

Glutamic protease distribution is limited to filamentous fungi

Andrew H. Sims^a, Nigel S. Dunn-Coleman^b, Geoffrey D. Robson^a,
Stephen G. Oliver^{a,*}

^a School of Biological Sciences, The Michael Smith Building, University of Manchester, Oxford Road, Manchester M13 9PT, UK

^b Genencor International Inc., 925 Page Mill Road, Palo Alto, CA 94304, USA

Received 17 June 2004; received in revised form 20 July 2004; accepted 18 August 2004

First published online 11 September 2004

Edited by M.Y. Galperin

Abstract

Glutamic proteases are a distinct, and recently re-classified, group of peptidases that are thought to be found only in fungi. We have identified and analysed the distribution of over 20 putative glutamic proteases from all fungal species whose genomes have been sequenced so far. Although absent from the Saccharomycetales class, glutamic proteases appear to be present in all other ascomycetes species examined. A large number of coding regions for glutamic proteases were also found clustered together in the *Phanerochaete chrysosporium* genome, despite apparently being absent from three other species of Basidiomycota.

© 2004 Federation of European Microbiological Societies. Published by Elsevier B.V. All rights reserved.

Keywords: Protease; Glutamic; Genome; Comparative genomics; Sequence; Filamentous fungi

1. Introduction

The glutamic protease family, recently re-classified as a sixth catalytic type of peptidase (family G1) in the MEROPS database (<http://www.merops.sanger.ac.uk>) currently contains peptidases from five species of Ascomycota. Previously known as the A4 family of aspartic endopeptidases, recent analysis of the molecular structure and catalytic mechanism has identified these enzymes as a novel protease family, the Eqolisins, a name derived from the active-site residues, glutamic acid (E) and glutamine (Q) [1]. Members of this newly recognised family of peptidases have a previously undescribed β -sandwich as a tertiary fold and a unique

catalytic dyad consisting of glutamine and glutamate residues which, respectively, activate the nucleophilic water and stabilise the tetrahedral intermediate on the hydrolytic pathway [1]. The only previously isolated examples of glutamic proteases (previously designated acidic or aspartic endopeptidases) are from *Scytalidium lignicolum*, *Aspergillus niger*, *Cryphonectria parasitica* (chestnut blight fungus), *Talaromyces emersonii* and *Sclerotinia sclerotiorum*, all filamentous fungal species of the Ascomycota phylum. Of these species, only *A. niger* has a completed genome sequence, although this is not publicly available (see, <http://www.dsm.com> and www.genencor.com). However, a growing number of ascomycete genomes have been sequenced [2–8] and the first from a basidiomycete (*Phanerochaete chrysosporium* [9]) has been publicly released. Nevertheless, many fungal open reading frames (ORFs) still have unknown functions. Comparative genomics is increasingly being used to examine gene conservation and can be used for improved gene prediction [7,10]. Here, we have used

* Corresponding author.

E-mail addresses: andrew.h.sims@stud.man.ac.uk (A.H. Sims), steve.oliver@man.ac.uk (S.G. Oliver).

¹ Tel.: +44 161 275 1578; fax: +44 161 275 5082.

a comparative genomics approach to look at the conservation and evolutionary distribution of genes for this interesting and newly characterised protease family in the fungi.

2. Methods

All predicted open reading frames from each species (in Table 1) with 100 or more amino acids, and beginning with a methionine start codon, were searched for amino-acid sequences with significant similarity to previously identified glutamic proteases using InterProScan [11] and BLAST searches. GenBank was also searched for homologs to previously identified glutamic proteases. Amino-acid sequences were downloaded from the sources listed in Table 1. The coding sequences for predicted proteases were manually compared using ClustalW alignment [12]. Hierarchical clustering of amino-acid sequence alignments was performed using the MultAlin program [13] with default (blosum62) matrix. The SignalP program [14] was used to look for predicted prepro signal sequences.

3. Results and discussion

The glutamic proteases are quite distinct from previously characterised proteases. This is highlighted by the fact that there are only a handful of significantly similar sequences currently in GenBank (BLASTp e -value ≤ 0.01), and these represent either the small number of previously isolated glutamic proteases or hypothetical proteins encoded by ORFs from recently sequenced genomes of filamentous fungal species. These hypothetical proteins and other predicted ORFs were included in this study.

