Plant pathogens: how can molecular genetic information on plant pathogens assist in breeding disease resistant crops

Kim Hammond-Kosack[3], Martin Urban, Thomas Baldwin, Arsalan Daudi, Jason Rudd, John Keon, John Lucas, Kerry Maguire, Dmitry Kornyukhin[2], Hai-Chun Jing, Chris Bass and John Antoniw

Wheat Pathogenesis Programme, Plant-Pathogen Interactions Division, Rothamsted Research, Harpenden, Herts AL5 2JQ, UK
[2] N.I. Vavilov Research Institute of Plant Industry (VIR), 44 Great Morskaya Street, St. Petersburg 190000, Russian Federation
[3]Corresponding author. Fax +44(0) 1582 715981, E-mail address: kim.hammond-kosack@bbsrc.ac.uk

Abstract

Numerous phytopathogenic fungi and Oomycete species cause disease on crop plants. Over the past 17 years a combination of protein purification and forward and reverse genetics has lead to the identification of > 150 distinct pathogenicity and virulence genes required by pathogens to cause disease. Many conserved molecular mechanisms are now recognised that underpin the pathogenesis of very diverse species. In addition, > 20 pathogen effectors have been characterised which activate either non-host or race-specific resistance (*R*) gene mediated plant defence responses. Recently, full genome sequencing projects have been completed for eight phytopathogenic eukaryotic organisms and others are in progress. Through comparative genome analyses, the gene inventory and genome synteny of animal and plant pathogenic fungi and Oomycete species are being compared with each other and with other non-pathogenic organisms. This is providing exciting new information on the evolution of plant pathogens. For plant breeders, this wealth of molecular genetic information provides the opportunity to develop a suite of new approaches to attain durable disease control. For example, many pathogen targets have been revealed which potentially are accessible for intervention via plant cell derived products. Also, there is now the unprecedented capacity to monitor changes in plant pathogen populations resulting from the introduction of different resistant plant germplasms. A case study involving *Fusarium graminearum* and *F. culmorum,* which cause ear blight disease on wheat and other small grain cereals, is presented and illustrates how crop breeding can be assisted by molecular genetic information on the pathogen.

Media summary

Gene inventories and gene function tests for organisms causing disease provide breeders with unique information to help identify and deploy new sources of durable resistance.

Keywords

Genomic sequence, pathogen mutant analysis, candidate gene disruption, durable disease resistance, pathogen population dynamics, *Gibberella zeae*

Introduction

Growing monocultures of genetically uniform crop species over vast tracts of land is a current practice that frequently leads to severe outbreaks of disease. Disease epidemics lower both crop yield and quality and can compromise the safety of the end product. Concerns over global food and feed security continue to grow. In temperate regions, biotic stress to crops caused by numerous fungal and Oomycete pathogenic species is a persistent problem. A **plant pathogen** is defined as an organism that, to complete a part or all of its life cycle, grows inside the plant and in so doing has a detrimental effect on the plant (Agrios, 1997). Fungal and Oomycete pathogens have evolved specialist ways to invade plants. Some penetrate the plant surface directly using mechanical pressure or enzymatic attack. Others pass through natural plant openings, for example, stomata or lenticels, whilst many take advantage of previous wounds. Once inside the plant three main colonisation strategies are deployed to utilise the host plant as a substrate for pathogen growth and development. Biotrophic organisms ensure the plant cell remains

alive, necrotrophic organisms kill plant cells in advance of colonisation, whilst hemibiotrophic organisms initially keep host plant cells alive but then kill them at later stages of the infection. **Pathogenesis** is the term used to describe the sequence of processes from host and pathogen contact to development of the complete syndrome. A pathogen strain that causes disease is termed **virulent** (Hammond-Kosack and Jones, 2000; Dickinson, 2003). Pathogenic organisms vary greatly in the extent to which the attacked plant is colonised. For example, the potato late blight pathogen *Phytophthora infestans* and the rice blast pathogen *Magnaporthe grisea* can infect nearly every plant organ with devastating consequences (Sesma and Osbourn, 2004). Other pathogens are restricted to just specific plant organs. For example, *Mycosphaerella graminicola* causes blotch disease solely on wheat leaves whilst the black stem rust pathogen *Puccinia graminis tritici* is restricted to leaf and stem tissue on cereal hosts and its alternate host plant, the barbarry bush (*Berberis vulgaris*). *Tapesia* spp infections cause eyespot lesions solely at stem base of cereals. In contrast *Fusarium oxysporum,* post root infection is restricted to the vascular tissues. For many pathogens, there is a long **latent period**, frequently weeks, between initial infection and the first onset of visible disease symptoms. Finally, some pathogens have a very broad host range, for example the grey mould fungus *Botrytis cinerea*, causes disease on hundreds of plant taxa whilst others attack only a single plant species, for example, some powdery mildews and rusts are restricted to specific small grain cereal hosts while *Cladosporium fulvum* causes leaf mould exclusively on tomato plants.

Controlling disease epidemics in agricultural crops has traditionally involved the integration of several approaches. In recent years, it has become commonplace to compare the various disease control options based not only on their overall economic cost versus the increased yield and/or superior end product quality obtained, but also in terms of energy consumption, time taken, environmental impact and overall sustainability. **Good crop husbandry** remains at the forefront of disease control and includes ensuring a diverse plant species rotation, the timely plough or burning of the previous season's infected crop residues, to reduce inoculum, and the early or late sowing of crop to avoid the maximum potential infection period. However, deep plough and stubble burning are increasingly recognised to have negative environmental impacts, for example, soil and water erosion and the release of greenhouse gases. Since the 1940's, the application of **fungicide chemistries** either as a seed dressing or to the growing crop has become increasingly important in the control of plant disease (Jones and Clifford, 1983; Agrios, 1997). The discovery of broad spectrum chemistries with systemic as well as curative modes of action was a major breakthrough in the 1970's and was responsible for a dramatic increase in fungicide use in high-input agricultural systems. The precision of chemical applications to crops has been greatly increased by disease forecasting, disease epidemic modelling, field disease diagnostics and applicator design. However, fungicide synthesis and formulation as well as the repeated application of fungicides to crops are each time-consuming processes, require the use of adequate protective clothing and are also reliant upon non-renewable energy sources. Therefore in most situations, fungicides should only be considered a short term solution to disease control. In addition, the emergence of pathogen strains resistant to specific chemistries frequently occurs just a few years after the introduction of a novel chemistry. The inability to predict the potential lifespan of a new chemistry is a growing concern. For example, resistance to strobilurin fungicides in wheat powdery mildew field isolates appeared within 2-3 years of the commercial launch of this novel chemistry in Europe. Subsequently, in 2002, five years after launch, isolates of another important wheat pathogen, *Mycosphaerella graminicola* with resistance to strobilurins were identified in the UK and Ireland (Fraaije et al., 2003).

The breeding of **disease resistant plant genotypes** is the third major approach globally to disease control in agricultural crops. The 19th century plant collectors often noticed differences in the disease susceptibility of various accessions of plants collected from the wild. After the rediscovery of Mendel's work, plant breeders in the early 1900s recognised that resistance to plant pathogens was often inherited as a dominant or semi-dominant trait. However, not until the seminal genetic studies performed in Australia by Harold H. Flor on flax and the flax rust pathogen *Melampsora lini* in the 1940s, was the inheritance of not only plant resistance but also pathogen virulence finally elucidated and the 'gene-for-gene' model proposed (reviewed by Day, 1974). This model predicts that plant resistance will occur only when a plant possesses a dominant **resistance gene** (*R*) and the pathogen expresses the complementary functional **avirulence gene** (*Avr*). Today, the pathogen avirulence genes are referred to as **effector genes** because many are recognised to also have a role in the pathogen's disease causing

ability (van Dijk et al., 1999). The 'gene-for-gene' model holds true for most biotrophic plant-pathogen interactions. For pathogens that deploy host selective toxins for successful pathogenesis, a different genetic model is proposed. Plant resistance is achieved through either enzymatic detoxification, a modified target site or through loss or alteration of the toxin in the pathogen. However, plant resistance to some fungal and *Oomycete* pathogens is predominantly based on major **quantitative trait loci** (QTLs) (Bai and Shaner, 2004). The molecular identity of many different major *R* gene classes has been determined over the past thirteen years for both cereal and non-cereal species (reviewed by Dangl and Jones, 2001). This advance has led to the development of precise 'within-the-*R*-gene' molecular markers. These markers now permit plant breeders to introgress novel combinations of *R* genes into elite cultivars without requiring pathogen testing. The molecular basis underlying QTL based resistance is not yet known. In rice, QTLs conferring resistance to bacterial pathogens often reside in chromosomal regions rich in specific defence-related proteins (Wu et al, 2004). The deployment of resistant genotypes is an extremely attractive disease control option because it involves no additional crop production costs and frequently provides the crop with lifelong protection. As is the case with fungicide resistance (see above), the inability to predict the potential durability of a newly introduced R gene in the field is a recurrent problem of increasing concern. This is because most *R* genes confer resistance against specific Avr-expressing isolates of a single pathogenic species. Only rarely does an *R* gene confer resistance to all races of pathogen or to multiple pathogens. An example of the former type, is the recessively inherited barley *mlo* gene that confers resistance to barley powdery mildew (*Blumeria graminis* fsp *hordei* (reviewed by Panstruga and Schulze-Lefert, 2002). Various deployment strategies have been devised to extend the lifetime of individual race-specific *R* genes through the use of cultivar mixtures and multilines as well as their spatial and temporal rotations (reviewed by Dangl and Jones, 2001; Hammond-Kosack and Parker, 2003). These approaches are often used in combination with monitoring the pathogen population for changes in specific effector sequences to determine the range of *R* genes that should/should not be deployed in particular cropping sequences and geographical locations.

