Secretome of fungus-infected aphids documents high pathogen activity and weak host response

Grell, Morten Nedergaard; Jensen, Annette Bruun; Olsen, Peter Bjarke; Eilenberg, Jørgen; Lange, Lene

Published in:
Fungal Genetics and Biology

DOI (link to publication from Publisher):
10.1016/j.fgb.2010.12.003

Publication date:
2011

Document Version
Publisher's PDF, also known as Version of record

Link to publication from Aalborg University

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

Take down policy
If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.
Secretome of fungus-infected aphids documents high pathogen activity and weak host response

Morten N. Grella a,*, Annette B. Jensen b, Peter B. Olsen c, Jørgen Eilenberg b, Lene Lange a

a Department of Biotechnology and Chemistry, Section for Sustainable Biotechnology, Aalborg University, Lautrupvang 15, DK-2750 Ballerup, Denmark
b Department of Agriculture and Ecology, University of Copenhagen, Thorvaldensvej 40, DK-1871 Frederiksberg C, Denmark
c Department of Biotechnology and Chemistry, Section for Sustainable Biotechnology, Aalborg University, Lautrupvang 15, DK-2750 Ballerup, Denmark

* Corresponding author. Fax: +45 9940 2594.
E-mail address: mng@bio.aau.dk (M.N. Grell).

1. Introduction

Most species within the fungal order Entomophthorales (subphylum Entomophthoromycotina; Hibbett et al., 2007) are obligate pathogens of insects and arachnids (Jensen et al., 2006; Roy et al., 2006). Entomophthoralean fungi exhibit a highly specialized life cycle, including a parasitic growth phase in the host (Roy et al., 2006). Some species, like Pandora neaphidis and Entomophthora planchoniana, form protoplast-like cells in the hemocoeel (Butt et al., 1981; Freimoser et al., 2001), possibly to escape detection by the host immune system (Beauvais et al., 1989). After completion of the asexual life cycle, conidial spores are dispersed from the host to infect new susceptible individuals (Roy et al., 2006). Entomophthoralean fungi can cause epidemics in host populations, both in natural habitats and under crop conditions (Hesketh et al., 2010). This has led to the suggestion of using entomophthoralean species as biocontrol agents for pest insects, such as aphids (Hountondji, 2008; Jensen et al., 2008; Nielsen et al., 2005).

Although the life cycle of many entomophthoralean species is well known, a detailed molecular understanding of infection processes and host responses is still lacking. In this study, we describe for the first time the use of a unique method, Transposon Assisted Signal Trapping (TAST) (Becker et al., 2004), on cDNA libraries from field-sampled, interacting hosts and pathogens. We used TAST to preferentially select full-length cDNAs encoding secreted proteins. Other advantages of TAST are that it can be used on organisms for which no genome sequence data are available. More importantly, genes are identified without the bias introduced by methods that rely on known gene sequences or functional assays, allowing the discovery of truly novel genes. In principle TAST is simple, with the entire process of library production, screening, and clone selection performed in a standard Escherichia coli laboratory strain. Previous results using TAST on eukaryotic cDNA libraries indicate that eukaryotic signal peptides are processed correctly in E. coli (Lange and Schnorr, 2004; Hamann and Lange, 2006). This is presumably due to the highly similar physicochemical properties of prokaryotic and eukaryotic N-terminal signal peptides (Paetzel et al., 2002).

Our hypothesis was that we could use TAST to discover new and unique proteins by studying interactions between insects and their pathogens, sampled directly from the field. These discoveries should improve our understanding of the biology of host–pathogen interactions for use in biological control methods, and possibly lead to the discovery of proteins with industrial potential. We confirm our hypothesis here, by reporting on several novel glycoside hydrolases, lipases, proteases, and other proteins secreted during interactions between grain aphids (Sitobion avenae) and the entomophthoralean fungi E. planchoniana, P. neaphidis, and Conidiobolus obscurus, found by sequencing fewer than 600 TAST clones.
2. Materials and methods

2.1. Biomass collection

An infestation of grain aphids (Fig. 1A and B) was followed during one growth season in a Danish organically grown winter wheat field to determine the appropriate time for sampling potentially infected individuals. Species composition was determined by weekly sampling and incubation of 100 live aphids until spore discharge allowed morphological examination, as described in Jensen et al. (2008). When fungal infection prevalence increased, thousands of live aphids were brought into the laboratory (Fig. 2, arrows). These aphids were inspected under a dissecting microscope in the afternoons and evenings for up to five days, and aphids showing signs of entomophthoralean infection were quickly frozen in liquid nitrogen and stored at −80 °C until RNA extraction. Individuals in three different stages of infection were pooled for library construction: (1) just before fungal hyphae penetrated the cuticle from the inside, a transition that can be assessed from the outside by swelling of the body and a color change to brown–orange–pink (Fig. 1A); (2) just after penetration; and (3) at the onset of fungal sporulation.

The composite biomass for library production consisted of infected alate and apterous adult aphids (with and without wings) as well as infected nymphs.

