

Secreted subtilisin gene family in *Trichophyton rubrum*

Olivier Jousson^a, Barbara Léchenne^a, Olympia Bontems^a, Bernard Mignon^b,
Utz Reichard^c, Jachen Barblan^d, Manfredo Quadroni^d, Michel Monod^{a,*}

^aDermatology Service (DHURDV), Centre Hospitalier Universitaire Vaudois, Lausanne, Switzerland

^bDepartment of Infectious and Parasitic Diseases, Faculty of Veterinary Medicine, University of Liège, Belgium

^cDepartment of Bacteriology, University of Göttingen, Germany

^dProtein Analysis Facility, Institute of Biochemistry, University of Lausanne, Epalinges, Switzerland

Received 10 December 2003; received in revised form 15 March 2004; accepted 10 June 2004

Available online 18 August 2004

Received by B. Dujon

Abstract

Secreted proteases constitute potential virulence factors of dermatophytes. A total of seven genes encoding putative serine proteases of the subtilisin family (*SUB*) were isolated in *Trichophyton rubrum*. Based on sequence data and intron–exon structure, a phylogenetic analysis of subtilisins from *T. rubrum* and other fungi revealed a presumed ancestral lineage comprising *T. rubrum SUB2* and *Aspergillus SUBs*. All other *SUBs* (*SUB1*, *SUB3–7*) are dermatophyte-specific and have apparently emerged more recently, through successive gene duplication events. We showed that two subtilisins, Sub3 and Sub4, were detected in culture supernatants of *T. rubrum* grown in a medium containing soy protein as a sole nitrogen source. Both recombinant enzymes produced in *Pichia pastoris* are highly active on keratin azure suggesting that these proteases play an important role in invasion of keratinised tissues by the fungus. The set of deduced amino acid sequences of *T. rubrum SUB* ORFs allowed the identification of orthologous Subs secreted by other dermatophyte species using proteolysis and mass spectrometry.

© 2004 Elsevier B.V. All rights reserved.

Keywords: Dermatophytes; Phylogeny; Secreted proteases; Recombinant proteins

1. Introduction

Trichophyton rubrum is the most commonly observed dermatophyte isolated from humans in European countries. This species is especially dominant in onychomycosis with

a prevalence of approximately 80% (Monod et al., 2002a). In nail infections, *T. rubrum* destroys keratin and forms channels and sizeable lacunae in the nail plates. These are considerably larger than the hyphae within them, implying the occurrence of extracellular enzymatic activity. Substantial serine and metalloprotease activities were secreted by dermatophytes in a soy protein growth medium and a five-member metalloprotease (*MEP*) gene family encoding fungalysins was isolated in *T. rubrum* as well as in *T. mentagrophytes* and *Microsporum canis* (Jousson et al., 2004). Although secreted metalloprotease activity was rather dominant, the serine protease activity was about 25–50% of the total activity of the culture supernatant as attested by inhibition of this fraction by PMSF. One 34.7 kDa serine alkaline protease (Sanyal et al., 1985) and two serine keratinolytic proteases with molecular masses of 93 and 71 kDa (Asahi et al., 1989) were isolated and characterized in *T. rubrum*. These proteases were shown to be active as dimers with subunits of 44 and 36 kDa, respectively.

Abbreviations: PCR, polymerase chain reaction; *SUB*/Sub, subtilisin gene/protein; *MEP*/Mep, metalloprotease gene/protein; *ALP*/Alp, alkaline protease gene/protein; *AOX1*, alcohol oxydase gene; ORF, open reading frame; aa, amino acid(s); nt, nucleotide(s); bp, base pairs; kDa, kilodalton; µg, microgram; ng, nanogram; MP, maximum parsimony; NJ, neighbor joining; ML, maximum likelihood; PAM, point accepted mutation; TBR, tree bisection recognition; GTR, general time reversible; G, gamma distribution shape parameter; I, proportion of invariable sites; LC, liquid chromatography; MS, mass spectrometry; PMSF, phenyl methyl sulphonyl fluoride; YNB, yeast nitrogen base; SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis; w/v, weight/volume ratio; pI, isoelectric point.

* Corresponding author. Tel.: +41-21-3140376; fax: +41-21-3140378.

E-mail address: Michel.Monod@chuv.hospvd.ch (M. Monod).

However, there is no reported amino acid sequencing of the N-terminus of these proteins which would have allowed their identification and characterization.

A gene family encoding serine proteases of the subtilisin family was recently isolated in *M. canis* (Descamps et al., 2002). One gene (*SUB3*) encoded a 31.5-kDa subtilisin previously isolated and characterized from the culture supernatant of *M. canis* grown in a keratin medium (Mignon et al., 1998). We describe here a seven-member gene family encoding subtilisins in *T. rubrum*. We show that two subtilisins are secreted in a soy protein medium. Both enzymes, expressed as recombinant proteins in *P. pastoris*, are highly active on keratine azure. The set of deduced amino acid sequences of *T. rubrum SUB* ORFs allowed orthologous Subs secreted by other dermatophyte species to be identified using proteolysis and mass spectrometry.

2. Materials and methods

2.1. Strains and plasmids

Clinical isolates of *T. rubrum*, *T. mentagrophytes*, *M. canis* and *Arthroderma benhamiae* from patients at the University Hospital, Lausanne (Switzerland) were used. All strains were identified on the basis of macroscopic appearance, microscopic examination of the cultures and partial 28S ribosomal DNA sequencing (Ninet et al., 2003). From each of the following species, one isolate of *T. rubrum* (LAU862-01), *T. mentagrophytes* (LAU2434-02), *M. canis* (LAU709-03) and *A. benhamiae* (LAU2352-02) was chosen for protein extract analysis.

Escherichia coli LE392 was used for the propagation of the bacteriophage λ EMBL3 (Promega). All plasmid sub-cloning experiments were performed in *E. coli* XL-1 blue with plasmids pMTL21 (Chambers et al., 1988) and pUC19 (Sambrook et al., 1989). *SUB* cDNAs were cloned in plasmid pKJ113 (Borg-von Zepelin et al., 1998) and *Pichia pastoris* GS115 (Invitrogen) was used for transformation and production of recombinant proteins.

2.2. Growth media

T. rubrum, *T. mentagrophytes*, *A. benhamiae* and *M. canis* were grown on Sabouraud agar and liquid medium (Bio-Rad) or, to promote production of proteolytic activity, in liquid medium containing 0.2% soy protein (Supro 1711, Protein Technologies International), as previously described (Jousson et al., 2004).

2.3. SUB gene cloning

Previously constructed *T. rubrum* λ EMBL3 genomic and pSport6 cDNA libraries (Jousson et al., 2004) were used. Recombinant plaques (2×10^4) of the genomic libraries of *T. rubrum* were immobilized on GeneScreen nylon mem-

branes (NEN Life Science products). The filters were hybridized with ^{32} P-labelled DNA fragments under low-stringency conditions (Monod et al., 1994). All positive plaques were purified and the associated bacteriophage DNAs were isolated as described previously (Grossberger, 1987). Agarose gel electrophoresis of restricted recombinant bacteriophage λ EMBL3 DNA, Southern blotting and sub-cloning of hybridizing fragments from bacteriophages into pMTL21 or pUC19 were performed using standard protocols (Sambrook et al., 1989). DNA sequencing was performed by Microsynth (Balgach, Switzerland).