In this article we examine how acquiring molecular genetic information on fungal and Oomycete pathogens with widely differing pathogenic lifestyles can assist plant breeders to create a greater diversity of disease resistant genotypes that when deployed appropriately in crop monocultures will provide sustainable control of multiple pathogens.

Gathering novel types of molecular information on pathogens

Use of forward and reverse genetics

One approach used to identify the pathogen genes of functional importance in disease has been **forward genetics**. Random mutations are created throughout the fungal genome, mutants are identified with reduced/no disease causing ability and then the DNA sequences causing the mutant phenotype are recovered. Random insertional mutagenesis is one type of forward genetics and involves inserting a linearised plasmid containing a selectable marker, e.g. an antibiotic resistance gene, randomly into the genome of the pathogen. This technique was first successfully used by Parsons *et al.* (1987) to transform *Magnaporthe grisea*. Fungal colonies with a plasmid integrated into their genome are selected by growth on a medium containing the relevant antibiotic. Subsequently, a range of PCR and/or DNA cloning techniques are used to identify the fungal sequence disrupted by the inserted plasmid. A recent alternative is to use *Agrobacterium tumefaciens* to insert single T-DNAs randomly into the fungal genome. This method was adapted from plant transformation protocols and first shown to work for filamentous fungi by de Groot *et al.* (1998). In both approaches, confirmation of the relationship between the disrupted gene sequence and the mutant phenotype is undertaken by a further round of site-directed targeted gene disruption involving homologous recombination and the wild-type strain. In contrast, as more fungal and Oomycete genes are cloned and pathogen genomes sequenced, the importance of **reverse genetics** in identifying novel pathogenicity and virulence genes as well as conserved mechanisms underlying pathogenesis will steadily increase for a diverse array of animal and plant pathogens. Reverse genetics starts from a candidate gene sequence, which is selectively disrupted in the wild-type strain by homologous recombination or gene silencing and the resulting transformants are examined for alterations in their disease causing ability.

When investigating fungal pathogenicity it is initially important to characterise the mutant strain under various in vitro conditions to establish that growth rates and sexual/asexual development are near wild-type. Then, in the subsequent plant pathogenicity tests, it is critical to distinguish the exact stage in the infection process at which the growth and development of a non-pathogenic or reduced virulence mutant differs from the isogenic wild-type strain. Typically, a pathogenicity minus ($Path^-$) mutant is classified as non-penetrating of intact plant surfaces, non-proliferating post-invasion or both. To date, approximately 290 validated pathogenicity/virulence/effector genes have been described in studies involving fourty-one plant and eight animal attacking fungal and Oomycete pathogens. However, these figures are artificially inflated because they include the characterisation of functionally homologous genes in multiple pathogenic species. Once functional homologues are removed the number of distinct pathogenicity, virulence and effector genes is reduced to 161 unique sequences. These genes code for various cellular signalling proteins (for example, members of mitogen activated protein (MAP) kinase cascades, G protein alpha and beta sub-units, adenylate cyclase, protein kinase A, cyclophilin), transcription factors (for example, GATA-factor, PacA, STE12-like), secreted proteins (for example, hydrophobins), membrane spanning proteins (for example, tetraspanin, toxin pumps, ABC transporters), and various types of enzymes (for example, polyketide synthase, isocitrate lyase and superoxide dismutase) (reviewed by Idnurm and Howlett, 2001) (**Table 1**).

Table 1. Conserved fungal and Oomycete pathogenicity and virulence genes

Gene Function	Plant Pathogen	Animal pathogen
ABC transporter	*Gp, Mg, Mgr*	*Ca*
Adenylate cyclase	*Bc, Mg, Um*	*Ca, Cn*
Adenylate forming enzyme	*Ch, Cv, Fg*	
Chitin synthase	*Bc, Fo, Um*	*Wd*
Cyclophilin	*Bc, Mg*	
GATA-factor type regulator	*Cli, Mg*	*Ca*
G-protein α and / or β subunits	*Ct, Cp, Fo, Mg, Pi,*	
Isocitrate lyase	*Lm, Mg*	
MAP kinase (HOG1)	*Cp*	*Ca*
MAP kinase (Fus3/Kss1)	*Bc, Ch, Cl, Cp, Fg, Fo, Mg, Pt, Um*	
MAP kinase (Slt2)	*Cl, Fg, Mg*	*Ca*

MAP kinase kinase (MEK)	*Cg, Um*	
Polyketide synthase	*Cl*	*Af, Wd*
Protein kinase A (r and /or c)	*Ct, Mg, Um,*	*Cn*
Superoxide dismutase	*Bc*	*Ca, Cn*
Tetraspanin	*Bc, Mg*	
Toxin pump	*Cc, Ck*	
Transcription factor (Ste12-like)	*Mg*	*Cg, Cn*
Transcription factor (SNF1 family)	Cc, Fo	

Plant pathogens: Bc-Botrytis cinerea, , Cc-Cochliobolus carbonum, Cg-Colletotrichum gloeosporioides, Ch-Cochliobolus heterostrophus, Ck- Cercospora kikuchii, Cl-Colletotrichum lagenarium,Cli-Colletotrichum lindemuthianum, Cp- Cryphonectria parasitica, Ct-Colletotrichum trifolii, Cp-Claviceps purpurea, Cv- Cochliobolus victoriae, Fg-Fusarium graminearum, Fo-Fusarium oxysporum, Gp-Gibberella pulicaris, Lm-Leptosphaeria maculans, Mg-Magnaporthe grisea, Mgr-Mycosphaerella graminicola, Pi-Phytophthora infestans, Pt-Pyrenophora teres, Um-Ustilago maydis. Animal pathogens: Af-Aspergillus fumigatus, Ca-Candida albicans, Cg-Candida glabrata, Cn-Cryptococcus neoformans, Wd-Wangiella (Exophiala) dematitidis. Gene function: r-regulatory sub-unit, c-catalytic sub-unit.

Most of the molecularly characterised effector molecules that trigger plant resistance are secreted proteins but with minimal sequence homology. For example, the effectors of the tomato pathogens *Cladosporium fulvum* (Avr2, Avr4 and Avr9) and *Fusarium oxysporum* f.sp. *lycopersici* which trigger race-specific *R* gene mediated resistance are small, cysteine rich proteins that are related by protein structure and not sequence homology (Joosten and de Wit, 1999; Rep et al., 2004). In contrast, the small Pep-13 effector of *Phytophthora sojae* which triggers non-host resistance is formed as a result of cleavage from a 42kD secreted glycoprotein and triggers the resistance response in the non-host plant, parsley (Lee et al., 2004). For the rice blast fungus *M. grisea* two effector proteins identified to date appear to be cytoplasmically localised and highly distinct. The *Avr Pi-ta* gene activates the rice resistance gene *Pi-ta*. *Pi-ta* encodes a putative cytoplasmic receptor with a centrally localised nucleotide-binding site and leucine-rich domain (LRD). *Avr Pi-ta* appears to encode a secreted neutral zinc metalloprotease (Orbach et al., 2000). Investigations using both the yeast two-hybrid system and an in vitro binding assay suggest that a non-secreted version of the Avr protein AVR-Pita (176) binds directly to the Pi-ta LRD region inside the plant cell to initiate a Pi-ta-mediated defence response (Jia et al., 2000). The second *M. grisea* effector gene *ACE1,* activates the rice resistance gene *Pi33*. *ACE1* encodes a combined polyketide synthase (PKS)/ non-ribosomal peptide synthetase (NPRS) that is 4035 amino acids in length (Bohnert et al., unpublished). No information is available on whether the ACE1 and Pi33 proteins directly interact. The current working hypothesis is that ACE1 is involved in the synthesis of a novel metabolite that is subsequently recognised by plant cells and leads to the activation of Pi33-mediated resistance.

The exploration of fungal and Oomycete pathogenicity by combined forward and reverse genetic approaches has revealed some remarkable conservation of gene/protein function despite the great diversity of pathogenic lifestyles and host species attacked. In Table 1 are shown the nineteen genes which have been verified experimentally to be of functional importance to pathogenicity in two or more

organisms, whether attacking plants or animals. A number of conclusions can be drawn from Table 1. Firstly, all isolates with these single gene mutations are viable and thus permit plant infection assays to be performed. Secondly, the number of entries involved in signal transduction and transcription events is high (12 of 19 entries). Thirdly, there is considerable evidence for molecular conservation of function between plant and animal pathogens at the level of regulation of gene transcription, cAMP signalling, mitogen-activated protein (MAP) kinase phosphorylation, the production and secretion of secondary metabolites and coping with reactive oxygen species, predominantly superoxide anions, generated as part of host defence mechanism via NADPH oxidase activity (Hammond-Kosack and Parker, 2003). The entries in **Table 1** also reveal that only ten of the conserved entries have so far been demonstrated to be of functional importance in pathogens attacking animals.

Sequencing whole fungal genomes

The goal of the global Fungal Genome Initiative (FGI) is 'to provide the sequence of key organisms across the fungal kingdom and thereby lay the foundation for work in medicine, agriculture, and industry' (Birren et al., 2003). Initially in February 2002, fifteen fungal species were nominated for genome sequencing, and considerable genomic sequence is now available for eight species (**Table 2** and **Table 3**).