2.2. cDNA library construction

For extraction of total RNA, 100 mg composite biomass was homogenized under liquid nitrogen, suspended and incubated in 30 ml extraction buffer (4 M guanosine triphosphate, 0.5% w/v Na-laurylsarcosine, 25 mM Na–citrate pH 7.0, 0.1 M β-mercaptoethanol) at 65 °C for 15 min, followed by an additional incubation at room temperature for 30 min and centrifugation (17,000 × g for 2 min). The supernatant (26.5 ml) was mixed with 12 ml 5.7 M CsCl in a 40 ml Beckmann Ultra Clear tube and subjected to CsCl gradient centrifugation in a Beckmann ultracentrifuge (25,000 rpm for 24 h at room temperature). The resulting total RNA pellet was resuspended in 0.5 ml TE pH 7.5 and extracted with an equal volume of phenol:chloroform:isoamyl alcohol 25:24:1. After centrifugation at 12,000 g for 2 min, the aqueous phase was extracted again with an equal volume of chloroform:isoamyl alcohol 24:1. Finally, RNA was concentrated by ethanol precipitation and resuspension in 100 μl RNase-free water.

Poly(A)+ RNA was isolated from the total RNA preparation using the mTRAP Total Kit (Ambion) following the manufacturer’s procedure. The eluted mRNA was ethanol precipitated and resuspended in 10 μl RNase-free water at an approximate concentration of 1 μg/μl.

Double-stranded cDNA was synthesized with the SMART cDNA Library Construction Kit (Clontech). For first-strand synthesis, the Superscript II reverse transcriptase (Invitrogen) was used according to the SMART protocol. Second-strand synthesis was carried out as described in the SMART primer extension protocol, using the GeneAmp 2400/9600 compatible thermal cycler program. After purification on GFX spin columns (GE Healthcare), cDNA was prepared for cloning by SfiI digestion, followed by agarose gel size selection of cDNAs larger than approximately 750 bp and GFX–gel band purification. The plasmid cDNA library vector pMHas7l (kindly provided by K. Schnorr, Novozymes) was digested with SfiI and GFX–gel band purified. Vector and cDNA were concentrated by ethanol precipitation and resuspended in 20 and 50 μl water, respectively. One tenth of the cDNA preparation was mixed with 50 ng vector, two units T4 DNA ligase, and supplied 10× buffer (Promega) in a total volume of 10 μl and ligation performed overnight at 16 °C. Subsequently, ElectroMAX DH10B E. coli cells (Invitrogen) were transformed with the plasmid library by electroporation according to the manufacturer of the cells. Following transformation, the cells were diluted in 1 ml SOC medium, incubated at 37 °C, 225 rpm for 1 h, and selected on LB agar plates containing 50 μg/ml kanamycin at 37 °C. Pooled plasmid DNA was isolated from approximately 30,000 colonies.

2.3. Signal trapping

For transposon tagging of the cDNA library, 1 μg plasmid cDNA library was mixed with 50, 100, or 250 ng purified TnSig transposon (Becker et al., 2004) (pSigA4 kindly provided by K. Schnorr), 0.22 μg MuA transposase, and 5× MuA transposase reaction buffer (Finnzymes) in a total volume of 20 μl and incubated at 30 °C for 3 h, followed by inactivation of the transposase by incubation at 75 °C for 10 min. Subsequently, the tagged library was ethanol precipitated and resuspended in 10 μl Tris–HCl pH 8.0, and 2 μl transformed into DH10B E. coli cells (as above). For controls, aliquots were plated on LB kanamycin (5 μl; transformation control) and LB kanamycin/ampicillin (50 μl; transposition control). For isolation of signal-trapped clones, the rest of the 1-ml culture was plated on LB kanamycin/ampicillin/ampicillin and incubated at 30 °C for 2 days. Antibiotic concentrations on all plates...
were: kanamycin, 50 μg/ml; chloramphenicol, 10 μg/ml; ampicillin, 13 μg/ml. Colonies on the triple antibiotic plates were re-isolated on new triple antibiotic plates, as above, except with 100 μg/ml ampicillin, to ensure elimination of false positives. Clones surviving this final selection were isolated from all library/transposon combinations.

2.4. DNA sequencing and annotation

A number of signal-trapped clones were selected, inoculated into 96-well growth blocks with 1 ml/well Terrific Broth kanamycin (50 μg/ml)/chloramphenicol (10 μg/ml), and incubated at 37 °C, 250 rpm for 2 days. Plasmid DNA was isolated using QIAprep 96 Turbo Miniprep Kit (Qiagen) and sequenced using primers specific for TnSig; seqA2 (5'-AGC GTC GCC GGC GAT CC-3') and seqB (5'-TTA TTC GGT CGA AAA GGA TCC-3'). DNA sequencing was performed with an ABI 3730 XL automatic DNA sequencer using BigDye v. 3.1 terminators. Removal of vector sequence, was performed with an ABI 3730 XL automatic DNA sequencer using BigDye v. 3.1 terminators. Removal of vector sequence, and assembly of DNA sequences was done with the PhredPhrap package (Ewing et al., 1998), followed by manual inspection and assembly correction.