T. rubrum cDNAs were obtained by PCR using 200 ng of DNA prepared from 10^6 clones of the cDNA library. PCR were performed with homologous primers designed on DNA sequences of the different *SUB* genes.

2.4. Phylogenetic analyses

Nucleotide and amino acid dermatophyte *SUB* sequences were aligned using Clustal W (Thompson et al., 1994) as implemented in BioEdit Sequence Alignment Editor software (Hall, 1999). Sequences of genes encoding other fungal serine proteases were incorporated in the alignment, including alkaline proteinases from *Aspergillus fumigatus* (Accession numbers: *ALP1*: Z11580; *ALP2*: Y13338), *A. niger* (*pepC*: M96758; *pepD*: L19059), *A. oryzae* (*alpA*: S75278), *A. nidulans* (*prtA*: L31778), allergens from different *Trichophyton* species (*T. rubrum* Tri r 2: AF082515; *T. schoenleinii* Tri m 2: AJ430843; *T. mentagrophytes* Tri m 2: AJ430840), as well as sequences of vacuolar serine proteases from *Penicillium oxalixum* (AF243425), *P. chrysogenum* (AF263454), and *Saccharomyces cerevisiae* (P09232). Phylogenetic analyses of *SUB* genes and proteins were performed in PAUP* v4.0b10 (Swofford, 1998). DNA and protein sequences were analysed using MP, NJ, and ML methods. Modeltest 3.06 (Posada and Crandall, 1998) was used to select the appropriate model of substitution for ML and NJ analyses of nucleotide sequences. The Dayhoff PAM model of protein evolution was used to compute the distances between the amino acid sequences using the PROTDIST program implemented in BioEdit. Analyses were performed using heuristic search with TBR branch swapping algorithm. The reliability of internal branches was assessed using the bootstrap method (Felsenstein, 1988), with 1000 replicates. Phylogenetic trees were edited using TreeView (Page, 1996).

Phylogenetic analyses based on exon–intron structure of *SUB* genes were also performed. The first procedure used an intron presence/absence matrix as an input, which was analysed using MP. The second procedure consisted of a measure of gene structure similarity (Betts et al., 2001) that uses the presence, location, and phase of introns to estimate the exon–intron similarity.

$$S_G(a, b) = \frac{1}{2N_{\max}} \sum_{i=1}^{N_{\text{equiv}}} \left(\frac{1}{1 + e^{\gamma(di-\delta)}} + \varphi(a_i, b_i) \right)$$

S_G is the similarity measure for the two proteins (a and b); $N_{\max}$ is the largest number of introns found in either protein; N_{equiv} is the number of equivalent (homologous) introns; a_i and b_i are the i th equivalent intron positions in the two proteins; d_i is the difference in position of the introns within the two proteins (in amino acids); $\varphi(a_i, b_i)$ is 1 if the intron phases are the same and 0 if they are different; γ and δ are constants (0.2 and 30) optimised such that the sigmoid function is insensitive to small changes in intron positions (± 10 residues) (Betts et al., 2001). A matrix of distances calculated by this method was used to reconstruct a tree by NJ.

2.5. Production of recombinant Subs

Expression plasmids were constructed by cloning cDNA PCR products in the multiple cloning site of the *E. coli*–*P. pastoris* shuttle vector pKJ113. The PCR products were purified using a PCR purification kit (Roche Diagnostics) and digested by restriction enzymes for which a recognition site was previously designed at the 5' extremity of the primers (Table 1). *P. pastoris* GS115 (Invitrogen) was transformed by electroporation with 10 μ g of plasmid DNA linearized by *Eco*RI or *Sma*I. Transformants selected on histidine-deficient medium [1 M sorbitol, 1% (w/v) dextrose, 1.34% (w/v) yeast nitrogen base (YNB) without amino acids, 4×10^{-5} % (w/v) biotin, 5×10^{-3} % amino acids (i.e. 5×10^{-3} % (w/v) of each L-glutamic acid, L-methionine, L-lysine, L-leucine, L-isoleucine), 2% (w/v) agarose] were screened for insertion of the construct at the *AOX1* site on minimal methanol plates [1.34% (w/v) YNB without amino acids, 4×10^{-5} % (w/v) biotin, 0.5% (v/v) methanol, 2% (w/v) agarose]. The transformants unable to grow on media containing only

methanol as carbon source were assumed to contain the construct at the correct yeast genomic location by integration events in the *AOX1* locus displacing the *AOX1* coding region. They were grown to near saturation (OD=20 at 600 nm) at 30 °C in 10 ml of glycerol-based yeast media (0.1 M potassium phosphate buffer at pH 6.0, containing 1% (w/v) yeast extract, 2% (w/v) peptone, 1.34% (w/v) YNB without amino acids, 1% (v/v) glycerol and 4×10^{-5} % (w/v) biotin). Cells were harvested and resuspended in 2 ml of the same medium with 0.5% (v/v) methanol instead of glycerol and incubated for 2 days. The supernatant was then harvested and tested for protein production on SDS-PAGE gels. Salts and small molecular weight solutes were removed from 2.5 ml of *P. pastoris* culture supernatant by passing through a PD10 column (Pharmacia) using 20 mM Tris–HCl buffer, pH 7.5 before testing for proteolytic activity. Protein concentrations were measured by the method of Bradford using a commercial reagent (Bio-Rad). Supernatant of *P. pastoris* GS115 grown in the same conditions was used as a negative control for comparison.

2.6. Proteolytic assays

Subtilisin Carlsberg (*Bacillus licheniformis* subtilisin A) and proteinase K were from Sigma. *A. fumigatus* and *A. oryzae* alkaline proteases were prepared as previously described (Monod et al., 1991).

The proteolytic activity was measured in 50 mM Tris–HCl buffer at pH between 7.0 and 9.5 using resorufin-labeled casein (Roche Diagnostics), azocollagen (Sigma) and elastin-Congo red (Sigma) as described previously (Monod et al., 1991; Jousson et al., 2004). Using Azoalbumin, the reaction mixture contained 0.1% substrate, 50

