

Generation of a wheat leaf rust, *Puccinia triticina*, EST database from stage-specific cDNA libraries

GUANGGAN HU^{1†}, ROB LINNING¹, BRENT MCCALLUM², TRAVIS BANKS², SYLVIE CLOUTIER², YARON BUTTERFIELD³, JERRY LIU³, ROBERT KIRKPATRICK³, JEFF STOTT³, GEORGE YANG³, DUANE SMAILUS³, STEVEN JONES³, MARCO MARRA³, JACQUELINE SCHEIN³ AND GUUS BAKKEREN^{1*}

¹Pacific Agri-Food Research Centre, Agriculture and Agri-Food Canada, Highway 97, Summerland, BC, V0H 1Z0, Canada

²Cereal Research Centre, Agriculture and Agri-Food Canada, 195 Dafoe Road., Winnipeg, MB, R3T 2M9, Canada

³Michael Smith Genome Sciences Centre, 600 W 10th Avenue, Vancouver, BC, V5Z 4E6, Canada

SUMMARY

Thirteen cDNA libraries constructed from small amounts of leaf rust mRNA using optimized methods served as the source for the generation of 25 558 high-quality DNA sequence reads. Five life-cycle stages were sampled: resting urediniospores, urediniospores germinated over water or plant extract, compatible, interactive stages during appressorium or haustorium formation just before sporulation, and an incompatible interaction. mRNA populations were subjected to treatments such as full-length cDNA production, subtractive and normalizing hybridizations, and size selection methods combined with PCR amplification. Pathogen and host sequences from interactive libraries were differentiated *in silico* using cereal and fungal sequences, codon usage analyses, and by means of a partial prototype cDNA microarray hybridized with genomic DNAs. This yielded a non-redundant unigene set of 9760 putative fungal sequences consisting of 6616 singlets and 3144 contigs, representing 4.7 Mbp. At an E -value $\leq 10^{-5}$, 3670 unigenes (38%) matched sequences in various databases and collections but only 694 unigenes (7%) were similar to genes with known functions. In total, 296 unigenes were identified as most probably wheat and ten as rRNA sequences. Annotation rates were low for germinated urediniospores (4%) and appressoria (2%). Gene sets obtained from the various life-cycle stages appear to be remarkably different, suggesting drastic reprogramming of the transcriptome during these major differentiation processes. Redundancy within contigs yielded information about possible expression levels of certain

genes among stages. Many sequences were similar to genes from other rusts such as *Uromyces* and *Melampsora* species; some of these genes have been implicated in pathogenicity and virulence.

INTRODUCTION

Puccinia triticina Eriks. (syn. *P. recondita* Rob. Ex Desm. f. sp. *tritici* Eriks. and Henn.) is a basidiomycete, biotrophic, obligate plant pathogen and the causal agent of leaf or brown rust on wheat (*Triticum aestivum*). It is an economically important pathogen in all wheat-growing areas of the world. Yield losses caused by leaf rust over a large area can be up to 20%, while individual fields can be destroyed when the disease is severe prior to heading (Kolmer, 2005; Saari and Prescott, 1985; Samborski, 1985). Chemical methods can be used to control the disease, but the most cost-effective strategy to reduce the incidence of the disease has been through the development of resistant wheat cultivars. Classical genetic data have demonstrated gene-for-gene interactions underlying incompatibility between *P. triticina* and its host plant (Browder, 1971) and 56 leaf rust resistant genes or alleles (*Lr*) are currently recognized world-wide (<http://www.cdl.umn.edu>; McCallum and Seto-Goh, 2005). However, due to the emergence of new virulent races, resistant cultivars eventually become susceptible to the rust. Novel methods to obtain durable resistance are needed, which requires a better understanding of fungal development and the interaction with its host to identify new targets for disease control.

P. triticina has a complex life cycle which includes five different spore types and two hosts: wheat and meadow rue, *Thalictrum speciosissimum*, its alternate host on which it completes its sexual stage (Horton *et al.*, 2005; Littlefield, 1979). The asexual uredial stage on wheat is the economically important part of the life cycle which can progress from initial infection to sporulation

* Correspondence: Tel.: +1250 4946368; Fax: +1250 4940755, E-mail: BakkerenG@agr.gc.ca

† Present address: Michael Smith Laboratories, University of British Columbia, Vancouver, BC, V6T 1Z3 Canada.

within 10 days under warm, humid conditions, potentially leading to epidemics. During infection on wheat, the pathogen undergoes a high degree of morphological and physiological differentiation (Hu and Rijkenberg, 1998a). After landing on a wheat epidermis, the dikaryotic urediniospore germinates within hours and an emerging germ tube extends until a stomatal pore is encountered. An appressorium is formed on the stoma and an infection peg is produced to enter the plant. Topographical features of the stomatal cells and chemical signals may be involved in the infection structure formation (Collins and Read, 1997). Once the rust has forcibly gained entrance into the underlying cavity, the germ tube differentiates into a substomatal vesicle followed by a haustorial mother cell. This latter fungal structure makes the first significant contact with a host cell between 12 and 24 h post infection (hpi). This entails breaching the cell wall, forming an intercellular haustorium and establishing an intimate feeding relationship. In compatible interactions, the fungus ramifies through the tissue producing intercellular hyphae and intracellular haustoria (Hu and Rijkenberg, 1998b). During incompatible interactions, resistant reactions are mainly observed only after the first haustorium is formed (Niks and Dekens, 1991; Niks, 1983).

For plant fungal pathogens there is limited genome sequence information available, although the total genome sequence of several fungi has been made public and major new projects are underway or are being promoted (Yarden *et al.*, 2003; Fungal Genome Initiative, <http://www.broad.mit.edu/annotation/fungi/fgi/>; Microbial Genome Sequencing Projects by the US Department of Energy, <http://microbialgenome.org/organisms.shtml>; suggested Microbial Genome Sequencing Priority List of the American Phytopathological Society, <http://www.apsnet.org/>). As a prelude to genome sequencing, the generation of an expressed sequence tag (EST) database represents a relatively inexpensive and rapid approach for gene discovery. In addition, ESTs can be used to design molecular markers such as microsatellites or simple sequence repeats (SSRs), to design and construct expression microarrays, and can aid with genome annotations or be used in comparative approaches to reveal relatedness between species (Marra *et al.*, 1999; Skinner *et al.*, 2001; Yarden *et al.*, 2003). Gene discovery and expression profiling through the generation of an EST database is particularly useful for organisms with little or no genetic research history or those that are experimentally difficult to handle. EST collections from several plant fungal pathogens have been compiled (e.g. COGEME repository, <http://cogeme.ex.ac.uk/>; Soanes *et al.*, 2002). However, collections from basidiomycete fungi are under-represented (Abu *et al.*, 2004; Guettler *et al.*, 2003; Karlsson *et al.*, 2003; Lee *et al.*, 2002; Loftus *et al.*, 2005; Ospina-Giraldo *et al.*, 2000). Collections from basidiomycete plant pathogens are restricted to the corn smut *Ustilago maydis* (Sacadura & Saville 2003; Nugent *et al.*, 2004) and small sets from a number of rust fungi.

In rusts, despite detailed cytological research, knowledge about molecular mechanisms underlying infection and development within host tissue is only recently emerging (Hahn, 2000; Voegelé, 2006). The major factor limiting molecular genetic analyses is the essentially obligate nature of fungal development and the lack of a stable DNA transformation system, although progress has been made with respect to the latter (Barja *et al.*, 1998; Bhairi and Staples, 1992; Li *et al.*, 1993; Schillberg *et al.*, 2000; Webb *et al.*, 2005). An additional handicap is the size of the rust genomes; by flow cytometry, genome sizes of rust fungi were measured to range from 54 Mbp for *P. coronata* to 415 Mbp for *Uromyces appendiculatus*, while that of *P. triticina* was estimated at 124 Mbp (Eilam *et al.*, 1994). The genome size of *P. graminis* f. sp. *tritici* was determined to be 67 Mbp by reassociation kinetics (Backlund and Szabo, 1993; Leonard and Szabo, 2005). Despite these hurdles, initial genomics work on several rusts has yielded interesting results over the last few years. The best-studied is the bean rust, *U. fabae*, from which haustorium-specific cDNAs have been described (Hahn and Mendgen, 1997; Jakupovic *et al.*, 2006). The subsequent study of several newly discovered genes and their expression has yielded insight into the subtle biotrophic nature of the interaction with its host (Hahn *et al.*, 1997; Link *et al.*, 2005; Mendgen and Hahn, 2002; Struck *et al.*, 1996, 1998; Voegelé, 2006; Voegelé *et al.*, 2001, 2005; Wirsal *et al.*, 2001, 2004). From the cereal pathogen wheat stem rust, *P. graminis* f. sp. *tritici*, 68 cDNA clones were generated from germinating urediniospores (Liu *et al.*, 1993) and 203 unigenes were developed from infected wheat leaves using an SSH technique (Broecker *et al.*, 2006). Recently, 488 unigenes were described for the soybean rust *Phakopsora pachyrhizi* (Posada-Buitrago & Frederick 2005).

To gain more insight into the molecular basis of *P. triticinal* wheat biotrophism and to identify genes involved in pathogenicity and virulence, we have generated and analysed ESTs from 13 cDNA libraries constructed by using different techniques and representing four life-cycle stages including resting and germinated urediniospores and during appressorial and haustorial development; an existing library representing an incompatible interaction was mined for fungal sequences. Previous reports described small *P. triticina* EST collections from wheat-infected leaves generated through a cDNA-AFLP (Zhang *et al.*, 2003) and an SSH technique (Thara *et al.*, 2003) aimed at revealing *in-planta* expressed rust genes. A proteomics approach revealed up-regulated wheat and fungal proteins (Rampitsch *et al.*, 2006). Genes expressed during the interaction with the host are generally considered to be more relevant to understanding the molecular mechanism of disease development. This presents a challenge to construct fungal-specific cDNA libraries from 'interactive' stages or to discriminate between cDNA sequences from pathogen and host. Haustorial structures can be isolated from heavily infected tissues (Catanzariti *et al.*, 2006; Hahn and Mendgen, 1997) but the isolation procedure

Table 1 Overview of materials and methods used for construction of stage-specific cDNA libraries and number of sequence reads (see also supplementary tables online).

Stage	Library designation	Description of stage†	cDNA library preparation§	No. of reads¶ (quality)
Ptu	PT001*	resting urediniospores	RNA cap + singleT-primers, PCR, normalization	5019 (3917)
	PT005	resting urediniospores	singleT primer + G-tailing, PCR	
Ptg	PT002*†	urediniospores germinated on water + wheat extract, 16 hpi	RNA cap + singleT-primers, PCR, normalization	13 317 (11 000)
	PT003*	urediniospores germinated on water, 16 hpi	RNA cap + singleT-primers, PCR, normalization	
	PT006	urediniospores germinated on water, 16 hpi	singleT-primer + G-tailing, PCR	
	PT007	urediniospores germinated on water, 16 hpi	RNA cap + anchored T-set primers, PCR	
	PT008	urediniospores germinated on water, 2 + 16 hpi pooled	anchored T-set primers + G-tailing, PCR	
	PT011	urediniospores germinated on water, 2 + 16 hpi pooled	anchored T-set primers + G-tailing, PCR	
Pta	PT010†	appressorium stage at 5 hpi	anchored T-set primers + G-tailing, PCR	6319 (5988)
Pth	PT004*†	haustorial stage at 6 dpi	RNA cap + singleT-primers, PCR, normalization, subtraction	4519 (4002)
	PT009†	haustorial stage at 6 dpi	anchored T-set primers + G-tailing, PCR	
	PT012†	haustorial stage at 6 dpi	anchored T-set primers + G-tailing, PCR	
	PT013†	haustorial stage at 6 dpi	Stratagene adapter; no PCR	
Pti	TaLr1	haustorial stage at 24 hpi, incompatible interaction	Stratagene adapter; no PCR	7727**

*Libraries were produced from subtracted and normalized cDNAs and were excluded from the stage-specific abundance analysis.

