Studies on global transcriptional regulator EBR1 and genome-wide gene expression in the fungal plant pathogen *Fusarium graminearum*

Chunzhao Zhao

Thesis committee

Promotors
Prof. Dr P.J.G.M. de Wit
Professor of Phytopathology
Wageningen University

Prof. Dr D. Tang
Professor of Genetics and Developmental Biology
Chinese Academy of Sciences, Beijing, China

Co-promotor
Dr T.A.J. van der Lee
Senior Scientist, Plant Research International
Wageningen University

Other members
Prof. Dr B.J. Zwaan, Wageningen University
Prof. M. Rep, University of Amsterdam
Prof. Dr R.P. de Vries, Utrecht University
Dr G. Smant, Wageningen University

The research was conducted under the auspices of the Graduate School of Experimental Plant Sciences

Studies on global transcriptional regulator EBR1 and genome-wide gene expression in the fungal plant pathogen *Fusarium graminearum*

Chunzhao Zhao

Thesis
submitted in fulfillment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus
Prof. Dr A.P.J.Mol,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Tuesday 8 December 2015
at 8:30 a.m. in the Aula.

Chunzhao Zhao
Studies on global transcriptional regulator EBR1 and genome-wide gene expression in the fungal plant pathogen *Fusarium graminearum*,
167 pages.

PhD thesis, Wageningen University, Wageningen, NL (2015)
With references, with summary in English

ISBN 978-94-6257-599-8

Contents

Chapter 1

General introduction

Cereals have become indispensable as food for human beings. In the field, cereals are threatened by a variety of plant pathogens. Those pathogens that can cause severe diseases on cereals include *Magnaporthe oryzae*, *Puccinia graminis*, *Fusarium graminearum*, *Blumeria graminis* and *Mycosphaerella graminicola*, which cause rice blast, stem rust, Fusarium head blight (FHB), powdery mildew and septoria tritici blotch (STB), respectively (Dean et al., 2012). All these diseases contribute to significant annual losses in cereal production. Some of these pathogens including *Fusarium* species also threaten the health of humans due to contamination by mycotoxins produced by these pathogens. As plant pathogens are so destructive, understanding the interaction of these pathogens with their hosts and finding durable solutions to control these destructive pathogens are becoming more and more urgent. In this thesis, I studied *F. graminearum* and the interaction between this pathogenic fungus and its cereal hosts.

Fusarium species causing FHB

F. graminearum is a globally occurring plant pathogen that causes FHB on wheat (*Triticum aestivum*), barley (*Hordeum vulgare*), rye (*Secale cereale*), and oat (*Avena sativa*); it also causes Gibberella stalk and ear rot disease on maize (*Zea mays*) (Goswami and Kistler, 2004) (Figure 1). FHB of wheat was first described in 1884 in the United Kingdom and was considered as a major threat to wheat and barley cultivation at the beginning of the twentieth century. Ever since, FHB occurrence has increased worldwide, and outbreaks have now been reported in Asia, Canada, Europe and the USA (Windels, 2000). The predominant *Fusarium* species associated with FHB in small grain cereals in Europe are *F. graminearum*, *F. culmorum*, *F. avenaceum* and *F. poae* (Bottalico and Perrone, 2002). The pathogen poses a two-fold threat. First, infected cereals are significantly reduced in seed quality and yield. In 1993, FHB outbreaks in North America caused more than $1 billion in losses (McMullen et al. 1997). Secondly, infected kernels become contaminated with trichothecenes (e.g. deoxynivalenol (DON)) and estrogenic mycotoxins (e.g. zearalenone (ZEA)), making them unsuitable for food or feed (Placinta et al. 1999). DON is a potent inhibitor of protein synthesis in eukaryotic organisms and toxic to humans and animals upon inhalation and ingestion (Bennett and Klich, 2003; Bluhm et al., 2007). ZEA is the causative agent of hyperestrogenic syndrome in animals, which probably leads to reproductive disorders (Park et al., 2015).

