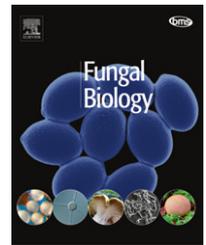




British Mycological
Society promoting fungal science

journal homepage: www.elsevier.com/locate/funbio



Trypsin-like proteins of the fungi as possible markers of pathogenicity

Aleksej G. DUBOVENKO^a, Yakov E. DUNAEVSKY^b, Mikhail A. BELOZERSKY^b,
Brenda OPPERT^c, Jeffrey C. LORD^{c,*}, Elena N. ELPIDINA^b

^aFaculty of Bioengineering and Bioinformatics, Moscow State University, Moscow 119991, Russia

^bA.N. Belozersky Institute of Physico-Chemical Biology, Moscow State University, Moscow 119991, Russia

^cUSDA ARS Grain Marketing and Production Research Center, 1515 College Ave., Manhattan, KS 66502, USA

ARTICLE INFO

Article history:

Received 29 May 2009

Received in revised form

19 November 2009

Accepted 23 November 2009

Available online 25 January 2010

Corresponding Editor: Michael Lorenz

Keywords:

Extracellular peptidases

Pathogens

Saprotrophs

Serine peptidase homologs

Trypsin-like peptidases

ABSTRACT

Sequences of peptidases with conserved motifs around the active site residues that are characteristic of trypsins (similar to trypsin peptidases, STP) were obtained from publicly-available fungal genomes and related databases. Among the 75 fungal genomes, 29 species of parasitic Ascomycota contained genes encoding STP and their homologs. Searches of non-redundant protein sequences, patented protein sequences, and expressed sequence tags resulted in another 18 STP sequences in 10 fungal species from Ascomycota, Basidiomycota, and Zygomycota. A comparison of fungi species containing STP sequences revealed that almost all are pathogens of plants, animals or fungi. A comparison of the primary structure of homologous proteins, including the residues responsible for substrate binding and specificity of the enzyme, revealed three groups of homologous sequences, all presumably from S1 family: trypsin-like peptidases, chymotrypsin-like peptidases and serine peptidases with unknown substrate specificity. Homologs that are presumably functionally inactive were predicted in all groups. The results in general support the hypothesis that the expression of trypsin-like peptidases in fungi represents a marker of fungal phytopathogenicity. A phylogenetic tree was constructed using peptidase and homolog amino acid sequences, demonstrating that all have noticeable differences and almost immediately deviate from the common root. Therefore, we conclude that the changes that occurred in STP of pathogenic fungi in the course of evolution represent specific adaptations to proteins of their respective hosts, and mutations in peptidase genes are important components of life-style changes and taxonomic divergence.

Published by Elsevier Ltd on behalf of The British Mycological Society.

Introduction

Fungi utilize both organic and inorganic derivatives as their nitrogen sources. Some fungi use organic nutrients from any source, and others have strictly defined substrate requirements (Siezen & Leunissen 1997). Fungi are divided into

saprotrophs that feed on dead organic matter and biotrophs that are either parasites or symbionts of other organisms. Under natural environmental conditions, fungi usually consume complex organic compounds that cannot enter the cell without preliminary modification. For this purpose, fungi synthesize extracellular hydrolases that convert high molecular

* Corresponding author. Tel.: +1 785 776 2705.

E-mail address: jeff.lord@ars.usda.gov

1878-6146/\$ – see front matter Published by Elsevier Ltd on behalf of The British Mycological Society.

doi:10.1016/j.funbio.2009.11.004

weight polymers to monomers capable of penetrating into the cell. In the course of evolution, peptidases are prominent among secreted fungal hydrolases because of the importance of proteins for nutritional purposes. Fungal peptidases vary greatly in the conditions in which their synthesis is induced and in biological functions.

Fungal life style and environment determine the array of enzymes that are expressed. Fungi have many trophic relationships with different groups of organisms. Peptidases that are secreted by fungi (especially pathogens) allow them to adapt to different living conditions. Fungi have expanded their repertoire of peptidases throughout evolution to take advantages of different protein sources. For instance, pathogenic fungi produce different enzymes (peptidases, in particular) to penetrate into the host organism depending on the defensive covering of the host (St. Leger et al. 1986; Freimoser et al. 2005).