Annotation was accomplished using BLASTX (http://blast.ncbi.nlm.nih.gov/Blast.cgi) (Altschul et al., 1997). Signal-trapped cDNAs were analyzed with SignalP v. 3.0 (http://www.cbs.dtu.dk/services/SignalP/) (Bendtsen et al., 2004; Nielsen et al., 1997) for secretory proteins having an N-terminal signal peptide or a signal anchor, and TMHMM v. 2.0 (http://www.cbs.dtu.dk/services/TMHMM-2.0/) (Krogh et al., 2001) for proteins carrying transmembrane helices.

2.5. Phylogenetic analysis

The phylogenetic analysis was performed on the Phylogeny.fr platform (Dereeper et al., 2008) and comprised the following steps: Selected signal-trapped cDNAs were used as queries to search the NCBI non-redundant protein sequence database for homologous, full-length sequences using BLAST. The cDNA encoded products and the retrieved sequences were aligned with T-Coffee (v. 6.85) using the following pair-wise alignment methods: the 10 best local alignments (Lalign_pair) and an accurate global alignment (slow_pair) (Notredame et al., 2000). Positions with gaps were removed from the alignment. The phylogenetic trees were reconstructed using the maximum likelihood method in the PhyML-aLRT program (v. 3.0) (Guindon and Gascuel, 2003). The default substitution model was selected assuming an estimated proportion of invariant sites (of 0.045 for the trypsin-like serine protease-tree and 0.036 for the lipase tree) and gamma-distributed rate categories to account for rate heterogeneity across sites. The gamma shape parameter was estimated directly from the data (gamma = 1.963) for the trypsin-like serine protease-tree and 1.795 for the lipase tree). Reliability for internal branching was assessed using the aLRT test (SH-Like) (Anisimova and Gascuel, 2006). Graphical representation and editing of the phylogenetic trees was performed with TreeDyn (v. 198.3) (Chevenet et al., 2006).

2.6. Genomic PCR

Isolates of P. neaphidis, E. planchoniana, and C. obscurus were grown in Grace’s Insect Culture Medium supplemented with 10% v/v fetal bovine serum. Protoplasts or hyphal bodies were collected by centrifugation for 10 min at 2000g, freeze-dried, and carefully crushed prior to DNA extraction with 200 μl extraction buffer (50 mM Tris–HCl pH 7.8, 50 mM EDTA, 150 mM NaCl, 2.5% w/v Na–lauroylsarcosine, 500 mM β-mercaptoethanol) and proteinase K (600 μg/ml; Sigma–Aldrich), and incubation at 65 °C for 4 h. The NaCl concentration of the solution was then adjusted to 1 M, and an equal volume of chloroform:octanol 24:1 added. After centrifugation at 12,000g for 2 min, the aqueous phase was extracted again with chloroform:octanol. Genomic DNA was precipitated with 0.6 volume isopropanol, rinsed with 70% ethanol, and, finally, dissolved in TE pH 7.8. Genomic DNA was extracted from field-collected grain aphids using the Nucleospin Plant II Kit with Lysisbuffer PL1 (Macherey–Nagel). For PCR, 1 μl genomic DNA diluted 1:100 in water was used in a 20 μl reaction with 1 u Taq DNA polymerase, 1× Taq Buffer with 1.5 mM MgCl2 (Ampliqon), 0.2 μM each dNTP, and 0.5 μM each primer (Table 1). Program: pre-incubation at 95 °C for 2 min; 25–40 cycles of 94 °C for 15 s, 50–65 °C for 15 s, and 72 °C for 30 s; final extension at 72 °C for 5 min. For the deduced insect genes HM001243–HM001246, no products were obtained (even after increasing extension time to 90 s), probably due to long introns (primers not included in Table 1).

3. Results and discussion

3.1. Production of a rich host–pathogen interaction cDNA library

To discover proteins secreted in the interaction between grain aphids and entomophthoralean fungi, we generated a host–pathogen interaction cDNA library. To ensure enrichment of the library with cDNAs specific to the host–pathogen interaction, we sampled aphids during a natural fungal epidemic. Aphid populations in arable land are typically infected with more than one entomophthoralean species (Jensen et al., 2008). On collection days, E. planchoniana was the most prevalent (15–27% of infected aphids), followed by P. neaphidis (4–8%), and with only a few percent infected by C. obscurus (Fig. 2). We selected infected aphids in the afternoon, as many entomophthoralean species kill their insect hosts during that time of day to prepare for active discharge of new infective spores in the evening and night (Krasnoff et al., 1995; Milner et al., 1984; Roy et al., 2006) (see Section 2.1). Thus, the infected individuals were selected at a stage when enzymes that assist in cuticle penetration from inside the insect host were likely to be produced.