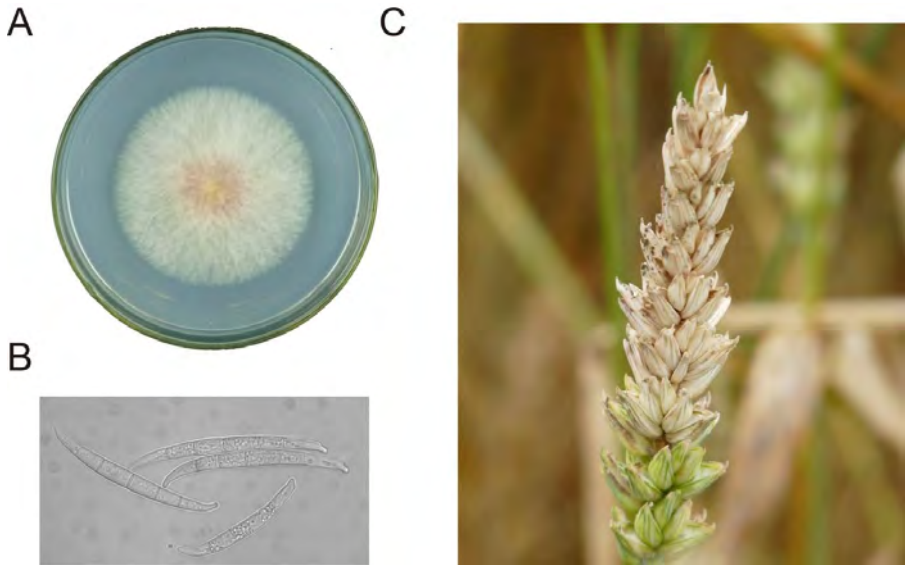


Figure 1. Morphology of *Fusarium graminearum* and Fusarium head blight (FHB) caused by this fungus. (A) Mycelium of *F. graminearum* growing on solid CM medium. (B) Macroconidia of *F. graminearum*. (C) Symptoms of FHB on wheat ears in the field.

Resistance of plants to FHB inherits as a quantitative trait. So far, no absolute resistance is available in commercially grown wheat, barley, maize and oat cultivars. Resistance to FHB in wheat is thought to be based on five components related to different stages of the infection process: type I resistance to initial infection, type II resistance to colonization or spread, type III resistance to kernel infection, type IV tolerance, and finally type V resistance against mycotoxin accumulation (Mesterhazy, 1995). It is important to identify the type and the level of resistance in breeding material to be transferred into commercial cultivars in order to develop effective breeding strategies to FHB (Mesterhazy, 1995). Presently, such information is not available for most of the breeding material. Chinese cultivar Sumai 3 and its derivatives have all type II resistance, which represent the highest level of resistance to FHB that is presently available. Usually FHB symptoms do not spread to uninoculated spikelets after point inoculation of Sumai 3, but sometimes FHB symptoms can be observed in nearby uninoculated spikelets under heavy disease pressure in the field (Bai and Shaner, 2004). The high cost of fungicides, their low efficacy, high chance of resistance development and the incomplete understanding of their effects on mycotoxin production hamper efficient chemical control strategies. In addition, factors that

influence disease development including host resistance and climate conditions and insufficient knowledge of the epidemiology make it difficult to control FHB (Trail et al., 2003). Thus it is necessary to get a better understanding of the interaction between the FHB causing *Fusarium* species and their hosts. Understanding host-, organ- and species-specificity in the different pathosystems can efficiently support future breeding strategies to obtain FHB-resistant cereal crops.

Virulence genes of *F. graminearum*

To understand the infection mechanism of the various cereal hosts by *F. graminearum*, the identification of genes that are associated with pathogenicity is essential. Generally, there are two strategies to identify genes that are involved in pathogenicity or virulence. A frequently used method involves the generation of a T-DNA or transposon insertion library of mutant *F. graminearum* strains that can be screened for loss of pathogenicity or reduced virulence; subsequently the disrupted genes in these mutants can be identified and studied. Another method involves the creation of targeted knock-outs of genes in *F. graminareum* that are putatively involved in virulence, and subsequently the knock-out mutants are analysed for changes in pathogenicity or virulence on their respective hosts.

So far, two high-throughput mutagenesis methods have been explored to identify mutants affected in virulence of *F. graminearum*. The first one involves *mimp1* transposon-mediated mutagenesis (Dufresne et al., 2008). Transposon *mimp1*, previously identified in *F. oxysporum*, is a miniature inverted-repeat transposable element (MITE) (Dufresne et al., 2007). Facilitated by the *impala* transposase, *mimp1* can move into new sites rich in TA nucleotides. *Mimp1* insertion in promoters or open reading frames of genes can disrupt the corresponding genes. The insertion sites of *mimp1* in new genes can be determined by tail-PCR technology. By this approach, 11 genes have been identified that are putatively involved in virulence (Dufresne et al., 2008). These genes include a transcription factor, a secreted protein and metabolite-related genes. The second high-throughput method exploits restriction enzyme-mediated integration (REMI) to generate random insertion mutants (Seong et al., 2005). Seong and co-workers analyzed the virulence of 6,500 REMI mutants of *F. graminearum* and identified 20 mutants showing reduced virulence on corn silk and 11 showing reduced virulence on both corn silks and flowering wheat heads.