Peptidases are divided into six main classes – serine, cysteine, threonine, aspartic, glutamic and metallopeptidases, according to the nature of functional groups in the active center of the enzyme molecule. Extracellular proteolytic enzymes of the fungi are represented to a large degree by serine peptidases. There are two major families of serine peptidases that are present in fungi: subtilisin (S8) and chymotrypsin (S1) families (Rawlings et al. 2008). The subtilisin family is similarly ubiquitous in bacteria and fungi. The chymotrypsin family includes trypsins and chymotrypsins from the subfamily S1A that are ubiquitous in animals, but less abundant in fungi species, and most of these enzymes in fungi are trypsins (EC 3.4.21.4). According to preliminary observations, the production of trypsin-like enzymes is characteristic of plant pathogens, whereas extracellular endoproteolytic activity of saprotrophs is provided mainly with subtilisin-like enzymes (St. Leger et al. 1997; Dunaevsky et al. 2001, 2006). These data suggested the existence of a correlation between trypsin genes and pathogenicity of the fungus. To examine this hypothesis, we undertook an analysis of all trypsin-like peptidases containing similar motifs of conserved amino acid sequences in the 75 publicly-available fungal genomes. Fungal trypsin-like sequences from other databases were also studied.

Materials and methods

Sequenced genomes from phyla Ascomycota, Basidiomycota, Zygomycota, and Microsporidia, deposited at the National Center of Biotechnology Information (NCBI) and genome of *Batrachochytrium dendrobatidis* from phylum Chytridiomycota deposited at DOE Joint Genome Institute (JGI), served as starting material for the search of fungi containing genes of trypsin-like peptidases. The search was performed using TBLASTN program from the BLAST cluster, and the sequence of a previously characterized trypsin from *Fusarium oxysporum* (Acc# P35049, Rypniewski et al. 1993) was used as the initial query. In addition, trypsin-like sequences from phyla Ascomycota, Basidiomycota, and Zygomycota were obtained from UniProt using the search system SRS (<http://srs.ebi.ac.uk/>) and BLASTP (Altschul et al. 1997) in databases including non-redundant protein sequences and patented protein

sequences, and these sequences were also used as queries. All sequences were also used to search the Expressed Sequence Tags (ESTs) at NCBI using TBLASTN. Multiple alignment of sequences was carried out using the Clustal W program (Thompson et al. 1994).

Phylogenetic analysis was used to elucidate evolutionary relationships between sequences of trypsin-like proteins. Phylogenetic trees of amino acid sequences were created using PHYLIP 3.67 cluster (Felsenstein 1981). The trees were constructed according to Neighbor-Joining, Maximum Parsimony and Maximum Likelihood algorithms, each using 1000 bootstrap replicates.

Results and discussion

The publicly-available sequenced genomes of 75 fungi species were examined for peptidases with conserved motifs around the active site residues His, Asp and Ser (HDS) similar to those found in trypsin peptidases (STP), as well as proteins referred to as peptidase ‘homologs’ (Yu et al. 2003) that lack hydrolytic functions because amino acid residues of the catalytic triad are substituted. The trypsin from *Fusarium oxysporum* was used as the initial query. Further searches with the enzyme from *Coccidioides immitis* revealed additional peptidase sequences in the genomes of *Coccidioides posadasii*, *Ajellomyces capsulatus*, *Arthroderma gypseum*, *Ascosphaera apis*, *Microsporum canis*, *Paracoccidioides brasiliensis*, *Trichophyton equinum* and *Uncinocarpus reesii*. Thirty-nine genes of peptidases and homologs were found in 30 representatives of Ascomycota subphylum Pezizomycotina, including 12 plant pathogens (18 genes), two fungi pathogens (three genes), as well as 14 animal pathogens (16 genes) and two saprotrophs, *Neurospora crassa* and *U. reesii*, which is a closest relative of the animal pathogenic species *C. immitis* and *C. posadasii*. Genes with conserved sequences similar to trypsins were not found in genomes of the remainder of Ascomycota (seven Pezizomycotina, 22 Saccharomycotina, three Schizosaccharomycetes), 11 Basidiomycota as well as in the two representatives of the Microsporidia. Additional searches in UniProt, including BLASTP in non-redundant and patented protein sequences, as well as a search of ESTs in GenBank, revealed another 18 partial or full sequences of STP genes in six fungal species from Ascomycota, one from Basidiomycota, and three from Zygomycota. These findings represented five plant pathogens, one fungus pathogen, and three animal pathogens. In addition, 22 sequences from 17 Ascomycota species were duplicates of STP genes previously found in the genome searches. It is noteworthy that searches with ESTs resulted in the identification of 27 presumably expressed sequences encoding potential STP or their homologs in fungi. One additional fungus, *Leptosphaeria maculans*, also was revealed to contain a STP sequence.

