The mitochondrial genome of the thermal dimorphic fungus *Penicillium marneffei* is more closely related to those of molds than yeasts

Patrick C.Y. Woo[a,1], Hongjun Zhen[b,1], James J. Cai[a], Jun Yu[b], Susanna K.P. Lau[a], Jian Wang[b], Jade L.L. Teng[a], Samson S.Y. Wong[a], Ronald H. Tse[a], Ran Chen[b], Huanming Yang[b], Bin Liu[b,*], Kwok-yung Yuen[a,c,**]

[a] *Department of Microbiology, The University of Hong Kong, University Pathology Building, Queen Mary Hospital, Hong Kong, PR China*
[b] *Beijing Genomics Institute, Chinese Academy of Sciences, Beijing Airport Industrial Zone B-6, Beijing, PR China*
[c] *HKU-Pasteur Research Centre, Hong Kong, PR China*

Received 8 September 2003; revised 23 October 2003; accepted 29 October 2003

First published online 18 November 2003

Edited by Lev Kisselev

Abstract **We report the complete sequence of the mitochondrial genome of *Penicillium marneffei*, the first complete mitochondrial DNA sequence of a thermal dimorphic fungus. This 35 kb mitochondrial genome contains the genes encoding ATP synthase subunits 6, 8, and 9 (*atp6*, *atp8*, and *atp9*), cytochrome oxidase subunits I, II, and III (*cox1*, *cox2*, and *cox3*), apocytochrome *b* (*cob*), reduced nicotinamide adenine dinucleotide ubiquinone oxireductase subunits (*nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5*, and *nad6*), ribosomal protein of the small ribosomal subunit (*rps*), 28 tRNAs, and small and large ribosomal RNAs. Analysis of gene contents, gene orders, and gene sequences revealed that the mitochondrial genome of *P. marneffei* is more closely related to those of molds than yeasts.**

Key words: Mitochondrial genome; *Penicillium marneffei*; Dimorphic fungus

1. Introduction

Penicillium marneffei is the most important thermal dimorphic fungus causing respiratory, skin and systemic mycosis in Southeast Asia [1–6]. Discovered in 1956 in hepatic abscesses of the Chinese bamboo rat *Rhizomys sinensis*, only 18 cases of human diseases were reported (in HIV-negative patients) until 1985 [7]. The appearance of the HIV pandemic, especially in Southeast Asian countries, saw the emergence of the infection as an important opportunistic mycosis in immunocompromized patients. About 10% of AIDS patients in Hong Kong are infected with *P. marneffei* [8]. In northern Thailand, penicilliosis is the third most common indicator disease of AIDS following tuberculosis and cryptococcosis [9]. Clinically, penicilliosis manifests as a systemic febrile illness, which results from intracellular infection of the reticuloendothelial cells by the yeast phase of the fungus and the associated inflammatory response of the host.

Despite its medical importance and its unusual thermal dimorphic capability, a large part of the ecology and epidemiology of *P. marneffei* remains unknown. The natural habitat of the fungus and its exact route of transmission have not been described. Studies of this fungus at the molecular level have been limited. Only one cell wall mannoprotein gene has been characterized and successfully used in serodiagnosis and prevention of this infection [10–14]. Based on the mitochondrial and spacer rRNA, which allowed investigators to suggest a strong phylogenetic connection with *Talaromyces* species [15], a PCR/hybridization assay was designed for molecular identification of this fungus in positive cultures [16].

P. marneffei is a model organism for understanding the molecular basis of thermal dimorphism. Given its propensity to cause disease in the AIDS patients, the genome of *P. marneffei* may also provide insights to its pathogenic mechanisms and its possible interactions with the immune system. Recently, we described a random analysis of 2303 random sequence tags from the genome of *P. marneffei* [17], which has laid down the foundation for the complete genomic sequencing project of this fungus. In 2002, the complete genome sequencing project of *P. marneffei* was started, and an approximately 4× coverage of the genome, which includes a contig that contains the complete sequence of the mitochondrial genome, has been achieved. In this article, we report this complete sequence of the mitochondrial genome of *P. marneffei*, which is the first complete mitochondrial DNA sequence of a thermal dimorphic fungus. Comparison of the mitochondrial genome of *P. marneffei* and those of yeasts and molds were also performed [18–21].

2. Materials and methods

2.1. Strain and DNA preparation

P. marneffei strain PM1 was isolated from an HIV-negative patient suffering from culture-documented penicilliosis in Hong Kong. The arthroconidia ('yeast form') of PM1 was used throughout the DNA sequencing experiments. Genomic DNA, including mitochondrial DNA, was prepared from the arthroconidia grown at 37°C. A single colony of the fungus grown on Sabouraud dextrose agar at 37°C was inoculated into yeast peptone broth and incubated in a shaker at 30°C for 3 days. Cells were cooled in ice for 10 min, harvested by centrifugation at 2000×*g* for 10 min, washed twice and resuspended in ice-cold 50 mM EDTA buffer (pH 7.5). 20 mg novazym/ml was added and incubated at 37°C for 1 h followed by digestion in a mixture of 1 mg proteinase K/ml, 1% *N*-lauroylsarcosine, and 0.5 M EDTA pH

*Corresponding author. Fax: (8610)-80498676.
**Fax: +852-28551241.
E-mail addresses: liub@genomics.org.cn (B. Liu), hkumicro@hkucc.hku.hk (K.-y. Yuen).

[1] These authors contributed equally to this work.

0014-5793/03/$22.00 © 2003 Published by Elsevier B.V. on behalf of the Federation of European Biochemical Societies.
doi:10.1016/S0014-5793(03)01307-3

9.5 at 50°C for 2 h. Genomic DNA was then extracted by phenol, phenol-chloroform, and finally precipitated and washed in ethanol. After digestion with RNase A, a second ethanol precipitation was followed by washing with 70% ethanol, air-dried and dissolved in 500 μl of TE (pH 8.0).

2.2. Library construction and sequence assembly

The *P. marneffei* mitochondrial genome was sequenced as part of the *P. marneffei* genome sequencing project. A genomic DNA (including mitochondrial DNA) library was made in pUC18 carrying inserts with sizes from 3.0 to 5.0 kb. DNA inserts were prepared by physical shearing using the sonication method. Phred/Phrap/Consed software package was used for sequence assembly and quality assessment [22–24]. The complete mitochondrial DNA genome was generated from assembly of 467 successful sequence reads (100 bp at Phred value Q20 [24,25]), which corresponded to an overall mitochondrial genome coverage of about 7×.