3.2. TAST screening confirmed the richness of the host–pathogen interaction cDNA library

The host–pathogen library was tagged with the TnSig transposon carrying a signal-less β-lactamase (Amp’ gene (Becker et al., 2004; Hamann and Lange, 2006). Only E. coli clones containing full-length cDNAs (with transposon inserts) encoding secreted or membrane-bound proteins were able to grow on ampicillin plates. A collection of 576 ampicillin-resistant E. coli colonies was selected for sequencing. Contig formation of sequences from the 576 clones resulted in assembly of 185 cDNAs that were used to search the NCBI non-redundant protein sequence and expressed sequence tag (EST) databases using BLAST. Of these, 130 gave significant BLAST hits. Twenty-six cDNAs were predicted to encode extracellular or membrane-bound enzymes from different enzyme classes (Fig. 3). Of these, 21 were fungal (11 from E. planchoniana, nine from P. neaphidis, and one from C. obscurus), two were of insect origin, and three were bacterial or bacterial-like, based on homology and/or PCR with gene-specific primers (Fig. 4). The apparent overrepresentation of secretome cDNAs from P. neaphidis suggested that this species was producing a complex and diverse secretome when aphids were collected (see Section 2.1). Selected fungal and insect signal-trapped cDNAs are in Table 2.

An evaluation of the list of signal-trapped cDNAs suggested that the TAST procedure was as successful at capturing insect cDNAs as fungal cDNAs because: (1) The most highly represented gene of the 576 selected clones (found 21 times) was predicted to encode a
Fig. 3. Diversity of secreted enzymes discovered by Transposon Assisted Signal Trapping (TAST).

secreted aphid protein of unknown function, based on high homology to pea aphid ESTs; (2) Several signal-trapped cDNAs in the insect part of the interaction library were predicted to encode silk-like and cuticle proteins (data not shown). Furthermore, TAST has been used successfully with pupae of the turnip moth, Agrotis segetum (M.N. Grell, A.B. Jensen, and L. Lange, unpublished results) and with other eukaryotes (Lange and Schnorr, 2004). Thus, we conclude that the low number of insect extracellular enzymes was not the result of a TAST bias for fungal signal peptides.

The next most prevalent genes in the collection of sequenced clones were a fungal subtilisin-like protease (E. planchoniana SPR1, EpSPR1; Table 2), found 14 times, and a gene of unknown origin and function, found 11 times. That two of the presumably most highly expressed secretome genes had no homology to proteins of known function reflects the limitations of our current molecular understanding of this highly specialized host–pathogen interaction.

3.3. The fungal pathogens produced host degrading enzymes

Three signal-trapped cDNAs were homologous to cuticle-degrading subtilisins. EpSPR1 (see Section 3.2) and P. neoaphidis SPR1 (PnSPR1) aligned most significantly with the subtilisin-like serine protease PR1A of the ascomycete entomopathogen Metarhizium anisopliae var. anisopliae. PnSPR2 aligned most significantly with bassiasin I, a PR1A homolog of another widespread ascomycete entomopathogen, Beauveria bassiana, respectively (Charnley, 2003; St. Leger et al., 1996a; Zhang et al., 2008). A number of clones deduced to encode cuticle-degrading enzymes, such as subtilisin-like serine proteases and chitinases, are essential for breaching the host surface, both during initial infection and during sporulation (Charnley, 2003). PR1A and CDEP-1 (96% identical to bassiasin I) have been shown to be pathogenicity determinants of M. anisopliae and B. bassiana, respectively (Charnley, 2003; St. Leger et al., 1996a; Zhang et al., 2008). A number of clones deduced to encode cuticle-degrading enzymes were identified in two EST studies with cultures of the entomophthoralean pathogens Conidiobolus coronatus and Zoophthora radicans, grown on insect cuticle and isolated from lepidopteran larvae, respectively (Freimoser et al., 2003; Xu et al., 2009). Our findings were substantiated by the identification of homologous clones in these libraries (Table 2, EST hits). We also identified a P. neoaphidis gene (PnTRY1) with 62% amino acid identity to a trypsin-like serine protease from Z. radicans (Table 2). In Z. radicans, this trypsin-like protease (ZrSP1) was implicated in pathogenesis, as it is expressed during infection (Xu et al., 2006). In M. anisopliae, trypsin-like proteases have been shown to be localized around appressoria and penetration pegs (St. Leger et al., 1996b) (see also Section 3.6).

Three identified genes, EpCHT1, EpCHT2, and PnCHT1, matched endochitinases of the glycoside hydrolase 18 (GH18) family (Cantarel et al., 2009) from basidiomycetes and oomycetes. A homolog of both EpCHT1 and PnCHT1 is expressed by C. coronatus grown on insect cuticle (Table 2, EST hits). The deduced product of PnCHD1, containing both a chitin-binding domain and an esterase/deacetylase domain according to the Conserved Domain
Pathogens were synthesizing cell wall material within the selected mycelium grown in culture (Table 2, EST hits). EpUtr2A and EpUtr2B included putative GH16-GPI-glucanosyl-homology to proteins that bind chitin fibrils within the cuticle are masked by proteins (Charnley, 2003). The reason for not identifying any C. obscurus cDNAs encoding penetration-specific enzymes is probably due to its low prevalence on collection days.