The analysis of all predicted peptidase or homolog sequences with conserved regions similar to those in trypsins was performed using multiple alignment (Supplementary material, Fig 3) and is presented in Table 1. The alignment revealed the residues responsible for substrate binding and specificity of the enzymes. These signature residues include DGG residues for trypsins and SGG or synonymic residues

Table 1 – Characteristics of STP sequences in fungi.

Data ^a	Species	Accession number ^b	Number of amino acid residues ^c	pI ^d	Activation residue ^e	Conserved Motives around active site residues ^f			Binding pocket residues ^g	Function	Pathogenicity
						TAA <u>H</u> C	<u>D</u> LAI	G <u>D</u> S <u>G</u> GP			
Phylum Ascomycota: Subphylum Pezizomycotina: Class Dothideomycetes: Subclass Pleosporomycetidae: Order Pleosporales ^h											
Genome	<i>Alternaria brassicicola</i>	ACIW01000550.1	215	9.22	N		<u>D</u> VAV		DGG	TLP ⁱ	Plant
		ACIW01000565.1	264	9.80	D		<u>D</u> LAI		DGG	TLP	
Protein + EST	<i>Cochliobolus carbonum</i>	Q00344	231	10.26	S		<u>D</u> IAI		DGG	TLP	Plant
EST	<i>Leptosphaeria maculans</i>	EB762832.1	223	9.99	D		<u>D</u> VAI		DGG	TLP	Plant
Protein + genome + EST	<i>Phaeosphaeria nodorum</i>	Q0URQ8	229	9.92	P		<u>D</u> VAI		DGG	TLP	Plant
		Q0UQI6	331	9.13	D	SA <u>A</u> H <u>C</u>	<u>D</u> IAI		DGG	TLP	
		O74696	299	9.95	P		<u>D</u> VAI		DGG	TLP	
Genome	<i>Pyrenophora tritici-repentis</i>	EDU44173.1	243	9.95	D		<u>D</u> IAI		DGG	TLP	Plant
Phylum Ascomycota: Subphylum Pezizomycotina: Class Eurotiomycetes: Subclass Eurotiomycetidae: Order Eurotiales											
Protein + genome	<i>Aspergillus clavatus</i>	A1CN69	228	4.66	A	TA <u>A</u> S <u>C</u>		G <u>D</u> K <u>G</u> GP	ATM	SPH ^j	Mammalian opportunistic
Genome	<i>Aspergillus flavus</i> (Gene_1)	AAIH02000071.1	222	6.14	A	TA <u>A</u> S <u>C</u>	<u>D</u> IAF	AD <u>Q</u> G <u>G</u> GP	GAG	SPH	Mammalian opportunistic
		AAIH02000079.1	218	4.69	A	TA <u>A</u> D <u>C</u>	<u>N</u> VAV	GD <u>H</u> G <u>G</u> GP	GRE	SPH	opportunistic
