The EMBO Journal vol.3 no.9 pp.2129–2136, 1984

The mitochondrial genome of the fission yeast *Schizosaccharomyces pombe*: highly homologous introns are inserted at the same position of the otherwise less conserved *cox1* genes in *Schizosaccharomyces pombe* and *Aspergillus nidulans*

B. Franz Lang

Institut für Genetik und Mikrobiologie der Universität München, Maria-Ward-Str. 1a, D-8000 München 19, FRG

Communicated by R.W. Davies

The DNA sequence of the second intron in the mitochondrial gene for subunit 1 of cytochrome oxidase (*cox1*), and the 3′ part of the structural gene have been determined in *Schizosaccharomyces pombe*. Comparing the presumptive amino acid sequence of the 3′ regions of the *cox1* genes in fungi reveals similarly large evolutionary distances between *Aspergillus nidulans, Saccharomyces cerevisiae* and *S. pombe*. The comparison of exon sequences also reveals a stretch of only low homology and of general size variation among the fungal and mammalian genes, close to the 3′ ends of the *cox1* genes. The second intron in the *cox1* gene of *S. pombe* contains an open reading frame, which is contiguous with the upstream exon and displays all characteristics common to class I introns. Three findings suggest a recent horizontal gene transfer of this intron from an *Aspergillus* type fungus to *S. pombe*. (i) The intron is inserted at exactly the same position of the *cox1* gene, where an intron is also found in *A. nidulans*. (ii) Both introns contain the highest amino acid homology between the intronic unassigned reading frames of all fungi identified so far (70% identity over a stretch of 253 amino acids). However, in the most homologous region, a GC-rich sequence is inserted in the *A. nidulans* intron, flanked by two direct repeats of 5 bp. The 37-bp insert plus 5 bp of direct repeat amounts to an extra 42 bp in the *A. nidulans* intron. (iii) TGA codons are the preferred tryptophan codons compared with TGG in all mitochondrial protein coding sequences of fungi and mammalia. In contrast, in structural genes of *S. pombe* exclusively TGG is used as tryptophan codon, whereas two TGA codons occur in the open reading frame of the second *cox1* intron.

Key words: Aspergillus nidulans/cox1 gene/DNA sequence/horizontal gene transfer/mitochondrial introns

Introduction

The mitochondrial genomes in all fungi studied so far contain genes interrupted by a variable number of long intervening sequences (for reviews, see Dujon, 1983; Grivell, 1983). Most of the introns have the coding capacity for one or other of the two types of conserved intronic unassigned reading frames (urfs) (e.g., Bonitz *et al.*, 1980; Davies *et al.*, 1983; Hensgens *et al.*, 1983; Helmer-Citterich *et al.*, 1983). Even in the small mitochondrial genomes of *Torulopsis glabrata* (Clark-Walker and Sriprakash, 1983) or *Schizosaccharomyces pombe* (Lang *et al.*, 1983; Lang and Wolf, 1984; Zimmer *et al.*, 1984) a number of introns was detected.

The reason why introns occur in fungal mitochondrial genes is still obscure. In species such as *Saccharomyces cerevisiae*, with its large mitochondrial DNA (mtDNA) and its long stretches of non-coding intergenic 'spacers' (Goursot *et al.*, 1982), introns might be readily tolerated. However, it is puzzling in fungi, which have an apparent preference for small mitochondrial genomes (e.g., the *S. pombe* strains or *T. glabrata*). In the *S. pombe* strain, analyzed here, one intron was found in *cob* and two introns in the *cox1* gene. Thus, ~25% of this small genome is occupied by introns (Lang *et al.*, 1983; Lang and Wolf, 1984). The analysis of the *S. pombe* introns and the construction of strains without them (possibly resulting in a 14-kbp mtDNA, which is even less than that of the small mtDNAs in mammalia) may give an answer to the question of why they are so strictly maintained in many species and whether or not they offer a selective advantage for the cell.

Provided that introns are involved in the control of gene expression, their presence does not seem to be gene-specific. It has been shown that different fungi have quite different sets of mosaic genes and, within a species, variant forms have been described in some cases with optional introns. For example, no *cob* intron is present in a *S. pombe* variant strain (Zimmer *et al.*, 1984) or in *T. glabrata* (Clark-Walker and Sriprakash, 1983). In *Neurospora crassa*, the *cox1* gene is not interrupted by introns (Burger *et al.*, 1982). In *S. cerevisiae* it is possible to delete introns from the *cob* gene without a dramatic effect on the expression of cytochrome *b* (Gargouri *et al.*, 1983). If introns play any role in gene expression or regulation, it should be rather limited in *S. cerevisiae*. The preference for insertion of introns in *cob* and *cox1* may only reflect that these genes have long coding regions.

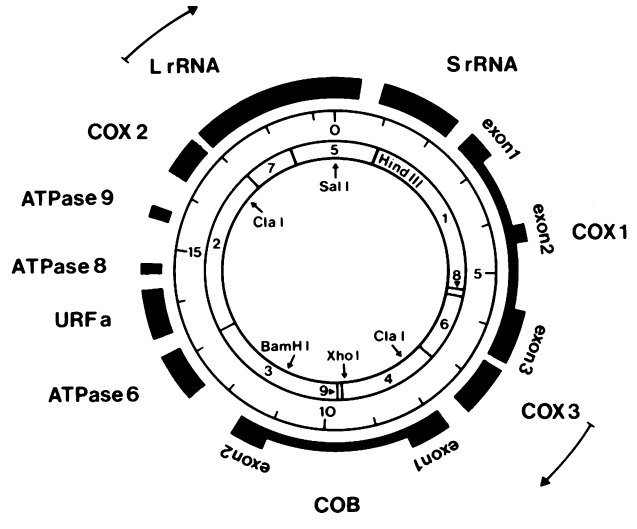

Fig. 1. Localization of the *cox1* gene on the circular map of the *S. pombe* mtDNA. *Cox1, cox2* and *cox3* refer to the genes coding for the three largest cytochrome oxidase subunits, *cob* to the apocytochrome *b* gene and urf*a* to a gene specific for *S. pombe* mtDNA (Lang *et al.*, 1983). The localization of genes is taken from Lang and Wolf (1984). Only minor corrections were made on the basis of the known DNA sequence of the *cox1* gene. The direction of transcription is indicated by the two arrows.

