The Predicted Secretome of the Plant Pathogenic Fungus
Fusarium graminearum: A Refined Comparative Analysis

Neil A. Brown, John Antoniw, Kim E. Hammond-Kosack*

Centre for Sustainable Pest and Disease Management, Department of Plant Pathology and Microbiology, Rothamsted Research, Harpenden, Hertfordshire, United Kingdom

Abstract

The fungus Fusarium graminearum forms an intimate association with the host species wheat whilst infecting the floral tissues at anthesis. During the prolonged latent period of infection, extracellular communication between live pathogen and host cells must occur, implying a role for secreted fungal proteins. The wheat cells in contact with fungal hyphae subsequently die and intracellular hyphal colonisation results in the development of visible disease symptoms. Since the original genome annotation analysis was done in 2007, which predicted the secretome using TargetP, the F. graminearum gene call has changed considerably through the combined efforts of the BROAD and MIPS institutes. As a result of the modifications to the genome and the recent findings that suggested a role for secreted proteins in virulence, the F. graminearum secretome was revisited. In the current study, a refined F. graminearum secretome was predicted by combining several bioinformatic approaches. This strategy increased the probability of identifying truly secreted proteins. A secretome of 574 proteins was predicted of which 99% was supported by transcriptional evidence. The function of the annotated and unannotated secreted proteins was explored. The potential role(s) of the annotated proteins including, putative enzymes, phytotoxins and antifungals are discussed. Characterisation of the unannotated proteins included the analysis of Pfam domains and features associated with known fungal effectors, for example, small size, cysteine-rich and containing internal amino acid repeats. A comprehensive comparative genomic analysis involving 57 fungal and oomycete genomes revealed that a small number of the predicted F. graminearum secreted proteins can be considered to be either species or sequenced strain specific.

Introduction

The intimacy of an association between a microbe and plant host is represented by foreign cells growing within plant tissue or even within living plant cells. Communication through the secretion of proteins and metabolites that are either taken up by the host or detected at the cell surface plays a pivotal role in determining the outcome of the interaction. Secreted proteins from animal-infecting malaria parasites and plant-infecting oomycete pathogens possess a conserved RxLR motif that facilitates protein secretion and uptake into the host cells resulting in the modulation of host transcription [1,2,3]. Fungi and oomycetes have convergently evolved a range of mechanisms to acquire nutrition from various habitats, including mutalistic, biotrophic, hemibiotrophic, necrotrophic and non-pathogenic saprophytic lifestyles. In fungi, no widely conserved translocation motif has been discovered, yet many small secreted proteins and metabolites are proven virulence factors [4]. However, a degenerative Y/F/WxC motif discovered in Blumeria graminis f. sp. hordei has been proposed to be conserved among intracellular non-necrotrophic ascomycetes [5]. Experimentally, secreted proteins termed ‘effectors’ that modulate the interaction between pathogenic microbes and hosts have been identified from all lifestyles. Examples include; the Avr and Ecp proteins from the tomato leaf mold fungus Cladosporium fulvum [6], the Tox proteins from the wheat glume blotch fungus Stagonospora nodorum [7], Avr10 and Avr1k1 from the barley powdery mildew fungus B. graminis f. sp. hordei [8], the SIX proteins from the vascular wilt fungus Fusarium oxysporum f. sp. lycopersici [9], the Avr-Pita and Pw1 proteins from the rice blast fungus Magnaporthe oryzae [10], Pep1 and Pit1/2 from the corn smut fungus Ustilago maydis [11], 3LysM from the wheat leaf blotch fungus Mycosphaerella graminicola [12] and Sp7 from the tomato mutant Glomeris intraradices [13].

Several apoplastic cysteine-rich and LysM containing fungal effectors have been shown to inhibit plant chitinases and/or bind chitin to prevent elicitation of pathogen associated molecular pattern (PAMP) triggered immunity (PTI) and thereby prevent the induction of host defences. These include the C. fulvum effectors Avr4/Ecp6 and M. graminicola 3LysM effector [12,14,15]. Recently the accelerated evolution of secreted proteins through internal amino acid repeats, which increase phenotypic plasticity, has been shown to influence the elicitation of the host’s defence response [16,17,18]. Several intracellular effectors contribute to virulence in a different way. For example, Sp7 and Pw1/2 are translocated to the
host nucleus where they influence host transcription [13,19]. Alternatively in *E. maydis*, Pep1 accumulates at the site of cell-to-cell passage and is essential to the establishment of infection [11], while clusters of effectors contribute to organ specificity [20]. Different again are the small necrotrophic effectors (Tox proteins) that induce host programmed cell death (PCD) to assist infection [7]. These examples from fungi with different *in planta* lifestyles demonstrate how different types of secreted fungal proteins define the outcome of an interaction between a microbe and its host.

Globally, the homoblastic ascomycete fungus *Fusarium graminearum* (Teleomorph: Gibberella zeae) is the predominant causal agent of *Fusarium* Ear Blight (FEB) disease, also referred to as Head Scab (www.scabusa.org). This disease affects most small grain cereal species including wheat, barley and maize and has been associated with up to 17 *Fusarium* species. Serious and repeated FEB outbreaks have been reported in all major wheat producing countries (www.faostat.fao.org) and consequently, the international maize and wheat improvement centre (CIMMYT) describes FEB as a major limiting factor to production [21]. This re-emergence is thought to be driven by changes in agronomic practices as well as to climatic changes. In addition to reducing grain yield and product quality, the crop is also contaminated with mycotoxins that are harmful to both animals and humans [22]. Due to the health threat, farmers in the EU and USA pay for their grain to be tested for the presence of the type B trichothecene mycotoxin deoxynivalenol (DON) and to determine that the levels are below the recommended safety guidelines (www.hgca.com, www.USDA.com) [23]. It is estimated that one infected ear per square metre of the wheat crop is sufficient for the DON concentrations in the grain to exceed safe levels.

*F. graminearum* was one of the first plant pathogens to be selected for full genome sequencing due to the growing global importance of the disease, the large number of cereal species infected and the growing health concerns. The sequenced *F. graminearum* genome of the PH-1 strain was found to be 36.1 Mb in size and due to the availability of a genetic map between the PH-1 and a second strain of USA origin (MNN0-676), this genomic sequence was immediately aligned to the four chromosomes [24]. Due to the homoblastic nature of *F. graminearum* and because an active repeat-induced point mutation system operates during each meiosis, the genome contains little repetitive DNA and no active transposable elements compared to related fungi [25]. The latest version of *F. graminearum* genome available from MIPS (version FG3.2) has a considerable amount of manual annotation incorporated and is predicted to encode 13,718 genes [26]. In the original genome analysis of ~11,600 genes [25], the TargetP defined secretome was predicted to account for approximately one tenth (1,442) of the predicted genes. Low level sequence coverage of a second *F. graminearum* strain, GZ3639, demonstrated the non-random distribution of nucleotide polymorphism in the genome, with hot spots of sequence variation occurring in sub-telomeric and central regions [25]. These highly variable ‘hot’ regions of the genome were found to be enriched for genes coding for predicted secreted proteins. An analysis of the genomic location of experimentally proven *F. graminearum* pathogenicity/virulence genes and homologues of verified pathogenicity/virulence genes from other species (www.PHbase.org) has revealed that most genes with this function resided in regions of low level recombination [27]. Their location in the ‘cooler’ parts of the genome has been suggested to protect them from gene loss [28]. The majority of these genes code for conserved intracellular proteins involved in signal transduction, such as the mitogen activated protein kinases (MAPKs), and represent ancient conserved signalling pathways recruited by pathogens to coordinate infection [29]. Within *F. graminearum* genomic regions found to exhibit high recombination frequencies [25] reside genes that encode for small and large sized secreted proteins. For instance, an abundance of plant cell wall degrading enzymes (PCWDEs) was identified. This type of genome positioning is hypothesised to assist the evolution of the pathogen in the rapidly changing arms race with its host. By contrast, in the oomycete *Phytophthora infestans* the vast majority of the predicted secretome, which is evolving rapidly, is located in regions of the genome where an abundance of transposon sequences reside [30]. A similar genome location, rich in transposon sequences, is now recognised to harbour the predicted secretome of the Ascomycete powdery mildew fungus, *B. graminis* f. sp. hordei [31].

