Alleles of the yeast mitochondrial \textit{var}1 gene, which encode a protein (\textit{var}1) associated with the small mitochondrial ribosomal subunit, contain one or two identical GC clusters within the coding region that are transcribed and retained in the putative \textit{var}1 mRNA (Zassenhaus, H. P., and Butow, R. A. (1984) \textit{J. Biol. Chem.} 259, 8417–8421). By comparing peptide fragments generated by defined chemical and enzymatic cleavages of the products of these alleles, we show that these GC clusters encode amino acids in the \textit{var}1 protein. First, there is a strict correlation between the presence of an optional GC cluster in the \textit{var}1 gene and a corresponding increase in size of the peptide that would contain the “extra” amino acids encoded by that GC cluster. Second, we find proline residues in specific peptides of \textit{var}1 that, from DNA sequence, would only be present if the GC clusters were translated. Thus, although the yeast mitochondrial genome contains 70–100 GC clusters similar to those in \textit{var}1, the \textit{var}1 protein is the only mitochondrial translation product now known to contain amino acids encoded by these elements. We have also examined predictions of \textit{var}1 secondary structure and find little resemblance to the secondary structures predicted for most other ribosomal proteins. Finally, our analysis suggests a significant conformational difference between the \textit{var}1 protein containing amino acids encoded by the optional GC cluster and the form of the protein lacking those amino acids.

The yeast mitochondrial \textit{var}1 gene encodes a polymorphic protein called \textit{var}1 that is found associated with the small (38 S) mitochondrial ribosomal subunit (1–4). We have identified numerous \textit{var}1 alleles among unselected yeast stocks, each encoding a different molecular weight form of the \textit{var}1 protein, which differ by the presence or absence of certain DNA sequences within the coding region of the gene (4–7). One such sequence is a 46-base pair palindromic G + C-rich cluster (GC cluster) identical in sequence to a GC cluster found within the coding region of all \textit{var}1 alleles examined thus far (4, 7).

Curiously, these GC clusters, when present together in a \textit{var}1 allele, are arranged in opposite orientation 158 base pairs apart. Since these are the only GC clusters among the 70–100 such related sequences in the yeast mitochondrial genome known to be present in a major mitochondrial protein structural gene, it is of obvious interest to know if these unusual elements encode amino acids in the \textit{var}1 protein.

In the accompanying article (8), we showed that the GC clusters within \textit{var}1 are transcribed and present in stable RNAs, including the putative mRNA for the \textit{var}1 protein. However, since there is no reliable system presently available for the accurate \textit{in vitro} translation of defined mitochondrial RNAs, a demonstration that the GC clusters present in \textit{var}1 transcripts are translated into protein requires a direct analysis of the protein products made \textit{in vivo} or in isolated mitochondria. Here we describe the analysis of the products of two \textit{var}1 alleles that differ only by the presence or absence of the optional GC cluster. Our results show that both of the GC clusters within \textit{var}1 encode amino acids in the \textit{var}1 protein. Secondary structure predictions show, moreover, that, 1) \textit{var}1 does not resemble the structures predicted for any ribosomal proteins analyzed thus far, and 2) the addition of the 16 amino acids encoded by the optional GC cluster predicts significant structural changes in the protein.

**EXPERIMENTAL PROCEDURES**

**Strains**

Strain COP 161 (\textit{a}, \textit{ad lys tyr})\textsuperscript{1} contains the \textit{var}1[40.0] allele. According to our previous convention (9), this is designated \textit{a} + \textit{b} - with respect to inserts within \textit{var}1. A strain containing the \textit{var}1[41.8] allele (a + \textit{b} -) with the same nuclear background as COP 161 was constructed as follows: strain 5DSS (\textit{ura} \textit{p}) \textit{var}1[41.8] (a + \textit{b} -) was crossed to COP 19 (\textit{a ad lys tyr}) \textit{var}1[40.0] (a - \textit{b} -). Diploids issued from the cross containing the \textit{var}1[41.8] allele were identified by \textit{in vivo} labeling of mitochondrial translation products essentially as described by Strausberg and Butow (9). To switch the \textit{var}1[41.8] mitochondrial genome to a haploid COP 161 nuclear background, diploids containing the \textit{var}1[41.8] allele were sporulated and haploids with the appropriate mating type and nuclear markers were crossed to a \textit{kar} strain (\textit{a leu p}'). Haploid \textit{p} + cytoductants containing the COP 161 nucleus were identified and one isolate was selected and used for the studies described here.

