Expressed sequence tags from the flower pathogen *Claviceps purpurea*

BIRGITT OESER^{1,*}, FRANÇOIS BEAUSSART², THOMAS HAARMANN^{3,†}, NICOLE LORENZ¹, EVA NATHUES¹, YVONNE ROLKE^{2,‡}, JAN SCHEFFER¹, JANUARY WEINER^{2,§} AND PAUL TUDZYNSKI¹

¹Institut für Botanik, Westf. Wilhelms Universität Münster, Schlossgarten 3, D 48149 Münster, Germany

²Institute for Evolution and Biodiversity, Westf. Wilhelms Universität Münster, Hüfferstr. 1, D 48149 Münster, Germany

³AB Enzymes GmbH, Feldbergstrasse 78, D 64293 Darmstadt, Germany

SUMMARY

The ascomycete *Claviceps purpurea* (ergot) is a biotrophic flower pathogen of rye and other grasses. The deleterious toxic effects of infected rye seeds on humans and grazing animals have been known since the Middle Ages. To gain further insight into the molecular basis of this disease, we generated about 10 000 expressed sequence tags (ESTs)—about 25% originating from axenic fungal culture and about 75% from tissues collected 6–20 days after infection of rye spikes. The pattern of axenic vs. *in planta* gene expression was compared. About 200 putative plant genes were identified within the *in planta* library. A high percentage of these were predicted to function in plant defence against the ergot fungus and other pathogens, for example pathogenesis-related proteins. Potential fungal pathogenicity and virulence genes were found via comparison with the pathogen–host interaction database (PHI-base; <http://www.phi-base.org>) and with genes known to be highly expressed in the haustoria of the bean rust fungus. Comparative analysis of *Claviceps* and two other fungal flower pathogens (necrotrophic *Fusarium graminearum* and biotrophic *Ustilago maydis*) highlighted similarities and differences in their lifestyles, for example all three fungi have signalling components and cell wall-degrading enzymes in their arsenal. In summary, the analysis of axenic and *in planta* ESTs yielded a collection of candidate genes to be evaluated for functional roles in this plant–microbe interaction.

INTRODUCTION

In the early 18th century, it was thought (incorrectly) that the ergot disease of rye and other grasses was caused by insect stings (Rapilly, 2001). It was later discovered that the biotrophic fungus *Claviceps purpurea*, which belongs to a large group of flower-infecting fungi, was the actual disease incitant. Within the flower-infecting fungi, *C. purpurea* belongs to a group that includes all fungal pathogens specialized in colonizing the host inflorescence and entering the host's ovary through the gynoeceal pathway (Ngugi and Scherm, 2006).

The development of *Claviceps* on the plant is summarized in Fig. 1. In spring, wind-dispersed ascospores are produced from ripe overwintering sclerotia at the same time as their main host—winter rye—exposes its stigmas to be pollinated. Sometimes a fungal spore instead of pollen lands on the thin-walled stigma, which is easier to penetrate than the leaves. After germination and penetration, the fungus grows towards the bottom of the ovary and taps into the phloem sap that normally fosters the development of seeds, but now supports the growth of a white fungal mass (sphacelium) that produces conidia drenched in a sugary fluid (honeydew). Secondly, fungal resting structures (sclerotia) replace the normal grains, whose remains are often observable on top of the ripe sclerotia.

Claviceps purpurea primarily affects the quality of harvested rye. Its sclerotia can cause severe damage to grazing animals and bread-eating humans because of their ergot-alkaloid content. However, ergot-alkaloids are also produced intentionally because of their beneficial pharmaceutical activities (Krska and Crews, 2008; Schardl *et al.*, 2006). Ergot-alkaloids can be harvested from infected rye fields or produced in fermentors, as *Claviceps* can be grown in axenic culture. In axenic culture, sclerotia are never formed. Completion of the life cycle requires a living host. Therefore, *C. purpurea* can be called an ecologically obligate biotroph.

*Correspondence: E-mail: oeser@uni-muenster.de

†Present address: Georg-Speyer-Haus, Institute for Biomedical Research, Paul-Ehrlich-Strasse 42-44, D 60596 Frankfurt am Main, Germany.

‡Present address: Center for Functional Genomics, Institute of Molecular BioSciences, Massey University, Private Bag 11 222, Palmerston North 4442, New Zealand.

§Present address: Institute for Bioinformatics, Faculty of Medicine, Westf. Wilhelms Universität Münster, Von-Esmarch-Str. 54, D 48149 Münster, Germany.

Fig. 1 (a) Infection phases of *Claviceps purpurea* on *Secale cereale*. (I) Few visible symptoms, ovary > sphacelium; 1–5 days post-inoculation (dpi). (II) Honeydew and conidia production, ovary < sphacelium; 6–12 dpi. (III) Sclerotium development started, sphacelium > sclerotium; 13–20 dpi. (IV) Sclerotium matures, sphacelium < sclerotium; 20–50 dpi. Occurrence and relative amounts of alkaloids, lipids and carbohydrates are represented as grey bars according to Corbett *et al.* (1974) and our own observations. (b) Mature, ergot-infected rye head with sclerotia.

For a long time, we have been interested in the characterization of the axenic and pathogenic lifestyles of *Claviceps*, and in the identification of putative pathogenicity genes in *Claviceps* and putative defence genes in its host *Secale cereale* (Tudzynski and Scheffer, 2004). Until now, plant defence reactions towards ergot have not been reported to occur in the field.

For the efficient discovery of genes, we generated and analysed expressed sequence tags (ESTs) originating from axenic culture and from *in planta* tissue. ESTs are widely used to explore the transcriptome (Nagaraj *et al.*, 2007). Currently, there are over 59 million ESTs from about 1600 species deposited in the dbEST database (<http://www.ncbi.nlm.nih.gov/dbEST/>). COGEME holds ESTs from 15 phytopathogenic and three saprophytic fungal species (Soanes and Talbot, 2006).

Of 9984 initial ESTs recovered from *C. purpurea* (one-quarter from axenic culture and three-quarters from *in planta* tissue), about 88% could be analysed. They were clustered, assembled into contigs, BLASTed to search for interesting genes and submitted to EMBL (Table S1, see Supporting Information). The axenic and *in planta* ESTs were also compared with each other. In addition, all ESTs were spotted on to nylon membranes for macroarray analyses (for example, Rolke and Tudzynski, 2008).

RESULTS AND DISCUSSION

Generation, clustering and assembly of ESTs

Three non-normalized cDNA libraries were generated. The early *in planta* library, corresponding to phase I of infection, was discarded because of its low quality (see Experimental Procedures). Phase I shows little visible sign of infection and consists primarily of host tissue (Fig. 1). All clones sequenced from this library showed similarity to plant genes (data not shown). The axenic (ax) library originated from a liquid culture of *C. purpurea*. The *in planta* (ip) library corresponded to phase II and III of infection, but should contain genes expected to be expressed in

phase IV, as sclerotium development initiates in phase III. Indeed, cDNAs from six genes involved in alkaloid biosynthesis (easA, easD-G, easH1) (Schardl *et al.*, 2006) were found in the ip library.

Sanger sequencing was performed as single-pass reads from the 5' end of each clone. The average length of read was 610 bases. After vector removal and sequence trimming, 8789 of the initial 9984 ESTs remained. They were clustered in order to link individual ESTs to unique genes, represented by fully, incomplete, alternatively spliced or even falsely joined transcripts. Assembly resulted in contigs, representing subsets of sequences in clusters, which overlapped without the insertion of large gaps in individual ESTs. For example, cluster 149 contained two contigs (0346 and 1224). Contig 1224 contained an intron, whereas Contig 0346 did not. In total, there were 1261 clusters and 2783 singlets. Of the 4887 contigs, 1325 contained more than one sequence.

Contigs were compared via BLASTX ($E \leq E-5$) against the non-redundant protein database of the National Center for Biotechnology Information (NCBI nr): 36.5% had no match, 3.7% matched ribosomal sequences, 15.1% showed similarity to a hypothetical protein and 44.7% to a protein with known function. In addition, 29.7% had no BLASTX match to nr (NCBI) and no BLASTN match to dbEST (NCBI). These contigs may be unique to *Claviceps*.

Comparison of the axenic and *in planta* libraries

ESTs originating from two dissimilar growth conditions (ax and ip) will differ in their composition, as the fungus will react to its different surroundings, expressing some of the same and different genes. For example, this can be seen if the most abundant ax and ip ESTs are considered (Fig. 2). About one-half of the most abundant ax ESTs are housekeeping genes: protein metabolism is represented by two ribosomal proteins, a ubiquitin fusion protein and a peptidyl-prolyl *cis-trans*-isomerase, and carbohydrate metabolism by a pyruvate decarboxylase. With one excep-

Fig. 2 Transcript abundance in the axenic (ax) and *in planta* (ip) *Claviceps* expressed sequence tag (EST) libraries. Light grey (dark grey) bars represent the percentage of ax (ip) sequences in the EST contigs [ax (ip) ESTs in a contig/total ax (ip) ESTs $\times$ 100]. On the left (right) side, the most abundant ax (ip) EST clones are shown. The x-axis shows the contig number and the result of a BLASTX ($E \leq E-5$) search against the non-redundant protein database of the National Center for Biotechnology Information (NCBI nr). ?, no BLASTX match; hyp prt, hypothetical protein; mt, mitochondrial; PDC, pyruvate decarboxylase; PPlase, peptidyl-prolyl *cis-trans*-isomerase; ribosomal prt, ribosomal protein; THI, thiamine biosynthesis protein; ubiquitin fus prt, ubiquitin fusion protein; (a) BLASTX hit to same hyp prt, but different nucleotide sequences; (b) in *Aspergillus nidulans* necessary for sexual development (Han *et al.*, 2008).

tion, all highly abundant ip ESTs showed either similarity with a hypothetical protein or no similarity to a known protein. Among them, Contig 1186 may encode a protein that has recently been shown to be necessary for sexual development in *Aspergillus nidulans* (Han *et al.*, 2008). It is conserved in filamentous fungi and may function as a developmental regulator. The only abundant ip EST with a putative function is Contig 0977 encoding a hydrophobin. Hydrophobins are small fungal proteins with roles in adhesion, development and pathogenesis. They were also found abundantly in EST libraries of *Trichoderma harzianum* (Vizcaino *et al.*, 2006), *Laccaria bicolor* and *Pisolithus microcarpus* (Peter *et al.*, 2003). Interestingly, the most abundant ax and ip ESTs do not match a known protein. Yet, there are matches to EST collections of other filamentous fungi (data not shown). Some ESTs are abundant in the ax or ip library and not uncommon in the other library [for example, ESTs in Contigs 1180 (ax), 1184 and 0655 (ip)].

For a more detailed comparison, gene ontology (GO) terms were assigned to each EST (see Experimental procedures, Sequence analysis, GO annotation). The GO terms were counted separately for ax and ip under the aspects 'Biological process' and 'Molecular function', and significant differences were noted (see Tables 1 and 2).

Considering biological process (Table 1), there were more counts for carbohydrate and lipid metabolism within the ip ESTs than within the ax ESTs, which is consistent with the knowledge that both types of metabolism are high in phases II and III of infection (see Fig. 1). The higher ip count for cell communication factors reflects interactions of fungal and host cells, in contrast

with simpler fungal growth in culture. Surprisingly, the count for cell organization and biogenesis was higher in culture than *in planta*. Differences in terms related to proteins and peptides are not easily explained; lower ip counts for protein metabolism, amino acid and derivative metabolism, and protein biosynthesis corresponded to higher ip counts for proteolysis, peptidolysis and protein modification.

Table 2 summarizes the molecular function GO term counts. More catalytic activity terms were found in ip than in ax, reflecting the greater complexity of fungal growth *in planta*. A more detailed examination showed that, for example, isomerase activity counts were even lower for ip ESTs. In some cases, counts appeared to be contradictory: for example, ip counts for oxidoreductase activity were higher than ax counts, whereas the opposite was observed for antioxidant activity, perhaps because oxidoreductases include enzymes that produce reactive oxygen species (ROS), such as superoxide dismutases or NADPH oxidases, as well as enzymes that destroy ROS, such as catalases. This result might reflect a lower count of antioxidant activity. No evidence of an oxidative burst was observed in *Claviceps*-infected rye tissue (Tudzynski and Scheffer, 2004), allowing the fungus to forego increased antioxidant activity. The count for signal transduction activity was higher in ip, consistent with the more complex milieu.