†Interactive libraries; note that PT002 was considered in this category because of added wheat extract.

‡hpi = hours post inoculation; dpi = days post inoculation (see Experimental procedures).

§See Experimental procedures for detailed description

¶29 174 PT-reads yielded 26 723 quality sequences (92%) which have been deposited in GenBank dbEST (DN956226 to DN956230, EC387017 to EC387038, EC394906 to EC421413 and EC917356 to EC917543); I.D. prefix 'PT00X' refers to the library designation 'X' in this table and subsequent numbers refer to a particular clone number as a storage plate coordinate. The suffix '.BR' etc. refers to the primer used for sequencing (see Experimental procedures). Additional computational filtering and chimer-splitting (see Experimental procedures) yielded 24 907 sequences.

**7727 TaLr1 quality sequences yielding 7460 ESTs (due to overlap between some 5'- and 3'-reads) were generated in a separate study (B. Fofana *et al.*, unpublished data) and deposited previously in GenBank (accessions BQ619578 to BQ620868, BG903972 to BG910140).

risks generating major changes in expression profiles. Prior to the construction of cDNA libraries, the mRNA populations from infected plant tissues can be physically subtracted with unchallenged host RNA. Alternatively, sequenced cDNA clones can be distinguished on the basis of G+C content (Qutob *et al.*, 2000), codon bias (Hsiang and Goodwin, 2003; Li *et al.*, 2004; Maor *et al.*, 2003), through identification by regular DNA blots (Hahn and Mendgen, 1997; Rauyaree *et al.*, 2001; Thara *et al.*, 2003; Zhang *et al.*, 2003), using PCR (Sugui and Deising, 2002) and/or subtracted *in silico* with available host sequences (Kruger *et al.*, 2002).

RESULTS

A detailed report on the construction of the cDNA libraries from various life-cycle stages of fungal development, the outline of the EST sequencing, and the results for the normalization, subtraction and full-length procedures are given in supplementary materials Appendix S1. Table 1 summarizes the stages sampled and number of reads obtained.

Assembly of a unigene set of putative fungal origin

Due to the manner in which our cDNA libraries were constructed, some of the clones contained chimeric inserts; that is, certain inserts were composed of cDNA fragments originating from different mRNA molecules ligated erroneously 'head-to-head' or 'tail-to-tail' around *Apal* or *NotI* restriction sites used to generate the inserts. Potential chimeric inserts were identified, and subsequently separated *in silico* by finding sequences based on the presence of internal *Apal* or *NotI* sites; non-chimeric sequences erroneously separated in this manner were identified by virtue of their joint appearance in the same contig (see below) coupled with similar BLAST homology results and were consequently reassembled (see Experimental procedures).

DNA sequences derived from resting urediniospores (stage Ptu) and urediniospores germinated *in vitro* (stage Ptg) were considered purely fungal; BLAST comparisons to the 'cereal database' did not indicate plant contaminants (other than the occasional leaf rust sequences that match both fungal and cereal/

plant sequences with similar *E*-values, discussed below). However, sequences that were derived from 'interactive' stages (appressorial and haustorial) needed to be distinguished from genes that were 'most likely wheat'. To achieve this, we employed three methods. Two *in silico* analyses used similarity searches through BLAST and codon bias (see Experimental procedures). A third method employed comparative genome hybridizations (CGH) using a prototype cDNA microarray representing an initial 4484-unigene subset representing libraries PT001 to PT013 (G.B., unpublished data) and hybridizations to wheat and rust total genomic DNAs (data not shown). We did not want to discard sequences that were potentially fungal and therefore employed rather conservative criteria; unigenes that could not be assigned an origin by these tests were considered potentially fungal. A subtraction method based solely on G+C content (e.g. Qutob *et al.*, 2000) proved unsuitable because G+C content of 400 analysed wheat genes ranged broadly from 40 to 70 mol% with most around 54 mol%, basically encompassing that of the set of pure rust sequences which formed a much narrower Gaussian distribution with an average calculated content of 48.2 mol% (data not shown).

The cDNA library from an incompatible interaction between Thatcher Lr1 and race 1 is not expected to yield many fungal sequences as little fungal biomass will have accumulated in host tissue undergoing a hypersensitive response triggered within 24–48 hpi. Nevertheless, fungal ESTs from this stage will be very informative in shedding light on the pathogen response to an active HR resistance. The previously generated 7727 TaLr1 quality sequences (Table 1, B. Fofana *et al.*, unpublished data) yielded a unigene set of 3769 sequences (2772 singlets and 997 contigs), which was mined for potential fungal sequences using conservative criteria (see Experimental procedures). A first set of similarity searches to collected Gramineae and fungal databases extracted 651 putative fungal unigenes from this stage, Pti (28 most likely fungal, 408 possibly fungal and 215 without any matches at $E \leq 10^{-5}$). Including these 651, we started analyses with a total set of 25 558 quality sequences. A non-redundant set of 9760 unigenes, consisting of 2848 singlets and 6912 contigs, was initially generated by the CAP3 software (Huang and Madan, 1999). Contigs were manually inspected for proper assembly and 3768 contigs consisting of overlapping 5'- and 3'-reads from the same clone were classified as singlets. As a result, the final unigene set consisted of 6616 singlets and 3144 contigs with an average length of 485 bp.

Similarity to databases, annotation and classification of ESTs according to putative biological functions

Various BLAST similarity searches were performed with the complete putative fungal data set of 9760 unigenes resulting in matches and annotations with various levels of confidence.

Table 2 Categories of similarity alignments of *P. triticina* unigenes against different databases.

Similarity*	e-value	NR	FUD	EO
high	$\leq 10^{-20}$	594	407	1480
moderate	$10^{-5} \sim 10^{-20}$	667	858	2271
low	$10^{-3} \sim 10^{-5}$	309	1412	1978
total		1570 (16%)	2677 (27%)	5729 (59%)
no match	$> 10^{-3}$	8190 (84%)	7083 (73%)	4031 (41%)

*Amino acid sequence similarity based on the BLASTX software to NR (GenBank non-redundant protein) or the TBLASTX software to FUD (custom collection of fungal DNA sequences; supplementary Table S7) or EO (est_others) databases.

Table 3 Summary of stage-specific ESTs, similarity searches and annotation achieved.

Stages	Input ESTs	Unigenes (%)*	Hits (%)‡	Annotated (%)¶
Ptu	3917	920 (24)	372 (41)	214 (23)
Ptg	11 000	3711 (34)	1178 (32)	153 (4)
Pta	5988	2910 (49)	877 (30)	56 (2)**
Pth	4002	1631 (41)	985 (61)	428 (26)**
Pti	7727	651†	176 (27)§	127 (20)**
total set	25 558	9760 (38)	3670 (38)	1000 (10)**

*Number of unique sequences (singlets plus contigs) representing one transcribed gene are given with percentage of total number of quality sequences for that stage in parentheses; CAP3 compilation within the respective stage-specific sets; total set not equal to numbers for individual stages combined due to overlap of extra sequences not assigned to stages.

†Potential fungal unigenes extracted from TaLr1 set (according to criteria outlined in Experimental procedures).

‡Number of unigenes finding a match at $E \leq 10^{-5}$ in all available databases (excluding human and mouse ESTs) are given with percentage of unigenes in parentheses.

§Excluding data for 'est_others' database to correct for self-identification of previously deposited TaLr1 ESTs.

¶Number of unigenes annotated with putative gene function at $E \leq 10^{-5}$ are given with percentage of unigenes in parentheses.

**Includes unigenes most likely wheat (based in addition on annotation results) in stage Pta: 3; Pth: 211; Pti: 79; for total set: 296 (supplementary Table S2).

Table 2 shows the result of an initial analysis to assess the overall level of similarity in three major databases. It became immediately clear that the majority of sequences did not produce significant matches at $E \leq 10^{-5}$. To address the question of whether levels of similarity were associated with specific life-cycle stages, we pooled the quality sequences in each stage and compiled unigene sets within each of them (Table 3). Using this approach, roughly one-third of the unigenes matched other sequences at $E \leq 10^{-5}$; the stage Pth produced significantly more matches

(61%) but it is likely that our conservative criteria left too many host sequences in this set and because databases contain many more cereal (plant) than fungal sequences, this would yield a biased set. A low annotation rate for the unigenes from stages Ptg and Pta suggested the presence of many novel or highly divergent genes.

Using the BLASTn algorithm, the total unigene set revealed a small set of ten members representing rRNA 'contaminants'. A further 990 unigenes having similarity matches with $E \leq 10^{-5}$ according to translated BLAST analyses were organized into 19 main functional categories (based on the Munich Information Centre for Protein Sequences' FunCat system; <http://mips.gsf.de/projects/funcat>; Mewes *et al.*, 1999), assigning each unigene to a single, most-appropriate category; class 99, 'unclassified proteins', was omitted. Supplementary Table S1 presents the annotation of 694 unigenes (7%) most likely to be of fungal origin. Overall, the ESTs represented a broad range of genes encoding proteins involved in general metabolism, energy, protein synthesis and cell fate. In particular, genes involved in cell rescue, defence, and virulence are of interest because these are thought to play a role in pathogenicity and pathogen defence against its host. Several gene sequences from leaf rust-infected wheat plants (Thara *et al.*, 2003) deposited in the public databases almost perfectly matched our *P. triticina* ESTs; some had close homologues in flax rust, *Melampsora lini* (Catanzariti *et al.*, 2006), soybean rust, *Phakopsora pachyrhizi* (Posada-Buitrago & Frederick 2005), or bean rust, *Uromyces fabae* (Jakupovic *et al.*, 2006).

Most of the translated *P. triticina* ESTs yielding significant hits (at $E \leq 10^{-20}$) showed similarity to fungal protein sequences, particularly to those from the related plant pathogenic fungus *U. maydis* and other basidiomycetes such as the *Cryptococcus* species and the higher mushrooms, *P. chrysosporium* and *C. cinereus*. Interestingly, 42 sequences matched ESTs from *Populus* species, but it is likely that some of the cDNA libraries from this species have been generated from mixed source mRNAs including fungal-infected (*Venturia tremulae*) or mycorrhiza-harboring tissues (e.g. *Amanita muscaria*; Sterky *et al.*, 2004).