2.3. Sequence annotation

The putative ORFs in *P. marneffei* mitochondrial DNA were denoted by using Artemis, a free sequence viewer and annotation tool, with the genetic code of mold. Genes, in which the putative ORFs locate, were functionally assigned through the BLASTP search against fungal mitochondrion encoding proteins available in the GenBank database. Introns and rRNAs were mainly identified by BLASTN pairwise comparison of *P. marneffei* mitochondrial DNA with mitochondrial DNAs of *Aspergillus nidulans*, *Neurospora crassa*, *Saccharomyces cerevisiae* (Acc. NC_001224), *Schizosaccharomyces pombe* (Acc. NC_001326), *Podospora anserina* (Acc. NC_001329), *Allomyces macrogynus* (Acc. NC_001715), *Pichia canadensis* (Acc. NC_001762), *Candida albicans* (Acc. NC_002653), *Yarrowia lipolytica* (Acc. NC_002659), and *Candida glabrata* (Acc. NC_004691) [18–21,26]. The BLASTN results were viewed through ACT, a DNA sequence comparison viewer based on Artemis, and exon and intron boundaries were adjusted manually. The tRNAs were predicted by tRNAscan-SE 1.21 (http://www.genetics.wustl.edu/eddy/tRNAscan-SE/). The core structures of the group I introns were inferred by the program CITRON.

2.4. Phylogenetic analysis

The 11 genes that encode subunits of respiratory chain complexes (*cox1*, *cox2*, *cox3*, *cob*, *nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5*, and *nad6*) and the three that encode ATPase subunits (*atp6*, *atp8*, and *atp9*) in the *P. marneffei* mitochondrial genome and the corresponding genes in 24 other fungi with completed mitochondrial genomes were used to determine the phylogenetic relationships of *P. marneffei* to the other fungi. Phylogenetic trees were constructed using unambiguously aligned portions of concatenated amino acid sequences of these 14 protein coding genes by the maximum likelihood method in the PHYLIP package, because the corresponding *nad* genes are not present in *Schizosaccharomyces japonicus*, *Schizosaccharomyces octosporus*, *S. pombe*, *C. glabrata*, *Saccharomyces castellii*, *Saccharomyces servazzii*, and *S. cerevisiae*, and the maximum likelihood method is not as sensitive to lack of sequence information as the distance methods. A total of 3462 amino acid positions were included in the analysis.

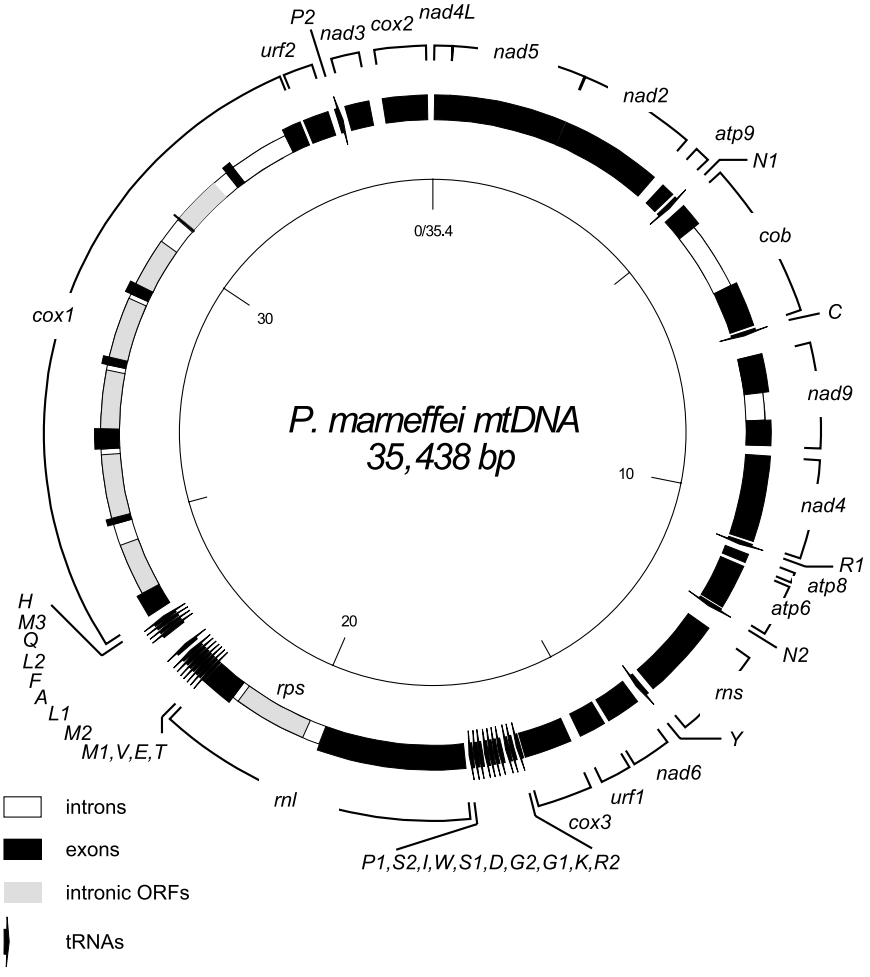

Fig. 1. Physical map of *P. marneffei* mitochondrial DNA. The map is based on an annotation of the reverse complement of Assembly 3 of the *P. marneffei* mitochondrial sequence determined by the *P. marneffei* Sequencing Project at the University of Hong Kong in collaboration with Beijing Genomics Institute of Chinese Academy of Sciences. Numbers in the inner circle are in kb. The sequence is numbered from the unique restriction enzyme *Cla*I site (AT|CGAT) (0/35.4), which is located just upstream to the *nad4L* gene and downstream to the *cox2* gene. Exons are shown in black, introns in white, and intronic ORFs in gray.

2.5. Mitochondrial DNA sequences in nuclear genome

Fragments of mitochondrial DNA sequences were searched for in the corresponding nuclear genomes in *P. marneffei*, *A. nidulans*, *N. crassa*, *S. cerevisiae*, and *S. pombe*. For each fungus, the corresponding mitochondrial DNA sequence was used as the query sequence to search against its own nuclear genome, using a published method for *S. cerevisiae* [27]. The mitochondrial and genomic DNA sequences of *A. nidulans* and *N. crassa* were downloaded from the *A. nidulans* database (http://www-genome.wi.mit.edu/annotation/fungi/aspergillus/) and *N. crassa* database (http://www-genome.wi.mit.edu/annotation/fungi/neurospora/) respectively, and those of *S. cerevisiae* and *S. pombe* were obtained from GenBank. For *P. marneffei*, a 4× coverage