Table 2. Websites for various fungal and Oomycete genome sequencing projects

Organism	URL (July 2004)
Ashbya gossypii	http://agd.unibas.ch/
Aspergillus fumigatus	http://www.sanger.ac.uk/Projects/A_fumigatus/
Aspergillus nidulans	http://www.broad.mit.edu/annotation/fungi/aspergillus/ http://www.tigr.org/tdb/e2k1/afu1/
Candida albicans	http://www-sequence.stanford.edu/group/candida/index.html
Coprinus cinereus	http://www.broad.mit.edu/annotation/fungi/coprinus_cinereus/
Coccidioides immitis	http://www.broad.mit.edu/annotation/fungi/coccidioides_immitis/
Cryptococcus neoformans Serotype A	http://www.broad.mit.edu/annotation/fungi/cryptococcus_neoformans/ http://www.tigr.org/tdb/e2k1/cna1/
Fusarium graminearum	http://www.broad.mit.edu/annotation/fungi/fusarium/
Magnaporthe grisea	http://www.broad.mit.edu/annotation/fungi/magnaporthe/
Neurospora crassa	http://www.broad.mit.edu/annotation/fungi/neurospora/

Phanerochaete chrysosporium	http://genome.jgi-psf.org/whiterot1/whiterot1.home.html
Phytophtora infestans	http://www.sanger.ac.uk/Projects/P_infestans/
Phytophthora ramorum	http://genome.jgi-psf.org/ramorum1/ramorum1.home.html
Phytophtora sojae	http://genome.jgi-psf.org/sojae1/sojae1.home.html
Pneumocystis carinii	http://www.sanger.ac.uk/Projects/P_carinii/
Saccharomyces cerevisiae	http://www.yeastgenome.org/
Schizosaccharomyces pombe	http://www.sanger.ac.uk/Projects/S_pombe/
Ustilago maydis	http://www.broad.mit.edu/annotation/fungi/ustilago_maydis/

The original species list included three experimentally tractable and important plant pathogens, *Magnaporthe grisea* (rice blast), *Fusarium graminearum* (ear blight/head scab/crown rot) and *Ustilago maydis* (corn smut). Recognising the power of comparative approaches, the FGI steering committee identified in June 2003 a further forty-four new fungi as immediate targets for sequencing with the emphasis on clusters of related species. The fungi the clusters will be developed around are *Candida*, *Aspergillus, Cryptococcus, Coccidioides, Penicillium, Fusarium, Histoplasma, Neurospora* and *Schizosaccharomyces* (Birren, 2003). Through other initiatives, sequenced genomes of the fungal pathogens *Stagnospora nodorum* (R. Oliver, pers comm) and *Botrytis cinerea* (M. Lebrun, pers comm) will also become available in 2005. The genomes of fungi are relatively modest in size (7-40 Mb) and most contain few repeats. This makes them ideal targets for whole-genome shotgun sequencing. The sequencing of several Oomycete genomes has also been co-ordinated by a global community effort and will include, *Phytophthora infestans*, *P. sojae* and an *Arabidopsis thaliana* attacking isolate of the downy mildew pathogen *Hyaloperonospora parasitica* (J. Benyon, UK, pers comm) (Slusarenko and Schlaich, 2003). Other organisms for which complete or near complete genomic sequence are available include the yeast–like plant pathogen *Ashbya gossypii*, human (*Homo sapiens*), fruit fly (*Drosophila melanogaster*), worm (*Caenorhabditis elegans*), a cluster of *Streptomyces* species, rice (*Oryza sativa*), the dicotyledonous model plant *Arabidopsis thaliana*, numerous bacteria and viruses as well as the human parasites, *Cryptosporidium parvum* and *Plasmodium falciparum*.

Table 3. Whole genome sequence information

Fungus	Genome size (MB)	Coverage	Predicted Genes	Contig > 2kb	Release date	Chromo-somes	Reference
Ashbya gossypii	9.2	complete	4,718	7	March 2004	7	Dietrich *et al*., 2004
Aspergillus fumigatus	32	10	na	19	July 2003	8	Unpubl.

Aspergillus nidulans	30	13	9,541	248	Feb 2003 V1.0	8	Unpubl.
Candida albicans	15	11	7,677	266	May 2002	7	Jones *et al.,* 2004
Candida dubliniensis	16	na	na	na	na	na	Unpubl.
Coprinus cinereus	38	10	na	431	June 2003V1.0	13	Unpubl.
Coccidioides immitis	29	10	na	215	April 2004	4	Unpubl.
Cryptococcus neoformans	20	na	na	na	Status June 04	na	Unpubl.
Fusarium graminearum	36	11	11,640	511	Sept 03	4	Unpubl.
Magnaporthe grisea	40	7	12,061	2275	Aug 02	7	Unpubl.
Neurospora crassa	40	>10	10,082	163	Apr 03	7	Galagan *et al.,* 2003
Phanerochaete chrysosporium	30	10.5	na	349 (>3kb)	2004 V1.0	10	Martinez et al., 2004
Saccharomyces cerevisiae	12	complete	6,000	12	May 1997	16	Goffeau et al., 1997
Schizosaccaromyces pombe	13.8	8	4,824	3	Feb 2002	3	Wood *et al.,* 2002
Ustilago maydis	20	10	6522	248	April 2004 2.0	23	Unpubl.

na – not available

Benefits of pathogen genomic sequences to crop breeding

Why is a complete gene inventory from non-pathogenic as well as animal and plant pathogenic fungal and Oomycete species of benefit to crop breeding? Also, why is it important to know which of these genes function in pathogenicity, virulence or as effectors? The answer is provided by comparative genome analyses performed between pathogenic and non-pathogenic species as well as within groups of

diverse pathogenic species. For example, how did parasitism evolve from a saprophytic lifestyle? Or, how did saprophytes evolve from parasites? (Lewis, 1973; Cooke and Whipp, 1980) Is pathogenicity due to the presence of novel genes? Or is pathogenicity the result of gene loss? Alternatively, is the gene inventory the same between pathogenic and non-pathogenic species, but there has been selection on how the genes are regulated? Do plant attacking species possess the same repertoire of genes as animal attacking species? The answers to these questions are only just beginning to emerge, as each newly sequenced genome is compared with the others available, and forward and reverse genetics experiments in multiple pathogen species reveal additional genes and pathways required for pathogencity and/or host plant recognition. In many species, typically up to 50% of the predicted genes are annotated as either of 'unknown function' or homologous to a gene predicted to exist in another organism that is also annotated of 'unknown function'. Despite this current paucity in functional information, within the next five years it is anticipated that many novel genes and gene families required by pathogens to cause disease will be identified through this comparative genome approach. For example, *Ashbya gossypii* which attacks cotton bolls was the first sequenced fungal plant pathogen (Dietrich et al., 2004). *A. gossypii* diverged from yeast (*S. cerevisiae*) more than a 100 million years ago. The two genomes differ substantially in GC content, 52% for *A. gossypii* and 38% for *S. cerevisiae.* However, ninety-five percent of the 4718 *A. gossypi* protein-coding sequences were also found in the *S. cerevisiae* genome and the majority (4281 ORFs) were at syntenic locations. Only 175 *A. gossypii* protein-coding genes (3.7%) showed homology but not synteny with *S. cerevisiae*, and 262 lack homology (5.6%). Based on the different hypotheses outlined above, this genomic information now permits molecular geneticists to select several series of candidate pathogenicity, virulence or effector genes. Reverse genetics is then used to explore the function of each gene sequence both in *A. gossypii* and other tractable species.

Comparative genome analyses can also provide answers to questions that explore the actual gene organisation on the chromosomes and relate this to gene function. For example, does gene clustering occur? For fungal and Oomycete genomes the answer is definitely yes. Therefore, are there differences in the extent of gene clustering between pathogenic and free-living species? Are the various pathogenicity, virulence and effector genes of pathogenic organisms randomly distributed across the genome or do they reside within these gene clusters? Is the GC content in the gene clusters abnormal which might indicate possible acquisition by horizontal gene transfer. In plant and animal attacking pathogenic bacterial species, the presence of 'pathogenicity gene islands' is very common (reviewed by Kim and Alfano (2002). In pathogenic fungi, this phenomenon has only once been described for two pathogenicity genes which occur on a supernumerary chromosome of the pea pathogen *Nectria haematococca* (Han et al., 2001). What is the function of gene clusters in pathogenic fungi and Oomycetes? Do they synthesise secondary metabolites required for disease causing ability or do they function at another lifecycle phase in the absence of the host plant? (reviewed by Keller and Hohn, 1997). Comparative genome analyses will also reveal the degree of inter-organism genomic synteny. In cases where the ORF sequence identity between species is less than 30%, synteny is particularly useful in the identification of highly diverged gene orthologs (Dietrich et al., 2004). Also, by linking the existing fungal and Oomycete genetic maps to the newly emerging genome sequence information, by using previously sequenced molecular markers, it will also be possible to determine whether specific regions of the genome have evolved to specialise in plant-associated processes. By comparing strains of the same species contrasting in virulence, it should then be possible to explore how such specialist chromosomal regions might be maintained. For example, in the Oomycete pathogen *Phytophthora infestans* the tight genetic linkage of three distinct *Avr* genes has recently been described (van der Lee et al., 2001). Syntenic chromosomal regions containing different *Avr* gene sequences have also been shown to exist in *Phytophthora infestans* (*Avr3a-1*) and *Hyaloperonospora parasitica* (*AtrNd*) (J. Benyon and P. Birch, unpublished) (Rehmany et al., 2003).

Comparative genome analyses can be linked to information on the pattern of expression of the organism's entire gene and protein complement, ie. the transcriptome and proteome, respectively. In the first instance, the transcription approach taken was to sequence libraries of expressed sequence tags (ESTs) from a number of distinct and biologically important stages in an organism's lifecycle. At the COGEME (2004) website, 54,083 unisequences have been collated based on ESTs deposited for thirteen fungal and two Oomycete plant pathogens as well as three free living organisms (Soanes et al., 2002). The entire COGEME database can be queried using a user friendly query page, by keyword and

sequence searches. Alternatively, in silico cDNA library subtraction experiments can be performed for *M. grisea* to identify biological event specific transcripts. In a newly funded phase of this project, the COGEME database will be linked with both the yeast Genome Information Management System (GIMS) object database that integrates many different types of 'omics data and the Central *Aspergillus* Data Repository (CADRE) both curated by the University of Manchester, UK, (Cornell et al., 2003, CADRE, 2004) and will be accessible via the GRID (S. Oliver and N. Talbot, pers comm.).