Annotation results for the remaining 296 unigenes suggested that they could be derived from host tissues. Based on all available criteria, they were deemed most likely host sequences and are given in supplementary Table S2. Some entries yielded annotations from fungal species, but in these cases a sequence from wheat or a related cereal with unknown function provided a better match. The advantage of constructing and analysing cDNA libraries from 'interactive tissues' in both compatible and incompatible interactions is the discovery of host genes potentially involved in accommodating or resisting fungal development. Changes in the transcription levels of such genes upon host infection would have to be verified, which has been done by comparing a compatible and an incompatible interaction using microarrays and quantitative RT-PCR (S.C. *et al.*, unpublished

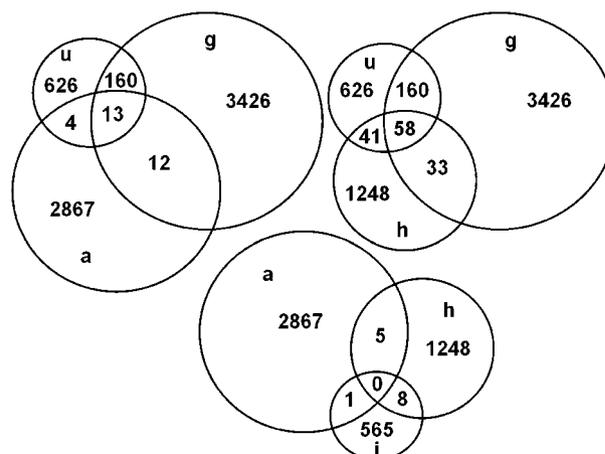


Fig. 1 Venn diagram depicting numbers of unigenes uniquely found in each stage and shared among stages. u, resting urediniospores (Ptu); g, urediniospores germinated over water (Ptg); a, appressorium stage on compatible leaf (Pta); h, haustorial stage in compatible wheat (Pth); i, incompatible interaction at 24 hpi (Pti). The sizes of the circles reflect the pool size of unigenes for each stage: u: 803, g: 3611, a: 2896, h: 1380, and i: 574.

data). Wheat genes potentially involved in disease progression as judged from their elevated occurrence in the Pth stage include a carbonic anhydrase (Contig5720), a glycine decarboxylase (Contig6034), a fumarylacetoacetate hydrolase (Contig5861), a ubiquitin-conjugating enzyme E2 (Contig5874), an aquaporin-like protein (Contig6585), a metacaspase (Contig6141), a metallothionein (Contig6001) and an α -tubulin protein (Contig5647). Of interest is a group of resistance gene-like proteins (supplementary Table S2, category 32) that might be induced in an incompatible interaction, as well as a heat shock protein (Hsp70, TaLr.Contig419). In general, we identified many wheat genes involved in energy housekeeping, photosynthesis, protein synthesis (ribosomal proteins) and protein degradation (ubiquitination), some of which have been shown to be differentially expressed upon host infection and subsequent defence initiation (B. Fofana *et al.*, unpublished data).

Abundance of ESTs and comparison between stages

The large set of unigenes allowed a search for sequences common to several stages and to identify gene sets that are possibly unique to certain stages. Figure 1 shows a Venn diagram summarizing this initial analysis. The highest number of common transcripts (160) was found between resting and germinated urediniospores, suggesting transcriptome overlap between these stages. Overall, however, little overlap was found between the various stages, suggesting the induction of rather different gene

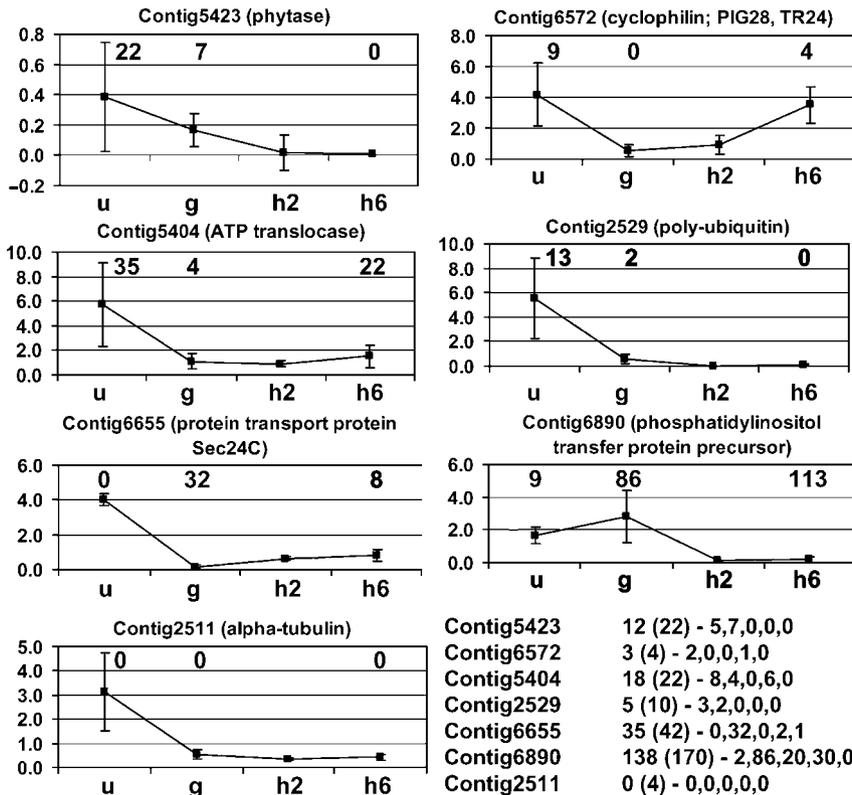


Fig. 2 Quantitative RT-PCR on selected genes to assess correlation of expression level with EST abundance in corresponding contigs. Relative values on y-axis normalized for the histone H4 gene (Contig 2841; Hu *et al.*, 2007), including standard deviation, are averaged among three independent samples for each stage, u: resting urediniospores; g: germinated urediniospores; h2 and h6: interactive haustorial stage on compatible wheat at 2 and 6 dai, respectively; note that stage h2 is not represented in the contig_info string. Numbers above the graph represent the abundance of ESTs normalized to stage 'g' and excluding ESTs from libraries PT001 to PT004 (i.e. total pool size ptu: 1836 ESTs, Ptg: 8093, Pth: 2160). Lower right panel gives the full contig_info string with the values in parentheses including ESTs from libraries PT001 to PT004 (see text and legend to supplementary Tables S1 and S2).

complements in these differentiated fungal structures. To assess which of the annotated genes were found in these various unique and overlapping stages, data from supplementary Table S1 were reanalysed to obtain supplementary Table S3. Among observed trends were a possibly increased number of genes involved in energy conversion, including six weak hits to fructose-bisphosphate aldolase, in germinated urediniospores. Also, in the haustorial stage, a large number of ribosomal protein genes was revealed, suggesting very active protein synthesis and metabolism, similar to what has been reported in earlier studies in the bean and wheat rust pathosystems (Broeker *et al.*, 2006; Jakupovic *et al.*, 2006; Thara *et al.*, 2003; Zhang *et al.*, 2003). Several transcripts for proteins involved in the biosynthesis of vitamins and some other compounds were only discovered in the haustorial stage, some of which have been identified as *in planta*-induced (PIG) genes in haustorial cDNA libraries from the bean rust system (Voegelé, 2006) and the leaf rust-wheat system (TR genes, Thara *et al.*, 2003; see Discussion). Very few fungal unigenes with significant hits specific to the Pti stage were recovered; of interest might be a P450 monooxygenase (TaLr.1170H09.F.pti, TaLr.1170H09.R.pti) possibly involved in detoxification of defence compounds, and a *U. maydis* homologue (TaLr.1174C01.R.pti), similar to Sec24C, which codes for a component in a coat protein complex responsible for COPII vesicle budding from the endoplasmic reticulum (Miller *et al.*,

2003). Sec24C was also identified with lower similarity, however, through Contig6655, which represented other stages as well and expression of which was analysed by qPCR (Fig. 2).

Within the collection of 9760 non-redundant unigenes, the overall level of sequence abundance could be assessed based on the numbers of ESTs included in the 3144 contigs. Supplementary Table S4 shows a list of the 35 most frequently occurring ESTs; many of these contigs did not yield annotations with high confidence levels nor had hits in the databases we used, indicating novel or highly diverged genes. Most ESTs in this project were produced by PCR-based approaches that were designed to ensure linear amplification of all cDNA molecules. However, there is a possibility that the abundance of ESTs may not necessarily represent the real transcript level. Because libraries PT001 to PT004 were produced from subtracted and normalized cDNAs and were therefore expected not to represent true mRNA levels, they were excluded from the stage-specific abundance analysis. Therefore, numbers for the respective stages were generated out of total pools of 1836 ESTs for Ptu, 8093 for Ptg, 5259 for Pta, 2160 for Pth and 651 for Pti. The data nevertheless allowed a preliminary look at possible gene expression levels based on transcript abundance in certain stage-specific libraries (supplementary Tables S1–3: column 'contig_info'). No statistical analysis was performed (such as described in Soanes and Talbot, 2005; Stekel *et al.*, 2000) but we assessed the correlation of the

'contig_info' with mRNA abundance in the stages similar to those from which the cDNA libraries were made, by performing semi-quantitative real-time PCR (qPCR) experiments on a select number of genes. Figure 2 illustrates that this correlation is roughly observed for several of the genes tested, suggesting that amplification of cDNAs by PCR for library construction does reflect the abundance of mRNA. However, anomalies were seen: in Contig6655, EST abundance did not seem to be correlated with the low transcript level in germinated urediniospores as measured by qPCR. Similarly, in Contig6890, no large transcript accumulation was measured by qPCR 6 days after infection. It is possible that the latter gene represents a wheat homologue of a very conserved gene given that no mixed-source sequences were apparent in the contig pile-up; however, the gene-specific primers did not produce a PCR product from wheat total genomic DNA. For both anomalies, no easy explanation can be offered but these results indicate that the presented 'contig_info strings' are mere 'indications of transcription rates'. In another example, Contig2511 (Fig. 2) consisted of four ESTs sequenced from library PT001, which was constructed from resting urediniospores using a normalizing procedure. Consequently, the 'contig_info string' contained no information yet an elevated level of mRNA was measured in this stage by qPCR correlated with its occurrence (four ESTs out of 1166 quality reads revealed in library PT001). The ratio, however, indicated that the normalization procedure might not have worked as anticipated for this library.

From the annotated contigs (supplementary Table S1) that had at least five EST members between two compared stages, we attempted to glean tentative differential gene expression values by comparing the respective EST abundance in normalized sets (supplementary Table S5). A few warrant special mention, such as a 'negative regulator of sulphur metabolism' sconCp (in particular Contig5823, among several contigs), most similar to a *Botryotinia fuckeliana* homologue. Many transcripts for this gene were found in resting urediniospores and upon germination but seemed significantly more abundant *in planta*. SconCp from *A. nidulans* was recently described to belong to the highly conserved family of *SKP1* genes essential for many cell functions, including cell cycle regulation (Piotrowska *et al.*, 2000). Two genes coding for a tropomyosin, involved in lateral binding and stabilizing actin cables (Contig103, Contig6817), appeared to be induced during spore germination, as was a gene similar to HIS4 (Contig6884) involved in histidine biosynthesis. A transcript (Contig5386) with weak similarity to infection structure protein Inf24p from *U. appendiculatus* seemed induced upon spore germination, probably in preparation for appressorium development whereas Contig2736, possibly coding for differentiation-related protein Inf8p from the same organism, seemed highly abundant in urediniospores, presumably as a result of its function during sporulation. A similar function in leaf rust urediniospore genesis and/or during its subsequent germination could also explain the

relative abundance of a homologue to conidiation-specific protein, con6p, from *N. crassa* (Contig6730). We encountered a number of ESTs with similarity to heat-shock protein family members (HSP40, 70, 78, 80 and 90) and their co-chaperones. These are important factors involved in alleviating stress on proteins (folding, protection, assisting mutated forms in attaining their function; Heitman, 2005). They were identified in several stages, sometimes in large numbers such as HSP90 (Contig2530), a homologue identified in a haustorium-specific cDNA library from the bean rust *U. viciae-fabae*. This latter HSP seemed similarly abundant in leaf rust haustoria and in germinating spores. This could indicate that the organism experiences stress during various developmental stages or suggests a common accessory, protein-stabilization role for this protein.

Some genes may have specific roles in certain stages. For example, an endopolygalacturonase possibly encoded by Contig731 and expressed in the appressorium stage is known to aid plant infections in other pathosystems by degrading host cell walls (Deising *et al.*, 1995). A phenylalanine ammonia lyase (Contig5951) seems significantly expressed in haustoria (supplementary Table S5) and, although known to be induced in plants upon pathogen attack, its function in fungi is not known but could indicate increased amino acid metabolism or a need for phenylpropanoid compounds. Very high mRNA abundance also specific to the haustorial stage was found for a metallothionein-like gene known to be expressed in organisms under oxidative and metal stress (Contig2474_part1); the same annotation was attributed to Contig23 and Contig6120, which were encountered only once in this stage. Similar results were reported for *in planta* stages for several other fungi (Broeker *et al.*, 2006; Ebbola *et al.*, 2004; Jakupovic *et al.*, 2006; Kim *et al.*, 2001; Sexton *et al.*, 2006). However, metallothioneins were also found in the resting urediniospore stage (PT00111b.F07.B7.ptu.chim1, homologue to PIG11, and PT0281.A23.C21.ptm.chim2) and in resting and germinated urediniospores as well as haustorial stages (Contig2745, homologue to PIG13).

