Yeast
Yeast 2002; **19**: 1221–1231.
Published online in Wiley InterScience (www.interscience.wiley.com). **DOI:** 10.1002/yea.912

Yeast Sequencing Report

The *kex2* gene from the dimorphic and human pathogenic fungus *Paracoccidioides brasiliensis*

Emerson J. Venancio[1,2], Bruno S. Daher[2], Rosângela V. Andrade[2], Célia M. A. Soares[3], Ildinete Silva Pereira[2] and Maria Sueli S. Felipe[2]*

[1] *Departamento de Ciências Patológicas, Centro de Ciências Biológicas, Universidade Estadual de Londrina, 86051–970 Londrina, PR, Brasil*
[2] *Laboratório de Biologia Molecular, Instituto de Ciências Biológicas, Universidade de Brasília, 70910–900 Brasília, DF, Brasil*
[3] *Laboratório de Biologia Molecular, Instituto de Ciências Biológicas, Universidade Federal de Goiás, 74001–970 Goiânia, GO, Brasil*

**Correspondence to: Maria Sueli S. Felipe, Laboratorio de Biologia Molecular, Instituto de Ciências Biológicas, ICC-Sul, Campus Universitário Darcy Ribeiro, Universidade de Brasília, 70910–900 Brasília, DF, Brasil. E-mail: msueli@unb.br*

Received: 26 April 2002
Revised: 25 July 2002

Abstract

Kexin-like protein is a component of the subtilase family of proteinases involved in the processing of proproteins to their active forms. Kexin-like proteins are also synthesized as a propeptide and this is involved in (auto)inhibition, correct folding and subcellular sorting of proteins. The kexin-like protein was described as the product of the *kex2* gene for *Aspergillus niger*, *Candida albicans*, *Saccharomyces cerevisiae*, *Yarrowia lipolytica* and other fungi. Disruption of the *kex2* gene in *C. albicans* and *Y. lipolytica* affects hyphae production and induces morphological cell defects, strongly suggesting a possible role of kexin-like proteins in dimorphism of human pathogenic fungi. In this work, we report the nucleotide sequence of the *kex2* gene cloned from the dimorphic and human pathogenic fungus *Paracoccidioides brasiliensis* (*Pbkex2*). An open reading frame (ORF) of 2622 bp was identified in the complete sequence, interrupted by only one intron of 93 bp. The 5′ non-coding region contains consensus sequences such as canonical TATA, CAAT boxes and putative motifs for transcriptional factors binding sites, such as HSE-like regulating genes involved in thermo-dependent processes; Xbp1, reported as a transcriptional factor that may control genes involved in cell morphology; and StuAp, which may regulate spore differentiation and pseudohyphal growth in fungi. In the 3′ non-coding region were observed the canonical motifs necessary for correct mRNA processing and polyadenylation. The deduced protein sequence consists of 842 amino acid residues, showing identity to kexin-like proteinases from *A. niger* (55%), *Emericella nidulans* (53%) and *C. albicans* (48%). Comparative sequence analysis of *P. brasiliensis* kexin-like protein reveals the presence of homologous regions related to a signal peptide, a propeptide, a subtilisin-like catalytic domain, a P domain, a S/T rich region and a transmembrane domain. A putative Golgi retrieval signal (YEFEMI) has also been found in the cytoplasmic tail. The complete nucleotide sequence of *Pbkex2* and its flanking regions have been submitted to GenBank database under Accession No. AF486805.

Keywords: ***Pbkex2* gene; kexin-like protein; proteinase; dimorphic fungus; human pathogen; *Paracoccidioides brasiliensis***

Introduction

Kexin-like proteins are Ca^{2+}-dependent transmembrane serine proteinases of the kexin subfamily and belong to the conserved subtilases family (Seidah and Chrétien, 1999). Other proteins of the kexin subfamily are the mammalian precursor convertases (PCs) and the furins (Seidah and Chrétien, 1999). Kexin-like proteinases recognize and process the amino-terminal portion of proproteins, denominated as propeptides, to their active forms (Bergeron *et al*., 2000). Most of these propeptides

are cleaved after KR or RR at positions P2 and P1, where P1 is the first residue and P2 is the second residue on the amino-terminal side of the cleavage site (Bergeron *et al*., 2000). The members of the kexin subfamily are synthesized with a signal peptide followed by a propeptide at the amino-terminus, a subtilisin-like catalytic domain and a P domain, which is essential for catalytic activity and stability. The carboxy-terminal region may or may not have an S/T-rich sequence, a C-rich sequence, an amphipathic region, a single transmembrane domain and a cytoplasmatic domain (Seidah and Chrétien, 1999). The *Saccharomyces cerevisiae* Kex2, furins and some PCs all contain a single transmembrane domain. The *S. cerevisiae* Kex2 proteinase contains in the cytoplasmic domain a Golgi retrieval signal necessary to its retention in the Golgi compartment (Wilcox *et al*., 1992). The first *KEX2* gene was identified in *S. cerevisiae* by Wickner and Leibowitz (1976). The functional characterization of *S. cerevisiae KEX2* was reported by Julius *et al*. (1984).

In fungi, kexin-like proteinases are described to be involved in processing many proteins, such as hydrophobin (McCabe and Van Alfen, 1999), alkaline extracellular phosphatase (AEP; Enderlin and Ogrydziak, 1994), extracellular lipase (Lip2p; Pignede *et al*., 2000), secretory aspartyl proteinases (Saps; Newport and Agabian 1997), mannoprotein (Hwp1; Staab *et al*., 1996), α-mating factor and killer toxin in *S. cerevisiae* (Fuller *et al*., 1988). In *C. albicans*, the disruption of two *kex2* alleles impaired hyphae production and induced morphological defects in the cell, which could be reversed by transformation of the null mutants with a plasmid containing wild-type *KEX2* gene (Newport and Agabian, 1997). Also, in *Yarrowia lipolytica*, disruption of the *XPR6* gene (a *KEX2*-like gene) had a marked effect on hyphae formation (Richard *et al*., 2001). These results suggest a role of kexin-like proteins in dimorphism, an important feature of human pathogenic fungi (Mitchell, 1998), although the precise molecular mechanisms connecting kexin-like proteinases and dimorphism have not yet been described.

The temperature-dependent dimorphic fungus *P. brasiliensis* is the aetiological agent of paracoccidioidomycosis (PCM), a human systemic mycosis endemic in Latin America, occurring mainly in Colombia, Venezuela and Brazil. It is estimated to infect approximately 10 million people in Latin America (McEwen *et al*., 1995), with about 2% developing the disease. In most cases, disease starts after inhalation of fungal propagules that can differentiate to the yeast form. The primary infection usually occurs in the lung, and may later disseminate to other organs (Montenegro and Franco, 1994).