Table 1
Gene content of *P. marneffei* mitochondrial genome

Genetic element	Localization (nt)	Size		Codons	
		bp	aa	Start	Stop
nad4L	26–295	270	89	ATG	TAA
nad5	295–2271	1977	658	ATG	TAA
nad2	2289–4028	1740	579	TTA	TAA
atp9	4216–4440	225	74	ATG	TAA
trna-asn1	4501–4580	80			
cob	Join: (4706–5098, 6270–7037)	2332	386	ATG	TAA
cob-i1-ORF	5099–5965	867	288	TTG[a]	TAA
trna-cys	7089–7159	71			
nad1	Join: (7532–8179, 8650–9081)	1550	359	ATA	TAA
nad4	9253–10716	1464	487	ATG	TAA
trna-arg1	10765–10835	71			
atp8	10945–11091	147	48	ATG	TAG
atp6	11158–11928	771	256	ATG	TAA
trna-asn2	11959–12029	71			
rns	12341–13721	1381			
trna-tyr	13831–13915	85			
nad6	14053–14637	585	194	ATG	TAA
URF1	14722–15177	456	151	ATG	TAA
cox3	15352–16161	810	269	ATG	TAA
trna-arg2	16190–16260	71			
trna-lys	16303–16374	72			
trna-gly1	16380–16450	71			
trna-gly2	16507–16577	71			
trna-asp	16592–16664	73			
trna-ser1	16670–16750	81			
trna-trp	16754–16824	71			
trna-ile	16856–16927	72			
trna-ser2	16929–17014	86			
trna-pro1	17021–17093	73			
rnl	Join: (17165–19688, 21361–21902)	4738			
rps	19987–21252	1266	421	ATG	TAA
trna-thr	21915–21985	71			
trna-glu	21991–22063	73			
trna-val	22067–22138	72			
trna-met1	22140–22210	71			
trna-met2	22214–22286	73			
trna-leu1	22291–22372	82			
trna-ala	22378–22449	72			
trna-phe	22518–22590	73			
trna-leu2	22900–22973	74			
trna-gln	22976–23048	73			
trna-met3	23057–23127	71			
trna-his	23132–23202	71			
cox1	join: (23339–23718, 24994–25099, 26298–26641, 27740–27875, 29012–29201, 30504–30553, 31652–31806, 32835–33159)	9821	561	ATT	TAA
cox1-i1-ORF	23720–24622	903	300	AAA[a]	TAA
cox1-i2-ORF	25100–26200	1101	366	AAA[a]	TAA
cox1-i3-ORF	26643–27647	1005	334	AAA[a]	TAA
cox1-i4-ORF	27876–28928	1053	350	TGA[a]	TAA
cox1-i5-ORF	29204–30043	840	279	TTA[a]	TAA
cox1-i6-ORF	30554–31384	831	276	ACA[a]	TAA
cox1-i7-ORF	31808–32629	821	273	AGA[a]	TAG
URF2	33223–33660	438	145	ATT	TAA
trna-pro2	33775–33862	88			
nad3	33955–34362	408	135	ATG	TAA
cox2	34591–35346	756	251	ATG	TAA

The *cob*, *nad1*, *rnl*, and *cox1* genes of *P. marneffei* are composed of two, two, two, and eight exons, respectively. The complete *cob*, *nad1*, and *cox1* ORFs and the 23S rRNA gene are obtained by joining nucleotides at the positions indicated. URF, unassigned reading frame; *cob-i1*-ORF, intronic ORF in the intron 1 of *cob* gene; *cox1-i[1-7]*-ORF, intronic ORFs in the intron 1-7 of *cox1* gene; *rns* and *rnl*, rRNA of the small and large ribosomal subunits respectively; *rps*, ribosomal protein of the small ribosomal subunit.

[a]Exact start codon could not be determined merely through sequence comparison.

of genomic DNA sequences was generated by our own whole genome sequencing project.

2.6. Nucleotide sequence accession number

The mitochondrial genome sequence of *P. marneffei* has been lodged within the GenBank sequence database under accession no. AY347307.

3. Results and discussion

3.1. Gene content and genome organization

The mitochondrial DNA of *P. marneffei* is a circular DNA molecule of 35438 bp (Fig. 1). The overall G+C content is 25%, with 24% in protein-coding genes. The genome encodes 28 tRNAs, the small and the large subunit rRNAs, the ribosomal protein of the small ribosomal subunit, 11 genes encoding subunits of respiratory chain complexes, and the three ATPase subunits (Table 1).

All genes are encoded by the same DNA strand. 63.6% of the genome is occupied by structural genes (40.5% corresponds to protein coding exons, 5.9% to the 28 tRNA genes, and 17.3% to the rRNA subunits), 8.8% by intergenic spacers that are 14–372 bp in size, and 32.4% by the 11 introns.

3.2. Protein coding genes

The *P. marneffei* mitochondrial genome contains 15 protein coding genes. These include genes encoding ATP synthase subunits 6, 8, and 9 (*atp6*, *atp8*, and *atp9*), the cytochrome oxidase subunits I, II, and III (*cox1*, *cox2*, and *cox3*), apocytochrome *b* (*cob*), the reduced nicotinamide adenine dinucleotide ubiquinone oxireductase subunits (*nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5*, and *nad6*), and the ribosomal protein of the small ribosomal subunit (*rps*). This set of protein coding genes is exactly the same as that in the *A. nidulans* mitochondrial genome. Furthermore, the gene order of the protein genes is the same as that in the *A. nidulans* mitochondrial genome, except for the *atp9* gene, which is located between the *cox1* and *nad3* genes in the *A. nidulans* mitochondrial genome, but between the *nad2* and *cob* genes in the *P. marneffei* mitochondrial genome (Fig. 2).

Concatenated amino acid sequences of the 14 protein coding genes in the mitochondrial genomes of *P. marneffei* and 24 other fungi were used for phylogenetic tree construction. The closest relatives of *P. marneffei* were *A. nidulans* and other molds, such as *P. anserina*, *N. crassa*, *Hypocrea jecorina*, and *Verticillium lecanii* (Fig. 3). On the other hand, the yeasts, such as the *Saccharomyces* species, *Schizosaccharomyces* species, *Candida* species, and *P. canadensis* were more distantly related to *P. marneffei*. This implied that phylogenetically the mitochondrial genome of *P. marneffei* is more related to those of molds than yeasts. This is in line with our previous observation and also results published by others, that when the chromosomal 18S rRNA genes or the internal transcribed spacers and 5.8S rRNA genes (ITS1-5.8S-ITS2) and mitochondrial small subunit rRNA genes were used for phylogenetic trees construction, the closest neighbors of *P. marneffei*, besides the other *Penicillium* species, were the *Aspergillus* species as well as other molds [15,17]. Furthermore, the same gene content and almost the same gene order in the mitochondrial genomes of *P. marneffei* and *A. nidulans* also implies that the mitochondrial genome is probably not related to the unique characteristic of thermal dimorphism of *P. marneffei*. Interestingly, *MP1*, the gene that encodes an abundant and highly immunogenic protein in *P. marneffei*, only has known homologs in *A. nidulans*, *A. fumigatus*, and *A. flavus*, but not in other fungi [10–12,28–31].