		AAIH02000079.1	217	4.98	A	TA <u>A</u> N <u>C</u>	<u>N</u> IAV	GD <u>H</u> G <u>G</u> GP	GRE	SPH	
Protein + genome + EST	<i>Aspergillus nidulans</i>	Q5BAR4	227	3.90	A	TA <u>G</u> H <u>C</u>	<u>D</u> ISI		DGG	TLP	Mammalian opportunistic
Protein + genome + EST	<i>Aspergillus oryzae</i>	Q2UH30	204	4.66	A	TA <u>A</u> D <u>C</u>	<u>N</u> VAV	GD <u>H</u> G <u>G</u> GP	GR-	SPH	Mammalian opportunistic
Protein + genome + EST	<i>Aspergillus terreus</i>	Q0CKN5	229	6.33	G	TT <u>A</u> E <u>C</u>	<u>N</u> VAV	G <u>D</u> M <u>G</u> GP	GLG	SPH	Mammalian opportunistic
Phylum Ascomycota: Subphylum Pezizomycotina: Class Eurotiomycetes: Subclass Eurotiomycetidae: Order Onygenales											
Protein + genome	<i>Ajellomyces capsulatus</i>	A6QWE4	309	5.24	Absence of N-terminal homology	TA <u>S</u> H <u>C</u>	<u>D</u> YIV	GW <u>S</u> G <u>G</u> GP	FGG	SP ⁱ	Mammalian
Genome	<i>Ajellomyces dermatitidis</i>	ACBU01000519.1	356	5.16	Absence of N-terminal homology	TA <u>S</u> H <u>C</u>	<u>D</u> YIV	GW <u>S</u> G <u>G</u> GP	FGS	SP	Mammalian
Genome	<i>Arthroderma gypseum</i>	ABQE01000053.1	321	6.20	Absence of N-terminal homology	TA <u>S</u> H <u>C</u>	<u>D</u> YVI	GW <u>S</u> G <u>G</u> GP	FGG	SP	Mammalian
Genome	<i>Ascosphaera apis</i>	AARE01004707.1	279	9.63	Absence of N-terminal homology	TA <u>S</u> H <u>C</u>	<u>D</u> YVV	GW <u>S</u> G <u>G</u> GP	FGG	SP	Insect
Protein + genome	<i>Coccidioides immitis</i>	Q1DS39	346	5.87	Absence of N-terminal homology		<u>D</u> YVV	GW <u>S</u> G <u>G</u> GP	FAG	SP	Mammalian
Genome	<i>Coccidioides posadasii</i>	ABBB01000107.1	346	5.73	Absence of N-terminal homology		<u>D</u> YVV	GW <u>S</u> G <u>G</u> GP	FAG	SP	Mammalian
Genome	<i>Microsporium canis</i> (<i>Arthroderma otae</i>)	ABVF01000045.1	345	7.87	Absence of N-terminal homology	TA <u>S</u> H <u>C</u>	<u>D</u> YVI	GW <u>S</u> G <u>G</u> GP	FGG	SP	Mammalian
Genome	<i>Paracoccidioides brasiliensis</i>	ABKH01000278.1	353	5.05	Absence of N-terminal homology	TA <u>S</u> H <u>C</u>	<u>D</u> YVI	GW <u>S</u> G <u>G</u> GP	FGG	SP	Mammalian