The 23 fungal genomes (listed in Table 1) were systematically searched for ORFs with sequence similarity to previously isolated examples of glutamic proteases. InterProScan [11] was used to search PROSITE, PRINTS, Pfam, ProDom and SMART for protein signatures in these ORFs (Table 2).

A total of 26 ORFs from seven species with completely sequenced genomes were identified as potentially encoding glutamic proteases, including one previously identified sequence (pepB from *A. niger*). Investigating the predicted ORFs from more than 20 fungal species showed that glutamic proteases were restricted to the higher ascomycetes and were not found in any of the Saccharomycetales sequenced to date (Table 1). Possible ORFs encoding glutamic proteases were found in all the other ascomycetes sequenced so far from the Eurotiales, Pezizales and Sordariales. The genomes of these species each contain 1–4 ORFs predicted to encode glutamic proteases; these genomes have approximately twice as

many predicted ORFs compared to members of the Saccharomycetales.

Further investigation of the predicted glutamic protease sequences (Table 2) showed a high level of conservation between species, with the glutamine and glutamic acid active-site residues resolved by Fujinaga et al. [1] conserved in the amino-acid sequences of most of them. Exceptions to this were two similar ORFs from *Neurospora crassa* and *Magnaporthe grisea* (NCU04205.1, MG09032.4) and two adjacent ORFs from *P. chrysosporium* (pc.12.97.1, pc.12.98.1). The sequences of these ORFs appear to have been correctly predicted and may represent genes which have lost functionality. The *P. chrysosporium* ORF, pc78.43.1, has a glutamine rather than a glutamic acid residue at the second active site position, which may still be functional. The additional sequences of putative glutamic proteases identified here can be used to create an improved Hidden Markov Model of this protein family for future analyses of other species. All of the ORFs characterised were predicted to encode signal peptides using SignalP [14] and many were seen to specify the conserved cysteine residues that form a disulphide bridge, which surrounds an aspartic acid residue that is conserved in all members of this family [1]. Within the glutamic protease family, there appear to be two distinct groups, with only half of the ORFs predicted to contain two C-terminal transglutaminase signatures (InterPro domain IPR008958) which catalyze the post-translational modification of proteins at glutamine residues [15].

The product of the *Fusarium graminearum* ORF, FG08196.12, has much greater amino-acid sequence similarity to other glutamic proteases (including conservation of the glutamic acid active-site residue), if it is considered to have just 1 exon (ignoring the predicted intron). This assumption is supported by a 99% match to Gz31371614, a full-length sequence from the Phytopathogenic Fungi and Oomycete EST Database (<http://cogeme.ex.ac.uk>) [16] and represents another example of the value of using alignments to cDNA sequences for correct gene prediction [17].

The largest number of ORFs encoding putative glutamic proteases was found in the genome of the white rot fungus *P. chrysosporium*, which contains 12 predicted sequences. This was unexpected as the Basidiomycota are evolutionarily distant relatives of the Ascomycota and no members of the glutamic protease family have been identified (using tBLASTn) in other members of the Basidiomycota which have been sequenced (but not yet annotated), namely *Cryptococcus neoformans* Serotype A, *Coprinus cinereus*, *Ustilago maydis* (www.broad.mit.edu/annotation/).

Looking more closely at the large number of predicted glutamic proteases in *P. chrysosporium*, it is clear that the sequences are very highly conserved. The nucleotide sequences of the ORFs are up to 94% identical

Table 1
Summary of glutamic protease distribution and genome sequence database websites

