Microbiology (2006), **152**, 1877–1890 **DOI** 10.1099/mic.0.28750-0

Genomic analysis of host–pathogen interaction between *Fusarium graminearum* and wheat during early stages of disease development

Rubella S. Goswami,[1] Jin-Rong Xu,[2] Frances Trail,[3] Karen Hilburn[4] and H. Corby Kistler[1,4]

Correspondence
H. Corby Kistler
hckist@umn.edu

[1]Department of Plant Pathology, University of Minnesota, St Paul, MN 55108, USA

[2]Department of Botany and Plant Pathology, Purdue University, West Lafayette, IN 47907, USA

[3]Department of Plant Biology and Department of Plant Pathology, Michigan State University, East Lansing, MI 48824, USA

[4]USDA ARS Cereal Disease Laboratory, University of Minnesota, St Paul, MN 55108, USA

Received 7 December 2005
Revised 30 January 2006
Accepted 8 February 2006

Fusarium graminearum strains responsible for causing the plant disease *Fusarium* head blight vary greatly in their ability to cause disease and produce mycotoxins on wheat. With the goal of understanding fungal gene expression related to pathogenicity, three cDNA libraries were created by suppression subtractive hybridization using wheat heads inoculated with a highly aggressive strain and either water or a less aggressive strain of this pathogen. Eighty-four fungal genes expressed during initial disease development were identified. The probable functions of 49 of these genes could be inferred by bioinformatic analysis. Thirty-five ESTs had no known homologues in current databases and were not identified by *ab initio* gene prediction methods. These ESTs from infected wheat heads probably represent *F. graminearum* genes that previously were not annotated. Four genes represented in one of these libraries were selected for targeted gene replacement, leading to the characterization of a two-component response regulator homologue involved in pathogenicity of the fungus. The mutants for this gene showed reduced sporulation and delayed spread of *Fusarium* head blight on wheat.

INTRODUCTION

Fusarium head blight (FHB), a disease of small grains first described in the late 19th century (Stack, 2003), has now become a devastating disease in many parts of the world. The major causal organisms of this disease are fungal pathgens that are members of the recently described *Fusarium graminearum* (Fg) species complex (O'Donnell *et al.*, 2004). The Fg species complex comprises strains belonging to at least nine biogeographically structured cryptic species that may differ significantly in aggressiveness on wheat and in mycotoxin production (Goswami & Kistler, 2005). The species *Fusarium graminearum* (O'Donnell *et al.*, 2004) [teleomorph *Gibberella zeae* (Schweinitz) Petch] is the most common causal agent of FHB in North America and many other regions of the world. FHB commonly affects wheat, barley and other small grains in both temperate and semi-tropical areas and has the capacity to destroy a potentially high-yielding crop within a few weeks of harvest (McMullen *et al.*, 1997). Additionally, infected grains often contain significant levels of trichothecene and oestrogenic mycotoxins, making the grain unfit for food or feed (McMullen *et al.*, 1997).

The current status of genomic research on *F. graminearum* recently has been reviewed (Goswami & Kistler, 2004). Initial studies on gene expression in *F. graminearum* involved analysis of expressed sequence tag (EST) libraries to study the expression patterns in *F. graminearum* grown under different culture conditions (carbon and nitrogen starvation) and at the reproductive phase in maturing perithecia (Trail *et al.*, 2003). *In planta* gene expression was studied using a similar approach (Kruger *et al.*, 2002). In the latter study, an EST database containing 4838 sequences was created using the inflorescence of the partially resistant wheat cultivar Sumai 3 infected with *F. graminearum*. However, the focus of that study was primarily on plant

Abbreviations: dai, days after inoculation; EST, expressed sequence tag; FHB, *Fusarium* head blight; REMI, restriction-enzyme-mediated insertion.

The GenBank/EMBL/DDBJ accession numbers for the sequences determined in this study are DV998659–DV998726 and DW005257–DW005273.

Three supplementary tables are available with the online version of this paper.

genes expressed during this interaction, and only 2 % of the non-redundant sequence set was attributable to the fungus.

Functional genomic studies in *F. graminearum* have mostly targeted genes potentially involved in pathogenicity and mycotoxin production. Mutants for eight such genes have been characterized using various techniques. The genes identified and verified by gene disruption or replacement include the *Tri5* gene of the trichothecene gene cluster encoding trichodiene synthase (Proctor *et al.*, 1995), two MAP kinase genes, *Mgv1* (Hou *et al.*, 2002) and *Map1*, also called *gpmk1* (Jenczmionka *et al.*, 2003; Urban *et al.*, 2003), a secreted lipase gene, *Fgl1* (Voigt *et al.*, 2005), a gene similar to *Cps1* from *Cochliobolus heterostrophus* (Lu *et al.*, 2003) and two polyketide synthase genes, *Zea1* (*PKS4*) and *Zea2* (*PKS13*), involved in the synthesis of the mycotoxin zearalenone (Gaffoor *et al.*, 2005; Kim *et al.*, 2005). Apart from these, genes encoding a predicted NADH : ubiquinone oxidoreductase, a putative b-ZIP transcription factor, a transducin β-subunit-like protein, as well as hydroxymethylglutaryl-CoA reductase (*HMR1*) and cystathionine beta-lyase (*CBL1*) have also been identified to be involved in pathogenicity through restriction-enzyme-mediated insertion (REMI) and directed mutagenesis (Seong *et al.*, 2005, 2006).

The release of the draft genome sequence assembly of this pathogen (http://www.broad.mit.edu/annotation/fungi/fusarium/) has allowed new approaches for genomic studies. However, there is still very limited information regarding pathogen genes expressed early in the development of disease within the host. This information is vital for understanding the host–pathogen interaction and characterizing genes involved in pathogenesis. To address this issue, we chose the technique of suppression subtractive hybridization (Diatchenko *et al.*, 1996). This technique has been used successfully to identify plant and fungal genes up-regulated in various pathosystems (Beyer *et al.*, 2002; Bittner-Eddy *et al.*, 2003; Cramer & Lawrence, 2004; Guilleroux & Osbourn, 2004; Lu *et al.*, 2004; Xiong *et al.*, 2001).

The objective of this study was to establish a comprehensive method for identification and functional analysis of novel genes potentially involved in the pathogenicity of the fungus. We specifically chose to identify pathogen genes that are expressed at early stages of the infection process by *F. graminearum* on wheat; the time point of 48 h after inoculation was chosen based on previous microscopic studies on infection (Pritsch *et al.*, 2000; Skadsen & Hohn, 2004) and reports on mycotoxin production (Evans *et al.*, 2000). Overall fungal gene expression was studied by preparing a subtraction library using the susceptible wheat cultivar Norm, inoculated with a highly aggressive *F. graminearum* strain or mock-inoculated with water. Considering the large variation in pathogenicity observed among members of the Fg species complex (Goswami & Kistler, 2005; O'Donnell *et al.*, 2000), an attempt also was made to identify fungal genes differentially expressed *in planta* during interactions between wheat and a highly aggressive or less aggressive strain. Four fungal genes found to be expressed early during infection were subsequently targeted for gene deletion in order to determine their effect on pathogenicity.

METHODS

Strains and inoculum preparation. *Fusarium* strains NRRL 31084 (PH-1) and NRRL 28303 were used in this study. NRRL 31084 is the highly aggressive *F. graminearum* strain whose complete genome sequence is available. NRRL 28303 belongs to the recently described species *F. asiaticum*, formerly known as lineage 6 of the Fg species complex (O'Donnell *et al.*, 2000, 2004). The strains were stored and macroconidia prepared as previously described (Goswami & Kistler, 2005; O'Donnell *et al.*, 1998). Comparisons of spore production between wild-type and transformants were made by counting the number of macroconidia 8 days after plating on mung-bean agar. Mean numbers of spores per ml from five replicates were compared using a two-tailed *t*-test (two-sample unequal variance).

Plant growth, inoculation, disease assessment and mycotoxin analysis. Wheat plants, cultivar Norm, were grown as previously described (Goswami & Kistler, 2005). For RNA extraction, wheat spikelets, at early-to-mid anthesis, were inoculated with approximately 10^4 conidia of strain NRRL 31084 or NRRL 28303 in 10 µl 0·01 % Triton 60. Mock inoculations were conducted in a similar manner with 0·01 % Triton 60 solution alone. After inoculation, wheat plants were placed in a humidity chamber for 48 h, after which inoculated wheat heads were collected and immediately frozen at −80 °C for RNA extraction. For disease assessment, inoculum was prepared and applied using the same procedure except that only a single central spikelet within each head was inoculated. After inoculation, the plants were placed in a humidity chamber for 72 h, and then transferred to a greenhouse maintained at approximately 27 °C. The wild-type strain NRRL 31084 was used as a positive control. Wheat heads inoculated with sterile 0·01 % Triton were used as negative controls. Inoculation treatments consisted of a total of ten plants, with five plants per pot and two pots per treatment. Each strain was tested three times. Head blight was rated 14 days after inoculation (dai) by counting the number of spikelets showing disease symptoms (necrosis and/or bleaching of palea/lemma) as described previously (Goswami & Kistler, 2005). A total of ten spikelets on each wheat head were evaluated for the presence of disease symptoms by scoring five spikelets above and four spikelets below the point of inoculation. Mean number of symptomatic spikelets on inoculated plants were compared using a two-tailed *t*-test (two-sample unequal variance). Inoculated spikelets also were collected, weighed and stored for mycotoxin analysis. Mycotoxin determination was conducted as previously described (Goswami & Kistler, 2005).