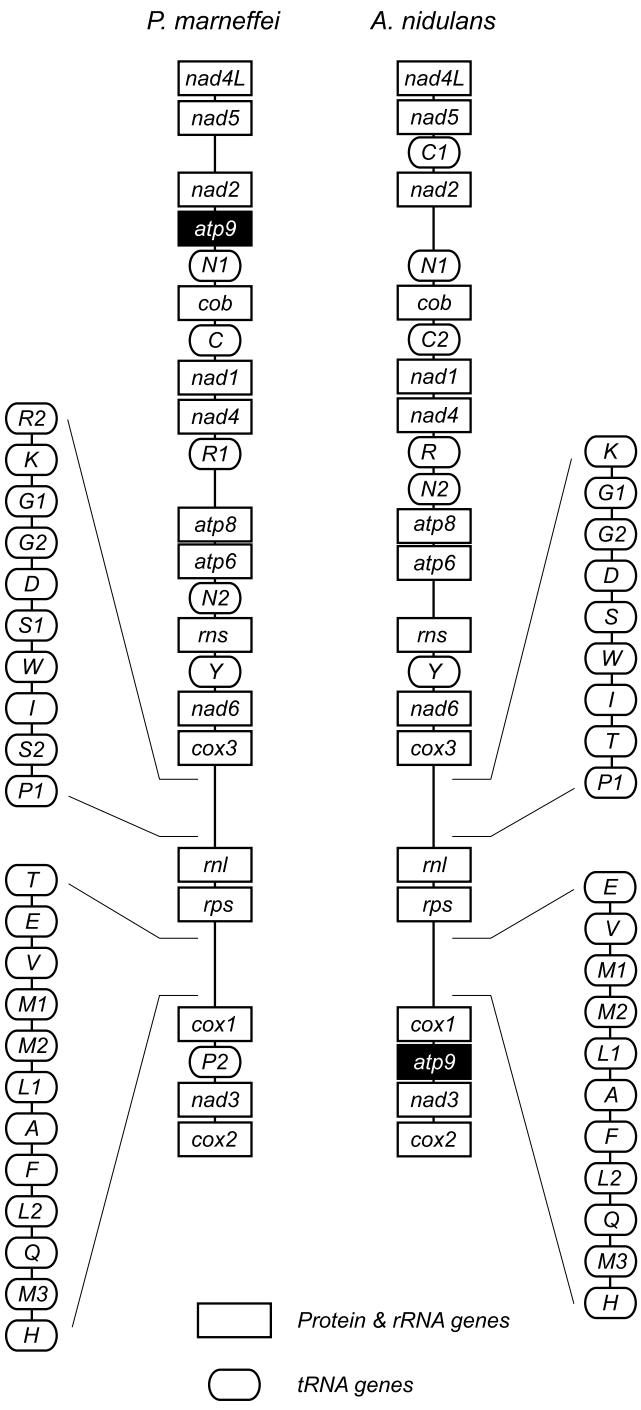

Fig. 2. Gene content and order comparison between *P. marneffei* mitochondrial DNA and *A. nidulans* mitochondrial DNA. The only exonic gene that has undergone gene rearrangement is *atp9*, which is highlighted in black background.

3.3. Genetic code and codon usage

Since the mitochondrial genome *P. marneffei* is phylogenetically closely related to those of molds and its gene content is the same as that of *A. nidulans*, the genetic code of the mito-