Several factors should be considered when comparing EST libraries by GO annotation. ESTs derived from the same gene could differ in length and thus appear to reflect differences between ax and ip libraries. Material for the ax library was collected from one flask, whereas material for the ip library was

GO ID	Definition	ax counts/%	ip counts/%	ip counts sign. diff.
GO:0005575	Cellular component	1.00	0.79	
GO:0019538	Protein metabolism	26.13	22.49	Down
GO:0006519	Amino acid and derivative metabolism	7.00	5.53	Down
GO:0006412	Protein biosynthesis (protein translation)	12.20	8.55	Down
GO:0006508	Proteolysis and peptidolysis	0.88	1.71	Up
GO:0006139	Nucleoside, nucleotide and nucleic acid metabolism	11.48	10.89	
GO:0005975	Carbohydrate metabolism	3.71	5.06	Up
GO:0006629	Lipid metabolism	2.27	5.56	Up
GO:0051186	Cofactor, cofactor metabolism	1.38	1.58	
GO:0009056	Catabolism	2.65	3.98	
GO:0019748	Secondary metabolism	0.39	0.77	
GO:0006118	Electron transport	1.22	1.13	
GO:0006091	Generation of precursor metabolites and energy	3.84	4.12	
GO:0007154	Cell communication	1.27	2.44	Up
GO:0016043	Cell organization and biogenesis	7.96	6.41	Down
GO:0006810	Transport	11.03	10.99	
GO:0050789	Regulation of biological process	2.32	2.23	
GO:0008104	Protein localization	1.71	1.98	
GO:0006464	Protein modification	1.02	2.9	Up
GO:0050896	Response to stimulus	0.28	0.52	
GO:0007049	Cell cycle	0.14	0.19	
GO:0008283	Cell proliferation	0.06	0.05	
GO:0007275	Developmental processes	0.03	0.07	
GO:0006955	Immunology, immune response		0.05	
GO:0046903	Secretion	0.03	0.01	

ax (ip) counts/%, ax (ip) counts as a percentage of total ax (ip) GO term counts for a specific GO ID; ip counts sign. diff., number of ip counts significantly different from number of ax counts for a specific GO ID; down, ax counts > ip counts; up, ax counts < ip counts.

collected over time from many different rye spikes, thus allowing for incidental exposure to physical factors, such as temperature change, that might affect transcription. Despite these caveats, it was clear that the number of ESTs without a predicted GO term was significantly higher in the ip library than in the ax library. Similarly, there were more ESTs matching a hypothetical protein, or with no BLASTX match, in the ip library than in the ax library. These observations suggest that novel genes may be expressed *in planta* by *Claviceps*.

Identification of plant genes in the ip library

Plant gene identification was complicated because the rye genome has not yet been sequenced. Each ax clone was considered *a priori* to be of fungal origin. Any ip clone similar to an ax clone, genomic *Clavicipitaceae* sequences or EST sequences from *Fusarium graminearum*, *Fusarium verticillioides* or *Magnaporthe grisea* (fungi related to *Claviceps*) was considered to be putatively fungal, whereas an ip clone similar to genomic *Triticeae* sequences or EST sequences from the grasses *S. cereale*, *Triticum aestivum*, *Hordeum vulgare* or *Oryza sativa* was putatively considered to be of host origin (see Experimental procedures, Sequence analysis, Assignment of plant or fungal origin for ip ESTs). Of the 6545 ip clones, 3032 (46%) could be classified. Two hundred and eight ip clones (191 contigs) may have originated

from the host plant (7% of the classifiable ones). All putative plant contigs were compared via BLASTX with the nr databank of NCBI (Table 3) to find possible functions. Forty-two contigs had no match in the nr databank, 60 matched a hypothetical protein and seven matched a ribosomal protein. One BLAST hit was the *S. cereale* large subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase, the plant CO₂-binding enzyme that catalyses the first step in the Calvin cycle, thus confirming its plant origin.

Contigs with putative functions in plant defence were noted, with the caveat that in no case was a defined role documented in this system. Some contigs contained components of the ubiquitin/26S proteasome protein degradation system; those found are not known to be required for resistance, but plants generally increase their ubiquitylation capacity after microbial attack (Dreher and Callis, 2007). Heat shock proteins (HSP20 of class I and HSP70) are considered to be part of a general stress response, whereas remorin (a plant-specific plasma membrane protein) responds to abiotic and biotic stimuli (Raffaele *et al.*, 2007). In addition, UDP-glucuronosyl transferases, which have a function in anthocyanin biosynthesis, have been suggested to play a role in stress tolerance (Gachon *et al.*, 2005). A UDP-glucuronosyl transferase was found to be a basal defence transcript in the interaction of wheat with *Puccinia striiformis* f. sp. *tritici* (Coram *et al.*, 2008). In addition, disulfide isomerase was found; it belongs to a protein family that responds in plants and

Table 1 Geneontology (GO) terms found in axenic (ax) and *in planta* (ip) expressed sequence tag (EST) clones—Biological process.

Table 2 Gene ontology (GO) terms found in axenic (ax) and *in planta* (ip) expressed sequence tag (EST) clones—Molecular function.

GO ID	Definition	ax counts/%	ip counts/%	ip counts sign. diff.
GO:0005488	Binding	18.26	15.81	
GO:0003676	Nucleic acid binding	12.70	10.17	Down
GO:0000166	Nucleotide binding	2.63	2.96	
GO:0005515	Protein binding	0.82	0.50	
GO:0003682	Chromatin binding	0.78	0.74	
GO:0005509	Calcium ion binding	0.52	0.32	
GO:0005102	Receptor binding	0.39	0.09	Down
GO:0008092	Cytoskeletal protein binding	0.04	0.23	
GO:0008289	Lipid binding		0.23	
	Other	0.39	0.56	
GO:0003824	Catalytic activity	55.34	70.09	Up
GO:0016740	Transferase activity	16.54	17.36	
GO:0016491	Oxidoreductase activity	14.51	23.23	Up
GO:0016787	Hydrolase activity	11.71	15.65	Up
GO:0016853	Isomerase activity	3.27	1.51	Down
GO:0016829	Lyase activity	2.24	3.68	Up
GO:0008233	Peptidase activity	2.20	3.81	Up
GO:0016874	Ligase activity	1.59	1.90	
GO:0004386	Helicase activity	0.78	1.17	
GO:0004721	Phosphoprotein phosphatase activity	0.60	0.92	
GO:0004518	Nuclease activity	0.22	0.38	
	Other	1.68	0.49	Down
GO:0005215	Transporter activity	7.32	7.20	
GO:0005386	Carrier activity	3.32	1.55	Down
GO:0015075	Ion transporter activity	2.20	1.94	
GO:0008565	Protein transporter activity	0.30	0.29	
GO:0015267	Channel or pore class transporter activity		0.07	
GO:0005216	Ion channel activity		0.07	
	Other	1.51	3.29	Up
GO:0004871	Signal transducer activity	2.20	3.52	Up
GO:0004872	Receptor activity	0.86	1.71	Up
	Other	1.34	1.81	
GO:0045182	Translation regulator activity	2.41	1.22	Down
GO:0030528	Transcription regulator activity	1.21	1.15	
GO:0016209	Antioxidant activity	0.65	0.31	Down
GO:0030234	Enzyme regulator activity	0.26	0.23	
GO:0003774	Motor activity	0.09		
GO:0005198	Structural molecule activity	12.27	9.41	Down
GO:0003735	Structural constituent of ribosome	10.81	8.10	Down
GO:0005200	Structural constituent of cytoskeleton	0.52	0.23	
	Other	0.95	1.08	

Main categories are written in bold letters, subcategories not. ax (ip) counts/%, ax (ip) counts as a percentage of total ax (ip) GO term counts for a specific GO ID; ip counts sign. diff., number of ip counts significantly different from number of ax counts for a specific GO ID; down, ax counts > ip counts; up, ax counts < ip counts.

animals to stress conditions associated with enhanced chaperone activity. The enzyme has been investigated in two interactions of a grass and a hemibiotrophic fungus (wheat—*Mycosphaerella graminicola* by Ray *et al.*, 2003; rice—*M. grisea* by Shim *et al.*, 2004). General response elements involving signalling were found: two calmodulins and a calmodulin-like protein. Calmodulins may play a role in defence-related calcium signalling (Yang and Poovaiah, 2003). In addition, a stearyl-ACP desaturase was found, an archetypical member of a family of soluble fatty acid desaturases. Stearyl-ACP desaturase mutants in *Arabidopsis thaliana* have been shown to be impaired in the jasmonic acid defence signalling pathway (Kachroo *et al.*, 2001,

2007). Also identified was *S*-adenosylmethionine decarboxylase, a key polyamine biosynthetic enzyme. Polyamines could be involved in triggering the hypersensitive response (HR) (Walters, 2003).

A connection for HR and programmed cell death (PCD) was suggested for two other contigs. Contig 3869 shows similarity to hypersensitive-induced reaction protein 3 from barley, which is increased 35-fold in mutants with a spontaneous HR phenotype (Rostoks *et al.*, 2003). Contig 3791 may encode a tritacin-like cysteine protease. Cysteine proteases have been connected with some features of plant defence, including PCD (Goodwin *et al.*, 2004; Matarasso *et al.*, 2005). Another putative protease shows

Table 3 *Claviceps* expressed sequence tag (EST) contigs of putative plant origin. Itemized are the results of BLAST searches against the non-redundant protein database of the National Center for Biotechnology Information (NCBI nr) and two self-assembled EST banks (flower, infected flower). Full data are only given for contigs with a meaningful hit. Contigs encoding proteins with a possible role in defence against pathogens are listed with reference.