The availability of EST and genomic sequence information has permitted the development of a number of species-specific EST and/or oligonucleotide arrays containing between 2,000 and 10,000 probes. Microarrays already exist for several plant fungal pathogens, for example, *M. grisea, U. maydis, M. graminicola, Blumeria graminis* and *F. graminearum.* Transcription profiling experiments allow global analyses of gene expression and thereby reveal specific transcriptome patterns associated with wild-type compatible interactions on susceptible host plants, as well as during pathogen growth on various artificial media used to mimic the in planta environment. Examples of the later include growth under low and high nitrogen and carbon condition, or the initial infection process, eg. spore germination, germling growth and appressorium formation on an artificial surface. An alternative approach is called **SAGE** (Serial Analysis of Gene Expression) which again relies on an abundance of available sequence information before interpretation of the datasets generated is possible (Irie et al., 2003). Transcriptome analysis experiments are also beginning to compare wild-type and non-pathogenic mutants in various biological scenarios. These experiments aim to identify the genes co-ordinately regulated by each verified pathogenicity, virulence and effector gene and provide an initial blueprint for the cellular networks of interacting partners operating in different cellular compartments. Other transcriptome analyses are exploring in parallel the patterns of pathogen and host gene expression on both susceptible and resistant plant genotypes. These types of experiments, which primarily involve interactions on Arabidopsis, rice, barley and wheat, should reveal important information on how different types of pathogen attempt to cope with both constitutive and inducible plant defence responses.

Targets for intervention

Which pathogen targets are accessible to intervention by plant cell-derived products and could be approached through plant breeding? Which are not? **Figure 1** shows the predicted cellular location of many of the verified pathogenicity, virulence and effector gene products. Fungal gene products localised to the cytoplasm, the various organelles and the nucleus are potentially the most difficult targets to access via delivery from a healthy plant cell. These intracellular compartments could potentially be reached if plant proteins, metabolites and molecules are actively transported into the healthy pathogen cell or can transfer through the pathogen cell wall and plasma membrane unaided because of inherent chemical properties, for example, hydrogen peroxide. In contrast, pathogen gene products secreted from the cell, forming the pathogen cell wall, adhering to the pathogen surface or integrated within the plasma membrane are more accessible to plant cell-derived products. These potential pathogen targets include secreted effectors, plant cell wall degrading enzymes, proteases, cell wall re-modelling enzymes, nascent and mature cell wall components as well as integral membrane and membrane associated proteins, for example, receptors, sugar transporters, amino acid and peptide transporters, xenobiotic transporters, sodium and proton pumps, ion channels, extracellular transporters and vesicles for trafficking (Figure 1). During normal pathogenesis many of these potential targets are likely to come into contact with a plethora of plant-derived products which may change at different stages of the infection process. From Table 1, the conserved list of pathogenicity and virulence genes, three of the nineteen entries would be predicted to be directly accessible to plant-derived molecules from healthy cells. These are toxin pumps, ABC transporters and tetraspanins. During incompatible interactions when host defences are activated in plant cells, the cell walls and membranes of exploratory hyphal tips as well as more mature fungal and Oomycete cells could be compromised sufficiently to permit access of many other types of plant-derived products to the intracellular pathogen targets.

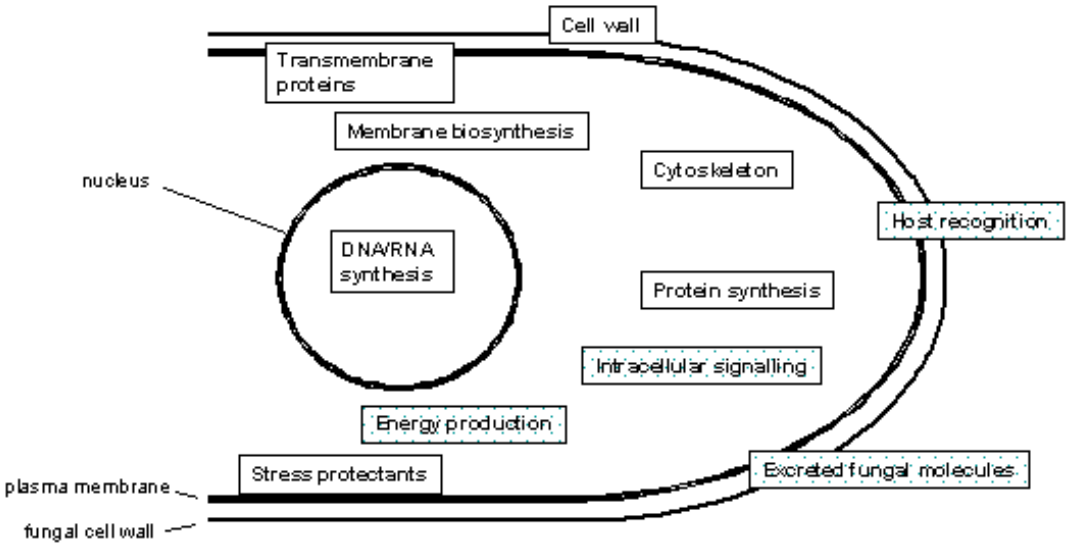

Figure 1. Pathogen targets accessible by plant-derived products and / or fungicides (FRAC, 2003)

For plant breeders to exploit information on readily accessible fungal and *Oomycete* targets for disease control purposes, a range of robust analysis tools and screening/selection methods will require development. For the tomato breeders this has already been achieved to identify novel sources of resistance to the leaf mould fungus *Cladosporium fulvum*. By injecting purified *C. fulvum* proteins or using a potato virus X-based gene expression system it was possible to identify additional *Lycopersicon* accessions that recognise and trigger a macroscopically visible disease resistance response to different *C. fulvum* effector proteins (Lauge et al., 2000). But how could one evaluate a collection of diverse plant germplasm accessions for the production of interfering or inhibitory compounds if the target is a fungal receptor, channel or transporter? One possible approach is to place a molecular tag on the target protein and use a transgenic pathogen strain with this protein reporter construct to monitor in vivo for aberrant protein expression or rapid target degradation. This would identify plant germplasms that interfere either directly or indirectly with the target. Alternatively, for some pathogen targets it may be possible to computer model the protein structure based on known targets from the pharmaceutical industry. This information would permit the most effective interfering chemistry to be predicted. Then, by establishing a specific metabolic screen it would be possible to identify plant germplasms enriched with this type of chemistry in the required plant organ. A third approach involves the transcriptome. Various companies have demonstrated that blocking a specific pathway, leads to transcriptional activation of the target gene to compensate for the loss in function. Therefore by using gene specific primers to several potential targets in a multiplexed quantitative reverse-transcriptase polymerase chain reaction or an EST/oligo microarray containing the various potential target pathogen sequences, it would be possible to identify plant germplasm that when inoculated lead to differential up- or down- regulation of one or more of the pathogen targets. Overall, the main challenge to applying each of the approaches outlined above to plant breeding is to devise robust screens and selections that can be undertaken in a high throughput format at modest costs to permit the surveying of entire crop plant germplasm collections.

Targets for monitoring

When a novel natural disease resistance source is introgressed into an elite commercial cultivar, even though this has taken the plant breeder many years to achieve, from the outset it is impossible to predict how many years effective disease control will be obtained in the field (Dangl and Jones, 2001). Likewise, the durability of a pathogen target cannot be predicted in advance. Therefore it is important to monitor pathogen populations for dynamic changes in the sequences of target genes, ie. pathogenicity, virulence and effector genes. This is best achieved by **archiving reference populations** of individual pathogen

species prior to the introduction of a new disease control option. These original archives can then be used to explore subsequent shifts in the pathogen population and additional pathogen entries procured each season whether a change is observed or not. In situations where a novel control approach is less effective, it is imperative that representative collections of the novel virulent pathogen isolates are made and compared at the molecular level with the reference set. For example, several effector gene sequences that trigger race-specific *R* gene-mediated resistance are now known for the economically important fungal pathogens *Magnaporthe grisea* (Jia et al., 2000, H. Bohnert and M-H Lebrun, unpublished), *Blumeria graminis* f..sp. *hordei* (barley powdery mildew*)* (C. Ridout and J. Brown, unpublished*)* as well as the *Oomycete* pathogens, *Phytophthora sojae* (Shan et al., 2004), *P. infestans* and *Hyaloperonospora parasitica* (downy mildew)(J. Benyon and P. Birch, unpublished). By comparing the effector genes sequences obtained from the new virulent isolates with the original pathogen reference collection it is possible to determine if molecular changes have occurred at the known effector sequence or at another unknown locus/loci. However, if the effector sequences are not known but genetic maps are available, the collected virulent isolates can be sexually crossed to avirulent isolates, the location of the mutant loci mapped and variation in open reading frames in the region explored by a technique called **PCR TILLING** (Targetting Induced Local Lesions In Genomes) (reviewed by Henikoff and Comai, 2003). This approach is considerably simplified for pathogenic organisms with full sequenced genomes because the candidate gene list can be developed immediately in the computer once the genetic map location has been determined. The PCR TILLING technique was originally developed for Arabidopsis. It is a relatively high throughput and facile way of identifying variations in a specific DNA sequence, typically 1000 bp in length, amongst pools of individuals, typically 4-16, without the use of DNA sequencing. In theory the PCR TILLING technique should work well for pathogens because its utility has already been demonstrated for diploid and polyploid organisms with significantly larger genomes.