(continued on next page)

Table 1 – (continued)

Data ^a	Species	Accession number ^b	Number of amino acid residues ^c	pI ^d	Activation residue ^e	Conserved Motives around active site residues ^f			Binding pocket residues ^g	Function	Pathogenicity
						TAAHC	DLAI	GDSGGP			
Genome	<i>Trichophyton equinum</i>	ABW101000361.1	349	5.46	Absence of N-terminal homology	TASHC	DYVI	GWSGGP	FGG	SP	Mammalian
Genome	<i>Uncinocarpus reesii</i>	AAIW01000252.1	342	5.85	Absence of N-terminal homology	TASHC	DYVV	GWSGGP	FAG	SP	Saprotroph
Phylum Ascomycota: Subphylum Pezizomycotina: Class Leotiomycetes: Subclass Helotiales: Order Helotiales											
Genome	<i>Botryotinia fuckeliana</i>	EDN23541.1	277	6.13	D		DIAL		DGG	TLP	Plant
Protein + genome	<i>Sclerotinia sclerotiorum</i>	A7EMI6	231	6.50	G		DIAL		DGG	TLP	Plant
Phylum Ascomycota: Subphylum Pezizomycotina: Class Sordariomycetes: Subclass Hypocreomycetidae: Order Hypocreales											
Protein	<i>Epichloe festucae</i>	B2CJ71	280	10.38	A	TAASC	DIAI	GDTGGP	SPR	CLPH ^j	Plant symbiont
Protein + genome + EST	<i>Fusarium graminearum</i> (<i>Gibberella zeae</i>)	Q4HV44	226	10.25	Q	TASHC	DVAL		DGG	TLP	Plant
Protein + genome + EST	<i>Fusarium oxysporum</i>	P35049	224	9.92	N				DGG	TLP	Plant
Protein	<i>Fusarium solani</i>	CAI95945.1 ^k	223	9.87	M		DVSI		DGG	TLP	Plant
Genome + EST	<i>Gibberella moniliformis</i> (<i>Fusarium verticillioides</i>)	AAIM02000163.1	224	10.43	N				DGG	TLP	Plant
Protein	<i>Metarhizium anisopliae</i>	DQ218240.1	218	10.09	N			GDSGXS	DGG	TLP	
Protein		Q6SV38 ^l	186	6.22	-	-	DVAI		DGG	TLP/TLPH ^j	Insect
Protein		Q6SV39 ^l	149	7.89	-	-	DVAI	GDIGGP	D-	TLPH	
Protein		Q6SV40 ^l	151	9.47	-	-	DIGI		N-	TLP	
Protein		Q6SV41 ^l	136	5.9	-	TAGHC	DIAI	-	-	TLP/TLPH	
Protein + EST		Q9Y7A9	226	4.82	F	TAGHC	DVAV		DGG	TLP	
Protein + EST		Q9Y842	227	4.62	F		DMAI		DGG	TLP	
Protein + EST		Q01136	225	5.16	F		DMAI		DGG	TLP	
Protein + EST		Q9Y843 ^l	186	8.29	M	SAGHC	DMAY	GDSGGS	CGT	CLP ⁱ	
Genome + EST	<i>Trichoderma atroviride</i>	ABDG01000170.1	152	6.08	E	TAGHC	DVAV		DGG	TLP	Fungi
Protein + EST	<i>Trichoderma harzianum</i>	Q8WZM5	229	5.05	D	TAGHC	DVGW		DGG	TLP	Fungi
		A4V8W4	226	6.10	D		DVAV		DGG	TLP	
Genome + EST	<i>Trichoderma reesei</i>	AAIL01000025.1	261	9.67	D		DVAV		DGG	TLP	Plant
Genome + EST genome	<i>Trichoderma virens</i> (Gene_1)	ABDF01000288.1	234	6.05	D	TAGHC	DIGV		DGG	TLP	Fungi
	<i>Trichoderma virens</i> (Gene_2)	ABDF01000047.1	219	9.20	D		DVAV		DGG	TLP	
Phylum Ascomycota: Subphylum Pezizomycotina: Class Sordariomycetes: Subclass Sordariomycetidae: Order Magnaporthales											
Protein + genome + EST	<i>Magnaporthe grisea</i>	A4QXG7	233	5.65	Y	TARHC	DWAI	GQSGGP	DGG	TLP	Plant
		A4RDZ8	222	6.51	D	TARHC	DWAI	GQSGGP	DIG	TLP	
		A4RGJ7	185	10.41	F	SAGHC	DMAF	GDSGGS	CGT	CLP	
Phylum Ascomycota: Subphylum Pezizomycotina: Class Sordariomycetes: Subclass Sordariomycetidae: Order Sordariales											
Protein + genome	<i>Neurospora crassa</i>	Q7RX08	722	5.60	Absence of N-terminal homology	TAGHC	DVAF	GPLGGP	SVK	CLPH	Saprotroph

system. A sequence found in *Aspergillus nidulans* encoded an active trypsin-like peptidase from TLP group with all active site residues necessary for catalytic function.

It is interesting that in *Aspergillus oryzae* and *Aspergillus terreus*, the active expression of enzyme homologs was observed, as these sequences were found in EST databases (Table 1). Therefore, although the homologs are presumably devoid of proteolytic activity due to substitutions in the catalytic triad residues, these proteins apparently are necessary for life activity of the fungus. The biological functions of peptidase homologs have not been elucidated. The most thorough investigation of these homologs was in insects, where it was speculated that plasma homologs participate in the innate immune response as cofactors for the activation of insect prophenoloxidase by prophenolactivating peptidase (Yu et al. 2003; Zhang et al. 2004; Kanost & Gorman 2008). However, peptidase homologs also have been speculated to be involved in the digestive response of insects to dietary inhibitors (Prabhakar et al. 2007).