Contig	BLASTX—nr (NCBI)	best meaningful BLASTX hit	BLASTX—flower EST bank*			BLASTX—infected flower EST bank*			Reference
			E-value	seqs/ctg	Best hit	E-value	Best hit	E-value	
0115	BMV2: β -amylase [At]		1.00E-26	2	TC304553_Os	1.00E-103			
2495	Vacuolar ATP synthase subunit d [NcIS]		3.00E-24	1	TC2533_Sc	0			
2870	SMC1 protein [Os]		1.00E-57	1	TC252752_Ta	0			
2900	Transcription factor X1 [Tm]		5.00E-146	1	TC149560_Hv	0			
3099	Copine-like protein [Os]		2.00E-72	1	TC292058_Os	1.00E-86			
3359	SMC1 protein [Os]		1.00E-94	1	TC252752_Ta	0			
3751	Putative ctp-like energy-dependent protease [Os]		9.00E-85	1	TC292797_Os	1.00E-141			
4314	DNA-directed RNA polymerase III largest subunit [Os]		8.00E-67	1	TC284638_Os	9.00E-53			
4624	Plectin-related protein-like [Os]		3.00E-65	1	TC287258_Os	6.00E-74			
0009	Calmodulin [Os]		4.00E-78	2	TC234402_Ta	0	TC234402_Ta	0	Signalling
0163	Triosephosphate isomerase, cytosolic [Sc]		6.00E-121	2	TC2431_Sc	0	TC248637_Ta	0	Yang and Poovaliah (2003)
0454	Profilin-1 [Hv]		2.00E-61	2	TC3163_Sc	0	TC247940_Ta	0	
0688	Glyceraldehyde 3-phosphate dehydrogenase B subunit [At]		4.00E-07	2	TC138635_Hv	1.00E-143	TC138635_Hv	1.00E-143	
0753	Histone H2A.7 histone [Ta]		2.00E-42	2	TC234709_Ta	0	TC235032_Ta	0	
0763	Digalactosyl diacylglycerol synthase 1 [Os]		2.00E-24	2	TC147860_Hv	1.00E-168	TC262935_Ta	1.00E-172	Ray <i>et al.</i> (2003); Shim <i>et al.</i> (2004)
0866	Protein disulphide isomerase [Os]		3.00E-124	2	TC139204_Hv	0	TC264921_Ta	0	Chaperone
0905	Aspartate aminotransferase [Os]		2.00E-68	2	TC232219_Ta	0	TC232219_Ta	0	
0947	S-Adenosylmethionine decarboxylase precursor [Ta]		0	2	TC3068_Sc	0	TC264559_Ta	0	Polyamine synthesis
1074	Adenosine kinase-like protein [Os]		8.00E-157	2	TC247506_Ta	0	TC247506_Ta	0	Walters (2003)
2212	Glycoprotein [Bi]		2.00E-59	1	TC2594_Sc	0	TC247933_Ta	0	
2215	Outer mitochondrial membrane protein porin [Ta]		1.00E-106	1	TC146872_Hv	0	TC250606_Ta	0	Protein degradation
2272	Polyubiquitin [Ss]		7.00E-45	1	TC234281_Ta	0	TC234209_Ta	0	Dreher and Callis (2007)
2279	Histone H4 [Ho]		1.00E-38	1	TC233179_Ta	0	TC250187_Ta	1.00E-151	
2365	MRP-like ABC transporter [Os]		1.00E-129	1	TC354124_Os	0	TC237754_Ta	0	
2369	Boron transporter [Os]		1.00E-22	1	TC252478_Ta	1.00E-173	TC252478_Ta	1.00E-173	
2392	DEAD/DEAH box helicase, putative [Os]		5.00E-123	1	TC307499_Os	0	TC271590_Ta	8.00E-52	
2415	Glutamine synthetase isoform GS1b [Ta]		4.00E-136	1	TC139434_Hv	0	TC264890_Ta	0	Dreher and Callis (2007)
2458	Polyubiquitin [Ss]		9.00E-102	1	TC2362_Sc	0	TC234234_Ta	0	Protein degradation
2560	LON1 protease [Ta]		2.00E-129	1	TC131081_Hv	0	TC235913_Ta	0	
2573	RuBisCO large subunit-binding protein subunit beta [Sc]		8.00E-120	1	TC139132_Hv	0	TC249610_Ta	0	
2683	Glucose-6-phosphate dehydrogenase [Ta]		6.00E-64	1	TC139350_Hv	0	TC251298_Ta	0	
2687	Metallothionein [Ta]		3.00E-07	1	TC246696_Ta	1.00E-127	TC246696_Ta	1.00E-127	Zhou <i>et al.</i> (2006)
2699	16.9-kDa class I heat shock protein [Ta]		1.00E-70	1	TC249526_Ta	0	TC140216_Hv	0	Vierling (1991); Wang <i>et al.</i> (2004)
2768	Elongation factor 2 [Os]		3.00E-116	1	TC3094_Sc	0	TC246958_Ta	0	Signalling
2771	Calmodulin TaCam1-1 [Os]		5.00E-77	1	TC330697_Os	0	TC237135_Ta	0	Yang and Poovaliah (2003)
2802	Superoxide dismutase 1, putative [Os]		9.00E-75	1	TC2466_Sc	0	TC262714_Ta	0	Mittler <i>et al.</i> (2004)
2837	α -Glucosidase II [Os]		2.00E-126	1	TC217731_Os	0	TC133712_Hv	0	
2901	AAA-superfamily of ATPases [Os]		1.00E-54	1	TC301332_Os	7.00E-55	TC271109_Ta	9.00E-86	
3014	Ubiquitin-like protein SMT3 [Os]		7.00E-34	1	TC2462_Sc	0	TC250842_Ta	0	Protein degradation
3101	26S proteasome subunit 4-like protein [Os]		6.00E-112	1	TC146596_Hv	0	TC146596_Hv	0	Dreher and Callis (2007)
3132	Remorin [Os]		1.00E-22	1	TC329748_Os	5.00E-68	TC252212_Ta	3.00E-67	Response to (a)biotic stimuli
3172	APETALA3-like protein [Hv]		2.00E-34	1	TC2593_Sc	1.00E-137	TC238728_Ta	1.00E-119	
3228	H-ATPase [Os]		3.00E-63	1	TC247104_Ta	0	TC248557_Ta	0	
3239	Cytosolic glutathione peroxidase [Tm]		2.00E-92	1	TC131656_Hv	0	TC246747_Ta	0	Mittler <i>et al.</i> (2004)
3288	Cysteine protease [Ta]		2.00E-11	1	TC232622_Ta	1.00E-88	TC232303_Ta	7.00E-91	
3374	Chitinase 1 precursor, type III [Os]		3.00E-111	1	TC147411_Hv	0	TC266129_Ta	0	van Loon <i>et al.</i> (2006)
3378	Ascorbate peroxidase [Hv]		2.00E-120	1	TC233791_Ta	0	TC232978_Ta	0	Mittler <i>et al.</i> (2004)

3477	Thioredoxin h isoform 2; HvTrxh2 [Hv]	2.00E-52	1	TC3164_Sc	0	TC267391_Ta	0	ROS scavenger	Mittler <i>et al.</i> (2004)
3502	Glucose-6-phosphate dehydrogenase [Ta]	6.00E-82	1	TC139350_Hv	0	TC251298_Ta	0		
3535	26S protease regulatory subunit 7 [Os]	4.00E-134	1	TC2796_Sc	0	TC246888_Ta	0	Protein degradation	Dreher and Callis (2007)
3545	Polyubiquitin 2 [Da]	3.00E-66	1	TC92012_Sb	0	TC232408_Ta	0	Protein degradation	Dreher and Callis (2007)
3559	Myo-inositol 1-phosphate synthase [Zm]	2.00E-69	1	TC2647_Sc	0	TC246900_Ta	0		
3706	Putative heat shock 70 KD protein, mitochondrial [Os]	1.00E-15	1	TC139412_Hv	1.00E-79	TC264160_Ta	1.00E-113	General stress response	Vierling (1991); Wang <i>et al.</i> (2004)
3714	Lipid transfer protein-like protein 2 precursor [Sc]	4.00E-35	1	TC265011_Ta	5.00E-67	TC234771_Ta	1.00E-75	PR protein	van Loon <i>et al.</i> (2006)
3791	Triticin β [Ta]	2.00E-38	1	TC147127_Hv	1.00E-150	TC233092_Ta	1.00E-156	Programmed cell death and others	Goodwin <i>et al.</i> (2004); Matarasso <i>et al.</i> (2005)
3848	Stearyl-ACP desaturase [Os]	1.00E-66	1	TC139428_Hv	0	TC235225_Ta	0	Signalling	Kachroo <i>et al.</i> (2007)
3858	Ubiquitin [Os]	8.00E-66	1	TC2357_Sc	0	TC249401_Ta	0	Protein degradation	Dreher and Callis (2007)
3869	Hypersensitive-induced reaction protein 3 [Hv]	8.00E-33	1	TC247941_Ta	1.00E-170	TC247941_Ta	1.00E-171	Hypersensitive response	Rostoks <i>et al.</i> (2003)
3997	Proteinase inhibitor-related protein [Ta]	3.00E-36	1	TC234467_Ta	1.00E-158	TC234468_Ta	1.00E-178	PR protein	van Loon <i>et al.</i> (2006)
4134	2-Oxoglutarate/malate translocator [Pm]	1.00E-56	1	TC3237_Sc	0	TC264258_Ta	0		
4137	Ribonucleotide reductase [Os]	6.00E-06	1	TC3660_Sc	8.00E-63	TC265616_Ta	1.00E-124		
4163	Glutathione-S-transferase, class Tau [Os]†	5.00E-75	1	TC326296_Os	4.00E-53	TC265501_Ta	0	ROS scavenger	Dixon <i>et al.</i> (2002)
4204	Histone H2B.2 [Os]	1.00E-39	1	TC3095_Sc	0	TC264367_Ta	0		
4217	Ubiquitin-conjugating enzyme [Ta]	1.00E-18	1	TC2608_Sc	0	TC232178_Ta	1.00E-135	Protein degradation	Dreher and Callis (2007)
4326	S-Adenosyl-L-homocysteine hydrolase [Hv]	9.00E-151	1	TC139066_Hv	0	TC264652_Ta	0		
4345	Small Ras-related GTP-binding protein [Ta]	1.00E-113	1	TC3107_Sc	0	TC234441_Ta	0		
4356	Elongation factor 1 β [Hv]	2.00E-84	1	TC131466_Hv	0	TC251778_Ta	0		
4388	Cap-binding protein [Os]	3.00E-36	1	TC251778_Ta	0	TC251778_Ta	0		
4390	Class III alcohol dehydrogenase enzyme [Os]	3.00E-94	1	TC263255_Ta	0	TC262835_Ta	0		
4528	Small nuclear ribonucleoprotein homologue [Os]	2.00E-39	1	TC341008_Os	1.00E-103	TC249742_Ta	0		
4529	Cytochrome P450 like_TBP [Nt]	2.00E-11	1	TC262719_Ta	3.00E-80	TC262719_Ta	1.00E-80		
4568	Ubiquitin carrier protein [Os]	5.00E-82	1	TC3331_Sc	0	TC247094_Ta	0		
4574	Endoplasmic reticulum chaperone precursor [Hv]	3.00E-29	1	TC264521_Ta	0	TC264521_Ta	0		
4673	Phospho-2-dehydro-3-deoxyheptonate aldolase [Os]	2.00E-11	1	TC234760_Ta	0	TC234760_Ta	0		
4680	Phytopsin precursor [Hv]	2.00E-139	1	TC3124_Sc	0	TC232293_Ta	0	Diverse	Schaller (2004)
4721	Alanine aminotransferase 2 [Hv]	1.00E-114	1	TC146731_Hv	0	TC264334_Ta	0		
4762	1,4- α -D-glucan 6- α -D-(1,4- α -D-glucanotransferase [Ta]	3.00E-90	1	TC139565_Hv	0	TC236593_Ta	0		
2956	Prolyl 4-hydroxylase α subunit homologues [Os]	2.00E-80	1			TC254333_Ta	0	Cell wall fortification	Deepak <i>et al.</i> (2007)
3042	FAD-linked oxidase family protein [At]‡	1.00E-92	1			TC140734_Hv	1.00E-105	ROS producer	Taler <i>et al.</i> (2004)
2431	Ubiquitin-associated protein [At]	9.00E-26	1					Protein degradation	Dreher and Callis (2007)
2913	Single-strand DNA repair-like protein [Tm]	2.00E-19	1						
3280	UPL4 (ubiquitin-protein ligase 4) [At]	7.00E-25	1					Protein degradation	Dreher and Callis (2007)
3537	UDP-glucuronosyl/UDP-glucosyltransferase family protein [Os]	5.00E-12	1					Stress response	Gachon <i>et al.</i> (2005)
3579	Elongation factor Tu, mitochondrial precursor [Ate]¶	4.00E-87	1						
3854	Calmodulin-like [Os]	2.00E-35	1					Signalling	Yang and Poovaiah (2003)
3944	Plastid division regulator MinE [Hv]	4.00E-20	1						
4168	Lipolytic enzyme, G-D-S-L family protein [Os]	2.00E-30	1						
4195	Thaumatin-like protein [Ta]	3.00E-69	1					PR protein	van Loon <i>et al.</i> (2006)
4539	Putative RNA apurinic site specific lyase [Os]	6.00E-21	1						

In addition: 7 contigs—ribosomal proteins; 57 contigs—hypothetical proteins (four hits only with infected flower EST bank); 38 contigs—no match (eight hits only with infected flower EST bank).

Organism abbreviations: At, *Arabidopsis thaliana*; Ate, *Aspergillus terreus*; Bi, *Bromus inermis*; Da, *Deschampsia antarctica*; Ho, *Hyacinthus orientalis*; Hv, *Hordeum vulgare*; Nc, *Neurospora crassa*; Nt, *Nicotiana tabacum*; Os, *Oryza sativa*; Pm, *Panicum miliaceum*; Sb, *Sorghum bicolor*; Sc, *Secale cereale*; Sp, *Sporobolus staphaneus*; Ta, *Triticum aestivum*; Tm, *Triticum monococcum*; Zm, *Zea mays*.

*Self-assembled EST banks (see Experimental procedures) ('TC identifier' abbreviated organism).

†Glutathione peroxidase activity/often induced by infection.

‡GO:0008891 glycolate oxidase activity.

§Probably joined cDNA.

¶Match for typing as 'putatively plant': WHEAT-146134, 1.00E-147.

similarity to phytepsin, an aspartic proteinase implicated in defence against microbial pathogens and in pollen–pistil interaction (Schaller, 2004). A more direct effect may involve the action of a prolyl-4-hydrolase, which hydroxylates proline-rich structural glycoproteins of the cell wall, contributing to the physical barrier against incoming pathogens. An increase in such glycoproteins was found in pearl millet infected with *Sclerospora* and in wheat or maize infected with *Fusarium* (Deepak *et al.*, 2007). ROS typify the plant–pathogen interaction (Mittler *et al.*, 2004; Torres *et al.*, 2006). Several putative ROS scavengers were found: superoxide dismutase, glutathione peroxidase, ascorbate peroxidase, thioredoxin and metallothionin. A class tau glutathione-S-transferase was also found, which may be a ROS scavenger, as some theta, pi and tau glutathione-S-transferases have been shown to have glutathione peroxidase activity (Dixon *et al.*, 2002). Only one possible ROS producer was found, a flavin adenine dinucleotide (FAD)-linked oxidase family protein, which has D-lactate dehydrogenase as well as glycolate oxidase activity according to GO classification. Glycolate oxidase, a key enzyme of photorespiration and an H₂O₂ producer, has been associated with the protection of melon against downy mildew (Taler *et al.*, 2004). Something similar was observed by Bohman *et al.* (2002) for *Leptosphaeria* infecting a chimeric *Brassica/Arabidopsis* plant. The finding of ROS-related EST contigs is interesting, because rye ovaries exhibit an oxidative burst after infection with a CREB-like transcription factor knockout mutant (Nathues *et al.*, 2004) and a small GTPase knockout mutant of *C. purpurea* (Scheffer *et al.*, 2005). Pathogenesis-related (PR) proteins (van Loon *et al.*, 2006) are common responses of plants to attack. Within the ip clones, there appear to be a PR-5 (a thaumatin-like protein), a PR-6 (a protease inhibitor of the Bowman–Birk type), a PR-8 (a chitinase type III) and a PR-14 (a non-specific lipid transfer protein).