A case study – Fusarium *ear blight of wheat*

Worldwide, *Fusarium* ear blight (FEB) infections of cereal crops cause considerable losses to grain quality and safety (Parry et. al., 1995; Windels, 2000). In the USA and China, this disease is known as *Fusarium* head scab (Scabusa, 2004) and Chimei Bing (pink head disease), respectively. *Fusarium* infections in small grain cereals and maize crops have been steadily increasing since the early 1990s due to changes to crop rotations, especially the increased use of cereal crops in the rotation, the introduction of maize into regions where previously only wheat was grown, and the use of low/minimal tillage practices that increase the levels of crop residues left on the soil surface (McMullen et al., 1997). Globally, *Fusarium* ear blight is caused by up to seventeen related haploid Ascomycete *Fusarium* species (Parry et. al., 1995). *F. graminearum* (sexual stage *Gibberella zeae*) is potentially the most damaging. *F. culmorum* is prevalent in NW Europe. Species specific DNA markers have been developed, to determine the most prevalent species in each geographical region (Schilling et al., 1996; Chandler et al., 2003).

Fusarium ear blight disease is primarily monocyclic with ear infections occurring when moist conditions prevail at anthesis and inoculum is available (Parry et al., 1995). Typically each cereal crop is vulnerable to infection for less than 21 days. Spore gain entry into the plant either by colonising the extruded anther and then the anther filament, or by penetrating directly the exposed ovary as each floret opens (Pugh and Johann, 1933; Kang and Buchenauer, 2000). No specialised fungal penetration structures have been reported. Anther tissue is rich in the metabolites choline and glycine betaine and these have been shown to alter the pattern of *Fusarium* hyphal branching (Strange et al., 1972; Pearce et al., 1976; Robson et al., 1995). Subsequent colonisation of florets, spikelets and the rachis involves a mixture of intercellular and intracellular growth. The advancing hyphal front colonises living plant tissue, which is subsequently killed. Microscopic examination of host cell walls in contact with *Fusarium* hyphae reveals localised degradation suggesting the action of polysaccharide-digesting enzymes. Considerable saprophytic aerial growth also occurs at all stages of the infection (Pugh, 1933; Kang and Buchenauer, 2000; Wanjiru et al., 2002). Infected wheat and barley ears bleach prematurely and masses of pink conidia form on infected spikelets under moist conditions. Grain from infected ears is often shrivelled (Parry et al., 1995) and can also be internally infected with fungal hyphae. Ascospores form within perithecia that develop on crop debris remaining on the soil surface after harvest (Pugh and Johann, 1933; Parry et al., 1995).

Floral infections by *Fusarium* species cause a lowering of grain quality and the developing grain becomes contaminated with various fungal mycotoxins, including the highly toxic type B trichothecene mycotoxins (Maresca et al., 2002). Three strain-specific profiles of trichothecene metabolites (**chemotypes**) have been identified within the B-trichothecene lineage of *Fusarium*: (i) nivalenol and acetylated derivatives (NIV chemotype), (ii) deoxynivalenol and 3-acetyldeoxynivalenol (3-ADON chemotype), and (iii) deoxynivalenol and 15-acetyldeoxynivalenol (15-ADON chemotype) (Miller et al., 1991; O'Donnell et al., 2000). NIV is generally regarded as more toxic to humans and animals than is DON (Ryu et al., 1988). Mycotoxin contamination of grain presents a serious health risk to humans and animals, and the EU is soon to legislate on the permitted DON levels in food and feed (Scholten et al., 2001). The proposed limits for DON and NIV will be 1-1.5ppm for unprocessed grain, 0.5-1ppm for flour, 0.5ppm for processed products and 0.1ppm for cereal based food sold for infants (European Commission). In the USA, equivalent limits are already in place. The brewing industry has zero tolerance for *Fusarium* mycelium and mycotoxins in cereal grain because these adversely affect the fermentation process to cause the phenomenon of 'gushing' beer. The cellular target site for DON mycotoxin is the peptidyl transferase protein in the ribosome. DON binding inhibits protein synthesis in eukaryotic cells (Cundliffe et al., 1974).

Many of the fungal genes involved in trichothecene production have been cloned (Desjardins et al., 1992; Brown et al., 2001; Lee et al., 2002). Earlier studies focussed on the regulation and biosynthesis of DON mycotoxin production from farnesyl pyrophosphate to trichothecene mycotoxins (**Figure 2)**. The initial committing step into the trichothecene biosynthetic pathway is catalysed by the enzyme trichodiene synthase to produce the intermediate trichodiene (reviewed by Brown *et al.*, 2001). The *TRI5* gene codes for trichodiene synthase. In three *Fusarium* species, at least ten of the trichothecene biosynthesis genes are located in close proximity to the *TRI5* gene, in a single gene cluster named the ***FgTRI5*-cluster** (**Figure 3).** This cluster contains the genes *TRI8*, *TRI3, TRI4, TRI6, TRI5, TRI10, TRI9, TRI11, TRI12, TRI13* and *TRI14* and spans less than 40Kb. However, four genes known to be required for trichothecene biosynthesis reside outside is *TRI* cluster (*TRI1*, *TRI15*, *TRI16* and *TRI101*) (Alexander *et al*., 2004).

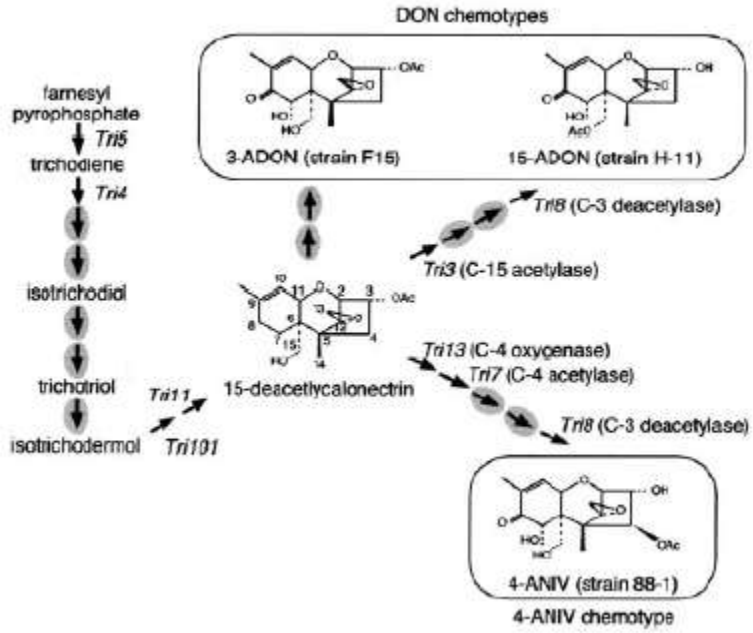

Figure 2. Trichothecene Biosynthesis (Kimura *et al.*, 2003)

Figure 3. The *FgTRI5* gene cluster of the sequenced strain PH-1

DON production has been shown to contribute to virulence of *F. graminearum* towards wheat, but is not an essential pathogenicity factor (Proctor et al., 1995; 2002). In addition, some natural *F. graminearum* and *F. culmorum* strains exist that do not produce DON mycotoxin, but are still able to cause ear blight disease. DON production is not constitutive and is specifically induced in planta or when *Fusarium* is grown under appropriate conditions in vitro (Doohan et al., 1999). A key step in trichothene mycotoxin induction, involves a zinc finger transcription factor encoded by the *TRI6* gene (Figure 3). No investigations have been reported on the plant genes and plant signals regulating DON induction. In vitro two–dimensional environmental profiling involving UK isolates of *F. culmorum* has revealed that production of both DON and NIV occurs under a far narrower range of conditions than overall hyphal growth (Hope and Magan, 2003).

Benefits to crop breeding - Molecular tools to predict different Fusarium chemotypes and the onset of mycotoxin production in planta

Complete sequencing of the trichothecene biosynthesis gene cluster (*FgTRI5*-cluster) in various 3-ADON, DON and 15-DON and NIV producing isolates of *F. graminearum* and *F. culmorum* has provided the possibility to predict the chemotype of an isolate based on the sequences of just four genes that reside toward the two ends of the *TRI5*-cluster. Disrupted or deleted *TRI8* and *TRI3* genes confers the 3-ADON chemotype, disrupted or deleted *TRI13* and *TRI8* genes confers the 15-ADON chemotype whilst functional copies of *TRI13* and *TRI7* genes are required for the NIV and 4-ANIV chemotypes, respectively (Lee et al., 2002; Chandler et al., 2003 and T. Ward, unpublished). Universal PCR primer sets have been developed as a diagnostic tool to chemotype isolates. For example, Jennings et al (2004) have recently reported on geographical differences in the incidence of various *F. culmorum* chemotypes in England and Wales between 1994 and 2002. Overall, DON chemotypes predominate but there was a greater proportion of NIV chemotypes in the warmer and moister south and west regions. Although resistance to *Fusarium* is both race and species non-specific, when screening wheat germplasm collections for FEB resistance it will now be feasible to ensure the inclusion of a range of *Fusarium* chemotypes in the breeding programme.

The *TRI5* gene encodes the enzyme trichodiene synthase and its activation predicts the onset of trichothecene mycotoxin production. The *TRI6* gene encodes the transcription factor which regulates *TRI5* gene induction. By generating various transgenic *Fusarium* reporter strains each expressing a different *TRI* promoter sequence fused to a reporter gene such as the beta-glucuronidase gene (*uidA*) or the green fluorescent protein (*GFP*), wheat genotype collections can be screened to identify accessions that fail to induce mycotoxin production. As each upstream step of the TRI6 regulatory pathway is identified, then this in turn could be used as a screening tool by plant breeders. Transgenic fungal reporter strains based on the *TRI13*, *TRI7* or *TRI8* promoter sequences could be used to screen for wheat genotypes that fail to induce production of the more toxic B-type NIV trichothecenes.