The mycelium to yeast transition of *P. brasiliensis* has been shown to be essential for establishment of the infection (Franco, 1987). Although the molecular mechanisms of dimorphism in *P. brasiliensis* are largely unknown, several genes have recently been described in *P. brasiliensis*, which may potentially be involved in dimorphism and/or pathogenicity (Silva *et al*., 1999; Nino-Vega *et al*., 2000; Izacc *et al*., 2001). More recently, Venancio *et al*. (2002) were able to identify three cDNA fragments (M73, M51 and M32) specific for the mycelial form and two others (MY-1 and MY-2) probably upregulated during M–Y transition. In this paper we describe the cloning, sequencing and structural analysis of the gene coding for a kexin-like protein from *P. brasiliensis*. In addition, the deduced structural motifs from *P. brasiliensis* kexin-like protein are compared with other fungal kexin-like proteinases.

Materials and methods

P. brasiliensis strain and growth conditions

The clinical isolate of *P. brasiliensis*, Pb01 (ATCC-MYA-826), was grown as a mycelium (22 °C) or a yeast (36 °C) on semisolid Fava–Neto's medium, as previously described (Silva *et al*., 1994).

Genomic library

The genomic library was constructed in λDASH II vector (Stratagene, USA). Briefly, *P. brasiliensis* total DNA fragments were partially digested with *Sau*3AI and separated on 1% agarose gels. Fragments around 17 kb in length were purified and cloned into the arms of the vector. The recombinant phages were obtained and amplified, resulting in a titre of 10^{12} pfu/ml (Pereira *et al*., 2000).

Amplification of M-Y2 DDRT–PCR fragment

The M-Y2 product was obtained by DDRT–PCR (Venancio *et al*., 2002). Briefly, cDNA fragments

were synthesized by reverse transcription (RT) reaction using H-T11A primer (5′-AAGCTTTTTT-TTTTTA-3′), 100 ng DNase treated total RNA (previously denatured at 65 °C for 10 min) from different times during the mycelium to yeast transition and SuperScript II RT (Gibco-BRL, USA). The cDNA fragments were submitted to amplification by PCR using primers H-T11A and H-AP3 (5′-AAGCTTTGGTCAG-3′). PCR products were run in a denaturing polyacrylamide gel and electrophoretic profiles were compared. The M-Y2 cDNA fragment was cut from the gel, reamplified using the same primers mentioned above and used as a probe.

Isolation, subcloning and sequencing of *kex2* gene from *P. brasiliensis*

The screening of the *kex2* gene was performed using the genomic library from *P. brasiliensis*. About 7×10^4 phage pfu were plated according to the standard procedures (Sambrook *et al*., 1989). The radiolabelled DDRT–PCR product M-Y2 was used as a probe. Hybridization was carried out in 45% formamide at 42 °C for 18 h and the membrane washed with 2 × SSC/0.1% SDS/30 min/50 °C, 1 × SSC/0.1% SDS/30 min/50 °C, 0.1 × SSC/ 0.1% SDS/30 min/50 °C and 0.1 × SSC/0.1% SDS/30 min/65 °C. After exposure to X-ray film, at −80 °C, six recombinant phages were identified. The phage DNA from the six clones were isolated and digested with *Eco*RI, *Bam*HI, and double-digested with *Eco*RI/*Bam*HI. Digestion fragments were analysed by Southern blot, also using as probe the DDRT–PCR product M-Y2 labelled using the AlkPhos direct kit (Amersham Pharmacia Biotech, USA). Hybridization and washing were performed as described by the suppliers. The CDP-Star reagent was used for signal detection (Amersham Pharmacia Biotech, USA) according to the manufacturer's instructions.

A DNA fragment of 5.5 kb generated from *Eco*RI digestion of clone G1 was selected and subcloned into dephosphorylated pUC 18 vector, predigested with *Eco*RI. Both strands of the 5.5 kb fragment were sequenced by primer walking, using the automated sequencer MegaBace 1000 (Amersham Pharmacia Biotech, USA), at the Molecular Biology Laboratory of Brasília University, Brasília-DF, Brazil.

Nucleotide and protein sequence analysis of *kex2* gene from *P. brasiliensis*

The nucleotide sequence obtained was analysed by the computational programs PHRED (Ewing *et al*., 1998), PHRAP and CONSED (Gordon *et al*., 1998). The search for data bank similarity was conducted using the BLAST search tools, using the interface web of the National Center for Biotechnology Information (NCBI) (**http://www.ncbi.nlm.nih.gov/**). Multiple alignments were generated using the program ClustalW (**http://www.ebi.ac.uk/clustalw**). Prediction of the promoter region was made using the Neural Network Promoter Prediction Program (**http://www.fruitfly.org/seq_tools/promoter.html**) and the TATA-like elements were mapped through the program HCtata (**http://l25.itba.mi.cnr.it/~webgene/wwwHC_tata.html**). The TFSEARCH (**http://www.cbrc.jp/research/db/TFSEARCH.html**) and MatInspector (**http://www.genomatix.de/cgi-bin/matinspector/matinspector.pl**) programs were used to identify putative transcription factor motifs, while the SignalP (**http://www.cbs.dtu.dk/services/SignalP/**) and SOSUI (**http://sosui.proteome.bio.tuat.ac.jp**) programs were used to find signal sequences for the secretion and putative transmembrane regions, respectively. The hydrophobic profile analysis was made according Kyte and Doolittle (1982), using the ProtScale (**http://ca.expasy.org/cgi-bin/protscale.pl**) program.

Results and discussion

Cloning and nucleotide sequencing of *kex2* gene from *P. brasiliensis*

A genomic library of *P. brasiliensis* was screened using as probe the M-Y2 cDNA fragment, previously reported by Venancio *et al*. (2002). Several positive clones were found; six of them were isolated and, following total DNA extraction, were analysed by single (*Eco*RI) or double digestion (*Eco*RI/*Bam*HI). Three distinct digestion patterns were obtained, generating a range of fragments of 0.4–10 kb (Figure 1A). The digested DNA was further analysed by Southern blotting, also using the M-Y2 cDNA fragment as a probe. Three different hybridization patterns were detected by *Eco*RI digestion of G1, G2 and G3 clones. The *Eco*RI/*Bam*HI double digestion also