The amino acid residue recognized by the proenzyme activating enzyme in mammalian trypsins is a basic amino acid residue (lysine or arginine), but this residue is not conserved in Ascomycota TLP. Instead, this residue is phenylalanine for most *M. anisopliae* peptidases, and for other Ascomycota species, this position may be occupied by glycine, alanine, aspartic acid, asparagine, serine, glutamine, glutamic acid, tyrosine, methionine, or proline. These results suggest that proenzyme activation due to autocatalysis or by the action of peptidases similar to enterokinase, as in mammalian trypsins, does not occur in fungal trypsins. In view of the substrate specificity of endopeptidases, in some cases activation may be performed by subtilisin-like enzymes (Balingier & Wells 1998), which are actively secreted by the fungi, including trypsin-containing fungi (Dunaevsky et al. 2001; Hu & St. Leger 2004). It is possible that the substitution of the basic amino acid residue for other residues in trypsin-like enzymes of fungi was successful in the course of evolution to eliminate the possibility of autocatalysis of the zymogen, and instead to require another specific enzyme for activation. This refined activation may be important for the regulation of the activity of fungal trypsin-like proteases, as well as the development of new functions correlated to pathogenesis. It must be noted that TLP from the phylum Zygomycota may be autoactivated because the activation residue is arginine. The SP in the species from order Onygenales do not contain the conserved N-terminal sequences found in trypsins and chymotrypsins, and so it is difficult to identify the activation residue.

Analysis of fungi containing STP genes indicated that all except three of the fungal species were pathogens of plants, animals or fungi (Table 1). Apparently, the interaction of fungal STP with the environment provides important specific functional advantages to those fungi expressing peptidases at higher levels, and possibly adapts them to a certain way of life, such as pathogenicity. The detailed analysis of the correlation between fungal pathogenicity and different groups of STP revealed that TLP were found primarily in 16 species of fungi pathogenic for plants (21 genes), three mycopathogens (five genes), three insect pathogens (nine genes) and only one opportunistic mammalian pathogen (one gene) (Fig 1B). CLP (five genes) were found in pathogens of plants and insects,

one plant symbiont, and one saprotroph. SP were derived primarily from pathogens of mammals (14 genes from 12 species) with the exception of one pathogen of an insect and one saprotroph. Accordingly, fungi possessing TLP and SP are mainly pathogens of different groups of hosts. Correlation of TLP and phytopathogenicity of fungi was demonstrated also in experimental studies (St. Leger et al. 1997; Dunaevsky et al. 2001, 2006).

Table 2 – Pathogenicity of fungal species with genomes that lack STP genes.

Ascomycota: Pezizomycotina	
<i>Aspergillus fumigatus</i>	Mammalian pathogen
<i>Blumeria graminis</i>	Plant pathogen
<i>Chaetomium globosum</i>	Mammalian pathogen
<i>Neosartorya fischeri</i>	Mammalian pathogen
<i>Penicillium marneffeii</i>	Mammalian pathogen
<i>Talaromyces stipitatus</i>	Non-pathogenic
Ascomycota: Saccharomycotina	
<i>Ashbya gossypii</i>	Plant pathogen
<i>Candida glabrata</i>	Mammalian pathogen
<i>Candida parapsilosis</i>	Mammalian pathogen
<i>Candida tropicalis</i>	Mammalian pathogen
<i>Clavispora lusitaniae</i>	Mammalian pathogen
<i>Candida albicans</i>	Mammalian pathogen
<i>Debaryomyces hansenii</i>	Non-pathogenic
<i>Kluyveromyces lactis</i>	Non-pathogenic
<i>Kluyveromyces waltii</i>	Non-pathogenic
<i>Lodderomyces elongisporus</i>	Non-pathogenic
<i>Pichia guilliermondii</i>	Mammalian pathogen
<i>Pichia stipitis</i>	Non-pathogenic
<i>Saccharomyces bayanus</i>	Non-pathogenic
<i>Saccharomyces castellii</i>	Non-pathogenic
<i>Saccharomyces cerevisiae</i>	Non-pathogenic
<i>Saccharomyces kluyveri</i>	Non-pathogenic
<i>Saccharomyces kudriavzevii</i>	Non-pathogenic
<i>Saccharomyces mikatae</i>	Non-pathogenic
<i>Saccharomyces paradoxus</i>	Non-pathogenic
<i>Saccharomyces pastorianus</i>	Non-pathogenic
<i>Vanderwaltozyma polyspora</i>	Non-pathogenic
<i>Yarrowia lipolytica</i>	Non-pathogenic
Ascomycota: Schizosaccharomycotina	
<i>Schizosaccharomyces japonicus</i>	Non-pathogenic
<i>Schizosaccharomyces octosporus</i>	Non-pathogenic
<i>Schizosaccharomyces pombe</i>	Non-pathogenic
Basidiomycota	
<i>Coprinopsis cinerea okayama</i>	Non-pathogenic
<i>Cryptococcus bacillisporus</i>	Mammalian pathogen
<i>Cryptococcus neoformans</i>	Mammalian pathogen
<i>Laccaria bicolor</i>	Non-pathogenic
<i>Malassezia globosa</i>	Mammalian pathogen
<i>Malassezia restricta</i>	Mammalian pathogen
<i>Moniliophthora perniciosa</i>	Plant pathogen
<i>Phanerochaete chrysosporium</i>	Non-pathogenic
<i>Postia placenta</i>	Non-pathogenic
<i>Puccinia graminis f. sp. tritici</i>	Plant pathogen
<i>Ustilago maydis 5</i>	Plant pathogen
Chytridiomycota	
<i>Batrachochytrium dendrobatidis</i>	Amphibian pathogen
Microsporidia	
<i>Encephalitozoon cuniculi</i>	Mammalian pathogen
<i>Enterocytozoon bieneusi</i>	Mammalian pathogen