To date, around 130 genes have been identified in *F. graminearum* that are involved in virulence (all genes are listed in Table 1). The first set of virulence-related genes are kinases. Wang and co-workers knocked out all 116 kinase-encoding genes identified in *F. graminearum* and found that 42 kinase-defective mutants showed reduced virulence or were no longer-pathogenic (Wang et al., 2011a). Among these kinases, MGV1 and GPMK1 were previously reported to be required for pathogenicity (Hou et al., 2002; Jenczmionka et al., 2003). Apart from reduced virulence, the *mgv1* mutant also exhibited reduced vegetative growth and a defect in development of perithecia. The *gpmk1* mutant showed a reduced production of conidia and sexual sterility. In addition, both *mgv1* and *gpmk1* mutants showed enhanced sensitivity to the antifungal defensin protein MsDef1, but not to MtDef4, indicating that these two MAP kinase signalling pathways play an important role in the regulation of the sensitivity of *F. graminearum* to MsDef1. The second set of virulence-related genes are transcription factors. In total, 657 putative transcription factors have been identified in *F. graminearum*. Son and co-workers knocked out all these transcription factor-encoding genes and identified 62 that are required for virulence (Son et al., 2011). An additional set of virulence genes is involved in the production of trichothecene. Although trichothecene is not required for initial infection of the wheat head, it is important for the spread of the fungus to other spikelets, allowing penetration of the rachis node of spikes (Bai et al., 2002; Desjardins et al., 2000; Jansen et al., 2005; Proctor et al., 1995). Knocking out the *Tri5* gene, which encodes trichodiene synthase involved in the first step of trichothecene biosynthesis, caused reduced virulence (Proctor et al., 1995). *Tri6* and *Tri10* encode two transcription factors that regulate the expression of *Tri* genes (Nasmith et al., 2011; Seong et al., 2009). Knocking out *Tri6* and *Tri10* also caused reduced virulence (Proctor et al., 1995; Seong et al., 2009). Also deletion of another regulator responsible for the expression of *Tri* genes and production of trichothecenes, *FgVe1*, caused reduced virulence (Merhej et al., 2012).

Other genes involved in virulence of *F. graminearum* include *RAS2*, *FSR1*, *FGL1*, *FgATG15*, *CBL1*, *MSY1*, *FBP1*, *HMR1*, *NPS6*, *GzGPA2*, *GzGPB1*, *MES1*, *NPC1*, *HDF1*, *FTL1*, *FgRrg-1*, *SSK1*, *FgTep1p* and *FgPtc1p* (Bluhm et al., 2007; Breakspear et al., 2011; Ding et al., 2009; Han et al., 2007; Jiang et al., 2011; Jiang et al., 2010; Li et al., 2011; Merhej et al., 2012; Nguyen et al., 2011; Oide et al., 2010; Oide et al., 2006; Rittenour and Harris, 2008; Seong et al., 2005; Seong et al., 2006; Shim et al., 2006; Voigt et al., 2005; Wang et al., 2011b; Yu et al., 2008; Zhang et al.,

2010). *RAS2* encodes a Ras GTPase, which was shown to be involved in the regulation of kinase activity of Gpmk1. Disruption of *RAS2* resulted in both reduced virulence and reduced hyphal growth (Bluhm et al., 2007). Mutant *fsr1*, which was identified by the screen of REMI mutants, is defective in female fertility, radial growth and virulence. It was hypothesized that FSR1 acts as a scaffold protein in signal transduction (Shim et al., 2006). *FGL1* encodes a secreted lipase, which exhibits extracellular lipolytic activity (Voigt et al., 2005). *FgATG15* encodes an autophagy-like lipase, which also exhibits lipolytic activity. More detailed studies showed that FgATG15 is required in lipid droplet degradation in response to starvation (Nguyen et al., 2011). Both lipases are required for the virulence of *F. graminearum*. *CBL1* and *MSY1* encode cystathionine beta-lyase and methionine synthase, respectively, both of which are involved in the synthesis of methionine. Disruption of either *CBL1* or *MSY1* resulted in reduced virulence and reduced growth on minimal medium, indicating that methionine synthesis is important for the virulence and aerial hyphal growth of *F. graminearum* (Seong et al., 2005). FBP1 is a versatile F-box protein, which is probably involved in the formation of the SCF^{FBP1} complex and subsequently controls the ubiquitin-mediated degradation of proteins associated with virulence (Han et al., 2007). *HMR1*, encoding a HMG-CoA reductase, is required for sterol and isoprenoid biosynthesis. Disruption of *HMR1* led to reduced growth and virulence (Seong et al., 2006). Non-ribosomal peptide synthetase, which is encoded by *NPS6* in *F. graminearum*, is a conserved virulence determinant in many plant pathogenic *Ascomycetes*. NPS6 is required for the biosynthesis of extracellular siderophores, which in turn supply iron to fungi (Oide et al., 2006). The G protein-mediated signalling pathway has been demonstrated to be important in virulence of many fungi. In *F. graminearum*, there are three putative Gα subunits, named GzGPA1, GzGPA2 and GzGPA3, and one Gβ subunit, GzGPB1. Targeted deletion of these genes showed that Δ*Gzgpa2* and Δ*Gzgpb1* mutants displayed reduced virulence, whereas the virulence of the Δ*Gzgpa1* and Δ*Gzgpa3* mutants was not significantly altered. Deletion of *GzGPB1* also resulted in the reduced hyphal growth (Yu et al., 2008). MES1, involved in the organization of proteins at the cell surface, is also important in virulence, which suggests that some virulence determinants may localize to a specialized site at the cell surface (Rittenour and Harris, 2008). NPC1 localizes to the vacuolar membrane and is required for sterol trafficking in *F. graminearum* (Breakspear et al., 2011). HDF1 is a putative histone deacetylase, which is not only involved in virulence, but also plays an important role in