Nucleic acid extraction. A revised version of the hot phenol method (Verwoerd *et al.*, 1989) was used for extraction of total RNA from infected wheat heads from each of the three treatments (inoculation with NRRL 31084, NRRL 28303 or water). All of the solutions and water used in the extraction procedure were rendered RNase free by mixing them with 0·1 % diethyl pyrocarbonate, incubating at 37 °C overnight and then autoclaving at 121 °C, 15 p.s.i. pressure for 30 min. Infected tissue (3–4 g) was ground in liquid nitrogen. A preheated 1 : 1 mixture of extraction buffer (0·1 M Tris/HCl pH 8; 0·1 M LiCl; 10 mM EDTA; 1 % SDS) and Tris-saturated phenol at 80 °C was added to the frozen powder followed by vigorous vortexing. Subsequently, half the buffer/phenol volume of chloroform was added and the mixture vortexed again. This was then centrifuged at 2500 ***g*** at room temperature for 30 min and the supernatant transferred to a 50 ml Teflon tube. One-third of the

volume of 8 M LiCl was added to the supernatant and precipitated on ice for more than 2 h. The solution was centrifuged at 12 000 ***g*** for 30 min at 4 °C. The pellet was then subjected to washes with 3 ml 2 M LiCl and 70 % ethanol at room temperature, each followed by 5 min centrifugation at 12 000 ***g***. The pellet was air-dried and dissolved in 2 ml RNase free water. Any undissolved debris was removed by another round of centrifugation, followed by precipitation using 0·1 vol. 3 M sodium acetate (pH 5·2) and 2·5 vols room temperature 100 % ethanol and incubation at −80 °C for 15 min. The clean pellet was then washed in 70 % ethanol, air-dried and dissolved in 100 µl RNase-free water. mRNA was prepared using the polyA Spin mRNA Isolation kit (New England Biolabs) according to the manufacturer's instructions. For RT-PCR, total RNA was isolated from infected wheat heads (cv. Norm) inoculated with NRRL 31084 or water at 0, 24, 48 and 72 h after inoculation using the RNeasy Plant Mini kit (Qiagen), following the manufacturer's instructions, and subjected to DNase (Roche Applied Science) treatment. DNA was extracted from lypholized mycelium of the two strains mentioned above using a protocol previously described (Rosewich *et al.*, 1999).

Library preparation. For the Fgr-S3/S4 library, suppression subtractive hybridization (SSH) was performed between the driver sequences derived from mock-inoculated wheat heads and tester sequences from wheat heads inoculated with NRRL 31084. The Fgr-S subtraction library was constructed using material from NRRL 31084-infected wheat as the tester and from NRRL 28303-infected wheat as the driver; the Fgr-S2 library was constructed using material from NRRL 28303-infected wheat as the tester and from NRRL 31084-infected wheat as the driver. The SSH procedure was performed using the PCR-Select cDNA Subtraction kit according to the manufacturer's directions (BD Biosciences-Clontech). The PCR product, enriched in differentially expressed genes from the Fgr-S3/S4 library, was cloned into the pGEM-T Easy vector (Promega) and the FgrS and Fgr-S2 libraries were cloned into pT-Adv vector using the AdvanTAge PCR Cloning kit (BD Biosciences-Clontech). After transforming DH5αMCR competent cells (Invitrogen), individual white colonies were picked into 96-well plates either manually or using a QBot robotic colony manipulator (Genetix).

DNA sequencing and bioinformatic analysis. Sequencing reactions for individual cDNA clones were performed using ABI BigDye version 3.0 chemistry (Applied Biosystems). Reaction products were ethanol-precipitated and run on an ABI3100 genetic analyser (Applied Biosystems). The CAP3 sequence assembly program based on a multiple sequence alignment method (Huang & Madan, 1999) was used to align the ESTs in each library and generate consensus sequences for contigs using the default parameters. Sequences have been deposited in the GenBank database under accession numbers DV998659–DV998726 and DW005257–DW005273.

The fungal origin of the contigs and singletons was determined based on BLASTN searches against the *F. graminearum* whole-genome sequence available at the Broad Institute website http://www.broad.mit.edu/annotation/fungi/fusarium/ (*F. graminearum* sequencing project). Sequences with *E*-values $<10^{-10}$ were considered to be derived from the pathogen unless they had better matches, according to BLASTX, with other sequences in GenBank. Comparisons also were made by conducting BLASTN searches against EST sequences from *F. graminearum* libraries created using the fungus grown on various culture media. This included EST sequences available in GenBank for libraries grown on a trichothecene induction medium, complex plant medium, cornmeal and simple substrates, and on nitrogen- and carbon-limited media (believed to provide conditions similar to those on the host). BLASTX comparisons were also made with *Magnaporthe grisea* (Dean *et al.*, 2005) and *Neurospora crassa* assembly version 3 (Galagan *et al.*, 2003), whose draft genome sequence assemblies are available. Putative functions of genes corresponding to the ESTs were ascertained by BLASTX matches to non-redundant protein sequences in GenBank (*E*-value $<10^{-5}$). In cases where the EST matched an ORF in the *F. graminearum* genome, the sequence of the entire ORF was used for the searches in GenBank, in order to better predict function of the gene. Gene predictions by the Broad Institute were primarily used for identification of ORFs corresponding to the ESTs. However, in cases where the ESTs matched regions with no predicted genes in the Broad Insititute database, predictions from the Munich Information Center for Protein Sequences *F. graminearum* genome database (Güldener *et al.*, 2006a) were considered.

Verifying fungal origin of genes and gene expression *in planta* by PCR and RT-PCR. Ten ESTs were selected for verification of their fungal origin; six ESTs (Fgr-S4_2_M02_T7, Fgr-S4_3_G23_T7, Fgr-S4_3_H08_T7, Fgr-S4_3_I05_T7, Fgr-S4_2_K04_T7 and Fgr-S4_3_G14_T7) matched the assembled *F. graminearum* genome sequence and four ESTs (Fgr-S4_1_J09_T7, Fgr-S4_1_K24_T7, Fgr-S4_1_D04_T7 and Fgr-S3_1_L11_T7) matched the excluded read files of the *Fusarium* sequencing project. Primers specific to each EST were designed using the WebPrimer available at http://seq.yeastgenome.org/cgi-bin/web-primer. Primers for a *F. graminearum* translation elongation factor 1α (EF1 and EF2; O'Donnell *et al.*, 2000) were used to generate positive controls. The primer sequences (see Supplementary Table S1, available with the online version of this paper) were used for amplification of the DNA sequences corresponding to the above-mentioned ESTs from fungal (NRRL 31084) and wheat (cv. Norm) DNA using the following cycling conditions: 94 °C for 2 min followed by 30 cycles of 94 °C for 1 min, 58 °C for 1 min, 72 °C for 1 min and a final extension for 10 min at 72 °C.

RT-PCR reactions were conducted for the genes FG08308.1 (*Abc2*), FG03405.1 (*Lyp1*), FG00215.1 (*Rrr1*) and FG02874.1 (*Zbc1*), corresponding to the ESTs Fgr-S4_3_G23, Fgr-S4_2_M02, Fgr-S4_3_I05 and Fgr-S4_3_G14 selected for targeted gene replacement. The EST/gene-specific primers mentioned above were also used for RT-PCR. The cycling parameters were as follows: one cycle of 94 °C for 2 min followed by 30 cycles of 94 °C for 1 min, 62 °C for 1 min, 68 °C for 3 min and one cycle at 68 °C for 10 min. The transformants generated after replacement of each of the genes were also initially tested using the primers in supplementary Table S1 and primers for the *hph* gene (HYG/F and HYG/R) (Supplementary Table S2). The same PCR conditions as described above were used for all the genes except the annealing temperature was lowered to 55 °C using the *hph* primers.

Identifying chromosomal positions of ESTs/contigs. The ESTs/contigs were located on the physical map by identifying the position of the contig that they matched according to BLASTN searches against the *F. graminearum* genome. Chi-squared analysis was used to analyse the distribution of these sequences on the chromosomes. The whole-genome size of 36 002 487 bp and projected size of each chromosome (Gale *et al.*, 2005) were used for calculating the expected number of genes on each chromosome.

Protoplast isolation and transformation. *F. graminearum* strain NRRL 31084 was maintained in 15 % glycerol at a spore density of 10^8 macroconidia ml^{-1} and stored at −80 °C. For protoplast preparation, 1 ml of the stock spore suspension was inoculated into 100 ml YEPD medium (yeast extract, 3·0 g; Bactopeptone, 10·0 g; glucose, 20·0 g; and distilled water to 1 l) and incubated for 12–14 h at room temperature with shaking at 175 r.p.m. Fungal protoplasts were prepared essentially as described previously (Hou *et al.*, 2002). Pure cultures were prepared by single-spore isolation from the putative transformant colonies growing on V8 juice medium and cultured on mung-bean agar (Evans *et al.*, 2000) medium containing 250 mg hygromycin l^{-1}. Macroconidia were harvested from these plates as mentioned above and used for pathogenicity tests.

Targeted gene replacement. The ligation PCR approach (Zhao *et al.*, 2004) developed for *M. grisea* was used for replacing *Abc2* and *Lyp1* in *F. graminearum* strain NRRL 31084. Primers used for each gene are presented in Supplementary Table S2. The split marker recombination procedure (Catlett *et al.*, 2003a) was used for gene replacement of *Rrr1* and *Zbc1* with modifications. The modifications included the use of the entire *hph* gene fragment amplified with the HYG/F and HYG/R primers for the fusion PCR step with the flanking regions using primers HY/R and YG/F located inside the *hph* gene along with the 1F and 2R primers for this step. Changes also were made to the HYG/F and HYG/R primers to fit regions flanking the gene in the plasmid pCX62 from which we amplified the gene (Supplementary Table S2). An annealing temperature of 60 °C was used for the first PCR cycle. An extra purification step of the first PCR product was added using a QIAquick PCR Purification kit (Qiagen). The cycling conditions for the second round of PCR were also modified from the original protocol and involved one cycle of 94 °C for 2 min followed by 30 cycles of 94 °C for 30 s, 55 °C for 30 s, 68 °C for 2·5 min and one cycle at 68 °C for 10 min.