© IRL Press Limited, Oxford, England.

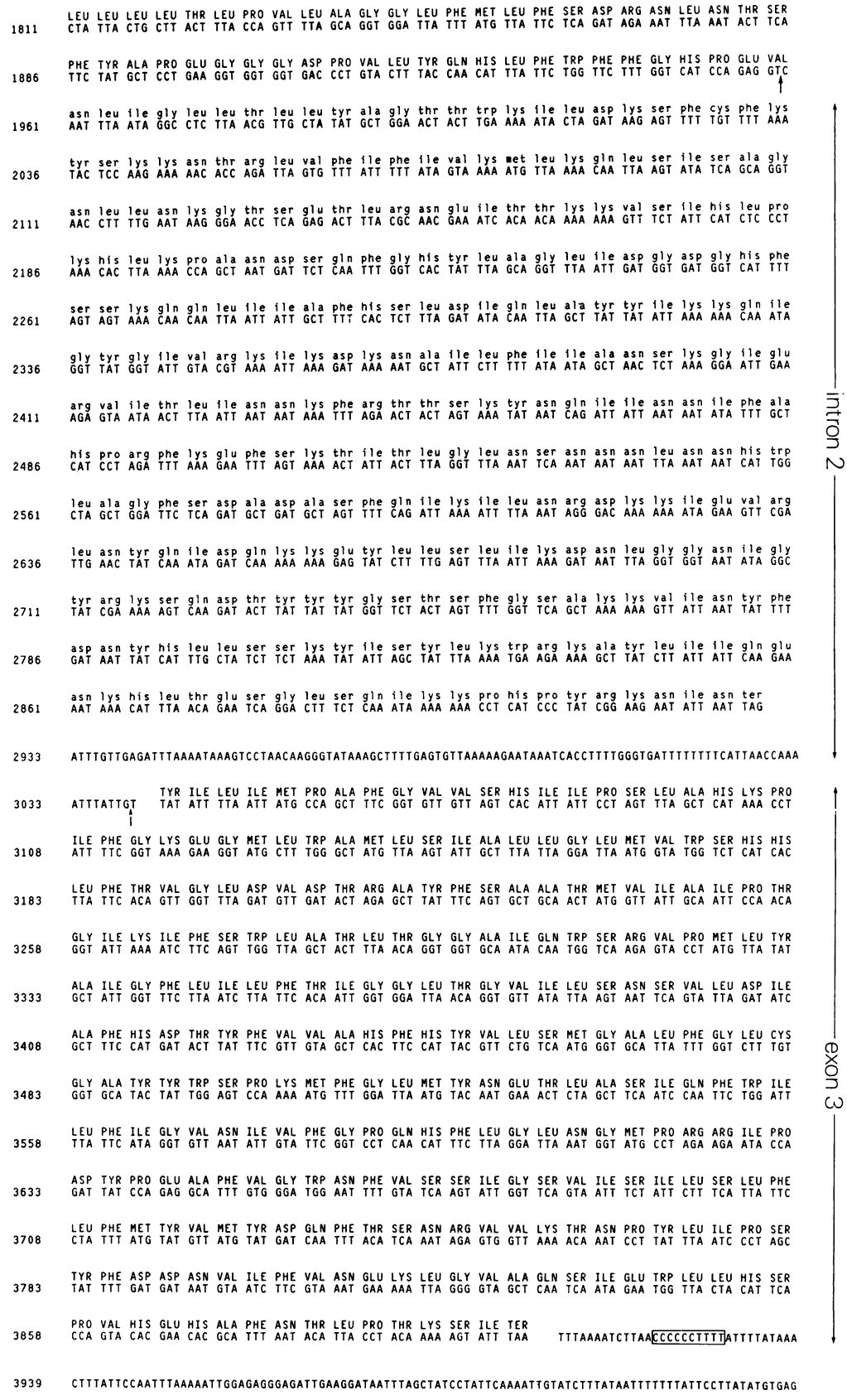

Fig. 2. Nucleotide sequence of the 3′ part of the *S. pombe cox1* gene, including the second intron. The sequence of the second half of the *cox1* gene is shown. Numbering of nucleotides starts at a methionine codon close to the putative amino terminus of the mature subunit 1. This methionine codon and that of the *N. crassa* gene at position −2 (Burger *et al.*, 1982) are homologous in position. A translation of codons as in *N. crassa* genes was assumed, based on comparisons of the codon usage in highly conserved regions between *N. crassa* and *S. pombe* and on the fact that the usage of most codons in *N. crassa* has been directly proven by amino acid sequencing (Burger *et al.*, 1982). The putative splice points of the intron are indicated by arrows. The amino acids of the intronic urf are written in small letters, that of the exons in capital letters. The oligo(TC) cluster at the 3′ end of the gene is boxed.