Natural disease resistance to Fusarium *floral infections*

In wheat, two main types of natural resistance mechanisms to *Fusarium* ear blight exist and other physiological types of resistance have been suggested (reviewed by Mesterhazy, 2002; Bai and Shaner, 2004). Type 1 resistance prevents initial ear infection. Type 2 resistance reduces the rate of disease

spread within an infected ear. Interestingly all sources of resistance in wheat to FEB are both pathogen race and pathogen species non-specific. Exotic wheat germplasm with Type 1 resistance exists but is rare and its genetic basis is not known. It therefore remains unclear whether the inability of wheat ears to prevent initial ear infection is due to a lack of recognition of *Fusarium,* the activation of only ineffective defence responses or active suppression of host defences by hyphae/toxins. Type 2 resistance is polygenic and is currently in commercial use or under selection in breeding programmes. Many major and minor quantitative trait loci (QTLs) linked to either type 1 or type 2 resistance have been identified on chromosomes 1B (type 1), 2AL, 2AS, 2BL, 2DL, 3A (type 1), 3BS (major QTL), 3D, 4B, 5A (type 1), 6AS, 6BS, 6D and 7A of wheat and a sub-set of these are associated with a reduction in DON levels in harvested grain, ie. 2B, 3BS, 5AS, 6B, and 7A in different genetic backgrounds (reviewed by Bai and Shaner, 2004). However, no resistance (*R*) gene sequences conferring type 2 resistance have been reported, although several map-based cloning projects are close to completion. Also the genetic relationship between pathogen resistance and low mycotoxin levels in harvested wheat grain is unclear. Germination of host genotypes in the presence of DON has proven an unsuccessful approach to select for resistant germplasm (B. Hollins (RAGT, UK) and P. Nicholson (JIC, UK, pers. com). Several plant genes up-regulated in one source of resistant germplasm post-FEB infection have been identified (Kruger et al., 2002). However, their exact role in conferring FEB disease resistance has not been reported.

F. graminearum *pathogenicity and virulence genes*

Molecular genetic analysis of the *Fusarium* ear infection process of cereal crop species is highly fragmentary compared with the model fungal species *Magnaporthe grisea*, *Ustilago maydis, Cryphonectria parasitica* (chestnut blight) and *Colletotrichum lindemuthianum* (anthracnose of bean). A molecular understanding of signalling events in *F. graminearum* controlling fungal penetration, cereal ear colonisation and spore formation would be extremely valuable. From a practical point of view, by preventing cereal tissue infection by *Fusarium* species, this would eliminate grain contamination and hence mycotoxin synthesis would no longer be a problem. From an evolutionary biology view point, molecular genetic analysis of the distinct infection biology of *F. graminearum* compared with other plant-attacking fungal species may provide fascinating generic insight into the maintenance and evolution of fungal pathogenicity. For example, *Fusarium oxysporum* species solely attack plant roots and exclusively colonise and sporulate within plant vascular tissue.

Recently, the targeted deletion of homologues of three previously known fungal pathogenicity genes has been reported for *F. graminearum* (Hou et al., 2002; Jenczmionka et al., 2003; Urban et al., 2003). *MGV1* encodes a mitogen-activated protein kinase and is homologous to the *MPS1* gene of the rice blast pathogen *Magnaporthe grisea*. *MPS1* is essential for *M. grisea* appressorium mediated penetration as well as post-invasive growth (Xu et al., 1998). In contrast, the *F. graminearum mgv1* mutant was still able to infect wheat ears although subsequent virulence is reduced (Hou et al., 2002). Most of the wheat spikelets inoculated with the *mgv1* mutant developed typical disease symptoms which spread into the adjacent spikelets but not the entire ear. The *mgv1* mutants produce highly reduced levels of DON mycotoxin in infected wheat ears. Recent work by Urban et al. (2003) has demonstrated by targetted gene disruption, that a second *F. graminearum* MAP kinase gene *MAP1* is absolutely required for the penetration of wheat florets and the stem base/root and also for invasive hyphal growth within plant tissue from a wound site. Hyphal colonisation by the *map1* mutant following inoculation on either intact or wounded wheat ears is restricted to the dehiscent anther tissue. The *map1* mutant is also compromised in deoxynivalenol (DON) mycotoxin production both in vivo and in vitro. The *Fusarium MAP1* gene is the functional equivalent of the *PMK1* gene of the rice blast pathogen *Magnaporthe grisea* and FUS3/KSS1 of *S. cerevisiae*. In *S. cerevisiae*, the FUS3/KSS1 pathway controls both the mating response and a part of the filamentous growth response following nitrogen starvation. Several signalling components have been identified in *S. cerevisiae* to act downstream of FUS3/KSS1 kinase. However, only a homologue of the transcription factor *Ste12* (Locus ID FG07310.1) has been identified in the sequenced *F. graminearum* genome (see below). Also, it remains unclear what the identity of the input signal(s) and receptor(s) are to this pathway in *F. graminearum. GzCPS1* encodes an adenylate forming enzyme and contains two AMP-binding domains (Lu et al., 2003). *F. graminearum cps1* mutants exhibit highly reduced virulence towards wheat ears causing only tiny yellow or brown spots to form on infected spikelets. The effect of the *cps1* mutation on trichthecene mycotoxin production was not reported. Targeted knock-out of

other conserved pathogenicity and virulence genes (see Table 1) will provide clues on the distinctness of pathogenesis of *F. graminearum* compared to other fungal pathogens.

The F. graminearum gene inventory

The *F. graminearum* genome was sequenced in 2003 to 11X coverage by the Broad Center for Genome Research, Boston, USA (Broad, 2004). The **PH-1 strain** sequenced represents the main genetic lineage distributed throughout the world (O'Donnell *et al.*, 2000). It also undergoes prolific sexual and asexual sporulation *in vitro*, is an extremely high mycotoxin producer (>300ppm) (chemotype DON:3-ADON: 15-ADON) (Hou et al., 2002) and is pathogenic to wheat, barley, rice and maize. The available genome sequence is present on just 43 supercontigs and covers 36 Mb. In addition in April 2004, a genetic map covering the four chromosomes, created using 162 sequenced markers, was successfully anchored to 99.6% of the assembled sequence (Gale and Kistler, 2004). The sequenced *F. graminearum* genome has been annotated by both the Broad Institute and MIPS in Germany (MIPS, 2004). The total gene number currently predicted by computer annotation is 11,620. In addition, the Broad Institute has calculated a mean gene density of 1 gene per 3100 bp, which is similar to other phytopathogenic fungal species, i.e. *M. grisea* at 1 gene per 3489 bp and *U. maydis* at 1 gene per 3018 bp, but considerably lower in gene density than *Saccharomyces cerevisiae* at 1 gene per 1914 bp. The Broad Institute also predicts that there are > 190 gene clusters within *F. graminearum*, where the gene density is greater than 10 ORFs per 25 kb stretch of DNA. Included amongst these are the mycotoxin producing *TRI5*-cluster on chromosome 2 (contig 1.159) and many containing polyketide synthase genes suggestive of the production of other secondary metabolites. It is possible that some gene clusters function during the saprophytic phase of the *F. graminearum* lifecycle and not during plant invasion. A collection of 4112 unisequences (EST contigs) are also available for the sequenced stain PH-1 (Trail et al., 2003, Cogeme, 2004). However, only 94.6% (e-value < -5) of these ESTs are found within the genomic sequence. The missing ESTs will aid the spanning of sequencing gaps in the genome. Based on the combined gene prediction by Broad and MIPS an Affymetrix oligonucleotide array for *F. graminearum* is under development and will be available in early 2005 to the global community for transcriptome analyses (C. Kistler, pers comm.).

Benefits to crop breeding – the F. graminearum gene inventory

A survey of the MIPS predicted *F. graminearum* gene inventory and further analyses by our own group now provide ample examples of the types of control targets discussed above. For example, in total 80 genes are predicted to be good homologues (e-value < -70) of verified pathogenicity and virulence genes (Baldwin, Antoniw and Hammond-Kosack, unpublished). MIPS currently predicts 1645 secreted proteins and 2471 transmembrane proteins that potentially would be accessible to plant-derived products. However, for the three MAP kinase cascades predicted (Urban and Hammond-Kosack, unpublished), two critical questions remain: (1) What are the input signal(s)? and (2) Are these signals of plant or fungal origin? The answers to both questions will influence whether these important pathogen signalling cascades could be targeted by plant breeders.

The linking of the four *Fusarium* chromosomes to the genomic sequence now permits distinct regions of the genome to be selected for exploring the natural diversity of *Fusarium* populations in different geographic locations. Previously, this had only been achieved using seven randomly selected single copy nuclear encoded genes (Ward et al., 2002). This new molecular information can be used by plant breeders to survey the entire genome for shifts in *Fusarium* population structure associated with the introduction of each novel resistance source into wheat. This type of analyses could also be linked to chemotype profiling of isolates (see above).

The existence of a near complete gene inventory now provides a tremendous resource for reverse genetics in *F. graminearum*. Recently a PCR based, split marker technique has been developed for high throughput gene replacement to test gene function by targeted homologous recombination (Catlett et al., 2003). It is anticipated that by applying knowledge gained from other pathogens on essential pathogenicity genes, this approach will soon identify a suite of *Fusarium* genes required for pathogenicity and /or virulence on wheat ear tissue and/or mycotoxin production as well as those which are not. For certain genes identified as functionally important for pathogenesis by *Fusarium*, it will then be possible to

create transgenic strains with specific pathways constitutively activated and thereby identify the downstream targets of each pathway. Some modified strains may exhibit heightened activation of plant defence responses and could be used to identify plant genotypes with elevated surveillance systems.