**In Vitro Mitochondrial Protein Synthesis**

Mitochondria were prepared and mitochondrial translation products were labeled by a modification of the method of McKee and Poyton (10). Cells, grown to mid-logarithmic phase in medium containing 1% yeast extract, 1% bactopeptone, and 2% galactose (YPGal) were harvested by centrifugation, washed once with cold distilled water and once with 0.1 M EDTA, pH 7.4. The cells were resuspended in 1 ml of 0.1 M EDTA, pH 7.4, per g, wet weight, to which 1/40 volume of 2-mercaptoethanol was added. After incubation for 15 min at 30 °C with shaking, the cells were harvested by centrifugation, resuspended

\textsuperscript{1}The abbreviations used are: \textit{p} + , respiratory competent cells; \textit{p} - , respiratory incompetent petite cells lacking mitochondrial DNA; SDS, sodium dodecyl sulfate; NTCB, nitrocyanotobenzoic acid.
in 2 ml/g of cells of 1.0 M sorbitol and 0.1 M EDTA, pH 7.4. One mg of zymolyase 60,000 was added per g of cells and the suspension gently shaken for 20 min at 30 °C. Spheroplasts were harvested by centrifugation at 3000 × g for 10 min, washed once with cold 1.2 M sorbitol, pH 7.4, and resuspended in one-fourth the initial culture volume of YPGal containing 1 M sorbitol. The cell suspension was gently shaken for 20 min at 30 °C. Spheroplasts were harvested by centrifugation at 30°C for 10 min at 1.0 M sorbitol and 0.1 M EDTA, pH 7.4. One mg of protein was incubated with gentle shaking for 30 min at 30°C. Cell debris was removed by centrifugation for 10 min at 3000 × g. The supernatant was recovered, recentrifuged, and a mitochondrial pellet was pelleted and resuspended in 0.6 M sorbitol, pH 7.0. The mitochondrial pellet was resuspended in 0.6 M sorbitol, pH 7.0, at about 12 mg of protein/ml. Approximately 1 mg of mitochondrial protein was incubated at room temperature for 45 min/ml of medium containing 20 mM Tris-Cl, pH 7.2, 150 mM KCl, 14 mM KH2PO4, 12.5 mM MgCl2, 5 mM α-ketoglutarate, 4 mM ATP, 0.5 mM GTP, 5 mM phosphoenolpyruvate, 0.6 M mannitol, 0.2 mg/ml of bovine serum albumin, 0.1 mg/ml of cycloheximide, 1% malonate, 0.1 mM amino acids (excluding the amino acids used for labeling), 6.3 units of pyruvate kinase, and either [35S]methionine or [3,2,3,4,5-3H]proline. Labeled mitochondria were recovered by centrifugation for 15 min in an Eppendorf microfuge.

Purification of Radiolabeled var1

After in vitro labeling, mitochondria were pelleted and resuspended to a protein concentration of 20 mg/ml in 50 mM Tris-Cl, pH 7.5. A one-tenth volume of freshly prepared 5% Brij 35 was added to the mitochondria and the suspension vortexed. The lysate was placed on ice for 10 min with occasional mixing. Large membrane material was removed by centrifugation for 15 min in an Eppendorf microfuge. The amber colored supernatant was then recentrifuged in either a Beckman airfuge for 15 min at 15 p.s.i. or in a Sorvall SS34 rotor for 15 min at 20,000 rpm. To the supernatant was added one-third volume of 1 M sucrose, 2 M KCl, 40 mM Tris-Cl, pH 7.4, containing one-tenth volume of 10% chloroform (neutralized) and 42.8% (v/v) of saturated (NH4)2SO4. After 30 min on ice, the precipitate was recovered by centrifugation at 14,500 × g for 16 min at 4°C. The pellet was resuspended in 0.6 M sorbitol, 2 mM EDTA, and 0.1% bovine serum albumin, 0.1 mg/ml of cycloheximide, 1 mM malonate, 0.1 mM amino acids (excluding the amino acids used for labeling), 6.3 units of pyruvate kinase, and either [35S]methionine or [3,2,3,4,5-3H]proline. Labeled mitochondria were recovered by centrifugation for 15 min in an Eppendorf microfuge.