EST variants and marker design

The comparison of constituent sequences in the contigs revealed variations among the members in many contigs. Identical changes seen in at least two independent clones in a contig, to exclude possible errors during cDNA library construction or sequencing, were analysed further. Figure 3 illustrates examples of contigs in which we observed 3'-end (Fig. 3A) and internal mRNA variation (Fig. 3B). Whether such variants originate from different members of gene families, illustrate alternative mRNA processing or represent polymorphic alleles coming from the two different parental nuclei must await analysis of the corresponding genomic sequences (Wahle and Ruedegger, 1999; Zhang and Min, 2005). The possibility of differential splicing, in particular in

et al., 2006). However, when considering the different stages we tested, the level of functional annotation varied more widely with very low numbers for the germinated urediniospore and appressorium stage libraries (Table 3). These highly differentiated and specialized stages could express very specific gene complements but since the number of overall database hits found for these stages is comparable with the other stages, it is more likely that they just express many genes of yet unknown function. It is also possible that our set contains many partial sequences that do not produce significant hits using the BLAST software.

It was revealing to see that various life-cycle stages seemed to express rather different gene complements (Fig. 1). Because our approach has probably resulted in the identification of the most prevalent mRNAs, expanding our non-comprehensive unigene set would undoubtedly reveal rarer transcripts and probably increase the occurrence of ESTs common to the various stages. Our results nevertheless suggest the occurrence of major transcriptome shifts during the profound developmental programmes the fungus undergoes during infection. Also of note was the high number of transcribed sequences obtained from resting urediniospores. Apparently, a large pre-existing pool of mRNA is available probably for essential functions during the early stages of germination. This was suggested for conidial germination in *A. nidulans* (Osherov & May, 2001). It is also possible that these mRNAs reflect (additionally) the transcriptome present at spore maturation and desiccation, therefore revealing proteins involved in these processes.

Ten unigenes matched 'glucose repressible protein GRG1' of which eight produced similar *E*-values (Contig147, Contig5538, Contig2506, Contig2539, Contig2764, Contig2793, Contig2664 and Contig5533). A more precise nucleotide comparison between these unigenes confirmed that their consensus sequences were related. Their divergence, however, caused the CAP3 program under the parameter settings used to assign them as originating from separate transcripts (data not shown). This is probably an example of a gene family whose members are possibly differentially transcribed between stages; for example, Contig2506 had 16 ESTs from stage Pt_u, six from Pt_g and three from Pt_h, which, given the total EST pool sizes in the respective stages (see legend to supplementary Table S1), seems significant. In contrast, Contig5533 harboured a single EST in stage Pt_u, three in stage Pt_g and two in stage Pt_h.

We identified two ESTs with similarity to the gEgh16 and gEgh16/MAS1 proteins from *B. graminis* (Contig5593 and Contig5424, respectively). In this fungus, both genes seem to constitute a large family potentially involved in plant interaction (Grell *et al.*, 2003). Similar to the situation in *B. graminis*, these two transcripts seemed more abundant in *P. triticina* urediniospores and seemed induced even further upon germination; this is corroborated by the finding of homologues in the soybean rust, *P. pachyrhizi* (Posada-Buitrago & Frederick, 2005).

In the interactive libraries we found several unigenes with similarity to sequences identified as *in planta*-induced genes (PIGs), in particular PIG1p (thiamine biosynthesis protein TH1p), PIG4p (thiazole biosynthetic enzyme), PIG16p (cytochrome P-450 monooxygenase), PIG17p and PIG6p in *U. fabae* (Hahn and Mendgen, 1997; Jakupovic *et al.*, 2006), and TR4 (thiamine biosynthesis protein, homologue to TH1p), TR29 (eF1 α), TR33 (microsomal signal peptidase) and TR40 (unknown) in *P. triticina* (Thara *et al.*, 2003). However, in our set, several homologues were clearly expressed in other stages as well, such as PIG11p and PIG13p (metallothioneins), PIG28p = TR24 (cyclophilin), TR21 (thiazole biosynthetic enzyme), TR14 (histone H4), TR27 (Gbeta-like protein), TR32 (mRNA polyadenylate-binding protein), TR34 (unknown), TR42 (ADP/ATP translocase-antiporter), TR51 (histone H2A) and TR56 (HSP70). These might be transcribed at a higher level *in planta* but must have roles during other developmental stages as well; for some of the homologues this has been corroborated in the bean rust (Jakupovic *et al.*, 2006) and stem rust systems (Broeker *et al.*, 2006). In this context, examination of the proteome of wheat leaves inoculated 9 days prior with *P. triticina*, identified 22 (possibly 25) consistently up-regulated fungal proteins out of 32 analysed. Six of the 22 were directly identified using the *P. triticina* EST database described in this study, one of which was similar to PIG4p from *U. fabae* (Rampitsch *et al.*, 2006). Several of the bean rust PIG genes are thought to contribute to the functionality of the haustoria and could therefore impinge on virulence, but no direct evidence for this is available (Voegelé, 2006).

Potential virulence and pathogenicity genes

Of great interest are potential virulence and pathogenicity genes, particularly those coding for products essential for early infection processes. Components of signal transduction cascades impinging on virulence have been the focus of many studies because they represent possible targets for fungicides. These include components of rather conserved pathways such as cAMP-dependent protein kinases and MAP kinases. We have identified homologues of several such components and a *P. triticina* MAPK1 (Contig2832) has been studied in detail (Hu *et al.*, 2007). Another pathway features calcineurin, an essential, serine/threonine-specific protein phosphatase that is controlled by calcium and calmodulin, acts through cyclophilin, and plays a role in morphogenesis and virulence of most fungi, including *C. albicans*, *C. neoformans*, *M. grisea* and *B. cinerea* (Viaud *et al.*, 2003). Many Ca²⁺-signalling proteins have been described in a comparative genomic analysis of *N. crassa*, *M. grisea* and *S. cerevisiae* (Zelter *et al.*, 2004). Some homologues involved in calcium binding and calmodulin-cyclophilin signalling were identified in *P. triticina* including a putative calcium transporting ATPase/translocase (Contig6835), a calcium/proton exchanger/transporter related to VCX1

(Contig2826), a calmodulin or a calmodulin-related protein (Contig2688, Contig2870, TaLr.1160A12.R.pti), a calmodulin-dependent protein kinase (Contig2859) and cyclophilin (Contig6572). Interestingly, all of these ESTs were obtained from resting urediniospore libraries except for the TaLr tag from stage Pti, which is similar to a haustorial-specific EST from *U. fabae*, and the cyclophilin tag which also appeared in stage Pth. Confirming these functions and their potential role in virulence will require further research.

Polyamines are positively charged molecules essential for cell growth and DNA replication in many organisms and interference in polyamine biosynthesis seems to reduce virulence of several phytopathogenic fungi (Bailey *et al.*, 2000; Guevara Olvera *et al.*, 1997; Mackintosh *et al.*, 1997; Rajam *et al.*, 1985). Both Odc and adenosylmethionine decarboxylase increase 50–100-fold in germinating conidia in *A. nidulans* (Stevens and Stevens, 1981) and *S. nodorum* odc1 mutants are polyamine auxotrophic and reduced in sporulation (Bailey *et al.*, 2000). Several ESTs involved in polyamine metabolism, ornithine decarboxylase Odc (Contig2661), acetylornithine aminotransferase (PT0065d.B01.BR.ptg), which catalyses the fourth step in ornithine and arginine biosynthesis, and ornithine aminotransferase CAR2 (PT0065c.G08.BR.ptg), which acts in the degradation of arginine, citrulline and ornithine, were identified. Odc seemed more abundant, particularly in the germinating urediniospore stage.

It is postulated that germinating urediniospores are sustained through fatty-acid/lipid metabolism (reviewed in Staples, 2000). Indeed, ultrastructural observations suggested that large lipid bodies in the cytoplasm of several spore types seemed to degrade during germination (Mendgen, 1984). Recent reports indicate the need for rapid increase of the acyl-CoA pool through lipolysis, i.e. peroxisomal fatty acid β -oxidation, probably involving carnitine-dependent metabolic activities, to build up the turgor necessary in the appressorium for penetration (Bhambra *et al.*, 2006; Ramos-Pamplona and Naqvi, 2006). Four unigenes identified as having acyl-coenzyme A dehydrogenase activities were generated from resting and germinating spore stages (Contig2609, Contig2901, Contig2773 and Contig5548). In addition, a homologue of mitochondrial carnitine/acylcarnitine translocase (Contig6743) seemed up-regulated upon urediniospore germination (supplementary Table S5). Lipid metabolism via the glyoxalate cycle requires malate synthase and it was recently shown that this enzyme was important for infection and vital for pathogenicity of *Stagonospora nodorum* (Solomon *et al.*, 2004). We sequenced five ESTs representing a glyoxysomal malate synthase (Contig6672) from the resting urediniospore stage indicating that this spore type is 'primed' for this metabolic activity. Moreover, Contig2729, sequenced at least three times from the same spore stage, codes for an isocitrate lyase which also has been shown to be important for spore germination and pathogenicity, possibly through peroxisomal metabolic pathways, in the blackleg fungus *Leptosphaeria*

maculans (Idnurm & Howlett, 2002) and the cucumber anthracnose fungus, *Colletotrichum lagenarium* (Asakura *et al.*, 2006); an EST having low similarity was also found in the incompatible stage (TaLr.1170H07.F.pti).

Interestingly, we identified a sequence with similarity to a member of a group of recently described 'haustorially expressed secreted proteins' (HESPs) in the flax rust, *M. lini* (Catanzariti *et al.*, 2006). Several of these 'HESPs', identified in a bioinformatics approach, are proven elicitors of the hypersensitive response in flax and co-segregate with known avirulence genes. Direct sequence comparison revealed that *P. triticina* Contig6739 is homologous to *M. lini*HESP379 (and *U. viciae-fabae* haustorium-specific cDNA clone Uf074; data not shown). It will be interesting to see whether this homologue elicits the HR in wheat and/or co-segregates with one of the many known avirulence genes in our populations.

Fungal components neutralizing host defence responses are virulence factors. Such host responses frequently involve the production of reactive oxygen species (ROS). The metabolism/detoxification of ROS is therefore essential for the establishment of the pathogen. We found a (manganese) superoxide dismutase precursor (Contig6132) in the haustorial stage. SOD was identified as a differentiation protein expressed in *U. appendiculatus* germings during early appressorium development (Lamboy *et al.*, 1995). Hydrogen peroxide as a byproduct of SOD must be detoxified by catalase, transcripts for which we discovered ESTs in the appressorial (Contig266), haustorial (Contig6007) and also germinated spore stages (PT0065d.F05.BR.ptg). Other antioxidants such as thioredoxin and glutathione can also enhance protection and detoxify xenobiotics, apart from fulfilling many other functions (Pocsi *et al.*, 2004). We found several ESTs coding for activities generating this potential, such as thioredoxin and glutathione reductase/lyase/hydrolase (glutaredoxin) and glutathione-S-transferase (Contig6033, PT00110c.D12.B7.ptu.chim1, Contig2851, PT0065b.C01.BR.ptg, PT0065b.F03.BR.ptg, PT0191.019.CPTR.ptm.chim2, PT0055a.F04.BR.ptu, Contig5460), but of these, Contig6033 was the only one uniquely found in the haustorial stage. Contig1886 (and PT0141.K04.C21.ptm) bore similarity to a β -glucosidase precursor from *U. viciae-fabae*, BGL1p, identified in a haustoria-specific library but later shown to be expressed in all early infection stages; it is involved in cellulose/cellobiose utilization but conceivably in defence because of its similarity to avenacinase with a potential phytoalexin-detoxifying activity (Haerter and Voegelé, 2004). Our unigene was discovered in germinated urediniospores and appressoria, which would not preclude a function in defence. Some fungal factors might be involved in evasion of the host response; *U. appendiculatus* changes the surface-exposed chitin moieties of its germ tubes and appressoria to chitosan using highly induced chitin deacetylases when penetrating host cells to prevent triggering immune responses (Deising

and Siegrist, 1995; Mendgen *et al.*, 1985). In *P. triticina* we identified several chitin synthases and a homologue of CDA2, a sporulation-specific chitin deacetylase and precursor (Contig6649, Contig2903) in germinated spores.