Interestingly, pathogenicity-related EST sequences were also found in the transcriptome analysis of the close relative of *C. purpurea*, *Neotyphodium* (Felitti *et al.*, 2006; <http://hornbill.cspg.latrobe.edu.au/cgi-bin/pub/endophyte/index.pl>). Some ESTs are found only in an interaction of *Neotyphodium* with a grass. Some most probably originate from the plant partner. As in the *Claviceps*–rye interaction, a Bowman–Birk-type wound-induced proteinase inhibitor and non-specific lipid transfer proteins were found. In addition, there are hits to germin, known to be involved in the cereal defence response against invading fungi (Lane, 2002), and the PR protein AOPR1 from *Asparagus officinalis*. Of course, *Neotyphodium* (the anamorph of *Epichloë*) is considered by some to be a harmless endophyte, despite the fact that its alkaloids poison grazing livestock.

Because *Claviceps* is a flower pathogen, it is intriguing that many PR proteins were found to be constitutively present in floral tissues. In general, one finds the expression of defence-related genes in flowers (Laitinen *et al.*, 2005; Liljeroth *et al.*, 2005; Tung

et al., 2005). Flowers would otherwise be extremely good places for a pathogen to attack, as pistils have no cuticle or only a disrupted one. Defence-related contigs were relatively abundant (19% of all putative plant contigs, 40% of all those with an assigned possible function). Host tissue in this disease is of flower origin, distinct because it must encourage pollen tube penetration but repel pathogens. To address this, all putative plant contigs were compared via BLASTN with two self-assembled EST banks of grasses. The 'flower' bank comprised ESTs from pollen, anthers, pollinated pistils, pollinated spikes, spikelets and panicles of millet, rye, barley, rice, wheat and maize. The 'infected flower' bank contained ESTs from spikes and heads of barley and wheat infected with the necrotroph *F. graminearum* (Experimental procedures, Sequence analysis). The largest group in Table 3 showed matches to ESTs in both banks, seeming to confirm the preparedness of flowers for pathogens. Most (79%) of all putative plant contigs showed significant similarity to the grass flower EST bank, which is consistent with our observations of infected rye spikes. Some contigs in Table 3 had no hit to either flower EST bank, but several of these contigs had a possible defence role (for example, Contigs 3537 and 4195). There were just two contigs with a hit only to the infected flower EST bank (Contigs 2956 and 3042). This interesting group can be enlarged by four contigs matching hypothetical proteins and eight contigs with no match at all. The last group is small, showing hits only to the flower EST bank and containing no contig with a possible role in defence.

Ergot is a replacement disease and our EST library does not contain clones from the first 5 days (phase I, Fig. 1), when the ovaries are barely invaded by the fungus. Therefore, few plant genes were expected within the library. Any plant genes found are interesting as the harvested part (infected ovary) included the rachilla, the site at which fungus and plant form a stable plant–pathogen interface. The defence genes found may be there because flowers by themselves are always on guard or because rye—contrary to current doctrine—recognizes the invading fungus and the fungus is able to suppress and balance the plant defences. In the interaction of *Ustilago* and maize, it was thought that classical defence pathways were not activated, but Doeblemann *et al.* (2008) have shown recently that, after the infection of seedlings, *U. maydis* is recognized early and triggers defence responses.

Identification of pathogenicity genes

To find potential pathogenicity and virulence genes in our EST collection, we compared the proteins of the pathogen–host interaction database (PHI-base; (Baldwin *et al.*, 2006; Winnenburg *et al.*, 2008) with our *Claviceps* ESTs. PHI-base contains curated information on pathogenicity, virulence and effector genes of mostly fungal pathogens. We proceeded on the assumption that the proof of pathogenicity/virulence function of

a gene in one fungus also suggests its pathogenicity/virulence function in other fungi (Baldwin *et al.*, 2006).

In accordance with this assumption, the search of PHI-base proteins yielded several already characterized *C. purpurea* molecules. Eight EST contigs were identical with already cloned genes (Contigs 0213, 0923, 2567, 2751, 1398, 0448, 1323 and 1513). Six of these were either pathogenicity or virulence factors: mitogen-activated protein (MAP) kinases cpmk1 and cpmk2, PAK kinase cla4, histidine kinase cphk2, polygalacturonase 2 and bZip transcription factor cptf1. Three EST contigs were related to already cloned genes (Contigs 2301, 0148 and 3103), two of which were pathogenicity factors (serine/threonine kinase cot1 and Rho GTPase cdc42).

The assumption is also valid when comparing fungal plant and animal pathogens. Sexton and Howlett (2006) found—based on the December 2006 PHI-base content and with a focus on ascomycetes—that there are parallels in the infection mechanisms used by fungal plant and animal pathogens. Many of those highlighted by Sexton and Howlett (2006) were also found as potential virulence or pathogenicity genes in the pool of *Claviceps* ax and ip ESTs (Table 4). During germination and during *in planta* growth, signalling components, such as MAP kinase, guanosine triphosphatases (GTPases) and cyclic adenosine monophosphate (cAMP)-dependent kinases, are active (Table 5, 'Signalling' section). Isocitrate lyase is thought to be involved in providing the pathogen with nutrition (Table 4, 'Primary metabolism' section). Some hydrophobins are needed for infection (Table 4, 'Hydrophobins and related' section), although that found in *C. purpurea* is a pentahydrophobin, already proven not to affect pathogenicity under the conditions investigated. The biosynthesis of melanin—a fungal cell wall component—is often important in the penetration stage (Table 4, 'Melanin synthesis' section). In addition, other fungal cell components, for example chitin or β -glucan, are important (Table 4, 'Fungal cell wall synthesis/modification' section). Hydrolytic enzymes are at least partly essential during infection (Table 4, 'Plant cell wall-degrading enzymes' and 'Proteases' sections). ROS-scavenging enzymes may restrict host-produced ROS damage or tune ROS-dependent signal transduction (Table 4, 'ROS and related' section). Our hypothesis is that *Claviceps* produces ROS, but keeps net ROS production at a low level (Nathues *et al.*, 2004).

Of the 127 EST contigs found (Table 4), 19 originate from the ax bank, 28 from both the ax and the ip bank, and 80 from the ip bank. Given that there are about three times more ip clones as ax clones, the frequency of hits among the ip contigs seems higher than expected. Table 4 shows the genes worth testing for roles in pathogenicity/virulence of *Claviceps*, yet functional analysis via knockout is necessary to validate the true pathogenicity/virulence genes.

A comparative analysis of EST collections between ascomycete pathogens and non-pathogens showed that ESTs from

certain functional categories are more highly represented in pathogenic than non-pathogenic species: for example, ESTs concerning cell rescue, defence, death and ageing (Soanes and Talbot, 2006). This category includes ESTs involved in PCD. As, in filamentous fungi, PCD is involved in development, differentiation and sometimes also pathogenicity (for example, in *M. grisea*: Liu *et al.*, 2007; Veneault-Fourrey *et al.*, 2006; in *Colletotrichum gloeosporioides*: Barhoom and Sharon, 2007; review: Ramsdale, 2008), we looked for potential PCD ESTs in our libraries (Table 5). We found 10, most of which are solely expressed *in planta*. Two are of particular interest: putative autophagy proteins Atg8 and Atg13. In yeast, Atg13 is linked to nutrient sensing via the protein kinase Tor (Meijer *et al.*, 2007). During infection, *C. purpurea* displays targeted growth towards the rachilla, which is connected to the phloem and provides an endless supply of sugars. Atg8 in *M. grisea* is required for pathogenicity. Infection requires high turgor pressure in appressoria, which relies on autophagic PCD of the germinating spores (Veneault-Fourrey *et al.*, 2006). 'Classical' appressorium formation was not observed in *Claviceps*, but turgor cannot yet be ruled out as a factor in the penetration process.

Genes involved in nutrient uptake were of special interest. Some biotrophs, for example *Uromyces*, produce haustoria as nutrient-absorbing organs, but these are not observed in *Claviceps*, which develops intracellular hyphae that may have haustorium function (Tenberge *et al.*, 1996). Although we did not make a library mainly isolated from cells at the plant–fungus interface, the plant–fungus interface was part of our collected tissue, and it was instructive to compare our ESTs with highly expressed haustorial genes in the well-investigated interaction of the biotroph *Uromyces fabae* with *Vicia faba* (Jakupović *et al.*, 2006) (see Table 6). Of the 35 *Ur. fabae* haustorial genes whose expression increased five times or more in rust-infected leaves vs. germinated uredospores, 14 (40%) showed a significant hit to a *Claviceps* EST contig. A similar search with haustorial EST clones from *Blumeria graminis* (an obligate grass biotroph) that are up-regulated five-fold *in planta* yielded no matches to *Claviceps* ESTs or to the *Uromyces* up-regulated haustorial genes (data not shown). This could be a reflection of a different host colonization mode. *Blumeria* has epicuticular hyphae with intracellular haustoria, whereas both *Uromyces* and *Claviceps* initially grow with intracellular hyphae and later with intercellular hyphae. The identification of genes encoding a sugar transporter and a mannitol dehydrogenase—via a hit to the haustorial rust genes—indicates the importance of sugars and sugar polyols in ergot disease. Sugar and sugar polyols are constituents of the honeydew produced in phase II of the infection. The enzyme mannitol dehydrogenase may have a secondary role as a defence molecule against ROS produced by the host (Solomon *et al.*, 2007). In addition, manganese superoxide dismutase may have an antioxidative role. Glutamine synthetase could play a

Table 4 *Claviceps* expressed sequence tag (EST) contigs with significant homology to pathogen–host interaction database (PHI-base) entries. The number and origin of the EST clones assembled in each contig are indicated: ax (ip) seqs, axenic (*in planta*) sequences.