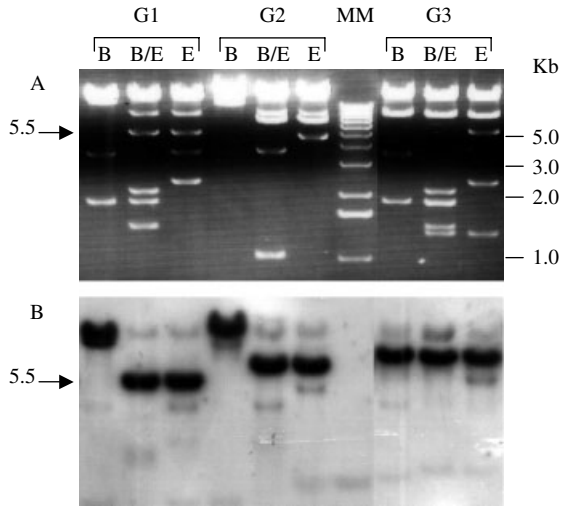

Figure 1. Southern-blot analysis of selected clones from *P. brasiliensis* genomic library. DNA from three genomic library clones (G1, G2, G3) was digested with *Bam*HI, *Eco*RI and *Bam*HI/*Eco*RI. The fragments were separated in a 0.8% agarose gel (A) and transferred to nylon membranes. Hybridization was performed using DDRT–PCR product M-Y2, labelled by the AlkPhos direct kit (Amersham Pharmacia Biotech, USA) (B). Three distinct digestion patterns are shown. An arrow indicates the DNA fragment chosen for subcloning. Molecular markers are depicted in lane MM

displayed three distinct band sizes as can be seen in Figure 1B. The data suggest that *P. brasiliensis kex2* gene (*Pbkex2*) sequences are located at least into three genomic arrangements. A 5.5 kb fragment, released by *Eco*RI digestion from G1 clone (arrow), was chosen for subcloning into pUC18 vector (Figure 1B, lane G1/E). The pC1 clone was obtained and the cloned DNA fragment was sequenced, showing a high identity to *kex2* genes (Enderlin *et al*., 1994; Jalving *et al*., 2000).

Analysis of *kex2* nucleotide sequence from *P. brasiliensis*

The complete nucleotide sequence of the *Pbkex2* gene, the 5′ and 3′ flanking regions and the deduced amino acid sequence are shown in Figure 2. A schematic representation of the genomic fragment containing the *Pbkex2* gene is also presented in Figure 3A, showing the open reading frame (ORF) position into the entire 5,5 kb DNA fragment corresponding to the clone G1 digested with *Eco*RI. The presence of sequences at the 5′ non-coding region that are similar to characteristic features of regulatory elements found in *S. cerevisiae*, *C. albicans* and *Y. lipolytica* and other microorganisms was also observed in a putative promoter region of 535 bp from the *Pbkex2* gene (Figure 2). The programs HC_TATA and Neural Network Promoter were used to determinate the position of TATA-like elements and transcription start points, respectively. We have detected a sequence of one CAAT box (−128), two TATA-like elements (−90 and −174), and two putative transcriptional start points (−157, −344). The TATA-box at position −174 and the transcriptional start point (−157) may correspond to a class of minimal promoters, considering that the distance of about 20–30 bp is commonly found in this class of promoters (Pedersen *et al*., 1999). The 5′ non-coding region has also been submitted to a search for transcriptional factor binding sites (MatInspector and TFSEARCH). Putative sequences related to transcriptional factors binding sites were observed in the *Pbkex2* gene promoter region. We have detected the presence of HSE-like motifs (−143); Xbp1 (−26, −82); StuAp (−100, −205); AbaA (−221) and GCR1 (−188). A HSE-like motif, a binding site for a heat shock factor, was detected in the promoter region of the *Pbkex2* gene, which may control thermoregulated genes (Santoro *et al*., 1998). Xbp1 is a stress-induced transcriptional repressor found in *S. cerevisiae* that contributes to changes in the cell cycle and cell morphology and is induced by heat shock, high osmolarity, oxidative stress, DNA damage and glucose starvation (Miled *et al*., 2001). StuAp regulates spore differentiation and ascosporogenesis in *Aspergillus nidulans* and may be involved in controlling pseudohyphal growth in budding yeast (Dutton *et al*., 1997). In *A. nidulans*, AbaA transcriptional factor is necessary for the terminal stages of conidiophore development (Ye *et al*., 1999). GCR1 mediates the activation of glycolytic enzymes (Deminoff and Santangelo, 2001). The relevance, functionality and regulatory significance of these motifs has not yet been demonstrated in *P. brasiliensis*.

The entire *Pbkex2* gene ORF comprises 2622 bp from the start to the termination codon. The ORF is interrupted by just one 93 bp intron (position 1779–1871), the presence and localization of which was confirmed by alignment of genomic DNA and cDNA sequences (data not shown). The intron of the *Pbkex2* gene is in a conserved position relative to the *kexB* gene from *A. niger*