Production of the water soluble, secreted trichothecene mycotoxin, deoxynivalenol (DON), is required by *F. graminearum* for full virulence on wheat ears, but not for full virulence on barley ears, maize cobs or the floral tissue of the model species *Arabidopsis thaliana* [32,33,34,35,36]. DON inhibits protein synthesis in eukaryotes and prevents polypeptide chain initiation or elongation by binding to the 60S ribosomal subunit [37]. In *F. graminearum* infections of wheat, the trichothecene mycotoxin genes within the *Tri* cluster are most highly expressed during symptomless infection [38]. Wheat ear infection by the non-DON producing *tr5* gene deficient mutant results in an enhanced defence response in the form of plant cell wall thickening adjacent to the invading *Fusarium* hyphae [39]. In the absence of DON production, the interaction between the two organisms at the infection front is altered. A macroscopically visible brown ring forms around the slowly expanding lesion on the glumes of wheat ears sprayed with *F. graminearum* [34]. Topoisomerase modulation of DNA topology has been demonstrated to regulate virulence gene expression, especially secreted proteins [40,41]. The top1 deficient *F. graminearum* strain was unable to colonise the wheat ear despite producing wild-type DON levels and infections were restricted to just below the surface of the floral brackets [42]. While the secreted lipase *fgp1*-deficient strain produced enhanced DON in *planta*, yet in wheat ears an extensive host cell browning reaction was evident in the tissue immediately beyond the confined *Fusarium* hyphae [43]. Collectively this implies that additional virulence factors, in combination with the secreted DON mycotoxin, promote symptomless infection and implicates a role for secreted proteins in *F. graminearum* pathogenicity.

In view of these findings and the recently identified symptomless phase of wheat ear infection where the *F. graminearum* hyphae advance exclusively extracellularly between the wheat cells [44], we decided to explore in detail the predicted secretome. This new study of the secretome could potentially give the first clues to what proteins are involved in the establishment/maintenance of symptomless infection, as well as the transition from extracellular to intracellular growth. Various bioinformatic tools that assist the prediction of fungal secretomes are available. These tools utilise different but highly complementary analytical approaches, namely the prediction of the presence of a signal peptide (SignalP/TargetP) and predicting the eventual cellular location of the mature protein (WolPSort). Used individually these approaches often predict non-secreted proteins as secreted, but when used in combination an increased accuracy of the prediction was anticipated. In this study we describe, in detail, a refined prediction and possible function of the *F. graminearum* secretome. In addition, a genomic comparison of the *F. graminearum* secretome with 57 similarly predicted fungal and oomycete proteomes, including other Fusaria and many pathogenic/non-pathogenic species, has been used to partition this predicted secretome into species specific, genera specific and highly conserved gene sets.
Materials and Methods

Bioinformatic analyses of the secretome

The FG3 version of the genome was downloaded from MIPS (http://mips.helmholtz-muenchen.de/genre/proj/FGDB/) in October 2009. The prediction of the refined F. graminearum secretome was based on the procedure described by Muller and colleagues [45] for U. maydis. We developed an automated secretome prediction pipeline based on this procedure using bash shell, AWK and Python scripts on a PC running Red Hat Linux 5.2. Initially all proteins with a Target P Loc = S (TargetP v1.1; http://www.cbs.dtu.dk/cgi-bin/nph-sw_request?targetp) and a Signal P D-score = Y (SignalP v3.0; http://www.cbs.dtu.dk/cgi-bin/nph-sw_request?signalp) were combined [46,47]. These were then scanned for transmembrane spanning regions using TMHMM (TMHMM v2.0; http://www.cbs.dtu.dk/cgi-bin/nph-sw_request?tmhmm) and all proteins with 0 TMs or 1 TM, if located in the predicted N-terminal signal peptide, were kept. GPI-anchor proteins were predicted by big-PI (http://mendel.imp.ac.at/gpi/cgi-bin/gpi_pred_fungi.cgi) [48]. ProtComp was also used to predict localization of the remaining proteins using the LocDB and PotLocDB databases (ProtComp v8.0; http://www.softberry.com). All proteins predicted as extracellular or unknown were kept in the final secretome dataset. Pfam analysis was done using the Pfam database (ftp://ftp.ncbi.nih.gov/pub/mmdb/cdd/) and the rpblast program in the NCBI blast using the Pfam database (ftp://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/). WolfPSort analysis was done using “runWolfPsortSummary fungi” in the WoLFSORT v0.2 package [49]. The number of cysteine residues within the mature peptide and the search for degenerative Y/F/WxC motifs were computed using custom Python scripts. The number of internal amino acid repeats was found using RADAR (http://www.cbi.ac.uk/Tools/Radar/) [50]. The detection of RNA transcripts for the 574 F. graminearum genes of interest was explored using Affymetrix gene expression data generated in several published in planta and in vitro investigations (Experiments FG1, FG2, FG15 and FG16) downloaded from www.PLEXdb.org.

Analysis of chromosome location alongside other key features of the F. graminearum genome

To inspect the position of individual or clusters of genes on the four F. graminearum chromosomes, the Fgra3Map tool was downloaded from www.Omnimapfree.org which displays a map of the complete F. graminearum genome (MIPS version 3.1). The Fgra3Map was used according to methods described [51].

Comparative analysis of the refined F. graminearum secretome

For the detailed follow up analyses, only proteins with a predicted signal peptide sequence and a value of extr 18 or greater from the WolfPSort analysis were used. The F. graminearum secretome was compared with 57 other fungal and oomycete genomes of pathogens varying in host range, tissue specificity and lifestyle as well as several exclusively saprophytic species (Table S1). The fungal and oomycete genomes and their predicted gene repertoires were downloaded from either the BROAD or JGI websites or from species specific websites maintained by various research communities. For the comparative analyses, the conservation, absence or expansion of the genes coding for the F. graminearum secreted proteins was explored by BLASTP analysis, determined at two levels of confidence, p<e^-5 and p<e^-40.

Results

The secretome of F. graminearum

In the original genome paper [25], the secretome was predicted from the FG1 gene call using only the TargetP software. In the current study we analysed an updated, refined FG3 gene call (13,937 proteins) of the F. graminearum genome in two phases. In the first phase (Figure 1A), designed to predict all possible secreted proteins, SignalP and TargetP were used to identify secreted proteins with signal peptides (1,853 proteins) and those predicted to contain GPI anchors (120 proteins) were identified. After removal of the signal peptide, any mature proteins that contained a transmembrane domain were excluded. An initial screen used ProtComp software to exclude proteins that were probably not located in the extracellular space. This produced a set of 1,369 secreted proteins (including those with GPI anchors). Phase 2 (Figure 1B), designed to identify proteins with a high probability of being secreted, contained more stringent conditions to further refine this set of proteins, discarding both those that did not begin with a methionine and small proteins where the mature proteins were shorter than 20 amino acids. At this stage, the 41 proteins with a TM domain predicted within the signal peptide sequence were also excluded. Similarly, all proteins predicted to contain a GPI-anchor were removed. A second software package (WolfPSort) that predicts the eventual location of proteins was used to only identify proteins that are secreted into the extracellular spaces (extracellular score >17). This resulted in a reduced set of 574 secreted proteins. In total 99% of the refined F. graminearum secretome is supported by transcriptional evidence from published in vitro and in planta investigations. Five of the fungal genes included within the refined secretome have not been assigned Affymetrix probe-sets and are therefore not supported by transcriptional evidence.