				Best PHI-base match				Other PHI-base matches					
	Contig	ax seqs	ip seqs	Putative function	PHI	Species*	Host†	Phenotype‡	HSP	E-value	# PHI	Hosts§	Phenotypes¶
				Signalling									
	3832		1	G α subunit	1013	<i>F. graminearum</i>	gr	vir	157	7.00E-40	7	2 gr/3 pl/2 an	vir/path/unaff
	3221		1	MAPKK kinase	1016	<i>F. graminearum</i>	gr/pl	path	408	E-115	5	3 gr/ 2an	vir/path
	0213		2	MAP kinase (cpmk1)	245††	<i>C. purpurea</i>	gr	path	274	E-145	15	7 gr/6 pl/2 an	path/vir/unaff
	0923	1	1	MAP kinase (cpmk2)	246††	<i>C. purpurea</i>	gr	vir	765	0	5	2 gr/2 pl/2 an	vir/path
	2567		1	PAK kinase	cla4***††	<i>C. purpurea</i>	gr	path	451	E-127	4	2 gr/3 an	vir/path/unaff/lethal
	2751		1	Histidine kinase	cphk2***††	<i>C. purpurea</i>	gr	vir	114	5.00E-26			
	2516		1	HOG1-type MAP kinase	153	<i>M. grisea</i>	gr	unaff	375	E-105	4	1 gr/2 pl/2 an	vir/unaff
	3498		1	PKA, catalytic subunit	36	<i>M. grisea</i>	gr	path	572	E-164	6	1 gr/2 pl/3 an	path/vir
	0806	5	2	G β subunit	334	<i>C. heterostrophus</i>	gr	path	90	2.00E-19	2	2 pl	vir
	2301		1	Serine/threonine kinase	158	<i>U. maydis</i>	gr	vir	218	1.00E-57	1††	1 gr	path
	0738		2	Guanyl nucleotide exchange factor	319	<i>U. maydis</i>	gr	vir	117	6.00E-27			
	1835	1		Adenylate cyclase	332	<i>C. lagenarium</i>	pl	path	198	3.00E-51	2	1 gr/1 pl	vir
	3604		1	PKA, regulatory subunit	231	<i>C. lagenarium</i>	pl	path	134	1.00E-32	1	1 gr	unaff
	0276		3	Rab/GTPase	339	<i>C. lindemuthianum</i>	pl	vir	286	5.00E-98			
	0148		3	Rho GTPase	270	<i>C. albicans</i>	an	path	300	4.00E-83	4††	2 gr	path/vir
	0124		2	CDC2-related/cydlin-dependent protein kinase	172	<i>C. albicans</i>	an	path	169	7.00E-43			
	1564	1		Type 2A-related protein phosphatase	378	<i>C. albicans</i>	an	vir	181	5.00E-47			
	3000		1	Phospholipase D	373	<i>C. albicans</i>	an	vir	228	2.00E-60			
	0270	1	2	Ras GTPase	182	<i>C. neoformans</i>	an	vir/path	293	5.00E-81	3	2 gr/1 an	vir/path
	1031	13	7	Cyclophilin	213	<i>C. neoformans</i>	an	vir	266	4.00E-73	2	1 gr/1 pl	vir
	1471	1		Calcineurin A catalytic subunit	89	<i>C. neoformans</i>	an	path	264	1.00E-71	1	1 an	vir
	3516		1	Calcineurin B regulatory subunit	474	<i>C. neoformans</i>	an	path	120	5.00E-31	1	1 an	vir
	2970		1	Rho guanyl-nucleotide exchange factor	675	<i>C. neoformans</i>	an	hypvir	276	7.00E-75			
	1118		2	Rheb GTPase	317	<i>A. fumigatus</i>	an	vir	228	2.00E-61			
	2239		1	Fatty acid oxygenase	494	<i>A. fumigatus</i>	an	hypvir	117	3.00E-27			
	3582		1	Inositolphosphorylceramide synthase	218	<i>C. neoformans</i>	an	vir	229	2.00E-61			
				Transporter									
	2879		1	ABC transporter	391/1018	<i>M. grisea</i>	gr	unaff	292	1.00E-79	3	2 gr/ 2pl	vir/unaff
	0188		2	Monocarboxylate permease	812	<i>M. grisea</i>	gr	vir	90	2.00E-19			
	4063		1	MFS transporter	737	<i>C. nicotianae</i>	pl	vir	151	7.00E-38	1	an	vir
	2267		1	Transporter	544	<i>B. cinerea</i>	pl	unaff	244	7.00E-66	2	1 gr/1 pl	vir/lethal
	1245	1	11	Mitochondrial carrier protein	254	<i>F. oxysporum</i>	pl	vir	545	E-156	1	1 gr	vir
	3299		1	Sugar transporter	689	<i>C. neoformans</i>	an	unaff	115	6.00E-27	1	1 pl	unaff
	2452		1	Voltage-gated chloride channel	286	<i>C. neoformans</i>	an	vir	123	4.00E-29			
	1365	1		Ferrichrome-type siderophore transporter	513	<i>C. albicans</i>	an	vir/unaff	214	1.00E-56			
	0997	1	2	Iron permease	162/485	<i>C. albicans</i>	an	path	176	3.00E-45			
	1015	1	1	Vacuolar H(+)-ATPase	435	<i>C. albicans</i>	an	path	139	5.00E-35			

β -Glucosidases
 β -Glucosidase
 β -Glucosidase
Fungal cell wall synthesis/modification

Table 4 Continued

				Best PHI-base match			Other PHI-base matches					
Contig	ax seqs	ip seqs	Putative function	PHI	Species*	Host†	Phenotype‡	HSP	E-value	# PHI	Hosts§	Phenotypes¶
Primary metabolism												
1678	1		Trehalose-6-phosphate synthase	322	M. grisea	gr	path	237	1.00E-63	2	1 pl/1an	vir/unaff
0098		2	Isocitrate lyase	305	M. grisea	gr	vir	172	3.00E-44	2	1 pl/1 an	vir
0472	1	1	S-Adenosylmethionine synthetase	877	M. grisea	gr	vir	367	E-103	1	1 an	path
3344		1	Imidazole glycerol phosphate dehydratase	121	M. grisea	gr	path	323	4.00E-90			
2470		1	β-Isopropyl-malate dehydrogenase	415	S. nodorum	gr	path	385	E-108	1	1 an	vir
2523		1	Glyoxylase I	414	S. nodorum	gr	unaff	333	1.00E-95			
1488	1		δ-Aminolevulinic acid synthase	595	S. nodorum	gr	vir	259	4.00E-70			
0837	5	1	Methionine synthase	442	F. graminearum	gr	vir	851	0			
2730		1	Homoserine O-acetyltransferase	355	F. graminearum	gr	vir	281	4.00E-77			
0049	1	1	Cystathionine β-lyases	443	F. graminearum	gr	vir	567	E-163			
0942		2	Fatty acid synthase β subunit	97	C. carbonum	gr	vir	253	E-121			
1368	1		Hexokinase	869	B. cinerea	pl	vir	486	E-139			
0857		2	Argininosuccinate lyase	200	F. oxysporum	pl	vir	311	2.00E-91			
4270		1	Alcohol oxidase	199	C. fulvum	pl	vir	153	3.00E-38			
3094		1	3-Ketoacyl-CoA thiolase	598	L. maculans	pl	vir	97	2.00E-21			
0853	1	1	Phosphoenolpyruvate carboxykinase	424	C. neoformans	an	vir	142	6.00E-35			
0432	2	2	Sulphate adenylyltransferase	265	C. neoformans	an	path	431	E-122			
2916		1	Acetolactate synthase	358	C. neoformans	an	path	192	5.00E-50			
4363		1	Methylcitrate synthase	447	A. fumigatus	an	vir	387	E-109			
3678		1	Fatty acid synthase alpha subunit	96	C. albicans	an	path	162	2.00E-48			
0666		3	Orotidine-5'-phosphate decarboxylase	506	S. cerevisiae	an	vir	92	2.00E-20			
Structural molecules												
0251	5	7	Woronin body major protein	356	M. grisea	gr	vir	260	4.00E-71			
0266		5	Cell wall protein	175	C. albicans	an	vir	100	3.00E-22			
3575		1	Septin	281	C. albicans	an	vir	263	7.00E-73			
0366		2	Septin	282	C. albicans	an	vir	141	6.00E-35			
Hydrophobins and related												
1323		5	Hydrophobin	291††	C. purpurea	gr	unaff	575	E-165	1	1 gr	vir
2720		1	snodprot1 homologue	860	M. grisea	gr	vir	430	E-122			
0273		6	Cerato-platanin	695	L. maculans	pl	unaff	127	3.00E-31			
Transcription factors												
1200		6	Transcription factor	776(802	M. grisea	gr	vir	131	9.00E-32			
1513	1		bZip transcription factor	344††	C. purpurea	gr	vir	365	E-102			
1606	1		bZIP transcription factor	444	F. graminearum	gr	vir	120	2.00E-30			
0516	1	1	Cross-pathway control protein	467	C. parasitica	pl	vir	150	9.00E-38	1	1 an	vir
4589		1	Ste12-like transcription factor	294	C. lagenarium	pl	path	113	4.00E-26	2	1 gr/1 pl	path
2296		1	Regulator of filamentous growth/virulence	475	C. albicans	an	path	90	3.00E-19			

*A. brassicicola, Altermaria brassicicola; A. fumigatus, Aspergillus fumigatus; B. cinerea, Botrytis cinerea; C. carbonum, Cochliobolus carbonum; C. albicans, Candida albicans; C. nicotianae, Cercospora nicotianae; C. fulvum, Cladosporium fulvum; C. purpurea, Claviceps purpurea; C. posadasii, Coccidioides posadasii; C. heterostrophus, Cochliobolus heterostrophus; C. gloeosporioides, Colletotrichum gloeosporioides; C. graminicola, Colletotrichum graminicola; C. lagenarium, Colletotrichum lagenarium; C. lindemuthianum, Colletotrichum lindemuthianum; C. parasitica, Cryphonectria parasitica; C. neoformans, Cryptococcus neoformans; F. graminearum, Fusarium graminearum; F. oxysporum, Fusarium oxysporum; G. moniliformis, Gibberella moniliformis; L. maculans, Leptosphaeria maculans; M. grisea, Magnaporthe grisea; S. cerevisiae, Saccharomyces cerevisiae; S. nodorum, Staenopora nodorum; T. virescens, Trichoderma virescens; U. maydis, Ustilago maydis; W. dermatidis, Wangiella dermatidis.

##EST related to already cloned *C. purpurea* gene.

Table 5 Expressed sequence tag (EST) contigs with homology to proteins involved in programmed cell death. The number and origin of the EST clones assembled in each contig are indicated: ax (ip) seqs, axenic (*in planta*) sequences.

Contig	ax seqs	ip seqs	Putative function	Identification†
0622	1	1	Metacaspase	BLASTX EST contig → nr NCBI
0639	1	1	Autophagy protein Atg20	BLASTX EST contig → nr NCBI
1509	1		Vacuolar fusion protein mon-1	TBLASTN autophagy/apoptosis proteins (http://www.uniprot.org) → EST contigs
2231		1	Autophagy protein Atg15	TBLASTN Mg/Nz/Gz proteins (Table 2, Meijer <i>et al.</i> , 2007) → EST contigs†
2460		1	Autophagy protein Atg29	TBLASTN Mg/Nz/Gz proteins (Table 2, Meijer <i>et al.</i> , 2007) → EST contigs†
2489		1	Bax inhibitor family protein	BLASTX EST contig → nr NCBI
2542		1	Autophagy protein Atg9	TBLASTN Mg/Nz/Gz proteins (Table 2, Meijer <i>et al.</i> , 2007) → EST contigs†
2652		1	Coat assembly protein SEC16	TBLASTN autophagy/apoptosis proteins (http://www.uniprot.org) → EST contigs
3115		1	Autophagy protein Atg8*	BLASTX EST contig → nr NCBI
3128		1	Protein-vacuolar targeting protein Atg18	BLASTX EST contig → nr NCBI
3309		1	Autophagy protein Atg8*	BLASTX EST contig → nr NCBI
4677		1	Autophagy protein Atg13	BLASTX EST contig → nr NCBI

*Contig 3315 contains introns, Contig 3309 does not.

†Gz, *Gibberella zeae*; Mg, *Magnaporthe grisea*; Nc, *Neurospora crassa*.

Table 6 *Claviceps* expressed sequence tag (EST) contigs with significant homology to haustorial *Uromyces fabae* genes highly expressed in rust-infected leaves. The number and origin of the EST clones assembled in each contig are indicated: ax (ip) seqs, axenic (*in planta*) sequences.

Contig	ax seqs	ip seqs	Putative function	Match with <i>U. fabae</i> haustorium: Clone identifier (Accession)*	HSP	E-value	Expression in <i>U. fabae</i> †
1298	7	6	Thiamine biosynthesis protein	THI1 (O00057)	468	1.00E-133	53.9
1162	1	12	Mannitol dehydrogenase	MAD1 (O00058)	154	6.00E-39	49.5
0359	4	5	Glutamine synthetase	Uf101 (DR010279)	146	1.00E-62	24.7
0200	1	5	Manganese superoxide dismutase	Uf058 (DR010236)	67	2.00E-27	17.8
0777	9	5	Thiamine biosynthesis protein	THI2p (CAB59856)	375	1.00E-105	13.5
0457	3	1	MFS monosaccharide transporter	HXT1 (CAC41332)	155	7.00E-39	10.1
3298		1	Regulatory protein cys-3	Uf053 (DR010231)	96	7.00E-21	9.9
0527	3	5	60s ribosomal protein L7a	Uf478 (DR010656)	209	3.00E-55	9.4
1031	13	7	Peptidyl-prolyl <i>cis-trans</i> -isomerase	PIG28 (AAB39880)	219	6.00E-59	8.7
3116		1	Cytochrome P450 monooxygenase	PIG16 (AAB39881)	91	1.00E-19	7.6
0037	1	1	40S ribosomal protein S28	Uf150 (DR010328)	109	3.00E-25	7.3
0825	1	2	60S ribosomal protein L44	Uf371 (DR010549)	238	3.00E-64	6.5
0345	3		Mitochondrial heat shock protein 60	Uf130 (DR010308)	138	1.00E-57	5.8
0724	8	3	S-phase specific ribosomal protein	Uf406 (DR010584)	188	3.00E-49	5

*TBLASTX hit for nucleotide sequences, TBLASTN hit for protein sequences. (Nucleotide sequences start with DR0 . . . All others are protein sequences.)

