/-	16.1.C
YDR512C	37	12	*EMI1*	0.36	0.46	+++	30	+/-	18.2.F
YDR519W	31	6	*FKB2*	0.84	0.84	++	30	l+/-	16.1.F
YDR522C	31	7	*SPS2*	0.86	0.83	++	30	l+/-	16.1.G
YDR523C	3	18	*SPS1*	0.27	0.67	+/-	30	-	19.5.c
YDR529C	31	8	*QCR7*	0.75	0.91	++	30	-	16.1.H
YDR533C	31	9		0.89	0.83	++	30	l+/-	16.2.A
YEL024W	15	7	*RIP1*	0.93	0.68	+++	30	-	4.3.d
YEL027W	15	8	*CUP5*	0.96	0.76	++	30	+/-	4.3.e
YEL027W	20	9	*CUP5*	0.65	0.92	+	30	+/-	4.3.e
YEL029C	34	12	*BUD16*	0.82	0.82	+++	30	l+/-	17.6.H
YEL051W	15	9	*VMA8*	0.95	0.79	++	30	+/-	4.3.g
YEL051W	20	10	*VMA8*	0.72	0.88	+++	30	+/-	4.3.g
YER014C-A	35	6	*BUD25*	0.68	0.72	+++	30	-	18.2.A
YER028C	32	4		0.78	0.82	++	30	l+/-	16.5.E
YER058W	15	10	*PET117*	0.97	0.63	++	30	-	4.4.e
YER058W	21	1	*PET117*	0.83	0.70	++	30	-	4.4.e
YER061C	15	11	*CEM1*	0.94	0.47	++	30	+/-	4.4.f
YER061C	cand 2	2	*CEM1*	0.22	0.44	+++	30	+/-	4.4.F
YER068W	15	12	*MOT2*	0.95	0.80	+++	30	+/-	4.4.g
YER077C	25	7		0.74	0.84	++	30	l+/-	11.3.C
YER083C	1	7	*RMD7*	0.76	0.68	++	30	+/-	4.5.c
YER087C-A	15	13		0.97	0.87	++++	30	l+/-	4.5.e
YER087W	3	19		0.77	0.64	+	30	-	19.6.b
YER093C-A	2	9		0.47	0.34	++	30	+/-	11.3.e
YER093C-A	4	11		0.52	0.35	+++	30	+/-	11.3.e
YER093C-A	cand 2	13		0.81	0.77	+++	30	+/-	11.3.E
YER141W	31	14	*COX15*	0.86	0.81	++	30	l+/-	16.3.C
YER145C	31	15	*FTR1*	0.75	0.75	+++	30	+/-	16.3.D
YER153C	4	13	*PET122*	0.77	0.52	+++	30	-	16.3.e
YER153C	cand 2	1	*PET122*	0.71	0.71	+++	30	-	16.3.E
YER178W	31	16	*PDA1*	0.65	0.67	++	30	+/-	16.3.H
YFL018C	33	7	*LPD1*	0.72	0.79	+++	30	+/-	17.3.B

ORF name	blot	lane	Gene name	COX % splicing	COB % splicing	rel. total RNA	temp (°C)	Glycerol growth	library grid location
YFR019W	37	3	*FAB1*	0.59	0.70	++++	30	l+/-	15.3.A
YFR033C	10	2	*QCR6*	0.73	0.65	++	***37***	t^s-	9.3.a
YFR034C	36	15	*PHO4*	0.57	0.62	+++	30	l+/-	9.3.B
YGL020C	33	4		0.78	0.78	+++	30	l+/-	17.2.A
YGL025C	22	11	*PGD1*	1.00	0.96	+	30	l+/-	9.6.B
YGL038C	33	5	*OCH1*	0.92	0.94	+++	30	-	17.2.B
YGL057C	22	12		0.77	0.77	+	30	l+/-	9.6.C
YGL058W	22	13	*RAD6*	0.97	0.96	+	30	l+/-	9.6.D
YGL070C	10	15	*RPB9*	0.81	0.86	+	***37***	t^s-	17.2.c
YGL071W	22	14	*RCS1*	0.98	0.99	+	30	+/-	9.6.G
YGL084C	23	1	*GUP1*	0.77	0.83	++	30	+/-	10.1.A
YGL095C	10	16	*VPS45*	0.76	0.86	+	***37***	-	17.2.d
YGL107C	31	10	*RMD9*	0.00	0.00	+	30	+/-	16.2.B
YGL115W	31	11	*SNF4*	0.69	0.87	+++	30	+/-	16.2.C
YGL135W	31	12	*RPL1B*	0.75	0.80	+	30	l+/-	16.2.E
YGL167C	31	13	*PMR1*	0.79	0.81	+++	30	+/-	16.2.G
YGL168W	4	9	*HUR1*	0.73	0.50	++	30	+/-	16.3.a
YGL206C	32	8	*CHC1*	0.81	0.91	+/-	30	+/-	16.6.E
YGL211W	5	4		0.81	0.75	+	30	+	8.3.e
YGL218W	19	5	*SRF1*	0.95	0.96	+	30	+/-	8.3.F
YGL218W	37	11	*SRF1*	0.60	0.89	+++	30	+/-	17.4.E
YGL219C	3	3	*MDM34*	0.69	0.37	+/-	30	+/-	14.5.f
YGL221C	5	5	*NIF3*	0.84	0.76	++	30	+	8.3.g
YGL223C	32	9	*COG1*	0.69	0.81	+++	30	+/-	16.6.F
YGL236C	19	6	*MTO1*	0.93	1.00	+	30	l+/-	8.4.A
YGL250W	19	8		0.94	0.90	+++	30	l+/-	8.4.C
YGR020C	19	9	*VMA7,*	0.92	0.72	+++	30	+/-	8.4.G
YGR036C	17	11	*CWH8,*	0.87	0.79	++	30	+/-	6.4.f
YGR061C	37	13	*ADE6*	0.62	0.71	+++	30	+/-	19.2.C
YGR062C	1	17	*COX18*	0.68	0.67	+	30	-	6.4.h
YGR062C	4	8	*COX18*	0.87	0.55	+++	30	-	6.4.h
YGR084C	17	12	*MRP13*	0.91	0.80	++	30	l+/-	6.5.b
YGR105W	36	4	*VMA21*	0.56	0.57	++++	30	+/-	6.5.f
YGR112W	36	5	*SHY1*	0.55	0.60	+++	30	-	6.5.g
YGR155W	10	12	*CYS4*	0.51	0.64	++	***37***	t^s-	15.1.c
YGR167W	35	3	*CLC1*	0.68	0.76	++	30	+/-	18.1.E
YGR174C	15	14	*CBP4*	0.93	0.69	++	30	+/-	4.6.b
YGR178C	15	15	*PBP1*	0.98	0.87	++++	30	l+/-	4.6.c

ORF name	blot	lane	Gene name	COX % splicing	COB % splicing	rel. total RNA	temp (°C)	Glycerol growth	library grid location
YGR183C	8	3	*QCR9*	**0.98**	**0.93**	++	***37***	t^s**+/-**	4.6.e
YGR183C	15	16	*QCR9*	**0.94**	**0.80**	++	30	t^s**+/-**	4.6.e
YGR188C	26	1	*BUB1*	**0.71**	**0.82**	++	30	**l+/-**	12.2.B
YGR222W	4	16	*PET54*	**0.27**	**0.71**	+	30	**-**	17.1.f
YGR222W	cand 1	14	*PET54*	**0.03**	**0.79**	++	30	**-**	17.1.F
YGR229C	10	14	*SMI1*	**0.70**	**0.66**	++	***37***	t^s**-**	17.1.g
YGR252W	35	9	*GCN5*	**0.84**	**0.75**	++	30	**+/-**	18.2.G
YGR255C	35	10	*COQ6*	**0.76**	**0.70**	++	30	**-**	18.2.H
YGR262C	33	8	*BUD32*	**0.84**	**0.77**	+++	30	**+/-**	17.3.C
YHL005C	1	9	*(MRP4)*	**0.80**	**0.72**	++	30	c^s**+/-**	5.1.f
YHL005C	9	8	*(MRP4)*	**0.05**	**0.17**	++	***20***	c^s**+/-**	5.1.f
YHL005C	cand 2	11	*(MRP4)*	**0.48**	**0.34**	+++	***20***	c^s**+/-**	5.1.F
YHL025W	25	8	*SNF6*	**0.86**	**0.91**	++	30	**+/-**	11.3.F
YHL038C	1	8	*CBP2*	**0.61**	**0.53**	+	30	**-**	5.1.a
YHL038C	4	5	*CBP2*	**0.78**	**0.02**	++	30	**-**	5.1.a
YHL038C	cand 1	5	*CBP2*	**0.24**	**0.24**	++	30	**-**	5.1.A
YHR008C	26	2	*SOD2*	**0.91**	**0.90**	+	30	**l+/-**	12.2.F
YHR009C	37	1		**0.71**	**0.62**	++++	30	**l+/-**	12.2.G
YHR013C	9	9	*ARD1*	**0.68**	**0.68**	+++	***20***	c^s**-**	5.2.a
YHR025W	26	3	*THR1*	**0.70**	**0.83**	++	30	**l+/-**	12.2.H
YHR026W	26	4	*PPA1*	**0.67**	**0.78**	+++	30	**+/-**	12.3.A
YHR037W	9	10	*PUT2*	**0.53**	**0.66**	+	***20***	c^s**-**	5.2.b
YHR038W	16	1	*RRF1*	**0.94**	**0.90**	+++	30	**l+/-**	5.2.c
YHR039C-B	25	1	*VMA10*	**0.88**	**0.83**	++	30	**+/-**	11.1.D
YHR041C	26	5	*SRB2*	**0.78**	**0.84**	+++	30	**l+/-**	12.3.B
YHR051W	16	2	*COX6*	**0.92**	**0.92**	++	30	**-**	5.2.d
YHR059W	26	6	*FYV4*	**0.74**	**0.83**	+++	30	**l+/-**	12.3.C
YHR060W	16	3	*VMA22*	**0.00**	**0.46**	+	30	**-**	5.2.e
YHR060W	21	2	*VMA22*	**0.76**	**0.70**	+++	30	**-**	5.2.e
YHR067W	26	7	*RMD12*	**0.77**	**0.80**	++	30	**+/-**	12.3.D
YHR075C	27	5	MRPS2	**0.87**	**0.84**		30	**+**	MRPS2
YHR100C	1	10		**0.75**	**0.67**	+	30	t^s**+/-**	5.2.g
YHR100C	4	14		**0.86**	**0.81**	++	30	t^s**+/-**	5.2.g
YHR100C	8	4		**0.00**	**0.00**	+	***37***	t^s**+/-**	5.2.g
YHR116W	1	11		**0.73**	**0.68**	+++	30	**-**	5.3.a
YHR117W	9	11	*TOM71*	**0.58**	**0.67**	+	***20***	c^s**+/-**	5.3.b
YHR155W	6	10		**0.92**	**0.89**	+++	30	c^s**+/-**	5.3.e
YHR155W	7	15		**0.67**	**0.68**	+++	***20***	c^s**+/-**	5.3.e

ORF name	blot	lane	Gene name	COX % splicing	COB % splicing	rel. total RNA	temp (°C)	Glycerol growth	library grid location
YHR179W	9	12	*OYE2*	**0.11**	**0.50**	+	***20***	c^s**+/-**	5.3.f
YHR179W	cand 2	7	*OYE2*	**0.08**	**0.53**	+++	***20***	c^s**+/-**	5.3.F
YHR189W	16	4	*PTH1*	**0.97**	**0.87**	++	30	**I+/-**	5.3.g
YHR194W	26	8		**0.75**	**0.82**	++	30	**I+/-**	12.3.E
YIL015C-A	24	8		**0.77**	**0.88**	++	30	**I+/-**	10.6.B
YIL018W	33	6	*RPL2B,*	**0.82**	**0.82**	+++	30	**I+/-**	17.2.H
YIL036W	24	6	*CST6*	**0.77**	**0.88**	++	30	**I+/-**	10.5.F
YIL047C	5	7	*SYG1*	**0.83**	**0.75**	++	30	**I+/-**	10.6.d
YIL060W	24	9		**0.80**	**0.93**	+++	30	**I+/-**	10.6.E
YIL070C	24	10	*MAM33*	**0.73**	**0.87**	++	30	**I+/-**	10.6.F
YIL093C	24	7	*RSM25*	**0.74**	**0.87**	++	30	**I+/-**	10.6.A
YIL098C	34	7	*FMC1*	**0.80**	**0.79**	+++	30	**I+/-**	17.5.h
YIL125W	29	5	*KGD1*	**0.75**	**0.89**	+	30	**I+/-**	14.3.A
YIL128W	2	18	*MET18*	**0.57**			30	**+/-**	14.3.B
YIL154C	29	6	*(IMP2)*	**0.69**	**0.92**	+	30	**I+/-**	14.3.C
YIL155C	37	8	*GUT2*	**0.68**	**0.71**	+++	30	**+/-**	14.3.D
YIL157C	29	7		**0.79**	**0.94**	+	30	**I+/-**	14.3.E
YIR005W	29	8	*IST3*	**0.91**	**0.88**	+++	30	**I+/-**	14.3.F
YIR009W	29	9	*MSL1*	**0.78**	**0.88**	+	30	**I+/-**	14.3.G
YIR021W	25	2	*MRS1*	**0.72**	**0.84**	++	30	**I+/-**	11.1.E
YIR024C	25	3	*GIF1*	**0.72**	**0.83**	++	30	**I+/-**	11.1.F
YJL003W	34	3		**0.72**	**0.85**	+++	30	**I+/-**	17.5.A
YJL004C	34	4	*SYS1*	**0.70**	**0.80**	+++	30	**I+/-**	17.5.B
YJL022W	28	3		**0.77**	**0.85**	++	30	**I+/-**	13.2.A
YJL029C	28	4	*VPS53*	**0.68**	**0.85**	++	30	**I+/-**	13.2.C
YJL046W	2	13	*YJL046W*	**0.41**	**0.50**	++	30	**+/-**	12.1.h
YJL075C	34	10		**0.77**	**0.86**	+++	30	**I+/-**	17.6.F
YJL094C	30	4	*KHA1*	**0.81**	**0.88**	++++	30	**I+/-**	15.3.B
YJL095W	8	13	*BCK1*	**0.94**	**0.93**	+++	***37***	t^s**+/-**	12.1.b
YJL101C	30	5	*GSH1*	**0.80**	**0.83**	+++	30	**+/-**	15.3.C
YJL117W	25	18	*PHO86*	**0.51**	**0.83**	++	30	**I+/-**	11.6.G
YJL120W	25	16		**0.62**	**0.86**	+++	30	**I+/-**	11.6.E
YJL121C	25	17	*RPE1*	**0.59**	**0.86**	+++	30	**I+/-**	11.6.F
YJL133W	5	3	*MRS3*	**0.81**	**0.73**	+++	30	**+**	7.3.f
YJL133W	cand 1	16	*MRS3*	**0.53**	**0.87**	+++	30	**+**	7.3.F
YJL140W	37	14	*RPB4*	**0.70**	**0.90**	++	30	**+/-**	19.2.D
YJL155C	36	7	*FBP26*	**0.60**	**0.70**	+++	30	**I+/-**	7.3.d
YJL166W	18	5	*QCR8*	**0.81**	**0.79**	+++	30	**-**	7.3.b

ORF name	blot	lane	Gene name	COX % splicing	COB % splicing	rel. total RNA	temp (°C)	Glycerol growth	library grid location
YJL175W	29	14		0.81	0.88	+++	30	l+/-	14.5.H
YJL176C	18	2	*SWI3*	0.87	0.83	++++	30	+/-	7.1.g
YJL180C	18	4	*ATP12*	0.00	0.00	++	30	-	7.2.h
YJL180C	cand 2	4	*ATP12*	0.73	0.41	+++	30	-	7.2.H
YJL180C	cand 2	5	*ATP12* ω-	0.37	0.27	++	30	-	7.2.H
YJL193W	18	3		0.90	0.89	+++	30	l+/-	7.2.g
YJL209W	36	6	*CBP1*	0.09	0.00	++	30	-	7.2.d
YJR018W	6	3		0.88	0.87	++	30	t^s-	13.2.f
YJR018W	10	7		0.66	0.64	++	***37***	t^s-	13.2.f
YJR033C	28	5	*RAV1*	0.75	0.82	+++	30	l+/-	13.2.D
YJR034W	31	18	*PET191*	0.67	0.77	++	30	l+/-	16.4.E
YJR040W	34	5	*GEF1*	0.79	0.80	+++	30	l+/-	17.5.C
YJR048W	28	6	*CYC1*	0.80	0.84	++	30	+/-	13.2.G
YJR074W	10	5	*MOG1*	0.62	0.72	+++	***37***	t^s+/-	12.5.c
YJR077C	1	18	*MIR1*	0.65	0.57	+/-	30	-	12.5.d
YJR077C	4	21	*MIR1*	0.80	0.53	++	30	-	12.5.d
YJR077C	cand 2	14	*MIR1*	0.91	0.75	+	30	-	12.5.D
YJR079W	26	9		0.64	0.71	+	30	l+/-	12.3.G
YJR090C	34	1	*GRR1*	0.77	0.87	+++	30	-	17.4.F
YJR102C	10	3	*VPS25*	0.67	0.77	++	***37***	t^s+/-	12.3.h
YJR118C	26	18		0.57	0.68	+	30	l+/-	12.5.G
YJR120W	2	14		0.55	0.87	+	30	-	12.5.h
YJR120W	4	20		0.67	0.56	++	30	-	12.5.h
YJR121W	27	1	*ATP2*	0.87	0.82		30	l+/-	12.6.A
YKL002W	16	10	*DID4*	0.86	0.69	++	30	-	5.6.d
YKL002W	21	3	*DID4*	0.58	0.58	++	30	-	5.6.d
YKL016C	16	11	*ATP7*	0.96	0.87	+	30	-	5.6.g
YKL016C	21	4	*ATP7*	0.00	0.00	++	30	-	5.6.g
YKL016C	27	4	*ATP7*	0.73	0.74		30	-	5.6.g
YKL037W	5	14		0.88	0.70	+++	30	t^s+/-	6.1.b
YKL037W	7	8		0.88	0.66	++	***37***	t^s+/-	6.1.b
YKL041W	8	6	*VPS24*	0.90	0.91	+++	***37***	t^s+/-	6.1.d
YKL053C-A	3	5	*MDM35*	0.87	0.05	++	30	+/-	15.2.b
YKL053C-A	4	22	*MDM35*		0.82	++	30	+/-	15.2.b
YKL054C	10	1	*VID31*	0.81	0.74	++	***37***	t^s+/-	6.1.e
YKL055C	17	1	*OAR1*	0.95	0.94	+++	30	l+/-	6.1.g
YKL080W	17	2	*VMA5*	0.90	0.90	+	30	+/-	6.1.h
YKL080W	21	5	*VMA5*	0.69	0.68	+++	30	+/-	6.1.h

ORF name	blot	lane	Gene name	COX % splicing	COB % splicing	rel. total RNA	temp (°C)	Glycerol growth	library grid location
YKL085W	17	3	*MDH1*	**0.95**	**0.85**	++++	30	**+/-**	6.2.a
YKL087C	17	4	*CYT2*	**0.82**	**0.65**	++	30	**-**	6.2.b
YKL087C	21	6	*CYT2*	**0.82**	**0.60**	+	30	**-**	6.2.b
YKL106W	17	5	*AAT1*	**0.85**	**0.20**	++++	30	**+/-**	6.2.c
YKL106W	21	7	*AAT1*	**0.69**	**0.19**	+++	30	**+/-**	6.2.c
YKL106W	cand 1	1	*AAT1*	**0.53**	**0.18**	+++	30	**+/-**	6.2.C
YKL109W	17	6	*HAP4*	**0.81**	**0.53**	+	30	**-**	6.2.d
YKL109W	21	8	*HAP4*	**0.68**	**0.54**	++	30	**-**	6.2.d
YKL118W	17	7		**0.98**	**0.78**	++++	30	**+/-**	6.2.f
YKL119C	17	8	*VPH2*	**0.92**	**0.73**	++	30	**+/-**	6.2.g
YKL119C	21	9	*VPH2*	**0.80**	**0.78**	++++	30	**+/-**	6.2.g
YKL137W	1	15		**0.75**	**0.70**	++	30	**+/-**	6.3.b
YKL138C	17	9	*MRPL31*	**0.92**	**0.89**	+++	30	**+/-**	6.3.c
YKL142W	27	6	MRP8	**0.68**	**0.77**		30	**+**	MRP8
YKL148C	17	10	*SDH1*	**0.91**	**0.82**	++	30	**+/-**	6.3.d
YKL167C	1	16	*MRP49*	**0.77**	**0.66**	++	30	**+/-**	6.3.h
YKL208W	23	3	*CBT1*	**0.72**	**0.23**	+	30	**-**	10.2.F
YKL212W	23	4	*SAC1*	**0.78**	**0.82**	+++	30	**+/-**	10.2.G
YKR001C	23	5	*VPS1*	**0.83**	**0.82**	++	30	**+/-**	10.2.H
YKR006C	10	17	*MRPL13*	**0.78**	**0.69**	++	***37***	**-**	17.2.e
YKR016W	2	7	*FMP13*	**0.76**	**0.90**	++	30	**+/-**	10.3.a
YKR021W	23	6		**0.83**	**0.84**	++	30	**+/-**	10.3.B
YKR036C	29	15	*CAF4*	**0.74**	**0.88**	++	30	**+**	14.6.B
YKR049C	23	7		**0.82**	**0.83**	++	30	**+**	10.3.D
YKR087C	26	10		**0.75**	**0.83**	++	30	**+**	12.4.B
YLL006W	11	4	*MMM1*	**0.79**	**0.91**	++	30	**+/-**	1.2.a
YLL009C	11	5	*COX17*	**0.83**	**0.91**	+++	30	**+/-**	1.2.b
YLL018C-A	26	11	*COX19*	**0.54**	**0.64**	+	30	**-**	12.4.C
YLL027W	35	2	*ISA1*	**0.64**	**0.82**	+++	30	**+/-**	18.1.A
YLL033W	11	6		**0.78**	**0.90**	++	30	**+/-**	1.2.c
YLL040C	11	7	*VPS13*	**0.68**	**0.89**	++	30	**+/-**	1.2.d
YLL041C	9	2	*SDH2*	**0.60**	**0.69**	+	***20***	**c^s-**	1.2.e
YLR025W	11	8	*SNF7*	**0.67**	**0.86**	+++	30	**+/-**	1.2.f
YLR027C	9	3	*AAT2*	**0.81**	**0.69**	++	***20***	**c^s+/-**	1.2.g
YLR038C	25	9	*COX12*	**0.86**	**0.88**	++	30	**-**	11.4.A
YLR056W	11	9	*ERG3*	**0.71**	**0.91**	+	30	**+/-**	1.2.h
YLR059C	11	10	*REX2*	**0.77**	**0.88**	++	30	**+/-**	1.3.a
YLR077W	1	2	*FMP25*	**0.86**	**0.64**	+	30	**+/-**	1.3.e

ORF name	blot	lane	Gene name	COX % splicing	COB % splicing	rel. total RNA	temp (°C)	Glycerol growth	library grid location
YLR083C	11	11	*EMP70*	**0.76**	**0.90**	++	30	l**+/-**	1.3.f
YLR087C	11	12	*CSF1*	**0.72**	**0.87**	+++	30	l**+/-**	1.3.g
YLR090W	6	6	*XDJ1*	**0.88**	**0.90**	+++	30	c^s**+/-**	1.3.h
YLR090W	7	11	*XDJ1*	**0.70**	**0.61**	+++	***20***	c^s**+/-**	1.3.h
YLR148W	35	5	*PEP3*	**0.65**	**0.78**	+++	30	**+/-**	18.1.G
YLR149C	16	6		**0.94**	**0.87**	+++	30	l**+/-**	5.5.b
YLR168C	16	7	*(MSF1)*	**0.91**	**0.82**	++	30	l**+/-**	5.5.c
YLR182W	16	8	*SWI6*	**0.93**	**0.80**	+++	30	l**+/-**	5.5.d
YLR201C	1	12	*FMP53*	**0.86**	**0.59**	+	30	**-**	5.5.g
YLR201C	4	1	*FMP53*	**0.92**	**0.85**	+	30	**-**	5.5.g
YLR202C	1	13		**0.76**	**0.73**	++	30	**-**	5.5.h
YLR202C	4	2		**0.89**	**0.66**	+	30	**-**	5.5.h
YLR203C	1	14	*MSS51*	**0.51**	**0.59**	+	30	**-**	5.6.a
YLR203C	4	10	*MSS51*	**0.71**	**0.12**	++	30	**-**	5.6.a
YLR203C	cand 1	12	*MSS51*	**0.51**	**0.74**	++	30	**-**	5.6.A
YLR218C	16	9		**0.94**	**0.85**	+++	30	l**+/-**	5.6.c
YLR232W	36	9		**0.63**	**0.79**	+++	30	l**+/-**	7.5.a
YLR234W	18	10	*TOP3*	**0.86**	**0.83**	++	30	l**+/-**	7.5.b
YLR235C	36	10		**0.65**	**0.80**	++	30	l**+/-**	7.5.C
YLR239C	18	11	*LIP2*	**0.79**	**0.74**	++	30	**+/-**	7.5.D
YLR240W	32	7	*VPS34*	**0.71**	**0.80**	++	30	**-**	16.6.D
YLR262C	18	12	*YPT6*	**0.76**	**0.74**	++	30	l**+/-**	7.5.G
YLR268W	18	13	*SEC22*	**0.83**	**0.81**	++	30	l**+/-**	7.5.H
YLR294C	18	14		**0.84**	**0.77**	++	30	l**+/-**	7.6.F
YLR295C	18	15	*ATP14*	**0.90**	**0.84**	++++	30	l**+/-**	7.6.G
YLR304C	18	16	*ACO1*	**0.84**	**0.79**	++	30	l**+/-**	7.6.H
YLR338W	32	6	*KRE21*	**0.86**	**0.84**	++	30	l**+/-**	16.6.B
YLR346C	5	8		**0.88**	**0.75**	+++	30	**+**	15.1.d
YLR358C	10	13		**0.77**	**0.60**	+++	***37***	t^s**-**	15.1.e
YLR368W	36	8	*MDM30*	**0.52**	**0.00**	+	30	**+/-**	7.4.c
YLR377C	18	6	*FBP1*	**0.87**	**0.83**	+++	30	**+/-**	7.4.d
YLR393W	18	7	*ATP10*	**0.50**	**0.32**	+++	30	**+/-**	7.4.f
YLR393W	cand 2	3	*ATP10*	**0.12**	**0.17**	+++	30	**+/-**	7.4.F
YLR395C	18	8	*COX8*	**0.89**	**0.83**	++++	30	l**+/-**	7.4.g
YLR396C	32	10	*VPS33*	**0.65**	**0.80**	+++	30	**+/-**	16.6.G
YLR417W	18	9	*VPS36*	**0.88**	**0.80**	++++	30	l**+/-**	7.4.h
YLR423C	26	12	*APG17*	**0.69**	**0.75**	++	30	l**+/-**	12.4.D
YLR425W	26	13	*TUS1*	**0.72**	**0.77**	+++	30	l**+/-**	12.4.E

ORF name	blot	lane	Gene name	COX % splicing	COB % splicing	rel. total RNA	temp (°C)	Glycerol growth	library grid location
YLR439W	30	6	*MRPL4*	**0.00**	**0.00**	+	30	**-**	15.3.D
YLR447C	26	14	*VMA6*	**0.58**	**0.69**	+++	30	**+/-**	12.4.F
YML009c	6	7	*MRPL39*	**0.83**	**0.91**	+++	30	**c^s+/-**	1.5.a
YML009c	7	12	*MRPL39*	**0.55**	**0.49**	+++	***20***	**c^s+/-**	1.5.a
YML009C	26	15	*MRPL39*	**0.74**	**0.82**	++	30	**c^s+/-**	12.4.G
YML009c	cand 2	8	*MRPL39*	**0.26**	**0.54**	+++	***20***	**c^s+/-**	1.5.A
YML030W	11	14		**0.71**	**0.88**	++	30	**I+/-**	1.4.g
YML061C	5	11	*PIF1*	**0.84**	**0.54**	+++	30	**t^s+/-**	1.4.f
YML061C	7	2	*PIF1*	**0.44**	**0.16**	++	***37***	**t^s+/-**	1.4.f
YML061C	cand 2	10	*PIF1*	**0.48**	**0.27**	+	***37***	**t^s+/-**	1.4.F
YML081C-A	26	16	*ATP18*	**0.59**	**0.67**	++	30	**I+/-**	12.4.H
YML089C	11	13		**0.67**	**0.76**	++	30	**I+/-**	1.4.c
YML090W	4	17	*(RPM2)*	**0.64**	**0.26**	++	30	**+/-**	11.4.g
YML090W	cand 2	15	*(RPM2)*	**0.21**	**0.30**	++	30	**+/-**	11.4.G
YML094W	25	11	*GIM5*	**0.80**	**0.91**	++	30	**I+/-**	11.4.H
YML110C	33	9	*COQ5*	**0.86**	**0.96**	+++	30	**-**	17.3.F
YML112W	33	10	*CTK3*	**0.76**	**0.82**	+++	30	**I+/-**	17.3.G
YML120C	25	12	*NDI1*	**0.73**	**0.91**	++	30	**I+/-**	11.5.A
YML129C	33	11	*COX14*	**0.69**	**0.80**	+++	30	**I+/-**	17.3.H
YMR021C	11	15	*MAC1*	**0.73**	**0.78**	+++	30	**-**	1.5.c
YMR021C	20	2	*MAC1*	**0.80**	**0.91**	++	30	**-**	1.5.c
YMR024W	1	3	*MRPL3*	**0.80**	**0.61**	+	30	**+/-**	1.5.e
YMR031W-A	11	16		**0.65**	**0.82**	++	30	**I+/-**	1.5.f
YMR035W	11	17	*IMP2*	**0.62**	**0.85**	++	30	**-**	1.5.g
YMR054W	31	17	*STV1*	**0.70**	**0.76**	++	30	**I+/-**	16.4.B
YMR058W	9	15	*FET3*	**0.68**	**0.72**	++	***20***	**c^s+/-**	16.4.d
YMR060C	10	4	*TOM37*	**0.70**	**0.84**	+	***37***	**t^s+/-**	12.5.a
YMR063W	25	4	*RIM9*	**0.33**	**0.42**	+++	30	**+/-**	11.2.B
YMR077C	8	12	*VPS20*	**0.95**	**0.92**	++	***37***	**+/-**	11.2.e
YMR123W	25	13	*PKR1*	**0.72**	**0.88**	+++	30	**I+/-**	11.5.C
YMR138W	9	14	*CIN4*	**0.78**	**0.75**	+	***20***	**c^s+/-**	15.1.g
YMR150C	12	2	*IMP1*	**0.23**	**0.76**	++	30	**-**	1.6.b
YMR150C	20	3	*IMP1*	**0.22**	**0.92**	++	30	**-**	1.6.b
YMR150C	cand 1	3	*IMP1*	**0.27**	**0.92**	++	30	**-**	1.6.B
YMR151W	12	1	*YIM2*	**0.65**	**0.86**	+++	30	**I+/-**	1.6.a
YMR167W	12	3	*MLH1*	**0.72**	**0.80**	++	30	**I+/-**	1.6.e
YMR188C	12	4	*MRPS17*	**0.88**	**0.91**	+++	30	**I+/-**	1.6.g
YMR193W	12	5	*MRPL24*	**0.79**	**0.89**	++	30	**I+/-**	2.1.a

ORF name	blot	lane	Gene name	COX % splicing	COB % splicing	rel. total RNA	temp (°C)	Glycerol growth	library grid location
YMR202W	12	6	*ERG2*	**0.88**	**0.91**	+++	30	l**+/-**	2.1.c
YMR207C	12	7	*HFA1*	**0.78**	**0.89**	+++	30	l**+/-**	2.1.d
YMR245W	12	8		**0.68**	**0.92**	++	30	l**+/-**	2.1.h
YMR256C	12	9	*COX7*	**0.41**	**0.93**	++	30	**-**	2.2.b
YMR256C	20	4	*COX7*	**0.75**	**0.90**	++	30	**-**	2.2.b
YMR257C	12	10	*PET111*	**0.54**	**0.88**	++	30	**-**	2.2.c
YMR257C	20	5	*PET111*	**0.61**	**0.91**	++	30	**-**	2.2.c
YMR257C	cand 1	10	*PET111*	**0.49**	**0.87**	++	30	**-**	2.2.C
YMR263W	35	12	*SAP30*	**0.26**	**0.26**	++++	30	**+/-**	18.3.E
YMR275C	5	10	*BUL1*	**0.88**	**0.71**	++	30	t^s**+/-**	2.2.e
YMR275C	7	1	*BUL1*	**0.75**	**0.67**	+	***37***	t^s**+/-**	2.2.e
YMR275C	9	4	*BUL1*	**0.77**	**0.71**	++	***20***	t^s**+/-**	2.2.e
YMR280C	12	11	*CAT8*	**0.66**	**0.92**	++	30	**+/-**	2.2.f
YMR282C	12	12	*AEP2*	**0.74**	**0.92**	+++	30	l**+/-**	2.2.g
YMR285C	12	13	*NGL2*	**0.64**	**0.89**	++	30	l**+/-**	2.2.h
YMR307W	37	5	*GAS1*	**0.70**	**0.68**	++++	30	**+/-**	2.3.e
YMR311C	25	10	*GLC8*	**0.84**	**0.92**	++	30	l**+/-**	11.4.C
YNL003C	9	16	*PET8*	**0.67**	**0.73**	++	***20***	c^s**+/-**	17.5.f
YNL037C	37	7	*IDH1*	**0.53**	**0.57**	++++	30	**+/-**	13.4.G
YNL052W	29	10	*COX5A*	**0.67**	**0.89**	++	30	**-**	14.4.A
YNL055C	6	1	*POR1*	**0.85**	**0.78**	+++	30	t^s**+/-**	18.1.d
YNL055C	7	9	*POR1*	**0.88**	**0.61**	++	***37***	t^s**+/-**	18.1.d
YNL064C	29	4	*YDJ1*	**0.68**	**0.82**	+++	30	l**+/-**	14.2.D
YNL067W	2	19	*RPL9B*	**0.72**			30	**+/-**	14.4.B
YNL070W	2	20	*TOM7*	**0.69**			30	**+/-**	14.4.C
YNL071W	29	11	*LAT1*	**0.80**	**0.85**	++	30	**+/-**	14.4.D
YNL080C	3	1		**0.57**	**0.62**	++	30	**+/-**	14.4.f
YNL080C	4	12		**0.77**	**0.54**	+	30	**+/-**	14.4.f
YNL097C	29	12	*PHO23*	**0.82**	**0.90**	+++	30	**+/-**	14.4.H
YNL117W	29	3	*MLS1*	**0.59**	**0.78**	++	30	**+/-**	14.2.B
YNL120C	3	2		**0.72**	**0.67**	+++	30	**+/-**	14.5.c
YNL133C	10	11	*FYV6*	**0.82**	**0.79**	+++	***37***	t^s**+/-**	14.1.h
YNL136W	10	10	*EAF7*	**0.47**	**0.53**	++	***37***	t^s**+/-**	14.1.g
YNL136W	cand 2	12	*EAF7*	**0.65**	**0.69**	+	***37***	t^s**+/-**	14.1.G
YNL138W	30	2	*SRV2*	**0.82**	**0.86**	+	30	l**+/-**	15.2.C
YNL139C	6	12	*RLR1*	**0.93**	**0.89**	+++	30	c^s**+/-**	14.1.f
YNL139C	7	17	*RLR1*	**0.89**	**0.68**	+++	***20***	c^s**+/-**	14.1.f
YNL169C	37	6	*PSD1*	**0.66**	**0.67**	++++	30	**+/-**	10.2.D

ORF name	blot	lane	Gene name	COX % splicing	COB % splicing	rel. total RNA	temp (°C)	Glycerol growth	library grid location
YNL237W	23	2	*YTP1*	**0.78**	**0.81**	++	30	l**+/-**	10.1.D
YNL252C	2	10	*MRPL17*	**0.71**	**0.63**	++	30	**+/-**	11.4.d
YNL259C	12	16	*ATX1*	**0.74**	**0.89**	++	30	l**+/-**	2.4.a
YNL280C	12	15	*ERG24*	**0.77**	**0.93**	++	30	l**+/-**	2.3.h
YNL296W	5	9		**0.90**	**0.73**	+	30	l**+/-**	18.2.b
YNL297C	12	14	*MON2*	**0.74**	**0.96**	++	30	l**+/-**	2.3.g
YNL315C	30	3	*ATP11*	**0.00**	**0.00**	++	30	**-**	15.2.D
YNL328C	6	8	*MDJ2*	**0.92**	**0.91**	+++	30	c^s**+/-**	2.3.f
YNL328C	7	13	*MDJ2*	**0.76**	**0.66**	+++	***20***	c^s**+/-**	2.3.f
YNR001C	28	12	*CIT1*	**0.72**	**0.84**	++	30	**+/-**	13.4.H
YNR006W	10	9	*VPS27*	**0.69**	**0.80**	++	***37***	t^s**-**	13.5.a
YNR018W	28	13		**0.69**	**0.85**	++	30	l**+/-**	13.5.B
YNR020C	2	15		**0.63**	**0.86**	++	30	**+/-**	13.5.c
YNR025C	2	16		**0.78**	**0.85**	++	30	**+/-**	13.5.e
YNR041C	28	14	*COQ2*	**0.54**	**0.79**	++	30	**-**	13.6.A
YNR042W	28	15		**0.82**	**0.85**	+++	30	l**+/-**	13.6.B
YNR045W	28	16	*PET494*	**0.44**	**0.80**	++	30	**-**	13.6.C
YNR052C	3	6	*POP2*	**0.66**	**0.55**	+/-	30	**+/-**	15.3.e
YNR052C	37	9	*POP2*	**0.72**	**0.72**	+++	30	**+/-**	15.3.E
YOL001W	14	1	*PHO80*	**0.75**	**0.94**	++	30	l**+/-**	3.1.a
YOL008W	1	4		**0.71**	**0.62**	+	30	**+/-**	3.1.d
YOL009C	14	2	*MDM12*	**0.55**	**0.98**	++	30	l**+/-**	3.1.e
YOL023W	1	5	*IFM1*	**0.74**	**0.66**	+++	30	**+/-**	3.1.f
YOL027C	14	3	*MDM38*	**0.59**	**0.95**	++	30	l**+/-**	3.1.g
YOL050C	5	1		**0.89**	**0.75**	+++	30	l**+/-**	3.2.c
YOL071W	14	4	*EMI5*	**0.75**	**0.91**	++++	30	**+/-**	3.2.d
YOL081W	14	5	*IRA2*	**0.82**	**0.92**	+++	30	l**+/-**	3.2.e
YOL095C	25	5	*HMI1*	**0.78**	**0.89**	++	30	l**+/-**	11.2.H
YOL096C	25	6	*COQ3*	**0.79**	**0.83**	++	30	**-**	11.3.A
YOL143C	35	1	*RIB4*	**0.64**	**0.79**	++	30	**-**	17.4.C
YOL148C	32	5	*SPT20*	**0.65**	**0.81**	++	30	**+/-**	16.5.G
YOR014W	12	17	*RTS1*	**0.76**	**0.93**	++	30	l**+/-**	2.4.b
YOR017W	12	18	*PET127*	**0.49**	**0.14**	++	30	l**+/-**	2.4.c
YOR017W	20	6	*PET127*	**0.26**	**0.12**	++	30	l**+/-**	2.4.c
YOR017W	cand 1	6	*PET127*	**0.27**	**0.03**	+++	30	l**+/-**	2.4.C
YOR022C	36	1		**0.64**	**0.68**	+++	30	l**+/-**	2.4.d
YOR026W	13	1	*BUB3*	**0.82**	**0.82**	+++	30	**+/-**	2.4.e
YOR036W	13	2	*PEP12*	**0.80**	**0.86**	+++	30	**+/-**	2.4.f

ORF name	blot	lane	Gene name	COX % splicing	COB % splicing	rel. total RNA	temp (°C)	Glycerol growth	library grid location
YOR037W	13	3	*CYC2*	**0.81**	**0.84**	+++	30	**+/-**	2.4.g
YOR065W	13	4	*CYT1*	**0.75**	**0.69**	++	30	**-**	2.5.b
YOR073W	13	5		**0.79**	**0.79**	++	30	l**+/-**	2.5.c
YOR125C	17	13	*CAT5*	**0.75**	**0.72**	+	30	**-**	6.6.a
YOR135C	17	14		**0.85**	**0.63**	++++	30	**+/-**	6.6.c
YOR136W	17	15	*IDH2*	**0.91**	**0.66**	++++	30	**+/-**	6.6.d
YOR147W	35	11	*MDM32*	**0.76**	**0.77**	++	30	l**+/-**	18.3.D
YOR196C	17	16	*LIP5*	**0.94**	**0.79**	++++	30	l**+/-**	6.6.g
YOR221C	18	1	*MCT1*	**0.82**	**0.99**	++	30	**+/-**	7.1.f
YOR271C	6	13		**0.90**	**0.88**	+++	30	c^s**+/-**	15.2.f
YOR271C	7	18		**0.90**	**0.70**	+++	***20***	c^s**+/-**	15.2.f
YOR290C	13	6	*SNF2*	**0.83**	**0.81**	+++	30	**+/-**	2.5.d
YOR305W	3	16		**0.00**	**0.19**	+/-	30	**+/-**	2.5.f
YOR305W	4	19		**0.73**	**0.67**	++	30	**+/-**	2.5.f
YOR331C	37	4		**0.75**	**0.74**	+++++	30	**+/-**	17.5.E
YOR332W	8	1	*VMA4*	**1.00**	**0.94**	++	***37***	t^s**+/-**	2.5.h
YOR333C	28	11		**0.73**	**0.84**	++	30	**+/-**	13.4.E
YOR334W	5	12	*MRS2*	**0.71**	**0.53**	+++	30	t^s**+/-**	2.6.a
YOR334W	7	3	*MRS2*	**0.60**	**0.68**	++	***37***	t^s**+/-**	2.6.a
YOR334W	cand 1	13	*MRS2*	**0.60**	**0.65**	+	***37***	t^s**+/-**	2.6.A
YOR339C	13	7	*UBC11*	**0.84**	**0.86**	++	30	l**+/-**	2.6.b
YOR350C	13	8	*MNE1*	**0.56**	**0.77**	++	30	**+/-**	2.6.c
YOR350C	20	7	*MNE1*	**0.40**	**0.85**	++	30	**+/-**	2.6.c
YOR350C	cand 1	4	*MNE1*	**0.41**	**0.85**	++	30	**+/-**	2.6.C
YOR354C	13	9	*MSC6*	**0.86**	**0.84**	++	30	l**+/-**	2.6.d
YOR358W	13	10	*HAP5*	**0.63**	**0.69**	++	30	**-**	2.6.f
YOR360C	13	11	*PDE2*	**0.87**	**0.89**	+++	30	l**+/-**	2.6.g
YPL002C	5	6	*SNF8*	**0.83**	**0.74**	++	30	l**+/-**	8.6.f
YPL024W	33	2	*NCE4*	**0.75**	**0.80**	+++	30	l**+/-**	17.1.D
YPL040C	19	15	*ISM1*	**0.88**	**0.80**	+++	30	l**+/-**	8.6.B
YPL042C	19	14	*SSN3*	**0.79**	**0.61**	+	30	l**+/-**	8.6.A
YPL045W	33	1	*VPS16*	**0.75**	**0.81**	+++	30	**+/-**	17.1.B
YPL060W	19	13	*LPE10*	**0.86**	**0.77**	++	30	l**+/-**	8.5.H
YPL065W	36	14	*VPS28*	**0.75**	**0.71**	+++	30	l**+/-**	8.5.G
YPL066W	19	12		**0.86**	**0.69**	+++	30	l**+/-**	8.5.F
YPL078C	19	11	*ATP4*	**0.89**	**0.81**	++	30	l**+/-**	8.5.C
YPL084W	19	10	*BRO1*	**0.92**	**0.81**	++	30	**+/-**	8.5.B
YPL098C	14	12	YPL098C	**0.52**	**0.88**	++	30	l**+/-**	3.4.h

ORF name	blot	lane	Gene name	COX % splicing	COB % splicing	rel. total RNA	temp (°C)	Glycerol growth	library grid location
YPL098C	cand 2	6	YPL098C	**0.43**	**0.69**	+++	30	l**+/-**	3.4.H
YPL132W	2	11	*COX11*	**0.78**	**0.80**	+	30	**+/-**	11.4.e
YPL135W	5	2	*ISU1*	**0.85**	**0.77**	+++	30	**+**	3.4.c
YPL148C	34	8	*PPT2*	**0.72**	**0.84**	+++	30	l**+/-**	17.6.A
YPL158C	22	10		**0.93**	**0.94**	+++	30	l**+/-**	9.6.A
YPL172C	14	11	*COX10*	**0.38**	**1.00**	++	30	**-**	3.4.a
YPL172C	cand 2	9	*COX10*	**0.47**	**0.77**	++	30	**-**	3.4.A
YPL174C	14	10	*NIP100*	**0.49**	**0.94**	++++	30	**+/-**	3.3.g
YPL174C	cand 1	11	*NIP100*	**0.95**	**0.88**	+++	30	**+/-**	3.3.G
YPL183W-A	3	4	*GON5*	**0.87**	**0.54**	+	30	**+/-**	15.1.h
YPL188W	9	7	*POS5*	**0.64**	**0.68**	+++	***20***	c^s**+/-**	3.3.f
YPL215W	9	6	*CBP3*	**0.66**	**0.73**	++	***20***	c^s**+/-**	3.3.e
YPL222W	6	9	*FMP40*	**0.92**	**0.90**	+++	30	c^s**+/-**	3.3.d
YPL222W	7	14	*FMP40*	**0.82**	**0.71**	+++	***20***	c^s**+/-**	3.3.d
YPL234C	14	9	*TFP3*	**0.61**	**0.93**	+++	30	**+/-**	3.3.c
YPL254W	14	8	*HFI1*	**0.40**	**0.96**	+/-	30	**+/-**	3.3.b
YPL262W	14	7	*FUM1*	**0.78**	**0.96**	++	30	**+/-**	3.3.a
YPL265W	14	6	*DIP5*	**0.87**	**0.92**	+++	30	l**+/-**	3.2.h
YPL268W	34	9	*PLC1*	**0.73**	**0.84**	+++	30	l**+/-**	17.6.B
YPL270W	9	5	*MDL2*	**0.65**	**0.63**	++	***20***	c^s**+/-**	3.2.g
YPR024W	37	10	*YME1*	**0.53**	**0.62**	+++	30	**+/-**	15.3.F
YPR036W	25	14	*VMA13*	**0.72**	**0.87**	+++	30	**+/-**	11.5.F
YPR066W	25	15	*UBA3*	**0.60**	**0.85**	++	30	**+/-**	11.6.A
YPR123C	6	11		**0.90**	**0.90**	+++	30	c^s**+/-**	9.1.c
YPR123C	7	16		**0.93**	**0.63**	+++	***20***	c^s**+/-**	9.1.c
YPR124W	22	1	*CTR1*	**0.93**	**0.92**	++	30	**-**	9.1.D
YPR133W-A	35	7	*TOM5*	**0.78**	**0.77**	+++	30	l**+/-**	18.2.C
YPR134W	2	6	*MSS18*	**0.85**	**0.91**	++	30	**+/-**	9.1.f
YPR134W	3	17	*MSS18*	**0.67**	**0.62**	+	30	**+/-**	9.1.f
YPR134W	cand 1	15	*MSS18*	**0.51**	**0.80**	++	30	**+/-**	9.1.F
YPR146C	22	2		**0.96**	**0.96**	++	30	l**+/-**	9.1.H
YPR166C	22	3	*MRP2*	**0.95**	**0.97**	+	30	l**+/-**	9.2.C
YPR173C	22	4	*VPS4*	**0.92**	**0.96**	+	30	l**+/-**	9.2.D
YPR191W	22	5	*QCR2*	**0.92**	**0.96**	++	30	l**+/-**	9.2.E
YPR197C	22	6		**0.92**	**0.98**	++	30	l**+/-**	9.2.F

APPENDIX D
Exon probed RNA blots from the splicing factor screen

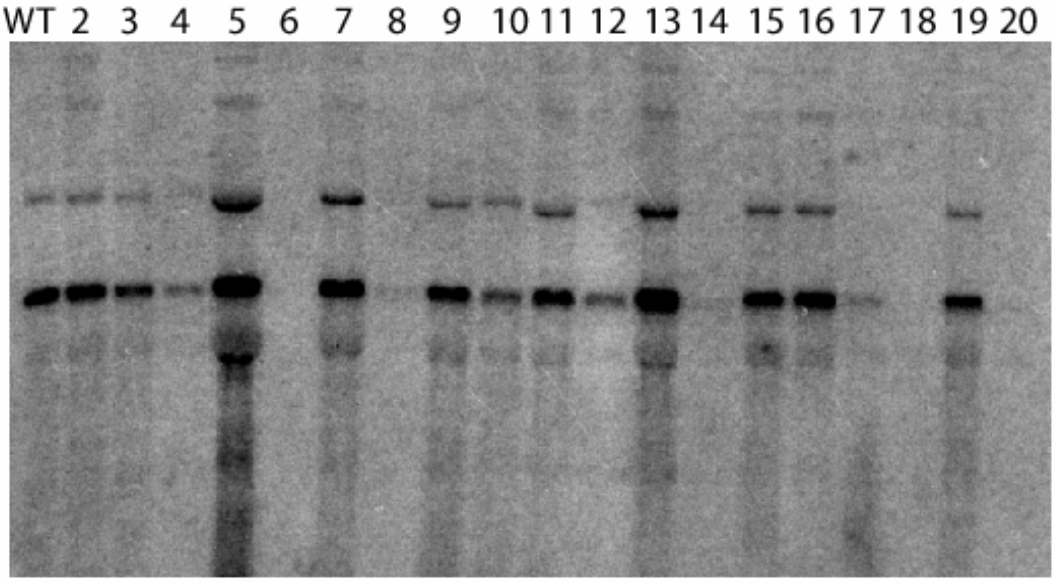
This appendix contains all 37 blots used to screen 476 yeast deletion strains for mitochondrial intron splicing defects.