Southern hybridization. Samples (5 μg) of DNA from each of the mutants and the wild-type were digested with restriction enzymes as follows: *Abc2* mutants, *Hin*dIII and *Afl*II; *Lyp1* mutants, *Hpa*I and *Sac*II; *Rrr1* mutants, *Hin*dIII; and *Zbc1* mutants, *Blp*I and *Spe*I. Southern hybridization for the *Abc2* mutants and wild-type was performed as described by Rosewich *et al.* (1999) except that the Decaprime DNA II Random Priming DNA Labelling kit from Ambion was used for labelling the probes. The AlkPhos Direct Labelling and Detection System with CDP-*Star* from Amersham Bioscience was used for Southern hybridization using the mutants for the other three genes, following the manufacturer's instructions with minor modification. Gene-specific fragments amplified from NRRL 31084 genomic DNA with the primers in Supplementary Table S2 were used as probes to ascertain replacement of the gene and identify those mutants where the *hph* gene had integrated at locations other than the target gene in the genome (ectopic mutants). The *hph* gene fragment used in the split marker protocol was used as a probe to verify integration of the *hph* gene.

RESULTS

Sequence analyses

Three cDNA libraries were constructed after suppression subtractive hybridization. The first library, designated Fgr-S3/S4, was created to characterize overall fungal gene expression during initial disease development. In this case RNA from wheat heads inoculated with the highly aggressive strain NRRL 31084 was used as the tester and RNA from mock-inoculated wheat heads was used as the driver. Five thousand white colonies were isolated after cloning the subtracted cDNA in pGEM-T, and an initial 1236 clones were sequenced. The mean size for the sequenced ESTs from this library was 0·53 kb. About 28 % of the ESTs from this library (348 clones) appeared to be of fungal origin, as determined by their BLASTN matches to the *F. graminearum* genome sequence. Of these, 145 ESTs matched the assembled genome sequence and could be assigned positions within the assembly. The remaining 203 EST sequences matched the excluded/unassembled reads for the genome sequence, which largely account for organellar and other low-complexity or repeated DNA of the pathogen. Since our focus was primarily on pathogen gene expression, further analysis was conducted using these 348 EST sequences. BLAST searches against the non-redundant nucleotide and protein databases in GenBank were conducted to predict the function of the corresponding genes. Fifty-eight per cent of the ESTs (204 total) corresponded to mitochondrial genes while 10 % (34 ESTs) corresponded to ribosomal proteins based on similarity to sequences in GenBank. The sequences with no significant BLASTN or BLASTX matches in GenBank accounted for 24 % (83) of the ESTs, while those sequences that had homologues in GenBank that were not mitochondrial or ribosomal genes represented 8 % or 28 ESTs.

The second and third subtraction libraries, named Fgr-S and Fgr-S2 were generated by using cDNA synthesized with RNA from wheat heads inoculated with a highly aggressive strain NRRL 31084 or RNA derived from wheat inoculated with the less aggressive strain NRRL 28303. The Fgr-S subtraction library was constructed using RNA from NRRL 31084-infected wheat as the tester and RNA from NRRL 28303-infected wheat as the driver; the Fgr-S2 library was constructed using RNA from NRRL 28303-infected wheat as the tester and RNA from NRRL 31084-infected wheat as the driver. The relative aggressiveness of strains had been determined by our initial studies (Goswami & Kistler, 2005). A total of 1361 clones from these libraries were sequenced: 588 from the Fgr-S library and 773 from the Fgr-S2 library. The mean EST size was 0·35 kb. Marked differences in genes represented in Fgr-S and Fgr-S2 were observed, as only 6·4 % of the ESTs were found in both the forward and reverse subtraction libraries. However, the number of fungal sequences found in each library was quite low. Only seven ESTs of fungal origin were identified from these libraries: three ESTs from the forward subtraction library and four from the reverse subtraction library.

Distribution of ESTs into contigs and singlets and their comparisons to ESTs from *F. graminearum* in culture and to the *M. grisea* and *N. crassa* genomes

The CAP3 sequence assembly program was used to group together redundant ESTs from each library (Fgr-S3/S4, Fgr-S and Fgr-S2) which had overlapping sequences. A consensus sequence was obtained for each contig, and every EST in the contig was considered to be a copy of the transcript from the same gene. The 1236 ESTs from the Fgr-S3/S4 library could be assembled into 182 contigs and 630 singlets. Of these, 22 contigs and 56 singlets matched the whole-genome assembly, and 45 contigs and 30 singlets matched the excluded reads of the *F. graminearum* genome sequence. Remaining contigs and singlets are presumed to correspond to wheat genes. The 78 contigs and singlets matching the *Fusarium* genome assembly were chosen for further analysis. In addition to searching against GenBank, these EST contigs and singlets were compared to the predicted proteins from the rice blast fungus *M. grisea* and the saprophyte *N. crassa*. Twenty contigs and 38 singlets that had no significant homologue in either of the two genomes initially were considered to be specific to *F. graminearum*. However, among

these, 12 contigs and 13 singlets matched predicted proteins from other organisms (Table 1). The EST contigs and singlets that also did not have any matches in GenBank are presumed to be specific to *F. graminearum* (Table 2). Those ESTs and contigs that had matches to proteins coded by *M. grisea* and *N. crassa* genomes are listed in Table 3. Comparisons also were made to EST sequences available in GenBank for libraries grown on trichothecene induction medium (Tag *et al.*, 2001), complex plant medium, cornmeal and simple substrates, and on nitrogen- and carbon-limited media. The ESTs/contigs that were identified that match EST sequences from the fungus grown in culture also are noted in Tables 1, 2 and 3.

ESTs from the Fgr-S and S2 libraries were assembled into 238 contigs and 271 singlets, of which only 1 contig and 5 singlets matched the *F. graminearum* assembled reads; no matches were found to the *F. graminearum* excluded reads. Of these, the sole contig and three of the singlets had BLASTX matches with both *M. grisea* and *N. crassa* genomes, whereas the other two, namely Fgr-S_8_D09 and Fgr-S_7_C07_T7, appeared to be specific to *F. graminearum* (Table 4).

Distribution of ESTs/contigs in the *F. graminearum* genome and identification of new ORFs

To determine whether *F. graminearum* genes expressed during infection were preferentially located on particular chromosomes, genes corresponding to ESTs and contigs from the Fgr-S3/S4, Fgr-S and Fgr-S2 libraries were located within the *F. graminearum* genome. The observed number of genes located on each chromosome was consistent with the distribution of genes that would be expected based solely on the size of the chromosome as a proportion of the genome as confirmed by chi-squared analysis (Table 5).

Thirty five ESTs/contigs were found to match regions of the genome that were not predicted to be genes by automated annotation methods employed by the Broad Institute, the *F. graminearum* Sequencing Project or the MIPS (Munich Information Center for Protein Sequences) *F. graminearum* database. Among these ESTs and contigs, nine match regions on scaffolds 23 and 42, now known to be part of the mitochondrial genome (J. C. Kennell, personal communication), and another eight match regions on scaffold 32, a sequence that most closely matches portions of the mitochondrial plasmid pCRY1 from the fungus *Cryophonectria parasitica*. Thus, errors in gene prediction on these scaffolds occurred because of the use of the standard genetic code rather than the mould mitochondrial genetic code for these regions. Nevertheless, 20 ESTs match regions in the assembled genome that have been assigned chromosomal locations and represent previously undescribed transcripts. One EST, Fgr-S4_3_M04, contains a 215 bp sequence found to be repeated 29 times in the *F. graminearum* genome. Interestingly, while the repeat contains no ORF encoding a potential protein of >60 amino acids, the core repeat region corresponds to the LTR of the retrotransposon Skippy from the related fungus *F. oxysporum*. Copies of these repeats are preferentially located near telomeres.

Gene replacement

Genes corresponding to four ESTs (Fgr-S4_3_G23, Fgr-S4_2_M02, Fgr-S4_3_I05 and Fgr-S4_3_G14) were selected for targeted gene replacement. These genes, FG08308.1 (*Abc2* – ATP binding cassette transporter 2), FG03405.1 (*Lyp1* – Lysine permease 1), FG00215.1 (*Rrr1* – Receiver response regulator 1) and FG02874.1 (*Zbc1* – Zinc finger binuclear cluster regulator 1), had BLASTX matches to an ABC transporter, a lysine permease, a putative two-component response regulator and the *alcR* regulatory gene respectively. Two genes, *Abc2* and *Rrr1*, had BLASTX matches with predicted proteins from both *M. grisea* and *N. crassa*. Based on BLASTX comparisons, the gene *Zbc1* appeared to encode a protein related to a zinc binuclear cluster regulatory protein similar to *alcR* from *Aspergillus nidulans* with sequence similarity to the pathogen *M. grisea*, but not to the non-pathogen *N. crassa*. The *Lyp1* gene, homologue of a lysine permease gene from *Saccharomyces cerevisiae*, did not have matches to either *M. grisea* or *N. crassa*. Based on these matches and their expression *in planta* it was suspected that these genes could be involved in pathogenicity or development of the fungus on the host; they were therefore selected for targeted gene replacement studies to test their influence on pathogenicity.

To confirm that these genes were indeed expressed *in planta*, RT-PCR was conducted using RNA collected from NRRL 31084 inoculated and mock (water)-inoculated wheat heads. The amplified products demonstrated that the four selected genes were expressed *in planta* on wheat heads inoculated with NRRL 31084 only. The expression of these genes was detected 24 h after inoculation and was observed at both time points, 48 and 72 h, thereafter. These results were consistent among all four genes and the gene for translation elongation factor 1α (EF-1α), used as a positive control (Fig. 1).