†Ratio of expression in infected leaves vs. expression in germinated uredospores.

role in nitrogen metabolism, and a sulphur regulatory protein cys-3 a role in sulphur metabolism. Two contigs were found which code for proteins involved in protein folding. Contig 1031 may encode a peptidyl-prolyl *cis-trans*-isomerase known to be a folding catalyst and Contig 0345 a mitochondrial HSP60, which participates in the folding of mitochondrial proteins. It would be interesting to determine the actual expression levels of these *Claviceps* ESTs. It might be possible to use them (via fusion to fluorescent proteins) to investigate whether intracellular ergot hyphae have haustorium function.

Comparison of *Claviceps* with other flower pathogens

Flower pathogens benefit from easy entry into the host through thin-walled petals and stigmas, readily available nutrients (for

example, nectar) and efficient spore dispersal via co-opted avenues that evolved for the distribution of pollen and seeds. However, the phenology of flower diseases is critical, as anthesis occurs in a narrow window of time, is short lived and is poised for rapid responses to environmental stresses. Three groups of flower pathogen have been described: (i) unspecialized pathogens, which generally do not produce survival structures on flowers; (ii) specialized pathogens, which infect ovaries through the stigma-style pathway and often have an obligate sexual cycle on the flower, and replace the host's inflorescence with resting structures; and (iii) specialized pathogens, which gain entry via the apical meristem and often grow for a long time endophytically in the host (Ngugi and Scherm, 2006).

Claviceps purpurea belongs to group (ii). The necrotroph flower pathogen *F. graminearum* belongs to group (i) and is, like

Claviceps, a member of the order Hypocerales. Even their hosts, barley and rye, are both members of the Triticaceae. Another member of group (ii) is the basidiomycete *U. maydis*, which has a biology and lifestyle similar to that of *Claviceps* (O'Connell and Panstruga, 2006; Ngugi and Scherm, 2006): for example, both are biotrophs and lack haustoria.

Fusarium graminearum and *U. maydis* are well-characterized organisms. Their genomes have been sequenced and annotated with respect to potential pathogenicity genes. A collection of enriched PHI-base sequences from both organisms was compared with the *Claviceps* ESTs and with each other. The *F. graminearum* collection also included about 400 sequences that are expressed solely during barley infection and located in regions enriched in single nucleotide polymorphisms. About 30% of these sequences are predicted to be part of the secretome (Cuomo *et al.*, 2007; Güldener *et al.*, 2006). This collection also includes the sequences described in Paper *et al.* (2007) as proteins identified at high reliability in infected wheat heads and genes, and found by Stephens *et al.* (2008) to be significantly up-regulated in both crown rot and head blight of wheat. The *U. maydis* collection also includes sequences published in the genome paper (Kämper *et al.*, 2006), as well as those predicted to be secreted proteins with or without enzymatic function (Mueller *et al.*, 2008)—assuming that secreted proteins are good candidates for being involved in the disease process. Many of the *Ustilago* sequences appear to be specific to *Ustilago*.

Results of the comparisons are found in Table S2 (see Supporting Information). The interpretation of the data is constrained as each dataset was collected under different conditions. That aside, the following observations were made.

1. No match to a *U. maydis*-specific sequence was found, which is not unexpected as most pathogens probably have both specific and general virulence factors.
2. Most matches were found between *C. purpurea*/*F. graminearum*, perhaps because of their close taxonomic relationship. Yet, despite belonging to two different taxonomic groups, more matches were noted between *U. maydis* and *C. purpurea* than between *U. maydis* and *F. graminearum*. This may reflect the similar lifestyles of corn smut pathogen and the ergot pathogen.
3. In all three fungi, signalling components and cell wall-degrading enzymes were commonly found, which is reasonable for plant pathogens in general. Yet, in each set of pairwise comparisons, different genes/proteins were prevalent: between *U. maydis* and *F. graminearum*, additional cell wall-degrading enzymes were found; between *F. graminearum* and *C. purpurea*, transporters for specific molecules, oxidoreductases/dehydrogenases and proteinases/peptidases were observed; and between *U. maydis* and *C. purpurea*, molecules involved in cell cycle control, transport

(different from the *F. graminearum*/*C. purpurea* matches) and protein assembly/folding were apparent. The latter observation may reflect the fine-tuning necessary in a biotrophic host–pathogen interaction.

Intriguingly, within the corn smut/ergot matches, a pathogen-related protein with similarity to the PR1 class of plant proteins was found. PR proteins are produced by plants to ward off attacking fungi. In particular, proteins similar to classes PR1 and PR5 have also been found recently in animals and fungi, but they have not been well characterized as yet. Mondego *et al.* (2008) have suggested that such proteins may limit microbial competition for the hemibiotrophic cacao pathogen *Moniliophthora perniciosa* during progression of witches' broom disease. The production of a PR1 protein by a pathogen may also be part of a strategy to overcome plant defences. Perhaps the plant is misled into prematurely sensing that the level of PR1 protein is sufficient for its protection and reacts too late to avoid a successful attack. As in many situations, the timing of the host's and pathogen's actions may be decisive for the outcome of the infection.

EXPERIMENTAL PROCEDURES

Fungal and plant material, and culture conditions for cDNA libraries

For the axenic cDNA library, *C. purpurea* strain 20.1 was grown for 3 days at 28 °C and 180 r.p.m. in Mantle medium (Mantle and Nisbet, 1976). Mycelium was harvested by centrifugation and lyophilized before RNA preparation. For *in planta* cDNA libraries, flowers of *S. cereale* cultivar Lo37-PxLo55-N (cytoplasmic male sterile) (KWS Lochow GmbH, Bergen, Germany) were inoculated with 5 µL of a suspension containing 2×10^6 conidia/mL. All ovaries (except for bases and tops) from inoculated spikes were harvested, submersed in liquid nitrogen and lyophilized before RNA preparation. For each day, from day 1 up to day 20 post-inoculation (1–20 dpi), ovaries from three rye spikes were collected.

cDNA library constructions

Extraction of total RNA from both infected rye ovaries and fungal mycelium was performed with the 'RNAagents® Total RNA Isolation System' from Promega, Mannheim, Germany, according to the instructions of the supplier. Total RNA from ovaries at 1–5 dpi was pooled for an early and from ovaries at 6–20 dpi for a late *in planta* cDNA library. mRNA was isolated from all three total RNAs with an 'Oligotex™ mRNA Mini Kit' from Qiagen, Hilden, Germany, according to the manufacturer's instructions. Three primary cDNA libraries (axenic, early and late *in planta*) were

produced with a 'pBluescript II cDNA Library Construction Kit' from Stratagene, La Jolla, CA, USA, following the instruction manual.

Clone isolation

About 100 clones of each cDNA library were plated and their pBluescript II inserts were sequenced from the 5' cDNA end as a quality check. Two of the three libraries passed the test. The early *in planta* library failed the test, because about 25% of the cDNA clones contained only the vector and another 25% had inserts that resisted sequencing (possibly as a result of a loss of the sequencing primer binding site). From the remaining ax and ip cDNA libraries, clones were plated; each was manually transferred into freezing medium (3 : 1 ratio of 2 × Luria–Bertani medium and glycerol), grown overnight at 37 °C and manually aliquoted into three parallel 96-well microtitre plates. The ax library was stored in 26 96-well microtitre plates (clones ax_001_A01 to ax_026_H12) and the late ip library in 78 96-well microtitre plates (clones ip_027_A01 to ip_104_H12). Two plates of one set were stored at –70 °C. The clones of one plate were sequenced.

DNA sequencing

The 5' ends of each cDNA clone were sequenced via Sanger sequencing using an Applied Biosystems 3730 capillary DNA sequencer by GreenTec GmbH (now Phytowelt GreenTechnologies GmbH, Nettetal/Cologne, Germany). Sequences were delivered in ab1 and fasta format.

Processing of sequences

Raw DNA sequences were trimmed with LUCY2 (Li and Chou, 2004). The output was screened for sequences containing putative internal linker sequences, indicating that two different cDNAs may have been ligated together. All potential coupled cDNAs were manually separated into two parts for the time of assembly and clustering (and assigning plant or fungal origin). These sequences were labelled afterwards with a 'j', for example ax_007jD12. Resulting sequences shorter than 100 bp were deleted from the project. About 1500 sequences rejected by LUCY2 were manually screened. All those longer than 100 bp and not consisting of short, monomeric stretches of Gs, As, Ts and Cs were accepted. All 343 included files were vector clipped, quality trimmed and indicated with an 'm' in the label, for example ax_006mF03.

The sequences were clustered with TGICL (Pertea *et al.*, 2003). To find appropriate parameters, a test set of 29 published *C. purpurea* genes was clustered under default conditions (–p 94 –l 30 –v 30) and under conditions (–p 88 –l 21 –v 900) which

resulted in clusters of very closely related genes and also clusters formed by spliced and unspliced versions of each gene. As ESTs are usually shorter than their corresponding genes, the final parameters chosen were –p 90 –l 30 –v 30.

SeqMan Pro (Lasergene 7.1 by DNASTAR, Madison, WI, USA) was used to assemble the EST sequences into contigs. Under default conditions, the assembly contained contigs with sequences in which large gaps (40–130 bp) had been introduced into sequences by the program to fit the sequence into the contig. Therefore, the assembly conditions were adjusted to Match Size = 50, Match Spacing = 50, Maximum Mismatch End Bases = 0, Use Repeat Handling = Yes, Fixed = 4 and Match Repeat Percentage = 150. To find at least a good part of the falsely joined sequences, contigs were checked for suspiciously small overlaps or long overhangs in the strategy view. Such contigs were subjected to a BLASTX search against nr of NCBI (http://www.ncbi.nlm.nih.gov/blast/Blast.cgi?PAGE=Translations&PROGRAM=blastx&BLAST_PROGRAMS=blastx&PAGE_TYPE=BlastSearch&SHOW_DEFAULTS=on). If the search revealed two regions with fundamentally different matches, the sequences of such a contig were split into two contigs. Table S1 shows which EST clones belong to which contig and cluster.

Sequence analysis

GO annotation (comparing ax and ip libraries)

GO terms of two aspects ('Molecular function' and 'Biological process') were assigned to all ax and ip ESTs via GOPET (–CONFIDENCE = 80) (Vinayagam *et al.*, 2006). For each aspect, the terms found were counted separately for ax ESTs and ip ESTs with the GO terms classifications counter (Hu *et al.*, 2007; <http://www.animalgenome.org/bioinfo/tools/countgo/>). The GO term counts for the ax and ip libraries were checked by the method of Audic and Claverie (1997) for significantly different counts (significance level = 0.01).

Assignment of plant or fungal origin for ip ESTs

All ax ESTs were organized into a fungal databank. A BLASTN search of all ip ESTs was performed against the ax databank. This was accomplished using SeqTools (Dr Søren Rasmussen, H. Lundbeck A/S, Copenhagen, Denmark, <http://www.seqtools.dk/>), which was also employed to build additional databanks and to perform BLAST searches with these databanks (against one of the self-assembled or NCBI databanks). All hits with an E-value ≤ 9E-100 were considered to be of putative fungal origin. Similarly, databanks were assembled from genomes and EST libraries of organisms closely related to *C. purpurea* or *S. cereale*, as neither genome has yet been sequenced:

- *Epichloë festucae* genome (no databank built, but use of BLASTN offered in *E. festucae* Genome project webpage; last

modified October 2000, <http://www.endophyte.uky.edu/index.php>;

- genomic Clavicipitaceae and Triticeae sequences ≥ 100 bp from NCBI (<http://www.ncbi.nlm.nih.gov/>, downloaded July 2007);
- *F. graminearum* genome (http://www.broad.mit.edu/annotation/genome/fusarium_graminearum/Home.html, March 2007);
- *F. verticillioides*, *M. grisea* cDNA sequences (<http://compbio.dfci.harvard.edu/tgi/fungi.html>; release 7.0 of *Fusarium* and release 5.0 of *Magnaporthe* libraries #0QV, #A36–39, #A3A, #CN6, #F3J, #F3P);
- *S. cereale*, *T. aestivum*, *H. vulgare* unique transcript sequences (<http://www.plantgdb.org/download/download.php?dir=/Sequence/ESTcontig>, version 157a for rye and barley, version 153a for wheat);
- *O. sativa* cDNA sequences (http://compbio.dfci.harvard.edu/tgi/cgi-bin/tgi/cat_download.pl?db=ricest, #GCH—a normalized whole-life-cycle cDNA library of rice, release 17.0).