The blots are listed in numerical order starting with blot 1. The lanes on each blot are labeled numerically. A asterisk on each blot indicates the location of spliced message RNA. The first three blots of a poor quality and many of those samples were repeated on later blots.

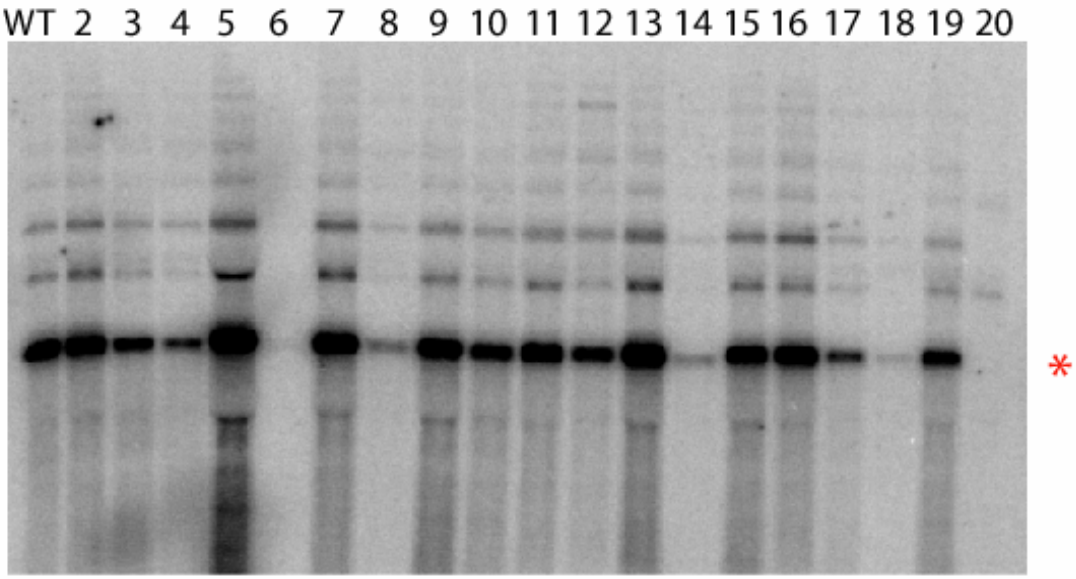
Following the blots, starting on page 264 is a table that lists which RNA samples were loaded in each lane along with a percent splicing calculation.

Northern blot 1

COX1 probe



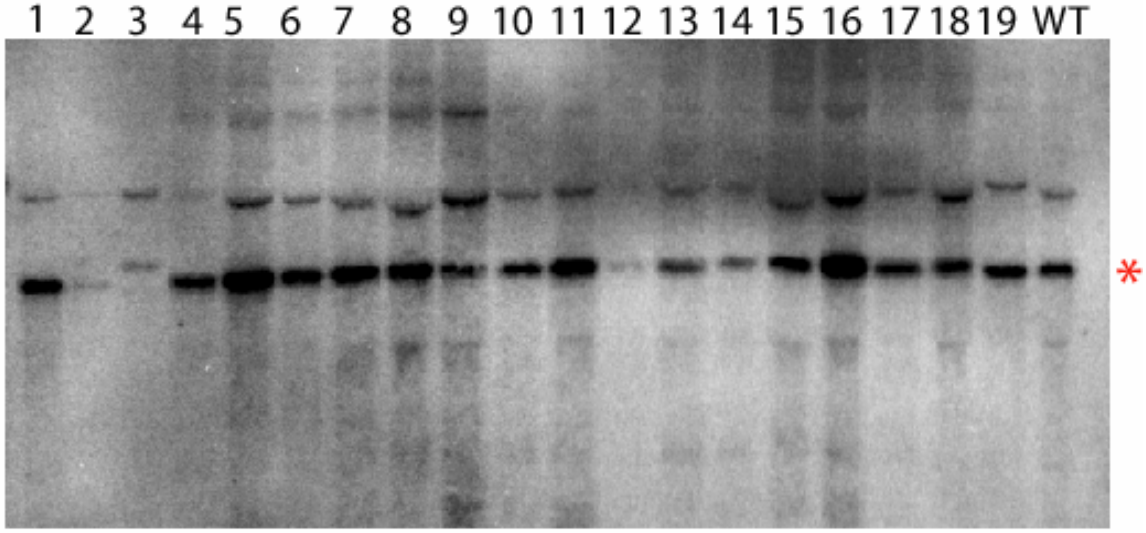
COB probe



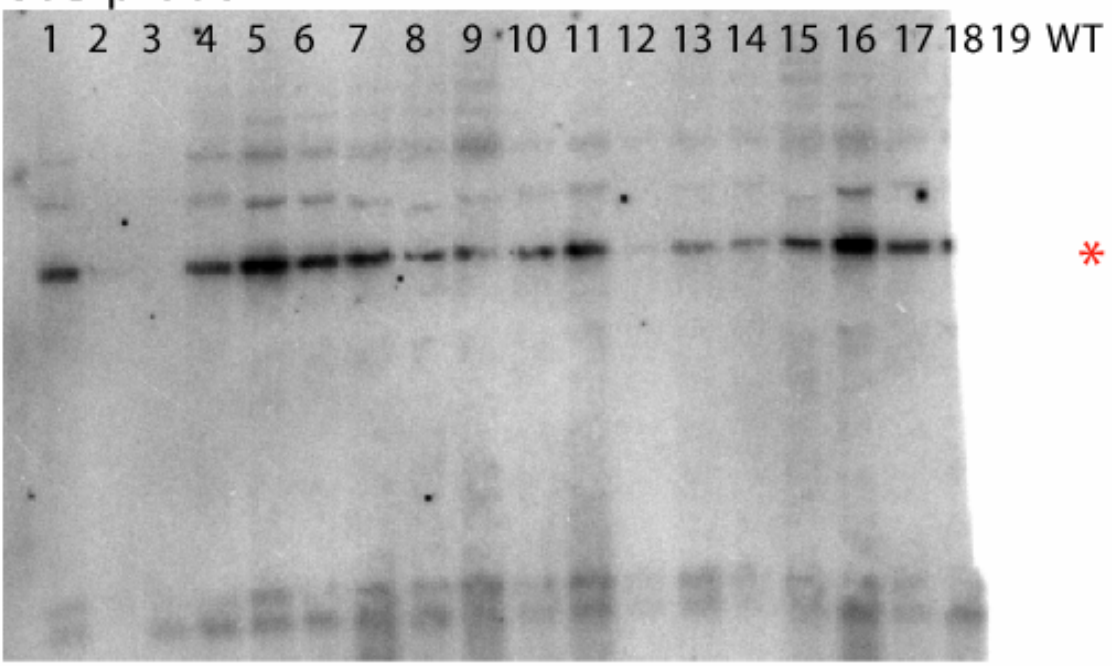
* indicates spliced message band

Northern blot 2

COX1 probe



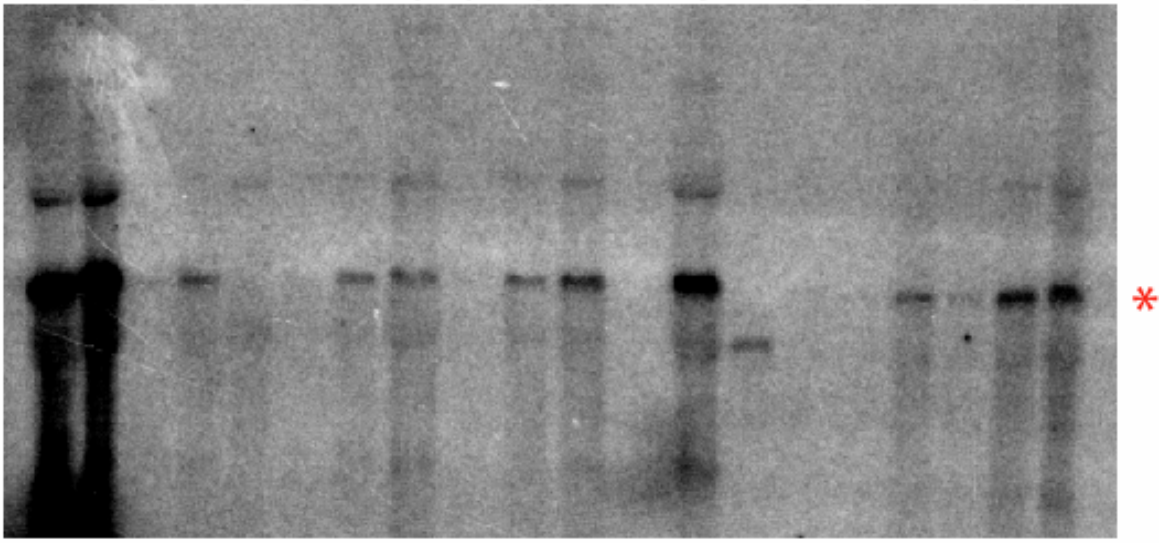
COB probe



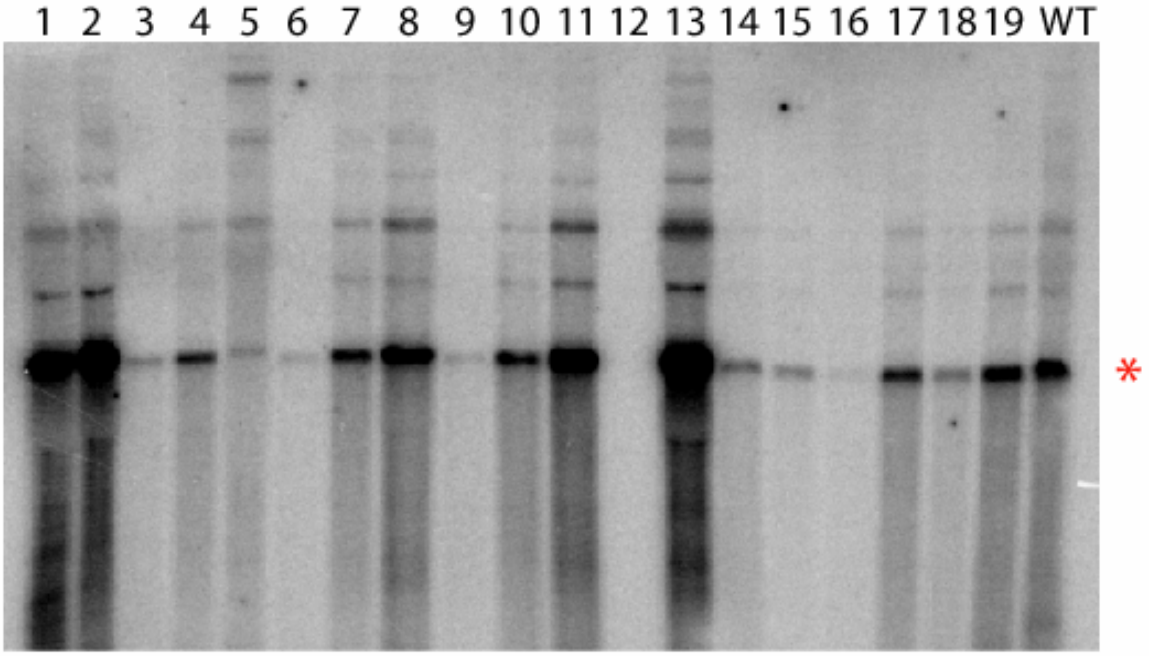
* indicates spliced message band

Northern blot 3

COX1 probe

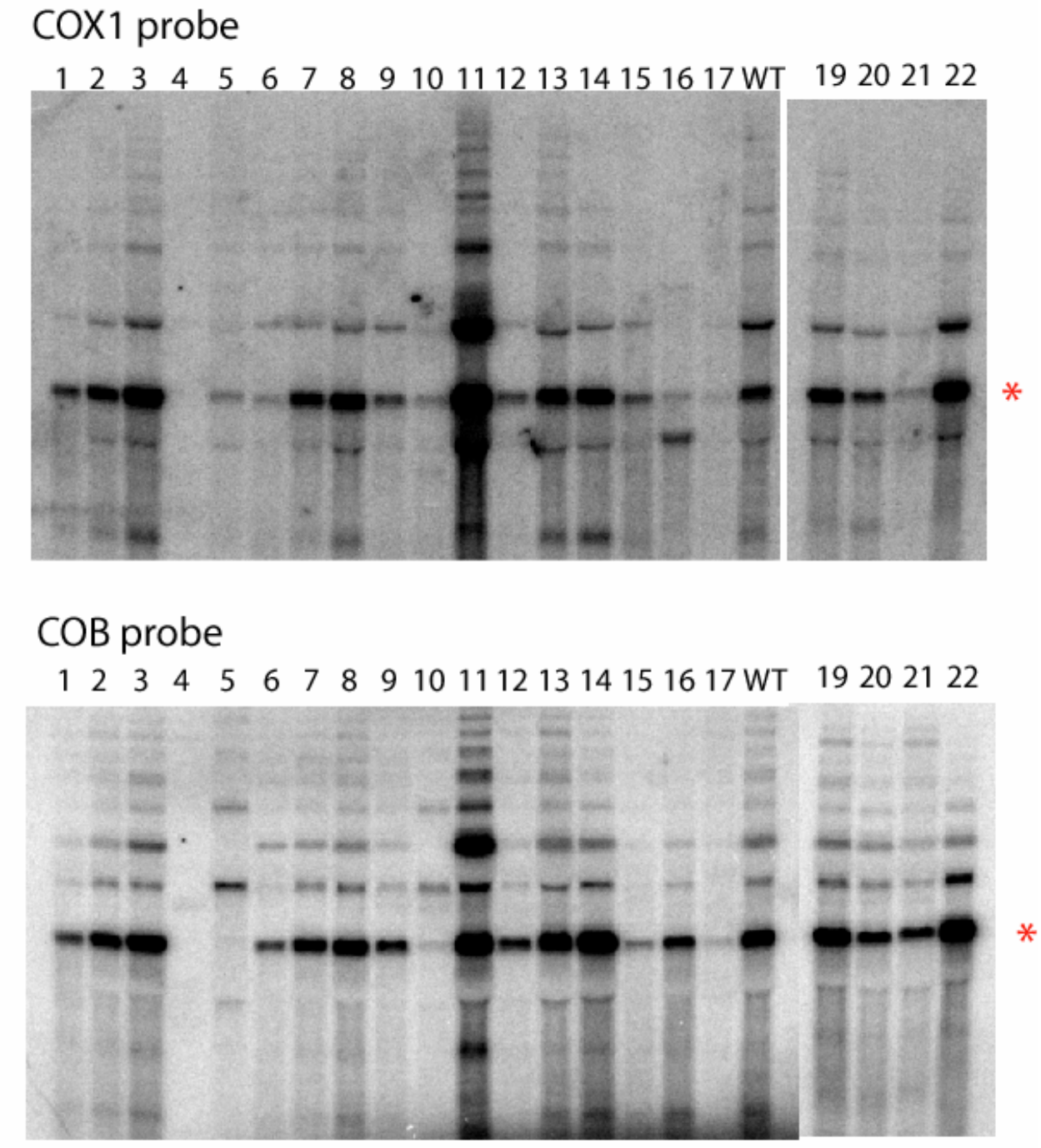


COB probe



* indicates spliced message band

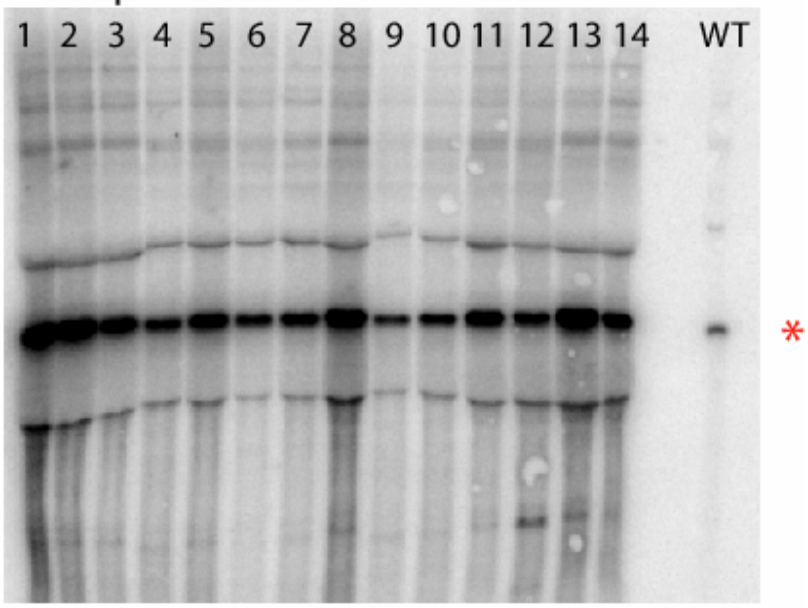
Northern blot 4



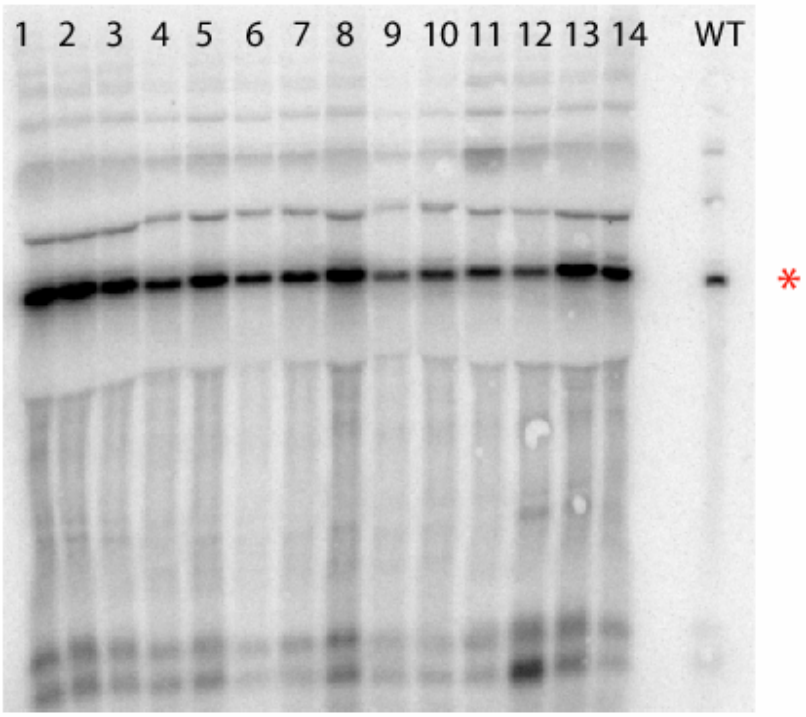
Blot 4 is essencially a repeat of the samples from blots 1-3, becaues blots 1-3 were poor.