Species	Class	Phylum	Glutamic proteases	Sequenced by	Website
<i>Ustilago maydis</i>	Ustilaginomycetes	Basidiomycota	0	Broad Institute	http://www.broad.mit.edu/
<i>Coprinus cinereus</i>	Hymenomycetes		0		
<i>Cryptococcus neoformans A</i>			0		
<i>Phanerochaete chrysosporium</i>			12	Joint Genome Institute [9]	http://www.jgi.doe.gov/
<i>Aspergillus fumigatus</i>	Eurotiales	Ascomycota	2	Consortium [21]	http://www.cadre.man.ac.uk/
<i>Aspergillus nidulans</i>			2	Broad Institute	http://www.broad.mit.edu/
<i>Aspergillus niger</i>			3	DSM	http://www.dsm.com/
				Integrated genomics/Genencor	http://www.integratedgenomics.com/
<i>Magnaporthe grisea</i>	Pezizales		2	Broad Institute	http://www.broad.mit.edu/
<i>Neurospora crassa</i>	Sordariales		2	Broad Institute [5]	
<i>Fusarium graminearum</i>			1	Broad Institute	
<i>Trichoderma reesei</i>			4	Joint Genome Institute	http://www.jgi.doe.gov/
<i>Candida albicans</i>	Saccharomycetales		0	Stanford Genome Technology [22]	http://www-sequence.stanford.edu/group/candida/
<i>Candida glabrata</i>		0	<i>Génolevures</i> [6]	http://cbi.labri.fr/Genolevures/	
<i>Yarrowia lipolytica</i>		0			
<i>Debaryomyces hansenii</i>		0			
<i>Kluyveromyces lactis</i>		0			
<i>Ashbya gossypii</i>		0		Biozentrum der Universität Basel [4]	http://agd.unibas.ch/
<i>Kluyveromyces waltii</i>		0		Broad Institute [8]	http://www.broad.mit.edu/
<i>Saccharomyces paradoxus</i>		0		Broad Institute [7]	
<i>Saccharomyces mikatae</i>		0			
<i>Saccharomyces bayanus</i>		0			
<i>Saccharomyces cerevisiae</i>	0		Consortium [2]	http://www.yeastgenome.org/	
<i>Schizosaccharomyces pombe</i>	0		Consortium [3]	http://www.genedb.org/genedb/pombe/	

Table 2
Known and hypothetical proteins thought to represent glutamic proteases

Species	ORF/GenBank Accession No.	Protein signatures				Active site residues		Disulphide bridge	
		IPR000250			IPR008958	Gln-53	Glu-136	Cys-47	Cys-127
		PD18627	PR00977	PF001828					
<i>Cryphonectria parasitica</i>	eapB (X83998)	X	X	X		Q	E	C	C
	eapC (X83997)	X	X	X	X	Q	E	A	Q
<i>Talaromyces emersonii</i>	(AF439998) ^a	X	X	X	X	Q	E	C	C
	(AF439999) ^a	X	X	X		Q	E	Y	G
<i>Sclerotinia sclerotiorum</i>	(AF221843)	X	X	X	X	Q	E	C	C
<i>Scytalidium lignicolum</i>	(AB038553)	X	X	X	X	Q	E	C	C
<i>Neurospora crassa</i>	NCU04205.1	X				R	E	T	C
	NCU05584.1	X	X			Q	E	W	A
<i>Magnaporthe grisea</i>	MG00311.4	X	X			Q	E	C	C
	MG09032.4	X	X	X		R	E	T	C
<i>Fusarium graminearum</i>	FG08196.1 ^c	X	X	X		Q	E	C	C
<i>Aspergillus fumigatus</i>	m03908	X	X	X	X	Q	E	C	C
	m04648	X	X	X		Q	E	C	C
<i>Aspergillus nidulans</i>	AN7515.1	X	X	X	X	Q	E	C	C
	AN3377.1	X	X	X		Q	E	Y	K
<i>Aspergillus niger</i>	pepB (RANI71962)	X	X	X	X	Q	E	C	C
	RANI69718	X	X	X	X	Q	E	C	C
	RANI73030	X	X	X		Q	E	Y	G
<i>Trichoderma reesei</i>	6583	X	X	X		Q	E	D	P
	1094	X	X	X		Q	E	A	A
	5103	X	X	X	X	Q	E	C	C
	8059	X	X	X		Q	E	C	C
<i>Phanerochaete chrysosporium</i>	pc.12.97.1	X	X	X		L	R	R	C
	pc.12.98.1	X	X	X		Q	G	C	C
	pc.123.21.1	X	X	X	X	Q	E	C	C
	pc.123.26.1	X	X	X		Q	E	C	C
	pc.123.27.1	X	X	X	X	Q	E	C	C
	pc.123.29.1	X	X	X	X	Q	E	C	C
	pc.16.69.1	X	X	X		Q	E	C	C
	pc.236.2.1	X	X	X	X	Q	E	C	C
	pc.16.88.1	X	X	X		Q	E	C	Y
	pc.78.43.1 ^b	X	X	X		Q	Q	C	C
	2nd.48.17.1 ^b	X	X	X	X	Q	E	C	C
	Pc.78.44.1	X	X	X	X	Q	E	C	C