For completeness, and to assist with follow up comparative analyses, the results for the predicted secretome with the larger size of 1,369 genes arising from phase A of the analysis are presented in Spreadsheet S1.

The MIPS annotation and functional classification was determined for the 574 secreted proteins present in the FG3 gene call (http://mips.helmholtz-muenchen.de/genre/proj/FGDB/). Of these, 278 proteins possessed information on protein function whilst 296 proteins were described as hypothetical or conserved hypothetical.

The chromosomal location of the genes encoding the 574 secreted proteins was compared to the recombination frequency across the four chromosomes using Fgra3Map software. Genes coding for secreted proteins were identified on all four chromosomes and were found to be preferentially located within subtelomeric regions and regions with a high recombination frequency (Figure 2). A similar distribution pattern had been noted in the original FG1 analysis [25]. Within this overall pattern, the annotated genes and the unannotated genes present in the refined secretome were equally represented in the high and low recombination regions of the genome. To inspect whether any F. graminearum genes that encode secreted proteins were organised in clusters the secretome was divided into genes that reside in regions of low or high recombination frequency and displayed on the genome. A few small clusters that demonstrated no clear conservation in function were identified in regions of low and high recombination (Figure 2). In total, 51.6% of these genes were annotated as either hypothetical or conserved hypothetical. Secretome clusters were small in size, containing three to nine genes and were coded for by either DNA strand. The clusters within areas of high recombination were sub-telomERICally located.
on chromosomes 1 and 3 as well as an interstitial hot spot on chromosome 2. Eight clusters resided in regions of low recombination and were closely located within a 97 and 495 Kb region of chromosomes 2 and 3, respectively. The 9th cluster was located in a ‘cool’ sub-telomeric region on chromosome 1. No clusters were found on chromosome 4.

Analysis of the proteins with a predicted function

Closer inspection of the 278 proteins possessing information on protein function revealed that 243 contained at least one Pfam domain. A sub-set of 171 proteins was predicted to be involved in the degradation of plant derived compounds. These were divided according to substrate specificity (Table 1, Tables S2 and S3). After excluding five fungal chitinases, the remaining 102 annotated proteins considered not to be involved in plant substrate degradation were organised according to their MIPS functional category. Each sub-set is described in turn.

Almost all host plant surfaces are coated by a waxy cuticle, which represents the first barrier to plant infection. The plant cell wall beneath consists of cellulose microfibrils cross-linked by an amorphous matrix of hemicellulose and pectin, often encased in lignin polymers as the plant matures. The *F. graminearum* secretome possesses an arsenal of secreted proteins and enzymes that target the plant cuticle and each of the cell wall components. This arsenal potentially involving up to 109 secreted proteins (Table 1 and Table S2). Thirty secreted proteins involved in the degradation of

---

**Figure 1. The bioinformatics pipelines used to predict the *F. graminearum* secretome.** (A) The total secretome and (B) the refine secretome. doi:10.1371/journal.pone.0033731.g001
cellulose were identified, four of which were predicted to bind cellulose, fourteen were predicted to target β-1,4 glucans and twelve to target the breakdown product cellobiose. Enzymes that modify the different polysaccharides which make up hemicellulose represent the largest group of secreted cell wall modifying proteins and this reflects the diversity in hemicellulose composition. The two major components of hemicelluloses, arabinose and xylan, were targeted by the greatest number of secreted proteins. Nine secreted enzymes were detected that degrade the phenolic polymer lignin and its crosslinks to hemicellulose, including laccases, peroxidases and ferulic acid esterases. Multiple pectate lyases and pectin esterases were found (n = 13) that breakdown pectin in the middle lamella and cell wall of the plant. Callose is a polysaccharide of β-1,3 glucan that exists in plasmodesmata, phloem sieve plates and is laid down in response to wounding or imminent pathogen attack. Nine enzymes that target callose were identified, including endo- and exo- β-1,3 glucosidases.

Beyond the plant surface and the cell wall, the rest of the plant cell consists of proteins, lipid, sugars and nucleic acids. In total, 37 protein digesting enzymes were identified (Table 1 and Table S3), and included multiple alkaline/neutral and serine/aspartyl proteinases as well as amino, carboxy and endo peptidase. In contrast, only three enzymes were identified that were predicted to breakdown starch into sugars suitable for uptake by the fungal cell. These were two amylases and a glucose dehydrogenase. A high number of secreted enzymes that target lipids were identified, including 15 triacylglycerol lipases. Therefore, *F. graminearum* secretes an array of proteins that possess the ability to degrade and utilise the plant cell in its entirety. In 2006, a comparative genome analysis including *Ustilago maydis* and various newly available completed fungal genomes, before the *F. graminearum* genome was

Figure 2. The macro- and micro-chromosome locations of the genes predicted to encode the refined *F. graminearum* secretome. (A) Gene distribution across the four chromosome (Chr 1–4), where each black vertical bars represents a single gene (n = 574), aligned next to a heat map for genetic recombination (red = high to blue = low, recombination frequency - upper row of each chromosome) displayed on Fgra3Map. Some *F. graminearum* genes that encode secreted proteins were organised in clusters. The secretome was divided into genes that reside in regions of low (blue bars) or high frequency recombination (red bars) and displayed on Fgra3Map. Details of the gene clusters coding for secreted proteins in low (B) and (C) high recombination regions. Clusters are presented in chromosome order, while the coloured arrows (secreted proteins) or white arrows (non-secreted proteins) represent gene orientation. Arrow length is proportional to gene length with the length of the scale bar representing 300 nucleotides. Genes are labelled with their respective FGSG identifiers.

doi:10.1371/journal.pone.0033731.g002
published [25], predicted the *F. graminearum* genome to encode 103 plant cell wall degrading enzymes [52].

The sub-set of 102 MIPS annotated secreted proteins not predicted to function in the degradation of plant cells were organised according to their MIPS functional category and scrutinised further (Table S4). Within this diverse selection of proteins, those involved in metabolism accounted for the greatest proportion (43%; Figure 3). These included acid/alkaline phosphatases, alcohol oxidases, a salicylate hydroxylase and four extracellular nucleases (Table 2).

The other MIPS functional categories that were highly represented within the annotated sub-set were proteins with binding functions or cofactor requirements and cell rescue, defence and virulence. Two fungal proteins similar to plant pathogenesis related (PR) proteins PR1 in *Nicotiana tabacum* and PR3K pathogenesis-related thaumatin family protein in *A. thaliana* were identified. The only protein to possess both nuclear export and localisation signals was FGSG_04685. This protein contained a dioxygenase domain that incorporates O2 into an unknown substrate. Two *F. graminearum* secreted proteins were predicted to be phytotoxic and were highly related (6e−40 and 9e−37) to the *S. nodorum* phytotoxin, Snodprot1 [53]. Four other *F. graminearum* secreted proteins were predicted to possess antifungal properties and three of these were related to the KP4 killer toxin from *U. maydis* [54]. Two detoxifying lactonohydrolases were also identified.