Northern blot 5

COX1 probe



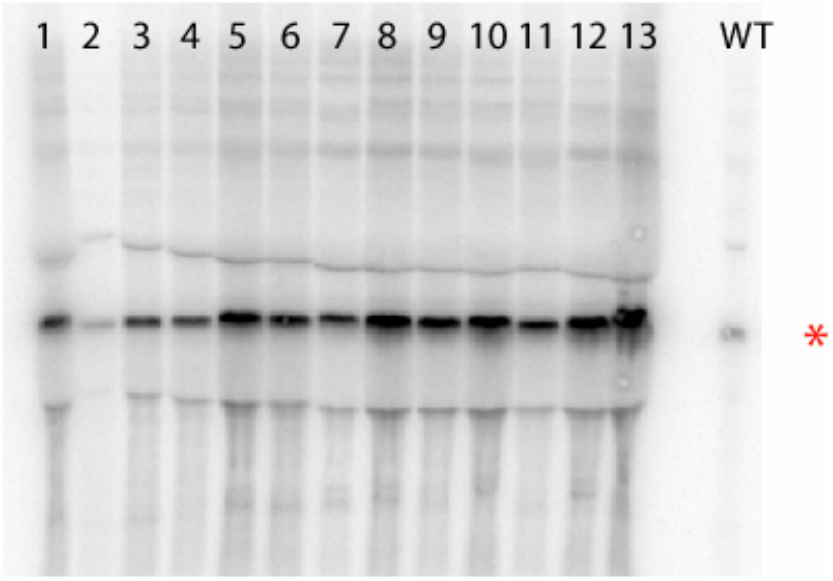
COB probe



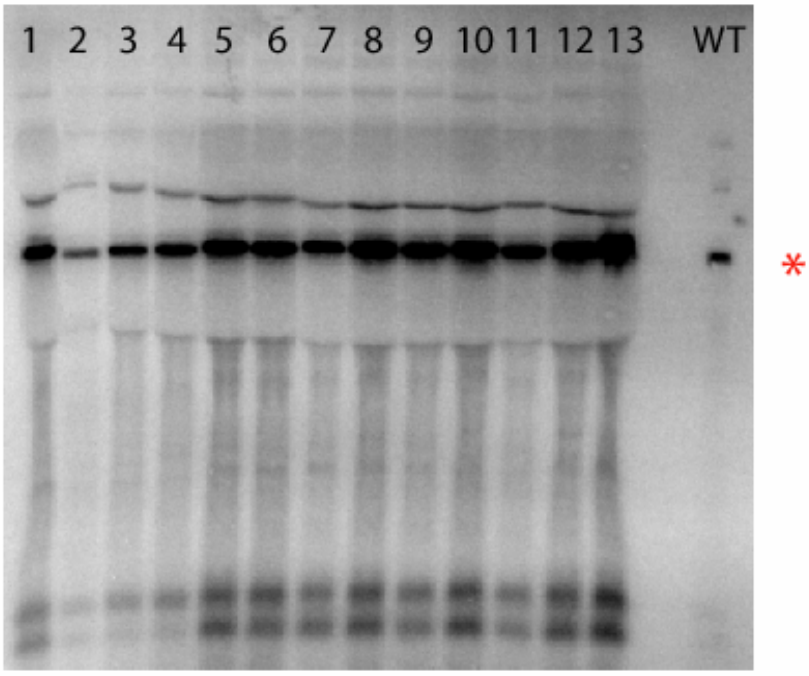
* indicates spliced message band

Northern blot 6

COX1 probe



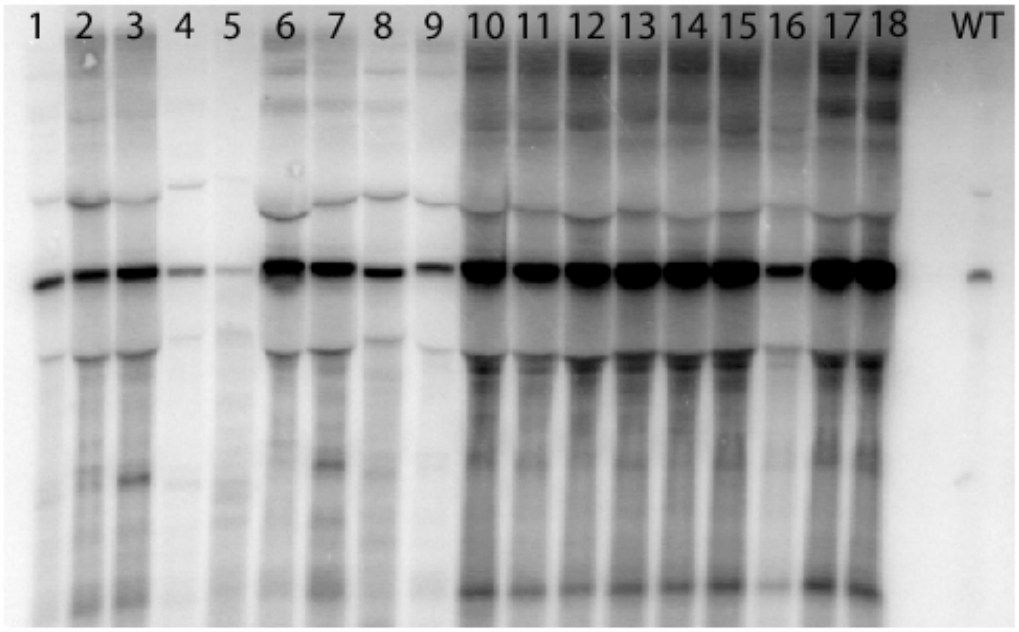
COB probe



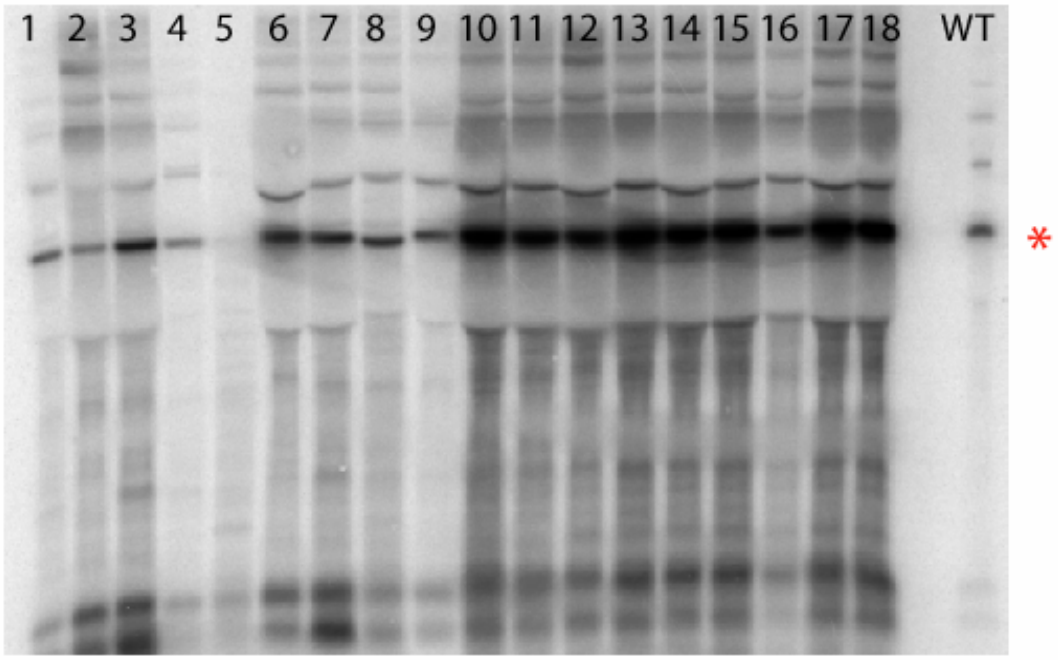
* indicates spliced message band

Northern blot 7

COX1 probe



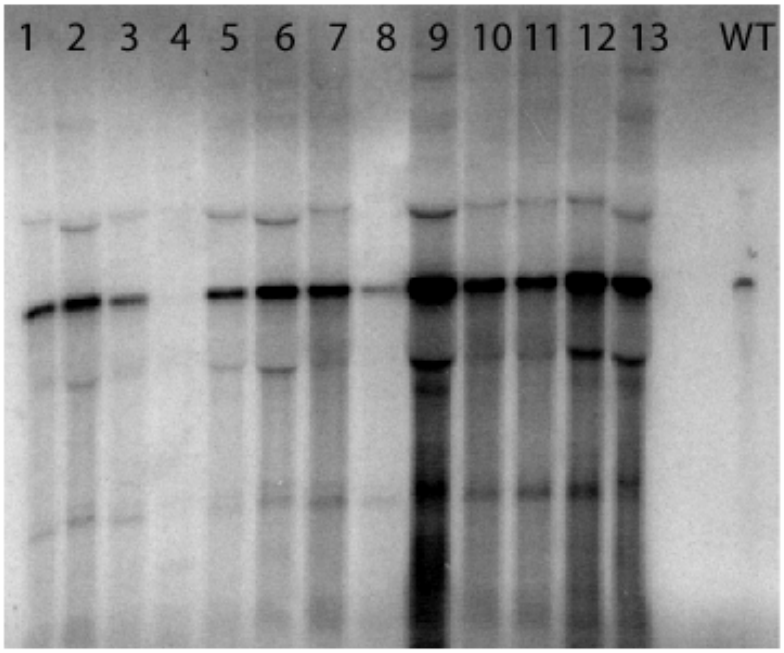
COB probe



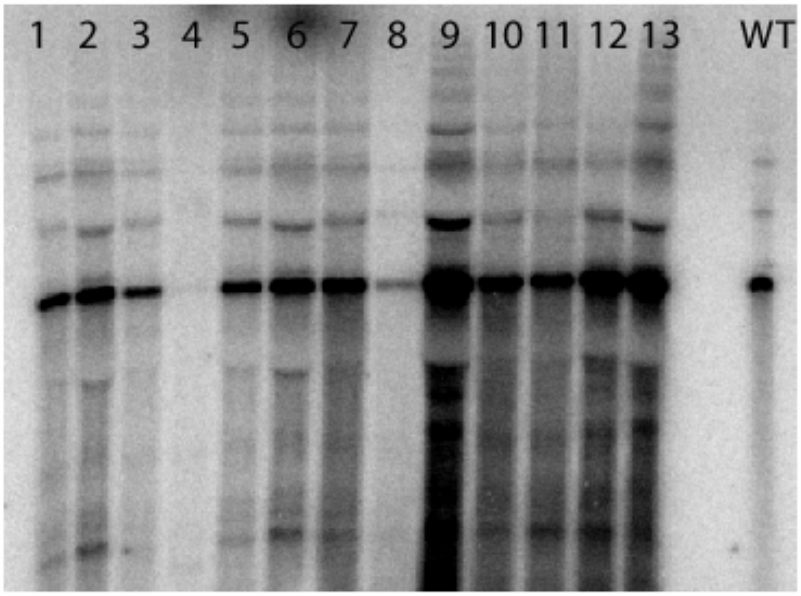
* indicates spliced message band

Northern blot 8

COX1 probe

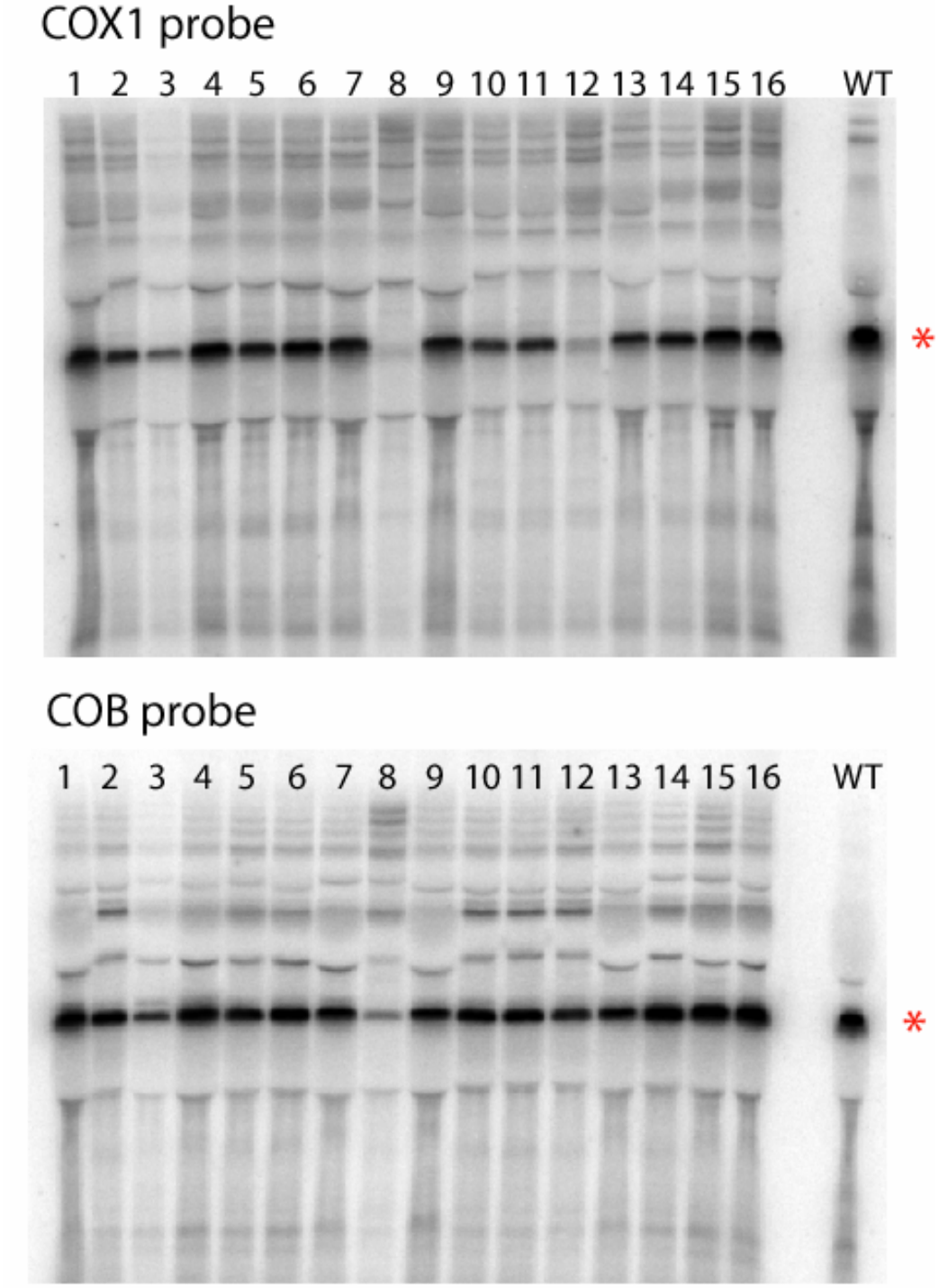


COB probe



* indicates spliced message band

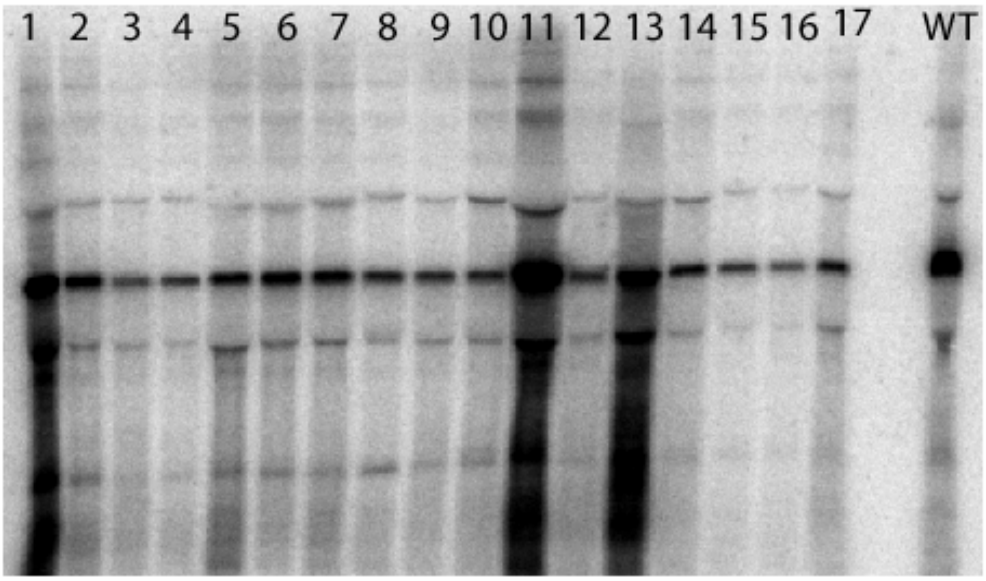
Northern blot 9



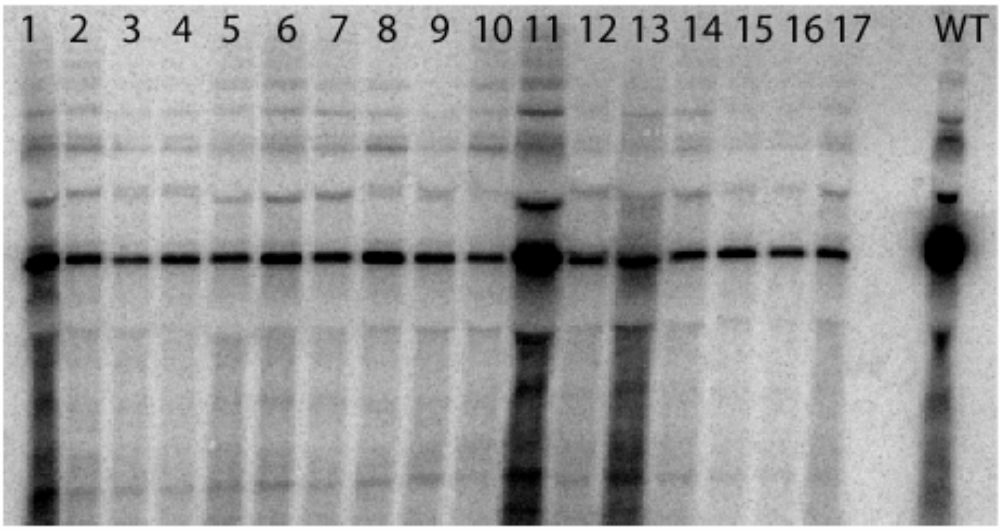
* indicates spliced message band

Northern blot 10

COX1 probe



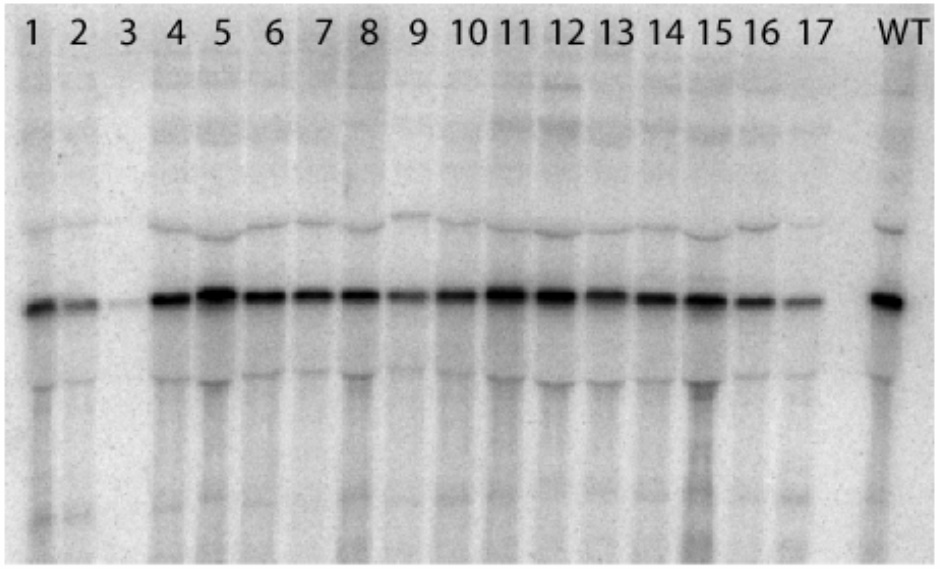
COB probe



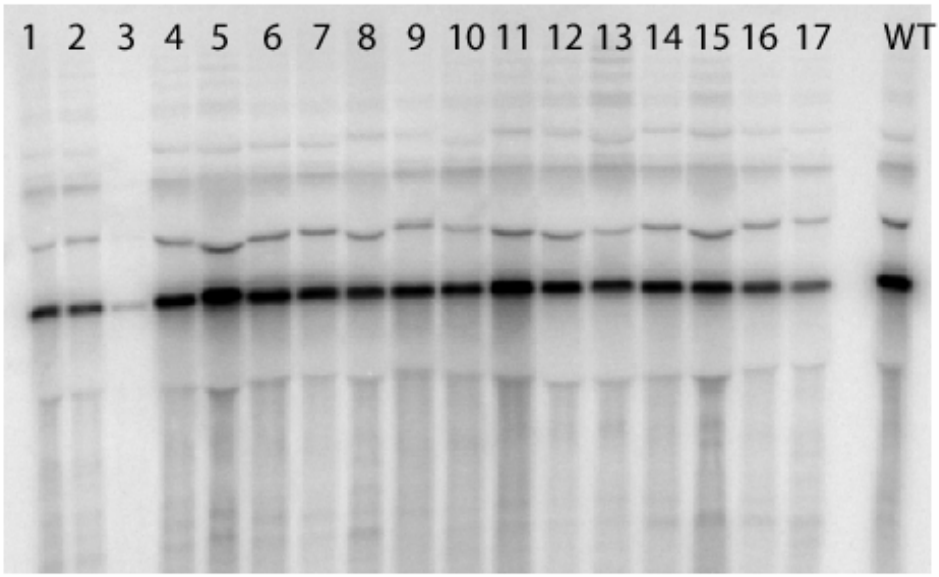
* indicates spliced message band

Northern blot 11

COX1 probe



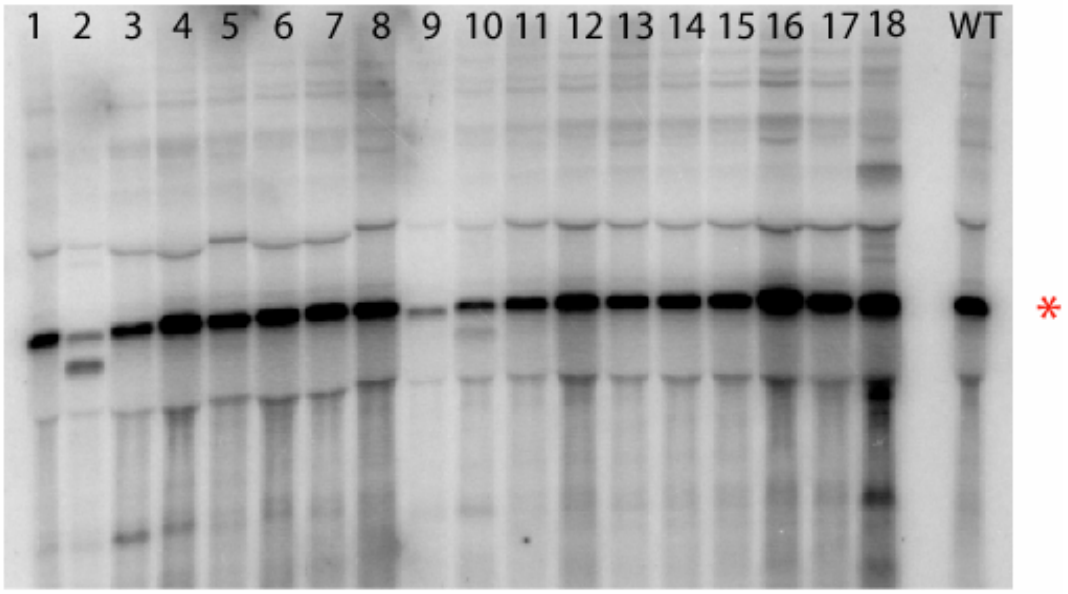
COB probe



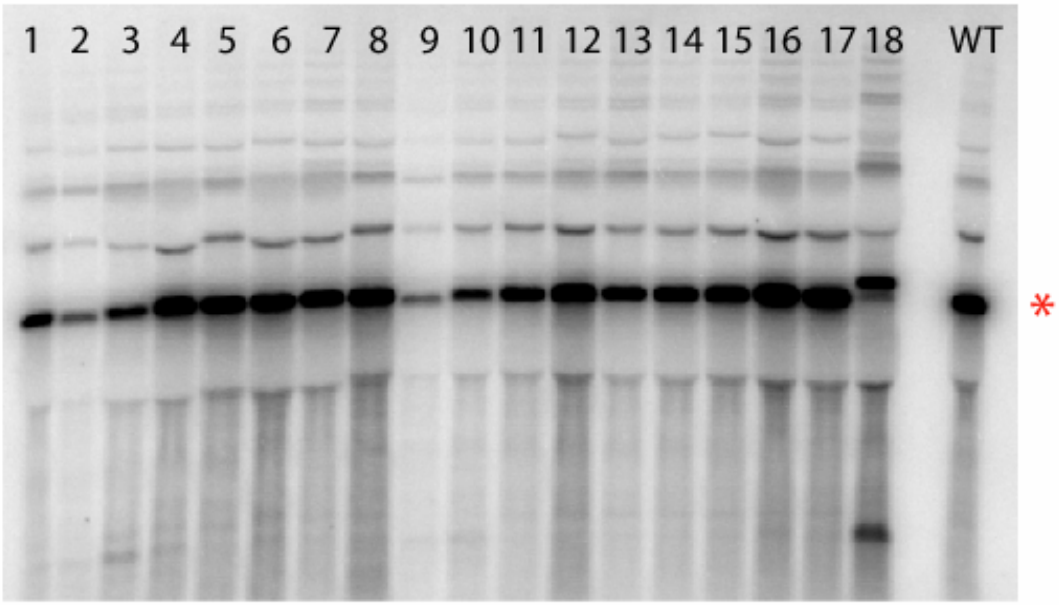
* indicates spliced message band

Northern blot 12

COX1 probe



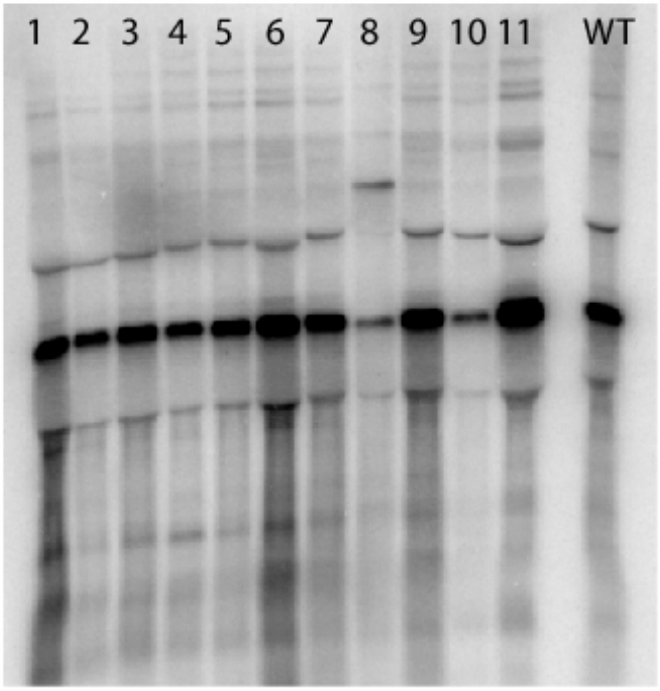
COB probe



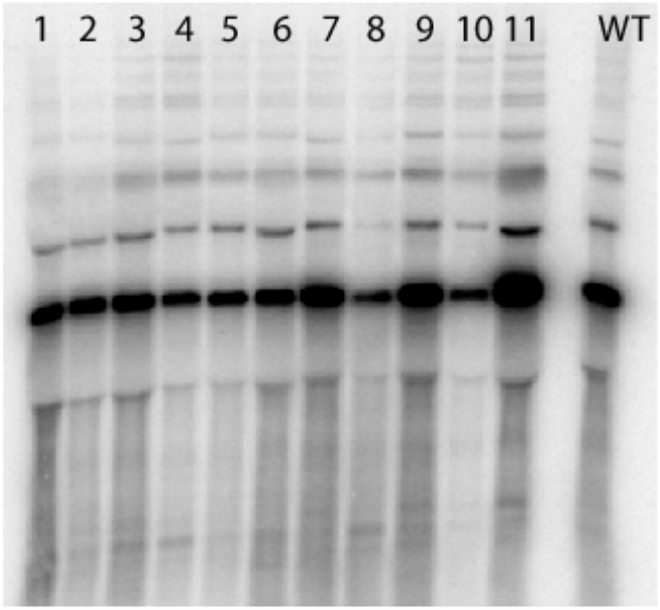
* indicates spliced message band

Northern blot 13

COX1 probe



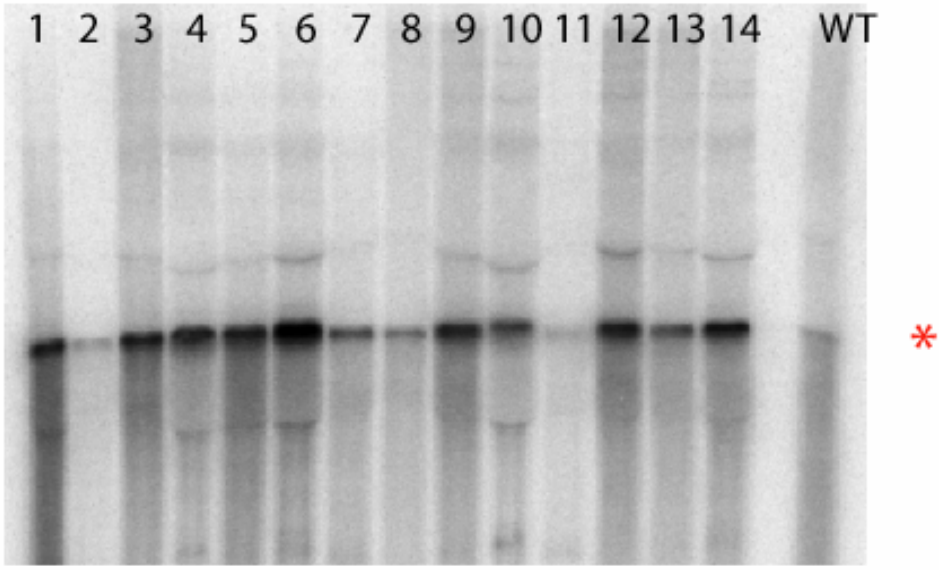
COB probe



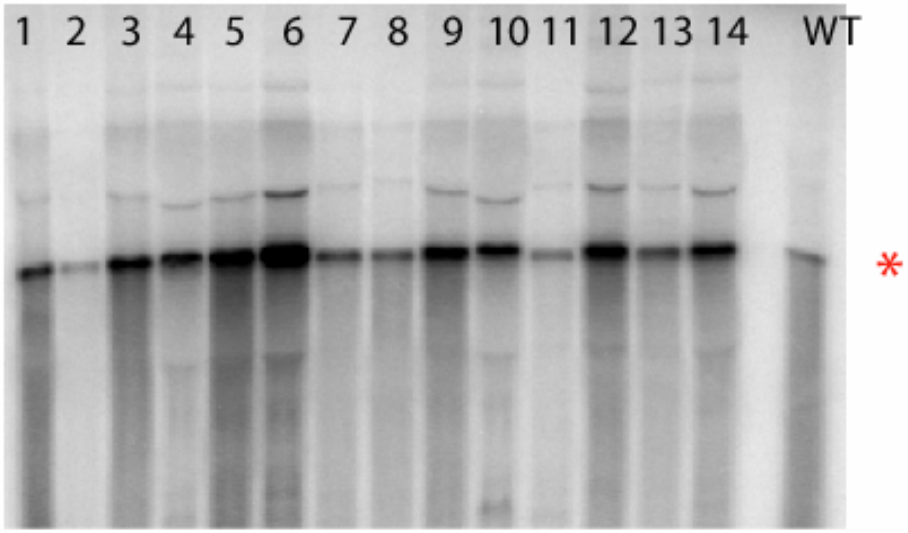
* indicates spliced message band

Northern blot 14

COX1 probe



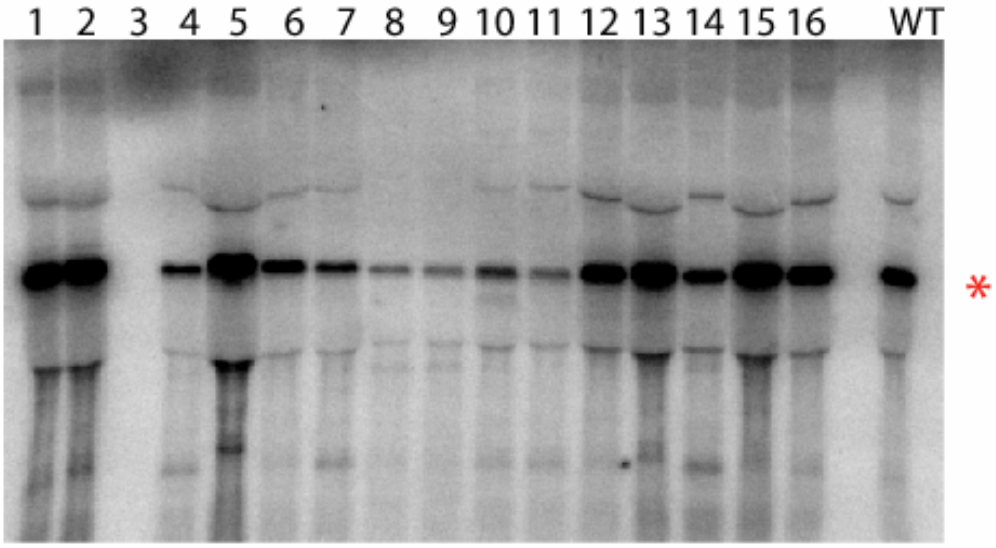
COB probe



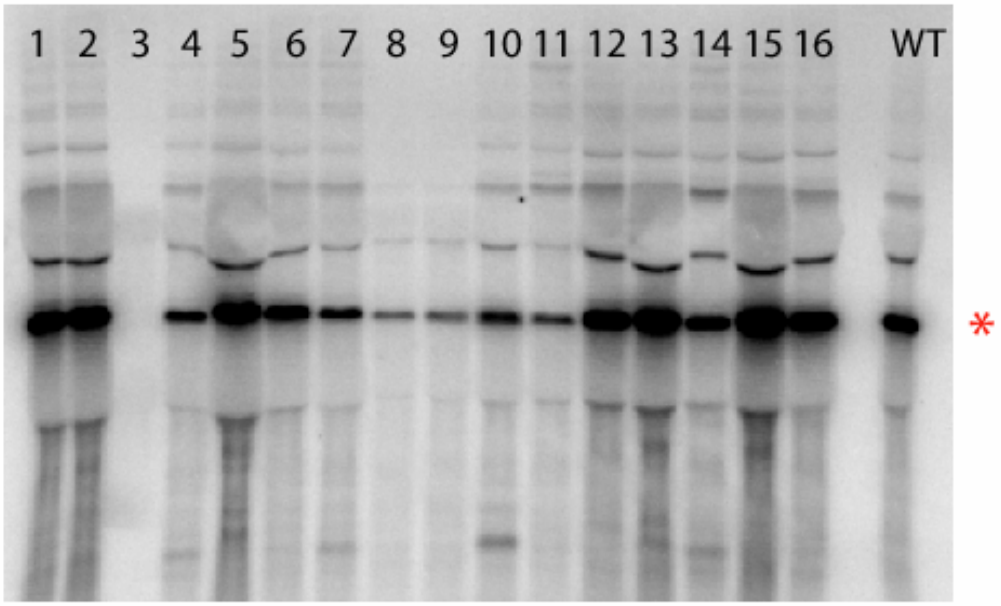
* indicates spliced message band

Northern blot 15

COX1 probe



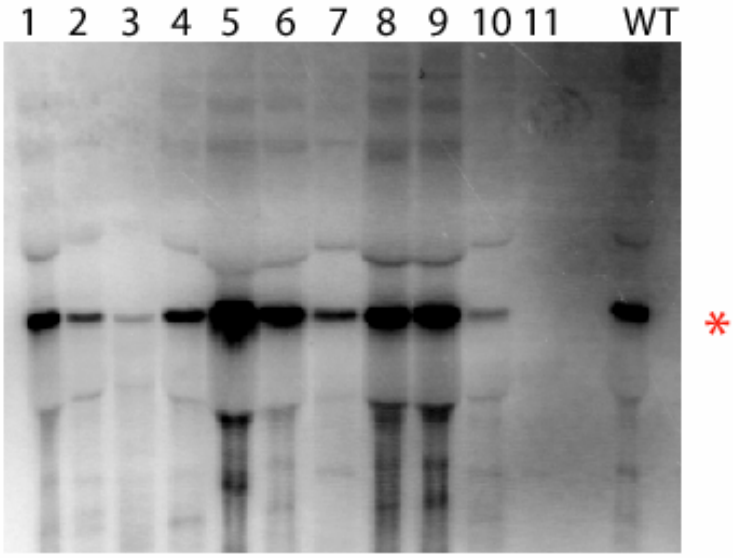
COB probe



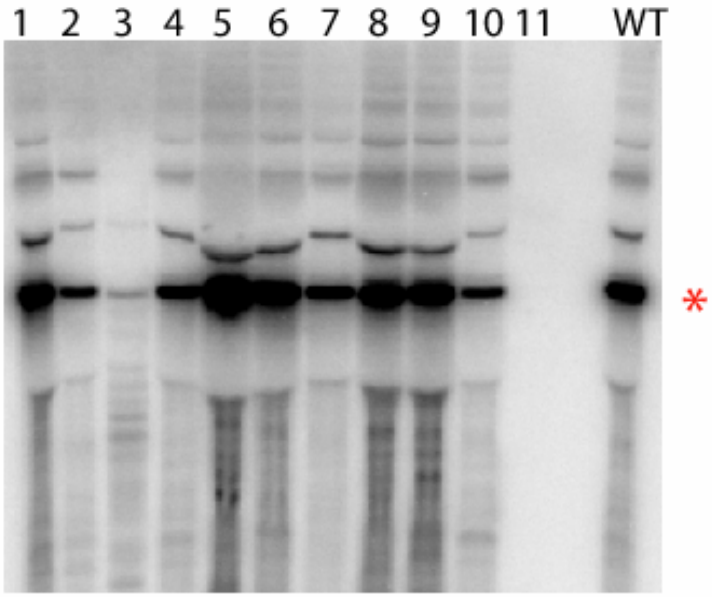
* indicates spliced message band

Northern blot 16

COX1 probe



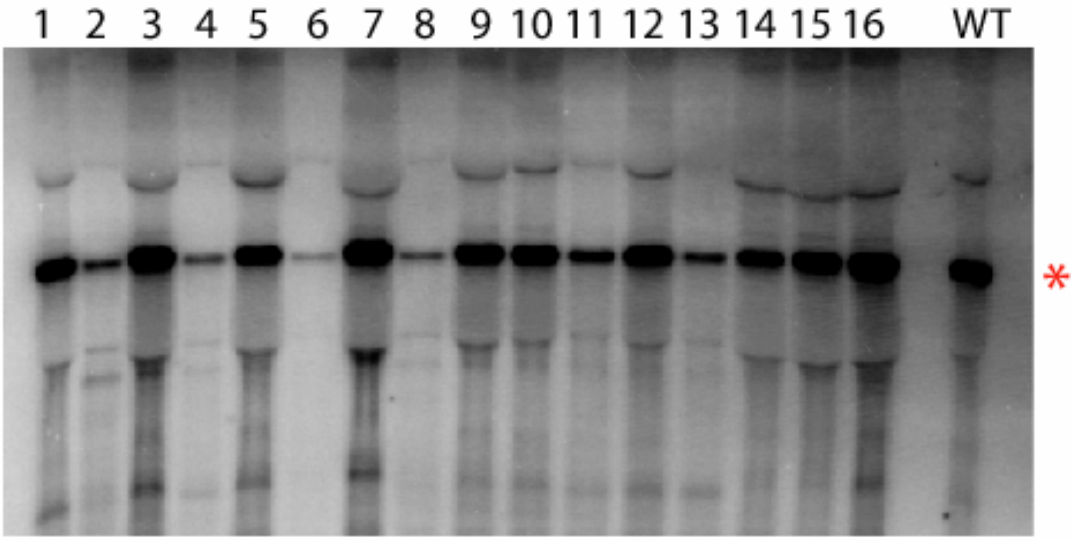
COB probe



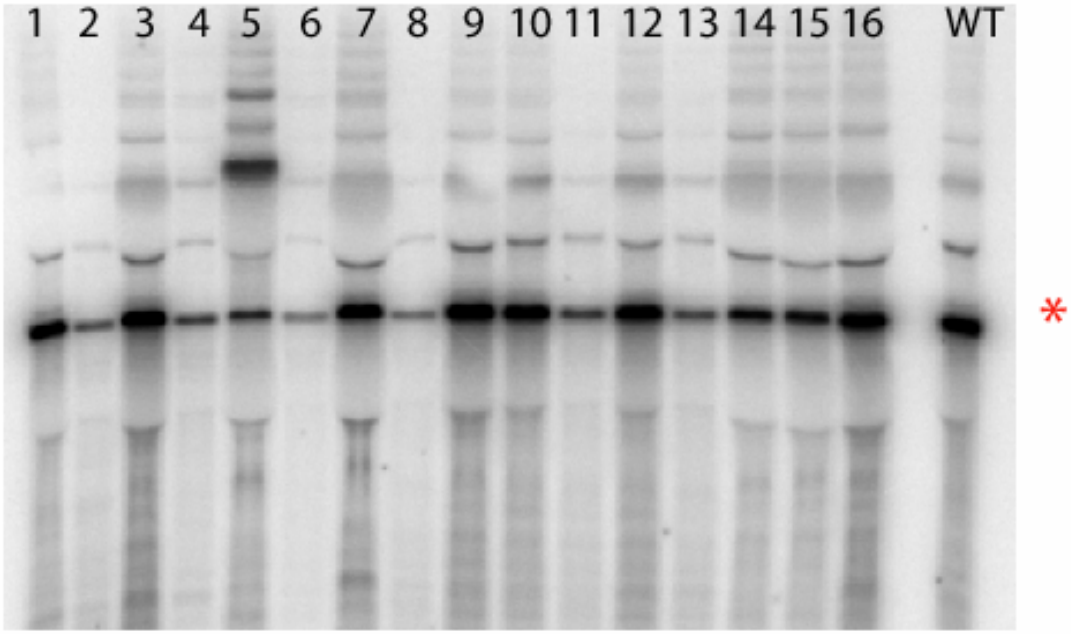
* indicates spliced message band

Northern blot 17

COX1 probe



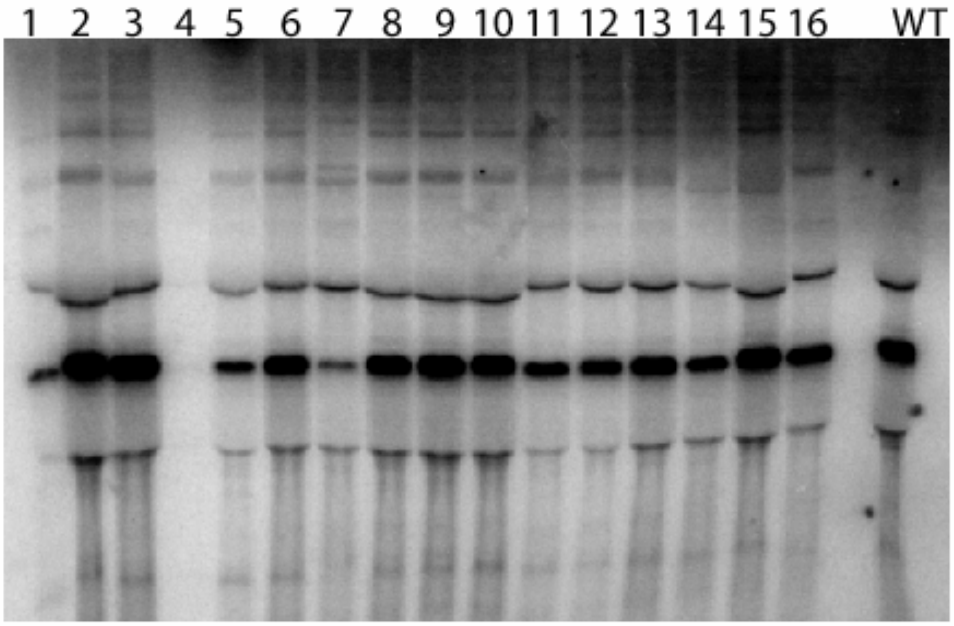
COB probe



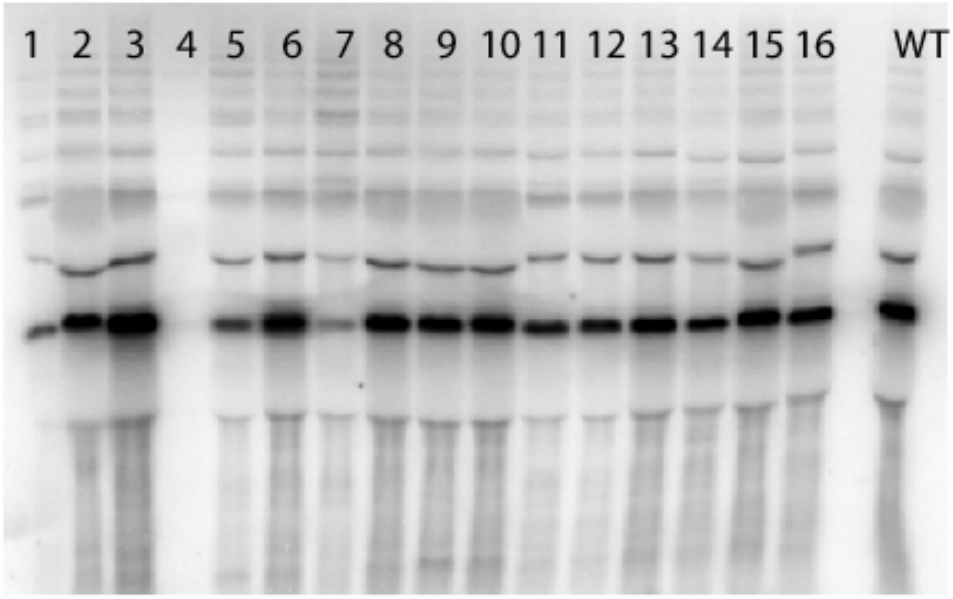
* indicates spliced message band

Northern blot 18

COX1 probe



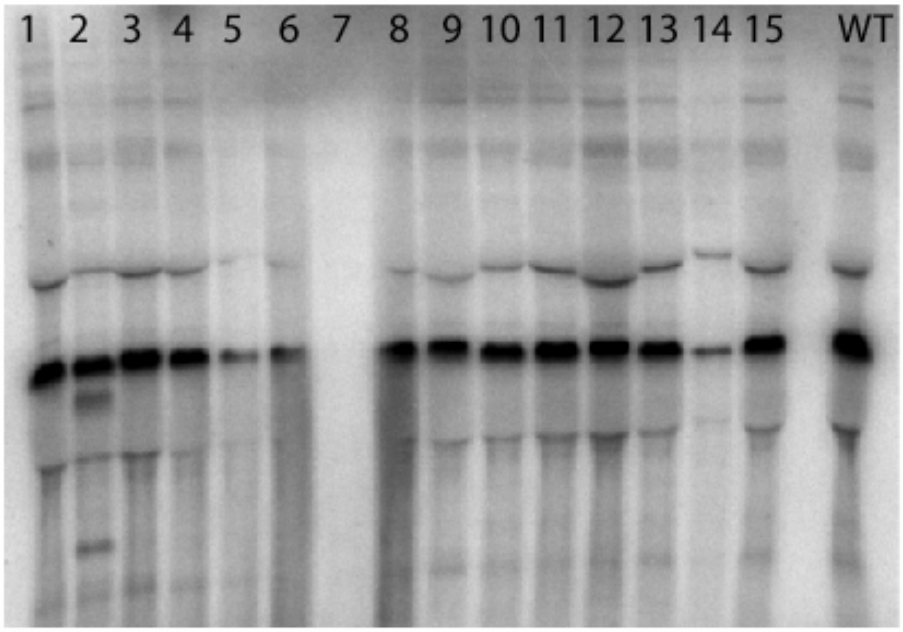
COB probe



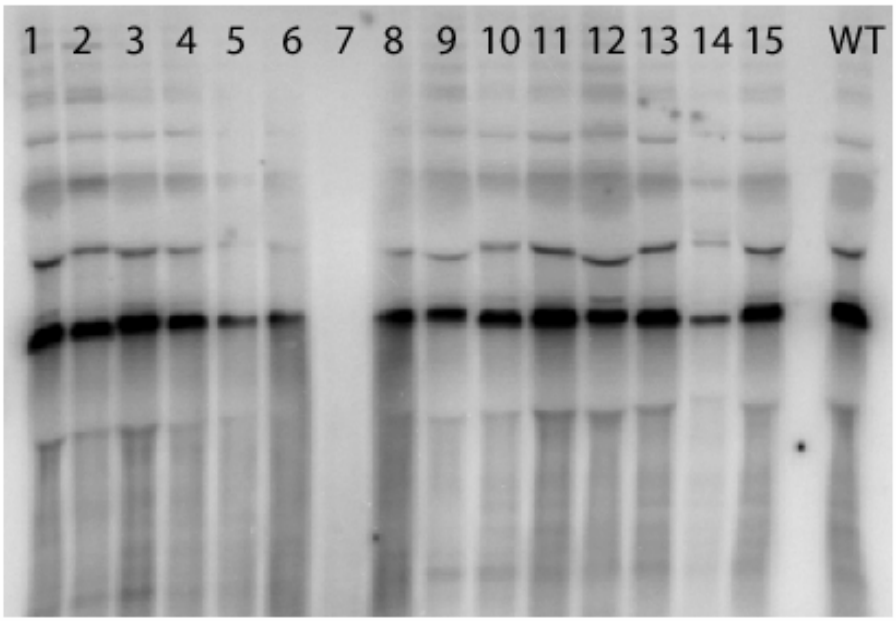
* indicates spliced message band

Northern blot 19

COX1 probe



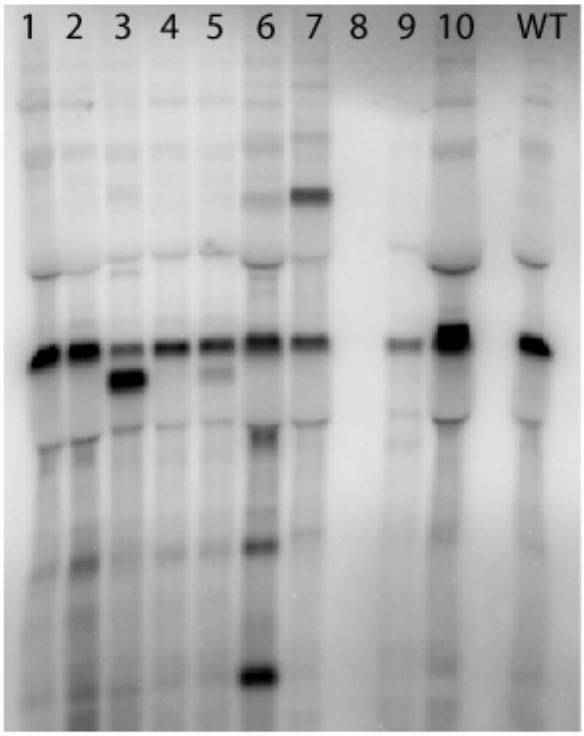
COB probe



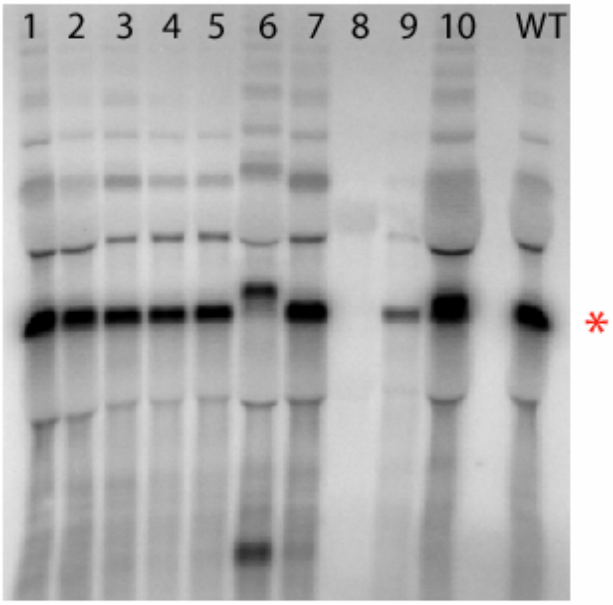
* indicates spliced message band

Northern blot 20

COX1 probe



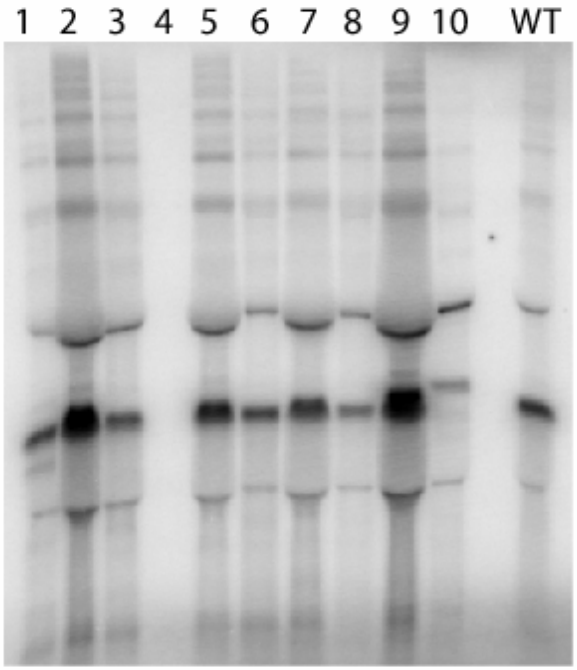
COB probe



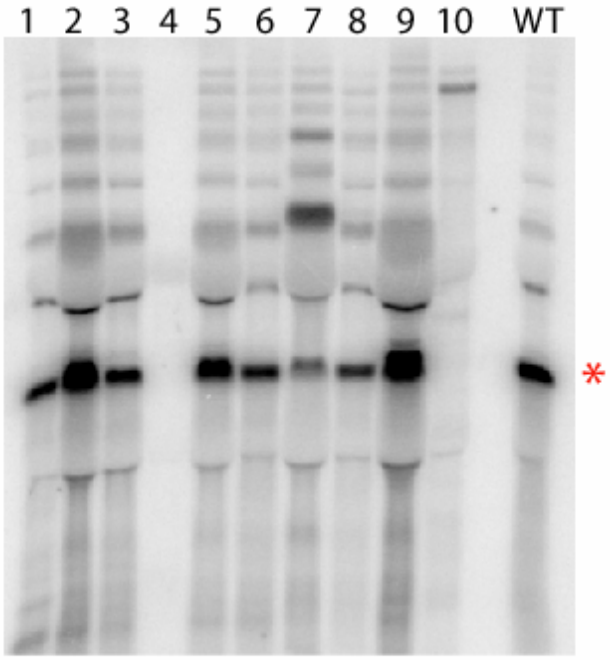
* indicates spliced message band

Northern blot 21

COX1 probe



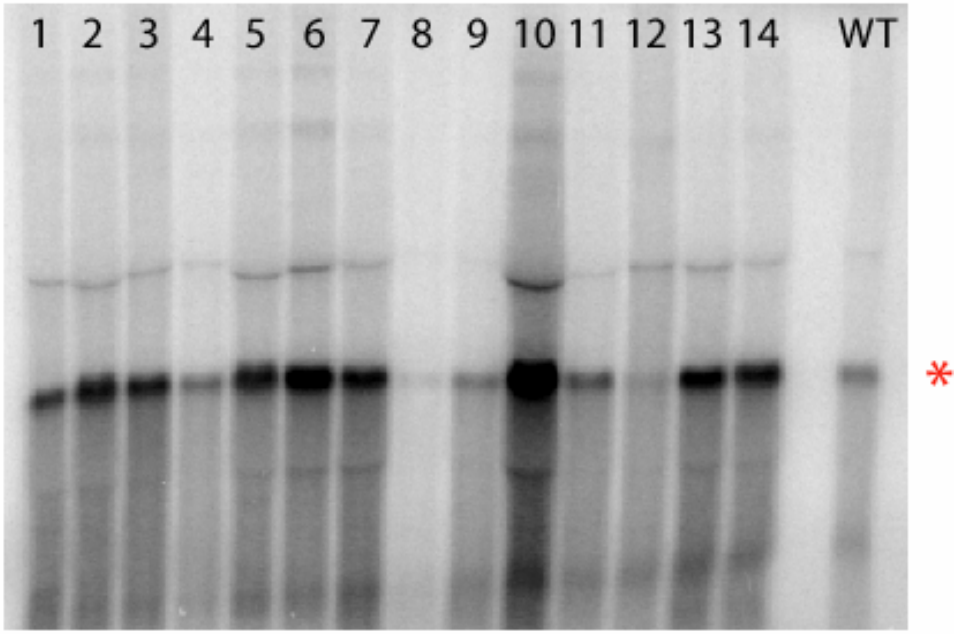
COB probe



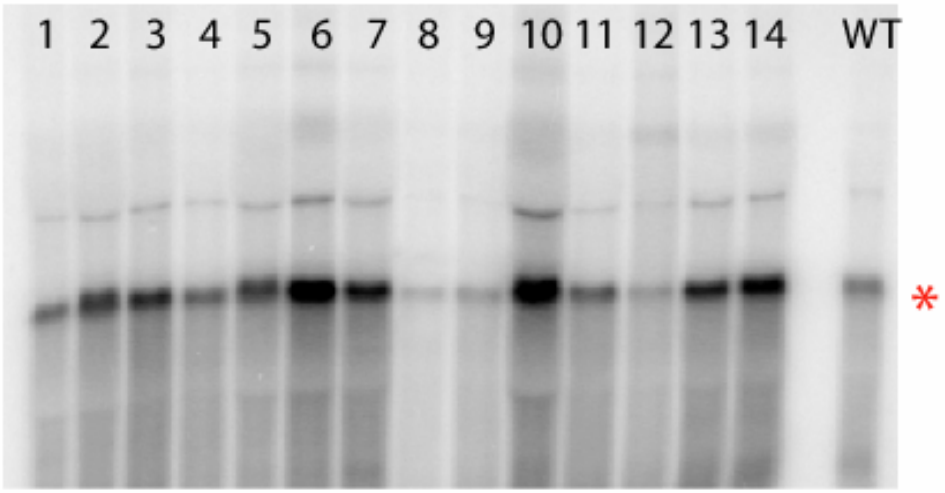
* indicates spliced message band

Northern blot 22

COX1 probe



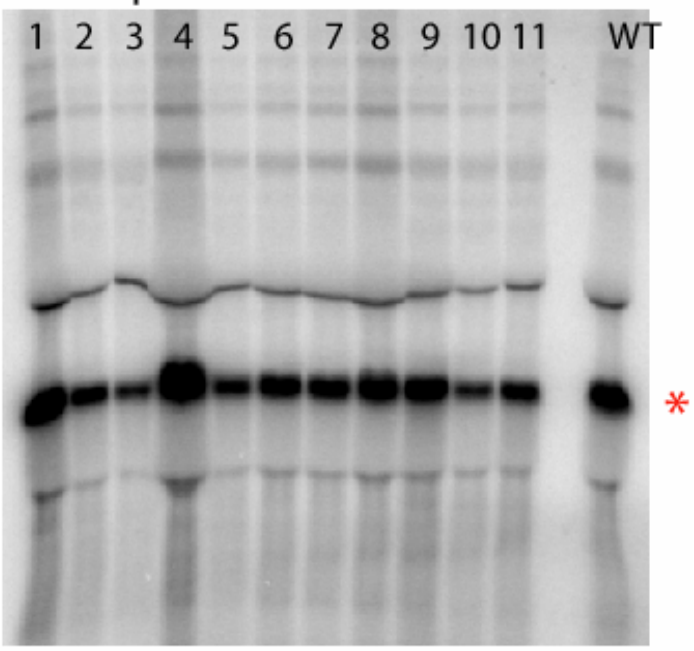
COB probe



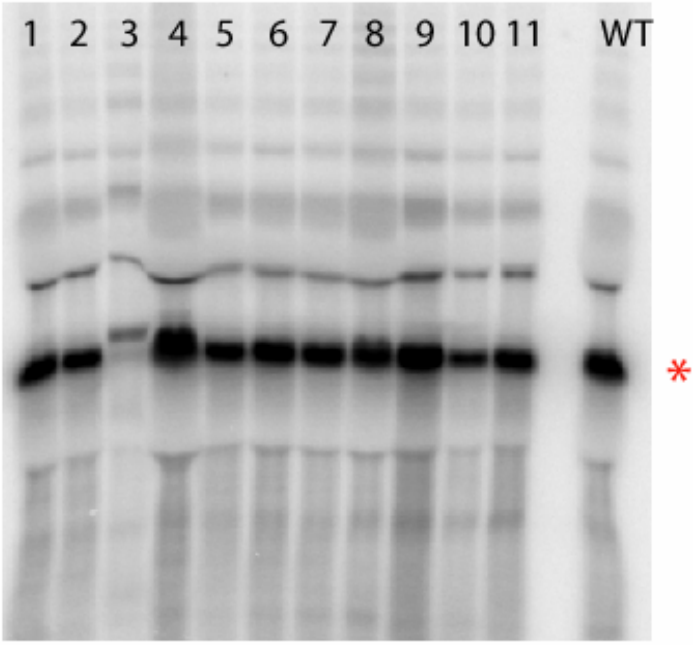
* indicates spliced message band

Northern blot 23

COX1 probe



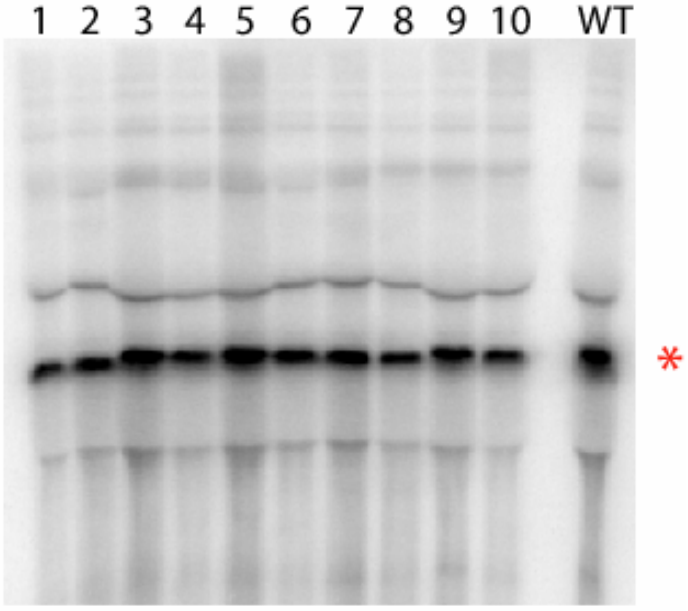
COB probe



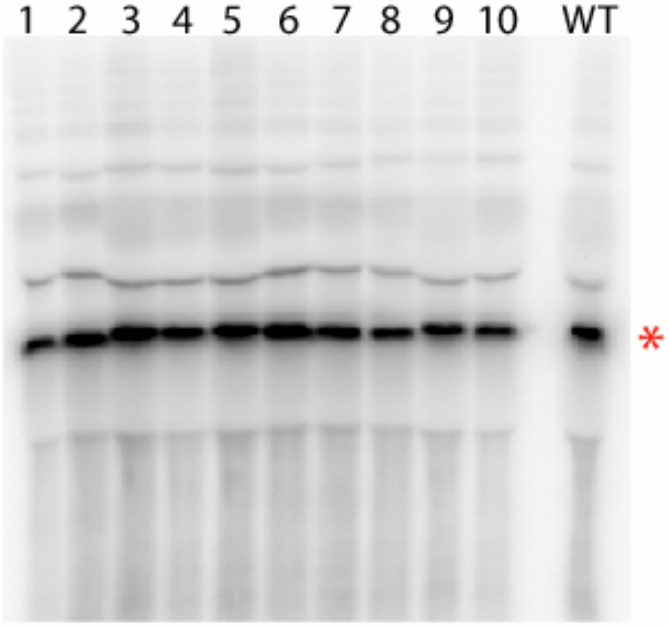
* indicates spliced message band

Northern blot 24

COX1 probe



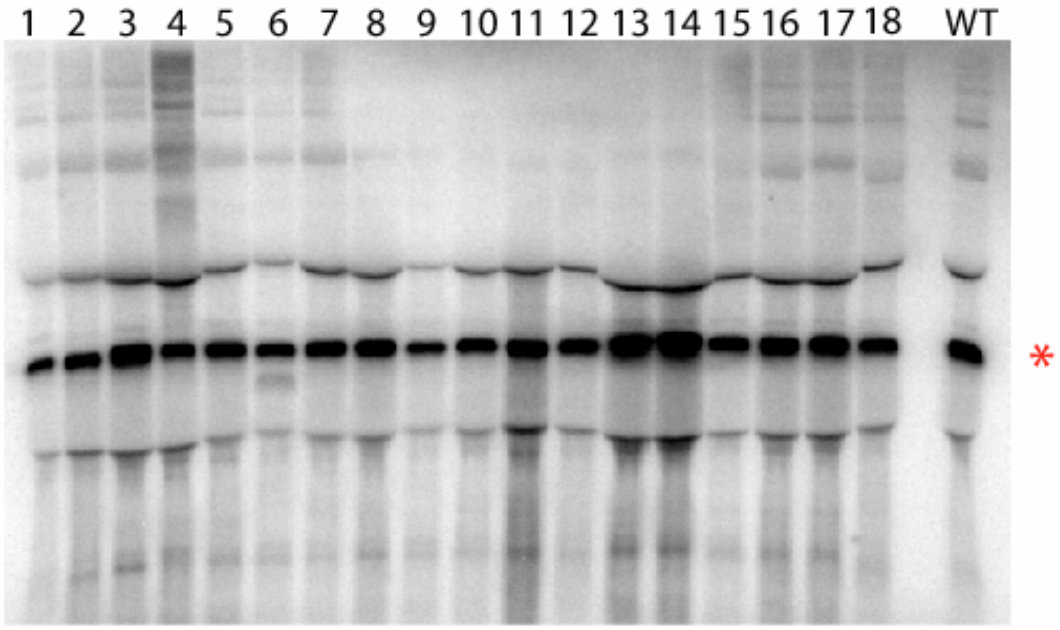
COB probe



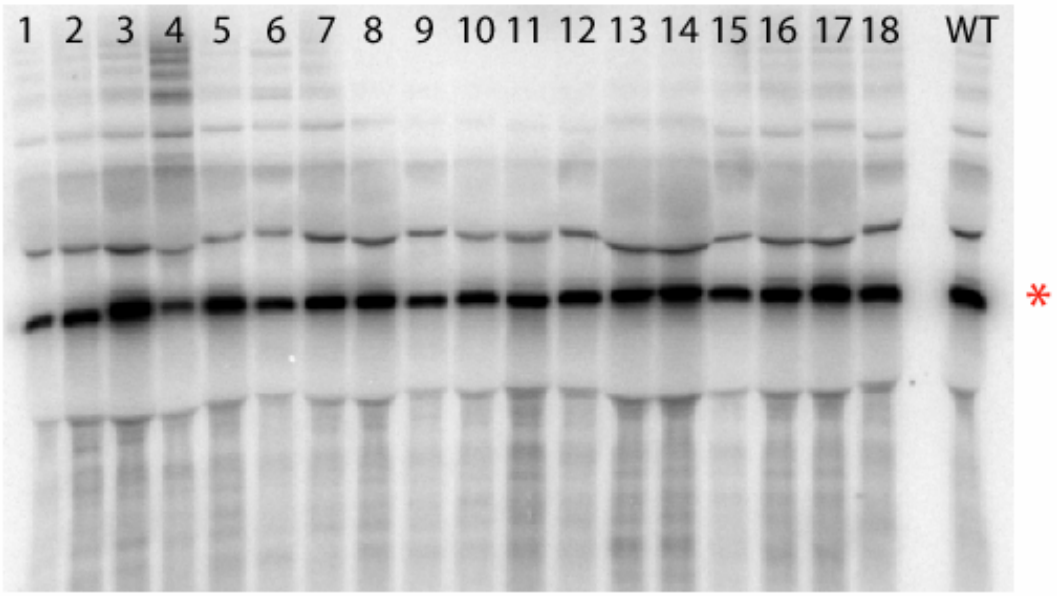
* indicates spliced message band

Northern blot 25

COX1 probe



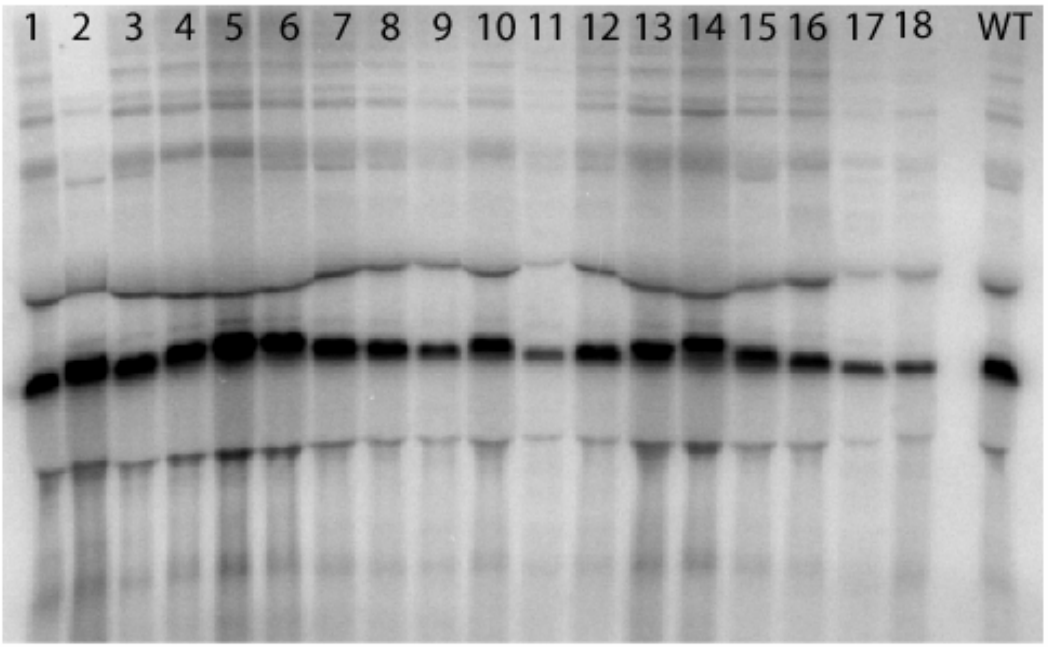
COB probe



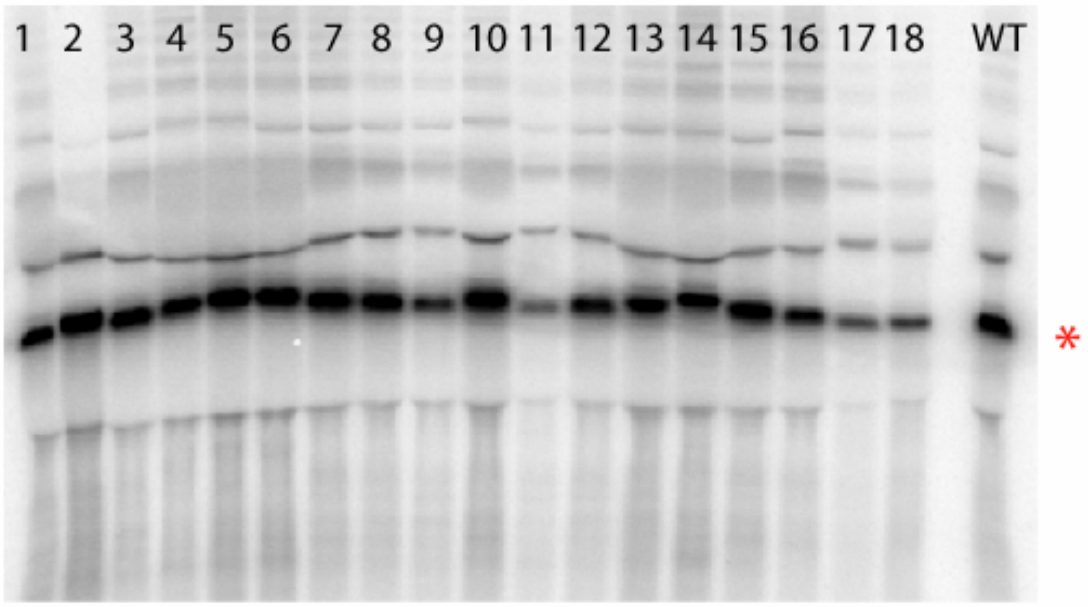
* indicates spliced message band

Northern blot 26

COX1 probe



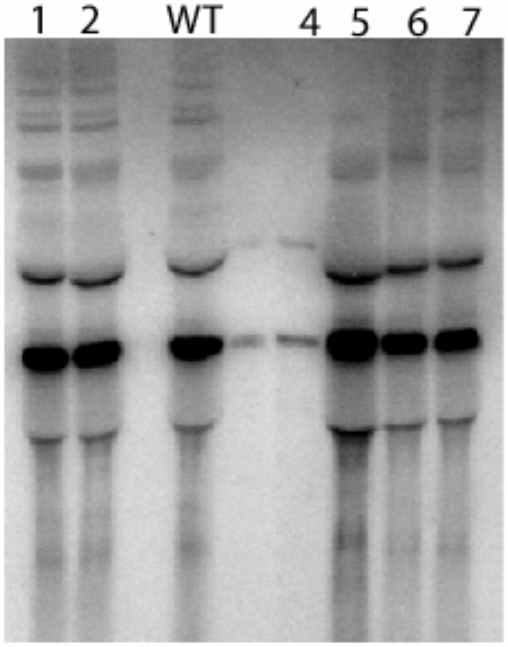
COB probe



* indicates spliced message band

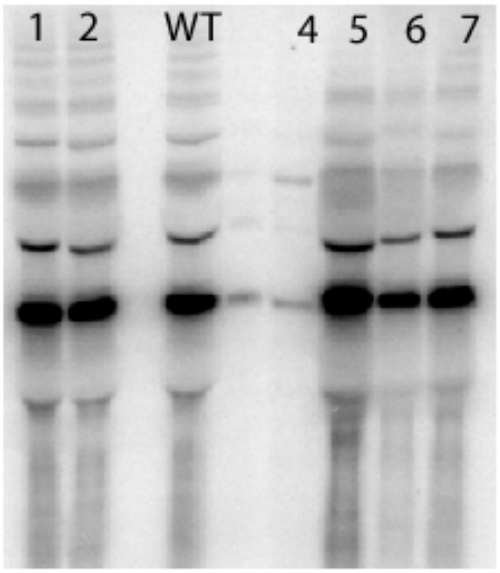
Northern blot 27

COX1 probe



*

COB probe



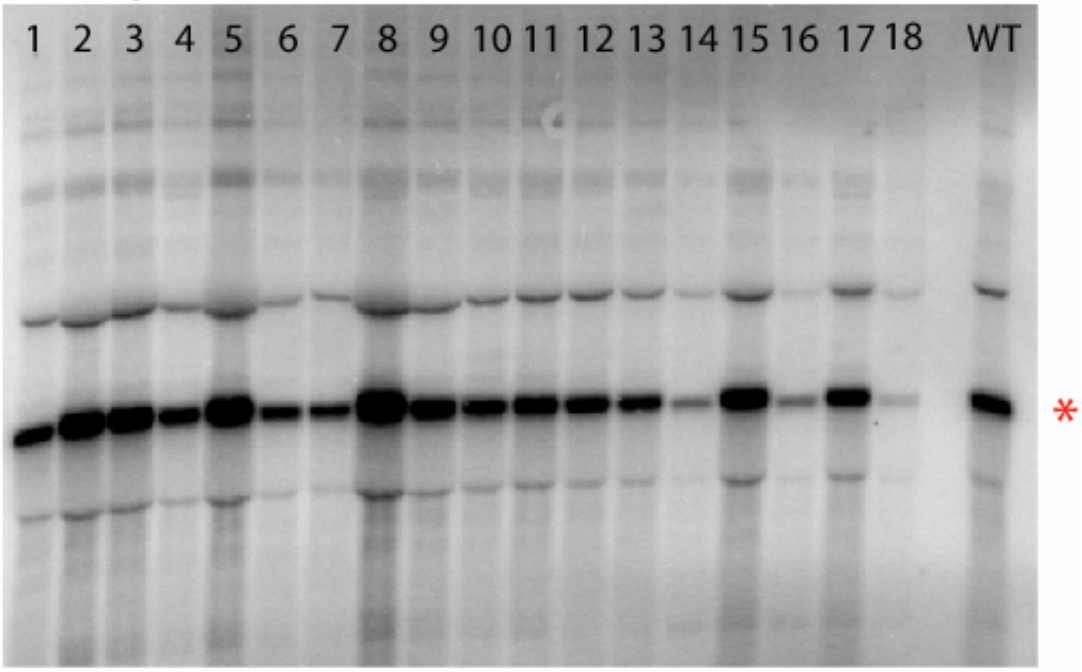
*

* indicates spliced message band

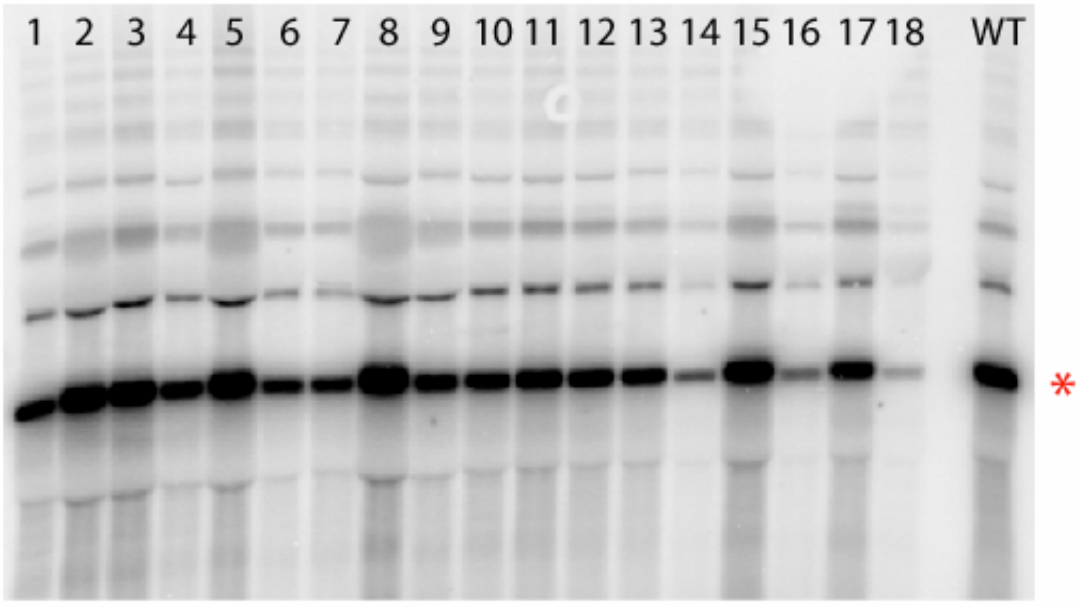
The lane to the left of wt does not contain a sample,
the bands in that lane are probably spill-over from the wt sample.