The molecular identity of QTL mediated resistance to *Fusarium* is not yet known. However, by using transgenic strains of *Fusarium* that produce or fail to produce DON mycotoxin, i.e. isogenic DON^{+} and DON^{-} strains, it should be possible to elucidate which QTLs confer a DON-mediated resistance response and which do not. It will be interesting to determine if type 1 resistance involves DON triggered events. By this approach, plant breeders can screen germplasm collections for wheat genotypes that confer resistance by other mechanisms and hopefully combine these sources to provide durable disease control. The imminent availability of the *F. graminearum* Affymetrix oligo array will allow identification of genes regulated during each resistance phenotype as well as by specific signalling pathways involved in pathogenicity. This will provide insight into the distinctiveness of each resistance source and give clues as to how the pathogen responds to the resistance mechanism. However, some of these experiments may be technically challenging because fungal biomass in ear tissue is extremely low immediately post-infection.

Overall perspectives

Each crop plant species is attacked by many different pathogens, each with a distinctive mode of pathogenesis. However, recent genome sequencing projects and forward and reverse genetic approaches to explore gene function and gene synteny have already revealed a surprising degree of gene conservation between pathogens with very diverse infection biology. The challenge for plant breeders is to harness this new molecular information on pathogens, not only to reveal novel sources of resistance in existing plant germplasm collections, but also to identify novel sources that can be successfully combined to provide durable disease control. This molecular information can also be used by plant breeders to evaluate the impact of each deployed resistance source on the overall pathogen population structure and thereby plan the successive introduction of appropriate resistant genotypes to sustain a high level of disease control.

Acknowledgements

Rothamsted Research receives grant aided support from the Biotechnology and Biological Sciences Research Council (BBSRC). DK is sponsored by a Rothamsted International Fellowship Award. TB and AD are in receipt of BBSRC sponsored CASE studentships, with the industrial partners Syngenta and RAGT, respectively. H-CJ is supported by the Department for the Environment, Food and Rural Affairs.

References

Agrios (1997) Plant Pathology, Academic Press, p635.

Alexander NJ, McCormick SP, Larson TM, Jurgenson JE (2004) Expression of *Tri15* in *Fusarium sporotrichioides.* Curr.Genet. 45, 157-162.
Arnold D, Pitman A and Jackson R (2003) Pathogenicity and other genomic islands in plant pathogenic bacteria. Mol.Plant Pathol. 4, 407-420.

Bai G and Shaner G (2004) Management and resistance in wheat and baley to *Fusarium* head blight. Ann. Rev. Phytopathol. 42, 135-161.

Birren B (2003) Fungal genome initiative – A white paper for fungal comparative genomics http://www.broad.mit.edu/annotation/fungi/fgi/FGI_whitepaper_Oct2003.

Broad (2004) Center for Genome Research. *Fusarium graminearum* genome. http://www.broad.mit.edu/cgi-bin/annotation/fusarium

Brown DW, McCormick SP, Alexander NJ, Proctor RH, Desjardins AE (2001) A genetic and biochemical approach to study trichothecene diversity in *Fusarium sporotrichioides* and *Fusarium graminearum.* Fungal Genet. Biol. 32, 121-133.
CADRE (2004) Central *Aspergillus* Data Repository. http://www.cadre.man.ac.uk/

Catlett NL, Lee BN, Yoder OC and Turgeon, BG (2003) Split-marker recombination for efficient targeted detection of fungal genes. Fungal Genet. News 50, 9-11.

Chandler EA, Simpson DR, Thomsett MA and Nicholson P (2003) Development of PCR assays to *Tri7* and *Tri13* trichothecene biosynthetic genes, and characterisation of chemotypes of *Fusarium graminearum*, *Fusarium culmorum* and *Fusarium cerealis*. Physiol. Mol. Plant Pathol. 62, 355-367.

COGEME (2004) Phytopathogenic Fungi and Oomycete EST Database. http://cogeme.ex.ac.uk

Cooke RC and Whipps JM (1973) The evolution of modes of nutrition and the origin of biotrophy. Biol. Rev. 55, 341-362.

Cornell M, Paton NW, Hedeler C, Kirby P, Delneri D, Hayes A and Oliver SG (2003) GIMS: an integrated data storage and analysis environment for genomic and functional data. Yeast 20, 1291-1306.

Cundliffe EM, Cannon M and Davies J (1974) Mechanism of inhibition of eukaryotic protein synthesis by trichothecene fungal toxins. P. Natl. Acad Sci USA 71, 30-34.

Dangl JL and Jones JDG (2001) Plant pathogens and integrated defence responses to infection. Nature 411, 826-833.

Day PR (1974) Genetics of host-parasite interaction. San Francisco, Calif., Freeman. 238p

de Groot MJA, Bundock P, Hooykaas PJJ and Beijersbergen AGM (1998). *Agrobacterium tumefaciens*-mediated transformation of filamentous fungi. Nat. Biotechnol. 16, 839-842.

Desjardins AE, Hohn TM and McCormick SP (1992) Effect of gene disruption on trichodiene synthase on the virulence of *Gibberella pulicaris*. Mol. Plant-Microbe Interact. 5, 214-222.

Dickinson M (2003) Molecular Plant Pathology. BIOS Scientific Publishers, London, UK. p244

Dietrich FS, Voegeli S, Brachat S et al (2004) The *Ashbya gossypii* genome as a tool for mapping the ancient *Saccharomyces cerevisiae.* genome. Science 304, 304-307.

Doohan FM, Weston G, Rezanoor HN, Parry DW and Nicholson P (1999) Development and use of a reverse transcription-PCR assay to study expression of *Tri5* by *Fusarium* species in vitro and in planta. Appl. Env. Microbiol. 65, 3850-3854.

Fraaije BA, Lucas JA, Clark WS and Burnett FJ (2003). QoI resistance development in populations of cereal pathogens in the UK. Proceedings of the BCPC Congress – Crop Science and Technology 2003, Volume 2, pp 689- 694.

FRAC (2003) Fungicide list (2) Sorted by mode of action. http://www.frac.info/publications/frac_list02.html

Han HY, Liu X, Benny U, Kistler HC and VanEtten HD (2001) Genes determining pathogenicity to pea are clustered on a supernumerary chromosome in the fungal plant pathogen *Nectria haematococca*. Plant J 25, 305-314.

Hammond-Kosack KE and Jones JDG (2000) Responses to Plant Pathogens In " Biochemistry and Molecular Biology of Plants" p 1102- 1156 Ed BB Buchanan, W Gruissem and RL Jones. Pp1367 American Society of Plant Physiology, Rockville, Maryland, USA.

Hammond-Kosack KE and Parker JE (2003) Deciphering plant-pathogen communication: Fresh perspectives for molecular resistance breeding. Cur. Opin. Biotechnol. 14, 177-193.

Henikoff S and Comai L (2003) Single-nucleotide mutations for plant functional genomics. Ann. Rev. Plant Biol. 54, 375-401.

Hope R and Magan N (2003) Two-dimensional environmental profiles of growth, deoxynivalenol and nivalenol production by *Fusarium culmorum* on a wheat-based substrate. Lett. App. Microbiol. 37, 70-74.

Hou ZM, Xue CY, Peng YL, Katan T, Kistler HC and Xu JR (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.

Gale and Kistler C (2004) http://www.broad.mit.edu/annotation/fungi/fusarium/maps.html

Galagan JE, Calvo SE, Borkovich KA et al., (2003) The genome sequence of the filamentous fungus *Neurospora crassa*. Nature 422, 859-868.

Goffeau A et al. (1997) The yeast genome directory. Nature 387, 1-105.

Idnurm A and Howlett BJ (2001) Pathogenicity genes of phytopathogenic fungi. Mol. Plant Pathol. 2, 241-255.

Irie T, Matsumura H, Terauchi R and Saitoh H (2003) Serial Analysis of Gene Expression (SAGE) of *Magnaporthe grisea*: genes involved in appressorium formation. Mol. Gen. Genet. 270, 181-189.

Jenczmionka NJ, Maier FJ, Losch AP, 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.

Jennings P, Coates ME, Turner JA, Chandler EA, Nicholson P (2004) Determination of deoxynivalenol and nivalenol chemotypes of *Fusarium culmorum* isolates from England and Wales by PCR assay. Plant Pathol. 53, 182-190.

Jia Y, McAdams SA, Bryan GT, Hershey HP and Valent B (2000) Direct interaction of resistance gene and avirulence gene products confers rice blast resistance. EMBO J. 19, 4004-4014.

Jones DG and Clifford BC (1983) Cereal diseases: their pathology and control. Chichester, John Wiley, 309p.

Jones T, Federspiel NA, Chibana H, et al. (2004) The diploid genome sequence of *Candida albicans* P. Natl. Acad Sci USA 101, 7329-7334.

Joosten MHAJ and de Wit PJGM (1999) The tomato - *Cladosporium fulvum* interaction: A versatile experimental system to study plant-pathogen interactions. Ann. Rev. Phytopathol. 37, 335-367.

Kang ZS and Buchenauer H (2000) Cytology and ultrastructure of the infection of wheat spikes by *Fusarium culmorum*. Mycological Res. 104, 1083-1093.

Kang ZS and Buchenauer H (2002). Studies on the infection process of *Fusarium culmorum* in wheat spikes: Degradation of host cell wall components and localization of trichothecene toxins in infected tissue. Eur. J. Plant Pathol. 108, 653-660.