A rough estimate puts the *P. triticina* genome coverage achieved through this project at over 4 Mbp (9760 unigenes with an average size of 485.3 bp), revealing an average G+C content of 48.2 mol%. The generated sequences could therefore represent 3–4% of the *P. triticina* genome estimated to contain 100–120 Mbp. Having generated over 9000 unigenes, we are uncertain what percentage coverage of the leaf rust transcriptome these ESTs might represent. Based on the analysis of complete fungal genome sequences, *U. maydis* has an estimated 6902 genes (Kamper *et al.*, 2006) whereas another basidiomycete, *P. chrysosporium*, contains a calculated 10 048 (Wymelenberg *et al.*, 2006). Among ascomycetes, *N. crassa* was recently estimated to possess 10 082 genes (Borkovich *et al.*, 2004) and the related plant pathogen *M. grisea*, 11 109 (Dean *et al.*, 2005). It is conceivable that macrocyclic rusts possess more genes than the aforementioned fungi because of their extraordinarily complex life cycles including a sexual stage on an alternate host.

To increase functionality of our EST database, we have made use of ESTIMA, a general EST management system (Kumar *et al.*, 2004), which manages gene ontology (GO; Camon *et al.*, 2004) information. Currently, we are in the process of extending the database by generating ESTs from a haustoria-specific stage (G.B., J. Fellers), a teliospore stage isolated from senescent wheat, and two spore stages (pycnio- and aeciospores) from the alternate host (G.B., J. Xu, Y. Anikster). *P. triticina* sequences from these collaborative projects, as well as those acquired from public resources, will be merged into a comprehensive collection for analyses and the annotation of the related stem rust (*P. graminis* f. sp. *tritici*) genome. The data presented here should contribute significantly to the generation of general genomic resources for the cereal rusts. Careful analyses of these resources should allow in time for the identification of genes essential for the disease process. Such fungal genes might prove to be their 'Achilles heel(s)', the targeting of which will help protect our cereal crops.

EXPERIMENTAL PROCEDURES

Plant materials and rust strains

Wheat cultivar Thatcher, which is susceptible to *P. triticina* race1 (designation BBB; Long and Kolmer, 1989) carrying avirulence gene *Avr1*, and near-isogenic line Thatcher *Lr1*, which is resistant to this race, were used in this study. Seedlings were grown in 10-cm pots maintained at 18 °C (night) and 20 °C (day) in a leaf rust-free greenhouse with a photoperiod of 12 h. Inoculation was done on 10-day-old seedlings as described (Hu and Rijkenberg, 1998a). Materials used to prepare cDNA libraries included: (1)

prefix Ptu—resting urediniospores; (2) Ptg—urediniospores germinating *in vitro* after dusting them on the surface of sterile double-distilled water in Petri dishes (for 6 h or pooled with those after 12 h)—for one library 'wheat extract' made from sterile-filtered, surface-sterilized 2-day-old wheat germlings, ground-up in water, was added; (3) Pta—germinating urediniospores on compatible Thatcher leaf surfaces, at the stage of appressorium formation, before penetration, at 5 hpi; (4) Pth—haustorium developmental stage before sporulation, at 6 days post inoculation (dpi); and (5) Pti—incompatible interaction between Thatcher *Lr1* and race 1 at 24 hpi; see Table 1.

Preparation of mRNA and cDNA library construction

Freshly harvested resting urediniospores, germinating urediniospores or infected plant tissues were ground to a fine powder in liquid nitrogen using a mortar and pestle. Appressoria at 5 hpi were embedded by spraying with a solution of 20% cellulose acetate in acetone, which solidified within minutes. This film was quickly peeled from the epidermis and frozen in liquid nitrogen. Total RNA was extracted in a solution of phenol-guanidine isothiocyanate ('TRI Reagent', BIO/CAN Scientific, Mississauga, ON, Canada) according to instructions provided by the manufacturer. RNA was precipitated overnight at 4 °C with 1/8 vol 10 M lithium chloride to remove polysaccharides and other contaminating substances. PolyA RNA was purified using biotinylated oligo-dT/streptavidin-coated magnetic beads ('PolyA-Tract'; Promega, Madison, WI). To allow for directionally cloning and 5'- or 3'-end sequencing of the cDNA products, we designed primers incorporating rare-cutting restriction enzyme sites. *NotI* is embedded at the 3'-end and *Apal* at the 5'-end of cDNAs. First-strand cDNA synthesis was performed as described (Ausubel *et al.*, 2003) using SuperScriptII RT (Invitrogen Canada Inc., Burlington, ON, Canada) and a polyT primer, dTNotI: 5'-GCGCCGCGGCCG-TTTTTTTTTTTTTTTT (restriction enzyme site underlined). Single strand cDNA-3'-ends were extended by poly(dG)-tailing. Samples in 25 µL were incubated at 94 °C for 3 min, cooled on ice after which 5 µL 5×TdT buffer, 1 µL 40 mM dGTP, and 1 µL (15 units) terminal deoxynucleotidyl transferase (Invitrogen) was added. After incubating at 37 °C for 1 h, the sample was heat-inactivated at 65 °C for 10 min.

Due to the biotrophic nature of the leaf rust fungus, only limited amounts of mRNA could be obtained, except from infected leaf tissues. We therefore employed a PCR-based protocol for cDNA library construction from limited starting material (Froussard, 1992; Revel *et al.*, 1995). This included the use of a combination of 'full-length' dTNotI (above) and dCApal (5'-GCCCCGCG-CCCTCCCCCCCC) primers at 0.2 µM each, and 'truncated' (tr) primers, trNotI (5'-GCGCCGCGGCCGCTT) and trApa1 (5'-GCCCCGGGGCCCTCC) at 1.8 µM each. To reduce the polyA/T tract to limit subsequent non-specific hybridization during

subtractions and slippage during sequencing (Wang *et al.*, 2000), we modified this protocol to replace the dTNotI primer with a set of five 'anchored' T primers (equal amounts of oligo-dT primers: 5'-GCGCCGCGGCCGCTTTTTTTTTTTA; 5'-GCGCCGCGGCCGCTTTTTTTTTTCA; 5'-GCGCCGCGGCCGCTTTTTTTTTTTCG; 5'-GCGCCGCGGCCGCTTTTTTTTTTTTG; and 5'-GCGCCGCGGCCGCTTTTTTTTTTCC) at 0.2 μ m total. PCR reactions were performed directly on first strand cDNAs in 50 μ L containing 60 mM Tris-HCl, pH 8.5, 17 mM ammonium sulphate, 6.5 mM MgCl₂, 6.7 μ M EDTA, 0.07% β -mercaptoethanol, 0.017% BSA, dNTPs at 0.5 mM, and 5 units Taq polymerase (Invitrogen), or 'Herculase' polymerase (Stratagene), or a mixture of 1 : 10 Pfu (Stratagene)/Taq polymerase (Invitrogen). The PCR protocol was designed to ensure linear amplification representative of the transcription profile (Revel *et al.*, 1995) and consisted of a first round: preheating at 95 °C for 10 min followed by 15 cycles of 2 min at 95 °C, 50 s at 68 °C and 6 min at 72 °C. For a second round the PCR mix was divided over four new tubes, 12.5 μ L in each, and 37.5 μ L of the same, fresh 1 \times PCR mix including all primers was added. The PCR was continued for an additional 15 cycles with a final 12 min-elongation step at 72 °C.

Full-length cDNA production

We modified an existing protocol (Maruyama and Sugano, 1994) based on the protocol from the GeneRacer Kit (Invitrogen): Calf Intestinal Alkaline Phosphatase (Invitrogen) was used to dephosphorylate all molecules except mRNA molecules protected by an intact 5'-m⁷Gppp-cap structure. Tobacco Acid Pyrophosphatase (Epicentre Technologies, Madison, WI) was used for subsequent de-capping of remaining mRNAs, leaving a 5'-phosphate. To these 5'-ends, an *Apa*I-restriction site-containing RNA-oligonucleotide, 5'-GCCCCGGGCCUCC (equivalent to the *trApa*I DNA primer), was ligated using T4 RNA ligase (Epicentre). Subsequent first strand cDNA synthesis and PCR were as described above.

Normalization

For normalization, 10/11 of a particular cDNA pool was randomly labelled with biotin according to the manufacturer's protocol (Biotin-Chem-link, cat#1-812-149, Boehringer Mannheim/Roche Diagnostics, Laval, Quebec, Canada). This fraction was mixed with the unlabelled 1/11, heated to 94 °C and hybridized for 24–48 h at 68 °C. Hybrids were removed by binding to streptavidin-coated magnetic beads (Promega) in three repeated steps. The remainder of the cDNA molecules were re-amplified for 15 cycles as above.

Subtraction

Plant host cDNAs from 'interactive libraries' (that is, fungus-infected wheat leaves) were subtracted by mixing in 10 \times excess (by weight), biotinylated cDNA molecules made from mRNA from

uninfected wheat leaves at the same developmental age. This mixture was heated to 94 °C and hybridized for 48 h at 68 °C. Hybrids were removed by binding to streptavidin-coated magnetic beads (Promega) in four repeated steps. The remainder of the cDNA molecules were re-amplified for 15 cycles as above.

Standard protocol

This comprised the Stratagene protocol with modifications. After the first strand synthesis using a polyT primer with an embedded *Xho*I restriction site, cDNAs were subjected to the second strand synthesis and addition of an *Eco*RI adapter. Ligation products were digested with *Xho*I and *Eco*RI and directionally cloned into the corresponding sites of pBlueScriptIIKS (Stratagene, La Jolla, CA).

PCR products were purified over columns (QIAquick PCR Purification Kit, Qiagen, Mississauga, ON, Canada) and digested with *Apa*I and *Not*I. They were then subjected to two rounds of size selection over 1% agarose gels in 0.5 \times TBE buffer (45 mM Tris-borate, 1 mM EDTA, pH 8.0) with intermittent direct isolation from agarose slices using a QIAquick Gel Extraction Kit (Qiagen). Fractions containing fragments of 0.4–1.0 kb, 1.0–2.0 kb, and > 2.0 kb were cloned separately and directionally in dephosphorylated pBlueScriptIIKS digested with *Not*I and *Apa*I. Ligation products were introduced into *E. coli* DH5 α or DH10B/r via electroporation. *Ta*Lr1 clones were generated using mRNA obtained from 14-day-old leaf tissue from the near-isogenic line Thatcher *Lr1*, 24 hpi with *P. triticina*, race 1. cDNA was prepared according to the standard Stratagene protocol using *Eco*RI and *Xho*I adapters and was cloned directionally into vector Lambda *Zap*II in *E. coli* strain XL0LR followed by mass-excision of the clones, now in plasmid vector pBK-CMV.