Northern blot 28

COX1 probe



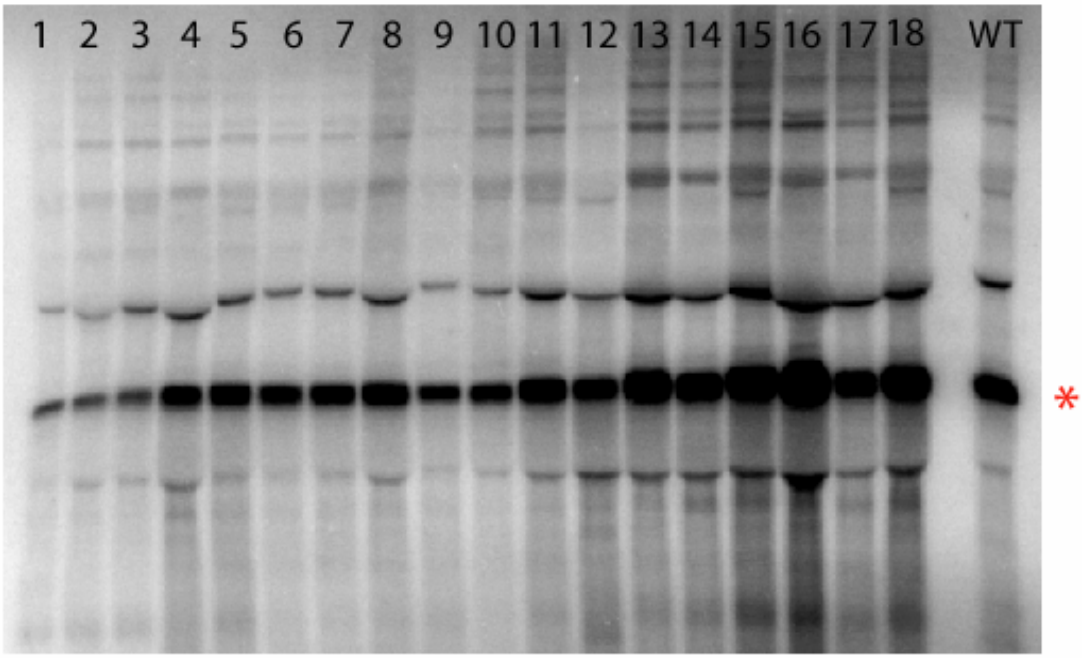
COB probe



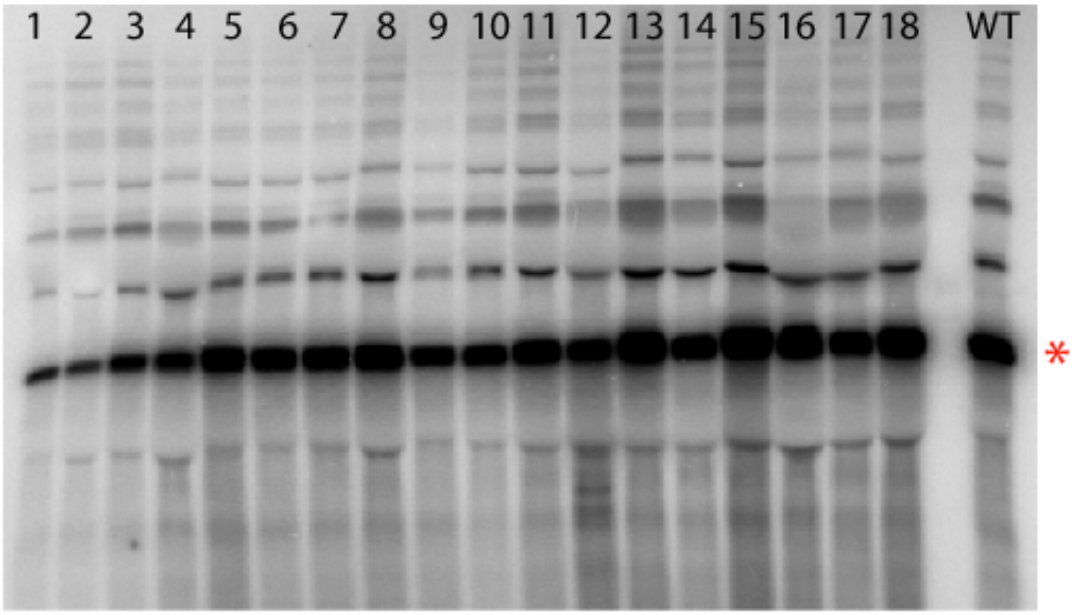
* indicates spliced message band

Northern blot 29

COX1 probe



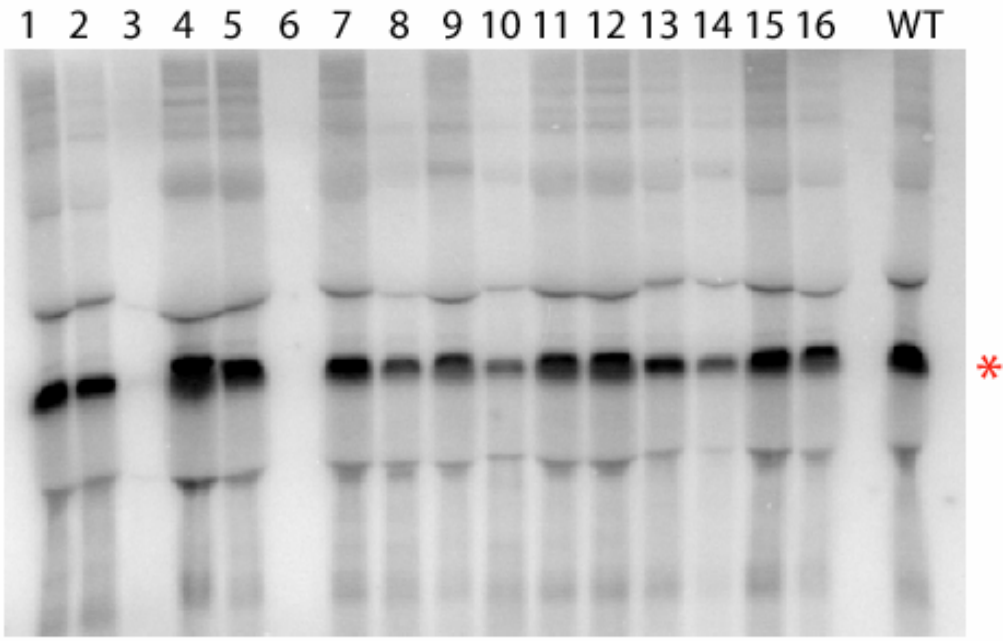
COB probe



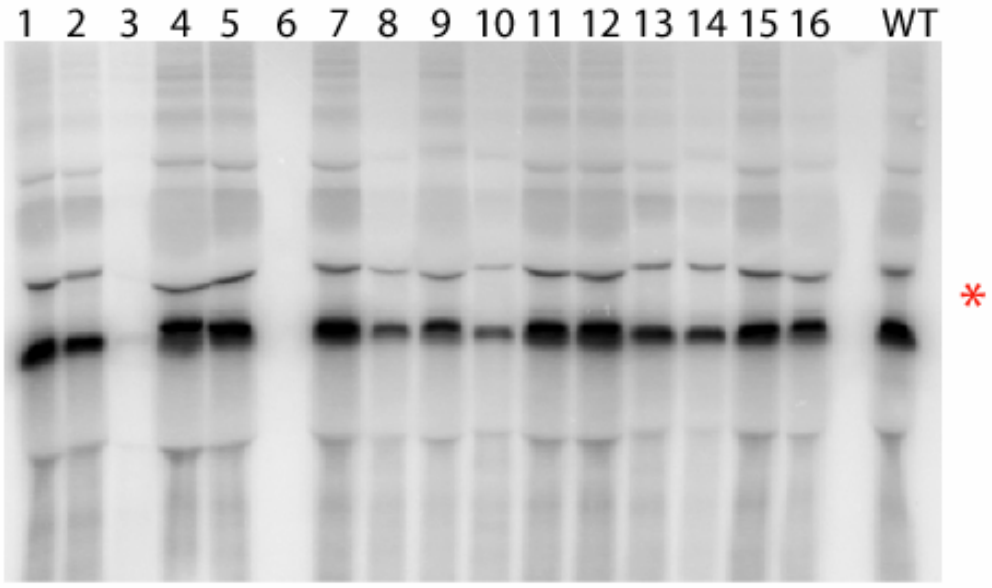
* indicates spliced message band

Northern blot 30

COX1 probe



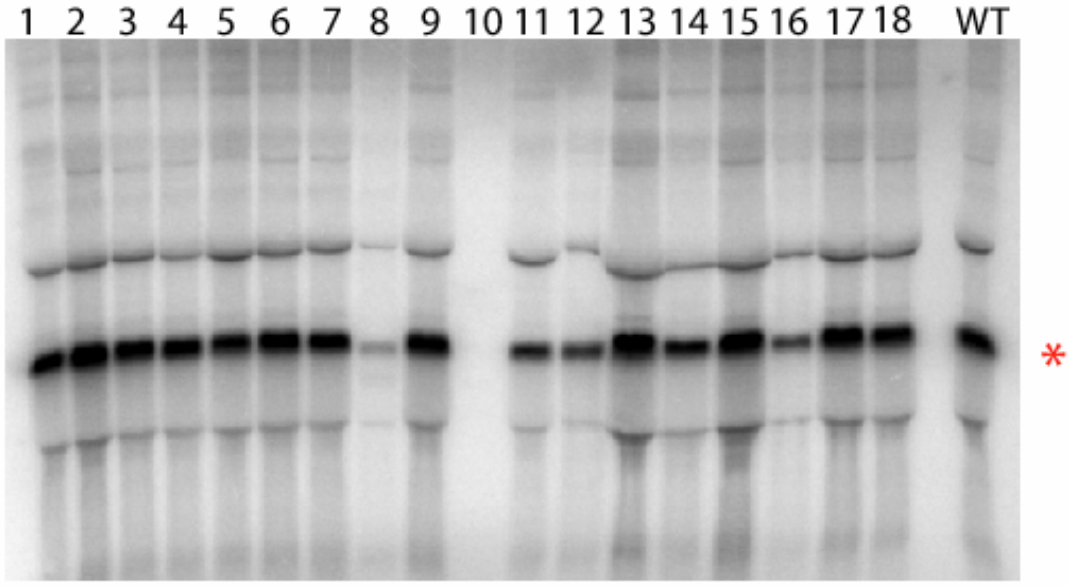
COB probe



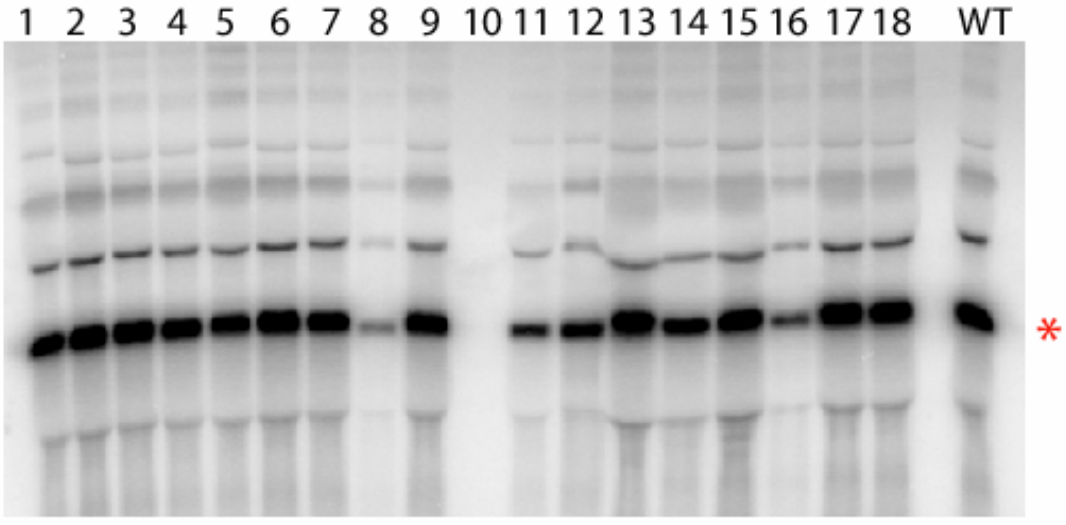
* indicates spliced message band

Northern blot 31

COX1 probe



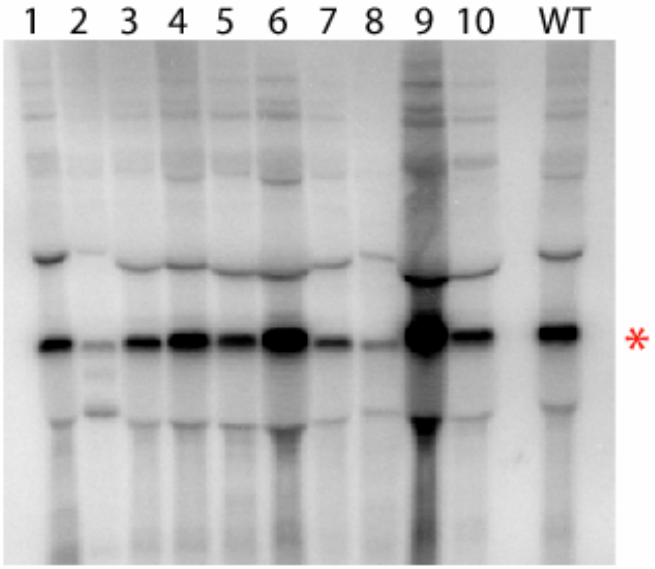
COB probe



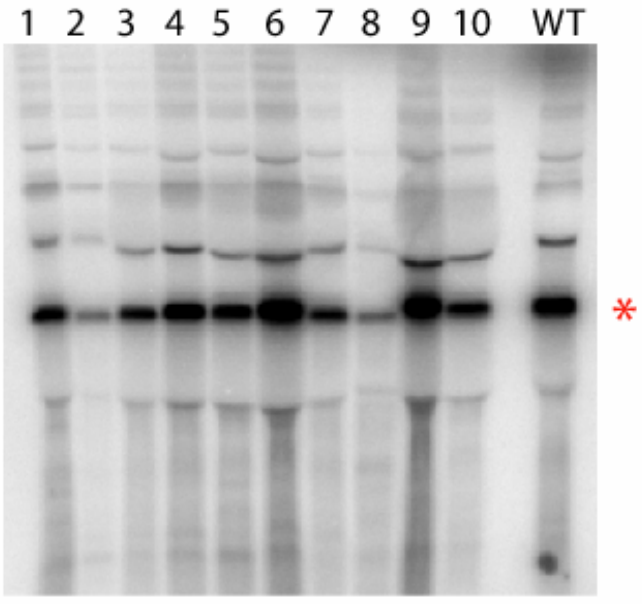
* indicates spliced message band

Northern blot 32

COX1 probe



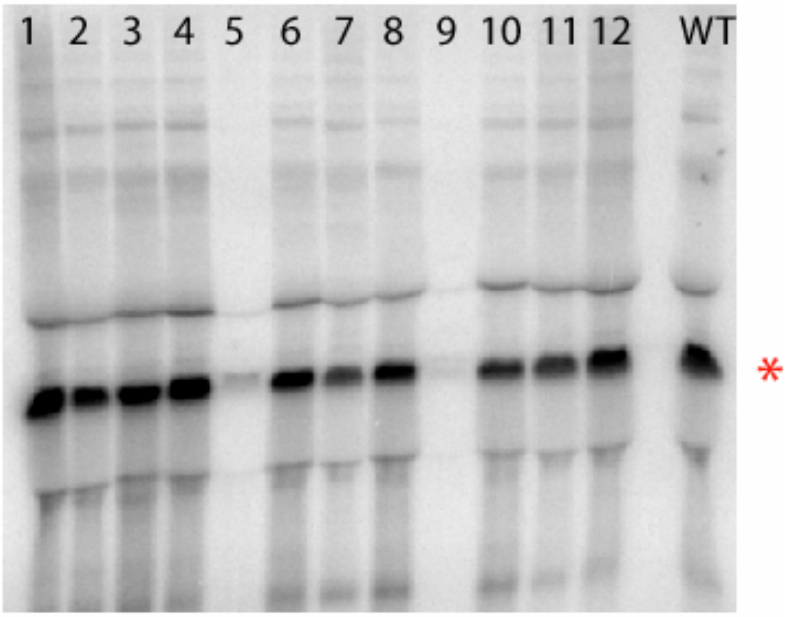
COB probe



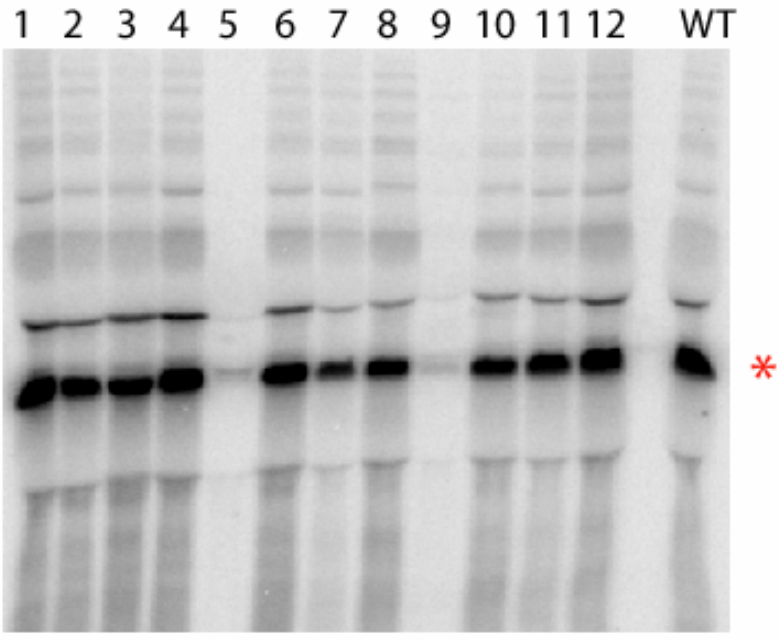
* indicates spliced message band

Northern blot 33

COX1 probe



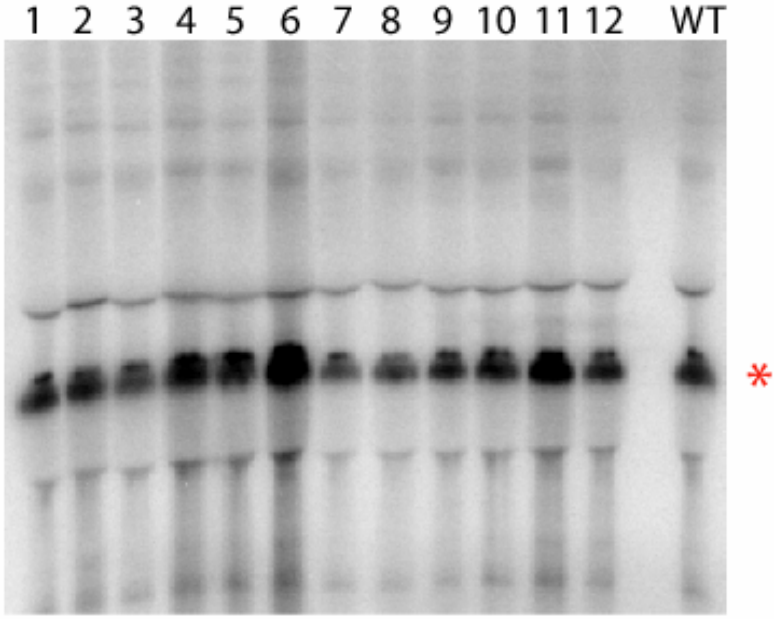
COB probe



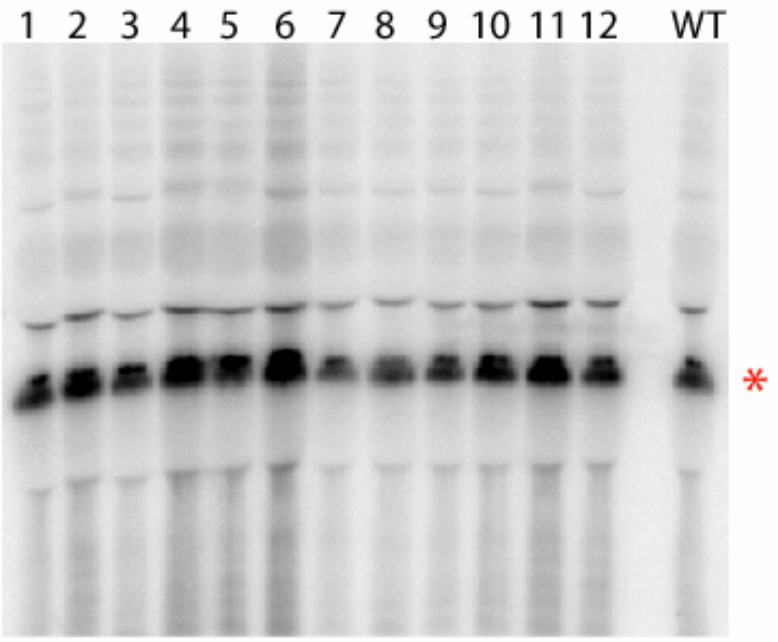
* indicates spliced message band

Northern blot 34

COX1 probe



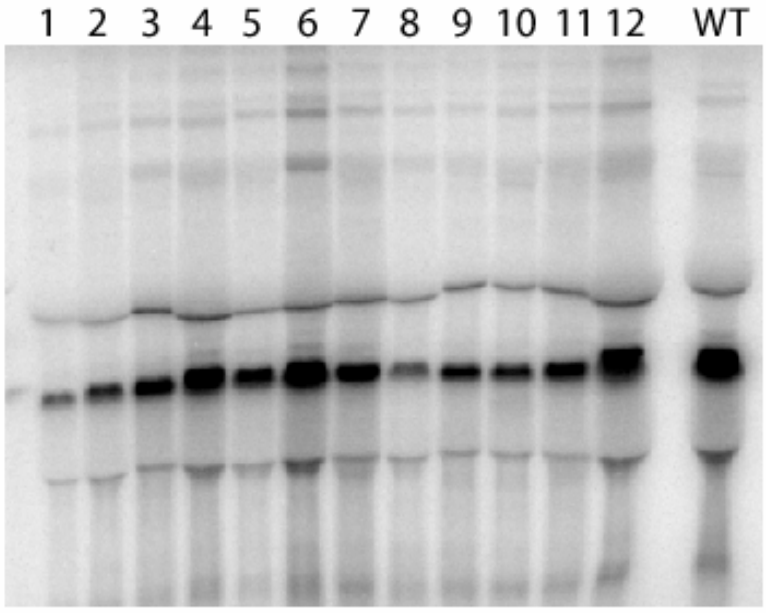
COB probe



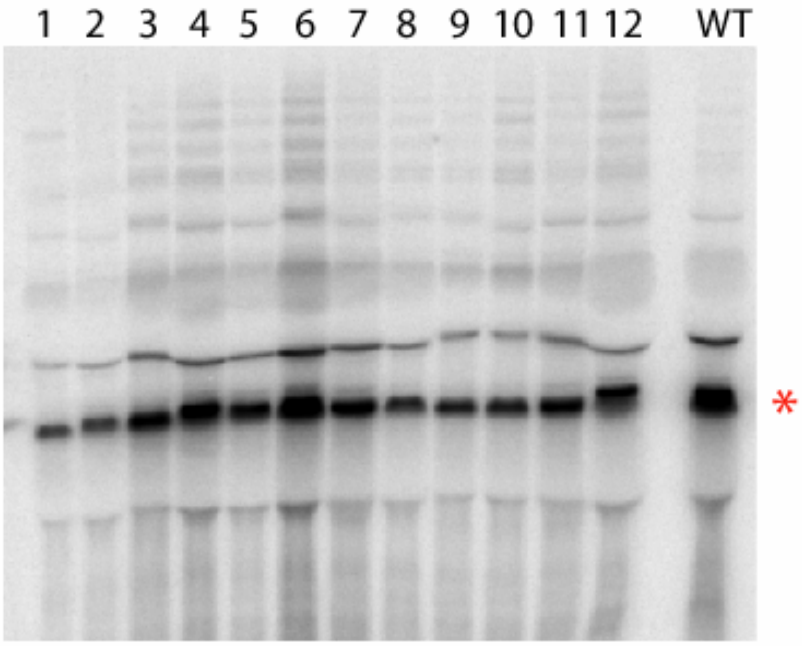
* indicates spliced message band

Northern blot 35

COX1 probe



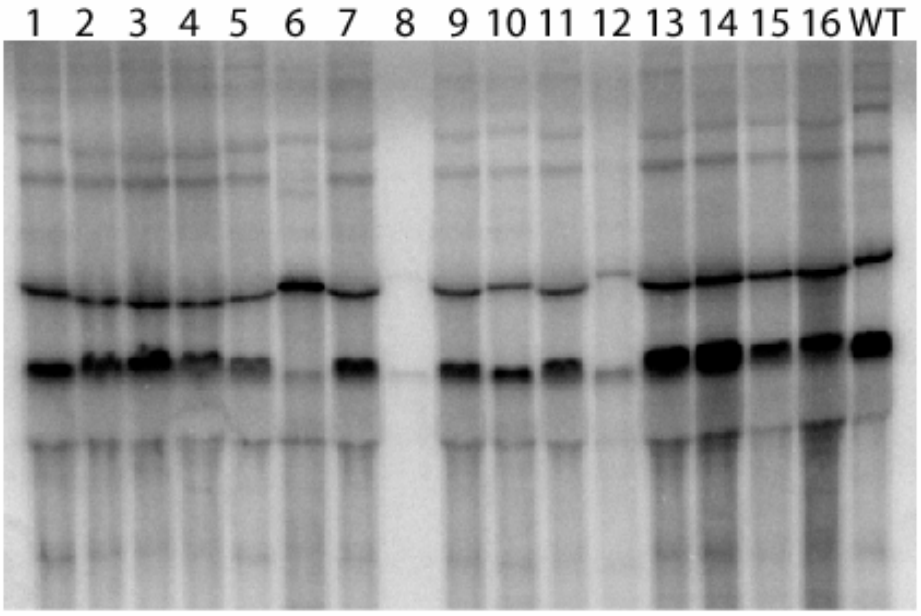
COB probe



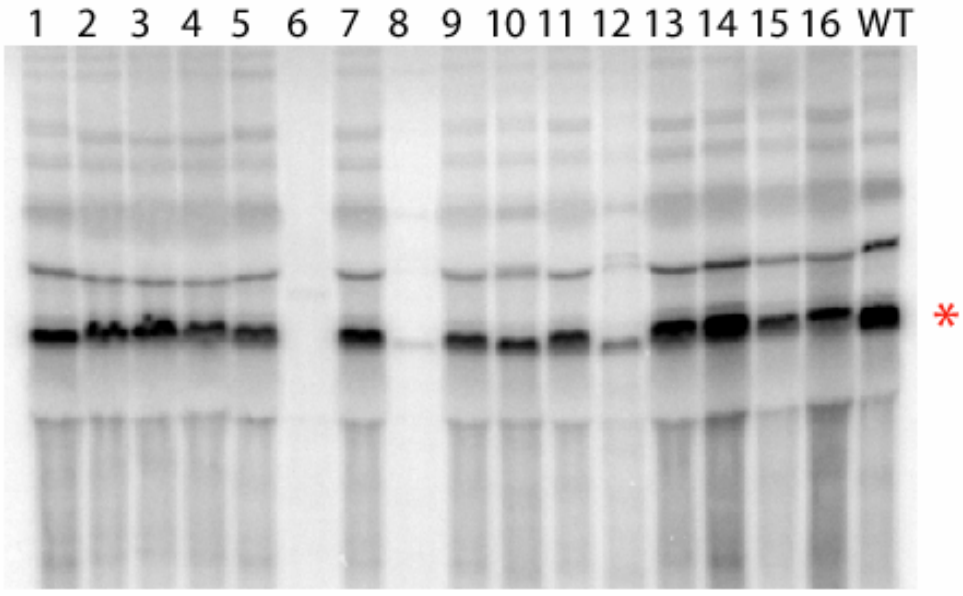
* indicates spliced message band

Northern blot 36

COX1 probe



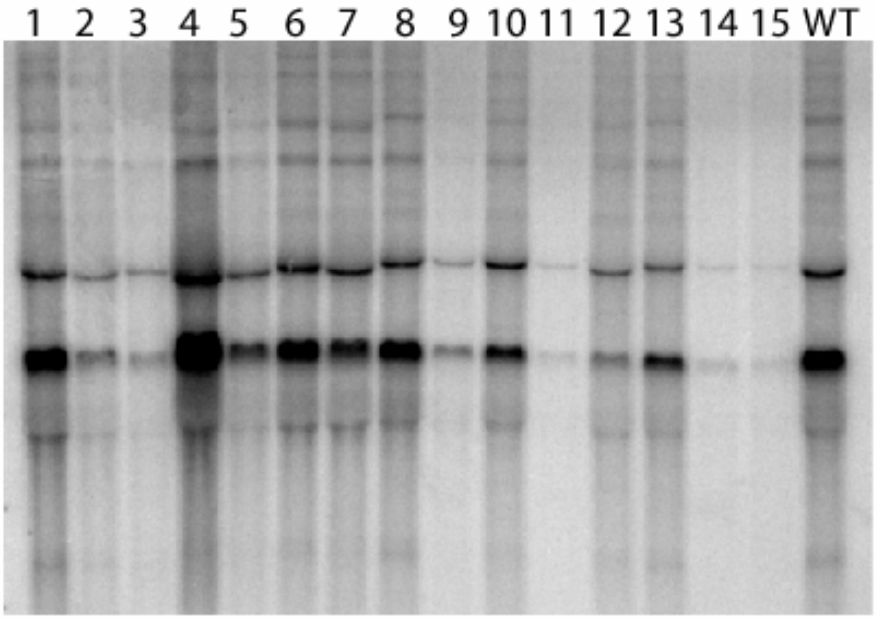
COB probe



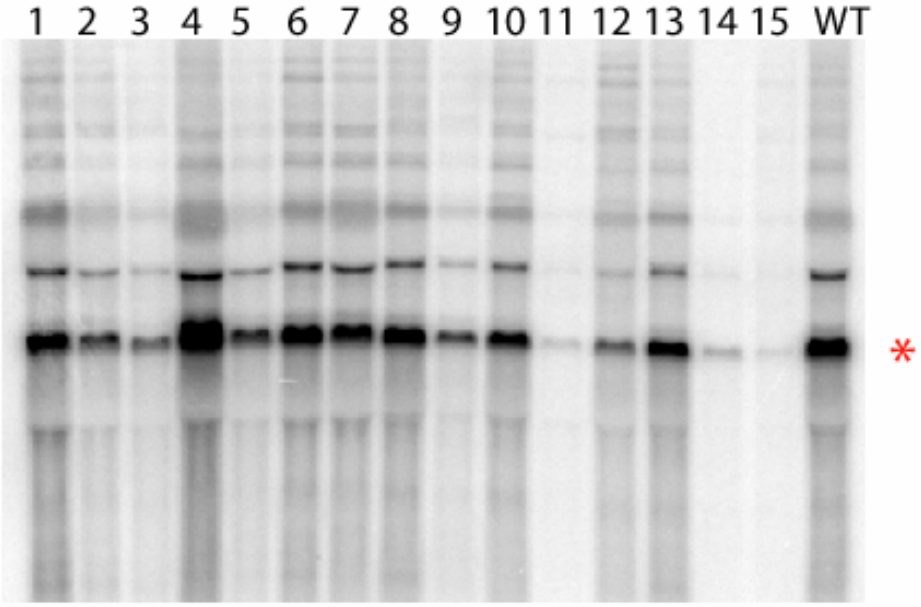
* indicates spliced message band

Northern blot 37

COX1 probe



COB probe



* indicates spliced message band

Blot legends listing ORF deletions orderd by blot, then by lane.

For a detailed explaination of the column headings see appendix B.

blot	lane	ORF name	Gene name	COX % splicing	COB % splicing	total RNA	growth temperature	library grid location
1	1		WT	0.88	0.72	+	30	wt
1	2	YLR077W	*FMP25*	0.86	0.64	+	30	1.3.e
1	3	YMR024W	*MRPL3*	0.80	0.61	+	30	1.5.e
1	4	YOL008W		0.71	0.62	+	30	3.1.d
1	5	YOL023W	*IFM1*	0.74	0.66	+++	30	3.1.f
1	7	YER083C	*RMD7*	0.76	0.68	++	30	4.5.c
1	8	YHL038C	*CBP2*	0.61	0.53	+	30	5.1.a
1	9	YHL005C	*(MRP4)*	0.80	0.72	++	30	5.1.f
1	10	YHR100C		0.75	0.67	+	30	5.2.g
1	11	YHR116W		0.73	0.68	+++	30	5.3.a
1	12	YLR201C	*FMP53*	0.86	0.59	+	30	5.5.g
1	13	YLR202C		0.76	0.73	++	30	5.5.h
1	14	YLR203C	*MSS51*	0.51	0.59	+	30	5.6.a
1	15	YKL137W		0.75	0.70	++	30	6.3.b
1	16	YKL167C	*MRP49*	0.77	0.66	++	30	6.3.h
1	17	YGR062C	*COX18*	0.68	0.67	+	30	6.4.h
1	18	YJR077C	*MIR1*	0.65	0.57	+/-	30	12.5.d
1	19	YDL107W	*MSS2*	0.78	0.69	++	30	13.3.d
1	20	YBR037C	*SCO1*	0.96	0.15	+	30	14.1.c
2	1		WT	0.84	0.93	+	30	wt
2	3	YDR197W	*CBS2*	0.17	0.00	+	30	8.2.d
2	4	YDR230W		0.75	0.86	+/-	30	8.3.a
2	5	YDR231C	*COX20*	0.85	0.82	++	30	8.3.b
2	6	YPR134W	*MSS18*	0.85	0.91	++	30	9.1.f
2	7	YKR016W	*FMP13*	0.76	0.90	++	30	10.3.a
2	8	YDR271C	*YDR271C*	0.69	0.52	+++	30	10.4.b
2	9	YER093C-A	*YER093C-A*	0.47	0.34	++	30	11.3.e
2	10	YNL252C	*MRPL17*	0.71	0.63	++	30	11.4.d
2	11	YPL132W	*COX11*	0.78	0.80	+	30	11.4.e
2	13	YJL046W	*YJL046W*	0.41	0.50	++	30	12.1.h
2	14	YJR120W		0.55	0.87	+	30	12.5.h
2	15	YNR020C		0.63	0.86	++	30	13.5.c
2	16	YNR025C		0.78	0.85	++	30	13.5.e
2	17	YBR037C	*SCO1*	0.66	0.74	+	30	14.1.c
2	18	YIL128W	*MET18*	0.57			30	14.3.B
2	19	YNL067W	*RPL9B*	0.72			30	14.4.B
2	20	YNL070W	*TOM7*	0.69			30	14.4.C

blot	lane	ORF name	Gene name	COX % splicing	COB % splicing	total RNA	growth temperature	library grid location
3	1	YNL080C		0.57	0.62	++	30	14.4.f
3	2	YNL120C		0.72	0.67	+++	30	14.5.c
3	3	YGL219C	*MDM34*	0.69	0.37	+/-	30	14.5.f
3	4	YPL183W-A	*GON5*	0.87	0.54	+	30	15.1.h
3	5	YKL053C-A	*MDM35*	0.87	0.05	++	30	15.2.b
3	6	YNR052C	*POP2*	0.66	0.55	+/-	30	15.3.e
3	7	YBR120C	*CBP6*	0.00	0.55	+	30	15.4.b
3	8	YDL032W	*YDL032W*	0.43	0.57	++	30	15.4.g
3	9	YDL033C	*MTO2*	0.51	0.73	+/-	30	15.4.h
3	10	YDL069C	*CBS1*	0.47	0.70	+	30	15.5.g
3	13	YDR493W	*FMP36*	0.77	0.69	++	30	16.5.f
3	16	YOR305W		0.00	0.19	+/-	30	2.5.f
3	17	YPR134W	*MSS18*	0.67	0.62	+	30	9.1.f
3	18	YDR523C	*SPS1*	0.27	0.67	+/-	30	19.5.c
3	19	YER087W	-	0.77	0.64	+	30	19.6.b
4	1	YLR201C	*FMP53*	0.92	0.85	+	30	5.5.g
4	2	YLR202C		0.89	0.66	+	30	5.5.h
4	3	YDR116C	*MRPL1*	0.88	0.63	+	30	4.1.f
4	5	YHL038C	*CBP2*	0.78	0.02	++	30	5.1.a
4	6	YDL033C	*MTO2*	0.54	0.38	+	30	15.4.h
4	7	YBR037C	*SCO1*	0.84	0.55	+	30	14.1.c
4	8	YGR062C	*COX18*	0.87	0.55	+++	30	6.4.h
4	9	YGL168W	*HUR1*	0.73	0.50	++	30	16.3.a
4	10	YLR203C	*MSS51*	0.71	0.12	++	30	5.6.a
4	11	YER093C-A	YER093C-A	0.52	0.35	+++	30	11.3.e
4	12	YNL080C		0.77	0.54	+	30	14.4.f
4	13	YER153C	*PET122*	0.77	0.52	+++	30	16.3.e
4	14	YHR100C		0.86	0.81	++	30	5.2.g
4	15	YDR195W	*REF2*	0.80	0.58	++	30	8.2.c
4	16	YGR222W	*PET54*	0.27	0.71	+	30	17.1.f
4	17	YML090W	*RPM2*	0.64	0.26	++	30	11.4.g
4	19	YOR305W		0.73	0.67	++	30	2.5.f
4	20	YJR120W		0.67	0.56	++	30	12.5.h
4	21	YJR077C	*MIR1*	0.80	0.53	++	30	12.5.d
4	22	YKL053C-A	*MDM35*		0.82	++	30	15.2.b

blot	lane	ORF name	Gene name	COX % splicing	COB % splicing	total RNA	growth temperature	library grid location
5	1	YOL050C		0.89	0.75	+++	30	3.2.c
5	2	YPL135W	*ISU1*	0.85	0.77	+++	30	3.4.c
5	3	YJL133W	*MRS3*	0.81	0.73	+++	30	7.3.f
5	4	YGL211W		0.81	0.75	+	30	8.3.e
5	5	YGL221C	*NIF3*	0.84	0.76	++	30	8.3.g
5	6	YPL002C	*SNF8*	0.83	0.74	++	30	8.6.f
5	7	YIL047C	*SYG1*	0.83	0.75	++	30	10.6.d
5	8	YLR346C		0.88	0.75	+++	30	15.1.d
5	9	YNL296W		0.90	0.73	+	30	18.2.b
5	10	YMR275C	*BUL1*	0.88	0.71	++	30	2.2.e
5	11	YML061C	*PIF1*	0.84	0.54	+++	30	1.4.f
5	12	YOR334W	*MRS2*	0.71	0.53	+++	30	2.6.a
5	13	YBR230C	*YBR230C*	0.91	0.77	+++	30	3.5.h
5	14	YKL037W		0.88	0.70	+++	30	6.1.b
6	1	YNL055C	*POR1*	0.85	0.78	+++	30	18.1.d
6	2	YCR081W	*SRB8*	0.82	0.78	+	30	13.1.h
6	3	YJR018W		0.88	0.87	++	30	13.2.f
6	4	YBR191W	*RPL21A*	0.83	0.90	+++	30	15.1.b
6	5	YAL008W	*FUN14*	0.90	0.88	+++	30	1.1.g
6	6	YLR090W	*XDJ1*	0.88	0.90	+++	30	1.3.h
6	7	YML009c	*MRPL39*	0.83	0.91	+++	30	1.5.a
6	8	YNL328C	*MDJ2*	0.92	0.91	+++	30	2.3.f
6	9	YPL222W	*FMP40*	0.92	0.90	+++	30	3.3.d
6	10	YHR155W		0.92	0.89	+++	30	5.3.e
6	11	YPR123C		0.90	0.90	+++	30	9.1.c
6	12	YNL139C	*RLR1,*	0.93	0.89	+++	30	14.1.f
6	13	YOR271C		0.90	0.88	+++	30	15.2.f
7	1	YMR275C	*BUL1*	0.75	0.67	+	37	2.2.e
7	2	YML061C	*PIF1*	0.44	0.16	++	37	1.4.f
7	3	YOR334W	*MRS2*	0.60	0.68	++	37	2.6.a
7	4	YCR081W	*SRB8*		0.65	+/-	37	13.1.h
7	6	YBR191W	*RPL21A*	0.84	0.66	+++	37	15.1.b
7	7	YBR230C	*YBR230C*	0.83	0.69	++	37	3.5.h
7	8	YKL037W		0.88	0.66	++	37	6.1.b
7	9	YNL055C	*POR1*	0.88	0.61	++	37	18.1.d
7	10	YAL008W	*FUN14*	0.71	0.67	+++	20	1.1.g
7	11	YLR090W	*XDJ1*	0.70	0.61	+++	20	1.3.h
7	12	YML009c	*MRPL39*	0.55	0.49	+++	20	1.5.a
7	13	YNL328C	*MDJ2*	0.76	0.66	+++	20	2.3.f
7	14	YPL222W	*FMP40*	0.82	0.71	+++	20	3.3.d
7	15	YHR155W		0.67	0.68	+++	20	5.3.e
7	16	YPR123C		0.93	0.63	+++	20	9.1.c
7	17	YNL139C	*RLR1,*	0.89	0.68	+++	20	14.1.f
7	18	YOR271C		0.90	0.70	+++	20	15.2.f

blot	lane	ORF name	Gene name	COX % splicing	COB % splicing	total RNA	growth temperature	library grid location
8	1	YOR332W	*VMA4*	1.00	0.94	++	37	2.5.h
8	2	YDR069C	*DOA4*	0.90	0.92	+++	37	3.6.h
8	3	YGR183C	*QCR9*	0.98	0.93	++	37	4.6.e
8	4	YHR100C		0.00	0.00	+	37	5.2.g
8	5	YCL008C	*STP22*	0.88	0.89	++	37	5.4.b
8	6	YKL041W	*VPS24*	0.90	0.91	+++	37	6.1.d
8	7	YBR227C	*MCX1*	0.94	0.91	++	37	3.5.g
8	9	YBL007C	*SLA1*	0.91	0.87	+++	37	9.4.b
8	10	YBL058W	*SHP1*	0.94	0.90	++	37	9.5.c
8	11	YDR270W	*CCC2*	0.93	0.93	++	37	10.4.a
8	12	YMR077C	*VPS20*	0.95	0.92	++	37	11.2.e
8	13	YJL095W	*BCK1*	0.94	0.93	+++	37	12.1.b
9	1	YAL026C	*DRS2*	0.82	0.74	+++	20	1.1.e
9	2	YLL041C	*SDH2*	0.60	0.69	+	20	1.2.e
9	3	YLR027C	*AAT2*	0.81	0.69	++	20	1.2.g
9	4	YMR275C	*BUL1*	0.77	0.71	++	20	2.2.e
9	5	YPL270W	*MDL2*	0.65	0.63	++	20	3.2.g
9	6	YPL215W	*CBP3*	0.66	0.73	++	20	3.3.e
9	7	YPL188W	*POS5*	0.64	0.68	+++	20	3.3.f
9	8	YHL005C	*(MRP4)*	0.05	0.17	++	20	5.1.f
9	9	YHR013C	*ARD1*	0.68	0.68	+++	20	5.2.a
9	10	YHR037W	*PUT2*	0.53	0.66	+	20	5.2.b
9	11	YHR117W	*TOM71*	0.58	0.67	+	20	5.3.b
9	12	YHR179W	*OYE2*	0.11	0.50	+	20	5.3.f
9	13	YDR150W	*NUM1*	0.66	0.69	+++	20	8.1.c
9	14	YMR138W	*CIN4*	0.78	0.75	+	20	15.1.g
9	15	YMR058W	*FET3*	0.68	0.72	++	20	16.4.d
9	16	YNL003C	*PET8*	0.67	0.73	++	20	17.5.f
9	17		WT	0.82	0.94	++	20	wt

blot	lane	ORF name	Gene name	COX % splicing	COB % splicing	total RNA	growth temperature	library grid location
10	**1**	YKL054C	*VID31*	**0.81**	**0.74**	++	37	6.1.e
10	**2**	YFR033C	*QCR6*	**0.73**	**0.65**	++	37	9.3.a
10	**3**	YJR102C	*VPS25*	**0.67**	**0.77**	++	37	12.3.h
10	**4**	YMR060C	*TOM37*	**0.70**	**0.84**	+	37	12.5.a
10	**5**	YJR074W	*MOG1*	**0.62**	**0.72**	+++	37	12.5.c
10	**6**	YCR081W	*SRB8*	**0.65**	**0.69**	++	37	13.1.h
10	**7**	YJR018W		**0.66**	**0.64**	++	37	13.2.f
10	**8**	YDL142C	*CRD1*	**0.71**	**0.81**	+	37	13.3.h
10	**9**	YNR006W	*VPS27*	**0.69**	**0.80**	++	37	13.5.a
10	**10**	YNL136W	*EAF7*	**0.47**	**0.53**	++	37	14.1.g
10	**11**	YNL133C	*FYV6*	**0.82**	**0.79**	+++	37	14.1.h
10	**12**	YGR155W	*CYS4*	**0.51**	**0.64**	++	37	15.1.c
10	**13**	YLR358C		**0.77**	**0.60**	+++	37	15.1.e
10	**14**	YGR229C	*SMI1*	**0.70**	**0.66**	++	37	17.1.g
10	**15**	YGL070C	*RPB9*	**0.81**	**0.86**	+	37	17.2.c
10	**16**	YGL095C	*VPS45*	**0.76**	**0.86**	+	37	17.2.d
10	**17**	YKR006C	*MRPL13*	**0.78**	**0.69**	++	37	17.2.e
10	**18**		WT	**0.91**	**0.90**	++	37	wt
11	**1**	YAL044C	*GCV3*	**0.68**	**0.84**	++	30	1.1.c
11	**2**	YAL039C	*CYC3*	**0.66**	**0.83**	++	30	1.1.d
11	**3**	YAL010C	*MDM10*	**0.83**	**0.99**	++	30	1.1.f
11	**4**	YLL006W	*MMM1*	**0.79**	**0.91**	++	30	1.2.a
11	**5**	YLL009C	*COX17*	**0.83**	**0.91**	+++	30	1.2.b
11	**6**	YLL033W		**0.78**	**0.90**	++	30	1.2.c
11	**7**	YLL040C	*VPS13*	**0.68**	**0.89**	++	30	1.2.d
11	**8**	YLR025W	*SNF7*	**0.67**	**0.86**	+++	30	1.2.f
11	**9**	YLR056W	*ERG3*	**0.71**	**0.91**	+	30	1.2.h
11	**10**	YLR059C	*REX2*	**0.77**	**0.88**	++	30	1.3.a
11	**11**	YLR083C	*EMP70*	**0.76**	**0.90**	++	30	1.3.f
11	**12**	YLR087C	*CSF1*	**0.72**	**0.87**	+++	30	1.3.g
11	**13**	YML089C		**0.67**	**0.76**	++	30	1.4.c
11	**14**	YML030W		**0.71**	**0.88**	++	30	1.4.g
11	**15**	YMR021C	*MAC1*	**0.73**	**0.78**	+++	30	1.5.c
11	**16**	YMR031W-A		**0.65**	**0.82**	++	30	1.5.f
11	**17**	YMR035W	*IMP2*	**0.62**	**0.85**	++	30	1.5.g
11	**18**		WT	**0.80**	**0.88**	++	30	wt