Protein Cleavages

Staphylococcus aureus V8 Protease—The protein sample was lyophilized to dryness and resuspended in 250 mM ammonium bicarbonate, pH 7.8, 10 mM EDTA; and 1.0% SLS. The samples were placed in a boiling water bath for 2 min. Four volumes of H2O and 20 μg of S. aureus V8 protease (11) were added and the samples incubated for 20–24 h at 37 °C. After digestion, the samples were heated for 2 min in electrophoresis sample buffer containing 0.05 M Tris-Cl, pH 6.8, 2% SDS, 2 mM EDTA, 1% 2-mercaptoethanol, and 10% glycerol, and stored at -20 °C until used.

Nitro-Genanobenzene Acid—The lyophilized protein sample was resuspended in 6.0 M guanidine HCl, 0.2 M Tris-Cl, pH 5.0, and 10 mM dithiothreitol. After incubation at 37 °C for 2 h, NTGB was added to a final concentration of 28 mg/ml, the pH rapidly readjusted to 8.0, and the samples incubated for 30 min at room temperature. Samples were then dialyzed overnight at 37 °C against 3000 volumes of 6 M guanidine HCl and 0.1 M sodium borate, pH 9.0. The samples were then dialyzed for several hours against 7000 volumes of 1% acetic acid. The sample was then analyzed directly or used for S. aureus V8 protease cleavage as described above.

Gel Electrophoresis

Samples were analyzed by electrophoresis on 30% SDS-polyacrylamide gels essentially as described previously (12). 35S-labeled material was visualized by autoradiography using Kodak XAR-5 film. H-labeled samples were visualized by fluorography.

Structure Predictions

Secondary structure predictions were made employing the method of Chou and Fasman (13, 14). This was done employing a program provided to us by Michael J. Kales (University of California, Berk-}

Translation of var1 GC Clusters

RESULTS

The locations of the amino acids that would be encoded by the var1 GC clusters are shown in Fig. 1 for two var1 alleles designated var1[40.0] and var1[41.8]. (We will refer to the protein products of these alleles as var1[40.0] and var1[41.8], respectively.) The GC cluster beginning at nucleotide 188 from the initiation codon in both of these alleles is called the common GC cluster since it is found in all var1 alleles examined thus far. The var1[41.8] allele contains an additional (optional) GC cluster beginning at nucleotide 393, downstream from the common GC cluster; both GC clusters have the identical sequence but are arranged in opposite orientation (Figs. 1 and 2). The nucleotide sequence and amino acids are
potentially encoded by these GC clusters is shown in Fig. 2. The optional GC cluster correlates with the a element described in previous studies (9) in which varl alleles containing the a element encode a varl protein that is about 1.8 kDa larger than the corresponding product of an allele lacking the a element. Other than the presence of the additional GC cluster in varl[41.8], there are no other differences we have observed in the varl coding region between these two alleles.

To determine whether these GC clusters encode amino acids in the varl protein, we have compared the peptide fragments obtained from enzymatic and chemical cleavages of radiochemically pure varl in which the varl protein was labeled in vitro in isolated mitochondria with either $^{35}$S methionine or $[^3H]$proline. Because of the biased distribution of GC-rich codons (varl is $\sim$93% A + T exclusive of the GC clusters), there is a biased distribution in the protein of some amino acids encoded by these sequences. In particular, proline residues in the NH$_2$-terminal half of the products of both alleles would be encoded exclusively by the GC clusters (Figs. 1 and 2). In contrast, varl contains 33 methionine residues that are distributed more or less uniformly throughout the protein (Fig. 1 and Ref. 4).

**Purification of Radiolabeled varl—**To obviate problems of dilution of labeled amino acids with large intracellular pools of unlabeled amino acid, mitochondrial translation products were labeled in vitro in isolated mitochondria. Fig. 3 shows that the major mitochondrial translation products labeled in vitro, including varl, correspond exactly to those labeled in vivo.

To prepare radiochemically pure varl protein from the in vitro labeled mitochondrial translation products, we took advantage of the fact that the varl protein is the only major yeast mitochondrial translation product that is not tightly bound to membranes (1, 2). Accordingly, varl can be selectively released from mitochondria with weak detergents, such as Brij 35, at detergent concentrations that lyse mitochondria but solubilize little of the mitochondrial membrane. Fig. 3 shows that at the ratio of 0.25 mg of Brij 35/mg of mitochondrial protein, varl is essentially the only labeled protein released.

**Enzymatic and Chemical Cleavage of the varl Protein—**The radiochemically labeled varl protein extracted with Brij 35 was concentrated by ammonium sulfate precipitation, equilibrated with 1% acetic acid, and lyophilized to dryness (see "Experimental Procedures"). Samples of this material were then used for cleavage reactions.