Table 1
Primers used for amplification of *T. rubrum* SUBs cDNAs

Gene	Amplification primers ^a	Orientation	Encoded amino acid sequence ^b	PCR product ^c (with cloning sites)
SUB1	GTTC↓TCGAGTTAATGCCGCCAAATCCTGTCT	5'–3'	(R)(V)NAAQILS	(47) <i>Xho</i> I
	GTA↓GATCCTTAGTTCCAGAAGCTGCTAAAC	3'–5'	FSSFVN*	(1515) <i>Bam</i> HI
SUB2	GGTTC↓TCGAGACCTCGTCCACAGCCTGAGCCG	5'–3'	(R)DLAPQPEP	(45) <i>Xho</i> I
	CTTG↓GATCCTCAGTAATACTTGGGCAGTTTGC	3'–5'	KLPKY*	(1275) <i>Bam</i> HI
SUB3	CTTG↓TCGACTTGATGCCGCGCTTCTCCACAAC	5'–3'	(R)(L)DARAFFHN	(45) <i>Sal</i> I
	CTTG↓GATCCTTATCGTCCACTTCCGTTGTAGAGG	3'–5'	LYNGSGR*	(1194) <i>Bam</i> HI
SUB4	CTTG↓TCGACTTGATGCCGCGCAGTCTTCAAG	5'–3'	(R)(L)DARAVFK	(45) <i>Sal</i> I
	CTTG↓GATCCTACTGGCCACTTCCGTTGTAGAG	3'–5'	LYNGSGQ*	(1200) <i>Bam</i> HI
SUB5	GTTC↓TCGAGTTGCACAGATCTTAAGTGTCGCC	5'–3'	(R)(V)AQILSVP	(54) <i>Xho</i> I
	CTTG↓GATCCTTAACGGCCACTGCCATTATAAAGC	3'–5'	LYNGSGR*	(1191) <i>Bam</i> HI
SUB6	GTGC↓TCGAGATGGTGCTAGAATCCTTGAGGCCGGT	5'–3'	(R)(D)GARILEAG	(51) <i>Xho</i> I
	GTTGC↓GGCCGCTTATTGCGCGTGCCGTTGTA	3'–5'	YNGSGK*	(1239) <i>Not</i> I
SUB7	GTTGC↓TCGAGCTGAGATCTTGAGACTCGCGCT	5'–3'	(R)(A)EILETRA	(57) <i>Xho</i> I
	GTTG↓GATCCTTACATGCCAGATCGGTTGTTGATGAGCTTGC	3'–5'	KLINNGSGM*	(1203) <i>Bam</i> HI

Amplification products were cloned in *E. coli*–*P. pastoris* shuttle vector pKJ113.

^a Vertical arrows indicate the cleavage site of restriction endonucleases; underlined nucleotides represent the recognition sequence.

^b In parentheses are shown amino acids encoded by the recognition site sequence and added to the N-terminal extremity of Sub prosequences; asterisks indicate stop codons.

^c The numbers in parentheses represent the nucleotide position of amplification primers on SUB cDNAs.

mM Tris–HCl buffer and 1.0 µg of enzyme. After incubation at 37 °C during 30 min, the undigested substrate was precipitated by trichloroacetic acid (5% final concentration) and separated from the supernatant by centrifugation. The absorbance of the mixture was measured at 400 nm. Keratin degradation was measured using keratin azure as a substrate (Sigma). The reaction mixture contained 10 µg of substrate, Tris–HCl buffer (50 mM at different pH values), and 2.5 µg of enzyme in a total volume of 500 µl. The degradation of keratin azure was estimated by measuring the absorbance at 595 nm of the supernatant after centrifugation. The assays were run in duplicate. For practical purposes, one arbitrary unit (U) of proteolytic activity was defined as that producing an absorbance of 0.01 per min.

2.7. Protein extracts analysis

Secreted proteins were precipitated from dermatophyte culture supernatants using trichloroacetic acid and purified using PlusOne SDS-PAGE Clean-Up Kit (Amersham). Extracts were analysed by SDS-PAGE (Laemmli, 1970) with a separation gel of 8.5% polyacrylamide. Gels were stained with Coomassie brilliant blue R-250 (Bio-Rad). Measures of band molecular mass and optical density were performed using the TotalLab software (Nonlinear Dynamics).

2.8. Protein identification by LC-MS/MS

Coomassie blue stained bands were excised from the SDS-PAGE gels and transferred to special 96-well plates (Perkin Elmer Life Sciences, Cambridge, United Kingdom). In-gel proteolytic cleavage with sequencing grade trypsin (Promega) was performed automatically in the robotic workstation Investigator ProGest™ (Perkin Elmer Life Sciences) according to the protocol of Wilm et al. (1996). Supernatants containing proteolytic peptides were analysed by LC-MS/MS on a SCIEX QSTAR Pulsar (Concord, Ontario, Canada) hybrid quadrupole-time of flight instrument equipped with a nanoelectrospray source and interfaced to an LC-Packings Ultimate (Amsterdam, Holland) HPLC system. Peptides were separated on a PepMap reversed-phase capillary C18 (75 µm ID × 15 cm) column at a flow rate of 200 nl min⁻¹ along a 52-min gradient of acetonitrile (0–40%). The instrument controlling Analyst software was used to perform peak detection and automatically select peptides for collision-induced fragmentation. Collections of non-interpreted collision-induced fragmentation spectra were bundled and used for searching an in-house built database of *T. rubrum* subtilisin sequences with the Mascot software (www.matrixscience.com) (Perkins et al., 1999). Only peptide sequences ranked as statistically significant were taken into account, and all peptide hits were manually examined for validation.

3. Results

3.1. SUB gene cloning

PCR products of *M. canis* SUB1, SUB2 and SUB3 genes were used as a probe to screen 2×10^4 plaques of the λEMBL3 *T. rubrum* genomic library corresponding to about 10 genome equivalents of the fungi. Restriction fragments hybridizing with the *M. canis* SUB1–3 probe were identified by Southern blotting, subcloned in pUC19 or pMTL21 and sequenced. Translated blast searches (<http://www.ncbi.nlm.nih.gov/BLAST/>) were performed for all sequences obtained and revealed five different sequences encoding parts of putative Sub proteins in *T. rubrum*. A preliminary phylogenetic analysis suggested that three sequences were the orthologs of *M. canis* SUB1, SUB2, and SUB3. The two remaining sequences apparently corresponded to new SUB genes and were subsequently named SUB4 and SUB5. However, none of these sequences corresponded to that of the previously cloned cDNA of Tri r 2 antigen encoding a putative subtilisin of *T. rubrum* (Woodfolk et al., 1998). Therefore, in a second step, 2×10^4 plaques of the genomic library were screened with a Tri r 2 PCR product. In addition to Tri r 2, one new *T. rubrum* gene encoding a putative Sub protein was found. These new SUB genes were named SUB6 and SUB7, respectively. No further new SUB paralogs were found in a third screening performed with the seven *T. rubrum* SUBs as probes. As a final result, a total of seven genes (SUB1–7) putatively encoding Sub proteins were found in *T. rubrum*.

3.2. Characterization of dermatophyte SUB genes

In order to determine the intron–exon structure of *T. rubrum* SUBs, primers pairs were used to amplify total DNA from a pool of 10^6 clones of the *T. rubrum* cDNA library as a target. The sequencing of PCR products and the comparison with genomic DNA sequences revealed three types of intron–exon structure among *T. rubrum* SUBs, which were compared with those of *Aspergillus* SUBs and proteinase K from *Tritirachium album* (Fig. 1). *T. rubrum* SUB1, SUB2, and SUB3–7 have two, four and three introns, respectively. The alignment of cDNA and protein sequences enabled us to determine the equivalence (homology) of introns among *T. rubrum* and *Aspergillus* SUBs. The phase of all homologous introns was conserved (Fig. 1).