Table I. Codon usage in the intronic *urf* of the second *cox1* intron of *S. pombe* and in part of the structural gene

Phe	UUU	15	(10)	Ser	UCU	8	(2)	Tyr	UAU	18	(11)	Cys	UGU	1	(1)
	UUC	1	(20)		UCC	1	(−)		UAC	1	(4)		UGC	–	(−)
Leu	UUA	21	(33)		UCA	6	(12)	Ter	UAA	–	(1)	Trp	UGA	2	(−)
	UUG	5	(−)		UCG	–	(−)		UAG	1	(−)		UGG	1	(9)
Leu	CUU	5	(5)	Pro	CCU	3	(10)	His	CAU	7	(8)	Arg	CGU	1	(−)
	CUC	2	(−)		CCC	1	(−)		CAC	3	(5)		CGC	1	(−)
	CUA	4	(4)		CCA	1	(8)	Gln	CAA	11	(6)		CGA	2	(−)
	CUG	–	(2)		CCG	–	(−)		CAG	2	(−)		CGG	1	(−)
Ile	AUU	22	(17)	Thr	ACU	10	(7)	Asn	AAU	23	(12)	Ser	AGU	10	(9)
	AUC	1	(6)		ACC	2	(−)		AAC	5	(−)		AGC	1	(1)
	AUA	14	(5)		ACA	3	(9)	Lys	AAA	35	(7)	Arg	AGA	5	(6)
Met	AUG	1	(13)		ACG	1	(−)		AAG	4	(−)		AGG	1	(−)
Val	GUU	3	(13)	Ala	GCU	12	(13)	Asp	GAU	12	(9)	Gly	GGU	12	(22)
	GUC	–	(−)		GCC	–	(−)		GAC	1	(1)		GGC	2	(−)
	GUA	3	(12)		GCA	2	(8)	Glu	GAA	6	(6)		GGA	5	(6)
	GUG	1	(2)		GCG	–	(−)		GAG	2	(2)		GGG	–	(1)

For an evaluation, the sequence, shown in Figure 2 was used. The codons occurring in the structural gene parts are written in brackets.

The spontaneous deletion of entire introns from a mosaic gene in *S. cerevisiae* occurs with a rate close to that of spontaneous back-mutation of point mutations ($\sim 10^{-8}$; Gargouri *et al.*, 1983; Labouesse and Slonimski, 1983). One might therefore expect that introns can be frequently lost during evolution. If so, how do intron sequences manage to survive and to propagate themselves among the fungal mitochondrial genomes?

There are two experimental approaches to solve these questions. One is to study intron variations of strains belonging to the same species. The other, which we followed, is to look for functionally related and highly homologous introns in different species. I will report here on the second *cox1* intron in *S. pombe* because it is unexpectedly homologous with an *A. nidulans cox1* intron (compare Waring *et al.*, accompanying paper). However, the analysis of the carboxy-terminal region of the structural gene displays size variation and low homology compared with *A. nidulans* and other fungal species (Burger and Werner, 1983; Bonitz *et al.*, 1980; de Jonge and de Vries, 1983). This demonstrates the evolutionary distance between structural genes of *S. pombe* and *A. nidulans*, in contrast to the high homology of their *cox1* introns.

Results and Discussion

Structure of the cox1 gene and DNA sequence of its second half, including a class I intron

Figure 1 shows a circular map of the *S. pombe* mtDNA, indicating the position of the *cox1* gene. Figure 2 presents the DNA sequence of the 3′ part of the *cox1* gene region, including a part of exon 2, intron 2 and exon 3. The open reading frame of the intronic urf is contiguous with the upstream exon 2 and stops with a TAG codon. The open reading frame of the structural gene (end of exon 3) stops with a TAA codon and is followed by an oligo(TC) cluster, characteristic for *S. pombe* 3′ ends of protein coding genes. This sequence was proposed to act as a processing signal (Lang *et al.*, 1983). The exon/intron boundaries, indicated in Figure 2 were deduced from a comparison with the structural gene sequence of other species. As the intron is inserted in a highly conserved domain of *cox1* (cf. Figures 5 and 6), and following the rule that the last base of the upstream exon is a pyrimidine, mostly a T, and the last base of the intron is a G (Davies *et al.*, 1983), there was only one possible insertion point.

TGA codons are found in the intronic urf but not in the structural gene

A remarkable difference between codon usage in the exon and intron sequence can be seen in Table I. In addition to unusual codons with G or C at the third position, frequently found in introns but rarely in exon sequences, two TGA codons are present in the intronic urf. (The same was found in the *cob* intron and in urf*a* of *S. pombe*, Lang *et al.*, in preparation.) Within the *cox1* exons however, as well as in all other *S. pombe* mitochondrial structural genes, no TGA codon is found but exclusively TGG (Lang *et al.*, 1983 and unpublished results). It is now generally accepted for all mammalian and fungal mitochondrial genomes studied so far that TGA codes for tryptophan rather than for a stop (Fox, 1979; Macino and Tzagoloff, 1979; Anderson *et al.*, 1982).