The deduced products of three signal-trapped cDNAs, EpLip1, PnLip1, and PnLip2, aligned most significantly with lipases of ciliates (Table 2) and cultivable zygomycetes in the order Mucorales. Lipases may be involved in degrading the epicuticle, but as a homolog of both EpLip1 and PnLip1 is expressed by Z. radicans in culture, but not C. coronatus grown on cuticle (Table 2, EST hits), these lipases are more likely nutritional lipases that degrade host lipids in the hemolymph (Hegedus and Khachatourians, 1988) or fat bodies (Butt et al., 1981) (see also Section 3.6). The product of the identified carboxypeptidase gene (EpMep1) (Table 2) is probably also involved in nutrient acquisition.

3.4. The fungal pathogens were in the process of cell wall regeneration

A number of fungal signal-trapped cDNAs were deduced to encode glycosyl transferases (EpCHS3, PnCHS3, and C. obscurus OCH1, CoOCH1) and cell wall glucanases (EpUTR2A, EpUTR2B, and PnGBP1) (Table 2). EpCHS3 and PnCHS3 showed homology to Saccharomyces cerevisiae chitin synthase 3, which in yeast is responsible for general chitin synthesis in the cell wall and chitosan synthesis in the spore (Shaw et al., 1991). The product of the third predicted glycosyl transferase gene (CoOCH1) had a central domain with homology to the Och1p-1,6-mannosyltransferase domain including a DXD sugar-binding motif, according to CDD. S. cerevisiae Och1p is involved in cell wall mannoprotein synthesis (Nakanishi-Shindo et al., 1993). The deduced products of EpUTR2A and EpUTR2B aligned significantly with cell wall glucanases of both filamentous fungi and yeasts, including S. cerevisiae Chh2p/Ur2p, a glycosylphosphatidylinositol (GPI)-anchored chitin transglycosidase that transfers chitin from chitin synthase 3 to β-1,6-glucan, to create cross-links during cell wall biosynthesis (Cabib et al., 2008). Both EpUTR2A and EpUTR2B included putative GH16-GPI-glucanosyltransferase and chitin-binding domains, according to CDD. PnGBP1 was also predicted to encode a GH16 family enzyme, but with homology to proteins that bind β-1,3-glucan, the major constituent of the entomophthoralean fungal cell wall (Latgé and Beauvais, 1987). UTR2 and GBP homologs are expressed by Z. radicans mycelium grown in culture (Table 2, EST hits).

Expression of these genes may indicate that entomophthoralean pathogens were synthesizing cell wall material within the selected aphid hosts. Together with the expression of penetration-specific genes, this implied they were in transition from the internal protoplast-like growth phase (Butt et al., 1981; Freimoser et al., 2001) to the cell wall-delimited hyphal growth phase that marks outward penetration and conidiophore development—the stages that were our primary target for discovery. Fig. 5 (upper) shows the predicted location of the discovered fungal gene products.

3.5. Do aphids possess an anti-fungal defense?

Only two signal-trapped cDNAs were predicted to encode extracellular enzymes of host origin (see Section 3.2). One of these (HM001243) showed homology to insect trypsin-like serine proteases that are part of the prophenoloxidase-activating system (Table 2), which is an important part of the invertebrate innate immune system. The system involves a serine protease cascade that is activated by the recognition of invading pathogens, mediated by β-1,3-glucan binding proteins, and eventually leads to melanization of the pathogen and release of toxic quinone intermediates (Cerdenius et al., 2008) (Fig. 5). However, phenoloxidases are also involved in wound healing/sclerotization (Cerdenius et al., 2008) and may be expressed merely as a reaction to injury from fungal penetration. Other genes that may have been expressed in response to pathogen invasion were two signal-trapped cDNAs with homology to the fruit fly lipocalin neural lazarillo (HM001245), and the moth lipocalins gallerin and bombyrin (HM001246) (Table 2). Lipocalins are small proteins with low sequence conservation that bind hydrophobic molecules. Neural lazarillo was implicated to be involved in stress tolerance (Hull-Thompson et al., 2009) and a related 23 kDa lipocalin from the fall webworm, Hyphantria cunea, in repair of brain tissue (Kim et al., 2005).