Gene replacement was successful for all four genes, as confirmed by PCR and Southern hybridization. For example, the 2·8, 1·0 and 0·4 kb *Hin*dIII fragments that hybridize to the probe for *Rrr1* in NRRL31084 (Fig. 2) are missing in four of the five transformants (labelled T1, T2, T4 and T5), indicating that for these strains, the *Rrr1* gene has been deleted. Strain T3 maintains the 1·0 and 0·4 kb internal fragments of the *Rrr1* gene, although it appears to be altered at the 2·8 kb flanking region. The 4·8 kb *Hin*dIII fragment corresponding to the intact gene for hygromycin resistance is found in all transformants but additional copies of the gene are also found in T3 and T4. A summary of the numbers of hygromycin-resistant colonies, deletion mutants and integration events not resulting in deletion for each of the four targeted genes is provided in Supplementary Table S3, available with the online version of this paper. Five strains with mutations in each of the four genes studied were tested for pathogenicity. For the genes *Lyp1*, *Rrr1* and *Zbc1*,

Table 1. ESTs/contigs from the Fgr-S3/S4 libraries without significant homology to *M. grisea* or *N. crassa* genomes but with homology to other sequences in GenBank

BLASTN against the *F. graminearum* assembled genome sequence					BLASTX against GenBank protein database		
EST/Contig ID*	*E*-value	Size (bp)	Contig ID	Gene ID	BLASTX hit	Organism	*E*-value
Fgr-S3_1_F21_T7	1·37E−15	637	1.417 (scaffold 7)	FG10081.1	COG1835: Predicted acyltransferases	*Methanosarcina barkeri* str. fusaro	8·00E−15
†Fgr-S3_1_M01_T7	3·33E−29	218	1.285 (scaffold 4)	FG06924.1	60S ribosomal protein L10A	*Chaetomium globosum*	1·00E−85
†Fgr-S3_1_O03_T7	2·23E−85	184	1.406 (scaffold 6)	FG09845.1	Delta 8-(E)-sphingolipid desaturase	*Pichia pastoris*	e−135
‡Fgr-S4_1_A19_T7	5·10E−77	186	1.488 (scaffold 23)	No predicted gene	NADH-ubiquinone oxidoreductase chain 4	*Hypocrea jecorina*	3·00E−19
‡Fgr-S4_1_G04_T7	7·02E−139	286	1.356 (scaffold 5)	No predicted gene	Mitochondrial I-PcI endonuclease	*Podospora curvicolla*	2·00E−15
†Fgr-S4_1_H04_T7	1·08E−50	807	1.23 (scaffold 1)	FG00560.1	Hypoxia-induced hypothetical protein 162	*Lycopersicon esculentum*	7·00E−14
Fgr-S4_1_H14_T7	0	811	1.194 (scaffold 2)	FG04651.1	Putative tetratricopeptide repeat (TPR)-containing protein	*Solanum demissum*	4·00E−11
Fgr-S4_2_J18_T7	0	558	1.456 (scaffold 8)	FG10943.1	Cell wall surface anchor family protein	*Streptococcus pneumoniae*	2·00E−15
Fgr-S4_2_M02_T7	0	690	1.154 (scaffold 2)	FG03405.1	Lysine permease	*Saccharomyces cerevisiae*	e−112
Fgr-S4_2_N13_T7	0	787	1.142 (scaffold 2)	FG02879.1	Hypothetical protein MTH632	*Methanothermobacter thermautotrophicus* str. Delta H	7·00E−60
Fgr-S4_2_O11_T7	3·79E−142	571	1.207 (scaffold 3)	FG05148.1	Putative amino acid transporter	*Schizosaccharomyces pombe*	e−123
‡Fgr-S4_3_B18_T7	4·54E−94	322	1.510 (scaffold 42)	No predicted gene	Mitochondrial protein Dod COI i8 grp ID	*Podospora anserina*	3·00E−06
Fgr-S4_3_H08_T7	1·42E−177	352	1.168 (scaffold 2)	FG03964.1	Type I polyketide synthase	*Botryotinia fuckeliana*	e−156
‡FgrS3S4Contig42 (2)	E−166	580	1.488 (scaffold 23)	No predicted gene	NADH dehydrogenase subunit 4	*Lecanicillium muscarium*	4·00E−55
FgrS3S4Contig45 (5)	0	832	1.498 (scaffold 32)	fgd498-30	Reverse transcriptase homologue	Barley	2·00E−28
FgrS3S4Contig66 (5)	0	877	1.489 (scaffold 23)	fgd489-10	COX1-I1	*Penicillium marneffei*	4·00E−12
‡FgrS3S4Contig79 (2)	0	604	1.498 (scaffold 32)	No predicted gene	COG3157: Hemolysin-coregulated protein (uncharacterized)	*Burkholderia fungorum* LB400	1·00E−17
FgrS3S4Contig100 (4)	E−154	833	1.70 (scaffold 1)	FG01380.1	Cytochrome *c* oxidase polypeptide I	*Hypocrea jecorina*	4·00E−18
‡FgrS3S4Contig120 (3)	0	688	1.498 (scaffold 32)	fgd498-40	DNA polymerase	*Cryphonectria parasitica*	4·00E−05
FgrS3S4Contig130 (4)	0	569	1.510 (scaffold 42)	FG11639.1	Cytochrome oxidase I intronic ORF 9	*Podospora anserina*	5·00E−54
FgrS3S4Contig141 (3)	0	776	1.498 (scaffold 32)	fgd498-30	I-PcI endonuclease	*Podospora curvicolla*	1·00E−15
FgrS3S4Contig152 (2)	0	980	1.510 (scaffold 42)	FG11639.1	Cytochrome oxidase I intronic ORF 9	*Podospora anserina*	2·00E−86
‡FgrS3S4Contig153 (2)	0	586	1.489 (scaffold 23)	No predicted gene	Cytochrome oxidase subunit III	*Trichoderma pseudokoningii*	5·00E−11
FgrS3S4Contig159 (2)	0	457	1.510 (scaffold 42)	fgd510-20 conserved hypothetical protein	Cytochrome oxidase subunits I-I3	*Penicillium marneffei*	4·00E−17
‡FgrS3S4Contig178 (2)	2·00E−75	379	1.488 (scaffold 23)	No predicted gene	NADH-ubiquinone oxidoreductase chain 4	*Hypocrea jecorina*	4·00E−39

*The numbers of ESTs in each contig are shown in parentheses after each contig ID.

†ESTs/contigs that match sequences from *F. graminearum* EST libraries in culture (Trail *et al.*, 2003).

‡ESTs/contigs that are considered to be new ORFs.

Table 2. ESTs/contigs unique to *F. graminearum* from the Fgr-S3/S4 libraries

BLASTN against the *F. graminearum* assembled genome sequence				
EST/Contig ID*	***E*-value**	**Size (bp)**	**Contig ID**	**Gene ID**
†Fgr-S3_1_B06_T7	4·93E−13	247	1.185 (scaffold 2)	fgd185-180
‡Fgr-S3_1_H05_T7	5·97E−54	436	1.452 (scaffold 8)	No predicted gene
‡Fgr-S3_1_H12_T7	1·50E−26	390	1.190 (scaffold 2)	No predicted gene
†‡Fgr-S3_1_J15_T7	2·47E−111	352	1.266 (scaffold 4)	No predicted gene
‡Fgr-S4_1_A11_T7	0	567	1.141 (scaffold 2)	No predicted gene
‡Fgr-S4_1_B06_T7	2·52E−56	125	1.26 (scaffold 1)	No predicted gene
‡Fgr-S4_1_D06_T7	1·28E−171	343	1.181 (scaffold 2)	No predicted gene
‡Fgr-S4_1_E01_T7	0	674	1.10 (scaffold 1)	No predicted gene
‡Fgr-S4_1_F16_T7	1·00E−27	407	1.364 (scaffold 6)	No predicted gene
‡Fgr-S4_1_H10_T7	0	499	1.185 (scaffold 2)	No predicted gene
‡Fgr-S4_1_O16_T7	0	803	1.320 (scaffold 4)	No predicted gene
‡Fgr-S4_1_P19_T7	0	610	1.465 (scaffold 9)	No predicted gene
Fgr-S4_2_B16_T7	0	596	1.459 (scaffold 8)	FG11116.1
‡Fgr-S4_2_F02_T7	0	621	1.30 (scaffold 1)	No predicted gene
‡Fgr-S4_2_F10_T7	0	635	1.242 (scaffold 3)	No predicted gene
‡Fgr-S4_2_G20_T7	0	428	1.395 (scaffold 6)	No predicted gene
†Fgr-S4_2_G24_T7	0	818	1.77 (scaffold 1)	No predicted gene
‡Fgr-S4_2_H08_T7	6·69E−163	770	1.354 (scaffold 5)	No predicted gene
Fgr-S4_2_L24_T7	0	696	1.357 (scaffold 5)	FG08818.1
Fgr-S4_2_N24_T7	0	713	1.208 (scaffold 3)	FG05178.1
‡Fgr-S4_2_O13_T7	8·63E−83	283	1.488 (scaffold 23)	No predicted gene
‡Fgr-S4_3_H19_T7	2·77E−123	341	1.498 (scaffold 32)	fgd498-10
‡Fgr-S4_3_I03_T7	0	546	1.498 (scaffold 32)	fgd498-40
†‡Fgr-S4_3_L01_T7	0	432	1.424 (scaffold 7)	No predicted gene
‡Fgr-S4_3_M04_T7	0	551	1.323 (scaffold 5)	No predicted gene
†‡FgrS3S4Contig12 (3)	0	455	1.488 (scaffold 23)	No predicted gene
‡FgrS3S4Contig15 (7)	E−174	368	1.488 (scaffold 23)	No predicted gene
‡FgrS3S4Contig20 (4)	E−112	522	1.488 (scaffold 23)	No predicted gene
‡FgrS3S4Contig74 (3)	3·00E−23	878	1.173 (scaffold 2)	No predicted gene
†FgrS3S4Contig81 (11)	0	737	1.498 (scaffold 32)	No predicted gene
FgrS3S4Contig83 (3)	2·00E−36	588	1.488 (scaffold 23)	No predicted gene
†FgrS3S4Contig86 (3)	0	1144	1.498 (scaffold 32)	fgd498-10
FgrS3S4Contig145 (2)	0	645	1.467 (scaffold 9)	FG11410.1