Quantitative RT-PCR

Two micrograms total RNA was treated with 2 units DNaseI (Amplification Grade, Invitrogen), reverse-transcribed with 200 units SuperscriptIII (Invitrogen), and remaining RNA digested with RNaseH. The first-strand cDNA samples were diluted to 12.5 ng/mL (based on input RNA). Quantitative RT-PCR was performed on an Mx3000P qPCR System (Stratagene) with primer pairs as indicated in supplementary Table S6; the length and uniqueness of the amplicons was verified by agarose gel electrophoresis and melting curve comparison within the analysis software. Transcript levels were derived from the accumulation of SYBR green fluorescence using the following components: 20 μ L volume containing 1 \times AmpliTaq Gold II buffer plus 0.5 unit AmpliTaq Gold (PE BioSystems), 2.5 mM MgCl₂, 0.2 mM each dNTPs, 150 nmol each primer, 7.5% glycerol, 3% DMSO, 1 : 40 000 dilution of SYBR green and 1 : 30 000 dilution of ROX as reference dye (both from Stratagene), and 25 ng cDNA. PCR conditions were: 10 min at 95 °C, 40 cycles of 30 s at 95 °C, 30 s at

55 °C, and 1 min at 72 °C, followed by a melting curve programme. The Ct values for each gene were converted to amounts of cDNA present in the sample relative to a standard curve of Ct values for each gene generated by serial dilutions of known quantities of total *P. triticina* genomic DNA using the same PCR cocktail. Amounts of product were subsequently normalized to amounts of histone H4 product.

DNA sequencing

Plasmids were isolated using a QIAprep Spin Miniprep Kit according to the manufacturer's protocol (Qiagen) and at the CRC using the Eppendorf Direct Bind Kit. At the GSC, clones in *E. coli* DH10B/r were picked on a Genetix Q-Pix into 384-well plates containing 80 µL 2×YT medium with 7.5% glycerol and plasmids were isolated from 96-well plate cultures using a modified alkaline lysis method. Inserts representing the 3'-end of the mRNA were sequenced using the standard M13 forward primer (F: 5' GTAAACGACGGCCAG, or B21/C21: 5' TGTAACGACGCGCCAGT), the T7 promoter primer (B7/C7: 5' AATACGACTCACTATAG) or primer TB (5'-(T)23 A/C/G) used to overcome stuttering and slippage during sequencing due to long polyT stretches. Inserts representing the 5'-end of the mRNA were sequenced using M13 reverse primers (BR/CR: 5' CAGGA-AACAGCTATGAC, or R or CPT: 5' AACAGCTATGACCATG). The BigDye Terminator Cycle Sequence chemistry was used for all reactions (Applied Biosystems, Foster City, CA). Sequences were run on an ABI310 Genetic Analyser (PARC), an ABI377 with 96-lane upgrade (CRC), and ABI3700s (GSC; all Applied Biosystems).

Bioinformatic analysis

Raw sequences were analysed and manipulated using 'Phred' (Ewing *et al.*, 1998a,b), which resulted in the identification and removal of poor quality regions (default settings) and vector sequences. Chimeric sequences were identified using a custom Perl script that identified inserts linked head-to-head and/or tail-to-tail through the respective *NotI* and/or *Apal* restriction sites used in cloning. The constituents of a chimeric sequence were separated as individual sequences each with the original sequence name plus the suffix '.chimx'; remaining ESTs of less than 50 bp in length were discarded. The extent of redundancy for each library preparation was ascertained using the CAP3 software (Huang and Madan, 1999), which was also used to compile the unigene set (default settings: overlap length cutoff > 30, identity > 75%, gap penalty factor of 6; varying these parameters experimentally did not yield significant changes in the unigene set).

DNA sequences from resting urediniospores and urediniospores germinated *in vitro* over water (stages Pt_u and Pt_g, respectively) were presumed to be 'pure' and of fungal origin.

Sequences from 'interactive' libraries (PT002, Pta, Pth and Pti) were subtracted *in silico* for contaminating wheat sequences by BLASTn similarity searches (Altschul *et al.*, 1997) against a large compiled database containing genomic and EST sequences from wheat (see supplementary Table S7). To extract putative fungal ESTs, sequences from 'interactive' libraries were also compared with our 'pure' rust sequences (Pt_u and Pt_g) as well as with a compiled custom fungal database consisting of collected non-interactive, publicly available fungal ESTs and fungal genomic sequences from genome sequencing projects (including genomic sequences from *P. graminis* released by the Broad Institute on 12 January 2007; see supplementary Table S7). *E*-values were compared using conservative criteria: hits to other 'pure' Pt sequences were considered Pt by definition. Sequences were considered Pt if matches to fungal sequences had a lower *E*-value than those to cereal sequences; hits to cereal sequences with $E \leq 10^{-100}$ were considered wheat contaminants; hits against cereal sequences with $10^{-100} < E \leq 10^{-20}$ were rated as probably fungal if they hit fungal sequences at $E \leq 10^{-50}$; hits to both fungal and cereal sequences with $e \geq 10^{-20}$ were considered as potentially fungal. Unigenes that did not match any sequences in the databases were considered potentially fungal. A second approach exploited the differences in codon usage (Li *et al.*, 2004) in a statistical method attributing 'most likely fungal' (f) or 'most likely wheat' (w) to unknown (parts of) ORFs of interactive ESTs by comparing them to wheat and leaf rust codon usage tables (prepared from sets of correctly annotated ORFs from our database). Pure wheat and leaf rust 'training sets' were correctly assigned 60 and 97% of the time, respectively. Thirdly, the results of hybridizations of a custom prototype cDNA microarray representing an initial set of 4484 unigenes to labelled wheat or *P. triticina* total genomic DNA were also used to assign the most likely origin of the ESTs (data not shown).

Sequences were split into stage-specific sets as well as compiled into a single pool of unigenes. Similarity searches were carried out using the standalone BLAST package (NCBI; Altschul *et al.*, 1997). Unigenes were compared via BLASTX against the NCBI non-redundant proteins (NR) and a custom-made database harbouring protein sequences from annotated fungal genome projects and the *S. cerevisiae* databases (FUP). TBLASTX comparisons against the COGEME set (<http://cogeme.ex.ac.uk>, October 2005: co), a custom database compiled from other public fungal ESTs and genomic sequences (FUD), and the non-human/mouse EST db, est_others (EO; see supplementary Table S7), were also carried out. The resultant reports were parsed for desired information via custom Perl scripts and used to populate a MySQL relational database.

Those similarities maintaining *E*-values $\leq 10^{-5}$ were further investigated with respect to protein function or structural classification. This was accomplished by comparing sequences via BLASTX against the Munich Information Center for Protein

Sequences (MIPS: <http://mips.gsf.de>) annotated *U. maydis* protein database and the corresponding MIPS functional category (FunCat) designation was then used as a basis for further classification. The parsed results of the BLAST reports against the informative NR, FUP and COGEME databases were then added and visually inspected in Microsoft Excel for final categorization.

ACKNOWLEDGEMENTS

We wish to thank Dr Barry Saville for sharing *U. maydis* EST sequences prior to publication, Dr Les Szabo for DNA probes and Mr Andrew Walsh and Mr Chad Matsalla for help with bio-informatic analyses.

REFERENCES

- Abu, S.M., Li, G. and Asiegbu, F.O. (2004) Identification of Heterobasidion annosum (S-type) genes expressed during initial stages of conidiospore germination and under varying culture conditions. *FEMS Microbiol. Lett.* **233**, 205–213.
- Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucl. Acids Res.* **25**, 3389–3402.
- Asakura, M., Okuno, T. and Takano, Y. (2006) Multiple contributions of peroxisomal metabolic function to fungal pathogenicity in *Colletotrichum lagenarium*. *Appl. Environ. Microbiol.* **72**, 6345–6354.
- Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G. and Struhl, K. (2003) *Current Protocols in Molecular Biology*. New York: John Wiley & Sons.
- Backlund, J.E. and Szabo, L.J. (1993) Physical characteristics of the genome of the phytopathogenic fungus *Puccinia graminis*. *Curr. Genet.* **24**, 89–93.
- Bailey, A., Mueller, E. and Bowyer, P. (2000) Ornithine decarboxylase of *Stagonospora (Septoria) nodorum* is required for virulence toward wheat. *J. Biol. Chem.* **275**, 14242–14247.
- Barja, F., Correa, A. Jr, Staples, R.C. and Hoch, H.C. (1998) Micro-injected antisense Inf24 oligonucleotides inhibit appressorium development in *Uromyces. Mycol. Res.* **102**, 1513–1518.
- Bhairi, S.M. and Staples, R.C. (1992) Transient expression of the beta-glucuronidase gene introduced into *Uromyces appendiculatus* uredospores by particle bombardment. *Phytopathology*, **82**, 986–989.
- Bhambra, G.K., Wang, Z.-Y., Soanes, D.M., Wakley, G.E. and Talbot, N.J. (2006) Peroxisomal carnitine acetyl transferase is required for elaboration of penetration hyphae during plant infection by *Magnaporthe grisea*. *Mol. Microbiol.* **61**, 46–60.
- Borkovich, K.A., Alex, L.A., Yarden, O., Freitag, M., Turner, G.E., Read, N.D., Seiler, S., Bell-Pedersen, D., Paietta, J., Plesofsky, N., Plamann, M., Goodrich-Tanrikulu, M., Schulte, U., Mannhaupt, G., Nargang, F.E., Radford, A., Selitrennikoff, C., Galagan, J.E., Dunlap, J.C., Loros, J.J., Catcheside, D., Inoue, H., Aramayo, R., Polymenis, M., Selker, E.U., Sachs, M.S., Marzluf, G.A., Paulsen, I., Davis, R., Ebbole, D.J., Zelter, A., Kalkman, E.R., O'Rourke, R., Bowring, F., Yeoman, J., Ishii, C., Suzuki, K., Sakai, W. and Pratt, R. (2004) Lessons from the genome sequence of *Neurospora crassa*: tracing the path from genomic blueprint to multicellular organism. *Microbiol. Mol. Biol. Rev.* **68**, 1–108.
- Broeker, K., Bernard, F. and Moerschbacher, B.M. (2006) An EST library from *Puccinia graminis* f. sp. *tritici* reveals genes potentially involved in fungal differentiation. *FEMS Microbiol. Lett.* **256**, 273–281.
- Browder, L.E. (1971) Pathogenic specialization in Cereal Rust Fungi, especially *Puccinia recondita* f. sp. *tritici*: Concepts, methods of study and application. *USDA Agric. Res. Serv. Techn Bull.*, No. **1432**, 51.
- Camon, E., Magrane, M., Barrell, D., Lee, V., Dimmer, E., Maslen, J., Binns, D., Harte, N., Lopez, R. and Apweiler, R. (2004) The Gene Ontology Annotation (GOA) Database: sharing knowledge in Uniprot with Gene Ontology. *Nucl. Acids Res.* **32**, D262–D266.
- Catanzariti, A.M., Dodds, P.N., Lawrence, G.J., Ayliffe, M.A. and Ellis, J.G. (2006) Haustorially expressed secreted proteins from flax rust are highly enriched for avirulence elicitors. *Plant Cell*, **18**, 243–256.
- Collins, T.J. and Read, N.D. (1997) Appressorium induction by topographical signals in six cereal rusts. *Physiol. Mol. Plant Pathol.* **51**, 169–179.
- Dean, R.A., Talbot, N.J., Ebbole, D.J., Farman, M.L., Mitchell, T.K., Orbach, M.J., Thon, M., Kulkarni, R., Xu, J.R., Pan, H., *et al.* (2005) The genome sequence of the rice blast fungus *Magnaporthe grisea*. *Nature*, **434**, 980–986.
- Deising, H., Rauscher, M., Haug, M. and Heiler, S. (1995) Differentiation and cell wall degrading enzymes in the obligately biotrophic rust fungus *Uromyces viciae-fabae*. *Can. J. Bot.* **73**, S624–S631.
- Deising, H. and Siegrist, J. (1995) Chitin deacetylase activity of the rust *Uromyces viciae-fabae* is controlled by fungal morphogenesis. *FEMS Microbiol. Lett.* **127**, 207–211.
- Ebbole, D.J., Jin, Y., Thon, M., Pan, H., Bhattarai, E., Thomas, T. and Dean, R. (2004) Gene discovery and gene expression in the rice blast fungus, *Magnaporthe grisea*: analysis of expressed sequence tags. *Mol. Plant-Microbe Interact.* **17**, 1337–1347.
- Eilam, T., Bushnell, W.R. and Anikster, Y. (1994) Relative nuclear DNA content of rust fungi estimated by flow cytometry of propidium iodide-stained pycniospores. *Phytopathology*, **84**, 728–735.
- Ewing, B., Hillier, L., Wendl, M. and Green, P. (1998a) Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res.* **8**, 175–185.
- Ewing, B., Hillier, L., Wendl, M. and Green, P. (1998b) Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res.* **8**, 186–194.
- Froussard, P. (1992) A random-PCR method (rPCR) to construct whole cDNA library from low amounts of RNA. *Nucl. Acids Res.* **20**, 2900.
- Grell, M.N., Mouritzen, P. and Giese, H. (2003) A *Blumeria graminis* gene family encoding proteins with a C-terminal variable region with homologues in pathogenic fungi. *Gene*, **311**, 181–192.
- Guettler, S., Jackson, E.N., Lucchese, S.A., Honaas, L., Green, A., Hittinger, C.T., Tian, Y., Lilly, W.W. and Gathman, A.C. (2003) ESTs from the basidiomycete *Schizophyllum commune* grown on nitrogen-replete and nitrogen-limited media. *Fungal Genet. Biol.* **39**, 191–198.
- Guevara Olvera, L., Xoconostle Cazares, B. and Ruiz Herrera, J. (1997) Cloning and disruption of the ornithine decarboxylase gene of *Ustilago maydis*: evidence for a role of polyamines in its dimorphic transition. *Microbiology*, **143**, 2237–2245.
- Haerter, A.C. and Voegelé, R.T. (2004) A novel beta-glucosidase in *Uromyces fabae*: feast or fight? *Curr. Genet.* **45**, 96–103.
- Hahn, M. (2000) The rust fungi. Cytology, physiology and molecular biology of infection. In: *Fungal Pathology*. (Kronstad, J.W., ed.), pp. 267–306. Dordrecht: Kluwer Academic Publishers.