All BLASTN hits against a fungal/plant databank with an E-value $\leq 9E-100$ were considered to be of putative fungal/plant origin. If there were hits of one EST with both fungal and plant databanks, the lowest E-value/highest HSP score were used as the decisive criteria. With all ESTs that could not be sorted, a BLASTX search against the nr databank of NCBI and a BLASTN search against non-human, non-mouse ESTs (EST_other) of NCBI were performed. The chosen criteria discriminate against smaller sequences (≤ 250 bp), whose BLASTN searches will not result in such low E-values. Yet, empirically, we found that, for a given average sequence (GC content between 50 and 60%), there is a linear relationship between sequence length and HSP score, if it is BLASTed to itself (query length $\times 1.98 \approx$ HSP score). There is also a linear relationship if the query sequence has 90% similarity with the sequence it is BLASTed against (query length $\times 1.17 \approx$ HSP score_{90%}). Therefore, small EST sequences were considered to be of putative fungal/plant origin when they achieved an HSP score higher than a calculated HSP score_{90%}. Finally, all putative plant ip clones were traced back to their contig. There was no occurrence of 'mixed' contigs. Table S1 shows the (putative) origin of each EST clone.

For comparison with the putative plant EST contigs, two EST banks were assembled from the TGI plant Gene Index Project (<http://compbio.dfci.harvard.edu/tgi/plant.html>): (i) a flower EST bank (59 273 sequences) comprised libraries 5474, #8GO, 5341, 7313 from *H. vulgare*, #DOH, #G8E, #IJA, #IJH, #ILF, #ILG, #IMK, #IOP, OS31, #IKR, #IKS, #IKT, #ILI, #ILJ, #ILS from *O. sativa*, #E14 from *Sorghum bicolor*, #5CC, #DF2 from *S. cereale*, 5459, #DEP, 5603 from *T. aestivum*, and ZM19 from *Zea mays*; (ii) an infected flower EST bank (11 939 sequences) comprised library #9JH from *H. vulgare* and #9HS, #9OO, #AS6, #AS7, #AS8, #DE6, 5584

from *T. aestivum*. Significant hits were those with an E-value $\leq 1E-38$.

Comparison with protein sequences in PHI-base (version 3.1, <http://www.phi-base.org/>)

Protein sequences collected in PHI-base were downloaded, arranged in a databank (together with four *C. purpurea* genes not yet part of PHI-base) and run in a TBLASTN search against the EST contigs of *C. purpurea* ($e < -15$). A hit was rejected if the contig had been classified previously to be of putative plant origin. If there was more than one PHI-base sequence originating from the same organism with a significant hit to the same contig, only the hit with the lowest E-value was listed in Table 4. In addition, each query and hit pair of this search was subjected to a BLASTP/BLASTX search to the nr databank of NCBI. If the result of these searches did not correspond, the hit was rejected after closer manual inspection of the hits.

Comparison with *U. fabae* genes highly expressed in haustoria

The comparison was similar to the comparisons with PHI-base sequences and secretome sequences. Selected for comparison were *U. fabae* genes published in Jakupovič *et al.* (2006), showing five-fold or greater up-regulation in rust-infected leaves vs. germinated uredospores. The accession numbers given in the publication refer to both nucleotide and protein sequences. The nucleotide sequences were compared with the *C. purpurea* EST contigs via TBLASTX, the protein sequences via TBLASTN.

Comparison with enriched PHI-base protein sequences of *U. maydis* or *F. graminearum* only

Additional *U. maydis* protein sequences were downloaded and subselected from the MIPS *Ustilago* ftp site (<ftp://ftpmips.gsf.de/ustilago/>, Umaydis_valid_orf_prot_140408.fas) according to the genes listed in supplementary Tables S1, S2 and S3 of the recently published secretome paper of *Ustilago* (Mueller *et al.*, 2008).

Additional *F. graminearum* sequences correspond to those listed as 'in planta only' in supplementary Table D in the publication of Güldener *et al.* (2006). The nucleotide sequences were downloaded from the PLEXdb site http://www.plexdb.org/modules/tools/get_seq.php (Choose microarray 'Fusarium'). To mass translate these nucleotide sequences correctly into protein sequences, they were BLASTed against Fgraminearum_valid_orf_prot_2007-05-10 from the MIPS *Fusarium* ftp site (<ftp://ftpmips.gsf.de/fusarium/>). Furthermore, all sequences listed in Table 2 (Paper *et al.*, 2007) and in Table S2 (Stephens *et al.*, 2008) were included in this *F. graminearum* sequence collection.

Searches were performed similar to that with PHI-base sequences. If there were two or more sequences with a significant hit to the same contig, only the best hit (that with the

lowest E-value) was listed in Table S2. In addition, hits were selected only if the result of a BLAST of sequence 1 from organism 1 against sequence 2 of organism 2 gave a result similar to the BLAST of sequence 2 from organism 2 against sequence 1 of organism 1 (reciprocal search).

ACKNOWLEDGEMENT

We thank Olen C. Yoder for critical reading of the manuscript, Erich Bornberg-Bauer for initial help in bioinformatic analysis and the Deutsche Forschungsgemeinschaft (SFB 629: 'Molecular cell dynamics') for funding.

REFERENCES

- Audic, S. and Claverie, J.M. (1997) The significance of digital gene expression profiles. *Genome Res.* **7**, 986–995.
- Baldwin, T.K., Winnenburg, R., Urban, M., Rawlings, C., Koehler, J. and Hammond-Kosack, K.E. (2006) The pathogen–host interactions database (PHI-base) provides insights into generic and novel themes of pathogenicity. *Mol. Plant–Microbe Interact.* **19**, 1451–1462.
- Barhoom, S. and Sharon, A. (2007) Bcl-2 proteins link programmed cell death with growth and morphogenetic adaptations in the fungal plant pathogen *Colletotrichum gloeosporioides*. *Fungal Genet. Biol.* **44**, 32–43.
- Bohman, S., Wang, M. and Dixelius, C. (2002) *Arabidopsis thaliana*-derived resistance against *Leptosphaeria maculans* in a *Brassica napus* genomic background. *Theor. Appl. Genet.* **105**, 498–504.
- Coram, T.E., Wang, M.N. and Chen, X.M. (2008) Transcriptome analysis of the wheat–*Puccinia striiformis* f. sp. *tritici* interaction. *Mol. Plant Pathol.* **9**, 157–169.
- Corbett, K., Dickerson, A.G. and Mantle, P.G. (1974) Metabolic studies on *Claviceps purpurea* during parasitic development on rye. *J. Gen. Microbiol.* **84**, 39–58.
- Cuomo, C.A., Güldener, U., Xu, J.-R., Trail, F., Turgeon, B.G., Pietro, A.D., Walton, J.D., Ma, L.-J., Baker, S.E., Rep, M., Adam, G., Antoniw, J., Baldwin, T., Calvo, S., Chang, Y.-L., DeCaprio, D., Gale, L.R., Gnerre, S., Goswami, R.S., Hammond-Kosack, K., Harris, L.J., Hilburn, K., Kennell, J.C., Kroken, S., Magnuson, J.K., Mannhaupt, G., Mauceli, E., Mewes, H.-W., Mitterbauer, R., Muehlbauer, G., Münsterkötter, M., Nelson, D., O'Donnell, K., Ouellet, T., Qi, W., Quesneville, H., Roncero, M.I.G., Seong, K.-S., Igor, V., Tetko, I.V., Urban, M., Waalwijk, C., Ward, T.J., Yao, J., Birren, B.W. and Kistler, H.C. (2007) The *Fusarium graminearum* genome reveals a link between localized polymorphism and pathogen specialization. *Science*, **317**, 1400–1402.
- Deepak, S., Shailasree, S., Kini, R.K., Hause, B., Shetty, S.H. and Mithofer, A. (2007) Role of hydroxyproline-rich glycoproteins in resistance of pearl millet against downy mildew pathogen *Sclerospora graminicola*. *Planta*, **226**, 323–333.
- Dixon, D.P., Laphorn, A. and Edwards, R. (2002) Plant glutathione transferases. *Genome Biol.* **3**, 3004.1–3004.10.
- Doehlemann, G., Wahl, R., Horst, R.J., Voll, L.M., Usadel, B., Poree, F., Stitt, M., Pons-Kuhnemann, J., Sonnenwald, U., Kahmann, R. and Kämper, J. (2008) Reprogramming a maize plant: transcriptional and metabolic changes induced by the fungal biotroph *Ustilago maydis*. *Plant J.* **56**, 181–195.
- Dreher, K. and Callis, J. (2007) Ubiquitin, hormones and biotic stress in plants. *Ann. Bot.* **99**, 787–822.
- Felitti, S., Shields, K., Ramsperger, M., Tian, T., Sawbridge, T., Webster, T., Logan, E., Erwin, T., Forster, J., Edwards, D. and Spangenberg, G. (2006) Transcriptome analysis of *Neotyphodium* and *Epichloe* grass endophytes. *Fungal Genet. Biol.* **43**, 465–475.
- Gachon, C.M.M., Langlois-Meurinne, M. and Saindrenan, P. (2005) Plant secondary metabolism glycosyltransferases: the emerging functional analysis. *Trends Plant Sci.* **10**, 542–549.
- Goodwin, P.H., Oliver, R.P. and Hsiang, T. (2004) Comparative analysis of expressed sequence tags from *Malva pusilla*, *Sorghum bicolor*, and *Medicago truncatula* infected with *Colletotrichum* species. *Plant Sci.* **167**, 481–489.
- Güldener, U., Seong, K.-Y., Boddu, J., Cho, S., Trail, F., Xu, J.-R., Gerhard Adam, G., Mewes, H.-W., Muehlbauer, G.J. and Kistler, H.C. (2006) Development of a *Fusarium graminearum* Affymetrix Gene-Chip for profiling fungal gene expression *in vitro* and *in planta*. *Fungal Genet. Biol.* **43**, 316–325.
- Han, K.-H., Kim, J.H., Moon, H., Kim, S., Lee, S.-S., Han, D.-M., Jahnga, K.-Y. and Chaea, K.-S. (2008) The *Aspergillus nidulans* *esdC* (early sexual development) gene is necessary for sexual development and is controlled by *veA* and a heterotrimeric G protein. *Fungal Genet. Biol.* **45**, 310–318.
- Hu, Z.-L., Bao, J. and Reecy, J. (2007) A gene ontology (GO) terms classifications counter. Plant & Animal Genome XV Conference, San Diego, CA, January 13–17, 2007.
- Jakupović, M., Heintz, M., Reichmann, P., Mendgen, K. and Hahn, M. (2006) Microarray analysis of expressed sequence tags from haustoria of the rust fungus *Uromyces fabae*. *Fungal Genet. Biol.* **43**, 8–19.
- Kachroo, A., Shanklin, J., Whittle, E., Lapchyk, L., Hildebrand, D. and Kachroo, P. (2007) The *Arabidopsis* stearyl-acyl carrier protein-desaturase family and the contribution of leaf isoforms to oleic acid synthesis. *Plant Mol. Biol.* **63**, 257–271.
- Kachroo, P., Shanklin, J., Shah, J., Whittle, E.J. and Klessig, D.F. (2001) A fatty acid desaturase modulates the activation of defense signaling pathways in plants. *Proc. Natl. Acad. Sci. USA*, **98**, 9448–9453.
- Kämper, J., Kahmann, R., Bölker, M., Ma, L.J., Brefort, T., Saville, B.J., Kronstad, J.W., Gold, S.E., Müller, O., Perlin, M.H., Wösten, H.A.B., de Vries, R., Ruiz-Herrera, J., Reynaga-Peña, C.G., Snetselaar, K., McCann, M., Pérez-Martín, J., Feldbrügge, M., Basse, C.W., Steinberg, G., Ibeas, J.I., Holloman, W., Guzman, P., Farman, M., Stajich, J.E., Sentandreu, R., González-Prieto, J.M., Kennell, J.C., Molina, L., Schirawski, J., Mendoza-Mendoza, A., Greilinger, D., Münch, K., Rössel, N., Scherer, M., Vranes, M., Ladendorf, O., Vincon, V., Fuchs, U., Sandrock, B., Shaowu S., Ho, E.C.H., Cahill, M.J., Boyce, K.J., Klose, J., Klosterman, S.J., Deelstra, H.J., Ortiz-Castellanos, L., Li, W., Sanchez-Alonso, P., Schreier, P.H., Häuser-Hahn, I., Vaupel, M., Koopmann, E., Friedrich, G., Voss, H., Schlüter, T., Margolis, J., Platt, D., Swimmer, C., Gnirke, A., Chen, F., Vysotskaia, V., Mannhaupt, G., Güldener, U., Münsterkötter, M., Haase, D., Oesterheld, M., Mewes, H.-W., Mauceli, E.W., DeCaprio, D., M. Wade, C.M., Butler, J., Young, S., Jaffe, D.B., Calvo, S., Nusbaum, C., James Galagan, J. and Birren, B.W. (2006) Insights from the genome of the biotrophic fungal plant pathogen *Ustilago maydis*. *Nature*, **444**, 97–101.
- Krska, R. and Crews, C. (2008) Significance, chemistry and determination of ergot alkaloids: a review. *Food Addit. Contam.* **25**, 722–731.