The alignment of deduced translation products showed that all *T. rubrum* Sub sequences display the catalytic triad (for instance Asp¹⁵⁸/His¹⁹⁰/Ser³⁴⁵ in Sub1) and the motifs around these residues characteristic of proteinases of the subtilisin family (Rawlings and Barrett, 1994). Among *T. rubrum* Subs, the aa sequence identity ranged from 29.4% (between Sub2 and Sub6) to 67.4% (between Sub3 and Sub4). The previously characterized N-terminus of mature Sub3 from *M. canis* (Mignon et al., 1998) and the analysis

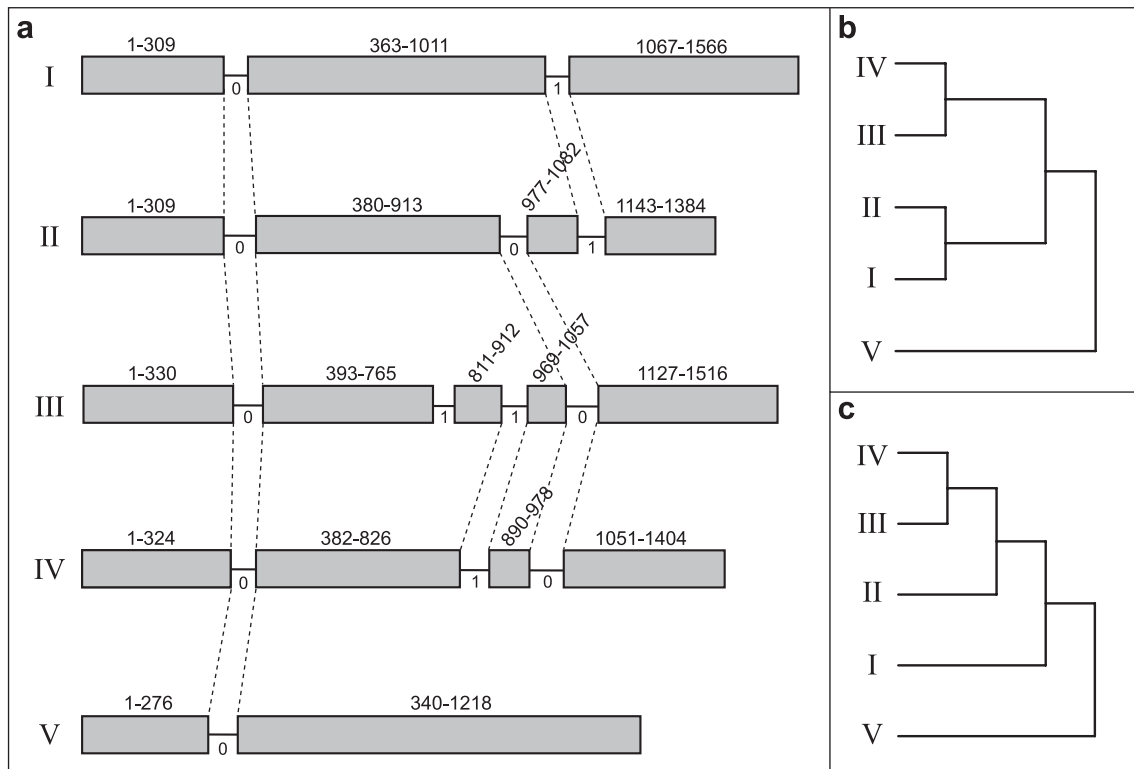


Fig. 1. (a) Intron–exon structure of *T. rubrum* and *Aspergillus* *SUB* genes. Roman numbers indicate the different types. Type I is for *T. rubrum* *SUB1*; type II: *T. rubrum* *SUB3–7*; type III: *T. rubrum* *SUB2*; type IV: *A. fumigatus* *ALP1*, *A. oryzae* *alpA*, *A. nidulans* *prtA*, and *A. niger* *pepD*. The structure of *T. album* proteinase K (type V) is shown for comparison. Dashed lines indicates homologous introns. Numbers above sequences correspond to the nucleotide positions. Numbers below sequences indicate the intron phase. Intron phases are defined as the relative positions of introns falling between codons (phase 0) or within a codon after the first (phase 1) or second (phase 2) nucleotide, respectively. (b) MP phylogenetic tree constructed on intron presence/absence. (c) NJ phylogenetic tree based on gene structure similarity.

of the sequence alignment suggested that dermatophyte Subs would be synthesized as preproteins with a propeptide of 96 to 103 amino acid residues. The mature domain generated after cleavage of the prosequences predicted a molecular mass ranging from 28.2 to 39.7 kDa (Table 2).

3.3. Phylogenetic analyses

The analysed *SUB* nucleotide data set comprised 1338 sites, among which 1051 are informative (78.5%). The amino acid data set consisted of 430 sites, among which

320 are informative (74.4%). Hierarchical likelihood ratio tests indicated that GTR+G+I was the most appropriate model of nucleotide substitution for subsequent analyses. Settings for the GTR+G+I model were as follows: base frequencies of 0.2058 (A), 0.3271 (C), 0.2549 (G), 0.2122 (T); substitution rates of 1.486 (A–C), 3.477 (A–G), 1.506 (A–T), 1.368 (C–G), 5.657 (C–T), 1.000 (G–T); a gamma distribution shape parameter (α) of 1.7943; 10.57 % of invariable sites. The analyses of subtilisin nucleotide and amino acid sequences using NJ, MP and ML phylogenetic methods produced the same tree topology (Fig. 2). Subtilisins isolated from dermatophytes appeared to be monophy-

Table 2
Main the characteristics of *SUB* genes and deduced translation products from *T. rubrum*

Gene	Genelength (nt)	Number of introns	GC content (%)	Preproprotein (aa)	Signal sequence (aa)	Proprotein (aa)	Mature domain of the protein (aa)	Theoretical molecular mass of the mature domain (kDa)	Calculated pI (preproprotein / mature domain)	GenBank accession number
<i>SUB1</i>	1628	2	55.8	504	19	97	388	39.720	5.57/4.98	AY343499
<i>SUB2</i>	1539	4	54.4	424	19	103	302	32.573	9.24/9.53	AY343500
<i>SUB3</i>	1415	3	54.5	397	19	97	281	28.324	8.71/9.25	AY343501
<i>SUB4</i>	1385	3	53.7	399	19	99	281	28.858	8.95/9.54	AY344481
<i>SUB5</i>	1384	3	50.2	396	20	96	280	29.021	9.67/10.28	AY344482
<i>SUB6</i>	1433	3	50.9	412	21	101	290	29.315	6.01/6.08	AF420485
<i>SUB7</i>	1387	3	55.4	400	21	98	281	28.186	5.52/4.47	AF407184

Theoretical molecular mass of the mature domain and the pI were calculated using VectorNTI suite 8 (InforMax).