*The numbers of ESTs in each contig are shown in parentheses after each contig ID.
†ESTs/contigs that match sequences from *F. graminearum* EST libraries in culture (Trail *et al.*, 2003).
‡ESTs/contigs that are considered to be new ORFs.

pathogenicity tests involved at least one successful gene deletion mutant and one transformant for which the targeted gene was not deleted (called ectopic mutants). Ectopic mutants were not obtained for the *Abc2* gene and so could not be included in the pathogenicity tests. All four *Rrr1* deletion strains exhibited reduced sporulation ($P<0·05$) as demonstrated by spore counts taken 8 dai (Fig. 3a) and also produced fewer symptomatic spikelets per wheat head ($P<0·05$), 7 dai (Fig. 3b), thereby indicating a delay in spread of the fungus. Nevertheless, the percentage of symptomatic spikelets observed 14 dai with *Rrr1* mutants was found not to be significantly different from the wild-type. The ectopic mutant *Rrr1*T3 was similar to the wild-type in both spore production ($P=0·92$) and number of symptomatic spikelets ($P=0·63$). Mutants of each of the other three genes did not differ from wild-type with respect to pathogenicity or sporulation under the conditions tested (data not shown).

DISCUSSION

In this study we investigated fungal gene expression during the initial stages of FHB disease development, 48 h after inoculation. Previously it has been shown that spores of *F. graminearum* used to inoculate wheat heads germinate and produce hyphae within 24 h of inoculation; however growth around stomatal rows and in subcuticular locations was not visible until 48–76 h after inoculation (Pritsch *et al.*,

Table 3. ESTs/contigs from the Fgr-S3/S4 libraries with homology to proteins predicted from the *M. grisea* and *N. crassa* genomes

BLASTN against the *F. graminearum* assembled genome sequence					BLASTX against GenBank protein database		
EST/Contig ID*	*E*-value	Size (bp)	Contig ID	Gene ID	BLASTX hit	Organism	*E*-value
F. graminearum* ESTs/contigs that have BLASTX matches to both *M. grisea* and *N. crassa							
†§Fgr-S3_1_B19_T7 and †§Fgr-S3_1_G01_T7	9·03E−13 and 2·21E−17	440 and 172	1.456 (scaffold 8)	FG10941.1	Ubiquitin/S27a fusion protein [imported]	*Neurospora crassa*	3·00E−58
Fgr-S3_1_H07_T7	6·72E−146	390	1.103 (scaffold 1)	FG01920.1	Minichromosome maintenance protein 3 homologue	*Schizosaccharomyces pombe*	0
Fgr-S4_1_I14_T7	0	538	1.308 (scaffold 4)	FG07369.1	MAPK cascade scaffold protein	*Mus musculus*	1·00E−49
Fgr-S4_1_P23_T7	0	650	1.37 (scaffold 1)	FG00862.1	Raffinose synthase family protein/seed imbibition protein	*Arabidopsis thaliana*	4·00E−38
Fgr-S4_2_D04_T7	0	607	1.190 (scaffold 2)	FG04421.1	SET domain-containing protein	*Arabidopsis thaliana*	2·00E−10
Fgr-S4_2_I13_T7	0	748	1.169 (scaffold 2)	FG04038.1	ATP(GTP)-binding protein	*Homo sapiens*	2·00E−80
Fgr-S4_2_O15_T7	0	754	1.60 (scaffold 1)	FG01211.1	Hypothetical PWWP domain coiled-coil protein	*Schizosaccharomyces pombe*	2·00E−07
Fgr-S4_3_B04_T7	0	438	1.452 (scaffold 8)	FG10884.1	Fructosamine-3-kinase-related protein	*Homo sapiens*	5·00E−24
Fgr-S4_3_E20_T7	0	752	1.56 (scaffold 1)	FG01118.1	Ubiquitin-like protein	*Lycopersicon esculentum*	2·00E−10
Fgr-S4_3_G23_T7	8·75E−176	351	1.330 (scaffold 5)	FG08308.1	Fum19p (ABC transporter)	*Gibberella moniliformis*	0
Fgr-S4_3_I05_T7	0	722	1.8 (scaffold 1)	FG00215.1	Putative two-component response regulator REC1p	*Gibberella moniliformis*	6·00E−38
Fgr-S4_3_L04_T7	0	699	1.169 (scaffold 2)	FG04042.1	Potential transmembrane protein	*Candida albicans* SC5314	9·00E−28
Fgr-S4_2_K04_T7	0	710	1.328 (scaffold 5)	FG08170.1	Related to D-amino acid oxidase	*Neurospora crassa*	9·00E−52
F. graminearum* ESTs/contigs that have BLASTX matches to only *M. grisea							
Fgr-S4_2_H21_T7	0	402	1.330 (scaffold 5)	FG08284.1	Predicted protein	*Magnaporthe grisea*	2·00E−85
Fgr-S4_3_G06_T7	0	688	1.50 (scaffold 1)	FG01033.1	Secretory component	*Saccharomyces cerevisiae*	1·00E−05
Fgr-S4_3_G14_T7	0	388	1.142 (scaffold 2)	FG02874.1	ALCR_EMENI regulatory protein AlcR	*Aspergillus nidulans* FGSC A4	5·00E−71
†‡FgrS3S4Contig13 (5)	4·00E−43	584	1.270 (scaffold 4)	No predicted gene	Predicted protein	*Magnaporthe grisea*	6·00E−24
Fgr-S4_3_D05_T7	0	796	1.166 (scaffold 2)	FG03829.1	Hypothetical protein MG03363.4	*Magnaporthe grisea*	6·00E−12
F. graminearum* ESTs/contigs that have BLASTX matches to only *N. crassa							
FgrS3S4Contig133 (2)	0	439	1.11 (scaffold 1)	FG00318.1	Related to DRPLA protein [imported]	*Neurospora crassa*	5·00E−23
Fgr-S4_1_O04_T7	0	510	1.16 (scaffold 1)	FG00416.1	Probable nitrate transport protein CrnA	*Neurospora crassa*	7·00E−73

*The numbers of ESTs in each contig are shown in parentheses after each contig ID.
†ESTs/contigs that match sequences from *F. graminearum*.
‡EST libraries in culture (Trail *et al.*, 2003).
§ESTs/contigs that are considered to be new ORFs.

Table 4. ESTs/contigs from the Fgr-S and Fgr-S2 libraries that matched the *F. graminearum* genome sequence

BLASTN against the *F. graminearum* assembled genome sequence					BLASTX against GenBank protein database		
EST/Contig ID	*E*-value	Size (bp)	Contig ID	Gene ID	BLASTX hit	Organism	*E*-value
Fgr-S_4_G12_T7	0	495	1.283 (scaffold 4)	FG06893.1	Protein component of the small (40S) ribosomal subunit	*Saccharomyces cerevisiae*	1·00E−48
Fgr-S_7_C07_T7	1·13E−115	247	1.224 (scaffold 3)	FG05454.1	Pyruvate dehydrogenase E1 component alpha subunit	*Pichia stipitis*	E−124
Fgr-S_8_D09_T7	0	568	1.421 (scaffold 7)	FG10135.1	Mixed-linked glucanase precursor (MLG1) related protein [imported]	*Neurospora crassa*	1·00E−06
Fgr-S2_3_D01_T7	3·52E−32	222	1.29 (scaffold 1)	FG00639.1	Tubulin alpha chain	*Neurospora crassa*	0
Fgr-S2_6_B03_T7	1·24E−29	313	1.500 (scaffold 34)	FG11626.1	Histone H2B	*Rosellinia necatrix*	2·00E−35
FgrSS2Contig 224 (3)	8·00E−48	624	1.354 (scaffold 5)	FG08768.1	Polyubiquitin/ribosomal protein CEP52	*Trypanosoma cruzi*	7·00E−68

2000). Moreover, studies on barley have revealed that while extensive colonization of brush hairs may occur by 48 h, colonization of glumes, awns and the outer lemma surface is still sparse at this timepoint (Skadsen & Hohn, 2004). The trichothecene mycotoxin DON accumulates in detectable amounts only 48 h after inoculation of barley (Evans *et al.*, 2000). Therefore, 48 h was considered an appropriate time point for investigating early stages of the host–pathogen interaction. Eighty-four putative fungal genes expressed during host–pathogen interaction between *F. graminearum* and wheat were identified. A predicted function for 49 of these genes was inferred by bioinformatic analysis.

Fifty-eight per cent of the fungal ESTs from the Fgr-S3/S4 library were found to be of mitochondrial origin, a much higher percentage than that from the fungus grown in culture (Trail *et al.*, 2003). This could indicate a general increase in mitochondrial RNA, which reflects increased mitochondrial activity and/or increased abundance of these organelles in the active disease state. Increase in levels of fungal mitochondrial RNA during plant infection has been reported previously for *M. grisea* on rice (Talbot & Tongue, 1998). The *F. graminearum* mitochondrial ESTs most frequently observed were related to either the NAD reductase or cytochrome *c* oxidase (COX) complexes. These results suggest that respiratory competence is likely to be critical for development of the fungus in the host tissues. This is further supported by the recent finding that reduced virulence in a REMI mutant of *F. graminearum* was caused by an insertion in a nuclear gene predicted to encode the NADH : ubiquinone oxidoreductase complex I (Seong *et al.*, 2005). The fungus may need enhanced mitochondrial activities to metabolize some plant molecules (for energy or detoxification) or simply to provide the energy for infectious growth. The mitochondrial antioxidant defence system also may be important for a necrotrophic pathogen like *F. graminearum*. Thus, while perhaps not specifically and uniquely involved in pathogenicity, these basic metabolic systems could prove to be effective targets for control of this pathogen.