*S. aureus* V8 protease was used to cleave varl at glutamic acid 105 (Fig. 1 and Ref. 4). varl (40.0) and (41.8) contain 1 other glutamic acid residue each at position 204 and 220, respectively (4); however, since these residues are adjacent to a proline, they will not be cleaved by *S. aureus* V8 protease (11).

For chemical cleavage, varl was reacted with NTCB. This reagent reacts selectively with cysteine residues (18) and cleavage occurs at alkaline pH after the modified residue, varl(40.0) and (41.8) each contain 1 cysteine residue at position 190 and 206, respectively (Fig. 1 and Ref. 4).

If translated, the amino acids encoded by common GC cluster would correspond to residues 63–78 in the products of both the varl[40] and varl[41.8] alleles, while those encoded by the optional GC cluster would correspond to residues 131–146 in the product of the varl[41.8] allele. Protein cleavage methods were chosen such that peptides potentially containing the amino acids encoded by two GC clusters could be separated from each other as well as from the COOH terminus of the protein. Cleavage of varl(40.0) and varl(41.8) by NTCB together with *S. aureus* V8 would produce three peptides each (Fig. 1): 1) identical NH$_2$-terminal peptides containing the amino acids encoded by the common GC cluster; 2) an internal peptide that would be larger in varl(41.8) by the 16 amino acids encoded by the optional GC cluster, and 3) identical COOH-terminal peptides. As indicated in Fig. 1, only those peptides containing the COOH-terminal portion of the protein would be labeled with $[^3H]$proline if the GC clusters were not translated, whereas all but one peptide (peptide B of varl(40.0), Fig. 2) will be labeled if the GC clusters are translated.

Fig. 4 shows the peptides generated by single and combined cleavage of varl(40.0) and varl(41.0) labeled in vitro with $^{[35]S}$methionine. These results are compared to the predicted pattern assuming the GC clusters are translated. The results shown are fully consistent with the location and translation of the optional GC cluster depicted in Fig. 1; all peptides from the varl[41.8] product that are predicted to contain amino acids encoded by this GC cluster are an estimated 1.5 kDa larger than the corresponding peptides from the varl[40.0] product. Although the resolution between the 13.8- and the

![Fig. 3. Fidelity of the in vitro translation system and purification of radiolabeled varl protein.](image-url)
13.6-kDa peptide in the double digest of the varl(41.8) protein is not complete, a doublet was evident as a much broader band in this region compared to the single digest of either varl form with S. aureus V8, or to the double digest of varl(40.0). Fig. 4 also shows the predicted and experimental results of S. aureus V8 and NTCB cleavage of [3H]proline-labeled proteins. Although the resolution of these peptides is not as great as those labeled with [35S]methionine, the results are fully consistent with the translation of both GC clusters: 1) the NH2-terminal half of both varl forms are labeled with [3H]proline, and 2) a doublet is evident in the 13-14-kDa region of the gel in the double digest of varl(40.0) but only a singlet is apparent in the varl(41.8) digest.