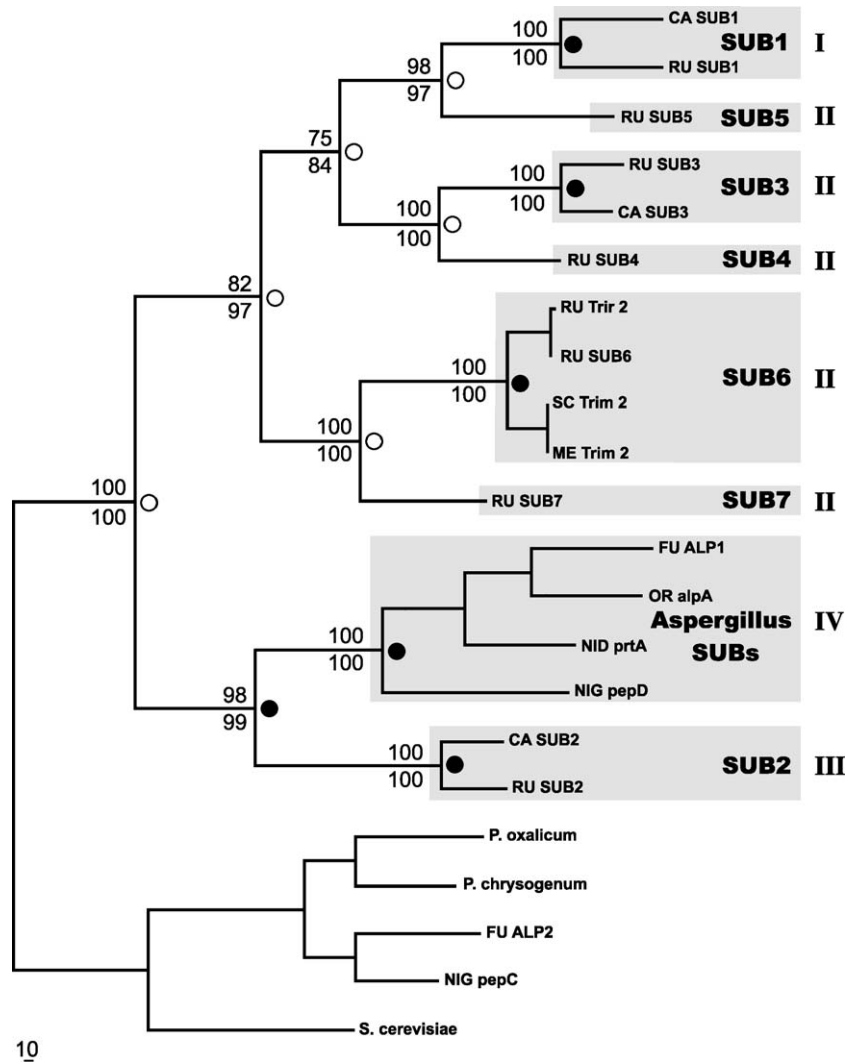


Fig. 2. Maximum Parsimony (MP) phylogenetic tree inferred from Sub amino acid sequences. Sequences of each Sub type are shown in shaded boxes. Abbreviations are: CA: *M. canis*; RU: *T. rubrum*; SC: *T. schoenleinii*; ME: *T. mentagrophytes*; FU: *A. fumigatus*; OR: *A. oryzae*; NID: *A. nidulans*; NIG: *A. niger*. Bootstrap values (1000 replicates) are given for amino acid (above branches) and nucleotide (below branches) sequences. Scale bar: number of steps. Open circles at nodes indicate gene duplication events producing paralogs; solid circles indicate speciation events producing orthologs. Roman numbers at right represent the intron–exon structure types from Fig. 1.

letic using vacuolar serine proteases of *Penicillium* spp., *Aspergillus* spp., and *S. cerevisiae* as an outgroup. The most basal and presumably ancestral clade of the ingroup includes *Aspergillus* secreted Subs, known as alkaline proteases (Alp) (e.g. Cheevadhanarak et al., 1991; Jaton-Ogay et al., 1992), as well as *T. rubrum* and *M. canis* Sub2.

Phylogenetic analyses based on intron data corroborated the relationships obtained with sequence data. The trees were arbitrarily rooted with intron–exon structure type V (from *T. album* proteinase K) (Fig. 1). The tree based on presence/absence of introns consisted of two sister groups including type I and II on one hand, and types III and IV on another hand (Fig. 1b). The tree based on intron structure showed an asymmetrical branching of types I and II, however, the group including types III and IV was recovered (Fig. 1c).

3.4. Subs identification by protein digestion and mass spectrometry

Among several tested protein sources, soy proteins were found to be particularly favourable for promoting proteolytic activity of dermatophytes (data not shown). The protein profile was similar for all isolates of a given species (Jousson et al., 2004). The proteins secreted by *T. rubrum*, as well as those of three other dermatophyte species for comparison (*T. mentagrophytes*, *A. benhamiae* and *M. canis*), were analysed using proteolysis and mass spectrometry (Table 3). *T. rubrum* Sub3 and Sub4 were identified as 33- and 31-kDa proteins, respectively, and the previously characterized *M. canis* Sub3 (Mignon et al., 1998) was unambiguously identified as a 33-kDa protein. Putative orthologous proteins of *T. rubrum* Sub3 showing

Table 3

Protein sequences matched to gel bands from Fig. 3 and their MASCOT scores

<i>T. rubrum</i>			<i>A. benhamiae</i>			<i>T. mentagrophytes</i>			<i>M. canis</i>		
Band	Protein	Score	Band	Protein	Score	Band	Protein	Score	Band	Protein	Score
–	–	–	Be40	Sub7	144	–	–	–	–	–	–
–	–	–		Sub3	54	–	–	–	–	–	–
–	–	–	Be37	Sub7	302	Me37	Sub3	60	–	–	–
–	–	–		Sub3	207	–	–	–	–	–	–
Ru33	Sub3	359	Be33	Sub3	458	Me33	Sub3	132	Ca33	Sub3	75
–	–	–	Be32	Sub3	331	–	–	–	–	–	–
Ru31	Sub4	450	Be31	Sub3	119	Me31	Sub4	52	–	–	–

The MASCOT score is defined as $-10 \cdot \log_{10}(P)$, where P is the absolute probability that the observed match is a random event. MASCOT also calculates a confidence threshold which corresponds to a 95% probability of correctness ($p=0.05$). For all spectral matching performed in this study, the threshold was at 12. All the matches reported were considerably above this value and all the peptide matches added up to calculate the score were manually validated.

compatible peptide profiles appeared to be the dominant Subs secreted in soy medium by the four dermatophyte investigated in this study. The presence of several bands corresponding to Sub3 in *A. benhamiae* and *T. mentagrophytes*

protein extracts (Fig. 3, Table 3) suggests the occurrence of post-translational protease processing, as previously demonstrated for *A. fumigatus* secreted alkaline protease (Kunert, 2001). A putative orthologous protein of *T. rubrum* Sub4 was detected in *T. mentagrophytes* but not in *A. benhamiae* and *M. canis* (Table 3). Interestingly, *A. benhamiae* 37- and 40-kDa bands showed compatible peptide profiles with that of *T. rubrum* Sub7. No other Subs were detected among the remaining 30- to 55-kDa proteins secreted by the four dermatophyte species in soy protein medium.

To demonstrate that *A. benhamiae* really secretes a Sub7 protein and to estimate the degree of aa sequence conservation of Sub orthologs, we amplified by PCR and sequenced *A. benhamiae* SUB genes. Oligonucleotides were designed on the base of the *T. rubrum* SUBs and *A. benhamiae* genomic DNA was used as a target. Seven *A. benhamiae* genes with deduced translation products highly similar to those of *T. rubrum* Sub1–7 (91–97% aa sequence identity) were found (Table 4).