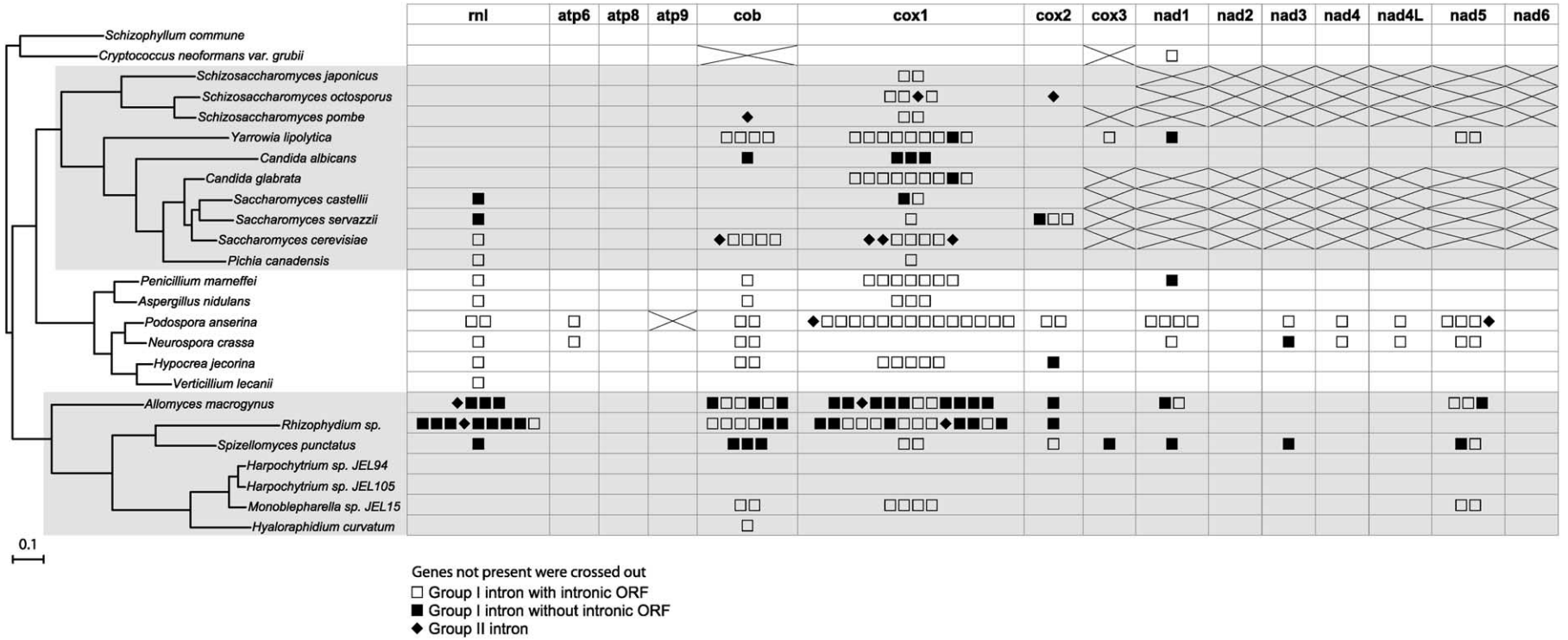

Fig. 3. Maximum likelihood tree showing phylogenetic relationships of *P. marneffei* to other fungi and distribution of group I and group II introns in the corresponding fungi. The tree was constructed using unambiguously aligned portions of concatenated amino acid sequences of the 14 protein-coding genes (*atp6*, *atp8*, *atp9*, *cob*, *cox1*, *cox2*, *cox3*, *nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5* and *nad6*). A total of 3462 amino acid positions were used for the inference with ProML. Sequences were obtained from GenBank: *A. macrogynus* (NC_001715), *A. nidulans* (CAA32799, CAA33481, AAA99207, AAA31737, CAA25707, AAA31736, CAA23994, P15956, CAA23995, CAA33116), *C. albicans* (NC_002653), *Candida glabrata* (NC_004691), *Cryptococcus neoformans var. grubii* (NC_004336), *Harpochytrium* sp. JEL105 (NC_004623), *Harpochytrium* sp. JEL94 (NC_004760), *Hyaloraphidium curvatum* (NC_003048), *Hypocrea jecorina* (NC_003388), *Monoblepharella* sp. JEL15 (NC_004624), *N. crassa* (CAA24041, CAA32799, AAA31961, CAA27029, CAA27418, AAA66053, AAA31959), *P. marneffei* (present study), *P. canadensis* (NC_001762), *P. anserina* (NC_001329), *Rhizophydium* sp. 136 (NC_003053), *Saccharomyces castellii* (NC_003920), *Saccharomyces cerevisiae* (NC_001224), *Saccharomyces servazzii* (NC_004918), *Schizophyllum commune* (NC_003049), *Schizosaccharomyces japonicus* (NC_004332), *Schizosaccharomyces octosporus* (NC_004312), *S. pombe* (NC_001326), *Spizellomyces punctatus* (NC_003052, NC_003061 and NC_003060), *Verticillium lecanii* (NC_004514), *Yarrowia lipolytica* (NC_002659). Some sequences of *A. nidulans* were downloaded from Fungal Mitochondrial Genome Project (http://megasun.bch.umontreal.ca/People/lang/FMGP/FMGP.html), and some sequences of *N. crassa* were downloaded from http://pages.slu.edu/faculty/kennellj/genbank.html. The scale bar indicates the branch lengths that were scaled in terms of expected numbers of amino acid substitutions.

Table 2
Codon usage in protein coding genes of *P. marneffei* mitochondrial genome

Codon	AA	Genes	ORFs	Codon	AA	Genes	ORFs	Codon	AA	Genes	ORFs	Codon	AA	Genes	ORFs
TTT	F	307	143	TCT	S	160	93	TAT	Y	191	180	TGT	C	24	21
TTC	F	66	13	TCC	S	1	5	TAC	Y	32	27	TGC	C	0	4
TTA	L	572	250	TCA	S	105	45	TAA	*	14	9	TGA	W	56	37
TTG	L	26	33	TCG	S	1	13	TAG	*	1	1	TGG	W	0	5
CTT	L	49	42	CCT	P	119	35	CAT	H	76	47	CGT	R	10	24
CTC	L	0	6	CCC	P	4	2	CAC	H	8	7	CGC	R	0	1
CTA	L	20	24	CCA	P	25	20	CAA	Q	83	75	CGA	R	0	1
CTG	L	0	4	CCG	P	4	3	CAG	Q	5	7	CGG	R	0	2
ATT	I	182	134	ACT	T	121	78	AAT	N	196	277	AGT	S	123	90
ATC	I	10	12	ACC	T	1	7	AAC	N	11	30	AGC	S	15	8
ATA	I	326	162	ACA	T	105	45	AAA	K	101	347	AGA	R	78	94
ATG	M	112	38	ACG	T	0	4	AAG	K	6	18	AGG	R	1	9
GTT	V	132	74	GCT	A	144	49	GAT	D	97	112	GGT	G	188	94
GTC	V	1	3	GCC	A	4	7	GAC	D	3	11	GGC	G	0	1
GTA	V	131	70	GCA	A	81	35	GAA	E	89	133	GGA	G	92	32
GTG	V	18	5	GCG	A	7	3	GAG	E	21	21	GGG	G	6	13

Numbers indicate the total numbers of codons in either identified protein coding genes or ORFs (including both free-standing URFs, intronic ORFs and RPS).

chondrial genome of *P. marneffei* is assumed to be the same as that of *A. nidulans* (Table 2).

There is a strong codon usage bias in exonic ORFs in the mitochondrial genome of *P. marneffei* towards codons ending in A or T. In fact, eight codons (CTC, CTG, ACG, TGC, TGG, CGC, CGG, and GGC) were not used at all, five codons (GTC, TCC, TCG, ACC, and AGG) were used only once, and nine codons (ATC, CCG, GCC, GCG, CAC, CAG, AGG, GAC, GGG) were used two to 10 times, in exonic ORFs. Moreover, this codon usage bias is also evident in the use of stop codon, where TAA is used as the stop codon in 14 genes, but TAG is only used in one gene.

3.4. tRNA genes

Twenty-eight tRNA genes were identified in the *P. marneffei* mitochondrial genome (Fig. 4). These are all located on the same DNA strand as the other genes. The set of mitochondrial tRNAs in *P. marneffei* is the same as that in *A. nidulans*. Furthermore, the sequences of the mitochondrial tRNA genes of *P. marneffei* are highly conserved with those of *A. nidulans*.

3.5. Other RNA genes

The genes that encode the 23S and 16S ribosomal RNAs of the large and small subunits of the ribosome (*rnl* and *rns*) were identified. Furthermore, a gene (*rps*), located within the intron of *rnl* (Table 1 and Fig. 5), that encodes the ribosomal protein of the small ribosomal subunit, which is also present in the *A. nidulans* mitochondrial genome, is also identified.

3.6. Group I introns

In *P. marneffei*, the *cox1* gene contains seven introns (PmCox1.1, PmCox1.2, PmCox1.3, PmCox1.4, PmCox1.5, PmCox1.6, and PmCox1.7), while the *cob* gene, *nad1* gene, and *rnl* gene contain one intron each (PmCob1.1, PmNad1.1, and PmRnl1.1 respectively). Each intron in the *cox1*, *nad1*, and *rnl* genes contains an ORF. The ORF in the *rnl* gene encodes the *rps* gene. The predicted secondary structures of two representative group I introns are depicted in Fig. 5. In both introns, the upstream exons end with a T and the introns end with a G, typical for most group I introns.

A comparison of the distribution of group I and group II introns in the 14 protein coding genes and *rnl* gene in the *P. marneffei* mitochondrial genome and that in the corresponding genes in the other 24 fungi is shown in Fig. 3. As a whole, the distribution of these introns in the genes encoded in the mitochondrial genome of *P. marneffei* concurs with those of the other fungi. The *cox1* gene, the gene that contains the largest number of self-splicing introns in other mitochondrial genomes, is also the gene that contains the largest number of self-splicing introns in the *P. marneffei* genome. The *cob* and *nad1* genes, the genes that also contain significant numbers of self-splicing introns, also possess one group I intron each in the *P. marneffei* mitochondrial genome.