**Secondary Structure**—We have not detected any significant amino acid sequence homology between varl and any prokaryotic or eukaryotic ribosomal proteins whose sequence is known. The possibility remained, however, that a relationship between varl and other ribosomal proteins could be detected by a comparison of predicted secondary structures. In addition, it was of interest to determine what structural effects the additional 16 amino acids encoded by the optional GC cluster might have on the protein. Accordingly, we have examined the secondary structural predictions for the varl(40.0) product employing the method of Chou and Fasman (13, 14). This analysis predicts a structure composed of 9% α-helix, 41% β-sheet, 40% β-turn, and 10% residual structure, with the bulk of the α-helix located at the NH2 terminus. This secondary structural composition is unusual for a small ribosomal subunit protein since these proteins have generally been found to be rich in α-helix with only a small amount of β-sheet (19-21). varl also has an unusually high amount of β-turn; it is roughly twice that of normal proteins and is due to the large amount of asparagine (31 mol %) (4). The detailed structural prediction for the varl(40.0) product in the region 109-145 is given in Fig. 5. The native structure is predicted to contain three β-turns (109–112, 131–134, and 142–145) and two sequences of β-sheet (113–130 and 135–141). The 16 amino acid insertion encoded by the optional GC cluster is predicted to have the structure a helix-β-sheet-α-turn. This segment is inserted in the middle of a β-turn of the varl(40.0) product (residues 131–134). The detailed structural prediction for residues 109–161 of the varl(41.8) product is also given in Fig. 5. Interestingly, residues 113–128 are predicted to be in an α-helix ± β-sheet conformational equilibrium, suggesting at least two conformers of this varl form. These two species are denoted Conformer I (with β-sheet segment) and Conformer II (with α-helix segment). Residues 109–161 of Conformer I has four regions of β-turn (109–112, 129–135, 144–148, and 150–161).
The hydrophobicity profile for the region corresponding to the insert site is given in Fig. 6, A and B, for varl(40.0) and varl(41.8), respectively. Considering the region 109-145 in Fig. 6A, three significant minima (ΔG trans approaching 0) of hydrophobicity and three significant maxima (ΔG trans > 1500 cal) of hydrophobicity are evident. In the corresponding region of varl(41.8) (109-161), four significant maxima and six significant minima are evident. In the region 132-147 (the position of the 16 amino acid insertion), it can be seen that a significant maximum of hydrophobicity and three significant minima of hydrophobicity are added to the native protein. Two of these minima are predicted to be β-turns. This analysis suggests a new region (residues 136-139) will be important in stabilizing tertiary structure and three new regions (residues 131-135, 140-141, and 144-147) will now be exposed to solvent. Hence, the dramatic changes in the hydrophobicity profile as well as changes in the secondary structure predicted by the method of Chou and Fasman (13, 14) suggest that the incorporation of the 16 amino acids encoded by the optional GC cluster will promote a significant change in conformation.

DISCUSSION

The accompanying article (8) showed that GC clusters in the varl coding region are transcribed and present in stable RNAs. Although it has not yet been possible to identify unambiguously the mRNA for the varl protein, there was no indication from these studies that these GC clusters are removed from varl transcripts by processing; thus, we could only infer from these data that the GC clusters in the varl coding sequences are expressed. The present work shows by direct analysis of varl proteins that the GC clusters are translated. First, we find a peptide in varl(41.8), which from DNA sequence is predicted to contain the extra amino acids encoded by the optional GC cluster, that is larger by the expected amount than the corresponding peptide from the product of an allele lacking that GC cluster. Second, we have shown the presence of proline residues in specific peptides generated from varl that would be present only if the GC clusters are translated.

The size and number of peptides generated by the S. aureus and NTCB cleavages shown in Fig. 4 provide further support for our assignment of the structural gene for the varl protein at the varl locus. This gene assignment was originally suggested by us on the basis of genetic mapping data (5, 22, 23) and documented more recently in a study showing a close correspondence between the experimentally determined amino acid composition of the varl protein with that predicted from the DNA sequence of the varl open reading frame (4). The distribution of [35S]methionine in the peptides generated from the NTCB and S. aureus V8 protease digestions of the
var1 proteins also supports our previous assignment of AUA as a methionine codon in the yeast mitochondrial genome (4) since peptides B and B' in Fig. 1 would not be labeled if AUG was the only codon that specified methionine.

Secondary Structure Analysis—Secondary structure prediction methods such as that of Chou and Fasman (13, 14) have been used extensively to predict the structures of the 30 and 50 S ribosomal subunit proteins of Escherichia coli (19–21, 24). In these studies, four different predictive methods have been employed, giving similar results. Independent physical measurements of the structures of these proteins in solution have been made in a few cases to test the predictions. For the S4, S8, L11, and L27 proteins of the E. coli ribosome, there is good correspondence between the overall secondary structure predicted from circular dichroism measurements and the secondary structure predicted from the primary sequence (19).

Although the absolute secondary structure of the entire protein is difficult to predict accurately in detail, local changes that result from, in this case the insertion of a sequence of amino acids, should be at least qualitatively predictable since, as noted above, the method of Chou and Fasman (13, 14) predicts turns with about 70–75% success rate.

As shown in Fig. 5 the Chou and Fasman method (13, 14) also predicts conformers of varl(41.8). These are distinguished in that in Conformer I, residues 113–128 are in the β-sheet conformation, while residues 113–128 in Conformer II are in the α-helix conformation. Similar conformational transitions have been proposed for glucagon (25). If this conformational equilibrium does exist, it is likely that the position of the equilibrium is shifted in favor of Conformer I since the transition to α-helix will disrupt an apparent three-strand antiparallel β-sheet. Based on the discussion of Richardson (26), such a transition would not be energetically favorable. In any event, the insertion of a 16-amino acid segment encoded by the optional GC cluster, with the secondary structure β-turn–β-sheet–β-turn in the middle of the β-turn, is likely to cause a significant conformational change.