Despite the stringency of the requisites of the predicted *F. graminearum* secretome several proteins believed to be intracellular were present. Two orthologs of *GEG16* from the powdery mildew fungi *B. graminis* f. sp. *hordei* and Gas1 and Gas2 from *M. oryzae* that function in pathogenicity and penetration [55] were predicted to be secreted. However, fluoroescently labelled Gas proteins localised to the cytosol of *M. oryzae* appressoria implying that the two *F. graminearum* *GEG16* homologues may not be secreted [55]. The trichothecene 3-O esterase code by *TRI8* (FGSG_03592), was also predicted to be secreted, but this biosynthetic enzyme is not detected in culture filtrates of *F. sporotrichioides* [56] suggesting that it is also not extracellular in *F. graminearum*.

**Analysis of the proteins with no predicted function**

The remaining 295 predicted secreted proteins that lacked annotation were analysed for internal amino acid repeats, high cysteine content and Pfam domains. The majority of these proteins, 190, are conserved. Protein functional domains can be predicted by sequence similarity. Using the Pfam database (http://pfam.sanger.ac.uk/) a total of 92 functional protein domains were found to be present within 73 secreted proteins, including 11 proteins with multiple Pfam domains (Table S5). The most abundant domain is Pfam04616 that belongs to the glycoside hydrolase 43 family (n = 5), which have been reported to have arabinofuranosidases, arabinamase and xylosidase activity. Other common domains included; a GDSL-like lipase domain (Pfam00675) involved in lipid metabolism (n = 4), a beta-lactamase domain (Pfam01444) associated with antibiotic resistance (n = 4), a nuclease/phosphatase family domain (Pfam0372) involved in intracellular signalling (n = 3), a necrosis inducing protein domain (Pfam05630) similar to the NPP1 protein from *P. infestans* [57] (n = 3), and a carbohydrate-binding domain (Pfam10528) found in fungal adhesins (n = 3).

The unannotated *F. graminearum* secreted proteins were screened at two thresholds, where the total number of cysteine residues represented greater than 5 or 10% of the mature protein. The majority (80%) of these predicted proteins had an even number of cysteine residues (Cys). At the >5% Cys threshold, 61 proteins were identified (31 conserved and 30 hypothetical proteins) (Table S6), whilst at the >10% Cys threshold, 11 proteins were identified (7 conserved and 4 hypothetical proteins) (Table 3). Four of these FGSG genes were not identified in the other *Fusarium* species (Table 3). Orthologs of the 29 *F. graminearum* proteins with 4, 6 or 8 cysteine residues were determined using BLASTP. Of these 29 *F. graminearum* proteins 20 were conserved amongst the Fusaria while seven had no strong hits. Eight, five and three proteins were conserved in other saprophytes (*A. flavus* species and *Neurospora crassa*), necrotrophs (*Botrytis cinerea*, *Cochliobolus heterostrophus*, *Sclerotinia sclerotiorum*, *S. nodorum* and *Pyrenophora tritici-repentis*) and hemibiotrophs (*M. oryzae*, *P. infestans* and *Verticillium albo-atrum*), respectively. The locus FGSG_03969 was atypical, coding for a somewhat larger mature protein at 482 amino acids in length of which 58 (12%) were cysteine residues and also contained 13 internal amino acid repeats.

All 574 sequences were inspected for the presence of the degenerative RxLR-dEER [1] and Y/F/WxG motifs [5] in close proximity to the predicted signal peptide sequence. No exact RxLR-dEER matches were found within the refined *F. graminearum* secretome. By contrast, a YsC motif was present in close proximity
to the predicted signal peptide cleavage site in five proteins. These were three conserved hypothetical proteins (FGSG_00260, FGSG_01815 and FGSG_03050), an endoglucanase (FGSG_02658) and a hypothetical protein (FGSG_13505). The unannotated portion of the predicted secreted proteins were screened for the presence of both perfect and imperfect internal repeats using the RADAR software [50]. This analysis identified 28 proteins ranging from 297–1862 amino acids in length of which five were also cysteine-rich (>5%). Multiple copies of nine of these proteins were found by BLASTP at two different thresholds, four at e\(^{-100}\) and nine at e\(^{-40}\) (Table 3). The 28 secreted proteins of *F. graminearum*, which were predicted to contain internal repeats, were highly conserved. For example; *F. oxysporum* f. sp. *lycopersici* possessed 19, *F. solani* 21, *F. verticillioides* 22 and *M. graminicola* 22 orthologous proteins (Table 3).

Gene family size for each *F. graminearum* encoded secreted protein of unknown function was determined by BLASTP. As anticipated, the majority of the *F. graminearum* gene families were larger in the other *Fusarium* species (*F. oxysporum* (43 copies) and *F. solani* (23 copies) as well as other saprophytic and soil dwelling organisms including 26 to 52 copies in Trichoderma spp., 30 copies in *A. nidulans*, and 22 copies in Chaetomium globosum. The conserved hypothetical, FGSG_08958, contained a nucleoside phosphorylase domain. Four genes were dramatically expanded in Phytophthora species with up to 19 copies identified (p<e\(^{-40}\)). Ten copies of a single gene FGSG_03708 were detected in *M. graminicola* and *M. fijiensis* (p<e\(^{-40}\)) (Spreadsheet S1).

**Comparison of the predicted *F. graminearum* secretome with a broad range of fungal and oomycete species**

A total of 57 genomes covering animal/plant pathogens, saprophytes and free living eukaryotic microbes were assembled. This list included 44 fungal and oomycete species (Table S1). The objective of this part of the study was to identify the *F. graminearum* specific secreted proteins, and the level of gene sequence

---

**Figure 3. The sub-set of MIPS annotated secreted *F. graminearum* proteins, which are predicted not to be involved in the degradation of plant cells.** The proteins were organised according to their MIPS functional category (http://mips.helmholtz-muenchen.de/genre/proj/FGDB/). N = 102.

doi:10.1371/journal.pone.0033731.g003
was conserved in the closely related saprophyte
66.03% of genes conserved, while only 56.97% of the secretome
most similarity to the
verticillioides
demonstrated 88.5% and 88% conservation, respec-
F. oxysporum
dissimilar to the
F. graminearum
showed the
S. nodorum
secretome was well conserved beyond
F. graminearum
secretome was most well conserved within the nine saprophytic
species obtained nutrition from dead plant material (78.05%).
the majority of the F. graminearum secretome was detected in all
species assessed (78.05%), while 5.4% of the
secretome was unique to these four Fusarium species (Table 4)
including, 22 hypothetical proteins, eight conserved hypothetical
proteins and a related cell wall mannoprotein. An additional
3.31% of the secretome was only found in
secretome was unique to these four
Fusarium
specificities. Each genome was screened for the presence of
F. graminearum
secretome homologues (Spreadsheet S1). The results
from this BLASTP analysis are reported at two levels of stringency,
however for clarity the results obtained at a p value < e-5 are
Table 2. A selection of MIPS annotated secreted F. graminearum proteins, which are not involved in the degradation of plant cells, but are associated with metabolism or pathogenicity.