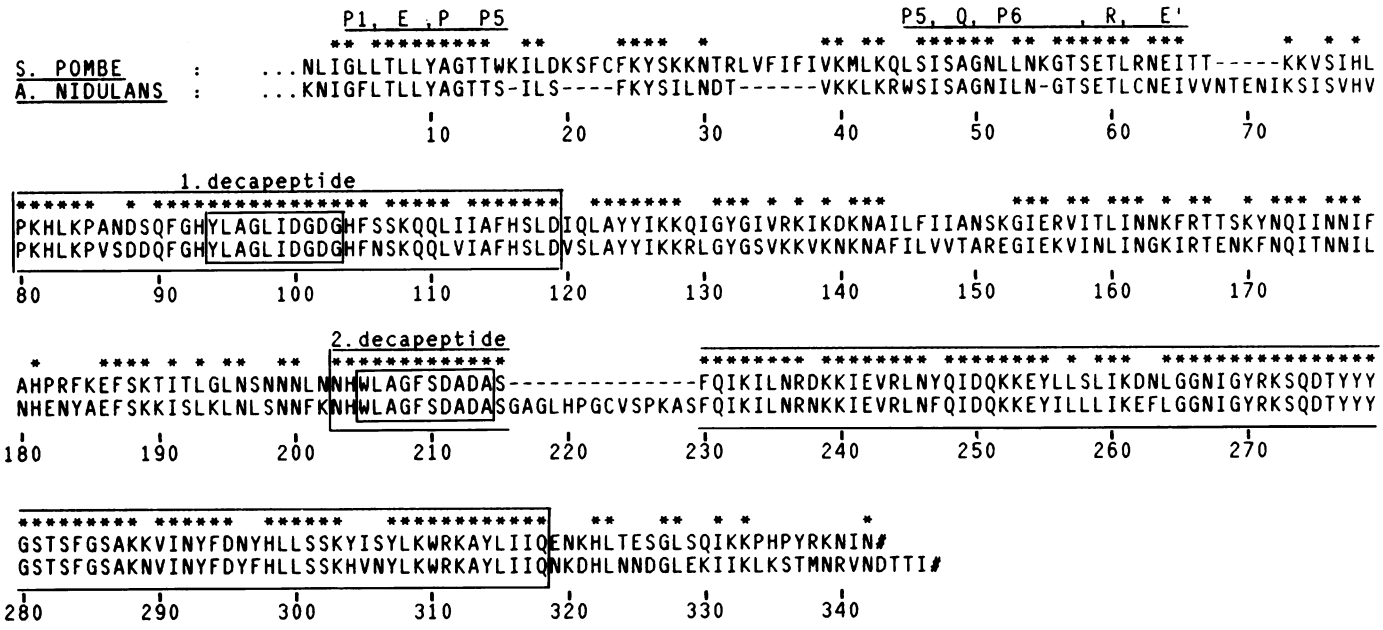

Fig. 3. Comparison of the amino acid sequences of the intronic urfs of *S. pombe cox1*, intron 2 and *A. nidulans cox1* intron 3. The sequence data of the *A. nidulans* intron are taken from Waring *et al.* (accompanying paper). Identical amino acids in this alignment are indicated by asterisks. The two identical decapeptide sequences and their conserved flanking regions are shown by boxing of respective sequences. Insertions and deletions of amino acids (–) have been introduced for an optimal alignment. Only amino acids of both intronic urfs, starting after the 5′ splice points are shown (stop codons: #). The nucleotide consensus sequences (E, P, Q, R, E′) and helical regions (P1, P5 and P6) contributing to the RNA secondary structures in this region of the intron (see Figure 4) are also indicated.

As there is an open reading frame in the *S. pombe* intron, which is homologous with the intronic urf in *A. nidulans*, the TGA codons in *S. pombe* also should be translated as tryptophan (e.g., the tryptophan codon in the second decapeptide sequence at position 205, in Figure 3).

The difference in codon usage may indicate a dual evolutionary origin of the structural genes and the second *cox1* intron, the *cob* intron and urf*a* in *S. pombe*. An import of these sequences from the nucleus is unlikely because of their relatively high AT content which is characteristic for mitochondrial sequences and because of the occurrence of TGA 'stop' codons. It would be more rational to assume that these 'TGA-type' sequences originate from mtDNA of other species.

However, what is the reason then that TGA codons are tolerated in intronic urfs, and how is this compatible with the properties of the mitochondrial tRNA-trp in *S. pombe*? The anticodon of tRNA-trp is in fact complementary to the TGG codon (5′...CCA...3′; Lang *et al.*, 1983) and pairing of the TGA codon to the tRNA will be much weaker (C-A pair in the wobble position) than that of the TGG codon. Therefore, one should expect a low efficiency of translation of TGA codon – if it is at all possible. A low efficiency of translation of intronic urfs would be acceptable for the cell, because they need not be translated as efficiently as structural genes. 'Maturases' (Lazowska *et al.*, 1980), the translation products of several *S. cerevisiae* introns, are only required at very low concentrations and are not detectable in wild-type strains (Bechmann *et al.*, 1981; Weiss-Brummer *et al.*, 1982; Dujardin *et al.*, 1982; Anziano *et al.*, 1982).

Further arguments in favour of a foreign origin of the second *cox1* intron of *S. pombe* come from a comparison with the third intron in the same gene of *A. nidulans* and are discussed below.

The second intron of the cox1 gene of S. pombe and the third intron in the same gene of A. nidulans are highly homologous

A systematic comparison of various class I introns with the second *S. pombe cox1* intron revealed that the latter clearly belongs to the same class because of: (i) the conservation of the two decapeptide sequences and some other minor amino acid homologies; (ii) the highly charged and basic character of the complete amino acid sequence; and (iii) the conservation of consensus sequences and the secondary structure of the RNA (Bonitz *et al.*, 1980; Waring *et al.*, 1982; Hensgens *et al.*, 1983; Davies *et al.*, 1982, 1983; Michel *et al.*, 1982). Even when comparing introns within the same species, amino acid homology between most intronic urfs of class I is restricted to the two conserved decapeptide sequences (Waring *et al.*, 1982; Davies *et al.*, 1983; Hensgens *et al.*, 1983). Practically, these homologies are in most cases only visible after an extensive computer evaluation. In contrast, in the region from the first decapeptide to the end of the open reading frame, the amino acid identity of the second *S. pombe* intron of the *cox1* gene and the third one in *A. nidulans* is 70%. 60%, the highest homology among 19 introns, was observed between introns of the same species, namely the *cob* intron 4 and *cox1* intron 4 of *S. cerevisiae*. Between the introns of *S. pombe* and *A. nidulans* a stretch of 40 triplets, including the first decapeptide sequence and of 102 triplets in the region of the second decapeptide (except an insertion of 14 triplets in *A. nidulans*) display 88% identical amino acid positions (Table II). Because of the perfect alignment in this region the extra sequence of *A. nidulans* can be defined as a 37-bp insert flanked by two 5-bp direct repeats. The repeats are reminiscent of those found at the borders of some insertion elements of *S. cerevisiae* (e.g., Farabaugh and