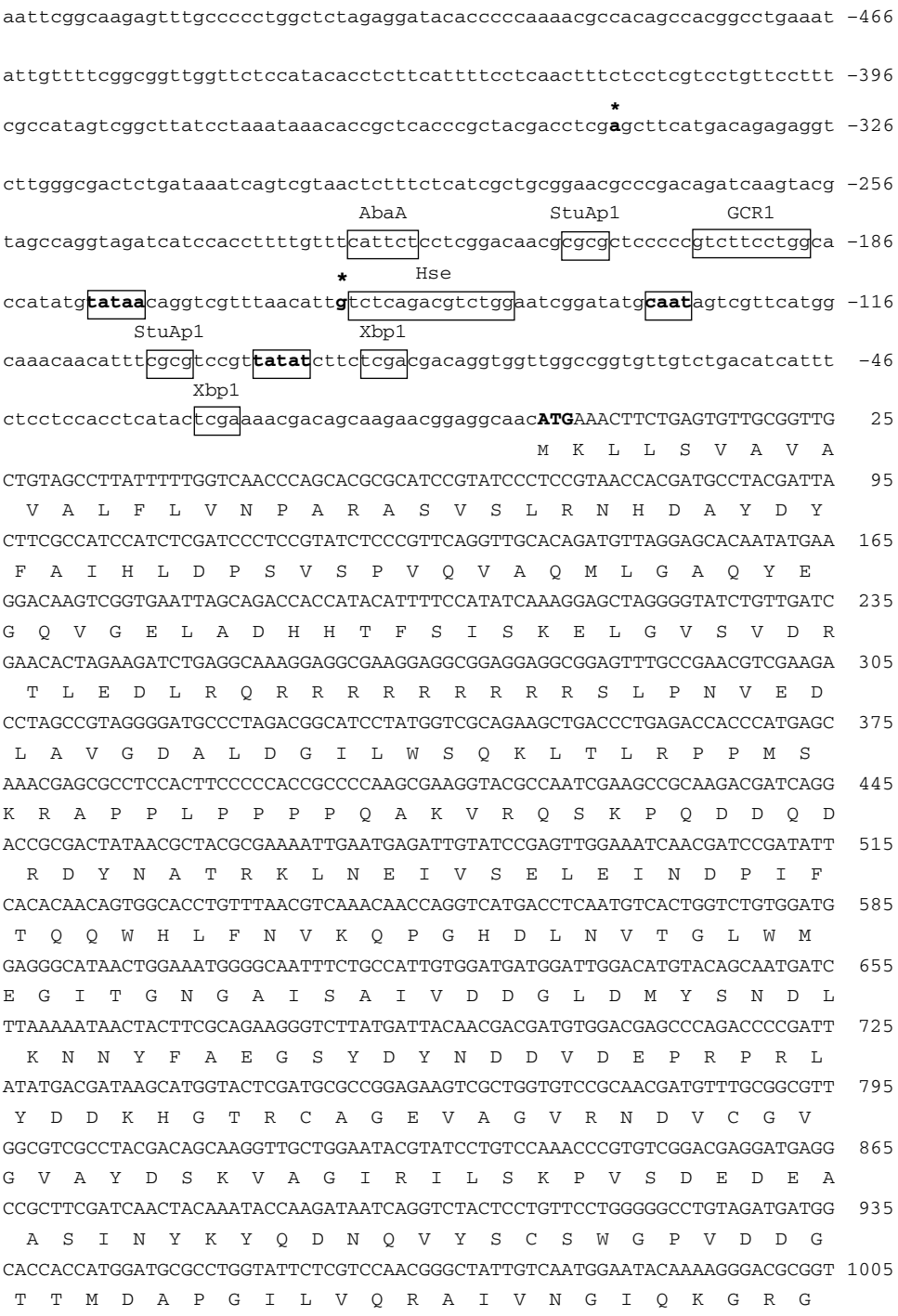

Figure 2. Nucleotide and deduced amino acid sequences of the *P. brasiliensis* gene (*Pbkex2*). The *Eco*RI fragment from clone G1 was sequenced. An open reading frame of 2622 bp and the predicted 842 amino acids residues are shown. The exon sequence is indicated by upper case letters, while the intron and 5′/3′ non-coding regions are in lower case letters. The junctions GT–AG and Lariat consensus sequence are underlined. Initiation and termination codes are in bold. Putative regions related to TATA box, CAT box, 'efficient element' (EE), 'positioning element' (PE) and polyadenylation site (PS) are in bold and boxed. Two predicted transcriptional start points are labelled with asterisks. Putative motifs for transcriptional regulation sites are boxed, with their names appearing above


```
GGGAAAGGCTCCGTGTATGTTTTCGCCGCCGGCAACGGCGCTCTACATGAAGATAACTGCAACTTTGATG 1075
G  K  G  S  V  Y  V  F  A  A  G  N  G  A  L  H  E  D  N  C  N  F  D  G
GCTACACGAACAGCATTTACAGTGTCACTGTGGGTGCAATTGATCACAATGACAAGCACCCATATTATTC 1145
  Y  T  N  S  I  Y  S  V  T  V  G  A  I  D  H  N  D  K  H  P  Y  Y  S
TGAGCCGTGCTCTGCGCAACTCGTGGTGACGTATAGCAGCGGAGGGGGGAATGCTATACACACCACCGAC 1215
  E  P  C  S  A  Q  L  V  V  T  Y  S  S  G  G  G  N  A  I  H  T  T  D
GTTGGTGTGAATACTTGCACTACGAAACATGGTGGTACGTCTGCAGCTGGACCGTTGGTCGTTGGGGTTG 1285
V  G  V  N  T  C  T  T  K  H  G  G  T  S  A  A  G  P  L  V  V  G  V  V
TGGCACTTGCGCTGAGCGTCCGCCCGGAGCTGACCTGGCGTGACGTGCAATATATTCTATTGGAGACGGC 1355
  A  L  A  L  S  V  R  P  E  L  T  W  R  D  V  Q  Y  I  L  L  E  T  A
GATCCCCGTCAATTTAAATGAAAGCTATTGGCAGGATACTGCTACTGGGAAGAAATTCAGTCACGATTAC 1425
 I  P  V  N  L  N  E  S  Y  W  Q  D  T  A  T  G  K  K  F  S  H  D  Y
GGATACGGGAAAGTAGACGCTTATTCTGCTGTCCATCTAGCCATGACGTGGAAATTGGTCAAACCACAAG 1495
G  Y  G  K  V  D  A  Y  S  A  V  H  L  A  M  T  W  K  L  V  K  P  Q  A
CTTGGCTACACTCCCCGTGGTTGCAAGTTTATGCCGATATCCCGCAGGGAGACAAGGGACTTGCTAGTAG 1565
  W  L  H  S  P  W  L  Q  V  Y  A  D  I  P  Q  G  D  K  G  L  A  S  S
CTTTGAAGTCACTAAGGAGCTGTTGGTGAGGAATAACGTTGAGCGACTTGAACATGTTACGCTTACGATG 1635
 F  E  V  T  K  E  L  L  V  R  N  N  V  E  R  L  E  H  V  T  L  T  M
AATATCAACCACACGCGCCGTGGTGATTTAAGTGTGGAACTCCGCAGTCCCACTGGAGCCGTCAGCTATC 1705
N  I  N  H  T  R  R  G  D  L  S  V  E  L  R  S  P  T  G  A  V  S  Y  L
TAAGTACGACGAGGAAATTGGATGACCTCCGCGCTGGCTACGTTGATTGGACTTTTATGTCGCTCGTTCA 1775
  S  T  T  R  K  L  D  D  L  R  A  G  Y  V  D  W  T  F  M  S  L  V  H
TTGgtaagtgccaacaaaagccccccactcctgctatccaaccccaaaccagaagcattgaacatgaaa 1845
 W
atctaacgacagtcaacaaaactcagGGGTGAATCCGGCATCGGCAAATGGACCGTCATCGTCAAGGACA 1915
                           G  E  S  G  I  G  K  W  T  V  I  V  K  D  T
CAATCGTCAATGATTTCAAGGGTGTGTTCATCGACTGGCAACTCAGCCTGTGGGGCGAAGCCATCGACGC 1985
  I  V  N  D  F  K  G  V  F  I  D  W  Q  L  S  L  W  G  E  A  I  D  A
AGATATCCAAGGCCTCCATCCTCTACCAGACGAACATGACCACGATCACAACACCGAACTTGCACCCGTA 2025
 D  I  Q  G  L  H  P  L  P  D  E  H  D  H  D  H  N  T  E  L  A  P  V
GCCACCACAACGATTATCTCCAGCGCAACCTCTTCGAAGCCCACCGCCTCCGTCGTCCCAACAGACCACA 2125
A  T  T  T  I  I  S  S  A  T  S  S  K  P  T  A  S  V  V  P  T  D  H  I
TCGATCGCCCCGTCAACACTAAACCCACTGGAACACAAAGCCCTTCTCAACCATCAGCATCATCACAACC 2195
  D  R  P  V  N  T  K  P  T  G  T  Q  S  P  S  Q  P  S  A  S  S  Q  P
AACACCCACTTCAACATCCGACAGTTTCCTCCCCTCCTTCTTCCCCACCTTCGGCGTCTCCAAGCGTACG 2265
 T  P  T  S  T  S  D  S  F  L  P  S  F  F  P  T  F  G  V  S  K  R  T
CAGGTCTGGATCTACGGCTCCCTCACGCTTATTCTCATCTTCTGTTCCACGCTGGGTACATACTTCCTCA 2335
Q  V  W  I  Y  G  S  L  T  L  I  L  I  F  C  S  T  L  G  T  Y  F  L  I
TCCAGCGGCGGAAGCGCATCCGCAGCAATCTACACGATGACTACGAGTTTGAGATGATTGACGATGAGAG 2405
  Q  R  R  K  R  I  R  S  N  L  H  D  D  Y  E  F  E  M  I  D  D  E  S
CGACGAAACGGCTCCGCTTAGTACGGGCGGGCATGTTGCTGGCGCTAGAACTGGGCGGAAGCAGCGAAGG 2475
 D  E  T  A  P  L  S  T  G  G  H  V  A  G  A  R  T  G  R  K  Q  R  R
AAGGGCGGAGAGTTGTATGATGCGTTTGCGGGGGGGAGTGATGAGGAGCTATTGAGCGATGGGGAGAGTG 2545
K  G  G  E  L  Y  D  A  F  A  G  G  S  D  E  E  L  L  S  D  G  E  S  E
```