| Function                  | FGSG_ID   | MIPS Annotation                                      |
|---------------------------|-----------|------------------------------------------------------|
| Acid/alkaline phosphatases| FGSG_04504| Related to acid phosphatase precursor                |
|                           | FGSG_05933| Related to acid phosphatase precursor                |
|                           | FGSG_06610| Related to alkaline phosphatase D precursor          |
|                           | FGSG_07608| Related to acid phosphatase precursor                |
|                           | FGSG_07678| Related to acid phosphatase Pho610                  |
| Salicylate hydroxylase    | FGSG_08116| Related to salicylate hydroxylase                    |
| Extracellular nucleases   | FGSG_02686| Related to ribonucleases                             |
|                           | FGSG_03379| Related to ribonucleases                             |
|                           | FGSG_11190| Probable ribonuclease T1                             |
|                           | FGSG_15003| Related to dnaase1 protein                           |
| Phospholipases            | FGSG_08150| Related to PLB1 - phospholipase B (lyso phospholipase)|
|                           | FGSG_11236| Related to non-hemolytic phospholipase C precursor   |
| PR-like proteins          | FGSG_03109| Related to plant PR-1 class of pathogen related proteins|
|                           | FGSG_08549| Related to pathogenesis-related protein PRSK (thauamatin family)|
| Antifungal proteins       | FGSG_00060| Related to KP4 killer toxin                         |
|                           | FGSG_00061| Related to KP4 killer toxin                         |
|                           | FGSG_00062| Related to KP4 killer toxin                         |
|                           | FGSG_04745| Related to antifungal protein                        |
| Pathogenicity related     | FGSG_00006| Related to gEgh 16 protein                          |
|                           | FGSG_09353| Related to gEgh 16 protein                          |
| Phytotoxin                | FGSG_10212| Probable SnodProt1 precursor                        |
|                           | FGSG_11205| Probable SnodProt1 precursor                        |
| Detoxification            | FGSG_03816| Probable lactonohydrolase                           |
|                           | FGSG_10675| Related to lactonohydrolase                         |

The basidomycete biotroph U. maydis demonstrated 38.85%
conservation, more than the ascomycete obligate biotroph B.
graminis f. sp. hordei at 30.14%. The oomycete hemibiotroph
P. infestans demonstrated 32.4% conservation. The yeast secretomes,
including the animal pathogen Candida albicans at 17.07%, and
non-pathogen S. cerevisiae at 14.81%, were less well
conserved.
Conservation of the F. graminearum secretome was subsequently
determined for the different sub-sets of species depending on host
tissue specificity including, animal, plant and cereal ear or cereal
leaf infecting pathogens (Table S1). The other Fusaria genomes
were excluded from these analyses. Conservation of the F.
graminearum secretome among 13 animal pathogens was surpris-
ingly high at 68.64%, however this was still lower than the
conservation with plant and cereal ear or cereal leaf infecting pathogens at 80.14%, 75% and 76.13%, respectively.
The genomes of plant interacting organisms were divided
to their mode of colonisation. The F. graminearum
secretome was most well conserved within the nine saprophytic
species that obtain nutrition from dead plant material (78.05%).
Conservation of the F. graminearum secretome among the seven
hemibiotrophs and seven necrotrophs was also high at 76.13% and
73.52%, respectively. The only class of plant pathogons within
which the F. graminearum secretome was poorly conserved was the
seven biotrophs (54.53%). This figure is substantially lower than
the level of conservation with the 13 animal pathogens. This result
we consider being somewhat artifactual and has been caused by

doi:10.1371/journal.pone.0033731.t002
The Fusarium graminearum Secretome
**Table 3.** Analysis of *F. graminearum* secretome proteins and genes with no predicted function for cysteine content, the presence of internal amino acid repeats, Pfam domains, gene family size and presence in other Fusaria species and *Mycosphaerella graminicola*.

| Locus_ID    | MPL* | NC  | %MP | RR    | Pfam | Twin e-40 | Fv     | Fol            | Fs             | Mg             |
|-------------|------|-----|-----|-------|------|-----------|--------|----------------|----------------|----------------|
| FGSG_00260  | 60   | 6   | 10  | -     | -    | 0        | 1.11e-28 | -              | 5.32e-23       | 9.8e-20        |
| FGSG_03969  | 482  | 58  | 12.03 | 13 | -    | 0 | 4.14e-141 | 5.25e-158 | 8.65e-106 | 2.31e-97       |
| FGSG_03599  | 77   | 10  | 12.99 | 3  | -    | 0 | 3.93e-21  | 5.20e-27  | -              | -              |
| FGSG_06712  | 129  | 16  | 12.4  | 3   | -    | 0 | 1.12e-27  | 3.01e-27  | 5.56e-28       | 6.08e-14       |
| FGSG_09066  | 64   | 8   | 12.5  | 2   | -    | 0 | 4.00e-14  | 6.05e-15  | 1.56e-12       | 1.07e-22       |
| FGSG_12214  | 79   | 10  | 12.66 | 2   | -    | 0 | 5.23e-47  | 2.37e-47  | -              | -              |
| FGSG_15142  | 71   | 8   | 11.27 | -   | -    | 0 | -         | -              | -              |
| FGSG_15437  | 53   | 8   | 15.09 | -  | -    | 0 | -         | -              | -              |
| FGSG_15448  | 53   | 8   | 11.27 | -  | -    | 0 | -         | -              | -              |
| FGSG_15448  | 53   | 8   | 11.27 | -  | -    | 0 | -         | -              | -              |
| FGSG_15448  | 53   | 8   | 11.27 | -  | -    | 0 | -         | -              | -              |

The *Fusarium graminearum* Secretome

PLoS ONE | www.plosone.org 9 April 2012 | Volume 7 | Issue 4 | e33731
the underrepresentation of ascomycete species within the biotroph sub-set. The only genome that is currently available for an ascomycete species that has a biotrophic lifestyle is B. graminis f. sp. hordei. This species forms abundant intracellular haustoria. The other six species of biotrophs were either basidiomycetes or oomycetes, whereas in the animal pathogens examined there were other six species of biotrophs were either basidiomycetes or oomycetes, whereas in the animal pathogens examined there were 10 ascomycetes. The two non-pathogens, S. cerevisiae and Schizosaccharomyces pombe, demonstrated a very poor level of conservation (22.13%) with 103 genes were considered to be F. graminearum specific and 179 genes were considered to be Fusarium specific.

**Comparison of the predicted secretome with published F. graminearum proteomic data sets**

A proteomic comparison of the secretome from F. graminearum grown in vitro and in planta, identified 122 extracellular proteins and according to Paper and colleagues [58], 68 of these proteins possessed a signal peptide. Only 14 of the 68 proteins identified were detected exclusively in planta, and these included a metallopeptidase, a KP4 killer toxin, a pectin lyase and an endoglucanase. A total of 68% of the proteins found in the proteomic study were also identified in our predicted secretome, which was designed to be extra stringent. In the proteomic study, nine proteins detected and found to have a signal peptide, were excluded from the bioinformatically predicted secretome generated by this study due to the stringency of the combined SignalP, TargetP and WollPSort analysis. The majority of the 68 proteins not detected in the predicted secretome were excluded during the WollPSort analysis. A cut off score of 18 had been used. For all 68 proteins to have been included a far lower WollPSort cut off score would have needed to have been used, and this would have raised considerably the potential number of false positives included within these analyses. Even when the WollPSort score was lowered to 17, this included three proteins where the probability scores from SignalP were only modest. The other 46 proteins detected in the proteomic study, which lacked signal peptides were all excluded from this detailed analysis by the WollPSort analysis. The authors of the proteomic study [58] concluded that the detection of these proteins may have arisen, because the Fusarium cell ruptured during sample preparation. For example, NADP-dependent oxidoreductase and elongation factor 1 are not known in other species to be extracellularly located.