Table 5. Distribution of 62 ESTs to the four *F. graminearum* chromosomes

Chromosome no.	Length (bp)*	Actual no. of genes	Expected no. of genes†
1	11 588 431	19	19·96
2	8 830 180	23	15·21
3	7 690 198	11	13·24
4	7 893 678	9	13·59

*Data from Gale *et al.* (2005). Total genome assembly size is 36 002 487 bp.

†The expected number of genes on each chromosome was calculated based on the number of genes expected to be present on the chromosome as determined by the proportion of the genome it represents (e.g. no. of genes in chromosome 1 = 11 588 431 × 62/36 002 487 = 19·96).

Several ESTs from our libraries match scaffold 32 of the *F. graminearum* genome assembly, which has not been assigned a chromosomal location (Gale *et al.*, 2005). Interestingly, the FgrS3S4Contig120 matching a region on scaffold 32 is homologous to a DNA polymerase from the mitochondrial plasmid pCry1 from *Cryphonectria parasitica* (Monteiro-Vitorello *et al.*, 2000). This result suggests that this scaffold may be associated with a mitochondrial plasmid previously not identified in *F. graminearum*, and warrants further investigation.

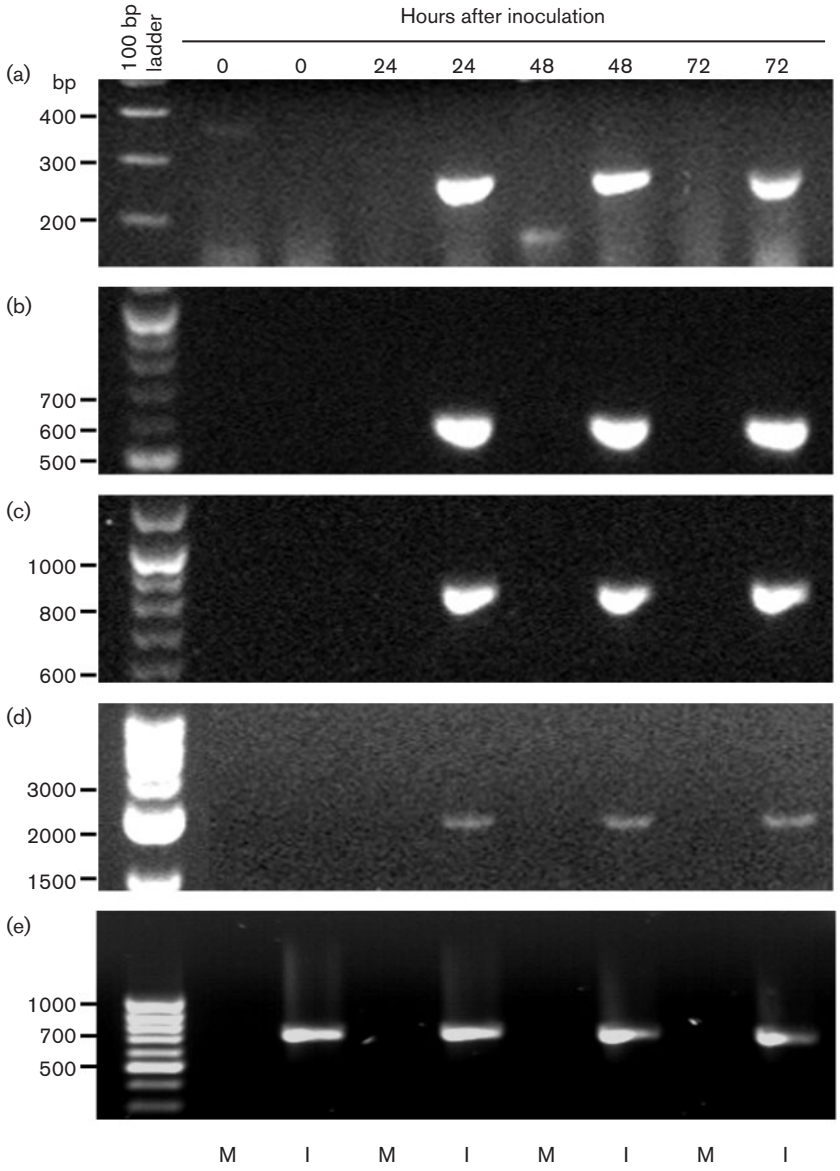

Fig. 1. Agarose gel electrophoresis of RT-PCR products using gene-specific primers (Supplementary Table S1) for the four genes *FgAbc2* (a), *FgLyp1* (b), *FgRrr1* (c) and *FgZbc1* (d), and the positive control EF-1α (e). RNA isolated from wheat heads inoculated with NRRL 31084 and water 0, 24, 48 and 72 h after inoculation was used as template for each of these primer pairs. RT-PCR products using RNA from water (M)- and NRRL 31084 (I)-inoculated wheat heads were loaded in alternate rows.

Within the Fgr-S3/S4 EST libraries were several predicted protein-coding sequences corresponding to genes that may be associated with pathogenicity. Among these is a gene encoding a polyketide synthase (PKS14 – FG03964.1). Polyketide synthases (PKSs) are large multifunctional enzymes involved in production of polyketide secondary metabolites, including toxins such as aflatoxin, T-toxin, fumonisin, patulin and ochratoxin (Kroken *et al.*, 2003). The *F. graminearum* gene identified as an EST during infection corresponds to a previously undescribed PKS orthologue of unknown function currently found only in *F. graminearum*. Expression of this gene has recently been described as 'grain-specific' since it has been found by RT-PCR to be expressed by the fungus during growth on rice and corn meal, but seemingly not during infection of wheat (Gaffoor *et al.*, 2005). Our results indicate that the gene may indeed be expressed during plant infection, and microarray data of the gene during infection of barley indicates this is the case (Güldener *et al.*, 2006b). The potential significance of the gene to pathogenicity remains to be determined.

An EST corresponding to a gene for an ABC transporter (FG08308.1) of unknown function also is represented in the Fgr-S3/S4 libraries. ABC transporters are membrane-associated proteins involved in transport of small molecules across membranes against a concentration gradient, coupled to the hydrolysis of ATP. Some fungal ABC transporters have been associated with plant pathogenicity (Fleissner *et al.*, 2002; Stergiopoulos *et al.*, 2003; Urban *et al.*, 1999), including a gene in the closely related species *F. culmorum* (Skov *et al.*, 2004); ABC transporters also have been shown to confer resistance to plant phytoalexins (Schoonbeek *et al.*, 2001). However, deletion of FG08308.1 in this study resulted in no loss of pathogenicity to wheat under the conditions tested.

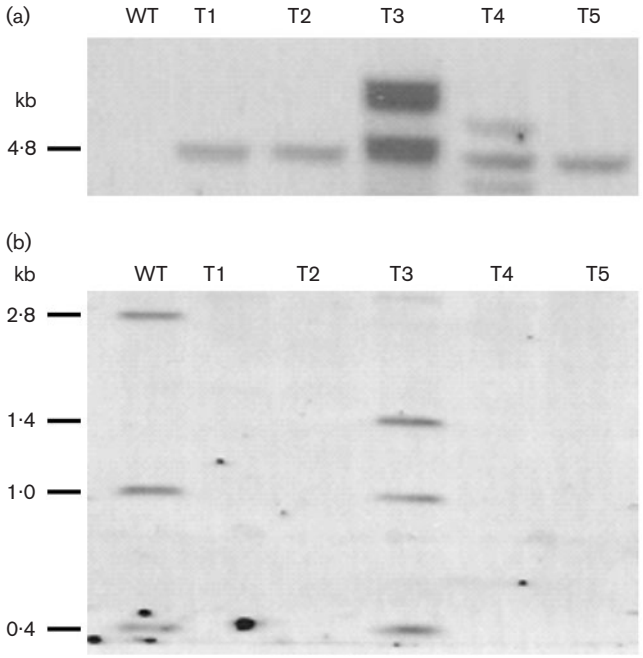

Fig. 2. Southern hybridization of *Hin*dIII-digested DNA from strain NRRL31084 (wild-type, WT) or transformants (T1–T5). (a) Blot probed with the labelled hygromycin gene fragment (*hph*), indicating that transformants T3 and T4 had additional ectopic copies of the gene. (b) Blot probed with labelled *Rrr1*, indicating that four transformants are *Rrr1* deletion mutants (T1, T2, T4, T5) whereas one transformant (T3) still has hybridization signals. The polymorphism observed in one *Hin*dIII band of T3 compared to wild-type indicated that the integration event had occurred in its flanking sequence.

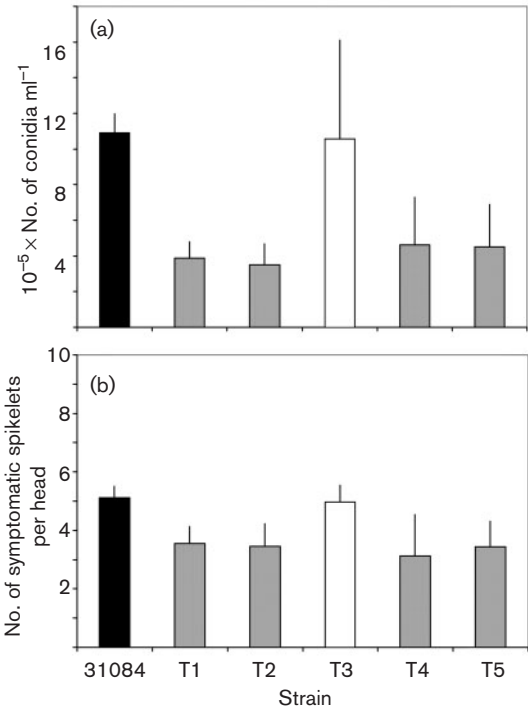

Fig. 3. Altered phenotypes of *Rrr1* deletion strains (T1, T2, T4 and T5) compared with wild-type strain NRRL 31084 and ectopic mutant T3. (a) Number of macroconidia ml^{-1} produced on 8-day-old cultures growing on mung-bean agar plates; (b) number of symptomatic spikelets per wheat head 7 dai with each strain. Error bars represent standard deviation among three replicates.