Bold sequences represent previously described genes in MEROPs and GenBank, other sequences can be found from websites listed in Table 1. Protein signatures identified with InterProScan from and PROSITE, PRINTS, Pfam, ProDom and SMART; IPR000250 Family Peptidase A4, PD018627 PRTA *A. niger*, PR00977 scytalidopepsin B, PF01828.7 Peptidase A4 family, IPR008958 transglutaminase. Positions of active site and disulphide bridge residues refer to *Scytalidium lignicolum* glutamic protease [1].

^a Partial sequence.

^b 1st exon and intron sequences of these ORFs were manually predicted based upon sequence alignments (Fig. 2).

^c one exon only (see text).

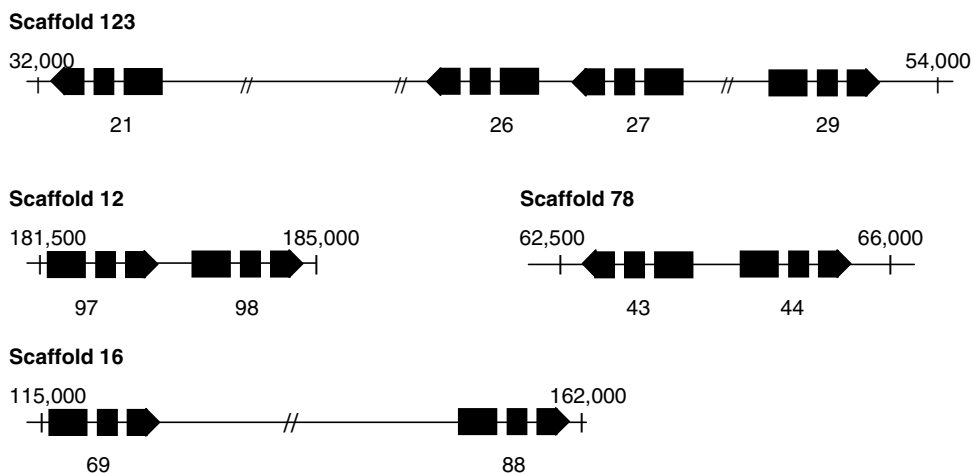


Fig. 1. Distribution of glutamic proteases in the *P. chrysosporium* genome. Lines represent genomic and intron sequences. Blocks represent exons in the direction indicated and numbers of the ORF in the scaffold are given below. Relative nucleotide positions in contigs are indicated but are not to scale.