**Discussion**

Communication through the secretion of proteins and metabolites frequently defines the outcome of the interaction between a host and a fungal symbiont, irrespective of their lifestyle [7,59,60]. During the formation of Fusarium Ear Blight disease an intimate host-pathogen association develops and an extended growth phase occurs in the apoplast, which is extracellular to the living wheat cells [44]. In the original analysis of the newly sequenced F. graminearum genome, only TargetP was used to predict the secretome [25]. Since this time, the gene call for F. graminearum has been considerably changed through the combined efforts of the BROAD and MIPS. In the current study, a refined F. graminearum secretome was predicted by the combination of multiple bioinformatic approaches. This strategy increased the probability of identifying truly secreted proteins. A secretome size of 574 proteins is predicted for F. graminearum, representing 4.2% of the predicted total gene repertoire. The cell biology of the different phases of wheat ear infection depicts a situation where Fusarium hyphae are exposed to different environments/distinct substrates, thereby causing transcriptional, proteinaceous and metabolic changes. Fusarium hyphae in the symptomless phase of infection are in close contact with live plant cells for two to three days [44]. During this prolonged latent period, communication between pathogen and host must occur. After several days the wheat cells die and are intracellularly colonised by the pathogen resulting in the development of visible disease symptoms and asexual sporulation. Transcriptional differences between the two phases of infection have been confirmed for the biosynthetic genes responsible for the virulence factor DON, which showed maximal TRI gene expression during symptomless infection [38].

**Table 3.** Conservation of the F. graminearum (Fg) genes, predicted to encode secreted proteins, among the 57 fungal genomes assessed and the presented according to overall species distribution, host range or lifestyle.

| Sub-sets | All secreted proteins | All secreted proteins |
|----------|----------------------|----------------------|
|          | e-5                  | e-40                 |
| Total gene number | 574          |                     |
| Fg specific      | 19 (3.31%)      | 103 (17.94%)       |
| All Fus          | 448 (78.05%)    | 365 (63.59%)       |
| Unique to all Fus| 31 (5.4%)       | 179 (31.19%)       |
| Fg + other non Fus| 482 (83.97%)  | 395 (68.82%)       |
| Animal pathogen  | 394 (68.64%)    | 289 (50.35%)       |
| Plant pathogen   | 460 (80.14%)    | 372 (64.81%)       |
| Cereal ear pathogen| 419 (73%)   | 337 (58.71%)       |
| Cereal leaf pathogen| 437 (76.13%) | 344 (59.93%)       |
| Budding yeast    | 127 (22.13%)    | 35 (6.1%)          |
| Biotroph         | 313 (54.53%)    | 156 (27.18%)       |
| Hemibiotroph     | 437 (76.13%)    | 361 (62.89%)       |
| Necrotroph       | 422 (73.52%)    | 330 (57.49%)       |
| Saprotroph       | 448 (78.05%)    | 354 (61.67%)       |

**Table 4.** Conservation of the F. graminearum (Fg) genes, predicted to encode secreted proteins, among the 57 fungal genomes assessed and the presented according to overall species distribution, host range or lifestyle.

| Sub-sets | All secreted proteins | All secreted proteins |
|----------|----------------------|----------------------|
|          | e-5                  | e-40                 |
| Total gene number | 574          |                     |
| Fg specific      | 19 (3.31%)      | 103 (17.94%)       |
| All Fus          | 448 (78.05%)    | 365 (63.59%)       |
| Unique to all Fus| 31 (5.4%)       | 179 (31.19%)       |
| Fg + other non Fus| 482 (83.97%)  | 395 (68.82%)       |
| Animal pathogen  | 394 (68.64%)    | 289 (50.35%)       |
| Plant pathogen   | 460 (80.14%)    | 372 (64.81%)       |
| Cereal ear pathogen| 419 (73%)   | 337 (58.71%)       |
| Cereal leaf pathogen| 437 (76.13%) | 344 (59.93%)       |
| Budding yeast    | 127 (22.13%)    | 35 (6.1%)          |
| Biotroph         | 313 (54.53%)    | 156 (27.18%)       |
| Hemibiotroph     | 437 (76.13%)    | 361 (62.89%)       |
| Necrotroph       | 422 (73.52%)    | 330 (57.49%)       |
| Saprotroph       | 448 (78.05%)    | 354 (61.67%)       |
secretion of DON is hypothesised to inhibit the plant’s ability to respond to infection by impeding protein synthesis [61]. Mechanisms in addition to DON mycotoxin may also be required to promote infection, implicating a role for the *F. graminearum* secretome [25]. Therefore, an in depth re-analysis of the secretome’s capabilities was undertaken.

The comparative genomics analysis of 57 fungal and oomycete genomes revealed a high level of secretome conservation among filamentous ascomycetes, irrespective of their mode of obtaining nutrition from plant or animal hosts or during a free living lifestyle. This high level of secretome conservation may reflect the ability of *F. graminearum* to survive both as a pathogen and as a saprophyte. The identification of 31 *Fusarium* specific and 25 *F. graminearum* specific secreted proteins, of which all were functionally unannotated proteins, may represent the conserved and unique protein-protein interactions that assist *Fusarium* pathogenicity. The predicted *F. graminearum* secretome, with a size of 1,569 from the initial analysis and 574 from the refined selection (Figure 1) appears larger than the *B. graminis* (248), and *U. maydis* (426) secretomes, but possibly slightly smaller than that of *M. oryzae* (739) [31,52,62]. However, the size of these fungal secretomes was predicted using slightly different approaches. Despite representing a large fungal secretome, the refined set of *F. graminearum* secreted proteins demonstrated less species specificity than the biotrophic pathogens *B. graminis* and in particular *U. maydis* where two thirds of the secreted proteins are species specific.

As previously noted, the *F. graminearum* secretome predominantly localises to hot spots of chromosomal recombination and subtelomeric regions [25] facilitating alterations to the secretome that could enable the pathogen to cope with changes in the host plant response. However, some genes predicted to code for a secreted protein were located in the intervening low or no recombination regions found on each of the four chromosomes. In addition, several small clusters of secreted proteins (ranging from 3–6 genes) were identified in regions of the genome located with either a low or a high level of recombination and in both sub-telomeric and more central locations. Unlike *U. maydis* [52] these *F. graminearum* clusters did not contain genes of similar function and did not represent gene duplication events.

In total 99% of the refined bioinformatic prediction of the *F. graminearum* secretome was supported by transcriptional evidence. During early *F. graminearum* infection *TRI* gene expression is up-regulated at the advancing hyphal front in the florets [63] and the rachis tissue [38]. Along with the array of secreted proteins, DON may inhibit the plant cells ability to detect or respond to infection. The small cysteine-rich secreted proteins, of which many contained internal amino acid repeats, and the additional pathogenicity related protein similar to a circumsperozoite that has been shown to inhibit protein synthesis in cells infected by malaria *Plasmodium* parasites [64] may also play a role in establishing wheat infection. An ability to obtain nutrition from the apoplast and possibly inhibit, or circumvent, plant defences is in agreement with the observed lack of physiological changes to the plant cells during this initial phase of infection [44].

After a latent period of infection wheat host cells die prior to, or at the same time as, *F. graminearum* hyphae penetrate host cells en masse [44]. Whether host cell death is induced by the plant in an attempt to limit infection, or by the fungus to obtain nutrition, remains unknown. The two small secreted proteins, related to Snodprot1 from *S. nodorum* that has proven phytotoxin activity [65] also contain the cerato-platanin Platt07249 domain. In *Ceratocystis fimbriata*, the Snodprot1 protein exists in the fungal cell wall and has been shown to induce host cell phytoalexin synthesis as well as necrosis [66]. In *M. oryzae* the Snodprot1 homologue is required for full virulence [67]. Several *F. graminearum* SNODPROT1 homologues were identified in the secretome that have been demonstrated to be transcribed during wheat ear [68]. From the extracellular location, the possible phytotoxins activity of these two small *F. graminearum* secreted proteins may play a role in the induction of host cell death. In wheat, infiltration of high concentrations of DON mycotoxin into healthy leaves has been shown to elicit hydrogen peroxide production and programmed cell death [69]. Therefore, the level or length of exposure to the mycotoxin could also be involved in the induction of host cell death. Interestingly, several lactonohydrolases were predicted in the secretome. A novel lactonohydrolase cloned from *Chonostalathes roseus* into *S. pombe* or *Escherichia coli* was able to detoxify the trichotheccene mycotoxin, zearalenone [70,71]. The localised secretion of a lactonohydrolase by *F. graminearum*, may therefore act as a self defence mechanism, in addition to the experimentally proven Tri101 protein [72].