blot	lane	ORF name	Gene name	COX % splicing	COB % splicing	total RNA	growth temperature	library grid location
12	**1**	YMR151W	*YIM2*	**0.65**	**0.86**	+++	30	1.6.a
12	**2**	YMR150C	*IMP1*	**0.23**	**0.76**	++	30	1.6.b
12	**3**	YMR167W	*MLH1*	**0.72**	**0.80**	++	30	1.6.e
12	**4**	YMR188C	*MRPS17*	**0.88**	**0.91**	+++	30	1.6.g
12	**5**	YMR193W	*MRPL24*	**0.79**	**0.89**	++	30	2.1.a
12	**6**	YMR202W	*ERG2*	**0.88**	**0.91**	+++	30	2.1.c
12	**7**	YMR207C	*HFA1*	**0.78**	**0.89**	+++	30	2.1.d
12	**8**	YMR245W		**0.68**	**0.92**	++	30	2.1.h
12	**9**	YMR256C	*COX7*	**0.41**	**0.93**	++	30	2.2.b
12	**10**	YMR257C	*PET111*	**0.54**	**0.88**	++	30	2.2.c
12	**11**	YMR280C	*CAT8*	**0.66**	**0.92**	++	30	2.2.f
12	**12**	YMR282C	*AEP2*	**0.74**	**0.92**	+++	30	2.2.g
12	**13**	YMR285C	*NGL2*	**0.64**	**0.89**	++	30	2.2.h
12	**14**	YNL297C	*MON2*	**0.74**	**0.96**	++	30	2.3.g
12	**15**	YNL280C	*ERG24*	**0.77**	**0.93**	++	30	2.3.h
12	**16**	YNL259C	*ATX1*	**0.74**	**0.89**	++	30	2.4.a
12	**17**	YOR014W	*RTS1*	**0.76**	**0.93**	++	30	2.4.b
12	**18**	YOR017W	*PET127*	**0.49**	**0.14**	++	30	2.4.c
12	**19**		WT		**0.93**	++	30	wt
13	**1**	YOR026W	*BUB3,*	**0.82**	**0.82**	+++	30	2.4.e
13	**2**	YOR036W	*PEP12*	**0.80**	**0.86**	+++	30	2.4.f
13	**3**	YOR037W	*CYC2*	**0.81**	**0.84**	+++	30	2.4.g
13	**4**	YOR065W	*CYT1*	**0.75**	**0.69**	++	30	2.5.b
13	**5**	YOR073W		**0.79**	**0.79**	++	30	2.5.c
13	**6**	YOR290C	*SNF2*	**0.83**	**0.81**	+++	30	2.5.d
13	**7**	YOR339C	*UBC11*	**0.84**	**0.86**	++	30	2.6.b
13	**8**	YOR350C	*MNE1*	**0.56**	**0.77**	++	30	2.6.c
13	**9**	YOR354C	*MSC6*	**0.86**	**0.84**	++	30	2.6.d
13	**10**	YOR358W	*HAP5*	**0.63**	**0.69**	++	30	2.6.f
13	**11**	YOR360C	*PDE2*	**0.87**	**0.89**	+++	30	2.6.g
13	**12**		WT		**0.84**	++	30	wt

blot	lane	ORF name	Gene name	COX % splicing	COB % splicing	total RNA	growth temperature	library grid location
14	1	YOL001W	*PHO80*	0.75	0.94	++	30	3.1.a
14	2	YOL009C	*MDM12*	0.55	0.98	++	30	3.1.e
14	3	YOL027C	*MDM38*	0.59	0.95	++	30	3.1.g
14	4	YOL071W	*EMI5*	0.75	0.91	++++		3.2.d
14	5	YOL081W	*IRA2*	0.82	0.92	+++		3.2.e
14	6	YPL265W	*DIP5*	0.87	0.92	+++	30	3.2.h
14	7	YPL262W	*FUM1*	0.78	0.96	++	30	3.3.a
14	8	YPL254W	*HFI1*	0.40	0.96	+/-	30	3.3.b
14	9	YPL234C	*TFP3*	0.61	0.93	+++	30	3.3.c
14	10	YPL174C	*NIP100*	0.49	0.94	++++	30	3.3.g
14	11	YPL172C	*COX10*	0.38	1.00	++	30	3.4.a
14	12	YPL098C	YPL098C	0.52	0.88	++	30	3.4.h
14	13	YBR185C	*MBA1*	0.46	0.95	++	30	3.5.d
14	14	YBR221C	*PDB1*	0.73	0.92	++	30	3.5.f
15	1	YDR120C	*TRM1*	0.94	0.87	++++	30	4.1.g
15	2	YDR129C	*SAC6*	0.96	0.88	++++	30	4.2.a
15	3	YDR350C	*TCM10*	0.59	0.99	++	30	4.2.d
15	4	YDR375C	*BCS1*	0.89	0.88	++	30	4.2.e
15	5	YDR388W	*RVS167*	0.96	0.82	++++	30	4.2.g
15	6	YDR393W	*SHE9*	0.95	0.82	+++	30	4.2.h
15	7	YEL024W	*RIP1*	0.93	0.68	+++	30	4.3.d
15	8	YEL027W	*CUP5*	0.96	0.76	++	30	4.3.e
15	9	YEL051W	*VMA8*	0.95	0.79	++	30	4.3.g
15	10	YER058W	*PET117*	0.97	0.63	++	30	4.4.e
15	11	YER061C	*CEM1*	0.94	0.47	++	30	4.4.f
15	12	YER068W	*MOT2*	0.95	0.80	+++	30	4.4.g
15	13	YER087C-A		0.97	0.87	++++	30	4.5.e
15	14	YGR174C	*CBP4*	0.93	0.69	++	30	4.6.b
15	15	YGR178C	*PBP1*	0.98	0.87	++++	30	4.6.c
15	16	YGR183C	*QCR9*	0.94	0.80	++	30	4.6.e
15	17		WT	0.95	0.87	++	30	WT
16	1	YHR038W	*RRF1*	0.94	0.90	+++	30	5.2.c
16	2	YHR051W	*COX6*	0.92	0.92	++	30	5.2.d
16	3	YHR060W	*VMA22*	0.00	0.46	+	30	5.2.e
16	4	YHR189W	*PTH1*	0.97	0.87	++	30	5.3.g
16	5	YCR009C	*RVS161*	0.93	0.83	++++	30	5.4.g
16	6	YLR149C		0.94	0.87	+++	30	5.5.b
16	7	YLR168C	*(MSF1)*	0.91	0.82	++	30	5.5.c
16	8	YLR182W	*SWI6*	0.93	0.80	+++	30	5.5.d
16	9	YLR218C		0.94	0.85	+++	30	5.6.c
16	10	YKL002W	*DID4*	0.86	0.69	++	30	5.6.d
16	11	YKL016C	*ATP7*	0.96	0.87	+	30	5.6.g

blot	lane	ORF name	Gene name	COX % splicing	COB % splicing	total RNA	growth temperature	library grid location
17	1	YKL055C	*OAR1*	0.95	0.94	+++	30	6.1.g
17	2	YKL080W	*VMA5*	0.90	0.90	+	30	6.1.h
17	3	YKL085W	*MDH1*	0.95	0.85	++++	30	6.2.a
17	4	YKL087C	*CYT2*	0.82	0.65	++	30	6.2.b
17	5	YKL106W	*AAT1*	0.85	0.20	++++	30	6.2.c
17	6	YKL109W	*HAP4*	0.81	0.53	+	30	6.2.d
17	7	YKL118W		0.98	0.78	++++	30	6.2.f
17	8	YKL119C	*VPH2*	0.92	0.73	++	30	6.2.g
17	9	YKL138C	*MRPL31*	0.92	0.89	+++	30	6.3.c
17	10	YKL148C	*SDH1*	0.91	0.82	++	30	6.3.d
17	11	YGR036C	*CWH8,*	0.87	0.79	++	30	6.4.f
17	12	YGR084C	*MRP13*	0.91	0.80	++	30	6.5.b
17	13	YOR125C	*CAT5,*	0.75	0.72	+	30	6.6.a
17	14	YOR135C		0.85	0.63	++++	30	6.6.c
17	15	YOR136W	*IDH2*	0.91	0.66	++++	30	6.6.d
17	16	YOR196C	*LIP5*	0.94	0.79	++++	30	6.6.g
17	17			0.90	0.84	++	30	WT
18	1	YOR221C	*MCT1*	0.82	0.99	++	30	7.1.f
18	2	YJL176C	*SWI3*	0.87	0.83	++++	30	7.1.g
18	3	YJL193W		0.90	0.89	+++	30	7.2.g
18	4	YJL180C	*ATP12*	0.00	0.00	++	30	7.2.h
18	5	YJL166W	*QCR8*	0.81	0.79	+++	30	7.3.b
18	6	YLR377C	*FBP1*	0.87	0.83	+++	30	7.4.d
18	7	YLR393W	*ATP10*	0.50	0.32	+++	30	7.4.f
18	8	YLR395C	*COX8*	0.89	0.83	++++	30	7.4.g
18	9	YLR417W	*VPS36*	0.88	0.80	++++	30	7.4.h
18	10	YLR234W	*TOP3*	0.86	0.83	++	30	7.5.b
18	11	YLR239C	*LIP2*	0.79	0.74	++	30	7.5.D
18	12	YLR262C	*YPT6*	0.76	0.74	++	30	7.5.G
18	13	YLR268W	*SEC22*	0.83	0.81	++	30	7.5.H
18	14	YLR294C		0.84	0.77	++	30	7.6.F
18	15	YLR295C	*ATP14*	0.90	0.84	++++	30	7.6.G
18	16	YLR304C	*ACO1*	0.84	0.79	++	30	7.6.H
18	17		WT	0.87	0.84	+++	30	WT

blot	lane	ORF name	Gene name	COX % splicing	COB % splicing	total RNA	growth temperature	library grid location
19	1	YDR148C	*KGD2*	0.90	0.78	++	30	8.1.B
19	2	YDR204W	*COQ4*	0.73	0.76	+	30	8.2.E
19	3	YDR225W	*HTA1*	0.89	0.91	++	30	8.2.G
19	4	YDR226W	*ADK1*	0.90	0.94	++	30	8.2.H
19	5	YGL218W		0.95	0.96	+	30	8.3.F
19	6	YGL236C	*MTO1*	0.93	1.00	+	30	8.4.A
19	8	YGL250W		0.94	0.90	+++	30	8.4.C
19	9	YGR020C	*VMA7,*	0.92	0.72	+++	30	8.4.G
19	10	YPL084W	*BRO1*	0.92	0.81	++	30	8.5.B
19	11	YPL078C	*ATP4*	0.89	0.81	++	30	8.5.C
19	12	YPL066W		0.86	0.69	+++	30	8.5.F
19	13	YPL060W	*LPE10*	0.86	0.77	++	30	8.5.H
19	14	YPL042C	*SSN3*	0.79	0.61	+	30	8.6.A
19	15	YPL040C	*ISM1*	0.88	0.80	+++	30	8.6.B
19	16		WT	0.89	0.87	+++	30	WT
20	1	YAL010C	*MDM10*	0.68	0.90	++	30	1.1.F
20	2	YMR021C	*MAC1*	0.80	0.91	++	30	1.5.c
20	3	YMR150C	*IMP1*	0.22	0.92	++	30	1.6.b
20	4	YMR256C	*COX7*	0.75	0.90	++	30	2.2.b
20	5	YMR257C	*PET111*	0.61	0.91	++	30	2.2.c
20	6	YOR017W	*PET127*	0.26	0.12	++	30	2.4.c
20	7	YOR350C	*MNE1*	0.40	0.85	++	30	2.6.c
20	8	YDR350C	*TCM10*	0.00	0.00	++	30	4.2.d
20	9	YEL027W	*CUP5*	0.65	0.92	+	30	4.3.e
20	10	YEL051W	*VMA8*	0.72	0.88	+++	30	4.3.g
20	11			0.74	0.91	++	30	WT
21	1	YER058W	*PET117*	0.83	0.70	++	30	4.4.e
21	2	YHR060W	*VMA22*	0.76	0.70	+++	30	5.2.e
21	3	YKL002W	*DID4*	0.58	0.58	++	30	5.6.d
21	4	YKL016C	*ATP7*	0.00	0.00	++	30	5.6.g
21	5	YKL080W	*VMA5*	0.69	0.68	+++	30	6.1.h
21	6	YKL087C	*CYT2*	0.82	0.60	+	30	6.2.b
21	7	YKL106W	*AAT1*	0.69	0.19	+++	30	6.2.c
21	8	YKL109W	*HAP4*	0.68	0.54	++	30	6.2.d
21	9	YKL119C	*VPH2*	0.80	0.78	++++	30	6.2.g
21	10	YDR197W	*CBS2*	0.10	0.09	+	30	8.2.D
21	11		WT	0.71	0.84	+	30	WT

blot	lane	ORF name	Gene name	COX % splicing	COB % splicing	total RNA	growth temperature	library grid location
22	1	YPR124W	*CTR1*	0.93	0.92	++	30	9.1.D
22	2	YPR146C		0.96	0.96	++	30	9.1.H
22	3	YPR166C	*MRP2*	0.95	0.97	+	30	9.2.C
22	4	YPR173C	*VPS4*	0.92	0.96	+	30	9.2.D
22	5	YPR191W	*QCR2*	0.92	0.96	++	30	9.2.E
22	6	YPR197C		0.92	0.98	++	30	9.2.F
22	7	YBL019W	*APN2*	0.96	0.97	++	30	9.4.E
22	8	YBL021C	*HAP3*	1.00	1.00	+	30	9.4.F
22	9	YBL045C	*COR1*	0.99	0.96	+	30	9.5.A
22	10	YPL158C		0.93	0.94	+++	30	9.6.A
22	11	YGL025C	*PGD1*	1.00	0.96	+	30	9.6.B
22	12	YGL057C		0.77	0.77	+	30	9.6.C
22	13	YGL058W	*RAD6*	0.97	0.96	+	30	9.6.D
22	14	YGL071W	*RCS1*	0.98	0.99	+	30	9.6.G
22	15		WT	0.97	0.98	+	30	WT
23	1	YGL084C	*GUP1*	0.77	0.83	++	30	10.1.A
23	2	YNL237W	*YTP1*	0.78	0.81	++	30	10.1.D
23	3	YKL208W	*CBT1*	0.72	0.23	+	30	10.2.F
23	4	YKL212W	*SAC1,*	0.78	0.82	+++	30	10.2.G
23	5	YKR001C	*VPS1*	0.83	0.82	++	30	10.2.H
23	6	YKR021W		0.83	0.84	++	30	10.3.B
23	7	YKR049C		0.82	0.83	++	30	10.3.D
23	8	YDR269C		0.75	0.77	+++	30	10.3.H
23	9	YDR276C	*PMP3*	0.85	0.84	+	30	10.4.C
23	10	YDR293C	*SSD1*	0.83	0.78	+	30	10.4.D
23	11	YDR316W		0.83	0.83	+	30	10.4.G
23	12		WT	0.77	0.86	++	30	WT
24	1	YDR323C	*PEP7*	0.71	0.86	++	30	10.5.A
24	2	YDR332W		0.73	0.88	+	30	10.5.B
24	3	YDR335W	*MSN5*	0.73	0.89	+++	30	10.5.C
24	4	YDR336W		0.74	0.88	+++	30	10.5.D
24	5	YDR337W	*MRPS28*	0.70	0.85	+++	30	10.5.E
24	6	YIL036W	*CST6*	0.77	0.88	++	30	10.5.F
24	7	YIL093C	*RSM25*	0.74	0.87	++	30	10.6.A
24	8	YIL015C-A		0.77	0.88	++	30	10.6.B
24	9	YIL060W		0.80	0.93	+++	30	10.6.E
24	10	YIL070C	*MAM33*	0.73	0.87	++	30	10.6.F
24	11		WT	0.82	0.90	+++	30	WT

blot	lane	ORF name	Gene name	COX % splicing	COB % splicing	total RNA	growth temperature	library grid location
25	**1**	YHR039C-B	*VMA10*	**0.88**	**0.83**	++	30	11.1.D
25	**2**	YIR021W	*MRS1*	**0.72**	**0.84**	++	30	11.1.E
25	**3**	YIR024C	*GIF1*	**0.72**	**0.83**	++	30	11.1.F
25	**4**	YMR063W	*RIM9*	**0.33**	**0.42**	+++	30	11.2.B
25	**5**	YOL095C	*HMI1*	**0.78**	**0.89**	++	30	11.2.H
25	**6**	YOL096C	*COQ3*	**0.79**	**0.83**	++	30	11.3.A
25	**7**	YER077C		**0.74**	**0.84**	++	30	11.3.C
25	**8**	YHL025W	*SNF6*	**0.86**	**0.91**	++	30	11.3.F
25	**9**	YLR038C	*COX12*	**0.86**	**0.88**	++	30	11.4.A
25	**10**	YMR311C	*GLC8*	**0.84**	**0.92**	++	30	11.4.C
25	**11**	YML094W	*GIM5*	**0.80**	**0.91**	++	30	11.4.H
25	**12**	YML120C	*NDI1*	**0.73**	**0.91**	++	30	11.5.A
25	**13**	YMR123W	*PKR1*	**0.72**	**0.88**	+++	30	11.5.C
25	**14**	YPR036W	*VMA13*	**0.72**	**0.87**	+++	30	11.5.F
25	**15**	YPR066W	*UBA3*	**0.60**	**0.85**	++	30	11.6.A
25	**16**	YJL120W		**0.62**	**0.86**	+++	30	11.6.E
25	**17**	YJL121C	*RPE1*	**0.59**	**0.86**	+++	30	11.6.F
25	**18**	YJL117W	*PHO86*	**0.51**	**0.83**	++	30	11.6.G
25	**19**		WT	**0.63**	**0.87**	++	30	WT
26	**1**	YGR188C	*BUB1*	**0.71**	**0.82**	++	30	12.2.B
26	**2**	YHR008C	*SOD2*	**0.91**	**0.90**	+	30	12.2.F
26	**3**	YHR025W	*THR1*	**0.70**	**0.83**	++	30	12.2.H
26	**4**	YHR026W	*PPA1*	**0.67**	**0.78**	+++	30	12.3.A
26	**5**	YHR041C	*SRB2*	**0.78**	**0.84**	+++	30	12.3.B
26	**6**	YHR059W	*FYV4*	**0.74**	**0.83**	+++	30	12.3.C
26	**7**	YHR067W	*RMD12*	**0.77**	**0.80**	++	30	12.3.D
26	**8**	YHR194W		**0.75**	**0.82**	++	30	12.3.E
26	**9**	YJR079W		**0.64**	**0.71**	+	30	12.3.G
26	**10**	YKR087C		**0.75**	**0.83**	++	30	12.4.B
26	**11**	YLL018C-A	*COX19*	**0.54**	**0.64**	+	30	12.4.C
26	**12**	YLR423C	*APG17*	**0.69**	**0.75**	++	30	12.4.D
26	**13**	YLR425W	*TUS1*	**0.72**	**0.77**	+++	30	12.4.E
26	**14**	YLR447C	*VMA6*	**0.58**	**0.69**	+++	30	12.4.F
26	**15**	YML009C		**0.74**	**0.82**	++	30	12.4.G
26	**16**	YML081C-A	*ATP18*	**0.59**	**0.67**	++	30	12.4.H
26	**17**	YAL012W	*CYS3*	**0.61**	**0.68**	+	30	12.5.B
26	**18**	YJR118C		**0.57**	**0.68**	+	30	12.5.G
26	**19**		WT	**0.85**	**0.90**	+	30	WT

blot	lane	ORF name	Gene name	COX % splicing	COB % splicing	total RNA	growth temperature	library grid location
27	**1**	YJR121W	*ATP2*	**0.87**	**0.82**		30	12.6.A
27	**2**	YDR027C	*LUV1*	**0.80**	**0.78**		30	12.6.F
27	**3**		WT	**0.79**	**0.80**		30	WT
27	**4**	YKL016C	*ATP7*	**0.73**	**0.74**		30	5.6.g
27	**5**		MRPS2	**0.87**	**0.84**		30	MRPS2
27	**6**		MRP8	**0.68**	**0.77**		30	MRP8
27	**7**		MRPS9	**0.68**	**0.78**		30	MRPS9
28	**1**	YBR269C		**0.75**	**0.87**	+	30	13.1.A
28	**2**	YCR065W	*HCM1*	**0.84**	**0.86**	+++	30	13.1.E
28	**3**	YJL022W		**0.77**	**0.85**	++	30	13.2.A
28	**4**	YJL029C	*VPS53*	**0.68**	**0.85**	++	30	13.2.C
28	**5**	YJR033C	*RAV1*	**0.75**	**0.82**	+++	30	13.2.D
28	**6**	YJR048W	*CYC1*	**0.80**	**0.84**	++	30	13.2.G
28	**7**	YDL106C	*PHO2*	**0.71**	**0.87**	+	30	13.3.C
28	**8**	YDL117W	*CYK3,*	**0.77**	**0.85**	+++	30	13.3.E
28	**9**	YDL118W		**0.74**	**0.81**	+++	30	13.3.F
28	**10**	YDL119C		**0.73**	**0.82**	++	30	13.3.G
28	**11**	YOR333C		**0.73**	**0.84**	++	30	13.4.E
28	**12**	YNR001C	*CIT1*	**0.72**	**0.84**	++	30	13.4.H
28	**13**	YNR018W		**0.69**	**0.85**	++	30	13.5.B
28	**14**	YNR041C	*COQ2*	**0.54**	**0.79**	++	30	13.6.A
28	**15**	YNR042W		**0.82**	**0.85**	+++	30	13.6.B
28	**16**	YNR045W	*PET494*	**0.44**	**0.80**	++	30	13.6.C
28	**17**	YBL099W	*ATP1*	**0.74**	**0.80**	++	30	13.6.F
28	**18**	YBL100C		**0.71**	**0.59**	++	30	13.6.G
28	**19**		WT	**0.83**	**0.88**	+	30	WT

blot	lane	ORF name	Gene name	COX % splicing	COB % splicing	total RNA	growth temperature	library grid location
29	**1**	YBR003W	*COQ1*	**0.74**	**0.81**	+	30	14.1.A
29	**2**	YBR026C	*ETR1*	**0.63**	**0.74**	++	30	14.1.B
29	**3**	YNL117W	*MLS1*	**0.59**	**0.78**	++	30	14.2.B
29	**4**	YNL064C	*YDJ1*	**0.68**	**0.82**	+++	30	14.2.D
29	**5**	YIL125W	*KGD1*	**0.75**	**0.89**	+	30	14.3.A
29	**6**	YIL154C	*(IMP2)*	**0.69**	**0.92**	+	30	14.3.C
29	**7**	YIL157C		**0.79**	**0.94**	+	30	14.3.E
29	**8**	YIR005W	*IST3*	**0.91**	**0.88**	+++	30	14.3.F
29	**9**	YIR009W	*MSL1*	**0.78**	**0.88**	+	30	14.3.G
29	**10**	YNL052W	*COX5A*	**0.67**	**0.89**	++	30	14.4.A
29	**11**	YNL071W	*LAT1*	**0.80**	**0.85**	++	30	14.4.D
29	**12**	YNL097C	*PHO23*	**0.82**	**0.90**	+++	30	14.4.H
29	**13**	YDR202C	*RAV2,*	**0.74**	**0.85**	++++	30	14.5.E
29	**14**	YJL175W		**0.81**	**0.88**	+++	30	14.5.H
29	**15**	YKR036C	*CAF4*	**0.74**	**0.88**	++	30	14.6.B
29	**16**	YBR289W	*SNF5*	**0.79**	**0.92**	++++	30	14.6.C
29	**17**	YCR047C	*BUD23*	**0.67**	**0.93**	+++	30	14.6.F
29	**18**	YCR053W	*THR4*	**0.77**	**0.93**	+++	30	14.6.G
29	**19**		WT	**0.84**	**0.92**	++	30	WT
30	**1**	YCR084C	*TUP1*	**0.72**	**0.84**	++	30	15.1.A
30	**2**	YNL138W	*SRV2,*	**0.82**	**0.86**	+	30	15.2.C
30	**3**	YNL315C	*ATP11*	**0.00**	**0.00**	++	30	15.2.D
30	**4**	YJL094C	*KHA1*	**0.81**	**0.88**	++++	30	15.3.B
30	**5**	YJL101C	*GSH1*	**0.80**	**0.83**	+++	30	15.3.C
30	**6**	YLR439W	*MRPL4*	**0.00**	**0.00**	+	30	15.3.D
30	**7**	YBR084W	*MIS1*	**0.81**	**0.85**	++	30	15.3.G
30	**8**	YBR127C	*VMA2*	**0.94**	**0.93**	++	30	15.4.C
30	**9**	YDL006W	*PTC1*	**0.67**	**0.85**	+++	30	15.4.E
30	**10**	YDL068W		**0.90**	**0.95**	++	30	15.5.F
30	**11**	YDR448W	*ADA2*	**0.78**	**0.82**	++	30	15.5.H
30	**12**	YDR455C		**0.80**	**0.82**	++	30	15.6.A
30	**13**	YDR456W	*NHX1*	**0.90**	**0.85**	++	30	15.6.B
30	**14**	YDR462W	*MRPL28*	**0.89**	**0.91**	++	30	15.6.C
30	**15**	YDR484W	*SAC2*	**0.77**	**0.82**	++	30	15.6.E
30	**16**	YDR495C	*VPS3*	**0.79**	**0.90**	+++	30	15.6.G
30	**17**		WT	**0.69**	**0.85**	++	30	WT

blot	lane	ORF name	Gene name	COX % splicing	COB % splicing	total RNA	growth temperature	library grid location
31	1	YDL056W	*MBP1*	0.87	0.91	++		15.5.D
31	2	YDR494W		0.80	0.82	++		15.6.F
31	3	YDR509W		0.83	0.86	++		16.1.A
31	4	YDR511W	*ACN9*	0.84	0.86	++		16.1.B
31	5	YDR512C	*EMI1*	0.72	0.75	++		16.1.C
31	6	YDR519W	*FKB2*	0.84	0.84	++		16.1.F
31	7	YDR522C	*SPS2*	0.86	0.83	++		16.1.G
31	8	YDR529C	*QCR7*	0.75	0.91	++		16.1.H
31	9	YDR533C		0.89	0.83	++		16.2.A
31	10	YGL107C	*RMD9*	0.00	0.00	+		16.2.B
31	11	YGL115W	*SNF4*	0.69	0.87	+++		16.2.C
31	12	YGL135W	*RPL1B*	0.75	0.80	+		16.2.E
31	13	YGL167C	*PMR1*	0.79	0.81	+++		16.2.G
31	14	YER141W	*COX15*	0.86	0.81	++		16.3.C
31	15	YER145C	*FTR1*	0.75	0.75	+++		16.3.D
31	16	YER178W	*PDA1*	0.65	0.67	++		16.3.H
31	17	YMR054W	*STV1*	0.70	0.76	++		16.4.B
31	18	YJR034W	*PET191*	0.67	0.77	++		16.4.E
31	19		WT	0.76	0.79	++		wt
32	1	YDL047W	*SIT4,*	0.78	0.82	++		16.5.A
32	2	YDL067C	*COX9*	0.50	0.84	+		16.5.B
32	3	YDR477W	*SNF1*	0.74	0.87	+++		16.5.C
32	4	YER028C		0.78	0.82	++		16.5.E
32	5	YOL148C	*SPT20,*	0.65	0.81	++		16.5.G
32	6	YLR338W	*KRE21*	0.86	0.84	++		16.6.B
32	7	YLR240W	*VPS34*	0.71	0.80	++		16.6.D
32	8	YGL206C	*CHC1*	0.81	0.91	+/-		16.6.E
32	9	YGL223C	*COG1*	0.69	0.81	+++		16.6.F
32	10	YLR396C	*VPS33*	0.65	0.80	+++		16.6.G
32	11			0.74	0.83	++		wt
33	1	YPL045W	*VPS16*	0.75	0.81	+++		17.1.B
33	2	YPL024W	*NCE4*	0.75	0.80	+++		17.1.D
33	3	YBL033C	*RIB1*	0.78	0.79	+++		17.1.H
33	4	YGL020C		0.78	0.78	+++		17.2.A
33	5	YGL038C	*OCH1*	0.92	0.94	+++		17.2.B
33	6	YIL018W	*RPL2B,*	0.82	0.82	+++		17.2.H
33	7	YFL018C	*LPD1*	0.72	0.79	+++		17.3.B
33	8	YGR262C	*BUD32*	0.84	0.77	+++		17.3.C
33	9	YML110C	*COQ5*	0.86	0.96	+++		17.3.F
33	10	YML112W	*CTK3*	0.76	0.82	+++		17.3.G
33	11	YML129C	*COX14*	0.69	0.80	+++		17.3.H
33	12	YAL016W	*TPD3*	0.70	0.75	+++		17.4.D
33	13			0.76	0.84	+++		WT

blot	lane	ORF name	Gene name	COX % splicing	COB % splicing	total RNA	growth temperature	library grid location
34	**1**	YJR090C	*GRR1*	**0.77**	**0.87**	+++		17.4.F
34	**2**	YCR044C	*PER1*	**0.70**	**0.85**	+++		17.4.H
34	**3**	YJL003W		**0.72**	**0.85**	+++		17.5.A
34	**4**	YJL004C	*SYS1*	**0.70**	**0.80**	+++		17.5.B
34	**5**	YJR040W	*GEF1*	**0.79**	**0.80**	+++		17.5.C
34	**6**	YDL185W	*TFP1*	**0.80**	**0.79**	+++		17.5.D
34	**7**	YIL098C	*FMC1*	**0.80**	**0.79**	+++		17.5.h
34	**8**	YPL148C	*PPT2*	**0.72**	**0.84**	+++		17.6.A
34	**9**	YPL268W	*PLC1*	**0.73**	**0.84**	+++		17.6.B
34	**10**	YJL075C		**0.77**	**0.86**	+++		17.6.F
34	**11**	YBR035C	*PDX3*	**0.79**	**0.82**	+++		17.6.G
34	**12**	YEL029C	*BUD16*	**0.82**	**0.82**	+++		17.6.H
34	**13**			**0.88**	**0.85**	+++		WT
35	**1**	YOL143C	*RIB4,*	**0.64**	**0.79**	++		17.4.C
35	**2**	YLL027W	*ISA1*	**0.64**	**0.82**	+++		18.1.A
35	**3**	YGR167W	*CLC1*	**0.68**	**0.76**	++		18.1.E
35	**4**	YCL007C	*CWH36*	**0.68**	**0.78**	++++		18.1.F
35	**5**	YLR148W	*PEP3*	**0.65**	**0.78**	+++		18.1.G
35	**6**	YER014C-A	*BUD25*	**0.68**	**0.72**	+++		18.2.A
35	**7**	YPR133W-A	*TOM5*	**0.78**	**0.77**	+++		18.2.C
35	**8**	YBR112C	*SSN6*	**0.73**	**0.72**	+++		18.2.D
35	**9**	YGR252W	*GCN5*	**0.84**	**0.75**	++		18.2.G
35	**10**	YGR255C	*COQ6*	**0.76**	**0.70**	++		18.2.H
35	**11**	YOR147W	*MDM32*	**0.76**	**0.77**	++		18.3.D
35	**12**	YMR263W	*SAP30*	**0.26**	**0.26**	++++		18.3.E
35	**13**			**0.86**	**0.84**	++++		WT

blot	lane	ORF name	Gene name	COX % splicing	COB % splicing	total RNA	growth temperature	library grid location
36	**1**	YOR022C		**0.64**	**0.68**	+++		2.4.d
36	**2**	YBR187W		**0.64**	**0.65**	++++		3.5.e
36	**3**	YDR136C		**0.62**	**0.69**	++++		4.2.b
36	**4**	YGR105W	*VMA21*	**0.56**	**0.57**	++++		6.5.f
36	**5**	YGR112W	*SHY1*	**0.55**	**0.60**	+++		6.5.g
36	**6**	YJL209W	*CBP1*	**0.09**	**0.00**	++		7.2.d
36	**7**	YJL155C	*FBP26*	**0.60**	**0.70**	+++		7.3.d
36	**8**	YLR368W	*MDM30*	**0.52**	**0.00**	+		7.4.c
36	**9**	YLR232W		**0.63**	**0.79**	+++		7.5.a
36	**10**	YLR235C		**0.65**	**0.80**	++		7.5.C
36	**11**	YDR162C	*NBP2*	**0.61**	**0.78**	+++		8.1.D
36	**12**	YDR176W	*NGG1*	**0.52**	**0.88**	+		8.1.F
36	**13**	YDR178W	*SDH4*	**0.66**	**0.66**	+++		8.1.G
36	**14**	YPL065W	*VPS28*	**0.75**	**0.71**	+++		8.5.G
36	**15**	YFR034C	*PHO4*	**0.57**	**0.62**	+++		9.3.B
36	**16**	YBL089W		**0.62**	**0.53**	+++		9.5.G
36	**17**			**0.67**	**0.72**	++		wt
37	**1**	YHR009C		**0.71**	**0.62**	++++		12.2.G
37	**2**	YBR044C	*TCM62*	**0.65**	**0.65**	++++	loosing mtDNA	14.1.D
37	**3**	YFR019W	*FAB1*	**0.59**	**0.70**	++++	loosing mtDNA	15.3.A
37	**4**	YOR331C		**0.75**	**0.74**	+++++		17.5.E
37	**5**	YMR307W	*GAS1*	**0.70**	**0.68**	++++		2.3.e
37	**6**	YNL169C	*PSD1*	**0.66**	**0.67**	++++		10.2.D
37	**7**	YNL037C	*IDH1*	**0.53**	**0.57**	++++		13.4.G
37	**8**	YIL155C	*GUT2*	**0.68**	**0.71**	+++		14.3.D
37	**9**	YNR052C	*POP2*	**0.72**	**0.72**	+++	loosing mtDNA	15.3.E
37	**10**	YPR024W	*YME1*	**0.53**	**0.62**	+++		15.3.F
37	**11**	YGL218W	*SRF1*	**0.60**	**0.89**	+++	loosing mtDNA	17.4.E
37	**12**	YDR512C	*EMI1*	**0.36**	**0.46**	+++		18.2.F
37	**13**	YGR061C	*ADE6*	**0.62**	**0.71**	+++		19.2.C
37	**14**	YJL140W	*RPB4*	**0.70**	**0.90**	++	loosing mtDNA	19.2.D
37	**15**	YDR470C	*UGO1*	**0.73**	**0.91**	++	loosing mtDNA	19.4.F
37	**16**			**0.67**	**0.74**	+++		wt

blot	lane	ORF name	Gene name	COX % splicing	COB % splicing	total RNA	growth temperature	library grid location
cand 1	**1**	YKL106W	*AAT1*	**0.53**	**0.18**	+++	30	6.2.C
cand 1	**2**	YDR197W	*CBS2*	**0.19**	**0.05**	++	30	8.2.D
cand 1	**3**	YMR150C	*IMP1*	**0.27**	**0.92**	++	30	1.6.B
cand 1	**4**	YOR350C	*MNE1*	**0.41**	**0.85**	++	30	2.6.C
cand 1	**5**	YHL038C	*CBP2*	**0.24**	**0.24**	++	30	5.1.A
cand 1	**6**	YOR017W	*PET127*	**0.27**	**0.03**	+++	30	2.4.C
cand 1	**7**	YDL069C	*CBS1*	**0.15**	**0.60**	++	30	15.5.G
cand 1	**8**	YDL032W	*YDL032W*	**0.40**	**0.89**	++	30	15.4.G
cand 1	**9**	YDL033C	*MTO2*	**0.66**	**0.86**	+++	30	15.4.H
cand 1	**10**	YMR257C	*PET111*	**0.49**	**0.87**	++	30	2.2.C
cand 1	**11**	YPL174C	*NIP100*	**0.95**	**0.88**	+++	30	3.3.G
cand 1	**12**	YLR203C	*MSS51*	**0.51**	**0.74**	++	30	5.6.A
cand 1	**13**	YOR334W	*MRS2*	**0.60**	**0.65**	+	37	2.6.A
cand 1	**14**	YGR222W	*PET54*	**0.03**	**0.79**	++	30	17.1.F
cand 1	**15**	YPR134W	*MSS18*	**0.51**	**0.80**	++	30	9.1.F
cand 1	**16**	YJL133W	*MRS3*	**0.53**	**0.87**	+++	30	7.3.F
cand 1	**17**		WT	**0.66**	**0.91**	++	30	WT
cand 2	**1**	YER153C	*PET122*	**0.71**	**0.71**	+++	30	16.3.E
cand 2	**2**	YER061C	*CEM1*	**0.22**	**0.44**	+++	30	4.4.F
cand 2	**3**	YLR393W	*ATP10*	**0.12**	**0.17**	+++	30	7.4.F
cand 2	**4**	YJL180C	*ATP12*	**0.73**	**0.41**	+++	30	7.2.H
cand 2	**5**		ATP12 omega-	**0.37**	**0.27**	++	30	7.2.H
cand 2	**6**	YPL098C	YPL098C	**0.43**	**0.69**	+++	30	3.4.H
cand 2	**7**	YHR179W	*OYE2*	**0.08**	**0.53**	+++	20	5.3.F
cand 2	**8**	YML009c	*MRPL39*	**0.26**	**0.54**	+++	20	1.5.A
cand 2	**9**	YPL172C	*COX10*	**0.47**	**0.77**	++	30	3.4.A
cand 2	**10**	YML061C	*PIF1*	**0.48**	**0.27**	+	37	1.4.F
cand 2	**11**	YHL005C	*(MRP4)*	**0.48**	**0.34**	+++	20	5.1.F
cand 2	**12**	YNL136W	*EAF7*	**0.65**	**0.69**	+	37	14.1.G
cand 2	**13**	YER093C-A	YER093C-A	**0.81**	**0.77**	+++	30	11.3.E
cand 2	**14**	YJR077C	*MIR1*	**0.91**	**0.75**	+	30	12.5.D
cand 2	**15**	YML090W	*RPM2*	**0.21**	**0.30**	++	30	11.4.G
cand 2	**16**		WT	**0.62**	**0.77**	++	30	WT

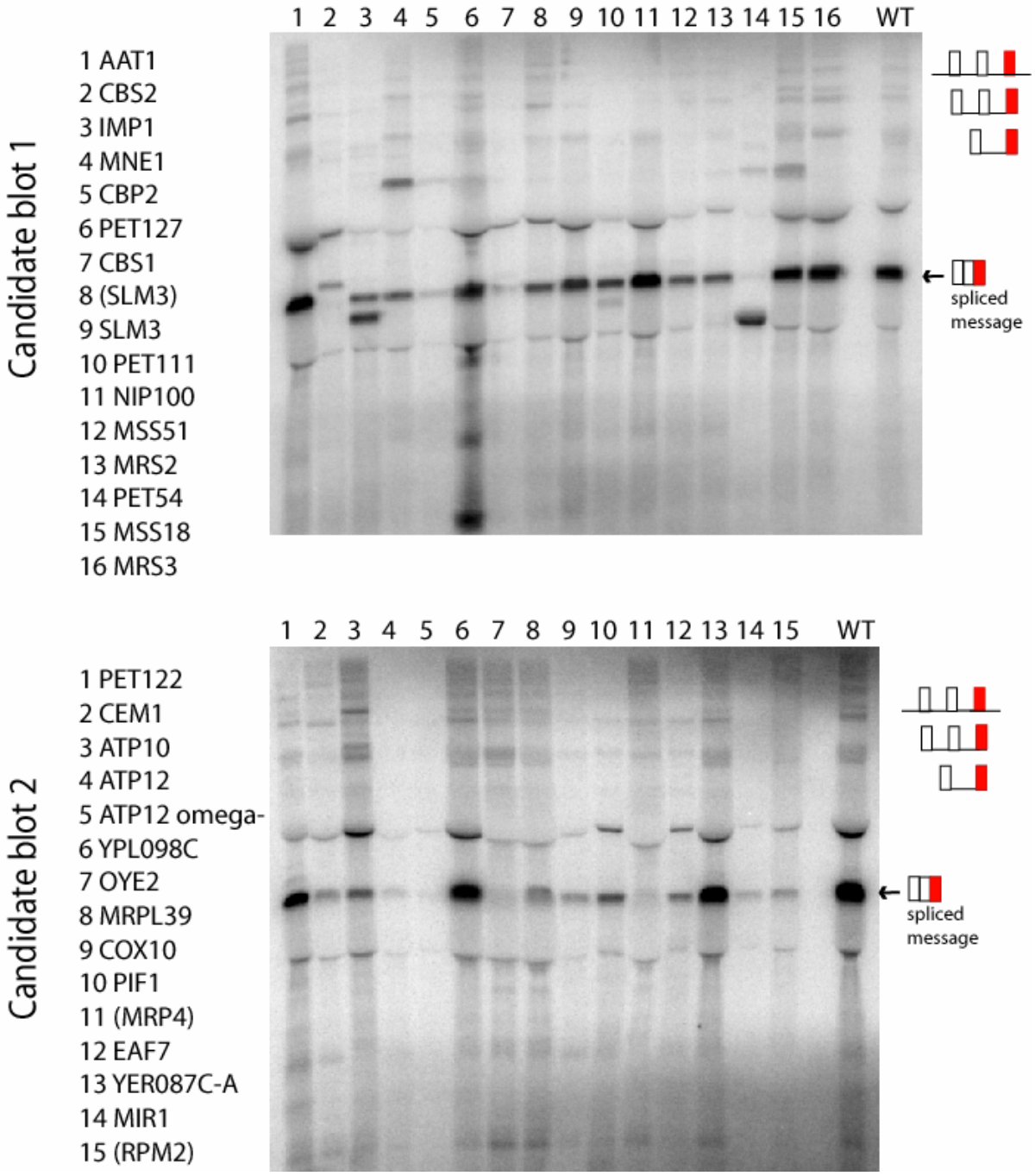
APPENDIX E
Intron probed RNA blots of candidate splicing factor strains

The individual intron probed blots are shown. Each blot contains a diagram of *COXI* or *COB* respectively with the probed intron indicated in ***bold***. Diagrams of the predicted RNA species found in each band is shown to the right of each blot with the probed sequence in red.

COX1 exon probed blots are shown first, followed the same blots probed for the seven introns found in *COXI*. The *COB* exon probed blots followed by the same blots probed for the six *COB* introns are shown next. Finally the same blots probed for omega intron is shown.