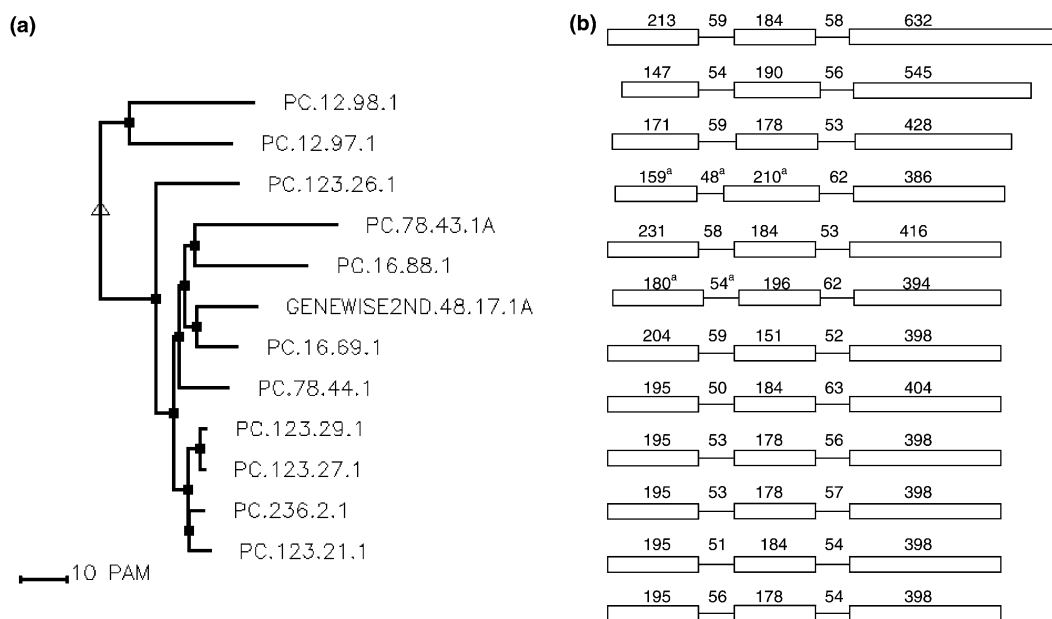


Fig. 2. Conservation of glutamic protease sequences in *P. chrysosporium*. (a) Hierarchical clustering of amino-acid sequence alignments. Relative evolutionary distances are shown in PAM units. (b) Comparison of gene structure showing relative nucleotide sizes of introns (lines) and exons (blocks). ^a Manually deduced from sequence alignments.

and, even if the introns are included in the analysis, an identity of 91% is still found. Moreover, there is a large degree of conservation in gene structure with almost identical sizes and positions of introns and exons (Fig. 2). In addition, the distribution of the ORFs is unusual with three pairs of genes adjacent to one another in the same scaffold and four within a 22-kb region of scaffold 123 (Fig. 1). Variation of the ORFs adjacent to the predicted glutamic proteases within the scaffolds and the quality control measures taken with the genome assem-

bly [9], suggest that these findings are not the result of sequencing or assembly errors. Although fewer in number, the glutamic proteases found in other species were not on the same contigs/scaffolds.

The high level of conservation and distribution of these ORFs within the *P. chrysosporium* genome may indicate local duplication events of fairly recent evolutionary origin. In *S. cerevisiae*, a number of gene families are telomere-associated [2]. The *P. chrysosporium* genome is believed to be divided into 10 chromosomes

[18] and the current genome assembly [9] contains 29.6 Mb of non-repetitive sequence in 349 scaffolds greater than 3 kb. Although the shotgun sequencing approach excluded telomeres from the assembly, the association (in scaffold 12) of two adjacent ORFs encoding glutamic proteases with *AADE*, a close homolog of the telomere-associated aryl-alcohol dehydrogenase (AAD) genes of *S. cerevisiae* [19], may suggest that this scaffold is close to a chromosome end. The occurrence of these closely related members of a gene family is in clear contrast to the specific and general situation of *N. crassa*, in which the repeat-induced point mutation system appears to restrict the sizes of gene families [5].

Glutamic proteases have been shown to be responsible for degrading recombinant proteins in *A. niger*. Silencing by antisense RNA expression or protease removal by gene disruption has resulted in increased yields [20]; further knock-outs of these protease genes identified here may also be advantageous. These data imply that glutamic proteases are not essential for fungal growth and the distribution of members of this gene family both within and between the sequenced fungal genomes demonstrates the greater variation of non-essential genes as compared to the ubiquitous occurrence of the core group of highly conserved orthologs in all species examined [16]. The increasing number and diversity of fungal species for which fully sequenced genomes are available [2–8] is facilitating investigations of gene conservation and the evolution of protein families. The glutamic proteases are interesting to study both in terms of their novel peptidase activity and their apparently unusual distribution among the filamentous fungi.

Acknowledgements

A.H.S is the grateful recipient of a CASE Studentship from the BBSRC and Genencor International. S.G.O. acknowledges the contribution of the COGEME Grant from the IGF Initiative of the BBSRC to his research in the bioinformatics of fungi.