Figure 2. Continued

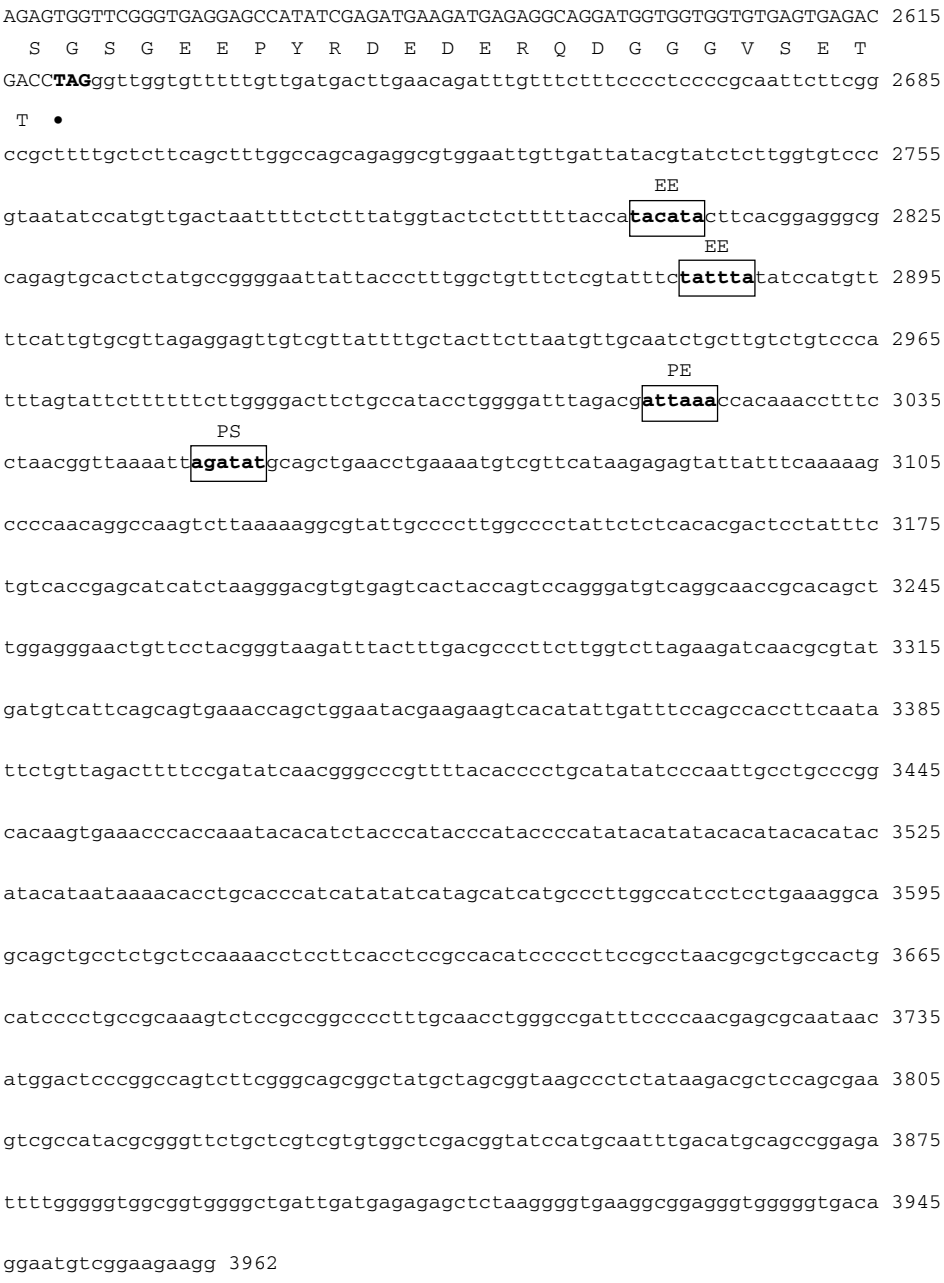

Figure 2. Continued

(Jalving *et al*., 2000) and *E. nidulans* (Accession No. AF291662) and contains the conserved motifs (Langford *et al*., 1984).

In general, the 3′ ends of mRNAs from *S. cerevisiae* present three important elements for mRNA processing: (a) a 'positioning element'; (b) an 'efficiency element'; and (c) a polyadenylation [poly(A)] site (Wahle and Ruegsegger, 1999). In *S. cerevisiae* a highly conserved 'positioning element' sequence is not normally observed, in contrast to mammalian AATAAA motifs. This sequence is necessary to direct the pre-mRNA cleavage at 10–30 nucleotides downstream. In the *Pbkex2* gene a putative 'positioning element' (ATTAAA) is found at 32 bp upstream from the poly(A) site. At the 3′ non-coding region of the