Once within dead plant tissue *F. graminearum* is predicted to secrete an array of PCWDEs and other enzymes, far more than many other fungal pathogens [25,73]. *F. graminearum* also appears to possess the capacity to utilise the plant cell in its entirety, which is in agreement of the observed phenotype of wheat rachis infection [44]. This ability of the secretome to breakdown the plant cell is probably essential for *F. graminearum* pathogenesis, but will be difficult to test experimentally because of the problem of genetic redundancy. The extensive repertoire of PCWDEs would also assist in the saprophytic phase of the *F. graminearum* lifecycle, which occurs post-harvest [74].

*F. graminearum* may be able to produce a range of antifungal proteins, including FGSG_04745 and four KP4 killer toxins. Their production could prevent additional colonisation by fungal competitors and protect the niche the *F. graminearum* hyphae have occupied. The trichothecene mycotoxins may also have some antifungal activity [75]. The *U. maydis* KP4 killer toxins provide antifungal activity by blocking calcium uptake thereby interfering with calcium signalling [76]. The increased production of antifungal proteins may be essential during late infection, reflecting the vulnerability of the dead plant tissue to further microbial colonisation.

The functional analysis of the secretome revealed the presence of a large set of extracellular proteins with a function in metabolism. This suggests *Fusarium* hyphae can manipulate or directly interfere with the plant’s metabolism. Acid and alkaline phosphatases are responsible for protein dephosphorylation, which is pivotal to cell signalling. Alcohol oxidases catalase the reaction between alcohol and O2 releasing an aldehyde and H2O2, which is an important plant signalling molecule. Salicylate hydroxylase is capable of degrading the plant defence signalling molecule, salicylic acid, which has been shown to be required for maintaining basal defence against *Fusarium* in the floral tissues of Arabidopsis [77,78] while a delay in salicylic acid signalling has also been associated with increased *Fusarium* susceptibility in wheat ears [79]. The extracellular nucleases indicate the potential to degrade DNA/RNA or interfere with nucleic acid function. Extracellular proteins involved in protein-binding were also highly represented.

Plant PR proteins are rapidly expressed upon the perception of pathogen attack [80]. The secretion by *Fusarium* of related PR proteins, such as PR1 and PR5K is intriguing. The *F. graminearum* PR1-like protein is conserved in *F. oxysporum*, *F. verticillioides*, *F. solani* and *M. oryzae*, while the PR5K-like protein in addition to the aforementioned species is widely conserved in *S. nodorum*, *S. sclerotiorum*, *P. triticis repentinis*, *Leptosphaeria maculans*, *T. reesei* and *N.
Refinements to the predicted secretome continue to evolve through the increased use of manual sequence induced host responses. The gene models for deep-RNA sequencing approach would be ideal. The later would separates the different phases of infection will be required. The use host species and different tissues, genome wide investigations [81]. However, in the M. graminicola wheat leaf interaction, which also switches from symptomless to symptomatic infection, the only NPP1 homologue was not required for full virulence [82]. Neither the consensus nor degenerative RxLR motif situated in the N terminus of the predicted proteins was identified in the refined secretome studied here. However, the five secreted protein identified with the YxC motif in close proximity to the signal peptide represents an interesting find that requires further investigation.

This study has greatly increased our understanding of the F. graminearum secretome and identified genes coding for secreted proteins that can be considered to be Fusarium conserved and F. graminearum specific. Once the genomic sequences of additional Fusaria species and strains and other fungal species are published, these secretome predictions can be further refined. In order to achieve a greater understanding of the transcriptional differences between the different phases of in planta infection in different plant host species and different tissues, genome wide investigations coupled with a synchronised biological assay that accurately separates the different phases of infection will be required. The use of the Fusarium Alyssum array [83] and/or a next generation deep-RNA sequencing approach would be ideal. The later would also give considerable information, in parallel, on the nature of the induced host responses. The gene models for F. graminearum continue to evolve through the increased use of manual sequence corrections [26]. This activity is likely to lead to further refinements to the predicted F. graminearum secretome.

Supporting Information

Table S1 The fungal and oomycete genomes included within the 57 species analysis.

Table S2 The sub-set of F. graminearum genes that code for secreted proteins involved in the degradation of the plant cuticle and cell wall, divided according to substrate specificity.

Table S3 The sub-set of F. graminearum genes that code for secreted proteins involved in the degradation of the plant cell, divided according to substrate specificity (starch, lipid and protein).

Table S4 The sub-set of F. graminearum genes that code for secreted proteins with no MIPS annotation, but contain conserved protein functional (Pfam) domains.

Table S5 The sub-set of F. graminearum genes that code for cysteine –rich (>5%) unannotated secreted proteins.

Table S6 The sub-set of F. graminearum genes that code for secreted proteins with no MIPS annotation, but contain conserved protein functional (Pfam) domains.

Spreadsheet S1

Acknowledgments

We would like to thank Alexandre Morais do Amaral, John Lucas, Jason Rudd, Martin Urban and Mansoor Saqi for commenting on earlier drafts of the manuscript. Martin Urban is also thanked for helping with the preparation of the accepted manuscript for final submission.

Author Contributions

Conceived and designed the experiments: NAB JA KEHK. Performed the experiments: NAB JA KEHK. Analyzed the data: NAB JA KEHK. Contributed reagents/materials/analysis tools: NAB JA KEHK. Wrote the paper: NAB JA KEHK.