A potential novel host defense is encoded by HM001244, based on homology to defensin-like cobatoxin-like peptides of moths (Table 2). However, we were unable to amplify the gene from S. avenueae (or any of the fungal species) using five different primer pairs, even though several insect anti-fungal defensins have been described (Aerts et al., 2008). The origin of this novel insect defensin therefore remains unconfirmed. A signal-trapped cDNA of confirmed host origin, SaCatb-16a, showed homology to cysteine proteases in the cathepsin B family (Table 2) (this is the other signal-trapped cDNA predicted to encode an extracellular enzyme of host origin). A cathepsin B from the social aphid Tuberaphis syriacus (cathepsin B-S) is specifically expressed in the midgut of soldiers and is a component of soldier venom (Kutsukake et al., 2004). However, the closest BLAST matches were cathepsin B-16A from Acr-ythisphoton pismum (pea aphid) and other nonsoldier-specific cathepsins (cathepsin B-Ns) (Rispe et al., 2008). Therefore, SaCatb-16A is most likely not a venomous protease, but may function as a digestive enzyme in the aphid midgut (Rispe et al., 2008). Despite the lack of confirmed host defense proteins, one of our fun-
<table>
<thead>
<tr>
<th>Accession No.</th>
<th>Gene family</th>
<th>Predicted function</th>
<th>SigP or TMH</th>
<th>Best BLASTX hit</th>
<th>Entomophthorales EST hits</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fungal</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HM001221</td>
<td>EpASP1</td>
<td>Vacular endopeptidase</td>
<td>–</td>
<td>Cryptococcus neoformans</td>
<td>XP_566887 60</td>
</tr>
<tr>
<td>HM001222</td>
<td>pepM14</td>
<td>Carboxypeptidase A/B</td>
<td>1–15</td>
<td>Acryothrophis pism</td>
<td>Zr</td>
</tr>
<tr>
<td>HM001223</td>
<td>pepP1</td>
<td>Trypsin-like serine protease</td>
<td>1–18</td>
<td>Zoophthora radicans</td>
<td>AAW31593 62</td>
</tr>
<tr>
<td>HM001224</td>
<td>pepS8A</td>
<td>Cuticle-degrading subtilisin-like protease (PR1A-like)</td>
<td>1–18</td>
<td>Metarhizium anisopliae var. anisopliae</td>
<td>CAC95042 46</td>
</tr>
<tr>
<td>HM001225</td>
<td>pepS8A</td>
<td>Cuticle-degrading subtilisin-like protease (PR1A-like)</td>
<td>1–20</td>
<td>Metarhizium anisopliae var. anisopliae</td>
<td>CAC95042 50</td>
</tr>
<tr>
<td>HM001226</td>
<td>pepS8A</td>
<td>Cuticle-degrading subtilisin-like protease (PR1A-like)</td>
<td>–</td>
<td>Beauveria bassiana</td>
<td>AAD29255 48</td>
</tr>
<tr>
<td>HM001227</td>
<td>GH16</td>
<td>Cell wall glucanase</td>
<td>1–16</td>
<td>Zygosaccharomyces rouxii</td>
<td>CAR28493 37</td>
</tr>
<tr>
<td>HM001229</td>
<td>GH16</td>
<td>Glucan binding protein</td>
<td>–</td>
<td>Aspergillus niger</td>
<td>XP_001391205 36</td>
</tr>
<tr>
<td>HM001230</td>
<td>GH18</td>
<td>Endochitinase</td>
<td>–</td>
<td>Actinophrys sol</td>
<td>BAC32349 52</td>
</tr>
<tr>
<td>HM001231</td>
<td>GH18</td>
<td>Endochitinase</td>
<td>–</td>
<td>Amanita muscaria</td>
<td>CAC35202 33</td>
</tr>
<tr>
<td>HM001232</td>
<td>GH18</td>
<td>Endochitinase</td>
<td>–</td>
<td>Postia placenta</td>
<td>EED84840 38</td>
</tr>
<tr>
<td>HM001233</td>
<td>GH17</td>
<td>Endoglucanase</td>
<td>1–18/1–23</td>
<td>Aphanomyces astaci</td>
<td>ACJ04636 30</td>
</tr>
<tr>
<td>HM001234</td>
<td>CE4</td>
<td>Chitin deacetylase</td>
<td>1–15</td>
<td>Sclerotinia sclerotiorum</td>
<td>XP_001598179 30</td>
</tr>
<tr>
<td>HM001235</td>
<td>Lipase_3</td>
<td>Triacylglycerol lipase</td>
<td>–</td>
<td>Paramecium tetraurela</td>
<td>XP_001455849 27</td>
</tr>
<tr>
<td>HM001236</td>
<td>Lipase_3</td>
<td>Triacylglycerol lipase</td>
<td>–</td>
<td>Tetrabymena thermophila</td>
<td>XP_001025189 31</td>
</tr>
<tr>
<td>HM001237</td>
<td>Lipase_3</td>
<td>Triacylglycerol lipase</td>
<td>–</td>
<td>Paramecium tetraurela</td>
<td>XP_001432795 28</td>
</tr>
<tr>
<td>HM001238</td>
<td>GT2</td>
<td>Chitin synthase</td>
<td>97–119/129–151/156–178</td>
<td>Laccaria bicolor</td>
<td>XP_001877307 69</td>
</tr>
<tr>
<td>HM001239</td>
<td>GT2</td>
<td>Chitin synthase</td>
<td>87–109/113–145/142–168</td>
<td>Rhizopus microsorps var. oligosporus</td>
<td>BAA34808 76</td>
</tr>
<tr>
<td>HM001240</td>
<td>GT2</td>
<td>Chitin synthase</td>
<td>9–26</td>
<td>Aspergillus niger</td>
<td>XP_001390011 76</td>
</tr>
<tr>
<td>HM001241</td>
<td>FA_desaturase</td>
<td>Δ-6 fatty acid desaturase</td>
<td>32–54/75–92/97–119</td>
<td>Mortierella alpina</td>
<td>BAA85588 51</td>
</tr>
<tr>
<td><strong>Insect</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HM001242</td>
<td>pepC</td>
<td>Endopeptidase</td>
<td>1–21/1–20</td>
<td>Acryothrophis pism</td>
<td>NP_001119617 87</td>
</tr>
<tr>
<td>HM001243</td>
<td>pepS1</td>
<td>Trypsin-like serine protease/Prophenoloxidase activating factor</td>
<td>1–18</td>
<td>Nasonia vitripennis</td>
<td>NP_001155060 62</td>
</tr>
<tr>
<td>HM001244</td>
<td>Defensin</td>
<td>Antimicrobial protein</td>
<td>1–16</td>
<td>Trichoplusia ni</td>
<td>ABV68874 38</td>
</tr>
<tr>
<td>HM001245</td>
<td>Lipocalin/CFABP</td>
<td>Lipocalin (neutral lazarillo-like)</td>
<td>1–16</td>
<td>Drosophila viridis</td>
<td>NP_002051716 28</td>
</tr>
<tr>
<td>HM001246</td>
<td>Lipocalin/CFABP</td>
<td>Lipocalin (gallerin/bombyrin-like)</td>
<td>1–16</td>
<td>Galleria mellonella/Bombyx mori</td>
<td>AAA85085/39 NP_001036872 36-96</td>
</tr>
</tbody>
</table>