EST sequence Fgr-S4_2_M02 corresponds to predicted gene FG03405.1 (*Lyp1*) that has strong sequence similarity to and conserved features of a lysine permease. Fungal nutritional requirements *in planta* are being increasingly recognized as factors essential for establishing pathogenic interactions with plants (Solomon *et al.*, 2003). Some plant-pathogenic fungi have been shown to be rendered non-pathogenic by mutations leading to deficiency in the biosynthesis of certain amino acids (Balhadère *et al.*, 1999; Namiki *et al.*, 2001). Amino acid transporters, especially lysine transporters, have been previously noted to be strongly up-regulated during plant infection (Hahn *et al.*, 1997). Despite these observations, deletion of *Lyp1* alone in *F. graminearum* resulted in no loss of pathogenicity under the conditions tested here. This could be due to the fact that the *F. graminearum* genome has many predicted amino acid permeases that perhaps are redundant in function so that deletion of any one of these genes would have no profound effect on pathogencity.

The EST sequence Fgr-S4_3_I05_T7 has a significant match to a gene encoding REC1, a putative two-component response regulator from *Gibberella moniliformis*. Two-component signal transduction pathways are phosphorelay mechanisms by which various organisms sense and adapt to their environment. They are generally composed of a histidine kinase (HK) and a response regulator (RR) domain, which are easily identifiable by sequence alignment (Borkovich *et al.*, 2004). They have been found in bacteria, slime moulds, plants and fungi but not yet in animals (Borkovich *et al.*, 2004; Catlett *et al.*, 2003b). These systems have been shown to be involved in virulence responses in fungi. For example, the *fos-1* gene from *Aspergillus fumigatus* (Clemons *et al.*, 2002) and the *cos-1* gene from *Candida albicans* (Selitrennikoff *et al.*, 2001) are two-component HKs that have been shown to be putative virulence factors. The eukaryotic two-component phosphorelay systems mostly contain hybrid HKs where both the HK and RR domains are contained in a single polypeptide (Catlett *et al.*, 2003b). These hybrid HKs are believed to function in multi-step phosphorelays where the phosphate is transferred from an RR domain of the hybrid HK to a second histidine residue and then to a second RR domain. The REC1-encoding gene was identified while searching for such potential downstream response regulators. It could not be related to other known fungal RRs or RR domains in hybrid HKs. Since *G. moniliformis* is taxonomically closer to *G. zeae* than the other fungi to which it has been previously compared (Catlett

et al., 2003b), we propose that this gene could potentially be a novel RR or part of a previously uncharacterized two-component phosphorelay mechanism.

Among the four genes selected for our mutation studies, *Rrr1* was found to play a role in sporulation and disease expression. Mutants for this gene exhibited significantly reduced ability to produce macroconidia on mung-bean agar medium as compared to the wild-type strain. They also produced a significantly lower percentage of symptomatic spikelets as compared to the wild-type 7 dai on the wheat head. The ectopic mutant retained its ability to produce conidia and spread to generate symptomatic spikelets at levels similar to the wild-type (Fig. 3). This suggests that the *Rrr1* gene is likely to play a role in pathogenicity and development of the fungus. As signal transduction pathways controlled by genes such as *Rrr1* are phosphorelay mechanisms by which organisms sense and adapt to their environment (Borkovich *et al.*, 2004), reduction in this ability due to deletion of the gene could account for the delay in spread of the mutants on the wheat heads as reflected by our results. The observation that the total number of symptomatic spikelets 14 dai was similar to that produced by the wild-type strain leads us to infer that this gene is likely to be involved in the initial response of the fungus to the host plant.

We have been successful in identifying several fungal genes expressed in the early stages of infection on wheat that have the potential to be involved in host–pathogen interaction. This information will help not only in understanding changes in overall gene expression in the pathogen during infection of the host, but also in identifying potential targets for development of disease control strategies. Surprisingly, a large percentage of sequences expressed *in planta* were not called genes by the *ab initio* gene prediction models used by the Broad Institute and MIPS. Forty-five per cent of the ORFs represented in the Fgr-S3/S4 library belong to this category. This suggests that currently available gene prediction models are not sufficient for identifying all *F. graminearum* genes and highlights the need for developing more appropriate models better suited to this organism. The ESTs from this library have thus been helpful in manual annotation of the *F. graminearum* genome and have supported the identification of 35 new ORFs. A more comprehensive study of fungal gene expression during plant infection is ongoing using a newly designed *F. graminearum* DNA microarray.

ACKNOWLEDGEMENTS

DNA sequencing was conducted at Purdue University under the direction of Phillip San Miguel. We are grateful to Jack Kennell for his help in evaluating the mitochondrial sequences and Gerhard Adam for his thoughtful suggestions regarding possible functions of genes identified in our libraries. We are thankful to Ulrich Güldener, Rodney Staggs and Wayne Xu for their assistance in bioinformatics analysis. We would like to thank Linda Harris for providing access to EST sequences from *F. graminearum* grown under various culture conditions. Thanks are due to Sam Gale for the immense help with gene replacement procedures. We also thank Liane Gale and Arianna Lindmann for technical assistance. We appreciate Jacolyn Morrison's help with artwork. The project was supported by the National Research Initiative of the USDA Cooperative State Research, Education and Extension Service, grant number 2004-35604-14327, and a grant from the US Wheat and Barley Scab Initiative to F. T., J.-R. X. and H. C. K. The mention of firm names or trade products does not imply that they are endorsed or recommended by the US Department of Agriculture over other firms or similar products not mentioned.

REFERENCES

Balhadère, P. V., Foster, A. J. & Talbot, N. J. (1999). Identification of pathogenicity mutants of the rice blast fungus *Magnaporthe grisea* by insertional mutagenesis. *Mol Plant Microbe Interact* **12**, 129–142.

Beyer, K., Jimenez, S. J., Randall, T. A., Lam, S., Binder, A., Boller, T. & Collinge, M. A. (2002). Characterization of *Phytophthora infestans* genes regulated during the interaction with potato. *Molecular Plant Pathology* **3**, 473–485.

Bittner-Eddy, P. D., Allen, R. L., Rehmany, A. P., Birch, P. & Beynon, J. L. (2003). Use of suppression subtractive hybridization to identify downy mildew genes expressed during infection of *Arabidopsis thaliana*. *Mol Plant Pathol* **4**, 501–507.

Borkovich, K. A., Alex, L. A., Yarden, O. & 36 other authors (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.

Catlett, N. L., Lee, B. N., Yoder, O. C. & Turgeon, B. G. (2003a). Split-marker recombination for efficient targeted deletion of fungal genes. *Fungal Genet Newsl* **50**, 9–11.

Catlett, N. L., Yoder, O. C. & Turgeon, B. G. (2003b). Whole-genome analysis of two-component signal transduction genes in fungal pathogens. *Eukaryot Cell* **2**, 1151–1161.

Clemons, K. V., Miller, T. K., Selitrennikoff, C. P. & Stevens, D. A. (2002). *fos*-1, a putative histidine kinase as a virulence factor for systemic aspergillosis. *Med Mycol* **40**, 259–262.

Cramer, R. A. & Lawrence, C. B. (2004). Identification of *Alternaria brassicicola* genes expressed *in planta* during pathogenesis of *Arabidopsis thaliana*. *Fungal Genet Biol* **41**, 115–128.

Dean, R. A., Talbot, N. J., Ebbole, D. J. & 32 other authors (2005). The genome sequence of the rice blast fungus *Magnaporthe grisea*. *Nature* **434**, 980–986.

Diatchenko, L., Lau, Y. F., Campbell, A. P. & 8 other authors (1996). Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. *Proc Natl Acad Sci U S A* **93**, 6025–6030.

Evans, C. K., Xie, W., Dill-Macky, R. & Mirocha, C. J. (2000). Biosynthesis of deoxynivalenol in spikelets of barley inoculated with macroconidia of *Fusarium graminearum*. *Plant Dis* **84**, 654–660.

Fleissner, A., Sopalla, C. & Weltring, K. M. (2002). An ATP-binding cassette multidrug-resistance transporter is necessary for tolerance of *Gibberella pulicaris* to phytoalexins and virulence on potato tubers. *Mol Plant Microbe Interact* **15**, 102–108.

Gaffoor, I., Brown, D. W., Plattner, R., Proctor, R. H., Qi, W. & Trail, F. (2005). Functional analysis of the polyketide synthase genes in the filamentous fungus *Gibberella zeae* (Anamorph *Fusarium graminearum*). *Eukaryot Cell* **4**, 1926–1933.

Galagan, J. E., Calvo, S. E., Borkovich, K. A. & 74 other authors (2003). The genome sequence of the filamentous fungus *Neurospora crassa*. *Nature* **422**, 859–868.

Gale, L. R., Bryant, J., Calvo, S. & 8 other authors (2005). Chromosome complement of the fungal plant pathogen *Fusarium graminearum* based on genetic and physical mapping and cytological observations. *Genetics* **171**, 985–1001.

Goswami, R. S. & Kistler, H. C. (2004). Heading for disaster: *Fusarium graminearum* on cereal crops. *Mol Plant Pathol* **5**, 515–525.

Goswami, R. S. & Kistler, H. C. (2005). Pathogenicity and *in planta* mycotoxin accumulation among members of the *Fusarium graminearum* species complex on wheat and rice. *Phytopathology* **95**, 1397–1404.