- Hahn, M. and Mendgen, K. (1997) Characterization of in planta-induced rust genes isolated from a haustorium-specific cDNA library. *Mol. Plant-Microbe Interact.* **10**, 427–437.
- Hahn, M., Neef, U., Struck, C., Gottfert, M. and Mendgen, K. (1997) A putative amino acid transporter is specifically expressed in haustoria of the rust fungus *Uromyces fabae*. *Mol. Plant-Microbe Interact.* **10**, 438–445.
- Heitman, J. (2005) A fungal achilles' heel. *Science*, **309**, 2175–2176.
- Horton, J.S., Bakkeren, G., Klosterman, S.J., Garcia-Pedrajas, M. and Gold, S.E. (2005) Genetics of morphogenesis in Basidiomycetes. In: *Applied Mycology and Biotechnology. Genes and Genomics* (Arora, D.K. and Berka, R., eds), pp. 353–422. Dordrecht: Elsevier.
- Hsiang, T. and Goodwin, P.H. (2003) Distinguishing plant and fungal sequences in ESTs from infected plant tissues. *J. Microbiol. Meth.* **1813**, 1–13.
- Hu, G., Kamp, A., Linning, R. and Bakkeren, G. (2007) Complementation of *Ustilago maydis* MAPK mutants by a wheat leaf rust, *Puccinia triticina* homolog; potential for functional analyses of rust genes. *Mol. Plant-Microbe Interact.* in press.
- Hu, G. and Rijkenberg, F.H.J. (1998a) Scanning electron microscopy of early infection structure formation by *Puccinia recondita* f. sp. *tritici* on and in susceptible and resistant wheat lines. *Mycol. Res.* **102**, 391–399.
- Hu, G.G. and Rijkenberg, F.H.J. (1998b) Ultrastructural studies of the intercellular hypha and haustorium of *Puccinia recondita* f. sp. *tritici*. *Phytopath. Z.* **146**, 39–50.
- Huang, X. and Madan, A. (1999) CAP3: a DNA sequence assembly program. *Genome Res.* **9**, 868–877.
- Idnurm, A. and Howlett, B.J. (2002) Isocitrate lyase is essential for pathogenicity of the fungus *Leptosphaeria maculans* to canola (*Brassica napus*). *Euk. Cell.* **1**, 719–724.
- Jakupovic, 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.
- Kamper, J., Kahmann, R., Bolker, M., Ma, L.-J., Brefort, T., Saville, B.J., Banuett, F., Kronstad, J.W., Gold, S.E., Muller, O., et al. (2006) Insights from the genome of the biotrophic fungal plant pathogen *Ustilago maydis*. *Nature*, **444**, 97–101.
- Karlsson, M., Olson, A. and Stenlid, J. (2003) Expressed sequences from the basidiomycetous tree pathogen *Heterobasidion annosum* during early infection of scots pine. *Fungal Genet. Biol.* **39**, 51–59.
- Kim, S., Ahn, I.P. and Lee, Y.H. (2001) Analysis of genes expressed during rice–*Magnaporthe grisea* interactions. *Mol. Plant-Microbe Interact.* **14**, 1340–1346.
- Kolmer, J.A. (2005) Tracking wheat rust on a continental scale. *Curr. Opin. Plant Biol.* **8**, 1–9.
- Kruger, W.M., Pritsch, C., Chao, S. and Muehlbauer, G.J. (2002) Functional and comparative bioinformatic analysis of expressed genes from wheat spikes infected with *Fusarium graminearum*. *Mol. Plant-Microbe Interact.* **15**, 445–455.
- Kumar, C.G., LeDuc, R., Gong, G., Roinishivili, L., Lewin, H.A. and Liu, L. (2004) ESTIMA, a tool for EST management in a multi-project environment. *BMC Bioinformatics*, **5**, 176–185.
- Lambooy, J.S., Staples, R.C. and Hoch, H.C. (1995) Superoxide dismutase: a differentiation protein expressed in *Uromyces* germlings during early appressorium development. *Exp. Mycol.* **19**, 284–296.
- Lee, S.H., Kim, B.G., Kim, K.J., Lee, J.S., Yun, D.W., Hahn, J.H., Kim, G.H., Lee, K.H., Suh, D.S., Kwon, S.T., Lee, C.S. and Yoo, Y.B. (2002) Comparative analysis of sequences expressed during the liquid-cultured mycelia and fruit body stages of *Pleurotus ostreatus*. *Fungal Genet. Biol.* **35**, 115–134.
- Leonard, K.J. and Szabo, L.J. (2005) Stem rust of small grains and grasses caused by *Puccinia graminis*. *Mol. Plant Pathol.* **6**, 99–111.
- Li, A., Altosaar, I., Heath, M.C. and Horgen, P.A. (1993) Transient expression of the beta-glucuronidase gene delivered into urediniospores of *Uromyces appendiculatus* by particle bombardment. *Can. J. Plant Pathol.* **15**, 1–6.
- Li, R., Rimmer, R., Buchwaldt, L., Sharpe, A.G., Seguin-Swartz, G., Couto, C. and Hegedus, D.D. (2004) Interaction of *Sclerotinia sclerotiorum* with a resistant *Brassica napus* cultivar: expressed sequence tag analysis identifies genes associated with fungal pathogenesis. *Fungal Genet. Biol.* **41**, 735–753.
- Link, T., Lohaus, G., Heiser, I., Mendgen, K., Hahn, M. and Voegelé, R.T. (2005) Characterization of a novel NADP(+)-dependent D-arabitol dehydrogenase from the plant pathogen *Uromyces fabae*. *Biochem. J.* **389**, 289–295.
- Littlefield, L.J. and Heath, M.C. (1979) Ultrastructure of rust fungi. Academic Press: New York.
- Liu, Z., Szabo, L.J. and Bushnell, W.R. (1993) Molecular cloning and analysis of abundant and stage-specific mRNAs from *Puccinia graminis*. *Mol. Plant-Microbe Interact.* **6**, 84–91.
- Loftus, B.J., Fung, E., Roncaglia, P., Rowley, D., Amedeo, P., Bruno, D., Vamathevan, J., Miranda, M., Anderson, I.J., Fraser, J.A., et al. (2005) The genome of the basidiomycetous yeast and human pathogen *Cryptococcus neoformans*. *Science*, **307**, 1321–1324.
- Long, D.L. and Kolmer, J.A. (1989) A North American system of nomenclature for *Puccinia recondita* f. sp. *tritici*. *Phytopathology*, **79**, 525–529.
- Mackintosh, C.A., Slater, L.A., McClintock, C.A., Walters, D.R., Havis, N.D. and Robins, D.J. (1997) Synthesis and antifungal activity of two novel spermidine analogues. *FEMS Microbiol. Lett.* **148**, 21–25.
- Maor, R., Kosman, E., Golobinski, R., Goodwin, P. and Sharon, A. (2003) PF-IND: probability algorithm and software for separation of plant and fungal sequences. *Curr. Genet.* **43**, 296–302.
- Marra, M., Hillier, L., Kucaba, T., Allen, M., Barstead, R., Beck, C., Blistain, A., Bonaldo, M., Bowers, Y., Bowles, L., et al. (1999) An encyclopedia of mouse genes. *Nat. Genet.* **21**, 191–194.
- Maruyama, K. and Sugano, S. (1994) Oligo-capping: a simple method to replace the cap structure of eukaryotic mRNAs with oligoribonucleotides. *Gene*, **138**, 171–174.
- McCallum, B.D. and Seto-Goh, P. (2005) Physiologic specialization of wheat leaf rust (*Puccinia triticina*) in Canada in 2002. *Can. J. Plant Pathol.* **27**, 90–91.
- Mendgen, K. (1984) Development and physiology of teliospores. In: *The Cereal Rusts* (Bushnell, W.R. and Roelfs, A.P., eds), pp. 375–398. New York: Academic Press.
- Mendgen, K. and Hahn, M. (2002) Plant infection and the establishment of fungal biotrophy. *Trends Plant Sci.* **7**, 352–356.
- Mendgen, K., Lange, M. and Bretschneider, K. (1985) Quantitative estimation of the surface carbohydrates on the infection structures of rust fungi with enzymes and lectins. *Arch. Microbiol.* **140**, 307–311.
- Mewes, H.W., Heumann, K., Kaps, A., Mayer, K., Pfeiffer, F., Stocker, S. and Frishman, D. (1999) MIPS: a database for genomes and protein sequences. *Nucl. Acids Res.* **27**, 44–48.
- Miller, E.A., Beilharz, T.H., Malkus, P.N., Lee, M.C., Hamamoto, S., Orci, L. and Schekman, R. (2003) Multiple cargo binding sites on the COPII subunit Sec24p ensure capture of diverse membrane proteins into transport vesicles. *Cell*, **114**, 497–509.