Keller NP and Hohn TM (1997) Metabolic pathway gene clusters in filamentous fungi Fungal Genet. Biol. 21, 17-29.
Kim JF and Alfano JR (2002) Pathogenicity islands and virulence plasmids of bacterial plant pathogens Curr Top Microbiol 264, 127-147. Kimura M, Tokai T, O'Donnell K, Ward TJ, Fujimura M, Hamamoto H, Shibata T and Yamaguchi I (2003) The trichothecene biosynthesis gene cluster of *Fusarium graminearum* F15 contains a limited number of essential pathway genes and expressed non-essential genes. FEBS Letters 539, 105-110.

Kruger WM, Pritsch C, Chao SM and Muehlbauer GJ (2002) Functional and comparative bioinformatic analysis of expressed genes from wheat spikes infected with *Fusarium graminearum*. Mol. Plant-Microbe Interact. 15, 445-455.

Lauge R, Goodwin PH, de Wit PJGM and Joosten MHAJ (2000) Specific HR-associated recognition of secreted proteins from *Cladosporium fulvum* occurs in both host and non-host plants. Plant J. 23, 735-745.

Lee J, Rudd JJ, Macioszek VK and Scheel D (2004) Dynamic changes in the localization of MAPK cascade components controlling pathogenesis-related (PR) gene expression during innate immunity in parsley. J. Biol. Chem. 279, 22440-22448.
Lee T, Han YK, Kim KH, Yun SH, Lee YW (2002) *Tri13* and *Tri7* determine deoxynivalenol- and nivalenol-producing chemotypes of *Gibberella zeae.* App. Env. Microbiol. 68, 2148-2154.

Lewis DH (1973) Concepts in fungal nutrition and the origin of biotrophy. Bio. Rev. 48, 261-278.

Lu SW, Kroken S, Lee BN, Robbertse B, Churchill ACL, Yoder OC and Turgeon BG (2003) A novel class of gene controlling virulence in plant pathogenic ascomycete fungi. P. Natl. Acad Sci USA 100, 5980-5985.

Maresca M, Mahfoud R, Garmy N and Fantini J (2002) The mycotoxin deoxynivalenol affects nutrient absorption in human intestinal epithelial cells. J. Nutrition 132, 2723-2731.

Martinez D, Larrondo LF, Putnam N, et al. (2004) Genome sequence of the lignocellulose degrading fungus *Phanerochaete chrysosporium* strain RP78. Nat. Biotechnol. 22, 695-700.

Mesterhazy A (2002) Role of deoxynivalenol in aggressiveness of *Fusarium graminearum* and *F. culmorum* and in resistance to *Fusarium* head blight. Eur. J. Plant Pathol. 108, 675-684.

Miller, JD, Greenhaugh R, Wang YZ and Lu M (1991) Trichothecene chemotypes of 3 Fusarium species. Mycologia 83, 121-130.

MIPS (2004) Munich information center for protein sequences.
http://mips.gsf.de/genre/proj/fusarium/Search/

O'Donnell K, Kistler HC, Tacke BK and Casper HH (2000) Gene genealogies reveal global phylogeographic structure and reproductive isolation among lineages of *Fusarium graminearum*, the fungus causing wheat scab. P. Natl. Acad Sci USA, 97, 7905-7910.

Orbach MJ, Farrall L, Sweigard JA, Chumley FG, Valent B (2000) A telomeric avirulence gene determines efficacy for the rice blast resistance gene *Pi-ta.* Plant Cell 12, 2019-2032.
Panstruga R and Schulze-Lefert P (2002) Live and let live: insights into powdery mildew disease and resistance. Mol. Plant Pathol. 3, 495-502.

Parry DW, Jenkinson P and McLeod L (1995) *Fusarium* ear blight (scab) in small grain cereals - a review. Plant Pathology 44, 207-238.

Parsons KA, Chumley FG and Valent B (1987) Genetic transformation of the fungal pathogen responsible for rice blast disease. Proc. Natl. Acad. Sci.(USA) 84, 4161-4165.

Pearce RB, Strange RN, and Smith H (1976) Glycinebetaine and choline in wheat: distribution and relation to infection by *Fusarium graminearum*. Phytochemistry 15, 953-954.

Proctor RH, Hohn TM and McCormick SP (1995). Reduced virulence of *Gibberella zeae* caused by disruption of a trichothecene toxin biosynthetic gene. Mol. Plant-Microbe Interact. 8, 593-601.

Proctor RH, Desjardins AE, McCormick SP, Plattner RD, Alexander NJ and Brown DW (2002) Genetic analysis of the role of trichothecene and fumonisin mycotoxins in the virulence of *Fusarium*. Eur. J. Plant Pathol. 108, 691-698.

Pugh WG and Johann H (1933) Factors affecting infection of wheat heads by *Gibberella saubinetti*. J. Agri. Research, 46, 771-797.

Rehmany AP, Grenville LJ, Gunn ND, Allen RL, Paniwnyk Z, Byrne J, Whisson SC, Birch PRJ, and Beynon JL (2003) A genetic interval and physical contig spanning the *Peronospora parasitica* (*At*) avirulence gene locus *ATR1Nd.* Fungal Genet. Biol. 38, 33-42.

Rep M, van der Does HC, Meijer M, van Wijk R, Houterman, PM, Dekker, HL, de Koster CG and Cornelissen BJC (2004) A small, cysteine-rich protein secreted by *Fusarium oxysporum* during colonization of xylem vessels is required for *I-3*-mediated resistance in tomato. Mol. Microbiol. in press.

Robson GD, Wiebe MG, Cunliffe B and Trinci APJ (1995) Choline-induced and acetylcholine-induced changes in the morphology of *Fusarium graminearum* – evidence for the involvement of the choline transport-system and acetycholinesterase. Microbiol. 141, 1309-1314.

Ryu JC, Ohtsubo K, Izumiyama N, Nakamura K, Tanaka, T, Yamamura H and Ueno, Y (1988) The acute and chronic toxicities of nivalenol in mice. Fund. Appl. Toxicol. 11, 38-47.

Scabusa (2004) US wheat and barley scab initiative. http://www.scabusa.org

Schilling AG, Moller EM, Geiger HH (1986) Polymerase chain reaction-based assays for species-specific detection of *Fusarium culmorum*, *F. graminearum*, and *F. avenaceum*. Phytopathol. 86, 515-522.

Scholten OE, Ruckenbauer P, Visconti A,van Osenbruggen WA and den Nijs APM (2001) Food Safety of Cereals: A chain-wide approach to reduce *Fusarium* Mycotoxins, Commission of the European Communities.

Sesma A and Osbourn AE (2004)The rice leaf blast pathogen undergoes developmental processes typical of root-infecting fungi. Nature

Shan WX, Cao M, Dan LU and Tyler BM (2004) The *Avr1b* locus of *Phytophthora sojae* encodes an elicitor and a regulator required for avirulence on soybean plants carrying resistance gene *Rps1b*. Mol. Plant-Microbe Interact. 17, 394-403.

Slusarenko AJ and Schlaich NL (2003) Downy mildew of *Arabidopsis thaliana* caused by *Hyaloperonospora parasitica* (formerly *Peronospora parasitica*). Mol. Plant Pathol. 4, 159-170.
Soanes DM, Skinner W, Keon J, Hargreaves J and Talbot N J (2002). Genomics of phytopathogenic fungi and the development of bioinformatic resources. Mol. Plant-Microbe Interact.15, 421-427.

Strange RN, Smith H and Majer JR (1972) Choline, one of to fungal growth stimulants in anthers responsible for the susceptibility of wheat to *Fusarium graminearum*. Nature 238, 103-104.

Trail F, Xu JR, San Miguel P, Halgren RG, Kistler HC (2003) Analysis of expressed sequence tags from *Gibberella zeae* (anamorph *Fusarium graminearum*). Fungal Genet. Biol. 38, 187-197.
Urban M, Mott E, Farley T and Hammond-Kosack K (2003) The *Fusarium graminearum MAP1* gene is essential for pathogenicity and development of perithecia. Mol. Plant Pathol. 4, 347-359.

van der Lee T, Robold A, Testa A, van't Klooster JW and Govers F (2001) Mapping of avirulence genes in *Phytophthora infestans* with amplified fragment length polymorphism markers by bulked segregant analysis. Genetics 157, 949-956.

van Dijk K, Fouts DE, Rehm AH, Hill AR, Collmer A and Alfano JR (1999) The Avr (effector) proteins HrmA (HopPsyA) and AvrPto are secreted in culture from *Pseudomonas syringae* pathovars via the Hrp (type III) protein secretion system in a temperature- and pH-sensitive manner. J. Bacteriol. 181, 4790-4797.

Wanjiru WM, Kang ZS and Buchenauer H (2002) Importance of cell wall degrading enzymes produced by *Fusarium graminearum* during infection of wheat heads. Eur. J. Plant Pathol., 108, 803-810.

Ward TJ, Bielawski JP, Kistler HC, Sullivan E and O'Donnell K (2002) Ancestral polymorphism and adaptive evolution in the trichothecene mycotoxin gene cluster of phytopathogenic *Fusarium*. P. Natl. Acad Sci USA 99, 9278-9283.

Wood V, Gwilliam R, Rajandream MA, et al. (2002) The genome sequence of *Schizosaccharomyces pombe*. Nature 415, 871-880.

Windels CE (2000) Economic and social impacts of *Fusarium* head blight: Changing farms and rural communities in the Northern Great Plains. Phytopathol. 90, 17-21.

Wu JL, Sinha PK, Variar M, Zheng KL, Leach JE, Courtois B and Leung H (2004) Association between molecular markers and blast resistance in an advanced backcross population of rice. Theor. Appl. Genet. 108, 1024-1032.

Xu JR, Staiger CJ, Hamer JE (1998) Inactivation of the mitogen-activated protein kinase Mps1 from the rice blast fungus prevents penetration of host cells but allows activation of plant defense responses. P. Natl. Acad Sci USA 95, 12713-12718.