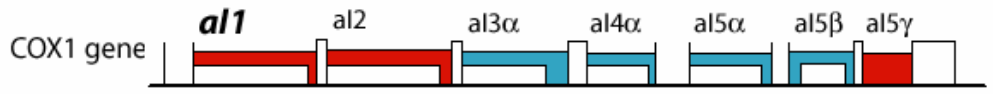
Candidate splicing factor Blots

COX1 probe



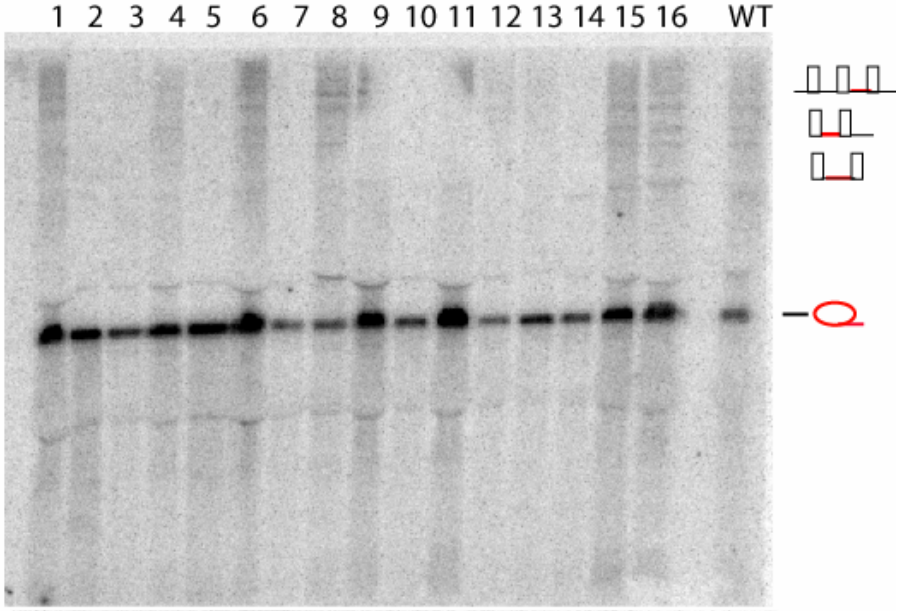
Candidate splicing factor Blots

al1 probe



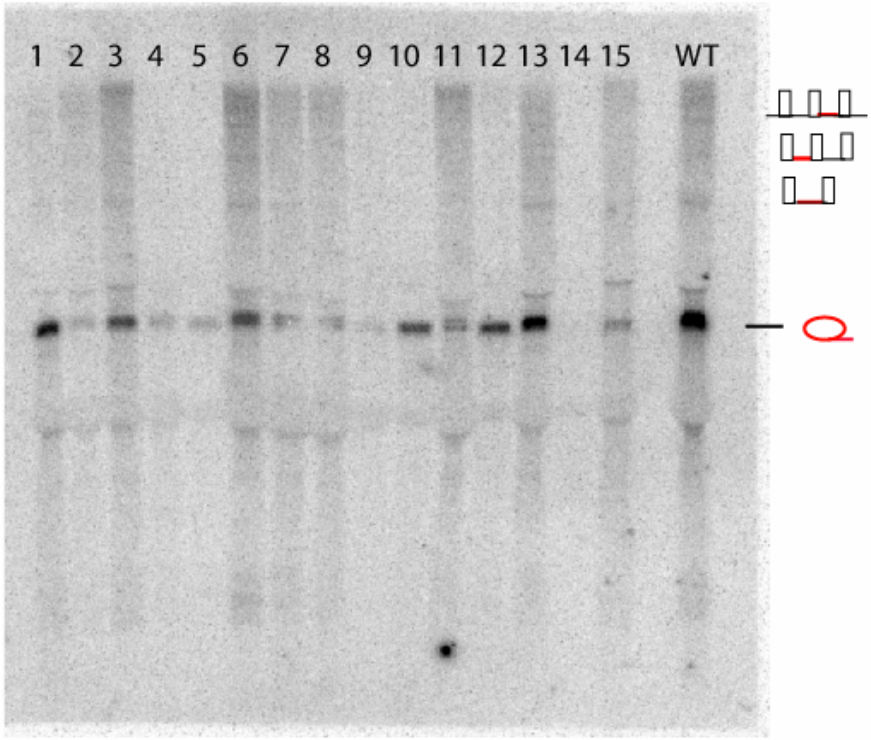
Candidate blot 1

1 AAT1
2 CBS2
3 IMP1
4 MNE1
5 CBP2
6 PET127
7 CBS1
8 (SLM3)
9 SLM3
10 PET111
11 NIP100
12 MSS51
13 MRS2
14 PET54
15 MSS18
16 MRS3



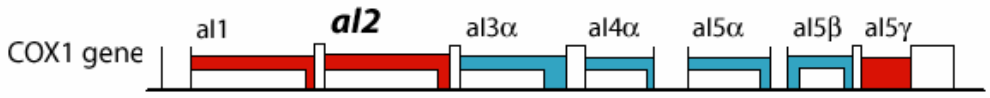
Candidate blot 2

1 PET122
2 CEM1
3 ATP10
4 ATP12
5 ATP12 omega-
6 YPL098C
7 OYE2
8 MRPL39
9 COX10
10 PIF1
11 (MRP4)
12 EAF7
13 YER087C-A
14 MIR1
15 (RPM2)



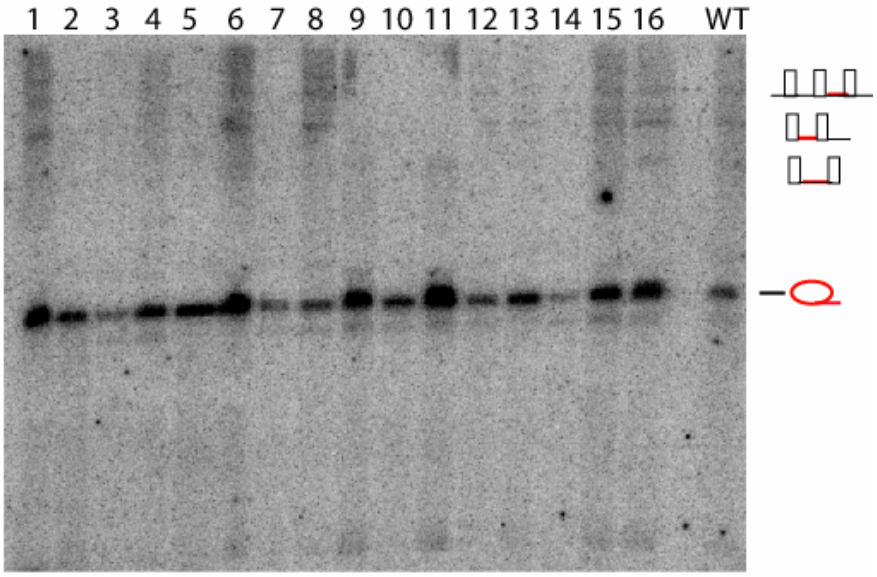
Candidate splicing factor Blots

aI2 probe



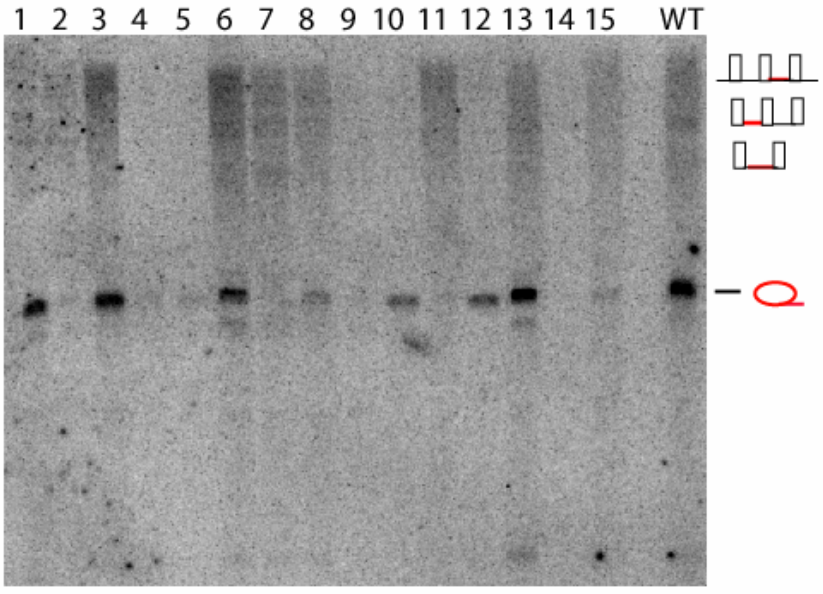
Candidate blot 1

1 AAT1
2 CBS2
3 IMP1
4 MNE1
5 CBP2
6 PET127
7 CBS1
8 (SLM3)
9 SLM3
10 PET111
11 NIP100
12 MSS51
13 MRS2
14 PET54
15 MSS18
16 MRS3



Candidate blot 2

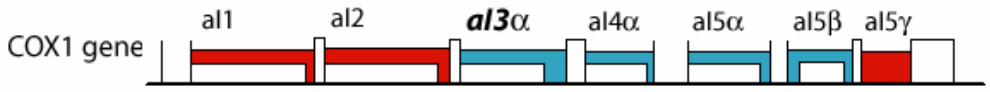
1 PET122
2 CEM1
3 ATP10
4 ATP12
5 ATP12 omega-
6 YPL098C
7 OYE2
8 MRPL39
9 COX10
10 PIF1
11 (MRP4)
12 EAF7
13 YER087C-A
14 MIR1
15 (RPM2)



Candidate splicing factor Blots

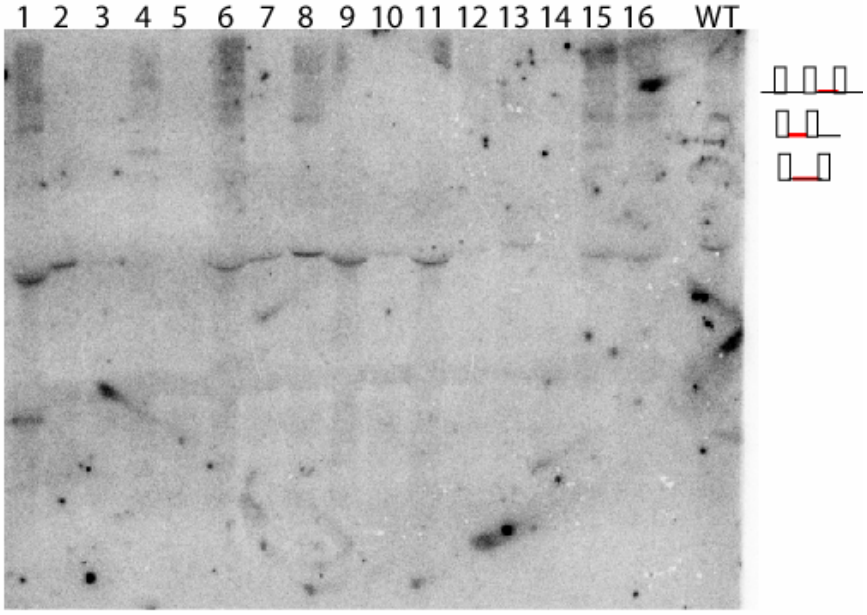
al3α probe

al3α free intron is not stable



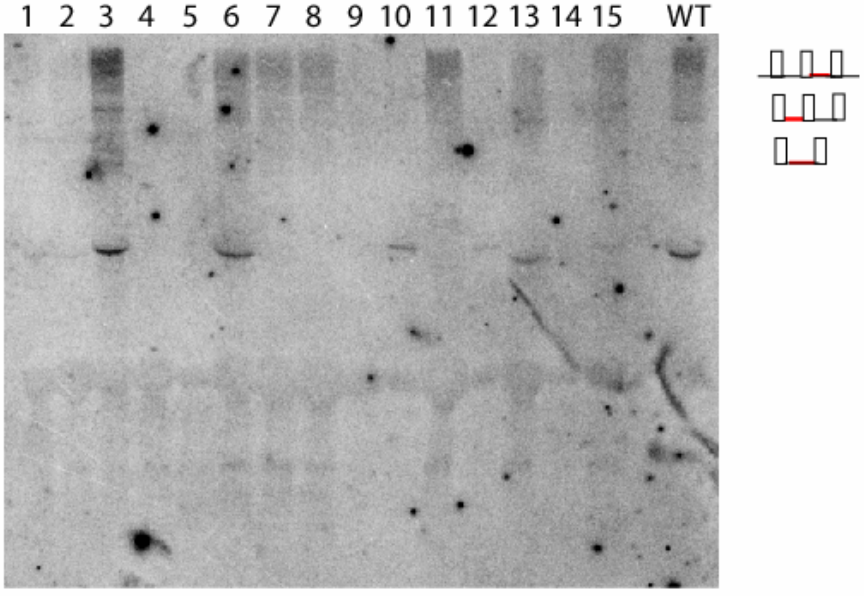
Candidate blot 1

1 AAT1
2 CBS2
3 IMP1
4 MNE1
5 CBP2
6 PET127
7 CBS1
8 (SLM3)
9 SLM3
10 PET111
11 NIP100
12 MSS51
13 MRS2
14 PET54
15 MSS18
16 MRS3



Candidate blot 2

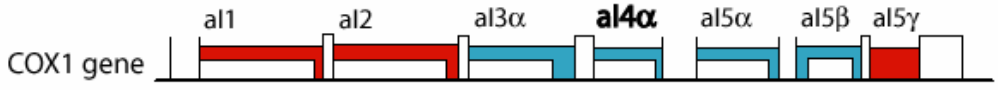
1 PET122
2 CEM1
3 ATP10
4 ATP12
5 ATP12 omega-
6 YPL098C
7 OYE2
8 MRPL39
9 COX10
10 PIF1
11 (MRP4)
12 EAF7
13 YER087C-A
14 MIR1
15 (RPM2)



Candidate splicing factor Blots

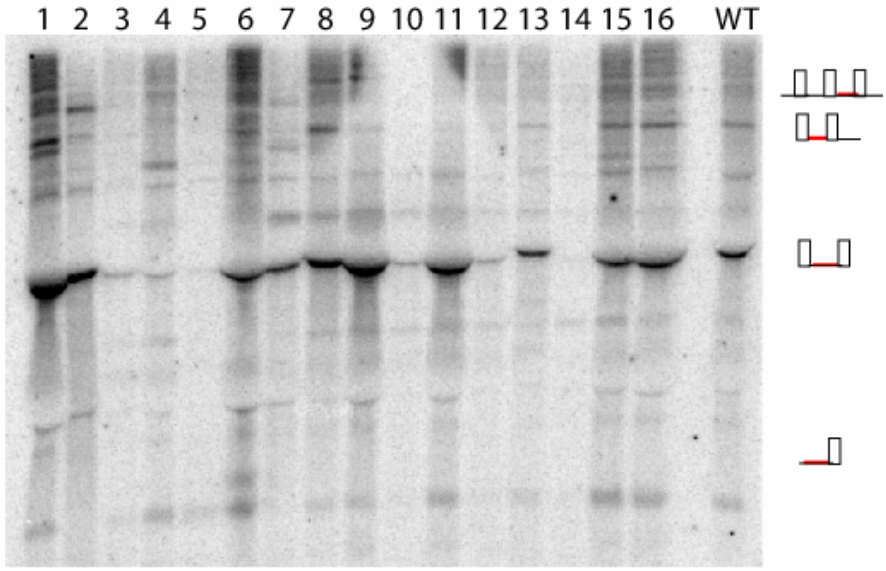
al4α probe

al4α free intron is not stable.



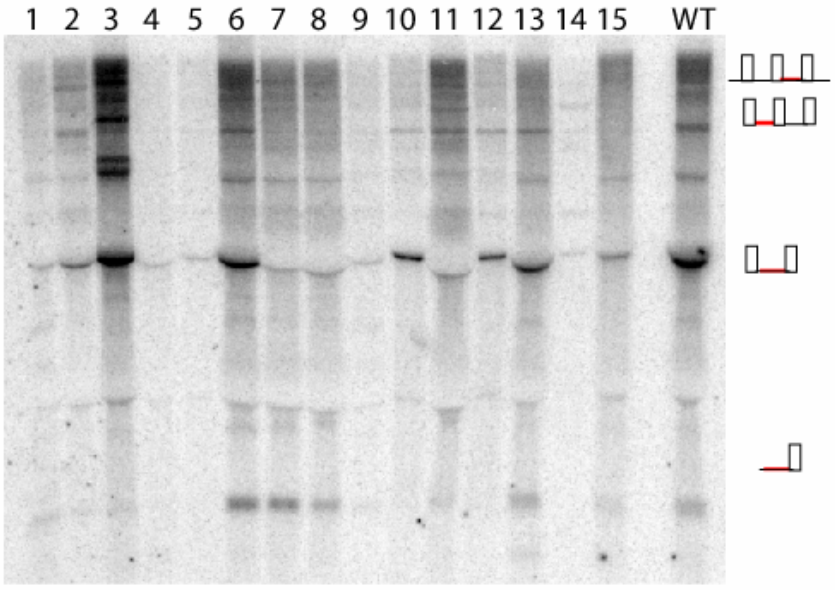
Candidate blot 1

1 AAT1
2 CBS2
3 IMP1
4 MNE1
5 CBP2
6 PET127
7 CBS1
8 (SLM3)
9 SLM3
10 PET111
11 NIP100
12 MSS51
13 MRS2
14 PET54
15 MSS18
16 MRS3



Candidate blot 2

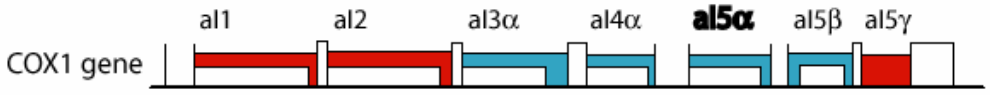
1 PET122
2 CEM1
3 ATP10
4 ATP12
5 ATP12 omega-
6 YPL098C
7 OYE2
8 MRPL39
9 COX10
10 PIF1
11 (MRP4)
12 EAF7
13 YER087C-A
14 MIR1
15 (RPM2)



Candidate splicing factor Blots

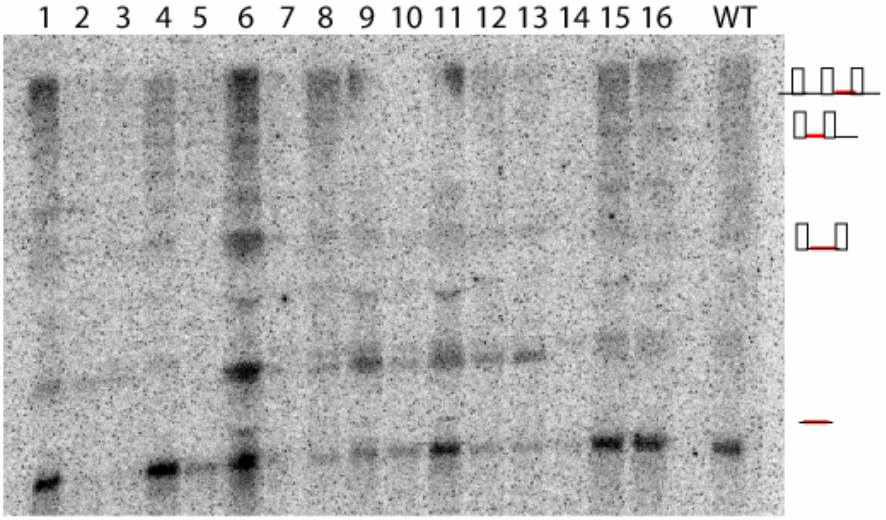
aI5α probe

aI4α free intron is not stable.



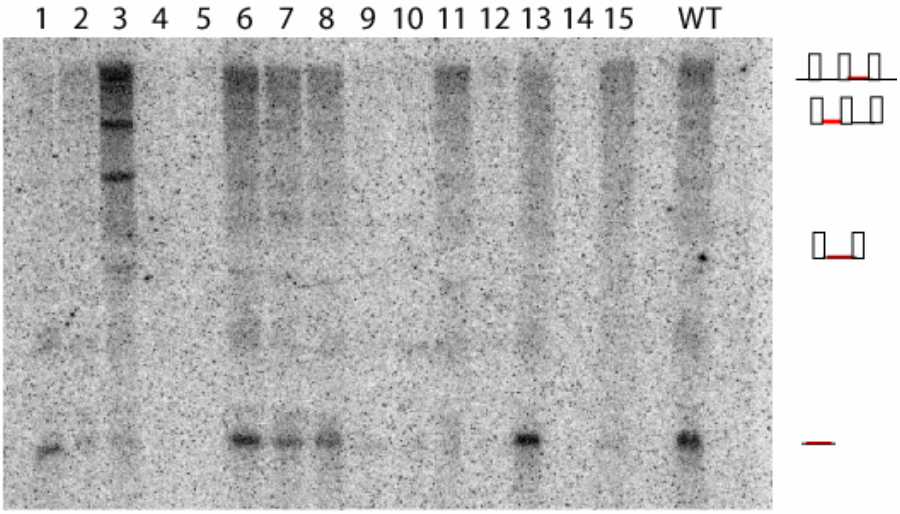
Candidate blot 1

1 AAT1
2 CBS2
3 IMP1
4 MNE1
5 CBP2
6 PET127
7 CBS1
8 (SLM3)
9 SLM3
10 PET111
11 NIP100
12 MSS51
13 MRS2
14 PET54
15 MSS18
16 MRS3



Candidate blot 2

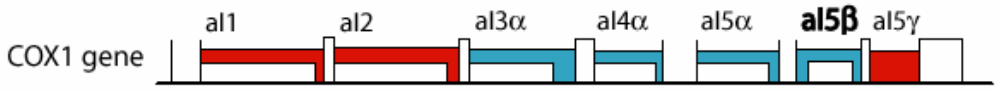
1 PET122
2 CEM1
3 ATP10
4 ATP12
5 ATP12 omega-
6 YPL098C
7 OYE2
8 MRPL39
9 COX10
10 PIF1
11 (MRP4)
12 EAF7
13 YER087C-A
14 MIR1
15 (RPM2)



Candidate splicing factor Blots

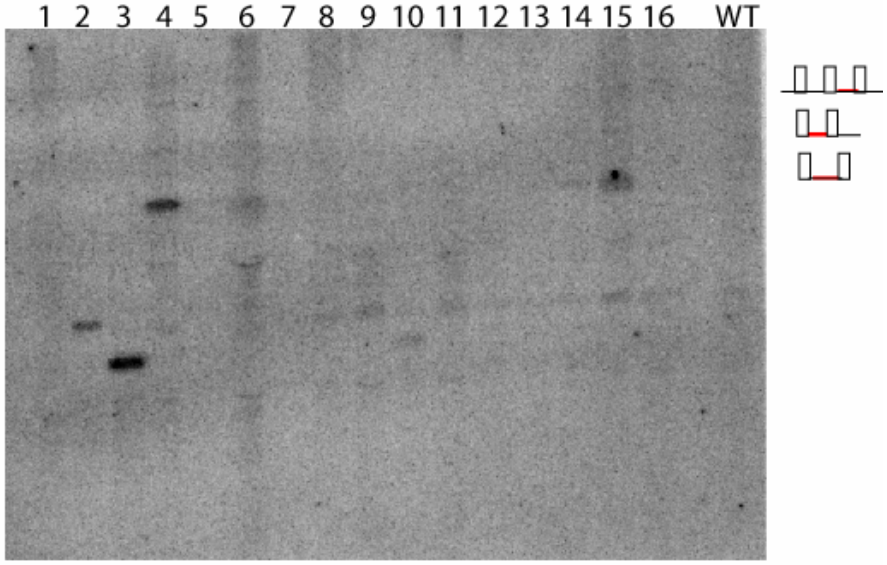
aI5β probe

aI5β free intron is not stable



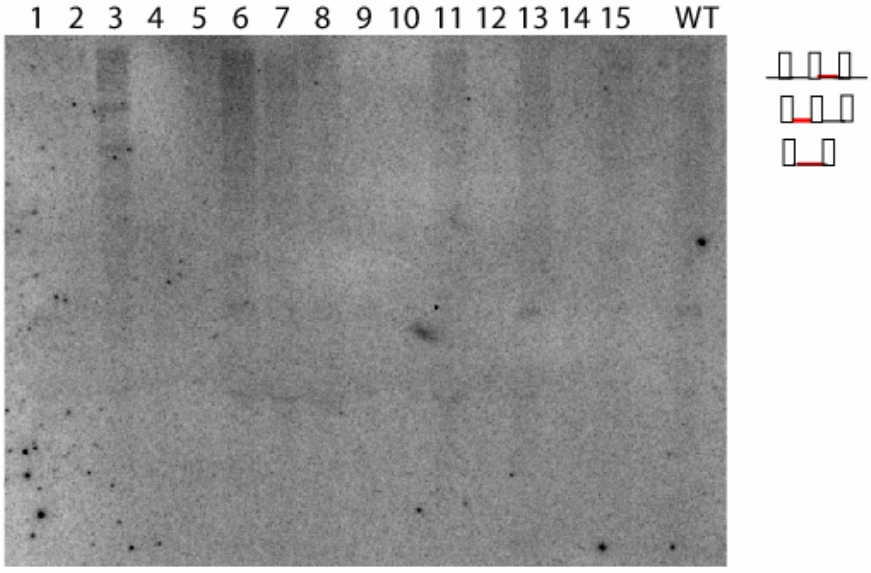
Candidate blot 1

1 AAT1
2 CBS2
3 IMP1
4 MNE1
5 CBP2
6 PET127
7 CBS1
8 (SLM3)
9 SLM3
10 PET111
11 NIP100
12 MSS51
13 MRS2
14 PET54
15 MSS18
16 MRS3



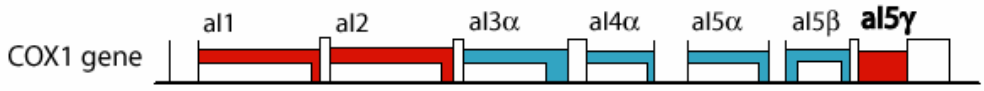
Candidate blot 2

1 PET122
2 CEM1
3 ATP10
4 ATP12
5 ATP12 omega-
6 YPL098C
7 OYE2
8 MRPL39
9 COX10
10 PIF1
11 (MRP4)
12 EAF7
13 YER087C-A
14 MIR1
15 (RPM2)



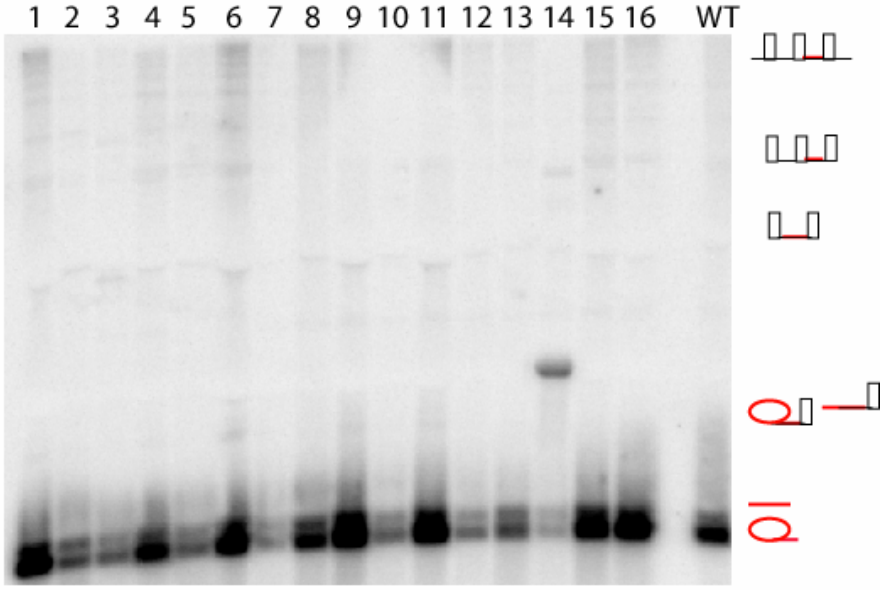
Candidate splicing factor Blots

aI5γ probe



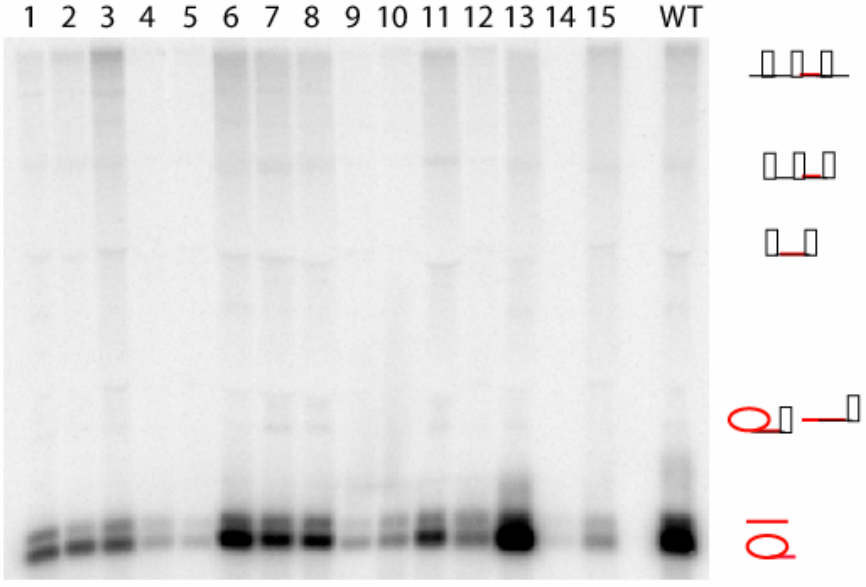
Candidate blot 1

1 AAT1
2 CBS2
3 IMP1
4 MNE1
5 CBP2
6 PET127
7 CBS1
8 (SLM3)
9 SLM3
10 PET111
11 NIP100
12 MSS51
13 MRS2
14 PET54
15 MSS18
16 MRS3



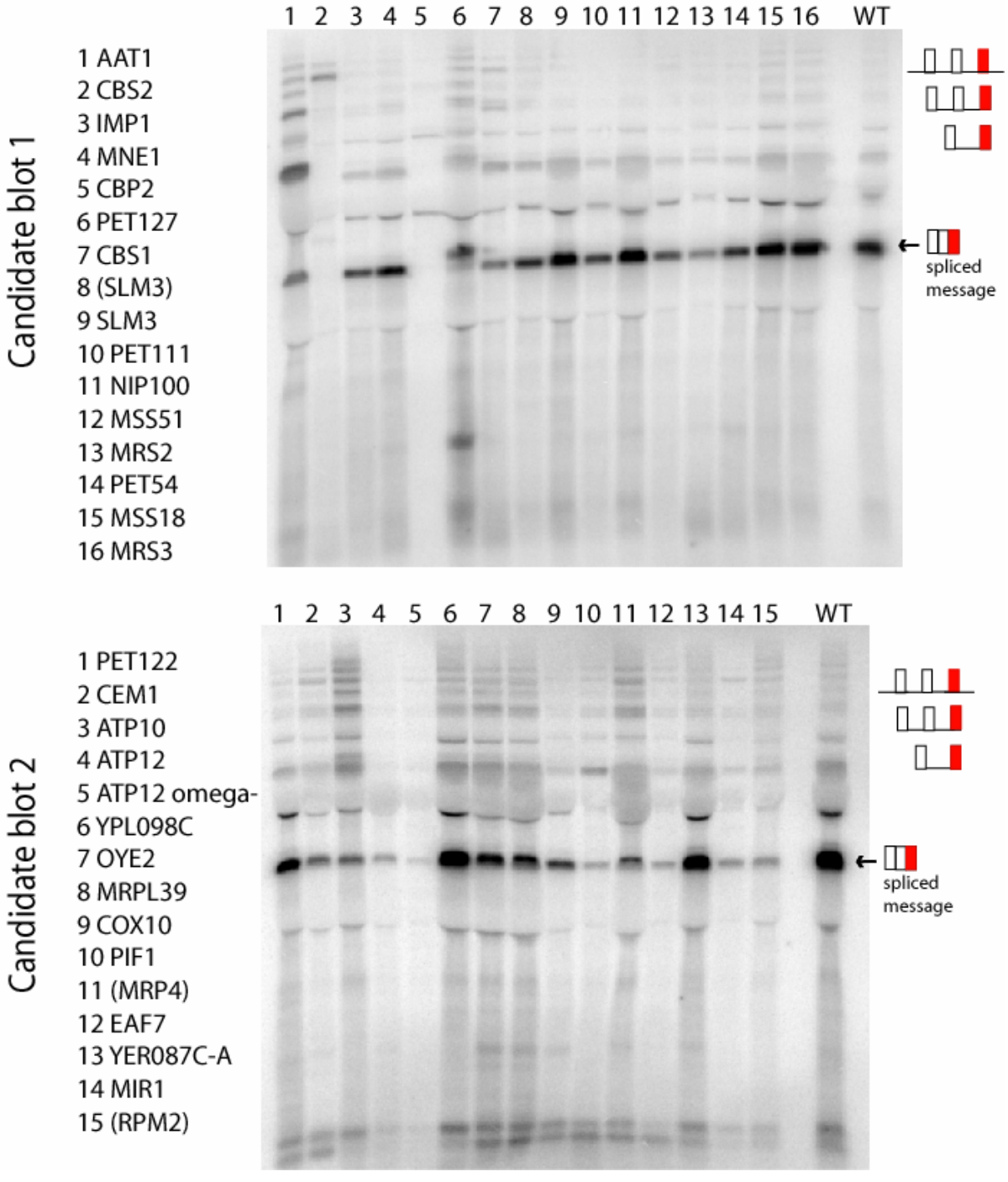
Candidate blot 2

1 PET122
2 CEM1
3 ATP10
4 ATP12
5 ATP12 omega-
6 YPL098C
7 OYE2
8 MRPL39
9 COX10
10 PIF1
11 (MRP4)
12 EAF7
13 YER087C-A
14 MIR1
15 (RPM2)



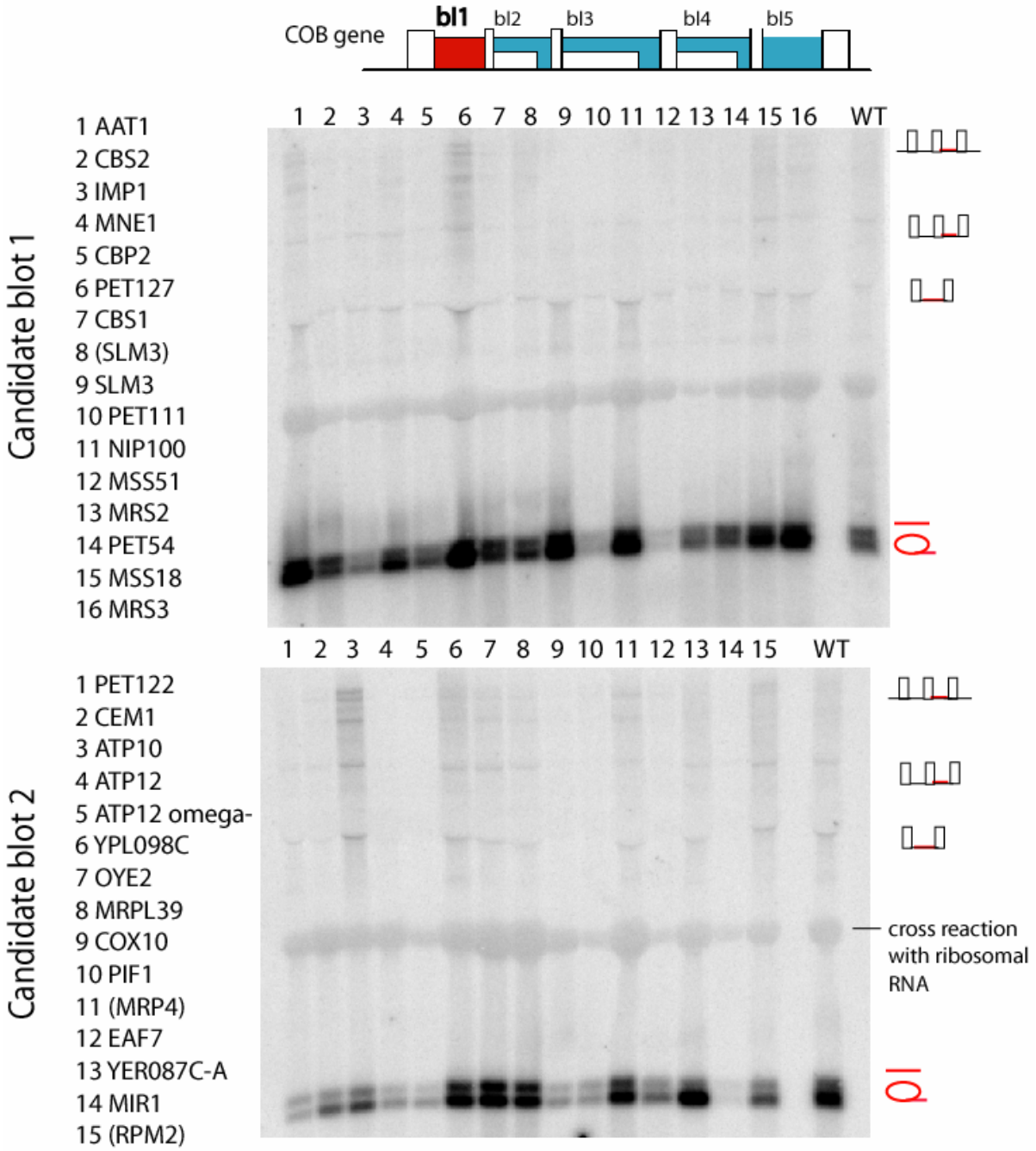
Candidate splicing factor Blots

COB probe

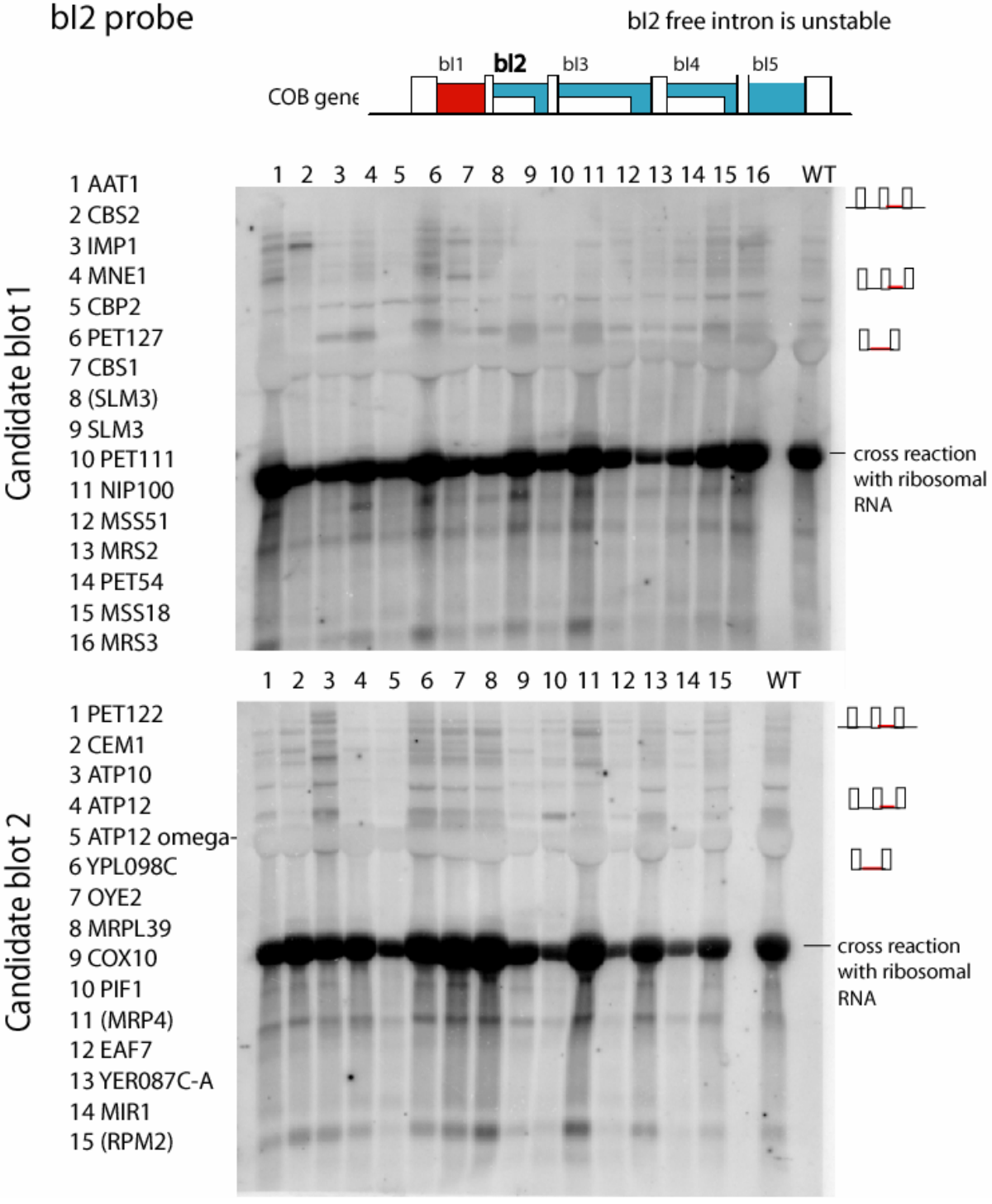


Candidate splicing factor Blots

bI1 probe



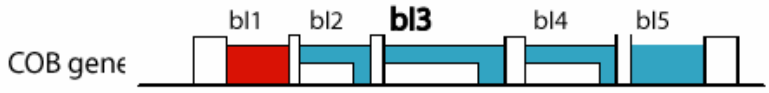
Candidate splicing factor Blots



Candidate splicing factor Blots

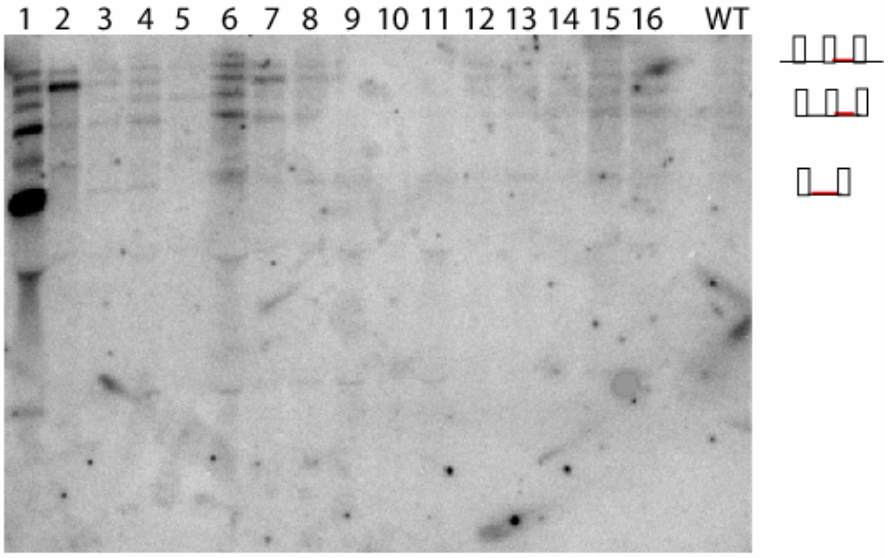
bI3 probe

bI3 free intron is unstable



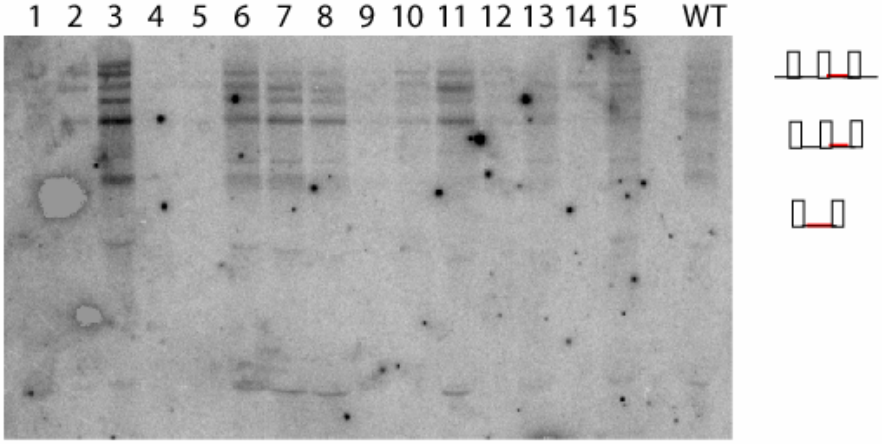
Candidate blot 1

1 AAT1
2 CBS2
3 IMP1
4 MNE1
5 CBP2
6 PET127
7 CBS1
8 (SLM3)
9 SLM3
10 PET111
11 NIP100
12 MSS51
13 MRS2
14 PET54
15 MSS18
16 MRS3



Candidate blot 2

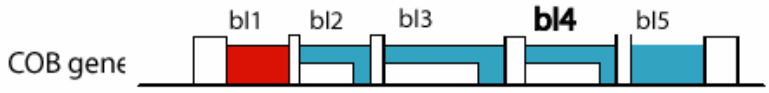
1 PET122
2 CEM1
3 ATP10
4 ATP12
5 ATP12 omega-
6 YPL098C
7 OYE2
8 MRPL39
9 COX10
10 PIF1
11 (MRP4)
12 EAF7
13 YER087C-A
14 MIR1
15 (RPM2)



Candidate splicing factor Blots

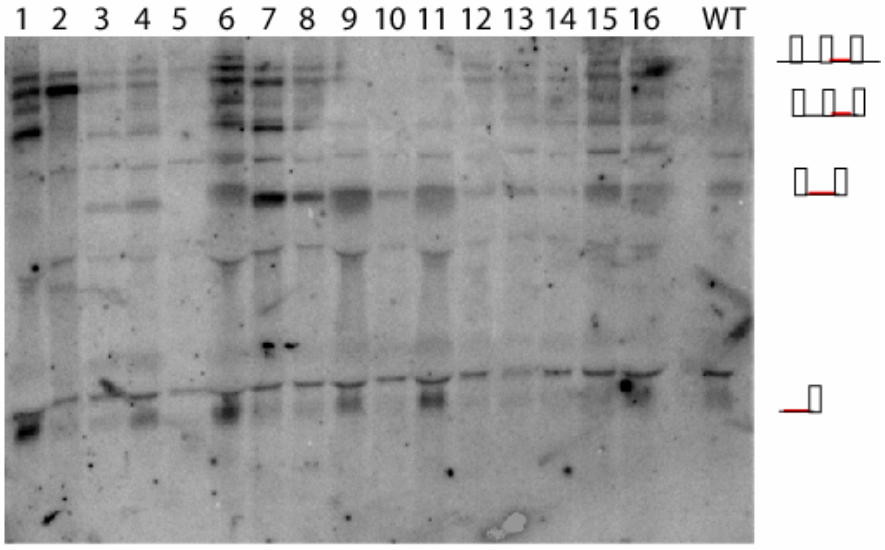
bI4 probe

bI4 free intron is unstable



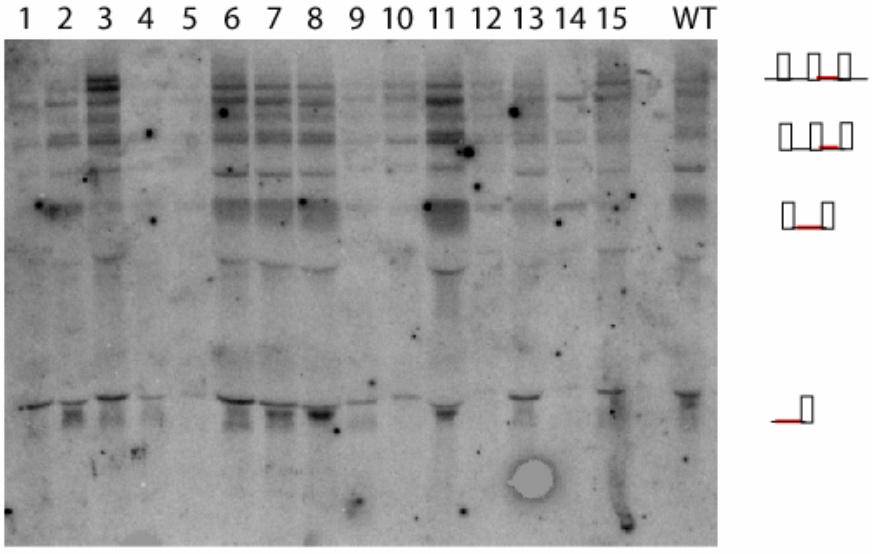
Candidate blot 1

1 AAT1
2 CBS2
3 IMP1
4 MNE1
5 CBP2
6 PET127
7 CBS1
8 (SLM3)
9 SLM3
10 PET111
11 NIP100
12 MSS51
13 MRS2
14 PET54
15 MSS18
16 MRS3



Candidate blot 2

1 PET122
2 CEM1
3 ATP10
4 ATP12
5 ATP12 omega-
6 YPL098C
7 OYE2
8 MRPL39
9 COX10
10 PIF1
11 (MRP4)
12 EAF7
13 YER087C-A
14 MIR1
15 (RPM2)



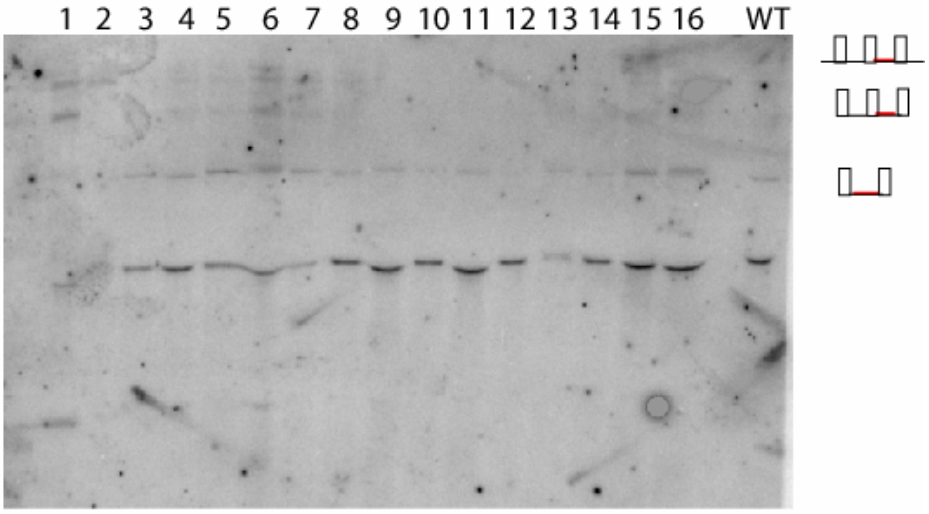
Candidate splicing factor Blots

bI5 probe

bI5 free intron is unstable

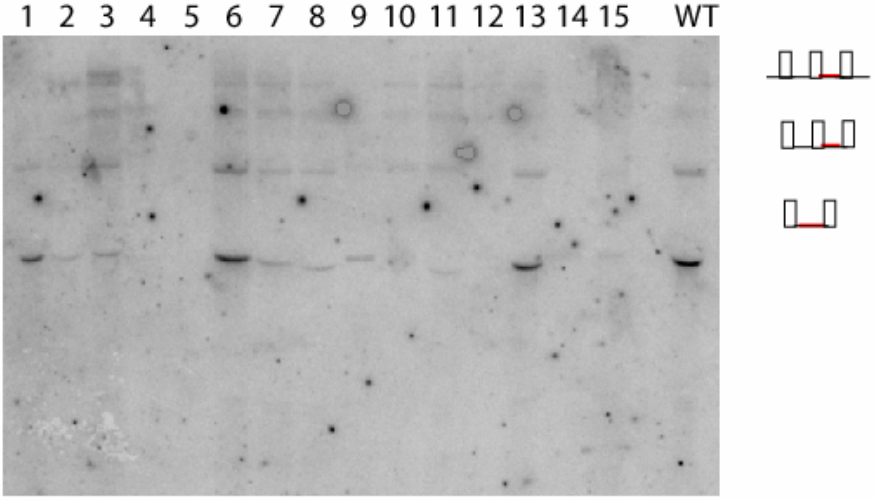
Candidate blot 1

1 AAT1
2 CBS2
3 IMP1
4 MNE1
5 CBP2
6 PET127
7 CBS1
8 (SLM3)
9 SLM3
10 PET111
11 NIP100
12 MSS51
13 MRS2
14 PET54
15 MSS18
16 MRS3



Candidate blot 2

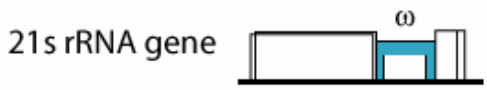
1 PET122
2 CEM1
3 ATP10
4 ATP12
5 ATP12 omega-
6 YPL098C
7 OYE2
8 MRPL39
9 COX10
10 PIF1
11 (MRP4)
12 EAF7
13 YER087C-A
14 MIR1
15 (RPM2)



Candidate splicing factor Blots

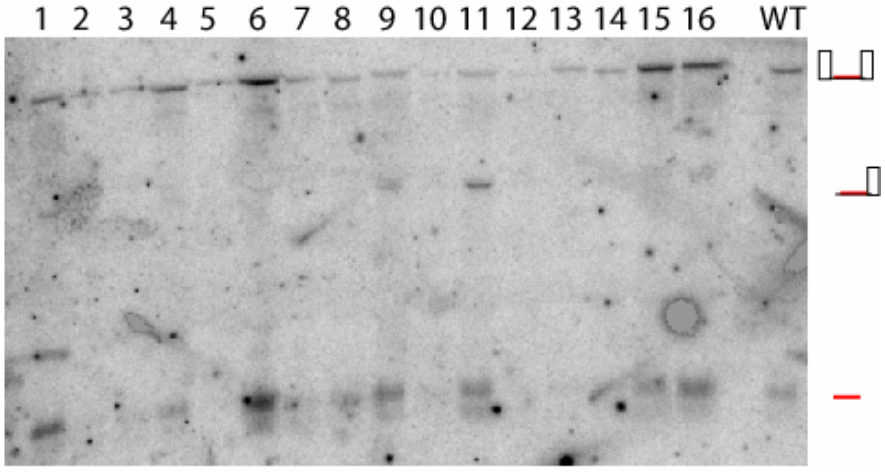
ω probe

ω free intron is stable



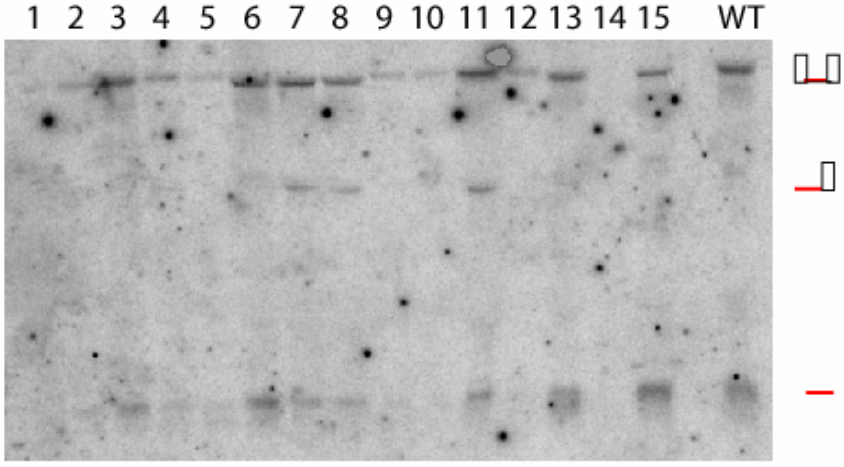
Candidate blot 1

1 AAT1
2 CBS2
3 IMP1
4 MNE1
5 CBP2
6 PET127
7 CBS1
8 (SLM3)
9 SLM3
10 PET111
11 NIP100
12 MSS51
13 MRS2
14 PET54
15 MSS18
16 MRS3



Candidate blot 2

1 PET122
2 CEM1
3 ATP10
4 ATP12
5 ATP12 omega-
6 YPL098C
7 OYE2
8 MRPL39
9 COX10
10 PIF1
11 (MRP4)
12 EAF7
13 YER087C-A
14 MIR1
15 (RPM2)



APPENDIX F
Descriptions and blot locations of candidate splicing factor genes

The following table lists the genes discussed in the results section of Chapter 4 along with a brief description of the known protein function, if any and on which blots the RNA samples may be found. Refer to Appendix D and Appendix E for the blot images.