- Laitinen, R.A.E., Immanen, J., Auvinen, P., Rudd, S., Alatalo, E., Paulin, L., Ainasonja, M., Kotilainen, M., Koskela, S., Teeri, T.H. and Elomaa, P. (2005) Analysis of the floral transcriptome uncovers new regulators of organ determination and gene families related to flower organ differentiation in *Gerbera hybrida* (Asteraceae). *Genome Res.* **15**, 475–486.
- Lane, B.G. (2002) Oxalate, germins, and higher-plant pathogens. *IUBMB Life*, **53**, 67–75.
- Li, S. and Chou, H.H. (2004) Lucy 2: an interactive DNA sequence quality trimming and vector removal tool. *Bioinformatics*, **20**, 2865–2866.
- Liljeroth, E., Marttila, S. and von Bothmer, R. (2005) Immunolocalization of defence-related proteins in the floral organs of barley (*Hordeum vulgare* L.). *J. Phytopathol.* **153**, 702–709.
- Liu, X.H., Lu, J.P., Zhang, L., Dong, B., Min, H. and Lin, F.C. (2007) Involvement of a *Magnaporthe grisea* serine/threonine kinase gene, MgATG1, in appressorium turgor and pathogenesis. *Eukaryotic Cell*, **6**, 997–1005.
- van Loon, L.C., Rep, M. and Pieterse, C.M.J. (2006) Significance of inducible defense-related proteins in infected plants. *Annu. Rev. Phytopathol.* **44**, 135–162.
- Mantle, P.G. and Nisbet, L.J. (1976) Differentiation of *Claviceps purpurea* in axenic culture. *J. Gen. Microbiol.* **93**, 321–334.
- Matarasso, N., Schuster, S. and Avni, A. (2005) A novel plant cysteine protease has a dual function as a regulator of 1-aminocyclopropane-1-carboxylic acid synthase gene expression. *Plant Cell*, **17**, 1205–1216.
- Meijer, W.H., van der Klei, I.J., Veenhuis, M. and Kiel, J.A.K.W. (2007) ATG genes involved in non-selective autophagy are conserved from yeast to man, but the selective Cvt and pexophagy pathways also require organism-specific genes. *Autophagy*, **3**, 106–116.
- Mittler, R., Vanderauwera, S., Gollery, M. and Van Breusegem, F. (2004) Reactive oxygen gene network of plants. *Trends Plant Sci.* **9**, 490–498.
- Mondego, J.M.C., Carazzolle, M.F., Costa, G.G.L., Formighieri, E.F., Parizzi, L.P., Rincones, J., Cotomacci, C., Carraro, D.M., Cunha, A.F., Carrer, H., Vidal, R.O., Estrela, R.C., Garcia, O., Thomazella, D.P.T., de Oliveira, B.V., Pires, A.B.L., Rio, M.C.S., Araújo, M.R.R., de Moraes, M.H., Castro, L.A.B., Gramacho, K.P., Gonçalves, M.S., Moura Neto, J.P., Góes Neto, A., Barbosa, L.V., Guiltinan, M.J., Bailey, B.A., Meinhardt, L.W., Julio, C.M., Cascardo, J.C.M. and Pereira, G.A.G. (2008) A genome survey of *Moniliophthora perniciosa* gives new insights into witches' broom disease of cacao. *BMC Genomics*, **9**, 548–573.
- Mueller, O., Kahmann, R., Aguilar, G., Trejo-Aguilar, B., Wu, A. and de Vries, R.P. (2008) The secretome of the maize pathogen *Ustilago maydis*. *Fungal Genet. Biol.* **45**, S63–S70.
- Nagaraj, S.H., Gasser, R.B. and Ranganathan, S. (2007) A hitchhiker's guide to expressed sequence tag (EST) analysis. *Brief. Bioinformatics*, **8**, 6–21.
- Nathues, E., Joshi, S., Tenberge, K.B., von den Driesch, M., Oeser, B., Bäumer, N., Mihlan, M. and Tudzynski, P. (2004) CPTF1, a CREB-like transcription factor, is involved in the oxidative stress response in the phytopathogen *Claviceps purpurea* and modulates ROS level in its host *Secale cereale*. *Mol. Plant-Microbe Interact.* **17**, 383–393.
- Ngugi, H.K. and Scherm, H. (2006) Biology of flower-infecting fungi. *Annu. Rev. Phytopathol.* **44**, 261–282.
- O'Connell, R.J. and Panstruga, R. (2006) Tête à tête inside a plant cell: establishing compatibility between plants and biotrophic fungi and oomycetes. *New Phytol.* **171**, 699–718.
- Paper, J.M., Scott-Craig, J.S., Adhikari, N.D., Cuomo, C.A. and Walton, J.D. (2007) Comparative proteomics of extracellular proteins *in vitro* and *in vivo* from the pathogenic fungus *Fusarium graminearum*. *Proteomics*, **7**, 3171–3183.
- Pertea, G., Huang, X., Liang, F., Antonescu, V., Sultana, R., Karamycheva, S., Lee, Y., White, J., Cheung, F., Parvizi, B., Tsai, J. and Quackenbush, J. (2003) TIGR Gene Indices clustering tools (TGICL): a software system for fast clustering of large EST datasets. *Bioinformatics*, **19**, 651–652.
- Peter, M., Courty, P.-E., Kohler, A., Delaruelle, C., Martin, D., Tagu, D., Frey-Klett, P., Duplessis, S., Chalot, M., Podila, G. and Martin F. (2003) Analysis of expressed sequence tags from the ectomycorrhizal basidiomycetes *Laccaria bicolor* and *Pisolithus microcarpus*. *New Phytol.* **159**, 117–129.
- Raffaele, S., Mongrand, S., Gamas, P., Niebel, A. and Ott, T. (2007) Genome-wide annotation of remorins, a plant-specific protein family: evolutionary and functional perspectives. *Plant Physiol.*, **145**, 593–600.
- Ramsdale, M. (2008) Programmed cell death in pathogenic fungi. *Biochim. Biophys. Acta*, **1783**, 1369–1380.
- Rapilly, F. (2001) Champignons des plantes: les premiers agents pathogènes reconnus dans l'histoire des sciences. Plant pathogenic fungi: the first identified pathogens in the history of science. *C. R. Acad. Sci. Gen.* **324**, 893–898.
- Ray, S., Anderson, J.M., Urmeev, F.I. and Goodwin, S.B. (2003) Rapid induction of a protein disulfide isomerase and defense-related genes in wheat in response to the hemibiotrophic fungal pathogen *Mycosphaerella graminicola*. *Plant Mol. Biol.* **53**, 741–754.
- Rolke, Y. and Tudzynski, P. (2008) The small GTPase Rac and the p21-activated kinase Cla4 in *Claviceps purpurea*: interaction and impact on polarity, development and pathogenicity. *Mol. Microbiol.* **68**, 405–423.
- Rostoks, N., Schmieder, D., Kudrna, D. and Kleinhofs, A. (2003) Barley putative hypersensitive induced reaction genes: genetic mapping, sequence analyses and differential expression in disease lesion mimic mutants. *Theor. Appl. Genet.* **107**, 1094–1101.
- Schaller, A. (2004) A cut above the rest: the regulatory function of plant proteases. *Planta*, **220**, 183–197.
- Schardl, C.L., Panaccione, D.G. and Tudzynski, P. (2006) Ergot alkaloids—biology and molecular biology. *Alkaloids Chem. Biol.* **63**, 45–86.
- Scheffer, J., Chen, C., Heidrich, P., Dickmann, M.B. and Tudzynski, P. (2005) A CDC42 homologue in *Claviceps purpurea* is involved in vegetative differentiation and is essential for pathogenicity. *Eukaryotic Cell*, **4**, 1228–1238.
- Sexton, A.C. and Howlett, B.J. (2006) Parallels in fungal pathogenesis on plant and animal hosts. *Eukaryotic Cell*, **5**, 1941–1949.
- Shim, K.S., Cho, S.K., Jeung, J.U., Jung, K.W., You, M.K., Ok, S.H., Chung, Y.S., Kang, K.H., Hwang, H.G., Choi, H.C., Moon, H.P. and Shin, J.S. (2004) Identification of fungal (*Magnaporthe grisea*) stress-induced genes in wild rice (*Oryza minuta*). *Plant Cell Rep.* **22**, 599–607.
- Soanes, D.M. and Talbot, N.J. (2006) Comparative genomic analysis of phytopathogenic fungi using expressed sequence tag (EST) collections. *Mol. Plant Pathol.* **7**, 61–70.
- Solomon, P.S., Waters, O.D.C. and Oliver, R.P. (2007) Decoding the mannitol enigma in filamentous fungi. *Trends Microbiol.* **15**, 257–262.
- Stephens, A.E., Gardiner, D.M., White, R.G., Munn, A.L. and Manners, J.M. (2008) Phases of infection and gene expression of *Fusarium*

- graminearum* during crown rot disease of wheat. *Mol. Plant-Microbe Interact.* **12**, 1571–1581.
- Taler, D., Galperin, M., Benjamin, I., Cohen, Y. and Kenigsbuch, D. (2004) Plant eR genes that encode photorespiratory enzymes confer resistance against disease. *Plant Cell*, **16**, 172–184.
- Tenberge, K.B., Homann, V., Oeser, B. and Tudzynski, P. (1996) Structure and expression of two polygalacturonase genes of *Claviceps purpurea* oriented in tandem and cytological evidence for pectinolytic enzyme activity during infection of rye. *Phytopathology*, **86**, 1084–1097.
- Torres, M.A., Jones, J.D.G. and Dangl, J.L. (2006) Reactive oxygen species signaling in response to pathogens. *Plant Physiol.* **141**, 373–378.
- Tudzynski, P. and Scheffer, J. (2004) *Claviceps purpurea*: molecular aspects of a unique pathogenic lifestyle. *Mol. Plant Pathol.* **5**, 377–388.
- Tung, C.W., Dwyer, K.G., Nasrallah, M.E. and Nasrallah, J.B. (2005) Genome-wide identification of genes expressed in *Arabidopsis* pistils specifically along the path of pollen tube growth. *Plant Physiol.* **138**, 977–989.
- Veneault-Fourrey, C., Barooah, M., Egan, M., Wakley, G. and Talbot, N.J. (2006) Autophagic fungal cell death is necessary for infection by the rice blast fungus. *Science*, **312**, 580–583.
- Vierling, E. (1991) The roles of heat-shock proteins in plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **42**, 579–620.
- Vinayagam, A., del Val, C., Schubert, F., Eils, R., Glatting, H., Suhai, S. and König, R. (2006) GOPET: a tool for automated predictions of Gene Ontology terms. *BMC Bioinformatics*, **7**, 161–167.
- Vizcaino, J.A., González, F.J., Suárez, M.B., Redondo, J., Heinrich, J., Delgado-Jarana, J., Hermosa, R., Gutiérrez, S., Monte, E., Llobel, A. and Rey, M. (2006) Generation, annotations and analysis of ESTs from *Trichoderma harzianum* CECT 2413. *BMC Genomics*, **7**, 193–207.
- Walters, D. (2003) Resistance to plant pathogens: possible roles for free polyamines and polyamine catabolism. *New Phytol.* **159**, 109–115.
- Wang, W.X., Vinocur, B., Shoseyov, O. and Altman, A. (2004) Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. *Trends Plant Sci.* **9**, 244–252.
- Winnenburg, R., Urban, M., Beacham, A., Baldwin, T.K., Holland, S., Lindeberg, M., Hansen, H., Rawlings, C., Hammond-Kosack, K.E. and Koehler, J. (2008) PHI-base update: additions to the pathogen host interaction database. *Nucleic Acids Res.* **36**, D572–D576.
- Yang, T.B. and Poovaiah, B.W. (2003) Calcium/calmodulin-mediated signal network in plants. *Trends Plant Sci.* **8**, 505–512.
- Zhou, G.K., Xu, Y.F., Li, J., Yang, L.Y. and Liu, J.Y. (2006) Molecular analyses of the metallothionein gene family in rice (*Oryza sativa* L.). *J. Biochem. Mol. Biol.* **39**, 595–606.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Table S1 Overview of expressed sequence tag (EST) clones. Listed are the contig and cluster association, the respective EMBL accession number and the (putative) origin. n.d., not determinable; put, putative.

Table S2 Comparison of (selected) sequences from three flower pathogens.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.