3.5. Production and characterization of recombinant *T. rubrum* Subs

T. rubrum Sub3 and Sub4 were produced as recombinant proteins from *P. pastoris* with a yield 40 µg/ml of culture supernatant. Recombinant Sub3 and Sub4 had the same electrophoretic mobility (31- and 33-kDa protein) as the native enzymes secreted by *T. rubrum* in soy protein

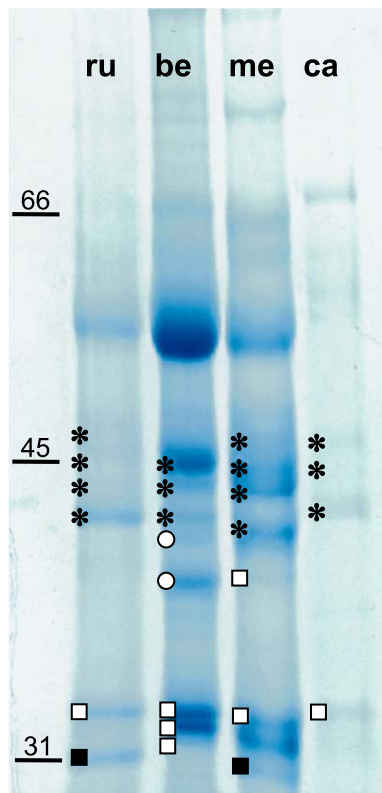


Fig. 3. Protein electrophoretic profiles from culture supernatant of *T. rubrum* (ru), *A. benhamiae* (be), *T. mentagrophytes* (me) and *M. canis* (ca). The proteins of 20 ml of dermatophyte culture supernatant were precipitated using trichloroacetic acid, purified, and resuspended in a total volume of 40 µl loading buffer. The 8.5% polyacrylamide SDS-PAGE gel was stained with Coomassie brilliant blue. Molecular mass markers are shown on the leftmost lane (bovine serum albumin, 66 kDa; ovalbumin, 45 kDa; bovine carbonic anhydrase, 31 kDa). Bands marked by a square or a circle correspond to Sub proteins identified by proteolysis and mass spectrometry (see Table 3). Open squares, solid squares and open circles correspond to Sub3, Sub4 and Sub7, respectively. None of the other bands produced a significant match with the *T. rubrum* Sub sequence database. Asterisks indicate Mep proteins (see Jousson et al., 2004).

Table 4

Pairwise amino acid sequence comparison between *T. rubrum* and *A. benhamiae* Subs

Gene	Sequence identity	Sequence similarity	<i>A. benhamiae</i> GenBank accession numbers
<i>SUB1</i>	0.96	0.97	AY437858
<i>SUB2</i>	0.92	0.95	AY437853
<i>SUB3</i>	0.95	0.99	AY437854
<i>SUB4</i>	0.93	0.99	AY437855
<i>SUB5</i>	0.97	0.99	AY437856
<i>SUB6</i>	0.91	0.94	AY437857
<i>SUB7</i>	0.96	0.98	AY437852

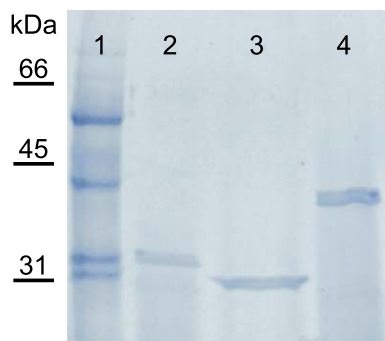


Fig. 4. Protein electrophoretic profiles from culture supernatant of *T. rubrum* (lane 1) and of *T. rubrum* recombinant Sub3, Sub4 and Sub5 (lanes 2–4) produced in *P. pastoris*. The gel was stained with Coomassie brilliant blue R-250. Molecular mass markers are shown on the leftmost lane. Ten microliters of *P. pastoris* culture supernatant were loaded in lanes 2–4.

medium (Fig. 4). Both enzymes were active in Tris–HCl buffer between pH 7.0 and 9.5 on different protein substrates (Fig. 5) with a broad peak of optimum activity between pH 7.5 and 9.0. The specific activity measured using resorufin-labeled casein at pH = 8.0 was 30 and 40 U/ μ g for recombinant Sub3 and Sub4, respectively.

Substrate specificity of Sub3 and Sub4 was compared to that of other proteases of the subtilisin family. In the presence of 1% β -mercaptoethanol, the measured enzyme activity on

keratin azure of each subtilisin was multiplied by a factor of 2 to 3. Sub4 was the most active enzyme on keratin azure. Sub3 showed also an appreciable activity on keratin azure comparable to that of subtilisin Carlsberg and proteinase K. The hair structure of 5 mg of keratin azure in 0.5 ml of Tris–HCl buffer was totally disrupted by 5 μ g of either Sub3 or Sub4 after 72 h of incubation at 37 °C. In comparison to subtilisin Carlsberg and proteinase K, Sub3 and Sub4 were less active on other protein sources, especially elastin, raising specificity towards keratinised substrates.

A 39-kDa recombinant protein was obtained with *T. rubrum* Sub5 but this protein was not proteolytically active on resorufin-labeled casein. No recombinant proteins were obtained for *T. rubrum* Sub1, Sub2, Sub6 and Sub7.

4. Discussion

A total of seven genes encoding serine proteases of the subtilisin family (Sub) were isolated in *T. rubrum*. Two subtilisins, Sub3 and Sub4, were detected in culture supernatants of *T. rubrum* grown in a medium containing soy protein as a sole nitrogen source. With an optimum pH of activity between 8.0 and 9.0, the previously isolated *T. rubrum* 34.7 kDa serine alkaline protease (Sanyal et al., 1985) and serine keratinolytic proteases active as dimers