ORF name	Gene name	Description of protein function	location blot	location lane
YBR230C		hypothetical ORF, no homology (v. weak atpase)	5 7	13 7
YDL067C	*COX9*	Subunit VIIa of cytochrome c oxidase,	32	2
YDL069C	*CBS1*	translational activator of COB mRNA	c 1	7
YDR197W	*CBS2*	Translational activator of COB mRNA	21 c 1	10 2
YDR204W	*COQ4*	ubiquinone (Coenzyme Q) biosynthesis. Forms complex with COQ3	19	2
YER058W	*PET117*	Protein required for assembly of cytochrome c oxidase	15 21	10 1
YER061C	*CEM1*	homology with beta-keto-acyl synthases	15 c 2	11 2
YGR222W	*PET54*	aids COX1 intron AI5 beta splicing; COX3 mRNA translation	4 c 1	16 14
YHL005C	*(MRP4)*	dubious ORF - overlapps with MRP4,a mito ribosomal protein that posesses tRNA binding activity	1 9 c 2	9 8 11
YHL038C	*CBP2*	required for splicing of COB aI5 intron	4 c 1	5 5

ORF name	Gene name	Description of protein function	location blot	location lane
YHR179W	*OYE2*	NADPH dehydrogenase (old yellow enzyme), isoform 1	9 c 2	12 7
YJL209W	*CBP1*	Protein required for COB mRNA stability or 5' processing	36	6
YJR077C	*MIR1*	mitochondrial, reduced mito proteins	4 c 2	21 14
YKL106W	*AAT1*	aspartate transaminase, mitochondrial	17 21 c 1	5 7 1
YKL208W	*CBT1*	processing of the 3' end of cytochrome b pre-mRNA, may also affect splicing	23	3
YLR393W	*ATP10*	chaprone: F1F0 ATPase complex assembly , slow growth w/ *nip100*	18 c 2	7 3
YML009c	*MRPL39*	Mitochondrial ribosomal protein of the large subunit	6 7 26 c 2	7 12 15 8
YML061C	*PIF1*	DNA helicase involved in repair and recombination of mitochondrial DNA	5 7 c 2	11 2 10
YMR063W	*RIM9*	Regulator of IME2	25	4
YMR150C	*IMP1*	Catalytic subunit of the mitochondrial inner membrane peptidase complex	12 20 c 1	2 3 3
YMR257C	*PET111*	translational activator for the COX2 mRNA, located in the mito inner membrane	12 20 c 1	10 5 10
YOL096C	*COQ3*	O-methyltransferase involved in ubiquinone (Coenzyme Q) biosynthesis	25	6
YOR017W	*PET127*	RNA stability/processing. Supresses pet122	12 20 c 1	18 6 6
YOR334W	*MRS2*	RNA splicing protein and member of the mitochondrial carrier family (MCF)	5 7 c 1	12 3 13
YOR350C	*MNE1*	similar to Lucilia illustris mitochondria cytochrome oxidase	13 20 c 1	8 7 4
YPL174C	*NIP100*	Nuclear import protein	14 c 1	10 11

APPENDIX G
Phase II: Petite strains

Collection of petite strians from the systematic deletion library.

All strains in this list were either initially tested to be gly^-, ρ^- ir, or appeared to be loosing mtDNA during phase I of this screen.

Prior to storage of these strains in fresh 96 well plates (Column: petite sub-grid) they were re-tested for growth on glycerol and the presence of mtDNA. (Column: re-tested phenotype).

The results of a cross to an intronless strain are also shown (Column: Result of I^o cross).

	Systematic ORF name	*Gene name*	General protein function	re-tested phenotype	Result of I^o cross	petite sub-grid		
						plate	col	row
1	YCR046C	*IMG1*	m ribosome	gly^- / ρ^-		1	1	B
2	YCR071C	*IMG2*	m ribosome	gly^- / ρ^-		1	1	C
3	YDR347W	*MRP1*	m ribosome	gly^- / ρ^-		1	1	D
4	YDL045W-A	*MRP10*	m ribosome	gly^- / ρ^-		1	1	E
5	YKL003C	*MRP17*	m ribosome	gly^- / ρ^-		1	1	F
6	YDR405W	*MRP20*	m ribosome	gly^- / ρ^-		1	1	G
7	YBL090W	*MRP21*	m ribosome	gly^- / ρ^-		1	1	H
8	YPL118W	*MRP51*	m ribosome	gly^- / ρ^-		1	2	A
9	YNL005C	*MRP7*	m ribosome	gly^- / ρ^-		1	2	B
10	YNL284C	*MRPL10*	m ribosome	gly^- / ρ^-		1	2	C

	Systematic ORF name	*Gene name*	General protein function	re-tested phenotype	Result of I^o cross	petite sub-grid plate	col	row
11	YDL202W	*MRPL11*	m ribosome	gly^- / ρ^-		1	2	D
12	YKR006C	*MRPL13*	m ribosome	gly^- / ρ^-	n/a	1	2	E
13	YLR312W-A	*MRPL15*	m ribosome	gly^- / ρ^-		1	2	F
14	YBL038W	*MRPL16*	m ribosome	gly^- / ρ^-		1	2	G
15	YKR085C	*MRPL20*	m ribosome	gly^- / ρ^-	$gly^{+?}$	1	2	H
16	YNL177C	*MRPL22*	m ribosome	gly^- / ρ^-	n/a	1	3	A
17	YOR150W	*MRPL23*	m ribosome	gly^- / ρ^-		1	3	B
18	YGR076C	*MRPL25*	m ribosome	gly^- / ρ^-		1	3	C
19	YBR282W	*MRPL27*	m ribosome	gly^- / ρ^-		1	3	D
20	YCR003W	*MRPL32*	m ribosome	gly^- / ρ^-		1	3	E
21	YMR286W	*MRPL33*	m ribosome	gly^- / ρ^-		1	3	F
22	YDR322W	*MRPL35*	m ribosome	gly^- / ρ^-		1	3	G
23	YBR122C	*MRPL36*	m ribosome	gly^- / ρ^-	gly^+	1	3	H
24	YBR268W	*MRPL37*	m ribosome	gly^- / ρ^-		1	4	A
25	YKL170W	*MRPL38*	m ribosome	gly^- / ρ^+	gly^+	1	4	B
26	YLR439W	*MRPL4*	m ribosome	gly^- / ρ^-		1	4	C
27	YLR439W	MRPL4	m ribosome	gly^- / ρ^-	n/a	1	4	D
28	YPL173W	*MRPL40*	m ribosome	gly^- / ρ^-		1	4	E
29	YJL096W	*MRPL49*	m ribosome	gly^- / ρ^-		1	4	F
30	YPR100W	*MRPL51*	m ribosome	gly^- / ρ^-		1	4	G
31	YHR147C	*MRPL6*	m ribosome	gly^- / ρ^-		1	4	H
32	YDR237W	*MRPL7*	m ribosome	gly^- / ρ^-		1	5	A
33	YJL063C	*MRPL8*	m ribosome	gly^- / ρ^-	$gly^{+/-}$	1	5	B
34	YGR220C	MRPL9	m ribosome	gly^- / ρ^-	n/a	1	5	C
35	YPL013C	*MRPS16*	m ribosome	gly^- / ρ^-		1	5	D
36	YGR165W	*MRPS35*	m ribosome	gly^- / ρ^-		1	5	E
37	YBR251W	*MRPS5*	m ribosome	gly^- / ρ^-		1	5	F
38	YMR158W	*MRPS8*	m ribosome	gly^- / ρ^-	$gly^{+/-}$	1	5	G
39	YOR158W	*PET123*	m ribosome	gly^- / ρ^-		1	5	H
40	YEL050C	*RML2*	m ribosome	gly^- / ρ^-		1	6	A
41	YER050C	*RSM18*	m ribosome	gly^- / ρ^-		1	6	B
42	YNR037C	*RSM19*	m ribosome	gly^- / ρ^-		1	6	C
43	YKL155C	*RSM22*	m ribosome	gly^- / ρ^-		1	6	D
44	YGL129C	*RSM23*	m ribosome	gly^- / ρ^-		1	6	E
45	YDR175C	*RSM24*	m ribosome	gly^- / ρ^-		1	6	F
46	YGR215W	*RSM27*	m ribosome	gly^- / ρ^-		1	6	G
47	YJR113C	*RSM7*	m ribosome	gly^- / ρ^-		1	6	H
48	YNL081C	*SWS2*	m ribosome	gly^- / ρ^+	gly^+	2	1	A
49	YDR042C	*(RSM10)*	m ribosome	gly^- / ρ^-	gly^+	2	1	C

	Systematic ORF name	*Gene name*	General protein function	re-tested phenotype	Result of I^o cross	petite sub-grid plate	col	row
50	YDR114C	(MRPL1)	m ribosome	gly^- / ρ^-		2	1	D
51	YDR115W	YDR115W	m ribosome	gly^- / ρ^+	gly^+	2	1	E
52	YGR219W	(MRPL9)	m ribosome	gly^- / ρ^-		2	1	F
53	YKL169C	(MRPL38)	m ribosome	gly^- / ρ^-		2	1	G
54	YNL184C	(MRPL19)	m ribosome	gly^- / ρ^-	gly^+	2	1	H
55	YNR036C	YNR036C	m ribosome	gly^- / ρ^-		2	2	A
56	YPR099C	(MRPL51)	m ribosome	gly^- / ρ^-		2	2	B
57	YHR011W	*DIA4*	m tRNA	gly^- / ρ^-		2	2	C
58	YPL104W	*MSD1*	m tRNA	gly^- / ρ^-		2	2	D
59	YOL033W	*MSE1*	m tRNA	gly^- / ρ^-		2	2	E
60	YPR047W	*MSF1*	m tRNA	gly^- / ρ^-		2	2	F
61	YNL073W	*MSK1*	m tRNA	gly^- / ρ^-		2	2	G
62	YGR171C	*MSM1*	m tRNA	gly^- / ρ^-		2	2	H
63	YHR091C	*MSR1*	m tRNA	gly^- / ρ^-		2	3	A
64	YKL194C	*MST1*	m tRNA	gly^- / ρ^-		2	3	B
65	YMR287C	*MSU1*	m tRNA	gly^- / ρ^-		2	3	C
66	YDR268W	*MSW1*	m tRNA	gly^- / ρ^-		2	3	D
67	YPL097W	*MSY1*	m tRNA	gly^- / ρ^-		2	3	E
68	YLR382C	*NAM2*	m tRNA	gly^- / ρ^-		2	3	F
69	YBL012C	(*FMT*1)	m tRNA	gly^- / ρ^-		2	3	G
70	YOR201C	*PET56*	m transc./ othr RNA	gly^- / ρ^-		2	3	H
71	YDR194C	*MSS116*	m transc./ othr RNA	gly^- / ρ^+		2	4	A
72	YPL029W	*SUV3*	m transc./ othr RNA	gly^- / ρ^-		2	4	B
73	YGL064C	*MRH4*	m transc./ othr RNA	gly^- / ρ^-		2	4	C
74	YBR163W	*DEM1*	m transc./ othr RNA	gly^- / ρ^+	gly^+	2	4	D
75	YOR199W	(PET56)	m transc./ othr RNA	gly^- / ρ^-	gly^+	2	4	E
76	YOR200W	(PET56)	m transc./ othr RNA	gly^- / ρ^-	gly^+	2	4	F
77	YFL036W	*RPO41*	m transc./ othr RNA	gly^- / ρ^-		2	4	G
78	YDL044C	*NAM1*	m transc./ othr RNA	gly^- / ρ^-	gly^+	2	4	H
79	YLR270W	*DCS1*	m transc./ othr RNA	gly^- / ρ^-	gly^+	2	5	A
80	YLR139C	*SLS1*	m transc./ othr RNA	gly^- / ρ^-		2	5	B
81	YPL005W	*AEP3*	m transc./ othr RNA	gly^- / ρ^-		2	5	C
82	YMR228W	*MTF1*	m transc./ othr RNA	gly^- / ρ^-		2	5	D
83	YMR064W	*AEP1*	m translation	gly^- / ρ^-	gly^+	2	5	E
84	YLR069C	*MEF1*	m translation	gly^- / ρ^-		2	5	F
85	YJL102W	*MEF2*	m translation	gly^- / ρ^-		2	5	G
86	YGL143C	*MRF1*	m translation	gly^- / ρ^-		2	5	H
87	YMR097C	MTG1	m translation	gly^- / ρ^-		2	6	A
88	YBL080C	*PET112*	m translation	gly^- / ρ^-		2	6	B

	Systematic ORF name	*Gene name*	General protein function	re-tested phenotype	Result of I^o cross	petite sub-grid plate	col	row
89	YLR067C	*PET309*	m translation	gly^- / ρ^-	gly^+	2	6	C
90	YMR066W	*SOV1*	m translation	gly^- / ρ^-		2	6	D
91	YOR187W	*TUF1*	m translation	gly^- / ρ^-		2	6	E
92	YKL134C	*OCT1*	m protein fold/proc.	gly^- / ρ^-		3	1	A
93	YFL016C	*MDJ1*	m protein fold/proc.	gly^- / ρ^-	n/a	3	1	B
94	YHR168W	*MTG2*	m protein fold/proc.	gly^- / ρ^-	$gly^{+/-}$	3	1	D
95	YER154W	*OXA1*	m protein fold/proc.	gly^- / ρ^-	gly^+	3	1	E
96	YGR101W	*PCP1*	m protein fold/proc.	gly^- / ρ^-		3	1	F
97	YDR079W	*PET100*	m protein fold/proc.	gly^- / ρ^-		3	1	G
98	YBL022C	*PIM1*	m protein fold/proc.	gly^- / ρ^-		3	1	H
99	YDL104C	*QRI7*	m protein fold/proc.	gly^- / ρ^-	gly^+	3	2	A
100	YLR369W	*SSQ1*	m protein fold/proc.	gly^- / ρ^-	gly^+	3	2	B
101	YMR089C	*YTA12*	m protein fold/proc.	gly^- / ρ^-	gly^+	3	2	C
102	YMR072W	*ABF2*	m DNA maint./repair	gly^- / ρ^-	gly^+	3	2	D
103	YER017C	*YTA10*	m protein fold/proc.	gly^- / ρ^-		3	2	E
104	YBR179C	FZO1	m DNA maint./repair	gly^- / ρ^-	n/a	3	2	F
105	YOR241W	*MET7*	m DNA maint./repair	gly^- / ρ^-	gly^+	3	2	G
106	YOR211C	*MGM1*	m DNA maint./repair	gly^- / ρ^-		3	2	H
107	YJR144W	*MGM101*	m DNA maint./repair	gly^- / ρ^-		3	3	A
108	YDR296W	*MHR1*	m DNA maint./repair	gly^- / ρ^-		3	3	B
109	YOR330C	MIP1	m DNA maint./repair	gly^- / ρ^-	n/a	3	3	C
110	YLL006W	*MMM1*	m DNA maint./repair	gly^- / ρ^-	gly^+	3	3	D
111	YHR120W	*MSH1*	m DNA maint./repair	**gly^+ / ρ^+**	n/a	3	3	E
112	YCR028C-A	*RIM1*	m DNA maint./repair	gly^- / ρ^-	$gly^{+/-}$	3	3	F
113	YGL218W	*SRF1 (MDM34)*	m stress resistance (mDNA maint./repair)	**gly^+ / ρ^+**	n/a	3	3	G
114	YDL198C	*YHM1*	m DNA maint./repair	gly^- / ρ^-	gly^+	3	3	H
115	YMR083W	*ADH3*	m respir./fermt.	gly^- / ρ^-	gly^+	3	4	A
116	YPL078C	*ATP4*	m respir./fermt.	gly^- / ρ^-		3	4	B
117	YDR298C	*ATP5*	m respir./fermt.	gly^- / ρ^-	gly^+	3	4	C
118	YNL315C	*ATP11*	m respir./fermt.	**gly^- / ρ^+**	n/a	3	4	D
119	YPL271W	*ATP15*	m respir./fermt.	gly^- / ρ^-		3	4	E
120	YDR377W	*ATP17*	m respir./fermt.	gly^- / ρ^-		3	4	F
121	YDR350C	*ATP22*	m respir./fermt.	gly^- / ρ^-	n/a	3	4	G
122	YPL059W	*GRX5*	m respir./fermt.	gly^- / ρ^-	gly^+	3	4	H
123	YER014W	*HEM14*	m respir./fermt.	gly^- / ρ^-		3	5	A
124	YDL181W	*INH1*	m respir./fermt.	gly^- / ρ^-	$gly^{+/-}$	3	5	B
125	YPR067W	*ISA2*	m respir./fermt.	gly^- / ρ^-		3	5	C
126	YMR267W	PPA2	m respir./fermt.	**gly^+ / ρ^+**	n/a	3	5	D

	Systematic ORF name	*Gene name*	General protein function	re-tested phenotype	Result of I^o cross	petite sub-grid plate	col	row
127	YDR529C	*QCR7*	m respir./fermt.	gly^- / ρ^-	n/a	3	5	E
128	YLR091W	YLR091W	m respir./fermt.	gly^- / ρ^-	gly^+	3	5	F
129	YGR257C	*MTM1*	m other	gly^- / ρ^-	n/a	3	5	G
130	YGL107C	RMD9	m other	**gly^+ / ρ^+**	n/a	3	5	H
131	YDR470C	UGO1	m other	gly^- / ρ^-	n/a	3	6	A
132	YJR122W	*CAF17*	m unknown fx	gly^- / ρ^-	$gly^{+/-}$	3	6	B
133	YGL240W	*DOC1*	m unknown fx	gly^- / ρ^-	gly^+	3	6	C
134	YOR205C	*FMP38*	m unknown fx	gly^- / ρ^-		3	6	D
135	YJL023C	*PET130*	m unknown fx	gly^- / ρ^-		3	6	E
136	YLR204W	*QRI5*	m unknown fx	gly^- / ρ^-		3	6	F
137	YDR065W	YDR065W	m unknown fx	gly^- / ρ^-		3	6	G
138	YGR102C	YGR102C	m unknown fx	gly^- / ρ^-	gly^+	3	6	H
139	YGR150C	YGR150C	m unknown fx	gly^- / ρ^-		4	1	A
140	YMR098C	YMR098C	m unknown fx	gly^- / ρ^-		4	1	B
141	YNL213C	YNL213C	m unknown fx	gly^- / ρ^-	$gly^{+?}$	4	1	C
142	YOR305W	YOR305W	m unknown fx	**gly^+ / ρ^+**	gly^+	4	1	E
143	YPR116W	YPR116W	m unknown fx	gly^- **/ ρ^+**	gly^+	4	1	F
144	YMR293C	*RNT1*	non-mito RNA	gly^- / ρ^-		4	1	G
145	YER087W	YER087W	non-mito RNA	gly^- / ρ^-		4	1	H
146	YJL140W	*RPB4*	non-mito RNA	gly^- / ρ^-	n/a	4	2	A
147	YDR507C	*GIN4*	non-mito known	gly^- / ρ^-	gly^+	4	2	B
148	YER122C	*GLO3*	non-mito known	gly^- / ρ^-	gly^+	4	2	C
149	YGL237C	HAP2	non-mito known	**gly^+ / ρ^+**	n/a	4	2	D
150	YBL021C	HAP3	non-mito known	**gly^+ / ρ^+**	n/a	4	2	E
151	YER110C	*KAP123*	non-mito known	gly^- / ρ^-	gly^+	4	2	F
152	YLR260W	*LCB5*	non-mito known	gly^- / ρ^-	gly^+	4	2	G
153	YPL031C	*PHO85*	non-mito known	gly^- / ρ^-		4	2	H
154	YOL100W	*PKH2*	non-mito known	gly^- / ρ^-	gly^+	4	3	A
155	YCR024C	*PMP1*	non-mito known	gly^- / ρ^-		4	3	B
156	YER070W	*RNR1*	non-mito known	gly^- / ρ^-	gly^+	4	3	C
157	YGR180C	*RNR4*	non-mito known	gly^- / ρ^-	gly^+	4	3	D
158	YJR004C	*SAG1*	non-mito known	gly^- / ρ^-		4	3	E
159	YGL095C	*VPS45*	non-mito known	gly^- / ρ^-	n/a	4	3	F
160	YAL048C	*GON1*	non-mito unknown	gly^- / ρ^-	gly^+	4	3	G
161	YLL033W	*GON2*	non-mito unknown	gly^- / ρ^-		4	3	H
162	YDL167C	*NRP1*	non-mito unknown	gly^- / ρ^-	gly^+	4	4	A
163	YNL170W	YNL170W	non-mito unknown	gly^- / ρ^-	gly^+	4	4	B

Petite strains that are able to maintain intronless mtDNA

The following table lists all the strains that when crossed to strain aMCC109/ Io, which contains the intronless Io mtDNA allele, became gly+. This list is not meant to be exhaustive –a negative result does not mean that a strain cannot maintain intronless mtDNA.

	Systematic ORF name	Gene name	General protein function	re-tested phenotype	Result of I^o cross	petite sub-grid plate	col	row
1	YBR122C	*MRPL36*	m ribosome	gly^- / ρ^-	gly^+	1	3	H
2	YKL170W	*MRPL38*	m ribosome	gly^- / ρ^+	gly^+	1	4	B
3	YJL063C	*MRPL8*	m ribosome	gly^- / ρ^-	$gly^{+/-}$	1	5	B
4	YMR158W	*MRPS8*	m ribosome	gly^- / ρ^-	$gly^{+/-}$	1	5	G
5	YNL081C	*SWS2*	m ribosome	gly^- / ρ^+	gly^+	2	1	A
6	YDR042C	*(RSM10)*	m ribosome	gly^- / ρ^-	gly^+	2	1	C
7	YDR115W		m ribosome	gly^- / ρ^+	gly^+	2	1	E
8	YNL184C	*(MRPL19)*	m ribosome	gly^- / ρ^-	gly^+	2	1	H
9	YBR163W	*DEM1*	m transc./ othr RNA	gly^- / ρ^+	gly^+	2	4	D
10	YOR199W	*(PET56)*	m transc./ othr RNA	gly^- / ρ^-	gly^+	2	4	E
11	YOR200W	*(PET56)*	m transc./ othr RNA	gly^- / ρ^-	gly^+	2	4	F
12	YDL044C	*NAM1*	m transc./ othr RNA	gly^- / ρ^-	gly^+	2	4	H
13	YLR270W	*DCS1*	m transc./ othr RNA	gly^- / ρ^-	gly^+	2	5	A
14	YMR064W	*AEP1*	m translation	gly^- / ρ^-	gly^+	2	5	E
15	YLR067C	*PET309*	m translation	gly^- / ρ^-	gly^+	2	6	C

	Systematic ORF name	Gene name	General protein function	re-tested phenotype	Result of I^o cross	petite sub-grid plate	col	row
16	YHR168W	*MTG2*	m protein fold/proc.	gly^- / ρ^-	$gly^{+/-}$	3	1	D
17	YER154W	*OXA1*	m protein fold/proc.	gly^- / ρ^-	gly^+	3	1	E
18	YDL104C	*QRI7*	m protein fold/proc.	gly^- / ρ^-	gly^+	3	2	A
19	YLR369W	*SSQ1*	m protein fold/proc.	gly^- / ρ^-	gly^+	3	2	B
20	YMR089C	*YTA12*	m protein fold/proc.	gly^- / ρ^-	gly^+	3	2	C
21	YMR072W	*ABF2*	m DNA maint./repair	gly^- / ρ^-	gly^+	3	2	D
22	YOR241W	*MET7*	m DNA maint./repair	gly^- / ρ^-	gly^+	3	2	G
23	YLL006W	*MMM1*	m DNA maint./repair	gly^- / ρ^-	gly^+	3	3	D
24	YCR028C-A	*RIM1*	m DNA maint./repair	gly^- / ρ^-	$gly^{+/-}$	3	3	F
25	YDL198C	*YHM1*	m DNA maint./repair	gly^- / ρ^-	gly^+	3	3	H
26	YMR083W	*ADH3*	m respir./fermt.	gly^- / ρ^-	gly^+	3	4	A
27	YDR298C	*ATP5*	m respir./fermt.	gly^- / ρ^-	gly^+	3	4	C
28	YPL059W	*GRX5*	m respir./fermt.	gly^- / ρ^-	gly^+	3	4	H
29	YDL181W	*INH1*	m respir./fermt.	gly^- / ρ^-	$gly^{+/-}$	3	5	B
30	YLR091W	*YLR091W*	m respir./fermt.	gly^- / ρ^-	gly^+	3	5	F
31	YJR122W	*CAF17*	m unknown fx	gly^- / ρ^-	$gly^{+/-}$	3	6	B
32	YGL240W	*DOC1*	m unknown fx	gly^- / ρ^-	gly^+	3	6	C
33	YGR102C	*YGR102C*	m unknown fx	gly^- / ρ^-	gly^+	3	6	H
34	YPR116W	*YPR116W*	m unknown fx	gly^- / ρ^+	gly^+	4	1	F
35	YDR507C	*GIN4*	non-mito known	gly^- / ρ^-	gly^+	4	2	B
36	YER122C	*GLO3*	non-mito known	gly^- / ρ^-	gly^+	4	2	C
37	YER110C	*KAP123*	non-mito known	gly^- / ρ^-	gly^+	4	2	F
38	YLR260W	*LCB5*	non-mito known	gly^- / ρ^-	gly^+	4	2	G
39	YOL100W	*PKH2*	non-mito known	gly^- / ρ^-	gly^+	4	3	A
40	YER070W	*RNR1*	non-mito known	gly^- / ρ^-	gly^+	4	3	C
41	YGR180C	*RNR4*	non-mito known	gly^- / ρ^-	gly^+	4	3	D
42	YAL048C	*GON1*	non-mito unknown	gly^- / ρ^-	gly^+	4	3	G
43	YDL167C	*NRP1*	non-mito unknown	gly^- / ρ^-	gly^+	4	4	A
44	YNL170W	*YNL170W*	non-mito unknown	gly^- / ρ^-	gly^+	4	4	B

APPENDIX H
Phenotypes of the mitochondrial ribosomal proteins.

Explanation of column headings:

ORF	Systematic name
Gene	common name
large or small	large or small ribosomal subunit
E. coli	Ecoli homolog
human	human mitochondrial homolog
Gly (pub)	SGD phenotype, (-) = growth defect on glycerol, (+) = probably no growth defect, (e) = essential gene note: this does not distinguish between slow growth and no growth on glycerol
gly/ρ this work	ability to grow on glycerol and presence of mtDNA (ρ^+ or ρ^-) as assessed in this study. note: gly^- does mean completely unable to grow on glycerol.
splicing phenotype	ability to splice the *COXI* and *COB* genes as assessed by northern blot analysis.

References for this table

The Saccharomyces Genome Database (SGD™) http://www.yeastgenome.org/

Saveanu C, *et al.* (2001) Identification of 12 new yeast mitochondrial ribosomal proteins including 6 that have no prokaryotic homologues. *J Biol Chem* 276(19):15861-7

Graack HR and Wittmann-Liebold B (1998) Mitochondrial ribosomal proteins (MRPs) of yeast. *Biochem J* 329 (Pt 3)():433-48

Gan X, *et al.* (2002) Tag-mediated isolation of yeast mitochondrial ribosome and mass spectrometric identification of its new components. *Eur J Biochem* 269(21):5203-14

	systematic ORF name	Gene	Subunit	homologs E. coli	homologs human	gly (pub)	gly/ρ this work	splicing phenotype
1	YDR347W	*MRP01*	S			-	gly$^-$ / ρ$^-$	
2	YDL045W-A	*MRP10*	S			-	gly$^-$ / ρ$^-$	
3	YKL003C	*MRP17*	S	S06	MRP-S6	-	gly$^-$ / ρ$^-$	
4	YBL090W	*MRP21*	S	S21	MRP-S21	-	gly$^-$ / ρ$^-$	
5	YPL118W	*MRP51*	S			-	gly$^-$ / ρ$^-$	
6	YBR251W	*MRPS05*	S	S05	MRP-S5	-	gly$^-$ / ρ$^-$	
7	YMR158W	*MRPS08*	S	S08		-	gly$^-$ / ρ$^-$	
8	YPL013C	*MRPS16*	S	S16	MRP-S16	-	gly$^-$ / ρ$^-$	
9	YGR165W	*MRPS35*	S			-	gly$^-$ / ρ$^-$	
10	YOR158W	*PET123*	S			-	gly$^-$ / ρ$^-$	
11	YJR113C	*RSM07*	S	S07	MRP-S7	-	gly$^-$ / ρ$^-$	
12	YER050C	*RSM18*	S	S18	IRP-S18(1-	-	gly$^-$ / ρ$^-$	
13	YNR037C	*RSM19*	S	S19		-	gly$^-$ / ρ$^-$	
14	YKL155C	*RSM22*	S			-	gly$^-$ / ρ$^-$	
15	YGL129C	*RSM23*	S		MRP-S29	-	gly$^-$ / ρ$^-$	
16	YDR175C	*RSM24*	S		MRP-S28	-	gly$^-$ / ρ$^-$	
17	YGR215W	*RSM27*	S			-	gly$^-$ / ρ$^-$	
18	YNL081C	*SWS2*	S	S13		-	gly$^-$ / ρ$^-$	
19	YNR036C		S	S12	MRP-S12	-	gly$^-$ / ρ$^-$	
20	YPR166C	*MRP02*	S	S14	MRP-S14	-	gly$^+$ / ρ$^+$	wt splicing partial (YHL005c)
21	YHL004W	*MRP04*	S	S02	MRP-S2	+	n/a	has defect at 20°C
22	YGR084C	*MRP13*	S			-	gly$^+$ / ρ$^+$	wt splicing
23	YMR188C	*MRPS17*	S	S17	MRP-S17	-	gly$^+$ / ρ$^+$	wt splicing
24	YDR337W	*MRPS28*	S	S15	MRP-S15	-	gly$^+$ / ρ$^+$	wt splicing
25	YIL093C	*RSM25*	S		MRP-S23	-	gly$^+$ / ρ$^+$	wt splicing
26	YFR049W	*YMR31*	S		MRP-S36	+	gly$^+$ / ρ$^+$	not analyzed
27	YHR075C	*MRPS02*	S			+	gly$^+$ / ρ$^+$	wt splicing
28	YBR146W	*MRPS09*	S	S09	MRP-S9	-	gly$^+$ / ρ$^+$	wt splicing
29	YNL306W	*MRPS18*	S	S11	MRP-S11	e		
30	YNL137C	*NAM9*	S			e		
31	YDR041W	*RSM10*	S	S10	MRP-S10	e		
32	YJR101W	*RSM26*	S			+	n/a	

	systematic ORF name	Gene	Subunit	homologs E. coli	homologs human	gly (pub)	gly/ρ this work	splicing phenotype
33	YCR046C	*IMG1*	L	L19	MRP-L19	-	gly$^-$ / ρ$^-$	
34	YCR071C	*IMG2*	L		MRP-L49	-	gly$^-$ / ρ$^-$	
35	YNL005C	*MRP07*	L	L27	MRP-L27	-	gly$^-$ / ρ$^-$	
36	YDR405W	*MRP20*	L	L23	MRP-L23	-	gly$^-$ / ρ$^-$	
37	YLR439W	*MRPL04*	L		MRP-L47	-	gly$^-$ / ρ$^-$	
38	YHR147C	*MRPL06*	L	L06		-	gly$^-$ / ρ$^-$	
39	YDR237W	*MRPL07*	L	L05		-	gly$^-$ / ρ$^-$	
40	YJL063C	*MRPL08*	L	L17	MRP-L17	-	gly$^-$ / ρ$^-$	
41	YGR220C	*MRPL09*	L	L03	MRP-L3	-	gly$^-$ / ρ$^-$	
42	YNL284C	*MRPL10*	L	L15	MRP-L15	-	gly$^-$ / ρ$^-$	
43	YDL202W	*MRPL11*	L	L10	MRP-L10	-	gly$^-$ / ρ$^-$	
44	YKR006C	*MRPL13*	L			-	gly$^-$ / ρ$^-$	
45	YLR312W-A	*MRPL15*	L			-	gly$^-$ / ρ$^-$	
46	YBL038W	*MRPL16*	L	L16	MRP-L16	-	gly$^-$ / ρ$^-$	
47	YKR085C	*MRPL20*	L			-	gly$^-$ / ρ$^-$	
48	YNL177C	*MRPL22*	L	L22	MRP-L22	-	gly$^-$ / ρ$^-$	
49	YOR150W	*MRPL23*	L	L13	MRP-L13	-	gly$^-$ / ρ$^-$	
50	YGR076C	*MRPL25*	L			-	gly$^-$ / ρ$^-$	
51	YBR282W	*MRPL27*	L		MRP-L41	-	gly$^-$ / ρ$^-$	
52	YCR003W	*MRPL32*	L	L32	MRP-L32	-	gly$^-$ / ρ$^-$	
53	YMR286W	*MRPL33*	L	L30	MRP-L30	-	gly$^-$ / ρ$^-$	
54	YDR322W	*MRPL35*	L		MRP-L38	-	gly$^-$ / ρ$^-$	
55	YBR122C	*MRPL36*	L	L31		+	gly$^-$ / ρ$^-$	
56	YBR268W	*MRPL37*	L		MRP-L54	-	gly$^-$ / ρ$^-$	
57	YKL170W	*MRPL38*	L	L14	MRP-L14	-	gly$^-$ / ρ$^-$	
58	YPL173W	*MRPL40*	L			-	gly$^-$ / ρ$^-$	
59	YJL096W	*MRPL49*	L	L21	MRP-L21	-	gly$^-$ / ρ$^-$	
60	YPR100W	*MRPL51*	L		MRP-L43	-	gly$^-$ / ρ$^-$	
61	YEL050C	*RML2*	L	L02	MRP-L2	-	gly$^-$ / ρ$^-$	
62	YDR115W		L			-	gly$^-$ / ρ$^-$	

	systematic ORF name	Gene	Subunit	homologs E. coli	homologs human	gly (pub)	gly/ρ this work	splicing phenotype
63	YPL183W-A	*GON5*	L			+	$gly^{+/-}$ / ρ^{+}	mild defect
64	YKL167C	*MRP49*	L			-	$gly^{+/-}$ / ρ^{+}	wt splicing
65	YDR116C	*MRPL01*	L	L01	MRP-L1	-	$gly^{+/-}$ / ρ^{+}	wt splicing
66	YMR024W	*MRPL03*	L		MRP-L44	+	$gly^{+/-}$ / ρ^{+}	wt splicing
67	YNL252C	*MRPL17*	L		MRP-L46	-	$gly^{+/-}$ / ρ^{+}	wt splicing
68	YMR193W	*MRPL24*	L	L28	MRP-L28	-	gly^{+} / ρ^{+}	wt splicing
69	YDR462W	*MRPL28*	L			-	gly^{+} / ρ^{+}	wt splicing
70	YKL138C	*MRPL31*	L			-	gly^{+} / ρ^{+}	wt splicing
71	YML009C	*MRPL39*	L	L33	MRP-L33	-	gly^{+cs} / ρ^{+}	mild defect at 20°C
72	YMR225C	*MRPL44*	L			+	gly^{+} / ρ^{+}	not analyzed
73	YNR022C	*MRPL50*	L			+	gly^{+} / ρ^{+}	not analyzed
74	YKL142W	*MRP08*	L			+	gly^{+} / ρ^{+}	wt splicing
75	YNL185C	*MRPL19*	L	L11	MRP-L11	+		
76	YML025C	*YML6*	L	L04	MRP-L4	e		
77	*YGL068W*		L	L07/L12	MRP-L7	e		

77 total yeast mitochondrial ribosomal genes

49 are rho minus
20 are rho plus
5 are essential
3 are unavailible

BIBLIOGRAPHY

Ackerman, S.H. and Tzagoloff, A. (1990)ATP 10, a yeast nuclear gene required for the assembly of the mitochondrial F1-F0 complex. *J.Biol.Chem.* 265 (17), 9952-9959.

Adams, P.L., Stahley, M.R., Gill, M.L., Kosek, A.B., Wang, J. and Strobel, S.A. (2004)Crystal structure of a group I intron splicing intermediate. *RNA.* 10 (12), 1867-1887.

Akins, R.A. and Lambowitz, A.M. (1987)A protein required for splicing group I introns in Neurospora mitochondria is mitochondrial tyrosyl-tRNA synthetase or a derivative thereof. *Cell* 50 (3), 331-345.

Amberg, D.C., Burke, D., Strathern, J.N., Burke, D. and Cold Spring, H.L. (2005)*Methods in yeast genetics*
a Cold Spring Harbor Laboratory course manual. Cold Spring Harbor, N.Y: Cold Spring Harbor Laboratory Press.

Anziano, P.Q., Hanson, D.K., Mahler, H.R. and Perlman, P.S. (1982)Functional domains in introns: trans-acting and cis-acting regions of intron 4 of the cob gene. *Cell* 30 (3), 925-932.

Bachl, J. and Schmelzer, C. (1990)Effect of deletions at structural domains of group II intron bI1 on self-splicing in vitro. *J.Mol.Biol.* 212 (1), 113-125.

Barrientos, A., Zambrano, A. and Tzagoloff, A. (2004)Mss51p and Cox14p jointly regulate mitochondrial Cox1p expression in Saccharomyces cerevisiae. *EMBO J.* 23 (17), 3472-3482.

Bassi, G.S., de Oliveira, D.M., White, M.F. and Weeks, K.M. (2002)Recruitment of intron-encoded and co-opted proteins in splicing of the bI3 group I intron RNA. *Proc.natl.Acad.Sci.U.S.A.* 99, 128-133.

Berget, S.M., Moore, C. and Sharp, P.A. (2000)Spliced segments at the 5' terminus of adenovirus 2 late mRNA. 1977. *Rev.Med.Virol.* 10 (6), 356-362.

Bernstein, E. and Allis, C.D. (2005)RNA meets chromatin. *Genes Dev.* 19 (14), 1635-1655.

Bonen, L. and Vogel, J. (2001)The ins and outs of group II introns. *Trends in Genetics* 17, 322-331.

Bonnefoy, N., Bsat, N. and Fox, T.D. (2001)Mitochondrial translation of Saccharomyces cerevisiae COX2 mRNA is controlled by the nucleotide sequence specifying the pre-Cox2p leader peptide. *Mol.Cell.Biol.* 21, 2359-2372.

Boulanger, S.C., Belcher, S.M., Schmidt, U., Dib-Hajj, S.D., Schmidt, T. and Perlman, P.S. (1995)Studies of point mutants define essential nucleotides in the domain 5 substructure of a group II intron. *Mol.Cell.Biol.* 15, 4479-4488.

Boulanger, S.C., Faix, P.H., Yang, H., Zhuo, J., Franzen, J.S., Peebles, C.L. and Perlman, P.S. (1996)Length changes in the joining segment between domains 5 and 6 of a group II intron inhibit self-splicing and alter 3' splice site selection. *Mol.Cell.Biol.* 16, 5896-5904.

Boule, J.B., Vega, L.R. and Zakian, V.A. (2005)The yeast Pif1p helicase removes telomerase from telomeric DNA. *Nature* 438 (7064), 57-61.

Bousquet, I., Dujardin, G., Poyton, R.O. and Slonimski, P.P. (1990)Two group I mitochondrial introns in the cob-box and coxI genes require the same MRS1/PET157 nuclear gene product for splicing. *Curr.Genet.* 18 (2), 117-124.

Boysen, J.H. and Mitchell, A.P. (2006)Control of Bro1-domain protein Rim20 localization by external pH, ESCRT machinery, and the Saccharomyces cerevisiae Rim101 pathway. *Mol.Biol.Cell* 17 (3), 1344-1353.

Brachmann, C.B., Davies, A., Cost, G.J., Caputo, E., Li, J., Hieter, P. and Boeke, J.D. (1998)Designer deletion strains derived from Saccharomyces cerevisiae S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast* 14 (2), 115-132.

Brown, N.G., Costanzo, M.C. and Fox, T.D. (1994)Interactions among three proteins that specifically activate translation of the mitochondrial COX3 mRNA in *Saccharomyces cerevisiae*. *Mol.Cell.Biol.* 14, 1045-1053.

Bui, D.M., Gregan, J., Jarosch, E., Ragnini, A. and Schweyen, R.J. (1999)The bacterial magnesium transporter CorA can functionally substitute for its putative homologue Mrs2p in the yeast inner mitochondrial membrane. *J.Biol.Chem.* 274 (29), 20438-20443.

Butcher, S.E. and Brow, D.A. (2005)Towards understanding the catalytic core structure of the spliceosome. *Biochem.Soc.Trans.* 33 (Pt 3), 447-449.

Butow, R.A., Henke, R.M., Moran, J.V., Belcher, S.C. and Perlman, P.S. (1996)Transformation of *Saccharomyces cerevisiae* mitochondria using the biolistic gun. In G.Attardi and A.Chomyn (eds) *Methods in Enzymology* (265-278)San Diego, N. Y., Boston: Academic Press, Inc.

Caprara, M.G., Lehnert, V., Lambowitz, A.M. and Westhof, E. (1996)A tyrosyl-tRNA synthetase recognizes a conserved tRNA-like structural motif i the group I intron catalytic core. *Cell* 87, 1135-1145.

Carignani, G., Groudinsky, O., Frezza, D., Schiavon, E., Bergantino, E. and Slonimski, P.P. (1983) An mRNA maturase is encoded by the first intron of the mitochondrial gene for the subunit I of cytochrome oxidase in S. cerevisiae. *Cell* 35, 733-742.

Cech, T.R. (1990)Self-splicing of group I introns. *Annu.Rev.Biochem.* 59, 543-568.

Chanfreau, G. and Jacquier, A. (1996)An RNA conformational change between the two chemical steps of group II self-splicing. *EMBO Journal* 15, 3466-3476.

Chao, M.Y. (2001)Genetic and biochemical studies of the group II intron aI2 RNP particle in the mitochondria of the bakers yeast Saccharomyces cerevisiae. Ph.D., University of Texas Southwestern Medical Center at Dallas.

Chapman, K.B. and Boeke, J.D. (1991)Isolation and characterization of the gene encoding yeast debranching enzyme. *Cell* 65 (3), 483-492.

Chevalier, B.S. and Stoddard, B.L. (2001)Homing endonucleases: structural and functional insight into the catalysts of intron/intein mobility. *Nucleic Acids Res.* 29 (18), 3757-3774.

Chu, V.T., Adamidi, C., Liu, Q., Perlman, P.S. and Pyle, A.M. (2001)Control of branch-site choice by a group II intron. *EMBO J.* 20, 6866-6876.

Chu, V., Liu, Q., Podar, M., Perlman, P. and Pyle, A. (1998)More than one way to splice an RNA: Branching without a bulge and splicing without branching in group II introns. *RNA* 4, 1186-1202.

Conde, J. and Fink, G.R. (1976)A mutant of Saccharomyces cerevisiae defective for nuclear fusion. *Proc.Natl.Acad.Sci.U.S.A* 73 (10), 3651-3655.

Conrad, H.M. (1987)Nuclear and mitochondrial mutations affecting mitochondrial RNA metabolism. xiv, 356 leaves, Ohio State University.

Conrad-Webb, H., Perlman, P.S., Zhu, H. and Butow, R.A. (1990)The nuclear SUV3-1 mutation affects a variety of post-transcriptional processes in yeast mitochondria. *Nucleic Acids Res.* 18 (6), 1369-1376.

Contamine, V. and Picard, M. (2000)Maintenance and integrity of the mitochondrial genome: a plethora of nuclear genes in the budding yeast. *Microbiol.Mol.Biol.Rev.* 64 (2), 281-315.

Costa, M., Deme, E., Jacquier, A. and Michel, F. (1997)Multiple tertiary interactions involving domain II of group II self-splicing introns. *J.Mol.Biol.* 267 (3), 520-536.

Costanzo, M.C. and Fox, T.D. (1990)Control of mitochondrial gene expression in Saccharomyces cerevisiae. *Annu.Rev.Genet.* 24, 91-113.

Costanzo, M.C., Seaver, E.C. and Fox, T.D. (1986)At least two nuclear gene products are specifically required for translation of a single yeast mitochondrial mRNA. *EMBO J.* 5 (13), 3637-3641.

Dai, L. and Zimmerly, S. (2003)ORF-less and reverse-transcriptase-encoding group II introns in archaebacteria, with a pattern of homing into related group II intron ORFs. *RNA.* 9 (1), 14-19.

Dib-Hajj, S.D., Boulanger, S.C., Hebbar, S.K., Peebles, C.L., Franzen, J.F. and Perlman, P.S. (1993)Domain 5 interacts with domain 6 and influences the second transesterification reaction of group II intron self-splicing. *Nucl.Acids.Res.* 21, 1797-1804.

Dickson, L., Huang, H.-R., Liu, L., Matsuura, M., Lambowitz, A.M. and Perlman, P.S. (2001)Retrotransposition of a yeast group II intron occurs by reverse splicing directly into ectopic DNA sites. *Proc.natl.Acad.Sci.U.S.A.* 98, 13207-13212.

Dieckmann, C.L., Pape, L.K. and Tzagoloff, A. (1982)Identification and cloning of a yeast nuclear gene (CBP1) involved in expression of mitochondrial cytochrome b. *Proc.Natl.Acad.Sci.U.S.A* 79 (6), 1805-1809.

Doherty, E.A. and Doudna, J.A. (2001)Ribozyme structures and mechanisms. *Annu.Rev.Biophys.Biomol.Struct.* 30, 457-475.

Doolittle, W.F. and Sapienza, C. (1980)Selfish genes, the phenotype paradigm and genome evolution. *Nature* 284 (5757), 601-603.

Doudna, J.A. and Cech, T.R. (2002)The chemical repertoire of natural ribozymes. *Nature* 418 (6894), 222-228.

Doudna, J.A. and Lorsch, J.R. (2005)Ribozyme catalysis: not different, just worse. *Nat.Struct.Mol.Biol.* 12 (5), 395-402.

Dujardin, G., Jacq, C. and Slonimski, P.P. (1982)Single base substitution in an intron of oxidase gene compensates splicing defects of the cytochrome b gene. *Nature* 298 (5875), 628-632.

Dujardin, G., Pajot, P., Groudinsky, O. and Slonimski, P.P. (1980)Long range control circuits within mitochondria and between nucleus and mitochondria. I. Methodology and phenomenology of suppressors. *Mol.Gen.Genet.* 179 (3), 469-482.

Dujon, B., Sherman, D., Fischer, G., Durrens, P., Casaregola, S., Lafontaine, I., De, M.J., Marck, C., Neuveglise, C., Talla, E., Goffard, N., Frangeul, L., Aigle, M., Anthouard, V., Babour, A., Barbe, V., Barnay, S., Blanchin, S., Beckerich, J.M., Beyne, E., Bleykasten, C., Boisrame, A., Boyer, J., Cattolico, L., Confanioleri, F., De, D.A., Despons, L., Fabre, E., Fairhead, C., Ferry-Dumazet, H., Groppi, A., Hantraye, F., Hennequin, C., Jauniaux, N., Joyet, P., Kachouri, R., Kerrest, A., Koszul, R., Lemaire, M., Lesur, I., Ma, L., Muller, H., Nicaud, J.M., Nikolski, M., Oztas, S., Ozier-Kalogeropoulos, O., Pellenz, S., Potier, S., Richard, G.F., Straub, M.L., Suleau, A., Swennen, D., Tekaia, F., Wesolowski-Louvel, M., Westhof, E., Wirth, B., Zeniou-Meyer, M., Zivanovic, I., Bolotin-Fukuhara, M., Thierry, A., Bouchier, C., Caudron, B., Scarpelli, C., Gaillardin, C., Weissenbach, J., Wincker, P. and Souciet, J.L. (2004)Genome evolution in yeasts. *Nature* 430 (6995), 35-44.