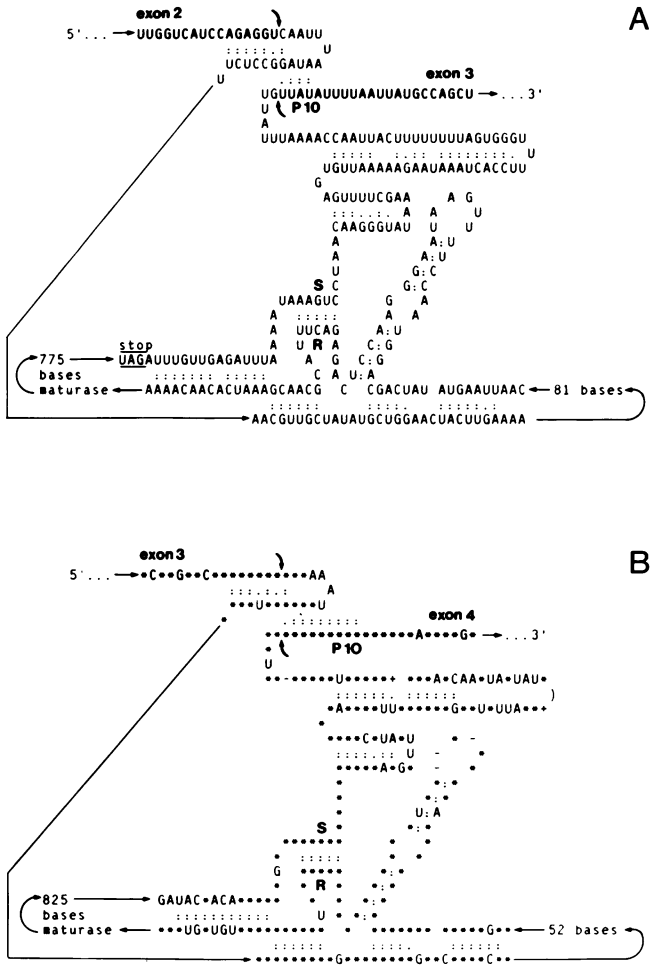

Fig. 4. Comparison of RNA secondary structures of the two highly homologous *cox1* introns of *S. pombe* and *A. nidulans*. **(A)** The RNA secondary structure of the second *S. pombe* intron was constructed according to the model of Davies *et al.* (1982). Arrows point to the two splice points. **(B)** In the *A. nidulans* secondary structure, nucleotides identical to **A** are indicated by asterisks and only nucleotide exchanges, an insertion (+) or deletions (−) are shown, in order to visualize compensatory exchanges within helical regions or stretches without pronounced conservation (the data of the secondary structure of the *A. nidulans* intron are taken from Waring *et al.*, accompanying paper). The position of the highly conserved R and S consensus sequences and the possible IGS pairing (denoted as P10) are depicted in bold letters.

Fink, 1980; Gafner and Philippson, 1980). The high GC content of the insert also suggests its non-mitochondrial, possibly nuclear origin (compare Waring *et al.*, preceding paper).

Type I introns have been detected at many different positions within various structural genes, but the highly homologous *S. pombe* and *A. nidulans cox1* introns are inserted exactly at the same position in the gene. Both the high homology of introns and the identical insertion point suggest a very close relationship between them. An analysis of the conservation of the *cox1* exon sequence will show that this is only the case for the introns (see Figure 6).

Two further examples have been reported, where introns are at the same position of genes. The sequence of these introns are either almost completely different (e.g. introns in the large rRNA of *N. crassa* and *S. cerevisiae*, Burke and RajBhandary, 1982; Dujon, 1980) or the homologies are less pronounced (*cob* introns of *S. cerevisiae* and *A. nidulans*; Waring *et al.*, 1981; Lazowska *et al.*, 1981).

Table II. Homologies of amino acids in the intronic urfs of the highly conserved *cox1* introns of *S. pombe* and *A. nidulans*

	Amino acid positions	Number of amino acids minus insert (216–229)	% homology
'LAGLI'-region (1. decapeptide)	80–119	40	88
'DADG'-region (2. decapeptide)	203–318	102	88
Complete reading frame, except region containing secondary structure	80–346	253	70

The calculations and the positions of amino acids are based upon the alignment shown in Figure 3. The homology of the first 88 amino acids has not been calculated because of constraint by the RNA secondary structures in this region rather than by amino acid conservation.

It is known from the analysis of various class I introns that certain regions of the intron RNAs can be folded into highly conserved secondary structures (Michel *et al.*, 1982; Davies *et al.*, 1982). In regions of the type I introns necessary to build the RNA secondary structures there is no apparent constraint to conserve amino acid sequences. The two blocks of conserved amino acids found in this region (Figure 3) reflect a high nucleotide conservation. From the strong conservation it is clear that the secondary structures of both intron RNAs are almost identical as shown in Figure 4. Remarkably, the pairing of an internal guide RNA sequence (IGS) (Davies *et al.*, 1982) is much weaker in *S. pombe* than in the homologous *A. nidulans* intron, indicating its minor role (if any) in splicing of *S. pombe* introns.

In the first *cox1* intron of *S. pombe*, which is also of class I, such an IGS is obviously dispensible for splicing and its function seems to be replaced by another mechanism (Lang *et al.*, in preparation).

The exon regions flanking the intron are of extraordinarily high homology

As shown in Figures 5 and 6, the second intron of the *cox1* gene of *S. pombe* and the third intron of *A. nidulans* lie at exactly the same position and within a highly conserved domain of the structural gene. An exon region of 60 bp, upstream of the introns, covers most of this highly conserved domain with a nucleotide homology of 82% and about the same percentage of identical amino acids (80%). Also 60 bp of the region downstream of the intron are very homologous. However, in this stretch there is only 45% amino acid homology, but 70% nucleotide homology. The relatively low amino acid homology of this stretch is in agreement with the generally low amino acid conservation in other species (see Figure 6). The much higher nucleotide homology may then result from the above postulated intron transfer: if the flanking exon sequences were co-transferred with the intron, subsequent nucleotide exchanges in this less conserved region could decrease the percentage of the amino acid homology to a higher degree than that of the nucleotide homology.