a Two first letters in name indicates origin: Ep, Entomophthora planchoniana; Ps, Pandora neaphidius; Co, Conidiobolus obscurus; Sa, Sitobion avenae; as determined by PCR (Fig. 4). –, no gene name assigned/PCR unsuccessful, probably due to long introns.

b Affiliation to family was deduced from BLASTX hit list. Family designation according to MEROPS (http://merops.sanger.ac.uk/): pepA (aspartic proteases), pepC (cysteine proteases), pepM (metalloproteases) and pepS (serine proteases); CAzy (http://www.ewb.ca/). (GH (glycoside hydrolases), CE (carbohydrate esterases) and GT (glycosyl transferases); PFAM (http://pfam.sanger.ac.uk/) and InterPro (http://www.ebi.ac.uk/interpro). HM001233 could not be assigned to a specific family. CFABP, cystolic fatty-acid binding protein.

c Deduced from BLASTX hit list.

d SigP, signal peptide (interval(s) starting with 1); TMH, transmembrane helix (other intervals). HM001240 was predicted to contain a signal anchor. Numbers according to deduced amino acid sequence. Some signal-trapped cDNAs lack a signal peptide interval due to incomplete sequencing (–). Two SigP intervals mean different cleavage site predictions by SignalP-NN and SignalP-HMM, respectively. The number of TMHs of transmembrane proteins is a minimum due to incomplete sequencing.

e Non-human, non-mouse EST database at NCBI was searched with the signal-trapped cDNA sequence using TBLASTX. Best hits from Conidiobolus coronatous (Cc) and Zoophthora radicans (Zr) EST libraries are indicated. –, no Entomophthorales hits.
gal findings indicated that the host interior is not an entirely benign environment: *E. planchoniana* expressed a gene (*EpASP1*) with high similarity to fungal vacuolar aspartyl endopeptidases (Table 2), including *S. cerevisiae* Pep4p. Pep4p is important for protein turnover after oxidative damage, and is thus involved in stress tolerance (Marques et al., 2006). An *EpASP1* homolog is expressed by *Z. radicans* in culture (Table 2, EST hits).

Expression of these genes indicated that the host was alive during pathogen vegetative growth and sporulation onset. However, the low number (or possible absence) of detected expressed host anti-fungal genes supports the hypothesis that protoplastic Entomophthorales can escape host immune detection (Beauvais et al., 1989). It also suggests that defense responses are induced only late in the infection cycle during hyphal cell wall regeneration, when host death is unavoidable, or that aphids possess only a limited anti-fungal defense, as suggested by Gerardo et al. (2010). The predicted location of the discovered host gene products is shown in Fig. 5 (lower).

### 3.6. Enzymes from entomophthoralean fungi are unique

To determine the uniqueness of enzymes secreted by the entomophthoralean fungi, we phylogenetically analyzed the predicted amino acid sequences of selected genes. After discovering a putative *P. neaphidis* ortholog (PnTry1) of the *Z. radicans* trypsin-like protease ZrSP1, we constructed a phylogenetic tree for comparison to trypsin-like proteases from other fungi (Fig. 6). The *P. neaphidis* and *Z. radicans* proteases group together on a strongly supported branch and form a sister group to a *C. coronatus* protease. The long branches between the *C. coronatus* protease and the *Pandora/Zoopltheta* protease group might occur because *C. coronatus* is among the most basal fungi within the monophyletic order Entomophthorales (Jensen et al., 1998). Together, the entomophthoralean proteases form a sister group to most proteases from ascomycete species, including two proteases from *M. anisopliae*, but excluding those from *Aspergillus* species, which form separate groups. Thus, the entomophthoralean trypsin-like proteases, including the one discovered here, appear to be unique from those of other fungi.