3.7. Mitochondrial DNA sequences in nuclear genome

Presence of mitochondrial DNA sequence fragments in the corresponding nuclear genomes of *P. marneffei*, *A. nidulans*, *N. crassa*, *S. cerevisiae*, and *S. pombe* were compared (Table 3). By using the same method of sequence similarity comparison used for *S. cerevisiae* [27], only 10 mitochondrial DNA sequence fragments were detected in the 4× coverage, representing 95%, nuclear genome sequences for *P. marneffei* (Table 4). This number of mitochondrial DNA sequence fragments in the corresponding nuclear genomes, as well as the

Table 3
Comparison of presence of mitochondrial DNA fragments in nuclear genomes

Fungi	Number of mitochondrial DNA fragments in nuclear genomes	Size of mitochondrial genomes (kb)	Size of nuclear genome (Mb)	Ratio of sizes of mitochondrial to nuclear genome (kb/Mb)
P. marneffei	10	35.4	~29.5	~1.20
A. nidulans	17	~33.2	~31	~1.07
N. crassa	21	~64.8	~43	~1.51
S. cerevisiae	34	85.7	12.1	7.08
S. pombe	21	19.4	13.8	1.41

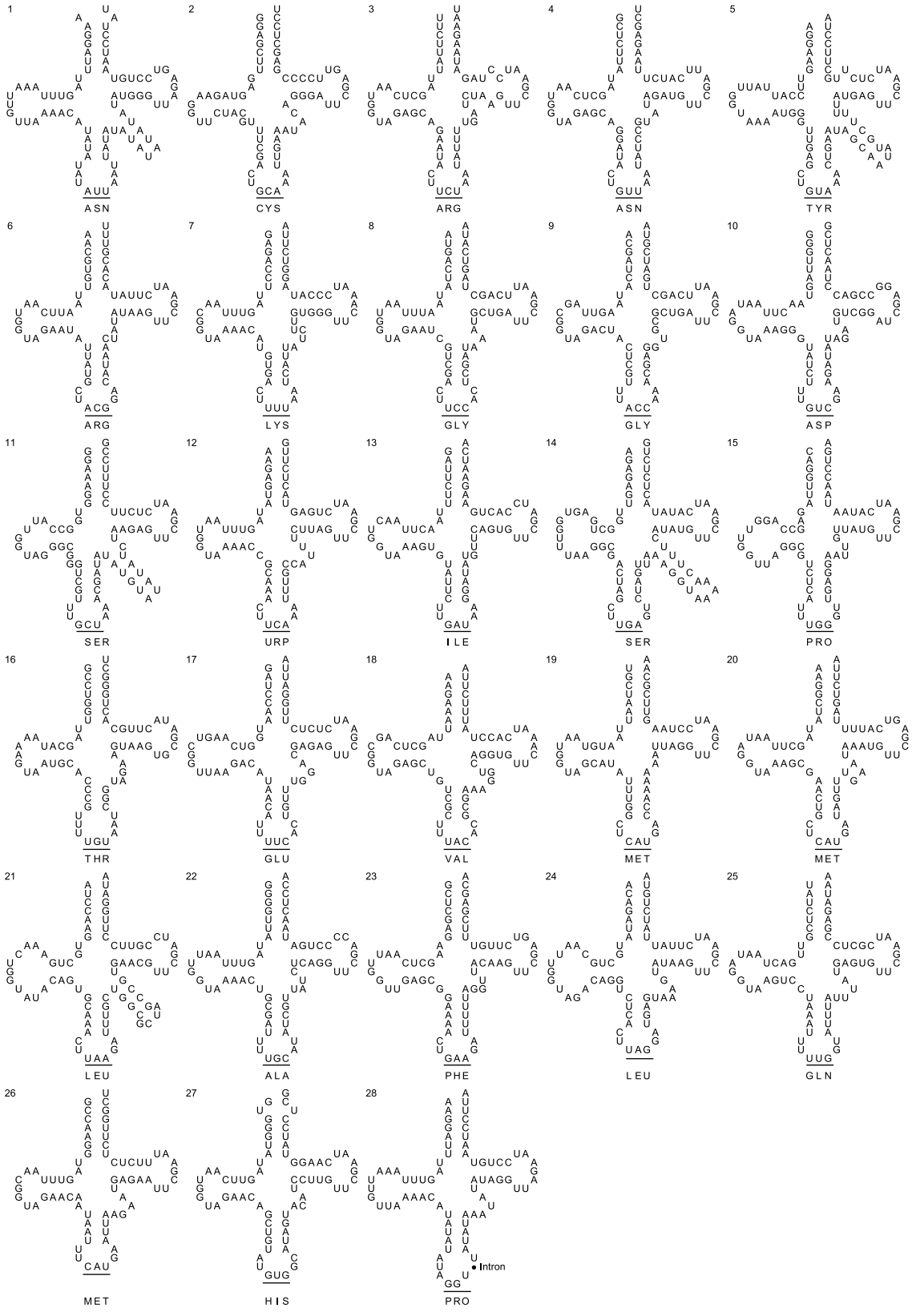

Fig. 4. Predicted clover-leaf structures of the 28 tRNAs encoded in the mitochondrial genome of *P. marneffei*. Anticodons are underlined and the corresponding amino acids are indicated. tRNAs are listed according to the order of their positions in the map in Fig. 1.