A comparison of the 3' part of various cox1 structural genes

The assumption that a horizontal intron transfer could have happened from an *Aspergillus*-type *cox1* intron to *S. pombe* was based on the observation that their structural genes are less conserved than the introns. A comparison of two structural genes has been already made for the ATPase subunit 6

gene (Lang *et al.*, 1983), indicating the considerable evolutionary distance between the *S. pombe* gene and any other fungal gene. A similar result is obtained when comparing the putative amino acid sequences of the carboxy-terminal region of *cox1* from *S. pombe* with that of different species (Bonitz *et al.*, 1980; Anderson *et al.*, 1982; Waring *et al.*, preceding paper; Burger and Werner, 1983). Blocks of homologous regions are interspersed by sequences of very low homology including multiple insertions or deletions (Figure 6). Especially for the alignment of fungal with human genes, amino acid

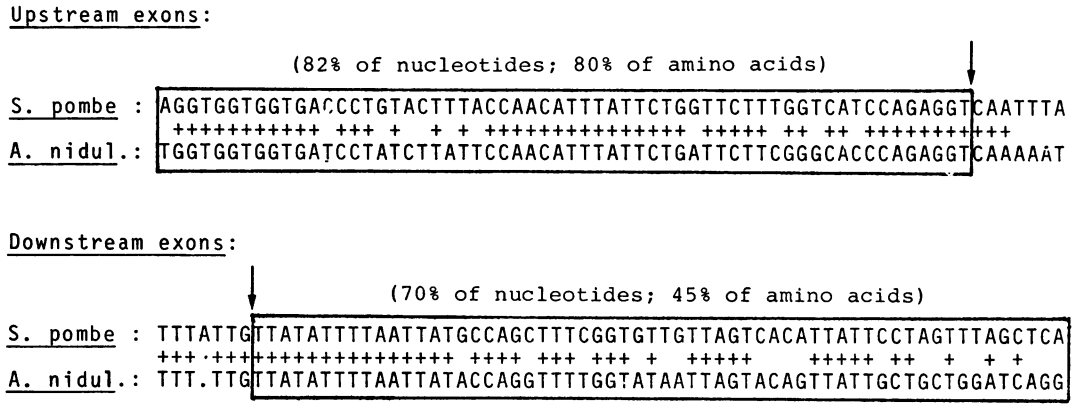

Fig. 5. Comparison of the nucleotide sequences of the *cox1* structural genes of *S. pombe* and *A. nidulans*, flanking the homologous introns. 60 nucleotides of the flanking exon regions (boxed) were aligned and the nucleotide homologies within the regions calculated. The respective amino acid homologies are given for comparison.