References

- [1] Fujinaga, M., Cherney, M.M., Oyama, H., Oda, K. and James, M.N. (2004) The molecular structure and catalytic mechanism of a novel carboxyl peptidase from *Scytalidium lignicolum*. Proc. Natl. Acad. Sci. USA 101, 3364–3369.
- [2] Goffeau, A., Barrell, B.G., Bussey, H., Davis, R.W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J.D., Jacq, C., Johnston, M., Louis, E.J., Mewes, H.W., Murakami, Y., Philippsen, P., Tettelin, H. and Oliver, S.G. (1996) Life with 6000 genes. Science 274, 546–567.
- [3] Wood, V., Gwilliam, R., Rajandream, M.A., Lyne, M., Lyne, R., Stewart, A., Sgouros, J., Peat, N., Hayles, J., Baker, S., Basham, D., Bowman, S., Brooks, K., Brown, D., Brown, S., Chillingworth, T., Churcher, C., Collins, M., Connor, R., Cronin, A., Davis, D., Feltwell, T., Fraser, A., Gentles, S., Goble, A., Hamlin, N., Harris, D., Hidalgo, J., Hodgson, G., Holroyd, S., Hornsby, T., Howarth, S., Huckle, E.J., Hunt, S., Jagels, K., James, K., Jones, L., Jones, M., Leather, S., McDonald, S., McLean, J., Mooney, P., Moule, S., Mungall, K., Murphy, L., Niblett, D., Odell, C., Oliver, K., O'Neil, S., Pearson, D., Quail, M.A., Rabinowitz, E., Rutherford, K., Rutter, S., Saunders, D., Seeger, K., Sharp, S., Skelton, J., Simmonds, M., Squares, R., Squares, S., Stevens, K., Taylor, K., Taylor, R.G., Tivey, A., Walsh, S., Warren, T., Whitehead, S., Woodward, J., Volckaert, G., Aert, R., Robben, J., Grymonprez, B., Weltjens, I., Vanstreels, E., Rieger, M., Schafer, M., Muller-Auer, S., Gabel, C., Fuchs, M., Dusterhoft, A., Fritzc, C., Holzer, E., Mostl, D., Hilbert, H., Borzym, K., Langer, I., Beck, A., Lehrach, H., Reinhardt, R., Pohl, T.M., Eger, P., Zimmermann, W., Wedler, H., Wambutt, R., Purnelle, B., Goffeau, A., Cadieu, E., Dreano, S., Gloux, S., Lelaure, V., Mottier, S., Galibert, F., Aves, S.J., Xiang, Z., Hunt, C., Moore, K., Hurst, S.M., Lucas, M., Rochet, M., Gaillardin, C., Tallada, V.A., Garzon, A., Thode, G., Daga, R.R., Cruzado, L., Jimenez, J., Sanchez, M., del Rey, F., Benito, J., Dominguez, A., Revuelta, J.L., Moreno, S., Armstrong, J., Forsburg, S.L., Cerutti, L., Lowe, T., McCombie, W.R., Paulsen, I., Potashkin, J., Shpakovski, G.V., Ussery, D., Barrell, B.G., Nurse, P. and Cerrutti, L. (2002) The genome sequence of *Schizosaccharomyces pombe*. Nature 415, 871–880.
- [4] Dietrich, F.S., Voegeli, S., Brachat, S., Lerch, A., Gates, K., Steiner, S., Mohr, C., Pohlmann, R., Luedi, P., Choi, S., Wing, R.A., Flavier, A., Gaffney, T.D. and Philippsen, P. (2004) The *Ashbya gossypii* genome as a tool for mapping the ancient *Saccharomyces cerevisiae* genome. Science 304, 304–307.
- [5] Galagan, J.E., Calvo, S.E., Borkovich, K.A., Selker, E.U., Read, N.D., Jaffe, D., FitzHugh, W., Ma, L.J., Smirnov, S., Purcell, S., Rehman, B., Elkins, T., Engels, R., Wang, S., Nielsen, C.B., Butler, J., Endrizzi, M., Qui, D., Ianakiev, P., Bell-Pedersen, D., Nelson, M.A., Werner-Washburne, M., Selitrennikoff, C.P., Kinsey, J.A., Braun, E.L., Zelter, A., Schulte, U., Kothe, G.O., Jedd, G., Mewes, W., Staben, C., Marcotte, E., Greenberg, D., Roy, A., Foley, K., Naylor, J., Stange-Thomann, N., Barrett, R., Gnerre, S., Kamal, M., Kamvysselis, M., Mauceli, E., Bielke, C., Rudd, S., Frishman, D., Krystofova, S., Rasmussen, C., Metzenberg, R.L., Perkins, D.D., Kroken, S., Cogoni, C., Macino, G., Catcheside, D., Li, W., Pratt, R.J., Osmani, S.A., DeSouza, C.P., Glass, L., Orbach, M.J., Berglund, J.A., Voelker, R., Yarden, O., Plamann, M., Seiler, S., Dunlap, J., Radford, A., Aramayo, R., Natvig, D.O., Alex, L.A., Mannhaupt, G., Ebbole, D.J., Freitag, M., Paulsen, I., Sachs, M.S., Lander, E.S., Nusbaum, C. and Birren, B. (2003) The genome sequence of the filamentous fungus *Neurospora crassa*. Nature 422, 859–868.
- [6] Feldmann, H. (2000) Genolevures – a novel approach to 'evolutionary genomics'. FEBS Lett. 487, 1–2.
- [7] Kellis, M., Patterson, N., Endrizzi, M., Birren, B. and Lander, E.S. (2003) Sequencing and comparison of yeast species to identify genes and regulatory elements. Nature 423, 241–254.
- [8] Kellis, M., Birren, B.W. and Lander, E.S. (2004) Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae*. Nature 428, 617–624.
- [9] Martinez, D., Larrondo, L.F., Putnam, N., Gelpke, M.D., Huang, K., Chapman, J., Helfenbein, K.G., Ramaiya, P., Detter, J.C., Larimer, F., Coutinho, P.M., Henriksat, B., Berka, R., Cullen, D. and Rokhsar, D. (2004) Genome sequence of the lignocellulose degrading fungus *Phanerochaete chrysosporium* strain RP78. Nat. Biotechnol. 22, 695–700.
- [10] Brachat, S., Dietrich, F.S., Voegeli, S., Zhang, Z., Stuart, L., Lerch, A., Gates, K., Gaffney, T. and Philippsen, P. (2003) Reinvestigation of the *Saccharomyces cerevisiae* genome annotation by comparison to the genome of a related fungus: *Ashbya gossypii*. Genome Biol. 4, R45.