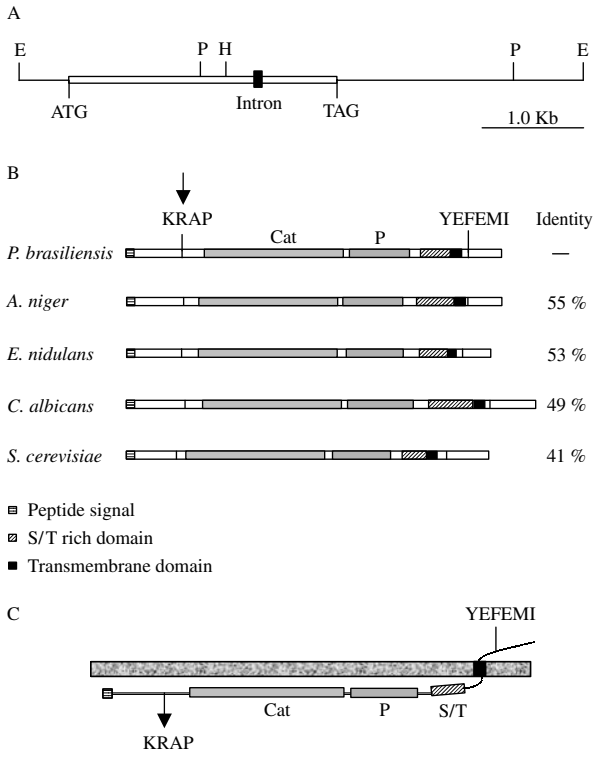

Figure 3. Schematic representation of *Pbkex2* gene and structural motifs comparison of fungal kexin-like proteinases. (A) Schematic representation of *P. brasiliensis kex2* genomic region. The intron is indicated by a black box. The position of restriction sites [*Eco*RI(E), *Pst*I(P), *Hin*dIII(H)] and the initiation/ stop codons are shown. (B) Comparison of fungus-like proteinases domains. The following domains are indicated: peptide signal, Cat, catalytic domain; P, P domain; S/T rich domain, transmembrane domain. The putative Golgi retention signal (YEFEMI) and the cleavage site (KRAP) are also indicated. (C) Representation of the predicted topology of the *P. brasiliensis* kexin-like protein. Only one transmembrane region is detected in the hydrophobic region at about 726–748 amino acid residues, close to the C-terminal end of the kexin-like protein from *P. brasiliensis*

Pbkex2 gene can be observed two sequences at positions 2806 (TACATA) and 2880 (TATTTA), that are similar to the TATATA sequence, the putative motif for 'efficiency element' (Wahle and Ruegsegger, 1999), as can be seen in Figure 2. The polyadenylation in mRNA from yeast occurs preferentially at PyA(n) sequences (Py = pyrimidine) about 20 nucleotides downstream from the 'positioning elements'. In mRNA from the *Pbkex2* gene the poly(A) tail is added at position 3053, as determinated by comparison of genomic and cDNA alignment sequences, exactly 32 nucleotides downstream from the putative 'positioning element'.

Analysis of Kex deduced amino acid sequence

The *Pbkex2* gene encodes an 842 amino acid protein, which shows a high degree of identity with kexin-like proteinases from the subtilases family and results in a 92 kDa predicted monomer. The highest identities observed were with kexin proteinases from *A. niger* (55%), *E. nidulans* (53%), *Kluyveromyces lactis* (52%), *C. albicans* (49%), and *Y. lipolytica* (48%). Analysis of the deduced amino acid sequence of *P. brasiliensis* kexin-like protein revealed the presence of five domains: (a) catalytic domain; (b) P domain; (c) Ser/Thr-rich domain; (d) transmembrane domain; and (e) cytoplasmic domain; which are found in kexin-like proteinases from other fungi. In Figure 3B we can observe the structural comparison among the kexin-like proteinases from *A. niger*, *E. nidulans*, *C. albicans*, *S. cerevisiae* and *P. brasiliensis* kexin-like protein. Figure 3C shows the representation of the predicted topology of Pbkex2 protein with only one transmembrane region (at amino acid residues 726–748), which was determinated using the computational program SOSUI. The first 20 amino acid residues comprise a signal peptide, which is probably cleaved at the RASV sequence, between A and S amino acid residues. The sequence of 21–127 amino acid residues probably corresponds to a propeptide. At the end of this propeptide sequence there are two basic residues (KR) that are present at cleavage site from kexin precursors (Figure 4). The catalytic domain, located at amino acid residues 167–481, also shows high identity to catalytic domains from other kexin-like proteinases (*A. niger*, 69%; *E. nidulans*, 66%; *Y. lipolytica*, 56%). In the catalytic domain of *P. brasiliensis*, kexin-like protein can be observed the presence of Asp (D), His (H) and Ser (S) catalytic triad (209, 247 and 419 amino acid residues) and a conserved Asn (N) residue at position 347, necessary to stabilize the oxyanion in the transitional state during the enzymatic cleavage by proteinases (Bryan *et al.*, 1986). A P domain-like sequence with 73% identity to the P domain from *A. niger* is present within 495–629 amino acid residues, which shows a RGD sequence, a reminiscent sequence for integrins recognition. The RGD sequence is found in furin/proprotein convertase (PC) and kexin-like