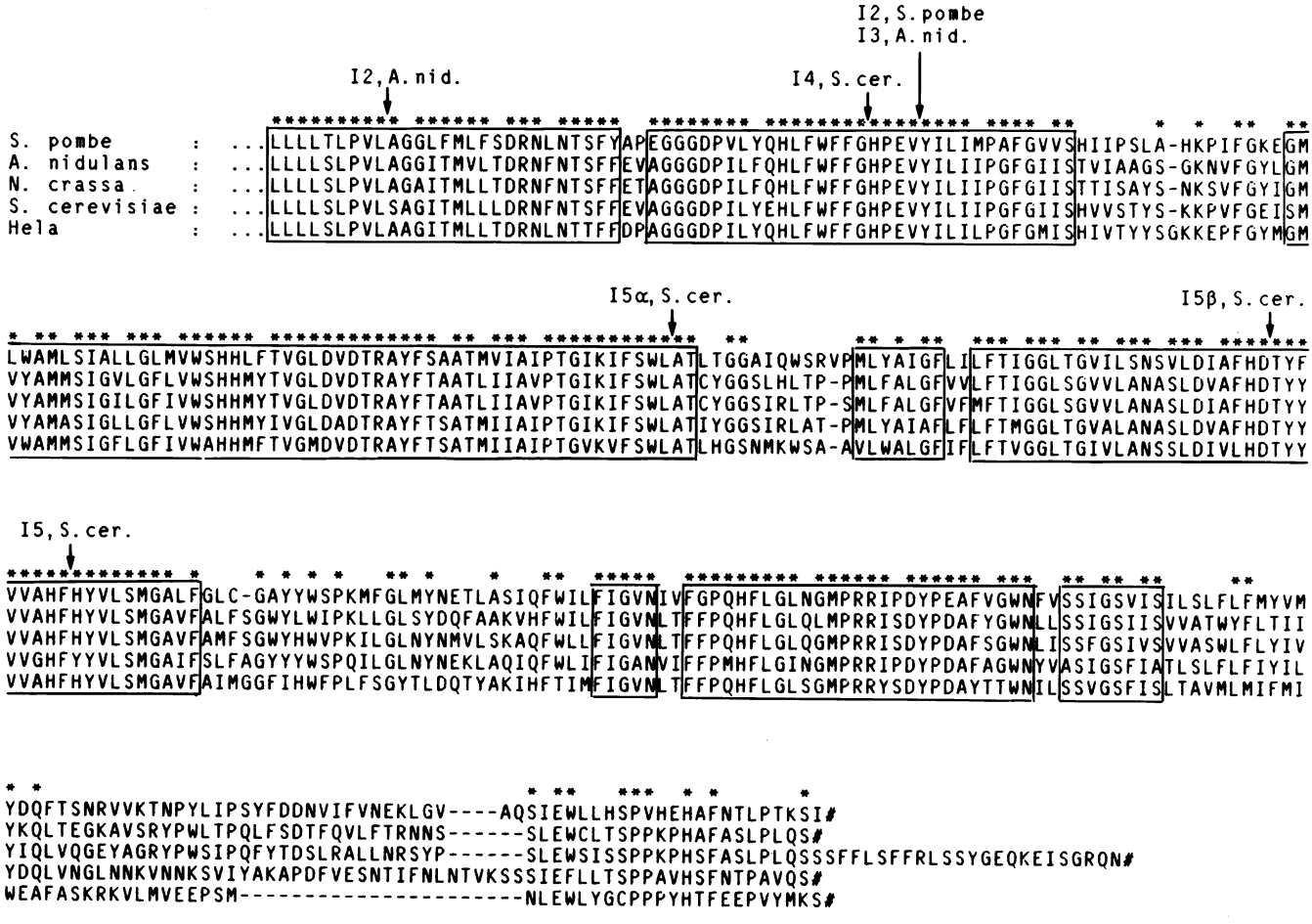

Fig. 6. Comparison of the putative amino acid sequences of *cox1* in various species. The amino acid sequences, shown in this alignment, were translated from respective sequences of the genes. The comparison includes the amino acids, homologous from position 198 to the carboxy-terminal of the mature *N. crassa* subunit (Burger and Werner, 1982). Other data are taken from Waring *et al.* (accompanying paper), Anderson *et al.* (1982), Bonitz *et al.* (1980) and Hensgens *et al.* (1983). In order to distinguish blocks of homologous regions, those positions were marked with an asterisk, having four or five identical amino acids. Blocks of homology are boxed, containing a minimum of five asterisks, interspersed by only single mismatches. The position of intron insertions are indicated by arrows. (Stop codons: #.)

Table III. Percentage of amino acid homology of *cox1* exons

	A. nidulans	*N. crassa*	*S. cerevisiae*	HeLa
S. pombe	58	57	60	53
A. nidulans	–	88	67	58
N. crassa	–	–	64	57
S. cerevisiae	–	–	–	54

The data are taken from the alignment shown in Figure 6.

insertions or deletions have to be assumed in the low homology regions. However, the small variation in length of the fungal and mammalian genes in this region does not support the idea of further small introns in the fungi as previously discussed (Bonitz *et al.*, 1980; de Jonge and de Vries, 1983). The size variations fall into the region of very low amino acid homology of the subunit shown to exist in the mature polypeptide in *N. crassa* by amino acid sequencing (Burger *et al.*, 1982). Thus it is more likely that the carboxy-terminal portion of the *cox1* gene is generally variable. The most significant deviation from all other fungal genes is found in the *N. crassa* gene, which has a 3′ extension of the reading frame of ~24 codons.

The study of the amino acid homology of the *cox1* exon regions (Figure 6 and Table III), leads to the same conclusion that was drawn from a comparison of the ATPase subunit 6 genes (Lang *et al.*, 1983). In terms of evolution, *S. pombe* mitochondrial sequences are far distant from any other equivalent fungal gene – with the exception of the second *cox1* intron.

Concluding remarks

Do the two *cox1* introns of *S. pombe* and *A. nidulans* result from a recent horizontal gene transfer? Three arguments can be brought forward, which suggest such a horizontal intron transfer: (i) the comparison of structural genes of *A. nidulans* and *S. pombe* clearly demonstrates that both fungi are not closely related, whereas the *cox1* introns are of an exceptionally high homology. (ii) TGA codons, preferred as a tryptophan codon in the mitochondrial genes of all other fungi and of mammalia are found only in two out of three *S. pombe* intronic urfs and in urf*a*, but not in the ubiquitous mitochondrial genes. The unique use of TGA codons in the introns would then indicate the direction of transfer: from an *Aspergillus*-type species to *S. pombe*. (iii) Exon sequences, directly adjacent to the highly homologous *cox1* introns in *S. pombe* or *A. nidulans*, respectively, have a remarkably high nucleotide homology.

Although introns are tolerated at many sites of various genes, one property seems to be common to most regions of intron insertion: they occur in highly conserved regions. For introns in the *cox1* gene, a relatively low homology is only found downstream of the *S. cerevisiae cox1* intron 5α. Two reasons may account for intron insertions in highly conserved regions. (i) Mutations in the intron can be eliminated in two ways: either by deletion of the intron or by reversion or suppression of such a mutation. The different possibilities of more or less precise intron deletions are considerably reduced if the intron lies in a highly conserved region of a gene. (ii) These introns, maintained in regions of high homology can be much more easily distributed *via* horizontal gene transfer, because the exon regions adjacent to the intron would provide the homology necessary for recombination events.

Materials and methods

Cloning

The DNA sequence has been obtained from clones pFM 111, pFM 261 and pFM 26, containing the first *Eco*RI fragment and the sixth and eighth *Hind*III fragments of the *S. pombe* mtDNA. A detailed description of plasmids has been published (Lang and Wolf, 1984). Part of the sequence has been also obtained from clone pFM 111-1, containing *Hind*III fragment 1 of the *S. pombe* mtDNA, which was subcloned from pFM 111 and was kindly provided by H. Trinkl.

DNA sequence analysis

The sequencing protocols of Maxam and Gilbert (1980) were followed with slight modifications: (i) DNA was recovered from gels by electroelution in dialysis bags, (ii) in addition to the 5′ end labelling with T4 polynucleotide kinase after dephosphorylation, 3′ ends were labelled using either DNA polymerase I and αdNTPs in fill-in reactions or terminal transferase and cordycepin (Roychoudhury *et al.*, 1979), (iii) the five reactions (A + G, G, A + C, T + C and C) were run on 0.3 mm thin sequencing gels and handled according to Garoff and Ansorge (1981). Both DNA strands were sequenced, using overlapping fragments.

Acknowledgements

The author thanks R.B. Waring, R.W. Davies and C. Scazzocchio for providing sequence data prior to publication. I am especially indebted to F. Kaudewitz for advice and the continuous support of the work. The helpful and critical comments of G. Burger, J. Kreike, T. Pillar, K. Wolf, R.J. Schweyen and F. Ahne on the manuscript are acknowledged. The skilful technical assistance of E. Praetzel is much appreciated. This work was supported by the 'Deutsche Forschungsgemeinschaft'.