have noticeable differences and almost immediately deviate from the common root, it is possible to identify a few branches corresponding to the taxonomic distribution of species, with a few exclusions. It was particularly interesting that all SP from the order Onygenales formed a single separate branch at the level of order. SP homologs from the genus *Aspergillus*, belonging to the same class but another order Eurotiales, were grouped in another branch. This branch did not include the only *Aspergillus* species containing TLP but not SP, *A. nidulans*. The TLP sequences demonstrated higher variability. All TLP from the genus *Fusarium*, as well as TLP from the genus *Trichoderma*, constituted separate branches, while multiple sequences from *M. grisea* as well as *M. anisopliae* were located in different branches. Overall, the tree, constructed according to a Maximum Likelihood algorithm, demonstrates that the fungal STP and their homologs have relatively low similarity. Thus, we conclude that the changes that occurred in STP of pathogenic fungi in the course of evolution represent specific adaptations to proteins of their respective hosts, and mutations in peptidase genes are important components of life-style changes and taxonomic divergence. Trees constructed according to Maximum Parsimony and Neighbor-Joining algorithms gave similar results (data not shown).

Investigations of Hu & St. Leger (2004) using hybridizations of trypsin gene fragments with genome DNA in 35 fungi species revealed the absence of trypsin genes in saprotrophs and their existence in a basidiomycete, an insect symbiont, a majority of zygomycetes and many ascomycetes, both plant and insect pathogens. According to the authors' opinion, the fragmentary distribution of trypsins among fungi indicates that their phylogenetic distribution may be greater in the early fungi than in modern ones. The authors specify that the bootstrap support in the trypsin tree is low for many groups of fungi, although it is not in contradiction with the evolutionary tree of fungi species. Investigations suggest that evolution proceeded in which the divergence of trypsins was largely correlated to the speciation of fungi, and the absence of trypsin in some fungi indicates the loss of the gene. Our data are in accordance with these conclusions.

Overall, our results support the hypothesis that the presence of trypsin or TLP gene(s) represents a marker of fungal phytopathogenicity, possibly excluding narrowly specialized pathogens. The results also indicate a correlation between presence of TLP in fungi and entomo- and mycopathogenicity, although at present there are limited data supporting this conclusion. The data on the presence of STP in mammals fall in two groups, lacking STP and containing predominantly another group of STP subgroup SP. Accordingly, no correlation between presence of STP and fungal pathogenicity to mammals was observed.

Acknowledgments

The research was performed with the support of ISTC grant #3455 and the Russian Foundation for Basic Research (grant # 10-04-00739). Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture.

Supplementary material

Supplementary material associated with this article is available in the online version at doi:10.1016/j.funbio.2009.11.004.