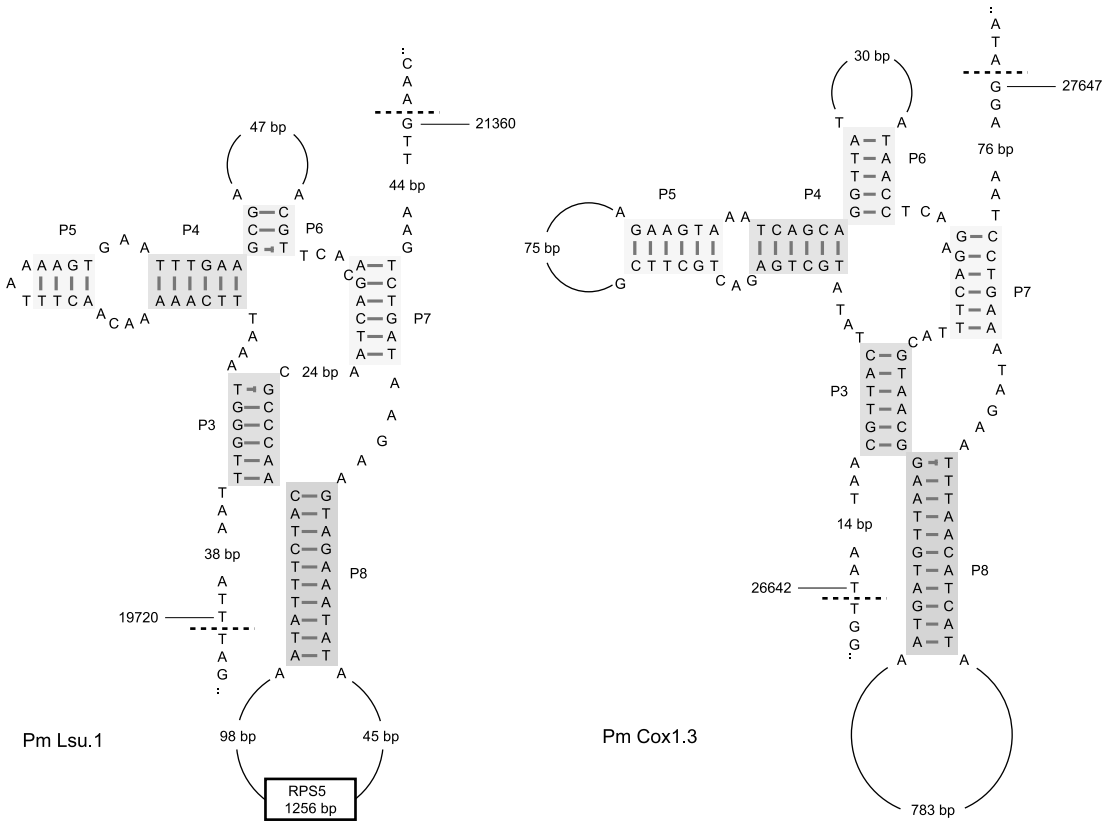

Fig. 5. Predicted secondary structures of two representative group I introns, Pm Rnl.1 and Pm Cox1.3, of *rnl* and *cox1* genes respectively, of *P. marneffei*. The exon/intron boundaries are represented by dotted lines. Base pairs are depicted by bars. The corresponding sizes of nucleotides not shown are indicated in bp. RPS5 gene is depicted by square box. The numbers correspond to the coordinates in the mitochondrial genome.

ratio of mitochondrial to nuclear genome size, was comparable to those found in *A. nidulans*, *N. crassa*, and *S. pombe* (Table 3). On the other hand, the number of mitochondrial DNA sequence fragments in the nuclear genome of *S. cerevisiae* was 34, which was about two times more than the other fungi. Although the relatively high ratio of mitochondrial to nuclear genome size of *S. cerevisiae* may partly explain this phenomenon, further studies would be necessary to elucidate the difference in the significance of these mitochondrial DNA fragments in the nuclear genomes for the different fungi.

4. Concluding remarks

Among the known mitochondrial genomes of fungi, the *P. marneffei* mitochondrial genome has an intermediate size. The replication origin of the *P. marneffei* mitochondrial genome is unknown. Despite the distinct biological property of thermal dimorphism in *P. marneffei*, its mitochondrial genome is much more closely related to those of molds, especially to that of *A. nidulans*, than to yeasts. The set of protein coding genes in the *P. marneffei* mitochondrial genome is exactly the same as that in the *A. nidulans* mitochondrial genome. Except for the *atp9* gene, the gene order of the protein genes is also the same as that in the *A. nidulans* mitochondrial genome. Furthermore, when concatenated amino acid sequences of 14 protein coding genes in the mitochondrial genomes of *P. marneffei* and 24 other fungi were used for phylogenetic tree construction, the closest relatives of *P. marneffei* were *A. nidulans* and other molds, whereas the yeasts were more distantly related.

Acknowledgements: This work was supported by the AIDS Trust Fund (MSS 083), Research Grant Council Grant (HKU 7363/03M),

Table 4
P. marneffei mitochondrial DNA sequences present in nuclear genome

Fragment number	Mitochondrial coordinates	Size of fragment (bp)	Location of fragment	BLAST score
1	9031..9069	39	*nad1*	9.00E-08
2	10182..10201	20	*nad4*	0.001
3	11622..11697	76	*atp6*	2.00E-15
4	13445..13465	21	*rrs*	2.00E-04
5	15158..15177	20	*rad6* < > *cox3*	0.001
6	18757..18776	20	*rnl*	0.001
7	25168..25187	20	*cox1*	0.001
8	31197..31216	20	*cox1*	0.001
9	32560..32580	21	*cox1*	2.00E-04
10	34510..34529	20	*nad3* < > *cox2*	0.001

and University Development Fund, Hong Kong. We thank Weiran Chen and Limin Liu for their help in the *P. marneffei* genome project and Dr. Frédérique Lisacek for his assistance in annotation.