References

Anderson,S., Bankier,A.J., Barrell,B.G., deBruijn,H.L., Coulson,A.R., Drouin,J., Eperon,I.C., Nierlich,D.P., Roe,B.A., Sanger,F., Schreier, P.H., Smith,J.H., Staden,R. and Young,I.G. (1982) in Slonimski,P.P., Borst,P. and Attardi,G. (eds.), *Mitochondrial Genes*, Cold Spring Harbor Laboratory Press, NY, pp. 5-43.

Anziano,P.O., Hanson,D.K., Mahler,H.R. and Perlman,P.S. (1982) *Cell*, **30**, 925-932.

Bechmann,H., Haid,A., Schweyen,R.J., Mathews,S. and Kaudewitz,F. (1981) *J. Biol. Chem.*, **256**, 3525-3531.

Bonitz,S.G., Coruzzi,G., Thalenfeld,B.E., Tzagoloff,A. and Macino,G. (1980) *J. Biol. Chem.*, **255**, 11927-11941.

Burger,G., Scriven,C., Machleidt,W. and Werner,S. (1982) *EMBO J.*, **1**, 1385-1391.

Burger,G. and Werner,S. (1983) in Schweyen,R.J., Wolf,K. and Kaudewitz, F. (eds.), *Mitochondria 1983*, Walter de Gruyter, Berlin-NY, pp. 331-342.

Burke,J.M. and RajBhandary,U.L. (1982) *Cell*, **31**, 509-520.

Clark-Walker,G.D. and Sriprakash,K.S. (1983) *EMBO J.*, **2**, 1465-1472.

Davies,R.W., Waring,R.B., Brown,T.A. and Scazzocchio,C. (1983) in Schweyen,R.J., Wolf,K. and Kaudewitz,F. (eds.), *Mitochondria 1983*, Walter de Gruyter, Berlin-NY, pp. 179-189.

Davies,R.W., Waring,R.B., Ray,J.A., Brown,T.A. and Scazzocchio,C. (1982) *Nature*, **300**, 719-724.

de Jonge,J.C. and de Vries,H. (1983) *Curr. Genet.*, **7**, 21-28.

Dujardin,G., Jacq,C. and Slonimski,P.P. (1982) *Nature*, **298**, 628-632.

Dujon,B. (1980) *Cell*, **20**, 185-197.

Dujon,B. (1983) in Schweyen,R.J., Wolf,K. and Kaudewitz,F. (eds.), *Mitochondria 1983*, Walter de Gruyter, Berlin-NY, pp. 1-24.

Farabaugh,P.J. and Fink,G.R. (1980) *Nature*, **286**, 352-356.

Fox,T.D. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 6534-6538.

Gafner,J. and Philippsen,P. (1980) *Nature*, **286**, 414-418.

Gargouri,A., Lazowska,J. and Slonimski,P.P. (1983) in Schweyen,R.J., Wolf,K. and Kaudewitz,F. (eds.), *Mitochondria 1983*, Walter de Gruyter, Berlin-NY, pp. 259-268.

Garoff,H. and Ansorge,W. (1981) *Anal. Biochem.*, **115**, 450-457.

Goursot,G., Mangin,M. and Bernardi,G. (1982) *EMBO J.*, **1**, 705-711.

Grivell,L.A. (1983) in Schweyen,R.J., Wolf,K. and Kaudewitz,F. (eds.), *Mitochondria 1983*, Walter de Gruyter, Berlin-NY, pp. 25-45.

Helmer-Citterich,M.H., Morelli,G., Nelson,M.A. and Macino,G. (1983) in Schweyen,R.J., Wolf,K. and Kaudewitz,F. (eds.), *Mitochondria 1983*, Walter de Gruyter, Berlin-NY, pp. 357-369.

Hensgens,L.A.M., Bonen,L., de Haan,M., van der Horst,G. and Grivell, L.A. (1983) *Cell,* **32**, 379-389.
Labouesse,H. and Slonimski,P.P. (1983) *EMBO J.,* **2**, 269-276.
Lang,B.F., Ahne,F., Distler,S., Trinkl,H., Kaudewitz,F. and Wolf,K. (1983) in Schweyen,R.J., Wolf,K. and Kaudewitz,F. (eds.), *Mitochondria 1983*, Walter de Gruyter, Berlin-NY, pp. 313-329.
Lang,B.F. and Wolf,K. (1984) *Mol. Gen. Genet.,* in press.
Lazowska,J., Jacq,C. and Slonimski,P.P. (1980) *Cell,* **22**, 333-348.
Lazowska,J., Jacq,C. and Slonimski,P.P. (1981) *Cell,* **27**, 12-14.
Macino,G. and Tzagoloff,A. (1979) *J. Biol. Chem.,* **254**, 4617-4623.
Maxam,A.M. and Gilbert,W. (1980) *Methods Enzymol.,* **65**, 449-560.
Michel,F., Jacquier,A. and Dujon,B. (1982) *Biochimie,* **64**, 867-881.
Roychoudhury,R., Tu,C.D. and Wu,R. (1979) *Nucleic Acids Res.,* **6**, 1323-1333.
Waring,R.B., Davies,R.W., Lee,S., Grisi,E., McPhail Berks,M. and Scazzocchio,C. (1981) *Cell,* **27**, 4-11.
Waring,R.B., Davies,R.W., Scazzocchio,C. and Brown,T.A. (1982) *Proc. Natl. Acad. Sci. USA,* **79**, 6332-6336.
Weiss-Brummer,B., Rödel,G., Schweyen,R.J. and Kaudewitz,F. (1982) *Cell,* **29**, 527-536.
Zimmer,M., Lückemann,G., Lang,B.F. and Wolf,K. (1984) *Mol. Gen. Genet.,* in press.

Received on 1 June 1984