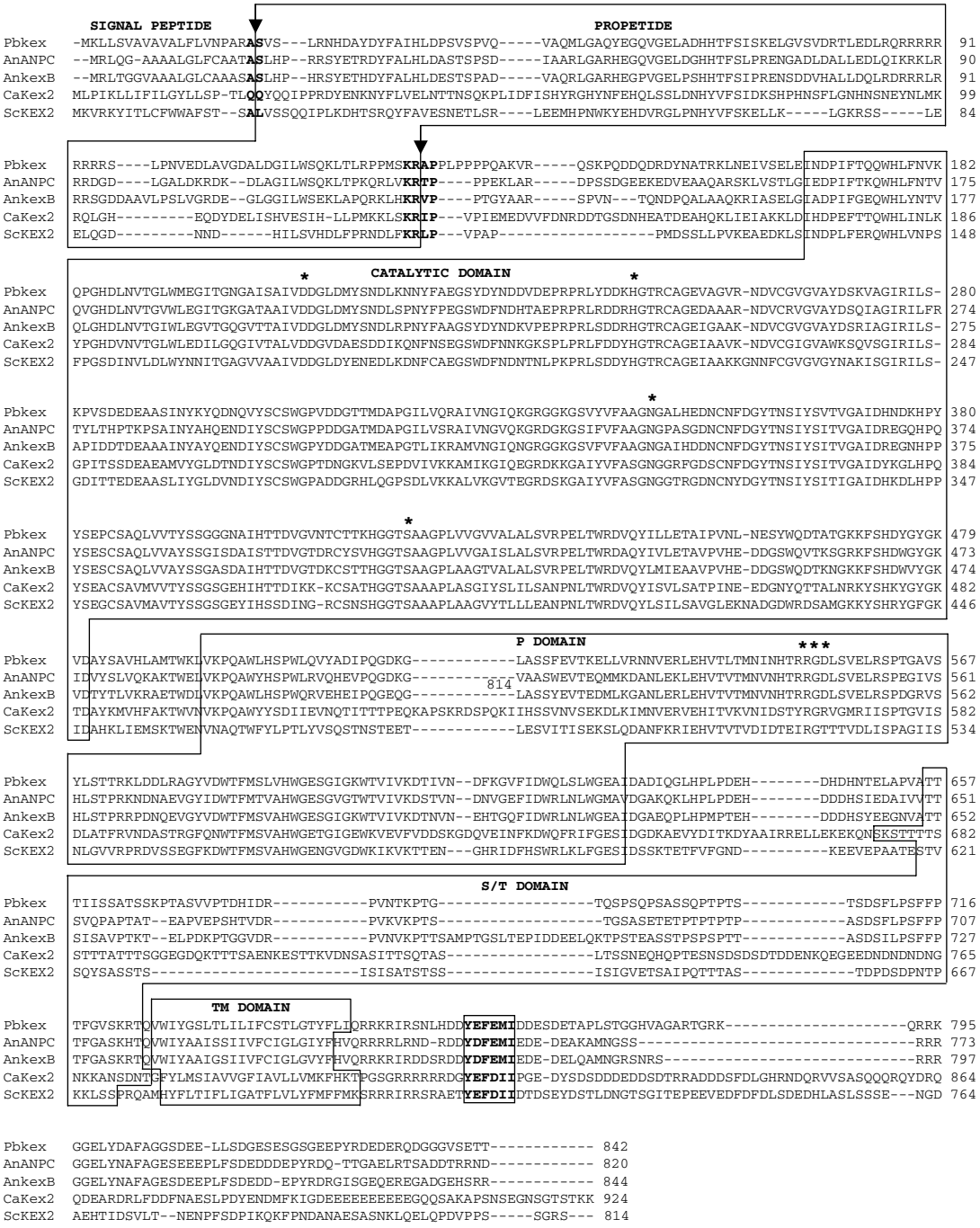

Figure 4. Multiple sequence alignment of fungal kexin-like proteinases. The figure compares amino acid sequences of: *P. brasiliensis* (*Pbkex2*, Accession No. AF486805); *E. nidulans* (AnANPC, Accession No. AF291662); *A. niger* (AnkexB, Accession No. Y18127); *C. albicans* (CaKex2, Accession No. AF022372); and *S. cerevisiae* (Sckex2, Accession No. M24201). Propeptide, catalytic, P, S/T and transmembrane (TM) domains are boxed. The catalytic Asp (D), Ser (S), His (H) and Asn (N) residues and RGD sequence are indicated by asterisks. The signal peptide and the propeptide cleavage sites are in bold and indicated by arrows. The Golgi retention signals are boxed and in bold

proteinases from *A. niger* and *E. nidulans*, but is not present in other kexin-like proteinases from *S. cerevisiae*, *C. albicans* and *K. lactis*. The kexin from fungi and PC7 from mammals have a C-terminal serine/threonine-rich (S/T) region, whose function is unknown. A similar S/T rich domain is identified between the P domain and transmembrane region in *P. brasiliensis* kexin-like protein (Figure 4). Finally, the presence of a consensus signal was observed (YEFEMI), a feature necessary to keep kexin-like proteinases in the Golgi network (Wilcox *et al*., 1992).

Dimorphism and proteinases

The kexin-like proteins of fungi are involved in several cellular processes, including dimorphism. In *C. albicans*, *KEX2* gene disruption produced several effects on cellular morphology and dimorphism (Newport and Agabian, 1997). The double mutant cells are aberrantly shaped, with some cells producing multiple buds and nuclei, considerably larger in size and tending to form aggregates. Chitin deposition was delocalized with focal deposition at the regions between cells, at bud scars and in cell extensions. These morphological defects were reversed when the strain was complemented with a plasmid containing wild-type *KEX2* gene from *C. albicans*. When submitted to dimorphic transition, *kex2/kex2* double mutant cells failed to produce hyphae, although the ability to form pseudohyphae was preserved. Komano and Fuller (1995) observed similar defects in chitin deposition in a *kex2* null mutant of *S. cerevisiae*, with their work also showing an actin delocalization. In *Y. lipolytica*, disruption/mutation of the *XPR6* (*KEX2* gene) resulted in a marked deficiency of hyphal formation, abnormal chitin deposition and cells aberrantly shaped, remaining attached after budding (Enderlin and Ogrydziak 1994; Richard *et al*., 2001). Kexin-like proteinases probably modify several proteins, including precursors of hyphae-specific proteins, or cell wall components (mannoproteins), or proteins with cell wall biogenesis activities [exo-β-(1–3)-glucanases] that prevent fungal dimorphism. However, as kexin-like proteinases have been involved in many cellular processes their precise mechanism of action in dimorphism remain unresolved.

Proteinase activity has been related to the virulence of some pathogenic fungi (De Bernardis *et al*., 2001). In *P. brasiliensis*, Carmona *et al*. (1995) characterized a SH-dependent serine proteinase activity, apparently involved in cleavage of components of the basal membrane. The function of kexin-like proteinase in virulence and/or dimorphism of *P. brasiliensis* are still to be elucidated.

Acknowledgements

We thank Dr Robert N. G. Miller for English review and M. F. L. Cesário and J. I. S de Santana for laboratory assistance. This work was supported by PACDT, CNPq and FUB, Brasília University, DF, Brazil. E. J. Venancio was supported by PICDT/ CAPES and University of Londrina (UEL), PR, Brazil.

References

Bergeron F, Leduc R, Day R. 2000. Subtilase-like pro-protein convertases: from molecular specificity to therapeutic applications. *J Mol Endocrinol* **24**: 1–22.

Bryan P, Pantoliano MW, Quill SG, Hsiao HY, Poulos T. 1986. Site-directed mutagenesis and the role of the oxyanion hole in subtilisin. *Proc Natl Acad Sci USA* **83**: 3743–3745.

Carmona AK, Puccia R, Oliveira MC, Rodrigues EG, Juliano L, Travassos LR. 1995. Characterization of an exocellular serine-thiol proteinase activity in *Paracoccidioides brasiliensis. Biochem J* **309**: 209–214.

De Bernardis F, Sullivan PA, Cassone A. 2001. Aspartyl proteinases of *Candida albicans* and their role in pathogenicity. *Med Mycol* **39**: 303–313.

Deminoff SJ, Santangelo GM. 2001. Rap1p requires Gcr1p and Gcr2p homodimers to activate ribosomal protein and glycolytic genes, respectively. *Genetics* **158**: 133–143.

Dutton JR, Johns S, Miller BL. 1997. StuAp is a sequence-specific transcription factor that regulates developmental complexity in *Aspergillus nidulans*. *EMBO J* **16**: 5710–5721.

Enderlin CS, Ogrydziak DM. 1994. Cloning, nucleotide sequence and functions of XPR6, which codes for a dibasic processing endoprotease from the yeast *Yarrowia lipolytica. Yeast* **10**: 67–79.