- [11] Zdobnov, E.M. and Apweiler, R. (2001) InterProScan – an integration platform for the signature-recognition methods in InterPro. *Bioinformatics* 17, 847–848.
- [12] Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) CLUSTALW: 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.
- [13] Corpet, F. (1988) Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* 16, 10881–10890.
- [14] Nielsen, H., Engelbrecht, J., Brunak, S. and Heijne, G. (1997) von Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng.* 10, 1–6.
- [15] Casadio, R., Polverini, E., Mariani, P., Spinozzi, F., Carsughi, F., Fontana, A., Polverino, de Laureto, Matteucci, G. and Bergamini, C.M. (1999) The structural basis for the regulation of tissue transglutaminase by calcium ions. *Eur. J. Biochem.* 262, 672–679.
- [16] Soanes, D.M., Skinner, W., Keon, J., Hargreaves, J. and Talbot, N.J. (2002) Genomics of phytopathogenic fungi and the development of bioinformatic resources. *Mol. Plant Microbe Interact.* 15, 421–427.
- [17] Sims, A.H., Gent, M.E., Robson, G.D., Dunn-Coleman, N. and Oliver, S.G. (2004) Combining transcriptome data with genomic and cDNA sequence alignments to make confident functional assignments for *Aspergillus nidulans* genes. *Mycol. Res.* 108, 853–857.
- [18] Reddy, C.A. and D'Souza, T.M. (1994) Physiology and molecular biology of the lignin peroxidases of *Phanerochaete chrysosporium*. *FEMS Microbiol. Rev.* 13, 137–152.
- [19] Delneri, D., Gardner, D.C. and Oliver, S.G. (1999) Analysis of the seven-member AAD gene set demonstrates that genetic redundancy in yeast may be more apparent than real. *Genetics* 153, 1591–1600.
- [20] Moralejo, F.J., Cardoza, R.E., Gutierrez, S., Lombrana, M., Fierro, F. and Martin, J.F. (2002) Silencing of the aspergillopepsin B (pepB) gene of *Aspergillus awamori* by antisense RNA expression or protease removal by gene disruption results in a large increase in thaumatin production. *Appl. Environ. Microbiol.* 68, 3550–3559.
- [21] Denning, D.W., Anderson, M.J., Turner, G., Latge, J.P. and Bennett, J.W. (2002) Sequencing the *Aspergillus fumigatus* genome. *Lancet Infect. Dis.* 2, 251–253.
- [22] Jones, T., Federspiel, N.A., Chibana, H., Dungan, J., Kalman, S., Magee, B.B., Newport, G., Thorstenson, Y.R., Agabian, N., Magee, P.T., Davis, R.W. and Scherer, S. (2004) The diploid genome sequence of *Candida albicans*. *Proc. Natl. Acad. Sci. USA* 101, 7329–7334.