References

[1] Yuen, K.Y., Wong, S.S., Tsang, D.N. and Chau, P.Y. (1994) Lancet 344, 444–445.
[2] Lo, C.Y., Chan, D.T., Yuen, K.Y., Li, F.K. and Cheng, K.P. (1995) Lupus 4, 229–231.
[3] Kwan, E.Y., Lau, Y.L., Yuen, K.Y., Jones, B.M. and Low, L.C. (1997) J. Paediatr. Child Health 33, 267–271.
[4] Chim, C.S., Fong, C.Y., Ma, S.K., Wong, S.S. and Yuen, K.Y. (1998) Am. J. Med. 104, 196–197.
[5] Wong, S.S., Siau, H. and Yuen, K.Y. (1999) J. Med. Microbiol. 48, 973–975.
[6] Wong, S.S., Woo, P.C. and Yuen, K.Y. (2001) Eur. J. Clin. Microbiol. Infect. Dis. 20, 132–135.
[7] Deng, Z.L. and Connor, D.H. (1985) Am. J. Clin. Pathol. 84, 323–327.
[8] Wong, K.H. and Lee, S.S. (1998) Singap. Med. J. 39, 236–240.
[9] Supparatpinyo, K., Khamwan, C., Baosoung, V., Nelson, K.E. and Sirisanthana, T. (1994) Lancet 344, 110–113.
[10] Cao, L., Chan, C.M., Lee, C., Wong, S.S. and Yuen, K.Y. (1998) Infect. Immun. 66, 966–973.
[11] Cao, L., Chen, D.L., Lee, C., Chan, C.M., Chan, K.M., Vanittanakom, N., Tsang, D.N. and Yuen, K.Y. (1998) J. Clin. Microbiol. 36, 3028–3031.
[12] Cao, L., Chan, K.M., Chen, D., Vanittanakom, N., Lee, C., Chan, C.M., Sirisanthana, T., Tsang, D.N. and Yuen, K.Y. (1999) J. Clin. Microbiol. 37, 981–986.
[13] Wong, S.S., Wong, K.H., Hui, W.T., Lee, S.S., Lo, J.Y., Cao, L. and Yuen, K.Y. (2001) J. Clin. Microbiol. 39, 4535–4540.
[14] Wong, L.P., Woo, P.C., Wu, A.Y. and Yuen, K.Y. (2002) Vaccine 20, 2878–2886.
[15] LoBuglio, K.F. and Taylor, J.W. (1995) J. Clin. Microbiol. 33, 85–89.
[16] Vanittanakom, N., Merz, W.G., Sittisombut, N., Khamwan, C., Nelson, K.E. and Sirisanthana, T. (1998) Med. Mycol. 36, 169–175.
[17] Yuen, K.Y., Pascal, G., Wong, S.S., Glaser, P., Woo, P.C., Kunst, F., Cai, J.J., Cheung, E.Y., Médigue, C. and Danchin, A. (2003) Arch. Microbiol. 179, 339–353.
[18] Foury, F., Roganti, T., Lecrenier, N. and Purnelle, B. (1998) FEBS Lett. 440, 325–331.
[19] Wood, V., Gwilliam, R., Rajandream, M.A., Lyne, M., Lyne, R., Stewart, A., Sgouros, J., Peat, N., Hayles, J., Baker, S., Basham, D., Bowman, S., Brooks, K., Brown, D., Brown, S., Chillingworth, T., Churcher, C., Collins, M., Connor, R., Cronin, A., Davis, P., Feltwell, T., Fraser, A., Gentles, S., Goble, A., Hamlin, N., Harris, D., Hidalgo, J., Hodgson, G., Holroyd, S., Hornsby, T., Howarth, S., Huckle, E.J., Hunt, S., Jagels, K., James, K., Jones, L., Jones, M., Leather, S., McDonald, S., McLean, J., Mooney, P., Moule, S., Mungall, K., Murphy, L., Niblett, D., Odell, C., Oliver, K., O'Neil, S., Pearson, D., Quail, M.A., Rabbinowitsch, E., Rutherford, K., Rutter, S., Saunders, D., Seeger, K., Sharp, S., Skelton, J., Simmonds, M., Squares, R., Squares, S., Stevens, K., Taylor, K., Taylor, R.G., Tivey, A., Walsh, S., Warren, T., Whitehead, S., Woodward, J., Volckaert, G., Aert, R., Robben, J., Grymonprez, B., Weltjens, I., Vanstreels, E., Rieger, M., Schafer, M., Muller-Auer, S., Gabel, C., Fuchs, M., Dusterhoft, A., Fritzc, C., Holzer, E., Moestl, D., Hilbert, H., Borzym, K., Langer, I., Beck, A., Lehrach, H., Reinhardt, R., Pohl, T.M., Eger, P., Zimmermann, W., Wedler, H., Wambutt, R., Purnelle, B., Goffeau, A., Cadieu, E., Dreano, S., Gloux, S., Lelaure, V., Mottier, S., Galibert, F., Aves, S.J., Xiang, Z., Hunt, C., Moore, K., Hurst, S.M., Lucas, M., Rochet, M., Gaillardin, C., Tallada, V.A., Garzon, A., Thode, G., Daga, R.R., Cruzado, L., Jimenez, J., Sanchez, M., del Rey, F., Benito, J., Dominguez, A., Revuelta, J.L., Moreno, S., Armstrong, J., Forsburg, S.L., Cerutti, L., Lowe, T., McCombie, W.R., Paulsen, I., Potashkin, J., Shpakovski, G.V., Ussery, D., Barrell, B.G., Nurse, P. and Cerrutti, L. (2002) Nature 415, 871–880.
[20] Brown, T.A., Waring, R.B., Scazzocchio, C. and Davies, R.W. (1985) Curr. Genet. 9, 113–117.
[21] Galagan, J.E., Calvo, S.E., Borkovich, K.A., Selker, E.U., Read, N.D., Jaffe, D., FitzHugh, W., Ma, L.J., Smirnov, S., Purcell, S., Rehman, B., Elkins, T., Engels, R., Wang, S., Nielsen, C.B., Butler, J., Endrizzi, M., Qui, D., Ianakiev, P., Bell-Pedersen, D., Nelson, M.A., Werner-Washburne, M., Selitrennikoff, C.P., Kinsey, J.A., Braun, E.L., Zelter, A., Schulte, U., Kothe, G.O., Jedd, G., Mewes, W., Staben, C., Marcotte, E., Greenberg, D., Roy, A., Foley, K., Naylor, J., Stange-Thomann, N., Barrett, R., Gnerre, S., Kamal, M., Kamvysselis, M., Mauceli, E., Bielke, C., Rudd, S., Frishman, D., Krystofova, S., Rasmussen, C., Metzenberg, R.L., Perkins, D.D., Kroken, S., Cogoni, C., Macino, G., Catcheside, D., Li, W., Pratt, R.J., Osmani, S.A., DeSouza, C.P., Glass, L., Orbach, M.J., Berglund, J.A., Voelker, R., Yarden, O., Plamann, M., Seiler, S., Dunlap, J., Radford, A., Aramayo, R., Natvig, D.O., Alex, L.A., Mannhaupt, G., Ebbole, D.J., Freitag, M., Paulsen, I., Sachs, M.S., Lander, E.S., Nusbaum, C. and Birren, B. (2003) Nature 422, 859–868.
[22] Ewing, B. and Green, P. (1998) Genome Res. 8, 186–194.
[23] Ewing, B., Hillier, L., Wendl, M.C. and Green, P. (1998) Genome Res. 8, 175–185.
[24] Gordon, D., Abajian, C. and Green, P. (1998) Genome Res. 8, 195–202.
[25] Paquin, B. and Lang, B.F. (1996) J. Mol. Biol. 255, 688–701.
[26] Koszul, R., Malpertuy, A., Frangeul, L., Bouchier, C., Wincker, P., Thierry, A., Duthoy, S., Ferris, S., Hennequin, C. and Dujon, B. (2003) FEBS Lett. 534, 39–48.
[27] Ricchetti, M., Fairhead, C. and Dujon, B. (1999) Nature 402, 96–100.
[28] Yuen, K.Y., Chan, C.M., Chan, K.M., Woo, P.C., Che, X.Y., Leung, A.S. and Cao, L. (2001) J. Clin. Microbiol. 39, 3830–3837.
[29] Chan, C.M., Woo, P.C., Leung, A.S., Lau, S.K., Che, X.Y., Cao, L. and Yuen, K.Y. (2002) J. Clin. Microbiol. 40, 2041–2045.
[30] Woo, P.C., Chan, C.M., Leung, A.S., Lau, S.K., Che, X.Y., Wong, S.S., Cao, L. and Yuen, K.Y. (2002) J. Clin. Microbiol. 40, 4382–4387.
[31] Woo, P.C., Chong, K.T., Leung, A.S., Wong, S.S., Lau, S.K. and Yuen, K.Y. (2003) J. Clin. Microbiol. 41, 845–850.