- Niks, R.E. (1983) Comparative histology of partial resistance and the nonhost reaction to leaf rust pathogens in barley and wheat seedling. *Phytopathology*, **73**, 60–64.
- Niks, R.E. and Dekens, R.G. (1991) Prehaustorial and posthaustorial resistance to wheat leaf rust in diploid wheat seedlings. *Phytopathology*, **81**, 847–851.
- Nugent, K.G., Choffe, K. and Saville, B.J. (2004) Gene expression during *Ustilago maydis* diploid filamentous growth: EST library creation and analyses. *Fungal Genet. Biol.* **41**, 349–360.
- Oshero, N. and May, G.S. (2001) The molecular mechanisms of conidial germination. *FEMS Microbiol. Lett.* **199**, 153–160.
- Ospina-Giraldo, M.D., Collopy, P.D., Romaine, C.P. and Royse, D.J. (2000) Classification of sequences expressed during the primordial and basidiome stages of the cultivated mushroom *Agaricus bisporus*. *Exp. Mycol.* **29**, 81–94.
- Piotrowska, M., Natorff, R. and Paszewski, A. (2000) SconC, a gene involved in the regulation of sulphur metabolism in *Aspergillus nidulans*, belongs to the SKP1 gene family. *Mol. Gen. Genet.* **264**, 276–282.
- Pocsi, I., Prade, R.A. and Penninckx, M.J. (2004) Glutathione, altruistic metabolite in fungi. *Adv. Microbial Physiol.* **49**, 1–76.
- Posada-Buitrago, M.L. and Frederick, R.D. (2005) Expressed sequence tag analysis of the soybean rust pathogen *Phakopsora pachyrhizi*. *Fungal Genet. Biol.* **42**, 949.
- Qutob, D., Hraber, P.T., Sobral, B.W.S. and Gijzen, M. (2000) Comparative analysis of expressed sequences in *Phytophthora sojae*. *Plant Physiol.* **123**, 243–253.
- Rajam, M.V., Weinstein, L.H. and Galston, A.W. (1985) Prevention of a plant disease by specific inhibition of fungal polyamine biosynthesis. *Proc. Natl Acad. Sci. USA*, **82**, 6874–6878.
- Ramos-Pamplona, M. and Naqvi, N.I. (2006) Host invasion during rice-blast disease requires carnitine-dependent transport of peroxisomal acetyl-CoA. *Mol. Microbiol.* **61**, 61–75.
- Rampitsch, C., Bykova, N.V., McCallum, B., Beimcik, E. and Ens, W. (2006) Analysis of the wheat and *Puccinia triticina* (leaf rust) proteomes during a susceptible host–pathogen interaction. *Proteomics*, **6**, 1897–1907.
- Rauyaree, P., Choi, W., Fang, E., Blackmon, B. and Dean, R.A. (2001) Genes expressed during early stages of rice infection with the rice blast fungus *Magnaporthe grisea*. *Mol. Plant Pathol.* **2**, 347–354.
- Revel, F., Renard, J.P. and Duranthon, V. (1995) PCR-generated cDNA libraries from reduced numbers of mouse oocytes. *Zygote*, **3**, 241–250.
- Saari, E.E. and Prescott, J.M. (1985) World distribution in relation to economic losses. In: *The Cereal Rusts, Vol. II: Diseases, Distribution, Epidemiology and Control* (Roelfs, A.P. and Bushnell, W.R., eds), pp. 259–298. Orlando, FL: Academic Press.
- Sacadura, N.T. and Saville, B.J. (2003) Gene expression and EST analyses of *Ustilago maydis* germinating teliospores. *Fungal Genet. Biol.* **40**, 47–64.
- Samborski, D.J. (1985) Wheat leaf rust. In: *The Cereal Rusts, Vol. II: Diseases, Distribution, Epidemiology and Control* (Roelfs, A.P. and Bushnell, W.R., eds), pp. 39–59. Orlando, FL: Academic Press.
- Schillberg, S., Tiburzy, R. and Fischer, R. (2000) Transient transformation of the rust fungus *Puccinia graminis* f. sp. *tritici*. *Mol. Gen. Genet.* **262**, 911–915.
- Sexton, A.C., Cozijnsen, A.J., Keniry, A., Jewell, E., Love, C.G., Batley, J., Edwards, D. and Howlett, B.J. (2006) Comparison of transcription of multiple genes at three developmental stages of the plant pathogen *Sclerotinia sclerotiorum*. *FEMS Microbiol. Lett.* **258**, 150–160.
- Skinner, W., Keon, J. and Hargreaves, J. (2001) Gene information for fungal plant pathogens from expressed sequences. *Curr. Opin. Microbiol.* **4**, 381–386.
- Soanes, D.M., Skinner, W., Keon, J., Hargreaves, J. and Talbot, N.J. (2002) Genomics of phytopathogenic fungi and the development of bioinformatic resources. *Mol. Plant-Microbe Interact.* **15**, 421–427.
- Soanes, D.M. and Talbot, N.J. (2005) A bioinformatic tool for analysis of EST transcript abundance during infection-related development by *Magnaporthe grisea*. *Mol. Plant Pathol.* **6**, 503–512.
- Solomon, P.S., Lee, R.C., Wilson, T.J. and Oliver, R.P. (2004) Pathogenicity of *Stagonospora nodorum* requires malate synthase. *Mol. Microbiol.* **53**, 1065–1073.
- Staples, R.C. (2000) Research on the rust fungi during the twentieth century. *Annu. Rev. Phytopathol.* **38**, 49–69.
- Stekel, D.J., Git, Y. and Falciani, F. (2000) The comparison of gene expression from multiple cDNA libraries. *Genome Res.* **10**, 2055–2061.
- Sterky, F., Bhalerao, R.R., Unneberg, P., Segerman, B., Nilsson, P., Brunner, A.M., Charbonnel-Campaa, L., Lindvall, J.J., Tandre, K., Strauss, S.H., *et al.* (2004) A *Populus* EST resource for plant functional genomics. *Proc. Natl Acad. Sci. USA*, **101**, 13951–13956.
- Stevens, L. and Stevens, E. (1981) Regulation of ornithine decarboxylase activity during the germination of conidia of *Aspergillus nidulans*. *FEMS Microbiol. Lett.* **11**, 229–232.
- Struck, C., Hahn, M. and Mendgen, K. (1996) Plasma membrane H⁺-ATPase activity in spores, germ tubes, and haustoria of the rust fungus *Uromyces viciae-fabae*. *Fungal Genet. Biol.* **20**, 30–35.
- Struck, C., Siebels, C., Rommel, O., Wernitz, M. and Hahn, M. (1998) The plasma membrane H⁺-ATPase from the biotrophic rust fungus *Uromyces fabae*: molecular characterization of the gene (*PMA1*) and functional expression of the enzyme in yeast. *Mol. Plant-Microbe Interact.* **11**, 458–465.
- Sugui, J.A. and Deising, H.B. (2002) Isolation of infection-specific sequence tags expressed during early stages of maize anthracnose disease development. *Mol. Plant Pathol.* **3**, 197–203.
- Thara, V.K., Fellers, J.P. and Zhou, J.M. (2003) In planta induced genes of *Puccinia triticina*. *Mol. Plant Pathol.* **4**, 51–56.
- Thomas, S.W., Rasmussen, S.W., Glaring, M.A., Rouster, J.A., Christiansen, S.K. and Oliver, R.P. (2001) Gene identification in the obligate fungal pathogen *Blumeria graminis* by expressed sequence tag analysis. *Fungal Genet. Biol.* **33**, 195–211.
- Viaud, M., Brunet-Simon, A., Brygoo, Y., Pradier, J.M. and Levis, C. (2003) Cyclophilin A and calcineurin functions investigated by gene inactivation, cyclosporin A inhibition and cDNA arrays approaches in the phytopathogenic fungus *Botrytis cinerea*. *Mol. Microbiol.* **50**, 1451–1465.
- Voegelé, R.T. (2006) *Uromyces fabae*: development, metabolism, and interactions with its host *Vicia faba*. *FEMS Microbiol. Lett.* **259**, 165–173.
- Voegelé, R.T., Hahn, M., Lohaus, G., Link, T., Heiser, I. and Mendgen, K. (2005) Possible roles for mannitol and mannitol dehydrogenase in the biotrophic plant pathogen *Uromyces fabae*. *Plant Physiol.* **137**, 190–198.
- Voegelé, R.T., Struck, C., Hahn, M. and Mendgen, K. (2001) The role of haustoria in sugar supply during infection of broad bean by the rust fungus *Uromyces fabae*. *Proc. Natl Acad. Sci. USA*, **98**, 8133–8138.

- Wahle, E. and Rueggsegger, U. (1999) 3'-End processing of pre-mRNA in eukaryotes. *FEMS Microbiol. Rev.* **23**, 277–295.
- Wang, S.M., Fears, S.C., Zhang, L., Chen, J.J. and Rowley, J.D. (2000) Screening poly (dA/dT)-cDNAs for gene identification. *Proc. Natl Acad. Sci. USA*, **97**, 4162–4167.
- Wang, X., Leclerc, C., Bakkeren, G. and McCallum, B.D. (2006) Development of simple sequence repeat markers for the genetic analysis of leaf rust [*Puccinia triticina*] in wheat. *Can. J. Plant Pathol.* **28**, 368.
- Webb, C.A., Szabo, L.J., Bakkeren, G., Garry, C., Staples, R.C., Eversmeyer, M. and Fellers, J.P. (2005) Transient expression and insertional mutagenesis of *Puccinia triticina* using biolistics. *Funct. Integr. Genomics*, **6**, 250–260.
- Wirsel, S.G.R., Voegelé, R.T., Baenninger, R. and Mendgen, K.W. (2004) Cloning of beta-tubulin and succinate dehydrogenase genes from *Uromyces fabae* and establishing selection conditions for their use in transformation. *Eur. J. Plant Pathol.* **110**, 767–777.
- Wirsel, S.G., Voegelé, R.T. and Mendgen, K.W. (2001) Differential regulation of gene expression in the obligate biotrophic interaction of *Uromyces fabae* with its host *Vicia faba*. *Mol. Plant-Microbe Interact.* **14**, 1319–1326.
- Wymelenberg, A.V., Minges, P., Sabat, G., Martinez, D., Aerts, A., Salamov, A., Grigoriev, I., Shapiro, H., Putnam, N. and Belinky, P. (2006) Computational analysis of the *Phanerochaete chrysosporium* v2.0 genome database and mass spectrometry identification of peptides in ligninolytic cultures reveal complex mixtures of secreted proteins. *Fungal Genet. Biol.* **43**, 343.
- Yarden, O., Ebbole, D.J., Freeman, S., Rodriguez, R.J. and Dickman, M.B. (2003) Fungal biology and agriculture: revisiting the field. *Mol. Plant-Microbe Interact.* **16**, 859–866.
- Zelter, A., Bencina, M., Bowman, B.J., Yarden, O. and Read, N.D. (2004) A comparative genomic analysis of the calcium signaling machinery in *Neurospora crassa*, *Magnaporthe grisea* and *Saccharomyces cerevisiae*. *Fungal Genet. Biol.* **41**, 827–841.
- Zhang, L., Meakin, H. and Dickinson, M. (2003) Isolation of genes expressed during compatible interactions between leaf rust (*Puccinia triticina*) and wheat using cDNA-AFLP. *Mol. Plant Pathol.* **4**, 469–477.
- Zhang, P. and Min, X.J. (2005) EST data mining and applications in fungal genomics. In: *Applied Mycology and Biotechnology. Genes and Genomics* (Arora, D.K. and Berka, R., eds), pp. 33–70. Dordrecht: Elsevier.

SUPPLEMENTARY MATERIAL

The following supplementary material is available for this article:

Appendix S1 Supplementary results: construction of cDNA libraries from different stages of fungal development

Table S1 *P. triticina* ESTs annotated with various levels of confidence and grouped according to functional categories.

Table S2 Pt ESTs judged most likely to be wheat host-specific, annotated with various levels of confidence and grouped according to functional categories.

Table S3 Annotated *P. triticina* ESTs encountered uniquely in certain stages.

Table S4 Most abundantly present *P. triticina* ESTs.

Table S5 Tentative differential expression of annotated genes between several combinations of two indicated stages.

Table S6 Quantitative RT-PCR oligonucleotide primers used in this study

Table S7 Public resources (with their respective Internet sites) used to generate our custom collections.

This material is available as part of the online article from:

<http://www.blackwell-synergy.com/doi/full/10.1111/j.1364-3703.2007.00406.x>

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