Similarly, a comparison of the hydrophobicity profile between varl(40.0) and varl(41.8) suggests also that the insertion of the 16-amino acid segment may cause a significant change in tertiary structure. Specifically, the 16-amino acid insertion results in three new turn regions and a significant new region of hydrophobicity. As discussed by Rose (15), maxima in the hydrophobicity profile may correspond to regions in the amino acid sequence that are involved in tertiary interactions, presumably at the interior of the protein. These interactions have been suggested to be important in stabilizing domain structure. Similarly, three new turn regions are predicted to exist, one of which is predicted to be in the β-turn configuration. Hence, these changes in the primary sequence are likely to have a significant effect on the tertiary structure of the varl protein.

Remarkably, in spite of such large predicted structural differences resulting from the additional amino acids encoded by the optional GC cluster, we have not observed any differences in mitochondrial protein synthesis or the ability of cells to grow on nonfermentable carbon sources that we can attribute to this polymorphism. Thus, at least for these parameters, it would appear that the varl protein is functionally indifferent to the additional 16 amino acids encoded by the optional GC cluster.

That the varl protein is essential to the maintenance of mitochondrial respiratory function is evident from the fact that mutations mapping within the coding region of the varl gene (27), and which affect the synthesis of varl mRNA (28), result in low levels of mitochondrial protein synthesis and a mit ‘phage’ phenotype. In addition, inhibition of mitochondrial protein synthesis in vivo, which has no effect on the assembly of the large mitochondrial ribosomal subunits, blocks assembly of the 38 S subunit presumably due to inhibition of varl synthesis (2).

Origins of varl—From both primary and secondary structural considerations, the varl protein is significantly different from most other ribosomal proteins for which such information is available. The overall base composition and organization of the varl coding sequences, however, are remarkably similar to intergenic yeast mitochondrial ‘spacer’ DNA; that is, it is a very A + T-rich DNA sequence (70% A + T) interspersed with GC clusters. Moreover, at least one GC cluster has been identified at the 3’ noncoding region of the ofl gene (29) that is identical in sequence to the varl GC clusters. Interestingly, all three of these GC clusters are flanked, 11 base pairs on each side, by the identical pure A + T sequence, AATAATAAAA-GC cluster-AATATTAATA (4, 7), suggesting that these GC clusters may have been ‘targeted’ to their present location. In other words, the presence of GC clusters within the varl gene may be a reflection of the apparent mobility of these elements, which are known to be extensively polymorphic among yeast mitochondrial genomes (30). If GC clusters are indeed some type of mobile element, then obviously only those genes in which insertions did not inactivate the product or result in some selective disadvantage would prevent the presence of a GC cluster to be tolerated. This appears now to be the case for varl and the genes encoding the 21 S and 15 S RNA genes (31–33). That the varl gene can undergo several major modifications without apparent loss of function of its product suggests that the function of the varl protein is limited to relatively specific regions of the protein.

In addition to the optional GC cluster, varl is represented by another polymorphic class characterized by a variable number of asparagine codons (AAT) within (AAT)n, repeats. These are found in two regions of the gene called b1 and b2 (4, 6). Besides these, there is a tract of (AAT)n that separates two α-helical domains in the NH2-terminal part of the protein. No polymorphisms have been associated with this tract. Interestingly, (AAT)n tracts are extensively represented in the A + T-rich intergenic spaces of the yeast mitochondrial genome. Thus, like the postulated role of introns in gene evolution (34), these sequences may have served as sites of recombination between independent domains, bringing these domains together to generate a new function. However, unlike ‘exon shuffling’ (34), these putative recombinational joints in varl are expressed in the final gene product.

This possibility for the origin of varl would be in contrast to the generally accepted view that present day mitochondrial genes derived from an endosymbiont genome that was supplied to emerging eukaryotic cell functions that conferred some selective advantage. Of the genes that entered as part of the endosymbiont, it is believed that most probably duplicated existing functions and were lost, while others were either transferred to the nucleus, or remained in the mitochondrial genome. The latter ‘remnants’ and any ‘new’ sequences that subsequently evolved would then constitute the present day mitochondrial gene complement. The yeast mitochondrial varl gene may then be such a new sequence that emerged later in the evolution of the mitochondrial genome.
REFERENCES


---

Translation of var1 GC Clusters

---