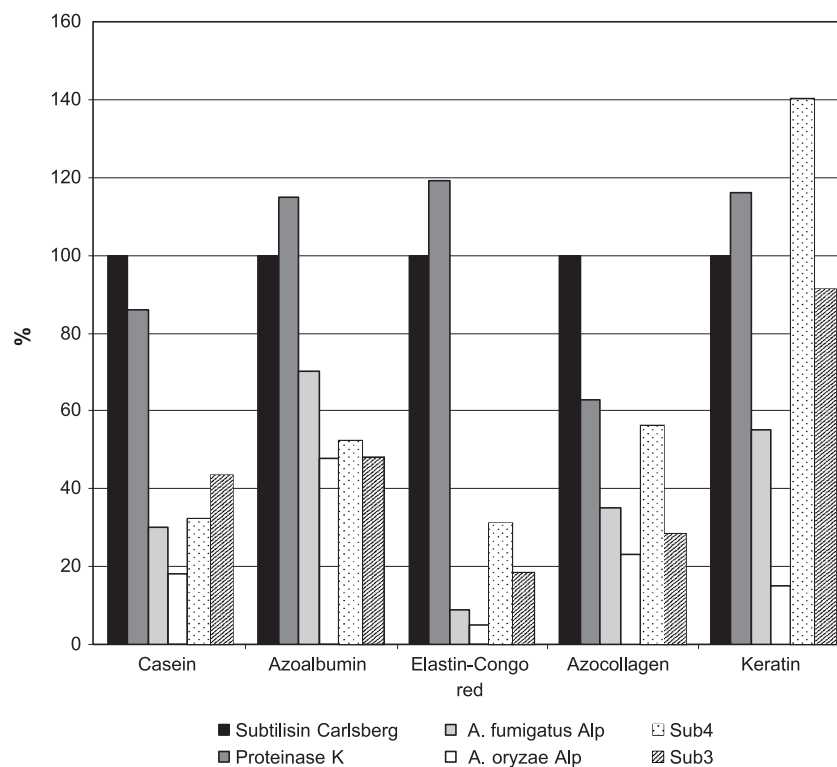


Fig. 5. Activities of *T. rubrum* Sub3 and Sub4, proteinase K, *A. fumigatus* and *A. oryzae* Alps relatively to subtilisin Carlsberg. Assays of activity were performed as described in Materials and methods at optimal pH of activity (9.0 for *A. fumigatus* Alp and 8.0 for the other proteases). The measured OD values were converted in amount of substrate digested by 1 μ g of enzyme per units of time. Resorufin-labeled casein (48 μ g), azoalbumin (17 μ g), elastin-Congo red (23 μ g), and azocollagen (330 μ g) were digested per minute by 1 μ g subtilisin Carlsberg, while 32 μ g keratin azure was digested per hour. These values were equated to 100% of activity.

with subunits of 36 or 44 kDa (Asahi et al., 1989) may be Sub3 for which the molecular mass was estimated as 33 kDa on SDS-PAGE. Recombinant Sub3 and Sub4 were more active on keratin azure in presence of β -mercaptoethanol. The cystine disulfide bridges, which are an important structural feature of native keratin, prevented the proteolytic degradation of the most compact keratinous tissues. Therefore, the additional cleavage of these disulfide bonds by β -mercaptoethanol made keratin azure more accessible for Sub3 and Sub4. In vivo, as demonstrated with *Microsporium gypseum*, the cleavage of these disulfide bonds occurs by sulfitolysis following secretion of sulfites by the fungus (Kunert, 1982; Kunert and Truper, 1986). Sub3 and Sub4 specificity towards keratin azure (Fig. 5) may reflect the high degree of specialisation of dermatophytes which develop only in keratinised tissues.

Phylogenetic analyses of *T. rubrum* subtilisin genes and proteins using sequences of vacuolar fungal serine proteases as an outgroup revealed a basal, presumably ancestral clade including *M. canis* and *T. rubrum* SUB2 and their putative orthologs from *A. fumigatus* (ALP1), *A. oryzae* (alpA), *A. nidulans* (prtA) and *A. niger* (pepD) (Fig. 2). Further phylogenetic analyses revealed that none of the other fungal serine proteases available in GenBank branched within the ingroup of the present Sub phylogeny (data not shown). This corroborates the taxonomic closeness of dermatophytes and *Aspergillus*, which belong to the same taxonomic group producing prototunicate asci in cleistothecia (class Eurotiomycetes). The orthology of SUB2 and *Aspergillus* secreted alkaline proteases was confirmed by the analyses based on intron–exon structure. Patterns of gene structure are independent of sequence data and have been widely used to gain insight into the evolution of multigenic families (e.g. Nau and Lipkowitz, 2003; Sanchez et al., 2003). *Aspergillus* ALP genes encoding secreted alkaline proteases (intron type IV) and dermatophyte SUB2 (intron type III) have three homologous (equivalent) introns, among which one is absent from all other SUB gene (Fig. 1).

In *A. fumigatus*, only two genes encoding secreted subtilisins (ALP1 and ALP2) (Jaton-Ogay et al., 1992; Reichard et al., 2000; Monod et al., 2002b) have been cloned. In addition, a blast search with subtilisin specific sequence patterns in the TIGR-database of the almost completed *A. fumigatus* genome (<http://www.tigr.org/tdb/e2k1/afu1/>) resulted in the finding of two other genes coding for subtilisins. Translation of presumed ORFs of one of the genes (ALP3) suggested an intracellular location of the protein as no signal peptide could be identified. On the other hand, the other enzyme (ALP4) revealed extensive sequence homology to the subtilisins of the kexin-family. The integration of *A. fumigatus* ALP3 and ALP4 in the SUB sequence alignment revealed that none of these sequences branched within the ingroup of the present SUB phylogeny (data not shown). Therefore, all *T. rubrum* SUBs with the exception of SUB2, appeared to be specific and do not have an ortholog in *Aspergillus*.

At least six SUB gene duplication events are required to explain the tree topology obtained. A primary duplication produced SUB2 on one hand, and the ancestral gene of SUB1 and SUB3–7 on another hand. Subsequent duplication events produced SUB6/SUB7, as well as SUB1/SUB5, and SUB3/SUB4 (Fig. 2). The translation products of dermatophyte SUB paralogs show a quite low degree of sequence conservation (29–67% aa sequence identity) compared to those of MEP paralogs (58–79% identity) (Jousson et al., 2004).

The set of *T. rubrum* Sub amino acid sequences allowed the identification of Subs secreted by other dermatophyte species using proteolysis and mass spectrometry. Orthologous proteases of dermatophytes share a high degree of aa sequence identity as shown for fungalisins (Jousson et al., 2004), and for *A. benhamiae* and *T. rubrum* Subs (Table 4). A protein ortholog to *T. rubrum* Sub7 was identified in *A. benhamiae* while Sub1, Sub2, Sub5 and Sub6 orthologs were not detected in culture supernatants of the four dermatophyte species investigated here. One could envisage that different Subs are secreted in other culture conditions or in vivo. In the growth conditions used here, the various Sub secretion patterns contrasted with those of Meps, which were found to be remarkably stable among dermatophyte species (Jousson et al., 2004). Indeed, the four species analysed here secreted substantial amounts of two orthologous Meps (Mep3 and Mep4) in soy protein medium.

Proteases other than the closely related Sub3 and Sub4 may have other functions than protein degradation during infection. The cDNAs obtained by PCR attested at least for a residual expression of other SUB genes in *T. rubrum*. Sub5 obtained as recombinant protein from *P. pastoris* was active neither on casein nor keratin azure and could be a more specialised protease having a narrow substrate specificity. Sub6, which is closely related to Sub7 and previously described under the name of Tri r 2 allergen is capable of inducing both immediate and delayed type hypersensitivity skin reactions (Woodfolk et al., 1998). Sub1 has a C-terminal extremity which is 104 aa longer than the other Subs and remarkably rich in proline (30 residues) like many parasite surface antigens (Dalrymple et al., 1993). Sub1, like Sub6, is a candidate for a role in the host immunomodulation during dermatophytosis.

The five MEP (Jousson et al., 2004) and seven SUB genes (present work) isolated in *T. rubrum* and in other dermatophytes contrast with the small number of genes encoding secreted endoproteases in *A. fumigatus*. This would suggest that the specialization of dermatophytes was accompanied by the development of a battery of specific proteases. The high keratinolytic activity of *T. rubrum* Sub3 and Sub4 suggests that these proteases play an important role in virulence. However, addressing the relevance of individual genes encoding proteases must await the development of methods for gene inactivation in these genetically less tractable organisms.

Acknowledgements

We thank Dr. Harold Pooley for the critical review of the manuscript and assistance with the English. This work was supported by the Swiss National Foundation for Scientific Research, grant 3100-043193. B. Mignon was supported by grant 3.4534.01 from FRSM in Belgium.