We further investigated how enzymes from the obligate pathogenic entomophthoralean fungi compare to enzymes produced by the ubiquitous, saprotrophic, and easily cultivable mucoralean zygomycetes. Both Entomophthorales and Mucorales were formerly placed in the phylum Zygomycota, but have recently been assigned to the subphyla *incertae sedis* Entomophthoromycotina and Mucoromycotina, respectively (Hibbett et al., 2007). We chose the three lipases (EpLip1, PnLip1, and PnLip2) for this analysis, because all aligned significantly (E-value <2e−11) with a number of lipases from the Mucorales genera *Rhizomucor* and *Rhizopus*. We included several lipases from other fungi and “protists” that also aligned significantly with the three lipases (Fig. 7). From the results, PnLip2 is clearly a sister to the group of selected mucoralean lipases, while PnLip1 and EpLip1 are on a different branch as a sister group to the selected basidiomycete lipases. The selected ascomycete and “protist” lipases form separate groups. PnLip1 and PnLip2 may thus represent two paralogous lipase clades. Neither PnLip1/EpLip1 nor PnLip2 appear to be closely related to any deposited sequences, and are separated from their nearest neighbors by long branches. An exception is the previously mentioned...
Fig. 6. Phylogenetic tree of trypsin-like proteases. P. neoaphidis trypsin-like serine protease PnTry1 of the present study (arrow) and its potential Z. radicans ortholog (ZrSP1) (AAW31593) are compared to full-length, trypsin-like proteases from other fungi. Entomophthoralean sequence group is indicated by a black vertical line. The tree was reconstructed using the maximum likelihood method in the PhyML-aLRT program. Reliability for internal branching is shown as a percentage at each node. Only values above 50% are shown. Bar indicates the number of amino acid substitutions. Each protein retrieved from the NCBI non-redundant sequence database is marked with species and database accession No. See Section 2.5 for additional details.

Fig. 7. Phylogenetic tree of lipases. Shown is the relationship of the entomophthoralean lipases identified in the present study (arrows) with selected full-length fungal and "protist" lipases. The six groups indicated are: (I) selected ascomycete lipases; (II) selected "protist lipases"; (III) selected mucoralean lipases; (IV) entomophthoralean lipase PnLip2; (V) selected basidiomycete lipases; and (VI) entomophthoralean lipases PnLip1 and EpLip1. The phylogenetic tree was reconstructed using the maximum likelihood method in the PhyML-aLRT program. Reliability for internal branching is shown as a percentage at each node. Only values above 50% are shown. Bar indicates the number of amino acid substitutions. Each protein retrieved from the NCBI non-redundant sequence database is marked with species and database accession No. See Section 2.5 for additional details.
Z. radicans EST clone (see Section 3.3; not included in the lipase tree), which is probably another member of the Pln1p/Epl1p clade. This analysis highlights the uniqueness of the entomophthoralean secreted enzymes compared to the mucoralean sequences.

4. Conclusions

Secretome-directed cDNA library screening shows potential as a state-of-the-art approach for discovering secreted proteins involved in complex interspecies interactions, especially for organisms whose genome is not yet sequenced. By mapping proteins involved in complex interspecies interactions, especially for organ-isms whose genome is not yet sequenced, we discovered several fungal gene products potentially involved in host cuticle penetration, nutrient acquisition, stress tolerance, and cell wall regeneration during conidiophore development. In contrast, grain aphids appeared to express only a few genes involved in response to pathogen invasion; in other words, they mounted only a limited immune response. This is consistent with the recently published pea aphid genome that showed that aphids have a limited immune repertoire (Gerardo et al., 2010). Interestingly, we identified a possible insect defensive molecule, although this was not unequivocally linked to the host.

Finally, the high diversity and phylogenetic uniqueness of the discovered hydrolytic enzymes indicate that pathogenic fungi, such as species within the order Entomophthorales, contain proteins whose characterization will add to our understanding of host–pathogen interactions, and lead to discovery of novel proteins with potential biotechnological applications.

Acknowledgments

The project was financially supported by The Danish Council for Strategic Research and The Danish National Research Foundation (Centre for Social Evolution). Line Sondt (Aalborg University) and Sanne Reib Jensen (Novozymes) are gratefully acknowledged for expert technical assistance. Novozymes is acknowledged for a royalty-free license to use the TAST technology for research purposes. M.N. Grell et al. / Fungal Genetics and Biology 48 (2011) 343–352

References

Nakanishi-Shindo, Y., Nakayama, K., Tanaka, A., Toda, Y., Jigami, Y., 1993. Structure of the N-linked oligosaccharides that show the complete loss of α-1, 6-polymannose outer chain from och1, och1 mnn1, and och1 mnn1 algl3 mutants of Saccharomyces cerevisiae. J. Biol. Chem. 268, 26338–26345.