References

1. Tyler BM (2009) Entering and breaking: virulence effector proteins of oomycete plant pathogens. Cellular Microbiology 11: 13–20.
2. Bhattacharjee S, Hiller NL, Lados K, Wu J, Kanneganti TD, et al. (2006) The malarial host-targeting signal is conserved in the Irish potato famine pathogen. PLoS Pathogens 2: 453–465.
3. Kale SD, Gu BA, Capelotto DGS, Dou DL, Feldman E, et al. (2010) External lipid PI3P mediates entry of eukaryotic pathogen effectors into plant and animal host cells. Cell 142: 284–295.
4. Deller S, Hammond-Kosack KE, Rudd J (2011) The complex interactions between host immunity and non-biotrophic fungal pathogens of wheat leaves. Journal of Plant Physiology 168: 63–71.
5. Godfrey D, Bohlenius H, Pedersen C, Zhang Z, Ennemersen J, et al. (2010) Powdery mildew fungal effector candidates share N-terminal Y/F/WxC-motif. BMC Genomics 11: 317.
6. Stergiopoulou I, de Wit PJGM (2009) Fungal Effector Proteins. Annual Review of Phytopathology 47: 253–263.
7. Tan KC, Oliver RP, Solomon PS, Moë LA CS (2010) Proteinaceous necrotrophic effectors in fungal virulence. Functional Plant Biology 37: 907–912.
8. Ridout CJ, Skamnioti P, Porritt O, Sacrison S, Jones JDG, et al. (2006) Multiple avirulence paralogues in cereal powdery mildew fungi may contribute to parasitic fitness and defeat of plant resistance. Plant Cell 18: 2402–2414.
9. Lievens R, Houterman PM, Rep M (2009) Effector gene screening allows unambiguous identification of Fusarium oxysporum f. sp. lycopersici races and discrimination from other formae speciales. FEMS Microbiology Letters 300: 201–215.
10. Valent B, Kiang CH (2010) Recent advances in rice blast effector research. Current Opinion in Plant Biology 13: 434–441.
11. Dochleimer G, van der Linde K, Assmann D, Schwambach D, Hof A, et al. (2009) Pep1, a secreted effector protein of Ustilago maydis, is required for successful invasion of plant cells. PLoS Pathogens 5(7): e1000290. doi: 10.1371/journal.ppat.1000290.
12. Marshall R, Kombrik A, Motteram J, Loza-Reyes E, Lucas J, et al. (2011) Analysis of two in planta expressed LysM effector homologs reveals novel functional properties and varying contributions to virulence on wheat. Plant Physiology 156: 756–769.
13. Kloppholz S, Kohan H, Requena N (2011) A secreted fungal effector of Glomus intraradices promotes symbiotic biotrophy. Current Biology 21: 1204–1209.
14. van Esse HP, Bolton MD, Stergiopoulou L, de Wit PJGM, Thonmina BPJ (2007) The chitin-binding Cladosporium fulvum effector protein Avr1 is a virulence factor. Molecular Plant-Microbe Interactions 20: 1102–1110.
15. de Jonge RS, van Esse HP, Kombrik A, Shinya T, Desaki Y, et al. (2010) Conserved fungal LysM effector Ep6 prevents chitin-triggered immunity in plants. Science 329: 953–955.
16. Verstrepen KJ, Jansen A, Lewitter F, Funk GR (2005) Intragenic tandem repeats generate functional variability. Nature Genetics 37: 986–990.
31. Spanu PD, Abbott JC, Amselem J, Burgis TA, Soanes DM, et al. (2010) Genome
32. Proctor RH, Hohn TM, McCormick SP, Desjardins AE (1995)
33. Proctor RH, Hohn TM, McCormick SP (1995) Reduced virulence of
27. Beacham A, Antoniw J, Hammond-Kosack KE (2009) A genomic fungal foray.
25. Cuomo CA, Gueldener U, Xu JR, Trail F, Turgeon BG, et al. (2007) The
36. Maier FJ, Miedaner T, Hadeler B, Felk A, Salomon S, et al. (2006) Involvement
35. Harris LJ, Desjardins AE, Plattner RD, Nicholson P, Butler G, et al. (1999)
24. Gale LR, Bryant JD, Calvo S, Giese H, Katan T, et al. (2005) Chromosome
21. Dublin HJ, Glichrist L, Reeves J, McNab A (1997) Fusarium Head Scab: Global
40. Dorman CJ, Corcoran CP (2009) Bacterial DNA topology and infectious
20. Skibbe DS, Doehlemann G, Fernandes J, Walbot V (2010) Maize tumors caused
19. Iqbal K, Desjardins AE, Fournier L, Brown AJ, Eberhard MG, et al. (2008) The
18. Nizamuddin Z, Fournier L, Snow R, Brown AJ (2008) Transduction of
17. Levdan D, Romano J, Shadkhan Y, Sharon H, Verstrepen KJ, et al. (2007)
16. Parakkal P, Blin N, Chou M, Saha S, Rives J, et al. (2009) A novel
15. Kidor F, Desjardins AE, Brown AJ (2007) The C2/C2 domain of a small-molecule
14. Brown NA, Urban M, van de Meene AM, Hammond-Kosack KE. (2010) The
13. Müller O, Schreier PH, Uhrig JF (2008) Identification and characterization of
12. Russinovitch Y, Shiller N, Bohnert H, Zilberman D, Lobkowitz A, et al. (2005)
11. Desjardins AE, Fournier L, Snow R, Brown AJ, Arroyo M, et al. (2008) The
10. Desjardins AE, Fournier L, Snow R, Brown AJ, Arroyo M, et al. (2008) The
9. Desjardins AE, Fournier L, Snow R, Brown AJ, Arroyo M, et al. (2008) The
8. Desjardins AE, Fournier L, Snow R, Brown AJ, Arroyo M, et al. (2008) The
7. Desjardins AE, Fournier L, Snow R, Brown AJ, Arroyo M, et al. (2008) The
6. Desjardins AE, Fournier L, Snow R, Brown AJ, Arroyo M, et al. (2008) The
5. Desjardins AE, Fournier L, Snow R, Brown AJ, Arroyo M, et al. (2008) The
4. Desjardins AE, Fournier L, Snow R, Brown AJ, Arroyo M, et al. (2008) The
3. Desjardins AE, Fournier L, Snow R, Brown AJ, Arroyo M, et al. (2008) The
2. Desjardins AE, Fournier L, Snow R, Brown AJ, Arroyo M, et al. (2008) The
1. Desjardins AE, Fournier L, Snow R, Brown AJ, Arroyo M, et al. (2008) The
72. McCormick SP, Alexander NJ, Trapp SE, Hohn TM (1999) Disruption of TRI101, the gene encoding trichothecene 3-O-acetyltransferase, from Fusarium sporotrichioides. Applied and Environmental Microbiology 65: 5252–5256.
73. Soanes DM, Richards TA, Talbot NJ (2007) Insights from sequencing fungal and oomycete genomes: What can we learn about plant disease and the evolution of pathogenicity? Plant Cell 19: 3318–3326.
74. Parry DW, Jenkinson P, McLeod I (1995) Fusarium ear blight (scab) in small grain cereals - a review. Plant Pathology 44: 207–238.
75. Lutz MP, Feichtinger G, Defago G, Duffy B (2003) Mycotoxigenic Fusarium and deoxynivalenol production repress chitinase gene expression in the biocontrol agent Trichoderma atroviride P1. Applied and Environmental Microbiology 69: 3077–3084.
76. Gage MJ, Brunn J, Fischer M, Sanders D, Smith TJ (2001) KP4 fungal toxin inhibits growth in Ustilago maydis by blocking calcium uptake. Molecular Microbiology 41: 773–785.
77. Cuzick A, Lee S, Gezan S, Hammond-Kosack KE (2008) NPR1 and EDS11 contribute to host resistance against Fusarium culmorum in Arabidopsis buds and flowers. Molecular Plant Pathology 9: 697–704.
78. Makandar R, Nalam V, Chaturvedi R, Jeannotte R, Sparks AA, et al. (2010) Involvement of salicylate and jasmonate signaling pathways in Arabidopsis interaction with Fusarium graminearum. Molecular Plant-Microbe Interactions 23: 861–870.
79. Ding LN, Xu HB, Yi HY, Yang LM, Kong ZX, et al. (2011) Resistance to hemibiotrophic F. graminearum infection is associated with coordinated and ordered expression of diverse defense signaling pathways. PLoS ONE 6(4): e19008. doi: 10.1371/journal.pone.0019008.
80. Sel J, Mathys J, de Coninck BMA, Cammue BPA, De Bolle MFC (2008) Plant pathogenesis-related (PR) proteins: A focus on PR peptides. Plant Physiology and Biochemistry 46: 941–950.
81. Qutob D, Kamoun S, Gijzen M (2002) Expression of a Phytophthora sojae necrosis-inducing protein occurs during transition from biotrophy to necrotrophy. The Plant Journal 32: 361–373.
82. Motteram J, Kuffer I, Deller S, Brunner F, Hammond-Kosack KE, et al. (2009) Molecular characterization and functional analysis of MgNLP, the sole NPP1 domain-containing protein, from the fungal wheat leaf pathogen Mycosphaerella graminicola. Molecular Plant-Microbe Interaction 22: 790–799.
83. Gueldener U, Seong KY, Bodhi J, Cho S, Traif F, et al. (2006) Development of a Fusarium graminearum Affymetrix GeneChip for profiling fungal gene expression in vitro and in planta. Fungal Genet Biol 43: 316–325.