REFERENCES

- Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ, 1997. Grapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* 25: 3389–3402.
- Ballinger MD, Wells JA, 1998. Subtilisin. In: Barrett AJ, Rawlings ND, Woessner JF (eds), *Handbook of Proteolytic Enzymes* Academic Press, New York, pp. 289–292.
- Dunaevsky YE, Golubeva EA, Gruban TN, Beliakova GA, Belozersky MA, 2001. Regulation of secretion of extracellular proteases by filamentous fungus *Botrytis cinerea* Fr. *Journal of Russian Phytopathological Society* 2: 39–44.
- Dunaevsky YE, Gruban TN, Beliakova GA, Belozersky MA, 2006. Extracellular proteinases of filamentous fungi as potential markers of phytopathogenesis. *Mikrobiologia* 75: 747–751 (Russ.).
- Felsenstein J, 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. *Journal of Molecular Evolution* 17: 368–376.
- Freimoser FM, Hu G, St. Leger RJ, 2005. Variation in gene expression patterns as the insect pathogen *Metarhizium anisopliae* adapts to different host cuticles or nutrient deprivation *in vitro*. *Microbiology* 151: 361–371.
- Freimoser FM, Screen SE, Bagga S, Hu G, St. Leger RJ, 2003. EST analysis of two subspecies of *Metarhizium anisopliae* reveals a plethora of secreted proteins with potential activity in insect hosts. *Microbiology* 149: 239–247.
- Hedstrom L, Szilagyi L, Rutter WJ, 1992. Converting trypsin to chymotrypsin: the role of surface loops. *Science* 255: 1249–1253.
- Hu G, St. Leger RJ, 2004. A phylogenomic approach to reconstructing the diversification of serine proteases in fungi. *Journal of Evolutionary Biology* 17: 1204–1214.
- Kanost MR, Gorman MJ, 2008. Phenoloxidases in insect immunity. In: Beckage N (ed), *Insect Immunology*. Academic Press, New York, pp. 69–96.
- Prabhakar S, Chen MS, Elpidina EN, Vinokurov KS, Smith CM, Oppert B, 2007. Molecular characterization of digestive proteinases and sequence analysis of midgut cDNA transcripts of the yellow mealworm, *Tenebrio molitor* L. *Insect Molecular Biology* 16: 455–468.
- Rawlings ND, Morton FR, Kok CY, Kong J, Barrett AJ, 2008. MEROPS: the peptidase database. *Nucleic Acids Research* 36: D320–D325. <http://merops.sanger.ac.uk>.
- Ross J, Jiang H, Kanost MR, Wang Y, 2003. Serine proteases and their homologs in the *Drosophila melanogaster* genome: an initial analysis of sequence conservation and phylogenetic relationships. *Gene* 304: 117–131.
- Rypniewski WR, Hastrup S, Betzel Ch, Dauter M, Dauter Z, Papendorf G, Branner S, Wilson KS, 1993. The sequence and X-ray structure of the trypsin from *Fusarium oxysporum*. *Protein Engineering* 6: 341–348.
- Screen SE, St. Leger RJ, 2000. Cloning, expression, and substrate specificity of a fungal chymotrypsin. Evidence for lateral gene transfer from an actinomycete bacterium. *Journal of Biological Chemistry* 275: 6689–6694.
- Siezen RJ, Leunissen JAM, 1997. Subtilases: the superfamily of subtilisin-like serine proteases. *Protein Science* 6: 501–523.
- St. Leger RJ, Charnley AK, Cooper RM, 1986. Cuticle-degrading enzymes of entomopathogenic fungi: synthesis in culture on cuticle. *Journal of Invertebrate Pathology* 48: 85–95.

-
- St. Leger RJ, Joshi L, Roberts DW, 1997. Adaptation of proteases and carbohydrases of saprophytic, phytopathogenic and entomopathogenic fungi to the requirements of their ecological niches. *Microbiology* **143**: 1983–1992.
- Thompson JD, Higgins DG, Gibson TJ, 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22**: 4673–4680.
- Yu X-Q, Jiang H, Wang Y, Kanost MR, 2003. Nonproteolytic serine proteinase homologs are involved in prophenoloxidase activation in the tobacco hornworm, *Manduca sexta*. *Insect Biochemistry and Molecular Biology* **33**: 197–208.
- Zhang G, Lu Z-Q, Jiang H, Asgari S, 2004. Negative regulation of prophenoloxidase (proPO) activation by a clip-domain serine proteinase homolog (SPH) from endoparasitoid venom. *Insect Biochemistry and Molecular Biology* **34**: 477–483.