Guilleroux, M. & Osbourn, A. (2004). Gene expression during infection of wheat roots by the 'take-all' fungus *Gaeumannomyces graminis*. *Mol Plant Pathol* **5**, 203–216.

Güldener, U., Mannhaupt, G., Münsterkötter, M., Haase, D., Oesterheld, M., Stümpflen, V., Mewes, H. W. & Adam, G. (2006a). FGDB: a comprehensive fungal genome resource on the plant pathogen *Fusarium graminearum*. *Nucleic Acids Res* **34**, D456–D458.

Güldener, U., Seong, K., Boddu, J. & 7 other authors (2006b). Development of a *Fusarium graminearum* Affymetrix GeneChip for profiling fungal gene expression *in vitro* and *in planta*. *Fungal Genet Biol* (in press).

Hahn, M., Neef, U., Struck, C., Göttfert, M. & 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.

Hou, Z. M., Xue, C. Y., Peng, Y. L., Katan, T., Kistler, H. C. & Xu, J. R. (2002). A mitogen-activated protein kinase gene (*MGV1*) in *Fusarium graminearum* is required for female fertility, heterokaryon formation, and plant infection. *Mol Plant Microbe Interact* **15**, 1119–1127.

Huang, X. Q. & Madan, A. (1999). CAP3: a DNA sequence assembly program. *Genome Research* **9**, 868–877.

Jenczmionka, N. J., Maier, F. J., Losch, A. P. & Schafer, W. (2003). Mating, conidiation and pathogenicity of *Fusarium graminearum*, the main causal agent of the head-blight disease of wheat, are regulated by the MAP kinase gpmk1. *Curr Genet* **43**, 87–95.

Kim, Y. T., Lee, Y. R., Jin, J. M., Han, K. H., Kim, H., Kim, J. C., Lee, T., Yun, S. H. & Lee, Y. W. (2005). Two different polyketide synthase genes are required for synthesis of zearalenone in *Gibberella zeae*. *Mol Microbiol* **58**, 1102–1113.

Kroken, S., Glass, N. L., Taylor, J. W., Yoder, O. C. & Turgeon, B. G. (2003). Phylogenomic analysis of type I polyketide synthase genes in pathogenic and saprobic ascomycetes. *Proc Natl Acad Sci U S A* **100**, 15670–15675.

Kruger, W. M., Pritsch, C., Chao, S. M. & 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.

Lu, S. W., Kroken, S., Lee, B. N., Robbertse, B., Churchill, A. C. L., Yoder, O. C. & Turgeon, B. G. (2003). A novel class of gene controlling virulence in plant pathogenic ascomycete fungi. *Proc Natl Acad Sci U S A* **100**, 5980–5985.

Lu, G., Jantasuriyarat, C., Zhou, B. & Wang, G. L. (2004). Isolation and characterization of novel defense response genes involved in compatible and incompatible interactions between rice and *Magnaporthe grisea*. *Theor Appl Genet* **108**, 525–534.

McMullen, M., Jones, R. & Gallenberg, D. (1997). Scab of wheat and barley: a re-emerging disease of devastating impact. *Plant Dis* **81**, 1340–1348.

Monteiro-Vitorello, C. B., Baidyaroy, D., Bell, J. A., Hausner, G., Fulbright, D. W. & Bertrand, H. (2000). A circular mitochondrial plasmid incites hypovirulence in some strains of *Cryphonectria parasitica*. *Curr Genet* **37**, 242–256.

Namiki, F., Matsunaga, M., Okuda, M., Inoue, I., Nishi, K., Fujita, Y. & Tsuge, T. (2001). Mutation of an arginine biosynthesis gene causes reduced pathogenicity in *Fusarium oxysporum* f. sp. *melonis*. *Mol Plant Microbe Interact* **14**, 580–584.

O'Donnell, K., Cigelnik, E. & Nirenberg, H. I. (1998). Molecular systematics and phylogeography of the *Gibberella fujikuroi* species complex. *Mycologia* **90**, 465–493.

O'Donnell, K., Kistler, H. C., Tacke, B. K. & Casper, H. H. (2000). Gene genealogies reveal global phylogeographic structure and reproductive isolation among lineages of *Fusarium graminearum*, the fungus causing wheat scab. *Proc Natl Acad Sci U S A* **97**, 7905–7910.

O'Donnell, K., Ward, T. J., Geiser, D. M., Kistler, H. C. & Aoki, T. (2004). Genealogical concordance between the mating type locus and seven other nuclear genes supports formal recognition of nine phylogenetically distinct species within the *Fusarium graminearum* clade. *Fungal Genet Biol* **41**, 600–623.

Pritsch, C., Muehlbauer, G. J., Bushnell, W. R., Somers, D. A. & Vance, C. P. (2000). Fungal development and induction of defense response genes during early infection of wheat spikes by *Fusarium graminearum*. *Mol Plant Microbe Interact* **13**, 159–169.

Proctor, R. H., Hohn, T. M. & McCormick, S. P. (1995). Reduced virulence of *Gibberella zeae* caused by disruption of a trichothecene toxin biosynthetic gene. *Mol Plant Microbe Interact* **8**, 593–601.

Rosewich, U. L., Pettway, R. E., Katan, T. & Kistler, H. C. (1999). Population genetic analysis corroborates the dispersal of *Fusarium oxysporum* f. sp. *radicis-lycopersici* from Florida and Europe. *Phytopathology* **89**, 623–630.

Schoonbeek, H., Del Sorbo, G. & De Waard, M. A. (2001). The ABC transporter BcatrB affects the sensitivity of *Botrytis cinerea* to the phytoalexin resveratrol and the fungicide fenpiclonil. *Mol Plant Microbe Interact* **14**, 562–571.

Selitrennikoff, C. P., Alex, L., Miller, T. K., Clemons, K. V., Simon, M. I. & Stevens, D. A. (2001). COS-I, a putative two-component histidine kinase of *Candida albicans*, is an in vivo virulence factor. *Med Mycol* **39**, 69–74.

Seong, K., Hou, Z. M., Tracy, M., Kistler, H. C. & Xu, J. R. (2005). Random insertional mutagenesis identifies genes associated with virulence in the wheat scab fungus *Fusarium graminearum*. *Phytopathology* **95**, 744–750.

Seong, K., Li, L., Kistler, H. C. & Xu, J.-R. (2006). Cryptic promoter activity of the HMR1 coding region in the wheat scab fungus *Fusarium graminearum*. *Fungal Genet Biol* **43**, 34–41.

Skadsen, R. W. & Hohn, T. A. (2004). Use of *Fusarium graminearum* transformed with *gfp* to follow infection patterns in barley and Arabidopsis. *Physiol Mol Plant Pathol* **64**, 45–53.

Skov, J., Lemmens, M. & Giese, H. (2004). Role of a *Fusarium culmorum* ABC transporter (FcABC1) during infection of wheat and barley. *Physiol Mol Plant Pathol* **64**, 245–254.

Solomon, P. S., Tan, K. C. & Oliver, R. P. (2003). The nutrient supply of pathogenic fungi; a fertile field for study. *Molecular Plant Pathology* **4**, 203–210.

Stack, R. W. (2003). History of Fusarium head blight with emphasis on North America. In *Fusarium Head Blight of Wheat and Barley*, pp. 1–34. Edited by K. J. Leonard & W. R. Bushnell. St Paul, MN: APS Press.

Stergiopoulos, I., Zwiers, L. H. & De Waard, M. A. (2003). The ABC transporter MgAtr4 is a virulence factor of *Mycosphaerella graminicola* that affects colonization of substomatal cavities in wheat leaves. *Mol Plant Microbe Interact* **16**, 689–698.

Tag, A. G., Garifullina, G. F., Peplow, A. W., Ake, C., Jr, Phillips, T. D., Hohn, T. M. & Beremand, M. N. (2001). A novel regulatory gene, *Tri10*, controls trichothecene toxin production and gene expression. *Appl Environ Microbiol* **67**, 5294–5302.

Talbot, N. J. & Tongue, N. (1998). High level expression of the *Magnaporthe grisea* mitochondrial small sub-unit rRNA during rice leaf colonization and rapid down-regulation prior to the onset of disease symptoms. *Physiol Mol Plant Pathol* **52**, 335–352.

Trail, F., Xu, J. R., San Miguel, P., Halgren, R. G. & Kistler, H. C. (2003). Analysis of expressed sequence tags from *Gibberella zeae* (anamorph *Fusarium graminearum*). *Fungal Genet Biol* **38**, 187–197.

Urban, M., Bhargava, T. & Hamer, J. E. (1999). An ATP-driven efflux pump is a novel pathogenicity factor in rice blast disease. *EMBO J* **18**, 512–521.

Urban, M., Mott, E., Farley, T. & Hammond-Kosack, K. (2003). The *Fusarium graminearum MAP1* gene is essential for pathogenicity and development of perithecia. *Molecular Plant Pathology* **4**, 347–359.

Verwoerd, T. C., Dekker, B. M. & Hoekema, A. (1989). A small-scale procedure for the rapid isolation of plant RNAs. *Nucleic Acids Res* **17**, 2362.

Voigt, C. A., Schafer, W. & Salomon, S. (2005). A secreted lipase of *Fusarium graminearum* is a virulence factor required for infection of cereals. *Plant J* **42**, 364–375.

Xiong, L. Z., Lee, M. W., Qi, M. & Yang, Y. N. (2001). Identification of defense-related rice genes by suppression subtractive hybridization and differential screening. *Mol Plant Microbe Interact* **14**, 685–692.

Zhao, X., Xue, C., Kim, Y. & Xu, J. R. (2004). A ligation-PCR approach for generating gene replacement constructs in *Magnaporthe grisea*. *Fungal Genet Newsl* **51**, 17–18.