Ewing B, Hillier L, Wendl MC, Green P. 1998. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res* **8**: 175–185.

Franco M. 1987. Host–parasite relationships in paracoccidioidomycosis. *J Med Vet Mycol* **25**: 5–18.

Fuller RS, Sterne RE, Thorner J. 1988. Enzymes required for yeast prohormone processing. *Ann Rev Physiol* **50**: 345–362.

Gordon D, Abajian C, Green P. 1998. Consed: a graphical tool for sequence finishing. *Genome Res* **8**: 195–202.

Izacc SMS, Gomez FJ, Jesuino RSA, *et al*. 2001. Molecular cloning, characterization and expression of the heat shock protein 60 gene from the human pathogenic fungus *Paracoccidioides brasiliensis. Med Mycol* **39**: 445–455.

Jalving R, van de Vondervoort PJ, Visser J, Schaap PJ. 2000. Characterization of the Kexin-like maturase of *Aspergillus niger*. *Appl Environ Microbiol* **66**: 363–368.

Julius D, Brake A, Blair L, Kunisawa R, Thorner J. 1984. Isolation of the putative structural gene for the lysine–arginine-cleaving endopeptidase required for processing of yeast prepro-α-factor. *Cell* **37**: 1075–1089.

Kyte J, Doolittle RF. 1982. A simple method for displaying the hydropathic character of a protein. *J Mol Biol* **157**: 105–132.

Komano H, Fuller RS. 1995. Shared functions *in vivo* of a glycosyl-phosphatidylinositol-linked aspartyl protease, Mkc7, and the proprotein processing protease Kex2 in yeast. *Proc Natl Acad Sci USA* **92**: 10752–10756.

Langford CJ, Klinz FJ, Donath C, Gallwitz D. 1984. Point mutations identify the conserved, intron-contained TACTAAC box as an essential splicing signal sequence in yeast. *Cell* **36**: 645–653.

McCabe PM, Van Alfen NK. 1999. Secretion of cryparin, a fungal hydrophobin. *Appl Environ Microbiol* **65**: 5431–5435.

McEwen JG, Garcia AM, Ortiz BL, Botero S, Restrepo A. 1995. In search of the natural habitat of *Paracoccidioides brasiliensis*. *Arch Med Res* **26**: 305–306.

Miled C, Mann C, Faye G. 2001. Xbp1-mediated repression of CLB gene expression contributes to the modifications of yeast cell morphology and cell cycle seen during nitrogen-limited growth. *Mol Cell Biol* **21**: 3714–3724.

Mitchell AP. 1998. Dimorphism and virulence in *Candida albicans*. *Curr Opin Microbiol* **1**: 687–692.

Montenegro MR, Franco M. 1994. Pathology. In *Paracoccidioidomycosis*, Franco M, Lacaz CS, Restrepo-Moreno A, Del Negro G (eds). CRC Press: Boca Raton, FL; 131–150.

Newport G, Agabian N. 1997. KEX2 influences *Candida albicans* proteinase secretion and hyphal formation. *J Biol Chem* **272**: 28954–28961.

Nino-Vega GA, Munro CA, San-Blas G, Gooday GW, Gow NA. 2000. Differential expression of chitin synthase genes during temperature-induced dimorphic transitions in *Paracoccidioides brasiliensis*. *Med Mycol* **38**: 31–39.

Pedersen AG, Baldi P, Chauvin Y, Brunak S. 1999. The biology of eukaryotic promoter prediction — a review. *Comput Chem* **23**: 191–207.

Pereira M, Felipe MS, Brigido MM, Soares CM, Azevedo MO. 2000. Molecular cloning and characterization of a glucan synthase gene from the human pathogenic fungus *Paracoccidioides brasiliensis*. *Yeast* **16**: 451–462.

Pignede G, Wang HJ, Fudalej F, Seman M, Gaillardin C, Nicaud JM. 2000. Autocloning and amplification of LIP2 in *Yarrowia lipolytica*. *Appl Environ Microbiol* **66**: 3283–3289.

Richard M, Quijano RR, Bezzate S, Bordon-Pallier F, Gaillardin C. 2001. Tagging morphogenetic genes by insertional mutagenesis in the yeast *Yarrowia lipolytica*. *J Bacteriol* **183**: 3098–3107.

Sambrook J, Fritsch EF, Maniatis T. 1989. *Molecular Cloning: a Laboratory Manual*. Cold Spring Harbor Laboratory Press: New York.

Santoro N, Johansson N, Thiele DJ. 1998. Heat shock element architecture is an important determinant in the temperature and transactivation domain requirements for heat shock transcription factor. *Mol Cell Biol* **18**: 6340–6352.

Seidah NG, Chretien M. 1999. Proprotein and prohormone convertases: a family of subtilases generating diverse bioactive polypeptides. *Brain Res* **848**: 45–62.

Silva SP, Borges-Walmsley MI, Pereira IS, Soares CM, Walmsley AR, Felipe MS. 1999. Differential expression of an hsp70 gene during transition from the mycelial to the infective yeast form of the human pathogenic fungus *Paracoccidioides brasiliensis*. *Mol Microbiol* **31**: 1039–1050.

Silva SP, Felipe MSS, Pereira M, Azevedo MO, Soares CMA. 1994. Phase transition and stage-specific protein synthesis in the dimorphic fungus *Paracoccidiodes brasiliensis*. *Exp Mycol* **18**: 294–299.

Staab JF, Ferrer CA, Sundstrom P. 1996. Developmental expression of a tandemly repeated, proline-and glutamine-rich amino acid motif on hyphal surfaces on *Candida albicans*. *J Biol Chem* **271**: 6298–6305.

Venancio EJ, Kyaw CM, Mello CV, *et al*. 2002. Identification of differentially expressed transcripts in the human pathogenic fungus *Paracoccidioides brasiliensis* by differential display. *Med Mycol* **40**: 45–51.

Wahle E, Ruegsegger U. 1999. 3′-End processing of pre-mRNA in eukaryotes. *FEMS Microbiol Rev* **23**: 277–295.

Wickner RB, Leibowitz MJ. 1976. Two chromosomal genes required for killing expression in killer strains of *Saccharomyces cerevisiae*. *Genetics* **82**: 429–442.

Wilcox CA, Redding K, Wright R, Fuller RS. 1992. Mutation of a tyrosine localization signal in the cytosolic tail of yeast Kex2 protease disrupts Golgi retention and results in default transport to the vacuole. *Mol Biol Cell* **3**: 1353–1371.

Ye XS, Lee SL, Wolkow TD, *et al*. 1999. Interaction between developmental and cell cycle regulators is required for morphogenesis in *Aspergillus nidulans*. *EMBO J* **18**: 6994–7001.