References

- Asahi, M., Linquist, R., Fukuyama, K., Apodaca, G., Epstein, W.L., McKerrow, J.H., 1989. Purification and characterization of major extracellular proteinases from *Trichophyton rubrum*. *Biochem. J.* 232, 139–144.
- Betts, M.J., Guigo, R., Agarwal, P., Russell, R.B., 2001. Exon structure conservation despite low sequence similarity: a relic of dramatic events in evolution? *EMBO J.* 20, 5354–5360.
- Borg-von Zepelin, M., Beggah, S., Boggian, K., Sanglard, D., Monod, M., 1998. The expression of the secreted aspartyl proteinases Sap4 to Sap6 from *Candida albicans* in murine macrophages. *Mol. Microbiol.* 28, 543–554.
- Chambers, S.P., Prior, S.E., Barstow, D.A., Minton, N.P., 1988. The pMTL nic- cloning vectors: I. Improved pUC polylinker regions to facilitate the use of sonicated DNA for nucleotide sequencing. *Gene* 68, 139–149.
- Cheevadhanarak, S., Renno, D.V., Saunders, G., Holt, G., 1991. Cloning and selective overexpression of an alkaline protease-encoding gene from *Aspergillus oryzae*. *Gene* 108, 151–155.
- Dalrymple, B.P., Peters, J.M., Goodger, B.V., Bushell, G.R., Waltisbuhl, D.J., Wright, I.G., 1993. Cloning and characterisation of cDNA clones encoding two *Babesia bovis* proteins with homologous amino- and carboxy-terminal domains. *Mol. Biochem. Parasitol.* 59, 181–189.
- Descamps, F., Brouta, F., Monod, M., Zaugg, C., Baar, D., Losson, B., Mignon, B., 2002. Isolation of a *Microsporum canis* gene family encoding three subtilisin-like proteases expressed in vivo. *J. Invest. Dermatol.* 119, 830–835.
- Felsenstein, J., 1988. Phylogenies from molecular sequences: inference and reliability. *Annu. Rev. Genet.* 22, 521–565.
- Grossberger, D., 1987. Minipreps of DNA from bacteriophage lambda. *Nucleic Acids Res.* 15, 6737.
- Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids Symp. Ser.* 41, 95–98.
- Jaton-Ogay, K., Suter, M., Cramer, R., Falchetto, R., Fatih, A., Monod, M., 1992. Nucleotide sequence of a genomic and cDNA clone encoding an extracellular alkaline protease of *Aspergillus fumigatus*. *FEMS Microbiol. Lett.* 92, 163–168.
- Jousson, O., L  chenne, B., Bontems, O., Capoccia, S., Mignon, B., Barblan, J., Quadroni, M., Monod, M., 2004. Multiplication of an ancestral gene encoding secreted fungalysin preceded species differentiation in the dermatophytes *Trichophyton* and *Microsporum*. *Microbiology* 150, 301–310.
- Kunert, J., Truper, H.G., 1986. Cystine catabolism in mycelia of *Microsporum gypseum*, a dermatophytic fungus. *Arch. Microbiol.* 145, 181–186.
- Kunert, J., 1982. Utilization of L- and DL-cystine by the fungus *Microsporum gypseum*. *Folia Microbiol.* 27, 390–394.
- Kunert, J., 2001. Further studies on the multiple forms of protease ALP of *Aspergillus fumigatus*. *Mycoses* 44, 307–310.
- Laemmli, U.K., 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680–685.
- Mignon, B., Swinnen, M., Bouchara, J.P., Hofinger, M., Nikkels, A., Pierard, G., Gerday, C., Losson, B., 1998. Purification and characterization of a 31.5 kDa keratinolytic subtilisin-like serine protease from *Microsporum canis* and evidence of its secretion in naturally infected cats. *Med. Mycol.* 36, 395–404.
- Monod, M., Togni, G., Rahalison, L., Frenk, E., 1991. Isolation and characterisation of an extracellular alkaline protease of *Aspergillus fumigatus*. *J. Med. Microbiol.* 35, 23–28.
- Monod, M., Togni, G., Hube, B., Sanglard, D., 1994. Multiplicity of genes encoding secreted aspartic proteinases in *Candida* species. *Mol. Microbiol.* 13, 357–368.
- Monod, M., Jaccoud, S., Zaugg, C., L  chenne, B., Baudraz, F., Panizzon, R., 2002a. Survey of dermatophyte infections in the Lausanne area (Switzerland). *Dermatology* 205, 201–203.
- Monod, M., Capoccia, S., L  chenne, B., Zaugg, C., Holdom, M., Jousson, O., 2002b. Secreted proteases from pathogenic fungi. *Int. J. Med. Microbiol.* 292, 405–419.
- Nau, M.M., Lipkowitz, S., 2003. Comparative genomic organization of the cbl genes. *Gene* 308, 103–113.
- Ninet, B., Jan, I., Bontems, O., L  chenne, B., Jousson, O., Panizzon, R., Lew, D., Monod, M., 2003. Identification of dermatophyte species by 28S ribosomal DNA sequencing with a commercial kit. *J. Clin. Microbiol.* 41, 826–830.
- Page, R.D.M., 1996. TREEVIEW: an application to display phylogenetic trees on personal computers. *Comput. Appl. Biosci.* 12, 357–358.
- Perkins, D.N., Pappin, D.J., Creasy, D.M., Cottrell, J.S., 1999. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 20, 3551–3567.
- Posada, D., Crandall, K.A., 1998. Modeltest: testing the model of DNA substitution. *Bioinformatics* 14, 817–818.
- Rawlings, D., Barrett, A.J., 1994. Families of serine peptidases. *Methods Enzymol.* 244, 19–61.
- Reichard, U., Cole, G.T., Hill, T.W., Ruechel, R., Monod, M., 2000. Molecular characterization and influence on fungal development of ALP2, a novel serine proteinase from *Aspergillus fumigatus*. *Zentralbl. Bakteriologie* 290, 549–558.
- Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. *Molecular cloning: a laboratory manual*, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- Sanchez, D., Ganformina, M.D., Gutierrez, G., MarOn, A., 2003. Exon–intron structure and evolution of the lipocalin gene family. *Mol. Biol. Evol.* 20, 775–783.
- Sanyal, A.K., Das, S.K., Banerjee, A.B., 1985. Purification and partial characterization of an extracellular proteinase from *Trichophyton rubrum*. *Sabouraudia* 23, 65–178.
- Swofford, D.L., 1998. PAUP*: Phylogenetic Analysis Using Parsimony (* and other methods) Sinauer Associates, Sunderland, MA.
- Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22, 4673–4680.
- Wilm, M., Shevchenko, A., Houthaeve, T., Breit, S., Schweigerer, L., Fotsis, T., Mann, M., 1996. Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry. *Nature* 379, 466–469.
- Woodfolk, J.A., Wheatley, L.M., Piyasena, R.V., Benjamin, D.C., Plattsmills, T.A., 1998. *Trichophyton* antigens associated with IgE antibodies and delayed type hypersensitivity. Sequence homology to two families of serine proteinases. *J. Biol. Chem.* 273, 29489–29496.