Dziembowski, A., Piwowarski, J., Hoser, R., Minczuk, M., Dmochowska, A., Siep, M., van der, S.H., Grivell, L. and Stepien, P.P. (2003)The yeast mitochondrial degradosome. Its composition, interplay between RNA helicase and RNase activities and the role in mitochondrial RNA metabolism. *J.Biol.Chem.* 278 (3), 1603-1611.

Edgell, D.R., Fast, N.M. and Doolittle, W.F. (1996)Selfish DNA: the best defense is a good offense. *Curr.Biol.* 6 (4), 385-388.

Eickbush, T.H. (2000)Molecular biology. Introns gain ground. *Nature* 404 (6781), 940-1, 943.

Ellis, T.P., Schonauer, M.S. and Dieckmann, C.L. (2005)CBT1 interacts genetically with CBP1 and the mitochondrially encoded cytochrome b gene and is required to stabilize the mature cytochrome b mRNA of Saccharomyces cerevisiae. *Genetics* 171 (3), 949-957.

Engelhardt, M.A., Doherty, E.A., Knitt, D.S., Doudna, J.A. and Herschlag, D. (2000)The P5abc peripheral element facilitates preorganization of the tetrahymena group I ribozyme for catalysis. *Biochemistry* 39 (10), 2639-2651.

Eskes, R., Liu, L., Ma, H., Chao, M.Y., Dickson, L., Lambowitz, A.M. and Perlman, P.S. (2000)Multiple homing pathways used by yeast mitochondrial group II introns. *Mol.Cell.Biol.* 20, 8432-8446.

Eskes, R., Yang, J., Lambowitz, A.M. and Perlman, P.S. (1997)Mobility of yeast mitochondrial group II introns: engineering a new site specificity and retrohoming via full reverse splicing. *Cell* 88, 865-874.

Faye, G. and Simon, M. (1983)Processing of the oxi-3 pre-messenger RNA in yeast. In R.J.Schweyen, K.Wolf and F.Kaudewitz (eds) *Mitochondria 1983: Nucleo-Mitochondrial Interactions; proceedings of a conference held in Schliersee, Germany, July 19-23, 1983* (433-439)Berlin; New York: de Gruyter.

Fedorova, O., Mitros, T. and Pyle, A.M. (2003)Domains 2 and 3 interact to form critical elements of the group II intron active site. *J.Mol.Biol.* 330 (2), 197-209.

Fedorova, O. and Pyle, A.M. (2005)Linking the group II intron catalytic domains: tertiary contacts and structural features of domain 3. *EMBO J.* 24 (22), 3906-3916.

Fedorova, O., Su, L.J. and Pyle, A.M. (2002)Group II introns: highly specific endonucleases with modular structures and diverse catalytic functions. *Methods* 28 (3), 323-335.

Foury, F. and Kolodynski, J. (1983)pif mutation blocks recombination between mitochondrial rho+ and rho- genomes having tandemly arrayed repeat units in Saccharomyces cerevisiae. *Proc.Natl.Acad.Sci.U.S.A* 80 (17), 5345-5349.

Foury, F., Roganti, T., Lecrenier, N. and Purnelle, B. (1998)The complete sequence of the mitochondrial genome of Saccharomyces cerevisiae. *FEBS Letters* 440, 325-331.

Fox, T.D. (1996)Genetics of mitochondrial translation. In J.W.B.Hershey, M.B.Mathews and N.Sonenburg (eds) *Translational Control* (733-758)Cold Spring Harbor: Cold Spring Harbor Laboratory Press.

Fox, T.D., Sanford, J.C. and McMullin, T.W. (1988)Plasmids can stably transform yeast mitochondria lacking endogenous mtDNA. *Proc.Natl.Acad.Sci.U.S.A* 85 (19), 7288-7292.

Gampel, A., Nishikimi, M. and Tzagoloff, A. (1989)CBP2 protein promotes in vitro excision of a yeast mitochondrial group I intron. *Mol.Cell Biol.* 9 (12), 5424-5433.

Gan, X., Kitakawa, M., Yoshino, K., Oshiro, N., Yonezawa, K. and Isono, K. (2002)Tag-mediated isolation of yeast mitochondrial ribosome and mass spectrometric identification of its new components. *Eur.J.Biochem.* 269 (21), 5203-5214.

Gavin, A.C., Bosche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J.M., Michon, A.M., Cruciat, C.M., Remor, M., Hofert, C., Schelder, M.,

Brajenovic, M., Ruffner, H., Merino, A., Klein, K., Hudak, M., Dickson, D., Rudi, T., Gnau, V., Bauch, A., Bastuck, S., Huhse, B., Leutwein, C., Heurtier, M.A., Copley, R.R., Edelmann, A., Querfurth, E., Rybin, V., Drewes, G., Raida, M., Bouwmeester, T., Bork, P., Seraphin, B., Kuster, B., Neubauer, G. and Superti-Furga, G. (2002)Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* 415 (6868), 141-147.

Giaever, G., Chu, A.M., Ni, L., Connelly, C., Riles, L., Veronneau, S., Dow, S., Lucau-Danila, A., Anderson, K., Andre, B., Arkin, A.P., Astromoff, A., El-Bakkoury, M., Bangham, R., Benito, R., Brachat, S., Campanaro, S., Curtiss, M., Davis, K., Deutschbauer, A., Entian, K.D., Flaherty, P., Foury, F., Garfinkel, D.J., Gerstein, M., Gotte, D., Guldener, U., Hegemann, J.H., Hempel, S., Herman, Z., Jaramillo, D.F., Kelly, D.E., Kelly, S.L., Kotter, P., LaBonte, D., Lamb, D.C., Lan, N., Liang, H., Liao, H., Liu, L., Luo, C., Lussier, M., Mao, R., Menard, P., Ooi, S.L., Revuelta, J.L., Roberts, C.J., Rose, M., Ross-Macdonald, P., Scherens, B., Schimmack, G., Shafer, B., Shoemaker, D.D., Sookhai-Mahadeo, S., Storms, R.K., Strathern, J.N., Valle, G., Voet, M., Volckaert, G., Wang, C.Y., Ward, T.R., Wilhelmy, J., Winzeler, E.A., Yang, Y., Yen, G., Youngman, E., Yu, K., Bussey, H., Boeke, J.D., Snyder, M., Philippsen, P., Davis, R.W. and Johnston, M. (2002)Functional profiling of the Saccharomyces cerevisiae genome. *Nature* 418 (6896), 387-391.

Gilbert, W. (1978)Why genes in pieces? *Nature* 271 (5645), 501.

Goda, S.K. and Minton, N.P. (1995)A simple procedure for gel electrophoresis and northern blotting of RNA. *Nucleic Acids Res.* 23 (16), 3357-3358.

Green-Willms, N.S., Butler, C.A., Dunstan, H.M. and Fox, T.D. (2001)Pet111p, an inner membrane-bound translational activator that limits expression of the Saccharomyces cerevisiae mitochondrial gene COX2. *J.Biol.Chem.* 276 (9), 6392-6397.

Gregan, J., Kolisek, M. and Schweyen, R.J. (2001)Mitochondrial Mg(2+) homeostasis is critical for group II intron splicing in vivo. *Genes Dev.* 15 (17), 2229-2237.

Grivell, L.A., rtal-Sanz, M., Hakkaart, G., de, J.L., Nijtmans, L.G., van, O.K., Siep, M. and van der, S.H. (1999)Mitochondrial assembly in yeast. *FEBS Lett.* 452 (1-2), 57-60.

Groudinsky, O., Dujardin, G. and Slonimski, P.P. (1981)Long range control circuits within mitochondria and between nucleus and mitochondria. II. Genetic and biochemical analyses of suppressors which selectively alleviate the mitochondrial intron mutations. *Mol.Gen.Genet.* 184 (3), 493-503.

Guo, H., Zimmerly, S., Perlman, P.S. and Lambowitz, A.M. (1997)Group II intron endonucleases use both RNA and protein subunits for recognition of specific sequences in double-stranded DNA. *EMBO J.* 16, 6835-6848.

Guo, W.W., Moran, J.V., Hoffman, P.W., Henke, R.M., Butow, R.A. and Perlman, P.S. (1995)The mobile group I intron 3 alpha of the yeast mitochondrial COXI gene encodes a 35-kDa processed protein that is an endonuclease but not a maturase. *J.Biol.Chem.* 270 (26), 15563-15570.

Hammond, S.M. (2005)Dicing and slicing: the core machinery of the RNA interference pathway. *FEBS Lett.* 579 (26), 5822-5829.

Harrington, L. (2003)Biochemical aspects of telomerase function. *Cancer Lett.* 194 (2), 139-154.

Haugen, P., Simon, D.M. and Bhattacharya, D. (2005)The natural history of group I introns. *Trends Genet.* 21 (2), 111-119.

Henke, R.M., Butow, R.A. and Perlman, P.S. (1995)Maturase and endonuclease functions depend on separate conserved domains of the bifunctional protein encoded by the group I intron aI4 alpha of yeast mitochondrial DNA. *EMBO J.* 14 (20), 5094-5099.

Hensgens, L.A.M., Arnberg, A.M., Rosendall, E., van der Horst, G., van der Veen, R., van Ommen, G.B. and Grivell, L.A. (1983)Variation, transcription, and circular RNAs of the mitochondrial gene for subunit 1 of cytochrome c oxidase. *J.Mol.Biol.* 164, 35-58.

Herbert, C.J., Dujardin, G., Labouesse, M. and Slonimski, P.P. (1988)Divergence of the mitochondrial leucyl tRNA synthetase genes in two closely related yeasts Saccharomyces cerevisiae and Saccharomyces douglasii: a paradigm of incipient evolution. *Mol.Gen.Genet.* 213 (2-3), 297-309.

Herrmann, J.M. and Funes, S. (2005)Biogenesis of cytochrome oxidase-sophisticated assembly lines in the mitochondrial inner membrane. *Gene* 354, 43-52.

Houman, F., Rho, S.B., Zhang, J., Shen, X., Wang, C.C., Schimmel, P. and Martinis, S.A. (2000)A prokaryote and human tRNA synthetase provide an essential RNA splicing function in yeast mitochondria. *Proc.Natl.Acad.Sci.U.S.A* 97 (25), 13743-13748.

Huang, H.R., Rowe, C.E., Mohr, S., Jiang, Y., Lambowitz, A.M. and Perlman, P.S. (2005)The splicing of yeast mitochondrial group I and group II introns requires a DEAD-box protein with RNA chaperone function. *Proc.Natl.Acad.Sci.U.S.A* 102 (1), 163-168.

Huang, H.-R., Chao, M.Y., Armstrong, B., Wang, Y., Lambowitz, A.M. and Perlman, P.S. (2003)The DIVa maturase binding site in the yeast group II intron aI2 is essential for intron homing but not for *in vivo* splicing. *Mol.Cell.Biol.* in press.

Huh, W.K., Falvo, J.V., Gerke, L.C., Carroll, A.S., Howson, R.W., Weissman, J.S. and O.S.E.K. (2003)Global analysis of protein localization in budding yeast. *Nature* 425, 686-691.

Hunte, C., Koepke, J., Lange, C., Rossmanith, T. and Michel, H. (2000)Structure at 2.3 A resolution of the cytochrome bc(1) complex from the yeast Saccharomyces cerevisiae co-crystallized with an antibody Fv fragment. *Structure.* 8 (6), 669-684.

Jacquier, A. (1990)Self-splicing group-II and nuclear pre-messenger RNA Introns - how similar are they. *Trends Biochem.Sci.* 15 (9), 351-354.

Jacquier, A. and Dujon, B. (1985)An intron-encoded protein is active in a gene conversion process that spreads an intron into a mitochondrial gene. *Cell* 41 (2), 383-394.

Jacquier, A. and Michel, F. (1987)Multiple exon-binding sites in class II self-splicing introns. *Cell* 50 (1), 17-29.

Jarrell, K.A., Dietrich, R.C. and Perlman, P.S. (1988a)Group II intron domain 5 facilitates a *trans*-splicing reaction. *Mol.Cell.Biol.* 8, 2361-2366.

Jarrell, K.A., Peebles, C.L., Dietrich, R.C., Romiti, S.L. and Perlman, P.S. (1988b)Group II intron self-splicing: alternative reaction conditions yield novel products. *J.Biol.Chem.* 263, 3432-39.

Jenkins, B.D. and Barkan, A. (2001)Recruitment of a peptidyl-tRNA hydrolase as a facilitator of group II intron splicing in chloroplasts. *EMBO J.* 20 (4), 872-879.

Jenkins, B.D., Kulhanek, D.J. and Barkan, A. (1997)Nuclear mutations that block group II RNA splicing in maize chloroplasts reveal several intron classes with distinct requirements for splicing factors. *Plant Cell* 9 (3), 283-296.

Johnson C.H. and McEwen J.E. (1997)Mitochondrial protein synthesis is not required for efficient excision of intron aI5beta from COX1 pre-mRNA in Saccharomycies cerevisiae. *Mol Gen Genet* 256 (1), 88-91.

Jones, E.W. and Fink, G.R. (1982)Regulation of amino acid and nucleotide biosynthesis in yeast. In J.N.Strathern, E.W.Jones and J.R.Broach (eds) *The Molecular biology of the yeast saccharomyces*
metabolism and gene expression (181-299)Cold Spring Harbor, N.Y: Cold Spring Harbor Laboratory.

Jurica, M.S. and Moore, M.J. (2003)Pre-mRNA splicing: awash in a sea of proteins. *Mol.Cell* 12 (1), 5-14.

Kahana, J.A., Schlenstedt, G., Evanchuk, D.M., Geiser, J.R., Hoyt, M.A. and Silver, P.A. (1998)The yeast dynactin complex is involved in partitioning the mitotic spindle between mother and daughter cells during anaphase B. *Mol.Biol.Cell* 9 (7), 1741-1756.

Kiss, T. (2002)Small nucleolar RNAs: an abundant group of noncoding RNAs with diverse cellular functions. *Cell* 109 (2), 145-148.

Koch, J.L., Boulanger, S.C., Dib-Hajj, S.D., Hebbar, S.K. and Perlman, P.S. (1992)Group II introns deleted for multiple substructures retain self-splicing activity. *Mol.Cell Biol.* 12 (5), 1950-1958.

Koll, H., Schmidt, C., Wiesenberger, G. and Schmelzer, C. (1987)Three nuclear genes suppress a yeast mitochondrial splice defect when present in high copy number. *Curr.Genet.* 12 (7), 503-509.

Krause, K., Lopes de, S.R., Roberts, D.G. and Dieckmann, C.L. (2004)The mitochondrial message-specific mRNA protectors Cbp1 and Pet309 are associated in a high-molecular weight complex. *Mol.Biol.Cell* 15 (6), 2674-2683.

Krause-Buchholz, U., Schobel, K., Lauffer, S. and Rodel, G. (2005)Saccharomyces cerevisiae translational activator Cbs1p is associated with translationally active mitochondrial ribosomes. *Biol.Chem.* 386 (5), 407-415.

Kreike, J., Bechmann, H., Van Hemert, F.J., Schweyen, R.J., Boer, P.H., Kaudewitz, F. and Groot, G.S. (1979)The identification of apocytochrome b as a mitochondrial gene product and immunological evidence for altered apocytochrome b in yeast strains having mutations in the COB region of mitochondrial DNA. *Eur.J.Biochem.* 101 (2), 607-617.

Kreike, J., Schulze, M., Ahne, F. and Lang, B.F. (1987)A yeast nuclear gene, MRS1, involved in mitochondrial RNA splicing: nucleotide sequence and mutational analysis of two overlapping open reading frames on opposite strands. *EMBO J.* 6 (7), 2123-2129.

Kreike, J., Schulze, M., Pillar, T., Korte, A. and Rodel, G. (1986)Cloning of a nuclear gene MRS1 involved in the excision of a single group I intron (bI3) from the mitochondrial COB transcript in S. cerevisiae. *Curr.Genet.* 11 (3), 185-191.

Kruger, K., Grabowski, P.J., Zaug, A.J., Sands, J., Gottschling, D.E. and Cech, T.R. (1982a)Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. *Cell* 31 (1), 147-157.

Kruger, K., Grabowski, P.J., Zaug, A.J., Sands, J., Gottschling, D.E. and Cech, T.R. (1982b)Self-splicing RNA: Autoexcision and autocyclization of the ribosomal RNA intervening sequence of tetrahymena. *Cell* 31 (1), 147-157.

Kumar, A., Agarwal, S., Heyman, J.A., Matson, S., Heidtman, M., Piccirillo, S., Umansky, L., Drawid, A., Jansen, R., Liu, Y. and et al. (2002)Subcellular localization of the yeast proteome. *Genes and Development* 16, 707-719.

Kurland, C.G. and Andersson, S.G. (2000)Origin and evolution of the mitochondrial proteome. *Microbiol.Mol.Biol.Rev.* 64 (4), 786-820.

Lambowitz, A.M. and Zimmerly, S. (2004)Mobile group II introns. *Annu.Rev.Genet.* 38, 1-35.

Lang, B.F., Ahne, F. and Bonen, L. (1985)The mitochondrial genome of the fission yeast Schizosaccharomyces pombe. The cytochrome b gene has an intron closely related to the first two introns in the Saccharomyces cerevisiae cox1 gene. *J.Mol.Biol.* 184 (3), 353-366.

Lazowska, J., Jacq, C. and Slonimski, P.P. (1980)Sequence of introns and flanking exons in wild-type and box3 mutants of cytochrome b reveals an interlaced splicing protein coded by an intron. *Cell* 22, 333-348.

Lazowska, J., Meunier, B. and Macadre, C. (1994)Homing of a group II intron in yeast mitochondrial DNA is accompanied by unidirectional co-conversion of upstream-located markers. *EMBO J* 13 (20), 4963-4972.

Lazowska, J., Szczepanek, T., Macadre, C. and Dokova, M. (1992)Two homologous mitochondrial introns from closely related Saccharomyces species differ by only a few amino acid replacements in their Open Reading Frames: one is mobile, the other is not. *C.R.Acad.Sci.III* 315 (2), 37-41.

Li, G.Y., Tian, G.L., Slonimski, P.P. and Herbert, C.J. (1996)The CBP2 gene from Saccharomyces douglasii is a functional homologue of the Saccharomyces cerevisiae gene and is essential for respiratory growth in the presence of a wild-type (intron-containing) mitochondrial genome. *Mol.Gen.Genet.* 250 (3), 316-322.

Li, W. and Mitchell, A.P. (1997)Proteolytic activation of Rim1p, a positive regulator of yeast sporulation and invasive growth. *Genetics* 145 (1), 63-73.

Liu, Q., Green, J.B., Khodadadi, A., Haeberli, P., Beigelman, L. and Pyle, A.M. (1997)Branch-site selection in a group II intron mediated by active recognition of the adenine amino group and steric exclusion of non-adenine functionalities. *J.Mol.Biol.* 267 (1), 163-171.

Longo, A., Leonard, C.W., Bassi, G.S., Berndt, D., Krahn, J.M., Hall, T.M. and Weeks, K.M. (2005)Evolution from DNA to RNA recognition by the bI3 LAGLIDADG maturase. *Nat.Struct.Mol.Biol.* 12 (9), 779-787.

Luan, D.D., Korman, M.H., Jakubczak, J.L. and Eickbush, T.H. (1993)Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for Non-LTR retrotransposition. *Cell* 72, 595-605.

Luban, C., Beutel, M., Stahl, U. and Schmidt, U. (2005)Systematic screening of nuclear encoded proteins involved in the splicing metabolism of group II introns in yeast mitochondria. *Gene* 354, 72-79.

Majumder, A.L., Akins, R.A., Wilkinson, J.G., Kelley, R.L., Snook, A.J. and Lambowitz, A.M. (1989)Involvement of tyrosyl-tRNA synthetase in splicing of group I introns in Neurospora crassa mitochondria: biochemical and immunochemical analyses of splicing activity. *Mol.Cell Biol.* 9 (5), 2089-2104.

Maleszka, R., Skelly, P.J. and Clark-Walker, G.D. (1991)Rolling circle replication of DNA in yeast mitochondria. *EMBO J.* 10 (12), 3923-3929.

Manthey, G.M., Przybyla-Zawislak, B.D. and McEwen, J.E. (1998)The Saccharomyces cerevisiae Pet309 protein is embedded in the mitochondrial inner membrane. *Eur.J.Biochem.* 255 (1), 156-161.

Marbois, B., Gin, P., Faull, K.F., Poon, W.W., Lee, P.T., Strahan, J., Shepherd, J.N. and Clarke, C.F. (2005)Coq3 and Coq4 define a polypeptide complex in yeast mitochondria for the biosynthesis of coenzyme Q. *J.Biol.Chem.* 280 (21), 20231-20238.

Marchler-Bauer, A. and Bryant, S.H. (2004)CD-Search: protein domain annotations on the fly. *Nucleic Acids Res.* 32 (Web Server issue), W327-W331.

Margossian, S., Li, H., Zassenhaus, H. and Butow, R. (1996)The DExH box protein Suv3p is a component of a yeast mitochondrial 3' -to-5' exoribonuclease that suppresses group I intron toxicity. *Cell* 84, 199-209.

Margulis, L. (1970)*Origin of eukaryotic cells;: Evidence and research implications for a theory of the origin and evolution of microbial, plant, and animal cells on the Precambrian earth*. Yale University Press.

Matlin, A.J., Clark, F. and Smith, C.W. (2005)Understanding alternative splicing: towards a cellular code. *Nat.Rev.Mol.Cell Biol.* 6 (5), 386-398.

Matsuura, M., Saldanha, R., Ma, H., Wank, H., Yang, J., Mohr, G., Cavanagh, S., Dunny, G.M., Belfort, M. and Lambowitz, A.M. (1997)A bacterial group II intron

encoding reverse transcriptase, maturase, and DNA endonuclease activities: biochemical demonstration of maturase activity and insertion of new genetic information within the intron. *Genes.Dev.* 11, 2910-2924.

McEwen, J.E., Hong, K.H., Park, S. and Preciado, G.T. (1993)Sequence and chromosomal localization of two PET genes required for cytochrome c oxidase assembly in Saccharomyces cerevisiae. *Curr.Genet.* 23 (1), 9-14.

McEwen, J.E., Ko, C., Kloeckner-Gruissem, B. and Poyton, R.O. (1986)Nuclear functions required for cytochrome c oxidase biogenesis in Saccharomyces cerevisiae. Characterization of mutants in 34 complementation groups. *J.Biol.Chem.* 261 (25), 11872-11879.

McGraw, P. and Tzagoloff, A. (1983)Assembly of the mitochondrial membrane system. Characterization of a yeast nuclear gene involved in the processing of the cytochrome b pre-mRNA. *J.Biol.Chem.* 258 (15), 9459-9468.

Michel, F. and Ferat, J.L. (1995)Structure and activities of group II introns. *Ann.Rev.Biochem.* 64, 435-461.

Michels, W.J., Jr. and Pyle, A.M. (1995)Conversion of a group II intron into a new multiple-turnover ribozyme that selectively cleaves oligonucleotides: elucidation of reaction mechanism and structure/function relationships. *Biochemistry* 34 (9), 2965-2977.

Mittelmeier, T.M. and Dieckmann, C.L. (1993)In vivo analysis of sequences necessary for CBP1-dependent accumulation of cytochrome b transcripts in yeast mitochondria. *Mol.Cell Biol.* 13 (7), 4203-4213.

Mohr, G. and Lambowitz, A.M. (2003)Putative proteins related to group II intron reverse transcriptase/maturases are encoded by nuclear genes in higher plants. *Nucleic Acids Res.* 31 (2), 647-652.

Mohr, G., Rennard, R., Cherniack, A.D., Stryker, J. and Lambowitz, A.M. (2001)Function of the Neurospora crassa mitochondrial tyrosyl-tRNA synthetase in RNA splicing. Role of the idiosyncratic N-terminal extension and different modes of interaction with different group I introns. *J.Mol.Biol.* 307, 75-92.

Mohr, G., Stryker, J. and Lambowitz, A.M. (2002)A DEAD-box protein functions as an ATP-dependent RNA chaperone in group I intron splicing. *Cell* 109, 769-779.

Mohr, S., Matsuura, M., Perlman, P.S. and Lambowitz, A.M. (2006)A DEAD-box protein alone promotes group II intron splicing and reverse splicing by acting as an RNA chaperone. *Proc.Natl.Acad.Sci.U.S.A* 103 (10), 3569-3574.

Moran, J.M., Zimmerly, S., Eskes, R., Kennell, J.C., Lambowitz, A.M., Butow, R.A. and Perlman, P.S. (1995)Mobile group II introns of yeast mitochondrial DNA are novel site-specific retroelements. *Mol.Cell.Biol.* 15, 2828-2838.

Moran, J.V., Wernette, C.M., Mecklenburg, K.L., Butow, R.A. and Perlman, P.S. (1992)Intron 5 alpha of the COXI gene of yeast mitochondrial DNA is a mobile group I intron. *Nucleic Acids Res.* 20 (15), 4069-4076.

Morin PJ, Subramanian GS and Gilmore TD. (1992)AAT1. a gene encoding a mitochondrial apsartate aminotransferase in Saccharomyces cereviseae. *Biochim Biophys Acta* 1171 (2), 211-214.

Muller, U.F. (2006)Re-creating an RNA world. *Cell Mol.Life Sci.*

Muroff, I. and Tzagoloff, A. (1990)CBP7 codes for a co-factor required in conjunction with a mitochondrial maturase for splicing of its cognate intervening sequence. *EMBO J.* 9 (9), 2765-2773.

Murray, H.L., Mikheeva, S., Coljee, V.W., Turczyk, B.M., Donahue, W.F., Bar-Shalom, A. and Jarrell, K.A. (2001)Excision of group II introns as circles. *Molecular Cell* 8, 201-211.

Naithani, S., Saracco, S.A., Butler, C.A. and Fox, T.D. (2003)Interactions among COX1, COX2, and COX3 mRNA-specific translational activator proteins on the inner surface of the mitochondrial inner membrane of *Saccharomyces cerevisiae*. *Mol.Biol.Cell* 14, 324-333.

Nielsen, H., Fiskaa, T., Birgisdottir, A.B., Haugen, P., Einvik, C. and Johansen, S. (2003)The ability to form full-length intron RNA circles is a general property of nuclear group I introns. *RNA.* 9 (12), 1464-1475.

Niemer, I., Schmelzer, C. and Borner, G.V. (1995)Overexpression of DEAD box protein pMSS116 promotes ATP-dependent splicing of a yeast group II intron in vitro. *Nucleic Acids Res.* 23 (17), 2966-2972.

Nilsen, T.W. (2003)The spliceosome: the most complex macromolecular machine in the cell? *Bioessays* 25 (12), 1147-1149.

Nunnari, J., Fox, T.D. and Walter, P. (1993)A mitochondrial protease with two catalytic subunits of nonoverlapping specificities. *Science* 262, 1997-2004.

O'Rourke, T.W., Doudican, N.A., Mackereth, M.D., Doetsch, P.W. and Shadel, G.S. (2002)Mitochondrial dysfunction due to oxidative mitochondrial DNA damage is reduced through cooperative actions of diverse proteins. *Mol.Cell Biol.* 22 (12), 4086-4093.

Ostheimer, G.J., Williams-Carrier, R., Belcher, S., Osborne, E., Gierke, J. and Barkan, A. (2003)Group II intron splicing factors derived by diversification of an ancient RNA-binding domain. *EMBO J.* 22 (15), 3919-3929.

Paukstelis, P.J., Coon, R., Madabusi, L., Nowakowski, J., Monzingo, A., Robertus, J. and Lambowitz, A.M. (2005)A tyrosyl-tRNA synthetase adapted to function in group I intron splicing by acquiring a new RNA binding surface. *Mol.Cell* 17 (3), 417-428.

Paul, M.F., Barrientos, A. and Tzagoloff, A. (2000)A single amino acid change in subunit 6 of the yeast mitochondrial ATPase suppresses a null mutation in ATP10. *J.Biol.Chem.* 275 (38), 29238-29243.

Peebles, C.L., Belcher, S.M., Zhang, M., Dietrich, R.C. and Perlman, P.S. (1993)Mutation of the conserved first nucleotide of a group II intron from yeast mitochondrial DNA reduces the rate but allows accurate splicing. *J.Biol.Chem.* 268, 11929-11938.

Peebles, C.L., Perlman, P.S., Mecklenburg, K.L., Petrillo, M.L., Tabor, J.H., Jarrell, K.A. and Chen, H.-L. (1986a)A self-splicing RNA excises an intron lariat. *Cell* 44, 213-223.

Peebles, C.L., Perlman, P.S., Mecklenburg, K.L., Petrillo, M.L., Tabor, J.H., Jarrell, K.A. and Cheng, H.-L. (1986b)A self-splicing RNA excises an intron lariat. *Cell* 44 (2), 213-223.

Peebles, C.L., Zhang, M., Perlman, P.S. and Franzen, J.S. (1995)Catalytically critical nucleotide in domain 5 of a group II intron. *Proc.Natl.Acad.Sci.U.S.A* 92 (10), 4422-4426.

Pereira de, S.A., Jubier, M.F., Delcher, E., Lancelin, D. and Lejeune, B. (1991)A trans-splicing model for the expression of the tripartite nad5 gene in wheat and maize mitochondria. *Plant Cell* 3 (12), 1363-1378.

Perlman, P. and Podar, M. (1996)Reactions catalyzed by group II introns in vitro. *Methods in Enzymology* 264, 66-86.

Pillar, T., Lang, B.F., Steinberger, I., Vogt, B. and Kaudewitz, F. (1983)Expression of the "split gene" cob in yeast mtDNA. Nuclear mutations specifically block the excision of different introns from its primary transcript. *J.Biol.Chem.* 258 (13), 7954-7959.

Piskur, J. (1994)Inheritance of the yeast mitochondrial genome. *Plasmid* 31 (3), 229-241.

Podar, M., Chu, V.T., Pyle, A.M. and Perlman, P.S. (1998)Group II intron splicing *in vivo* by first-step hydrolysis. *Nature* 391, 915-918.

Podar, M., Perlman, P.S. and Padgett, R.A. (1995)Stereochemical selectivity of group II intron splicing, reverse splicing, and hydrolysis reactions. *Mol.Cell.Biol.* 15, 4466-4478.

Podar, M. (1997)Biochemical, molecular and genetic investigations of the structure and mechanism of a group II intron catalytic RNA. Ph.D., University of Texas Southwestern Medical Center at Dallas.

Poyton, R.O. and McEwen, J.E. (1996)Crosstalk between nuclear and mitochondrial genomes. *Annu.Rev.Biochem.* 65, 563-607.

Prokisch, H., Scharfe, C., Camp, D.G. and et al. (2004)Integrative analysis of the mitochondrial proteome in yeast. *PLoS Biol* 2, E160.

Purnelle, B. and Goffeau, A. (1996)Nucleotide sequence analysis of a 40 kb segment on the right arm of yeast chromosome XV reveals 18 open reading frames including a new pyruvate kinase and three homologues to chromosome I genes. *Yeast* 12 (14), 1475-1481.

Qin, P.Z. and Pyle, A.M. (1998)The architectural organization and mechanistic function of group II intron structural elements. *Current Opinion in Structural Biology* 8, 301-308.

Reichert, A.S. and Neupert, W. (2004)Mitochondriomics or what makes us breathe. *Trends Genet.* 20 (11), 555-562.

Rho, S.B., Lincecum, T.L.Jr. and Martinis, S.A. (2002)An inserted region of leucyl-tRNA synthetase plays a critical role in group I intron splicing. *EMBO J.* 21, 6874-6881.

Rho, S.B. and Martinis, S.A. (2000)The bI4 group I intron binds directly to both its protein splicing partners, a tRNA synthetase and maturase, to facilitate RNA splicing activity. *RNA* 6, 1882-1894.

Rieger, K.J., Aljinovic, G., Lazowska, J., Pohl, T.M., Slonimski, P.P. and Aljinovic, G. (1997)A novel nuclear gene, CBT1, essential for mitochondrial cytochrome b formation: terminal processing of mRNA and intron dependence. *Curr.Genet.* 32 (3), 163-174.

RIS, H. and PLAUT, W. (1962)Ultrastructure of DNA-containing areas in the chloroplast of Chlamydomonas. *J.Cell Biol.* 13, 383-391.

Rodeheffer, M.S., Boone, B.E., Bryan, A.C. and Shadel, G.S. (2001)Nam1p, a protein involved in RNA processing and translation, is coupled to transcription through an interaction with yeast mitochondrial RNA polymerase. *J.Biol.Chem.* 276 (11), 8616-8622.

Roy, S.W. and Gilbert, W. (2006)The evolution of spliceosomal introns: patterns, puzzles and progress. *Nat.Rev.Genet.* 7 (3), 211-221.

Sambrook, J. and Russell, D.W. (2001)*Molecular cloning a laboratory manual*. Cold Spring Harbor, N.Y: Cold Spring Harbor Laboratory Press.

Sashital, D.G., Cornilescu, G., McManus, C.J., Brow, D.A. and Butcher, S.E. (2004)U2-U6 RNA folding reveals a group II intron-like domain and a four-helix junction. *Nat.Struct.Mol.Biol.* 11 (12), 1237-1242.

Schlatterer, J.C., Crayton, S.H. and Greenbaum, N.L. (2006)Conformation of the Group II intron branch site in solution. *J.Am.Chem.Soc.* 128 (12), 3866-3867.

Schmidt, U., Maue, I., Lehmann, K., Belcher, S.M., Stahl, U. and Perlman, P.S. (1998)Mutant alleles of the MRS2 gene of yeast nuclear DNA suppress mutations in the catalytic core of a mitochondrial group II intron. *J.Mol.Biol.* 282, 525-541.

Schmitt, M.E., Brown, T.A. and Trumpower, B.L. (1990)A rapid and simple method for preparation of RNA from Saccharomyces cerevisiae. *Nucleic Acids Res.* 18 (10), 3091-3092.

Schneider, A., Behrens, M., Scherer, P., Pratje, E., Michaelis, G. and Schatz, G. (1991)Inner membrane protease I, an enzyme mediating intra mitochondrial protein sorting in yeast. *EMBO J.* 10, 247-254.

Seraphin, B., Boulet, A., Simon, M. and Faye, G. (1987)Construction of a yeast strain devoid of mitochondrial introns and its use to screen nuclear genes involved in mitochondrial splicing. *Proc.Natl.Acad.Sci.U.S.A* 84 (19), 6810-6814.

Seraphin, B., Faye, G., Hatat, D. and Jacq, C. (1992)The yeast mitochondrial intron aI5 alpha: associated endonuclease activity and in vivo mobility. *Gene* 113 (1), 1-8.

Seraphin, B., Simon, M., Boulet, A. and Faye, G. (1989)Mitochondrial splicing requires a protein from a novel helicase family. *Nature* 337, 84-7.

Seraphin, B., Simon, M. and Faye, G. (1988)MSS18 a yeast nuclear gene involved in the splicing of intron aI5β of the mitochondrial *COX1* transcript. *EMBO J.* 7, 1455-64.

Sherman, F. (1963)Respiration-deficient mutants of yeast. I. Genetics. *Genetics* 48, 375-385.

Shukla, G.C. and Padgett, R.A. (2002)A catalytically active group II intron domain 5 can function in the U12-dependent spliceosome. *Mol.Cell* 9 (5), 1145-1150.

Sickman, A., Reinders, J., Wagner, Y., Joppich, C., Zahedi, R., Meyer, H.E., Schonfisch, B., Perschil, I., Chacinska, A., Guiard, B. and et al. (2003)The proteome of Saccharomyces cerevisiae mitochondria. *Proc.natl.Acad.Sci.U.S.A.* 100, 13207-13212.

Sigel, R.K., Sashital, D.G., Abramovitz, D.L., Palmer, A.G., Butcher, S.E. and Pyle, A.M. (2004)Solution structure of domain 5 of a group II intron ribozyme reveals a new RNA motif. *Nat.Struct.Mol.Biol.* 11 (2), 187-192.

Sigel, R.K., Vaidya, A. and Pyle, A.M. (2000)Metal ion binding sites in a group II intron core. *Nature Structural Biology* 7, 1111-1116.

Simon, M. and Faye, G. (1984)Steps in processing of the mitochondrial cytochrome oxidase subunit I pre-mRNA affected by a nuclear mutation in yeast. *Proc.Natl.Acad.Sci.U.S.A* 81 (1), 8-12.

Singh, R.N., Saldanha, R.J., D'Souza, L.M. and Lambowitz, A.M. (2002)Binding of a group II intron-encoded reverse transcriptase/maturase to its high affinity intron RNA binding site involves sequence-specific recognition and autoregulates translation. *J.Mol.Biol.* 318, 287-303.

Slonimski, P.P. and Tzagoloff, A. (1976)Localization in yeast mitochondrial DNA of mutations expressed in a deficiency of cytochrome oxidase and/or coenzyme QH2-cytochrome c reductase. *Eur.J.Biochem.* 61 (1), 27-41.

Steinmetz, L.M., Scharfe, C., Deutschbauer, A.M., Mokranjac, D., Herman, Z.S., Jones, T., Chu, A.M., Giaever, G., Prokisch, H., Oefner, P.J. and Davis, R.W. (2002)Systematic screen for human disease genes in yeast. *Nat.Genet.* 31, 400-404.

Storz, G., Altuvia, S. and Wassarman, K.M. (2005)An abundance of RNA regulators. *Annu.Rev.Biochem.* 74, 199-217.

Szczepanek, T. and Lazowska, J. (1996)Replacement of two non-adjacent amino acids in the S.cerevisiae bi2 intron-encoded RNA maturase is sufficient to gain a homing-endonuclease activity. *EMBO J.* 15 (14), 3758-3767.

Taanman, J.W. and Capaldi, R.A. (1992)Purification of yeast cytochrome c oxidase with a subunit composition resembling the mammalian enzyme. *J.Biol.Chem.* 267 (31), 22481-22485.

Tabak, H.F., Van der, H.G., Kamps, A.M. and Arnberg, A.C. (1987)Interlocked RNA circle formation by a self-splicing yeast mitochondrial group I intron. *Cell* 48 (1), 101-110.

Tabak, H.F., Van der, H.G., Osinga, K.A. and Arnberg, A.C. (1984)Splicing of large ribosomal precursor RNA and processing of intron RNA in yeast mitochondria. *Cell* 39 (3 Pt 2), 623-629.

Tilghman, S.M., Tiemeier, D.C., Seidman, J.G., Peterlin, B.M., Sullivan, M., Maizel, J.V. and Leder, P. (1978)Intervening sequence of DNA identified in the structural portion of a mouse beta-globin gene. *Proc.Natl.Acad.Sci.U.S.A* 75 (2), 725-729.

Tirupati, H.K., Shaw, L.C. and Lewin, A.S. (1999)An RNA binding motif in the Cbp2 protein required for protein-stimulated RNA catalysis. *J.Biol.Chem.* 274 (43), 30393-30401.

Tong, A.H., Lesage, G., Bader, G.D., Ding, H., Xu, H., Xin, X., Young, J., Berriz, G.F., Brost, R.L., Chang, M., Chen, Y., Cheng, X., Chua, G., Friesen, H., Goldberg, D.S., Haynes, J., Humphries, C., He, G., Hussein, S., Ke, L., Krogan, N., Li, Z., Levinson, J.N., Lu, H., Menard, P., Munyana, C., Parsons, A.B., Ryan, O., Tonikian, R., Roberts, T., Sdicu, A.M., Shapiro, J., Sheikh, B., Suter, B., Wong, S.L., Zhang, L.V., Zhu, H., Burd, C.G., Munro, S., Sander, C., Rine, J., Greenblatt, J., Peter, M., Bretscher, A., Bell, G., Roth, F.P., Brown, G.W., Andrews, B., Bussey, H. and Boone, C. (2004)Global mapping of the yeast genetic interaction network. *Science* 303 (5659), 808-813.

Toor, N., Hausner, G. and Zimmerly, S. (2001)Coevolution of group II intron RNA structures with their intron-encoded reverse transcriptases. *RNA.* 7 (8), 1142-1152.

Towpik, J. (2005)Regulation of mitochondrial translation in yeast. *Cell Mol.Biol.Lett.* 10 (4), 571-594.

Tzagoloff, A., Barrientos, A., Neupert, W. and Herrmann, J.M. (2004)Atp10p assists assembly of Atp6p into the F0 unit of the yeast mitochondrial ATPase. *J.Biol.Chem.* 279 (19), 19775-19780.

Tzagoloff, A. and Dieckmann, C.L. (1990)PET genes of Saccharomyces cerevisiae. *Microbiol.Rev.* 54 (3), 211-225.

Valencik, M.L., Kloeckener-Gruissem, B., Poyton, R.O. and McEwen, J.E. (1989)Disruption of the yeast nuclear PET54 gene blocks excision of mitochondrial intron aI5 beta from pre-mRNA for cytochrome c oxidase subunit I. *EMBO J.* 8 (12), 3899-3904.

Valencik, M.L. and McEwen, J.E. (1991)Genetic evidence that different functional domains of the PET54 gene product facilitate expression of the mitochondrial genes COX1 and COX3 in Saccharomyces cerevisiae. *Mol.Cell Biol.* 11 (5), 2399-2405.

van der Veen, R., Arnberg, A.C., van der Horst, G., Bonen, L., Tabak, H.F. and Grivell, L.A. (1986)Excised group II introns in yeast mitochondria are lariats and can be formed by self-splicing in vitro. *Cell* 44, 225-234.

van der Veen, R., Kwakman, J.H.J.M. and Grivell, L.A. (1987)Mutations at the lariat acceptor site allow self-splicing of a group II intron without lariat formation. *EMBO J.* 6, 1079-1084.

Van, D.E., Foury, F., Stillman, B. and Brill, S.J. (1992)A single-stranded DNA binding protein required for mitochondrial DNA replication in S. cerevisiae is homologous to E. coli SSB. *EMBO J.* 11 (9), 3421-3430.

Vogel, J. and Borner, T. (2002a)Lariat formation and a hydrolytic pathway in plant chloroplast group II intron splicing. *EMBO J.* 21 (14), 3794-3803.

Vogel, J. and Borner, T. (2002b)Lariat formation and a hydrolytic pathway in plant chloroplast group II intron splicing. *EMBO J.* 21, 3794-3803.

Vogel, J., Borner, T. and Hess, W.R. (1999)Comparative analysis of splicing of the complete set of chloroplast group II introns in three higher plant mutants. *Nucleic Acids Res.* 27 (19), 3866-3874.

Vogel, J., Hess, W. and Borner, T. (1997)Precise branch point mapping and quantification of splicing intermediates. *Nucleic Acids Research* 25, 2030-2031.

Wardleworth, B.N., Kvaratskhelia, M. and White, M.F. (2000)Site-directed mutagenesis of the yeast resolving enzyme Cce1 reveals catalytic residues and relationship with the intron-splicing factor Mrs1. *J.Biol.Chem.* 275 (31), 23725-23728.

Webb, A.E., Rose, M.A., Westhof, E. and Weeks, K.M. (2001)Protein-dependent transition states for ribonucleoprotein assembly. *J.Mol.Biol.* 309, 1087-1100.

Weeks, K.M. (1997)Protein-facilitated RNA folding. *Curr.Opin.Struct.Biol.* 7 (3), 336-342.

Weeks, K.M. and Cech, T.R. (1995a)Efficient protein-facilitated splicing of the yeast mitochondrial bI5 intron. *Biochemistry* 34, 7728-7738.

Weeks, K.M. and Cech, T.R. (1995b)Protein facilitation of group I intron splicing by assembly of the catalytic core and the 5' splice site domain. *Cell* 82, 221-230.

Weeks, K.M. and Cech, T.R. (1996)Assembly of a ribonucleoprotein catalyst by tertiary structure capture. *Science* 271, 345-348.

Weghuber, J., Dieterich, F., Froschauer, E.M., Svidova, S. and Schweyen, R.J. (2006)Mutational analysis of functional domains in Mrs2p, the mitochondrial Mg2+ channel protein of Saccharomyces cerevisiae. *FEBS J.* 273 (6), 1198-1209.

Wegierski, T., Dmochowska, A., Jablonowska, A., Dziembowski, A., Bartnik, E. and Stepien, P.P. (1998)Yeast nuclear PET127 gene can suppress deletions of the SUV3 or DSS1 genes: an indication of a functional interaction between 3' and 5' ends of mitochondrial mRNAs. *Acta Biochim.Pol.* 45 (4), 935-940.

Weiss-Brummer, B., Rodel, G., Schweyen, R.J. and Kaudewitz, F. (1982)Expression of the split gene cob in yeast: evidence for a precursor of a "maturase" protein translated from intron 4 and preceding exons. *Cell* 29 (2), 527-536.

Wiesenberger, G. and Fox, T.D. (1997)Pet127p, a membrane-associated protein involved in stability and processing of Saccharomyces cerevisiae mitochondrial RNAs. *Mol.Cell Biol.* 17 (5), 2816-2824.

Wiesenberger, G., Waldherr, M. and Schweyen, R.J. (1992)The nuclear gene MRS2 is essential for the excision of group II introns from yeast mitochondrial transcripts in vivo. *J.Biol.Chem.* 267 (10), 6963-6969.

Winkler, W.C. and Breaker, R.R. (2005)Regulation of bacterial gene expression by riboswitches. *Annu.Rev.Microbiol.* 59, 487-517.

Winter, A.J., Van der, H.G. and Tabak, H.F. (1988)Characterization of products derived from self-splicing of intron aI5 alpha which is located in the mitochondrial COX I gene of Saccharomyces cerevisiae. *Nucleic Acids Res.* 16 (9), 3845-3861.

Wright, R.M., Dircks, L.K. and Poyton, R.O. (1986)Characterization of COX9, the nuclear gene encoding the yeast mitochondrial protein cytochrome c oxidase subunit VIIa. Subunit VIIa lacks a leader peptide and is an essential component of the holoenzyme. *J.Biol.Chem.* 261 (36), 17183-17191.

Yang, J., Mohr, G., Perlman, P.S. and Lambowitz, A.M. (1998)Group II intron mobility in yeast mitochondria: target DNA primed reverse transcription activity of aI1

and reverse splicing into DNA transposition sites in vitro. *J.Mol.Biol.* 282, 505-523.

Yang, J., Zimmerly, S., Perlman, P.S. and Lambowitz, A.M. (1996)Efficient integration of an intron RNA into double-stranded DNA by reverse splicing. *Nature* 381, 332-335.

Zakharov, I.A. and Yarovoy, B.P. (1977)Cytoduction as a new tool in studying the cytoplasmic heredity in yeast. *Mol.Cell Biochem.* 14 (1-3), 15-18.

Zhang, L. and Doudna, J.A. (2002)Structural insights inton group II intron catalysis and branch-site selection. *Science* 295, 2084-2088.

Zhou, C., Yang, Y. and Jong, A.Y. (1990)Mini-prep in ten minutes. *Biotechniques* 8 (2), 172-173.

Zhu, H., Conrad-Webb, H., Liao, X.S., Perlman, P.S. and Butow, R.A. (1989)Functional expression of a yeast mitochondrial intron-encoded protein requires RNA processing at a conserved dodecamer sequence at the 3' end of the gene. *Mol.Cell Biol.* 9 (4), 1507-1512.

Zimmerly, S., Guo, H., Eskes, R., Yang, J., Perlman, P.S. and Lambowitz, A.M. (1995a)A group II intron RNA is a catalytic component of a DNA endonuclease involved in intron mobility. *Cell* 83, 529-538.

Zimmerly, S., Guo, H., Perlman, P.S. and Lambowitz, A.M. (1995b)Group II intron mobility occurs by target DNA-primed reverse transcription. *Cell* 82, 545-554.

Zimmerly, S., Hausner, G. and Wu, X. (2001)Phylogenetic relationships among group II intron ORFs. *Nucleic Acids Res.* 29 (5), 1238-1250.

VITAE

Tarah Michelle (Carter) Nyberg was born in Spokane, Washington, on March 19th, 1978, the daughter of Sandra Renee Carter and Timothy Antrim Carter. After completing her work at Mead High School, Spokane, Washington in 1996, she entered the University of Rochester in Rochester NY. She received the degree of Bachelor of Science with a major in Molecular Genetics from the University of Rochester in May, 2000. In August, 2000 she entered the Graduate School of Biomedical Sciences at the University of Texas Southwestern Medical Center at Dallas. She was awarded the degree of Doctor of Philosophy in August, 2006. In 1998, she married Timothy Jacob Nyberg. Their son, Isaac Yves Nyberg was born in 2004.

Permanent Address: 2100 Center Drive Apt 416
Vernon, TX 76384