regulation of developmental processes, including sexual reproduction and conidiation. HDF1 physically interacts with transducin beta-like protein FTL1 (Li et al., 2011). Similar to HDF1, FTL1 also plays a critical role in virulence (Ding et al., 2009). Response regulator FgRRG-1 is one of the elements of the high-osmolarity glycerol (HOG) pathway, mainly involved in virulence, osmotic stress response and fungicide resistance (Jiang et al., 2011). *SSK1*, encoding a histidine kinase, is also involved in the HOG pathway. Deletion of *SSK1* in *F. graminearum* resulted in reduced virulence, increased pigment production, enhanced asexual sporulation and decreased perithecium production (Oide et al., 2010). *FgTep1p* encodes a Tensin-like phosphatise 1 that is involved in the phosphatidylinositol-3 kinase signalling pathway (Zhang et al., 2010). *FgPtc1p* encodes a type-2C protein phosphatase (Jiang et al., 2010). Both FgTep1p and FgPtc1p are required for virulence and resistance to lithium toxicity. The functions of these virulence genes and the phenotypes of their mutant versions are listed in Table 1. From these studies, we can conclude that a variety of cellular processes, including signal transduction pathways, transcriptional regulatory networks and alteration of metabolism, all contribute to successful infection of *F. graminearum* on its hosts.

Secondary metabolite genes of *F. graminearum*

Secondary metabolites, including polyketides, non-ribosomal peptides, terpenes, have been identified in many fungi, including *Fusarium* spp. and *Aspergillus* spp. (Fox and Howlett, 2008). Secondary metabolites have been found that are important for the survival and pathogenicity of fungi. The most obvious secondary metabolites produced by fungi are pigments, which can be found in a variety of developmental structures, such as asexual and sexual spores, appressoria, sclerotia and sexual fruiting bodies (Yu and Keller, 2005). Studies have shown that pigments are generally required for survival of fungi under severe conditions, such as UV and ROS stress (Calvo et al., 2002; Coccia et al., 2001). It has also been shown that pigments can act as virulence factors in some fungi (Calvo et al., 2002). Another group of important secondary metabolites are mycotoxins, which are used by fungi to compete with saprophytes in various ecological niches or to increase their ability to cause diseases on their hosts (Proctor et al., 1995).

In *F. graminearum*, there are 15 *PKS* genes encoding polyketides and 19 NPRS genes encoding non-ribosomal peptides (Table 2). Of these, 9 *PKS* genes and 10 *NRPS* genes seem

specific to *F. graminearum*, indicating that these secondary metabolite genes could be newly evolved genes. So far, the well studied secondary metabolite genes in *F. graminearum* include *PKS4*, *PKS13*, *PKS9*, *PKS12*, *NRPS2* and *NRPS6*. *PKS4 and PKS13* are required for the production of ZEA (Gaffoor and Trail, 2006). *PKS9* is required for the production of three novel fusarielins (Sorensen et al., 2012). *PKS12* is essential for the production of aurofusarin, which is associated with the production of pigment (Malz et al., 2005). In addition, *NRPS2* is required for the production of ferricrocin (Tobiasen et al., 2007) and *NRPS6* for the biosynthesis of extracelluar siderophores (Oide et al., 2006). Of other *PKS* and *NRPS* genes, the final products are still unknown. Apart from polyketides and non-ribosomal peptides, one of the representative terpene-derived products in *F. graminearum* is trichothecene, the biosynthesis of which depends on 14 *Tri* genes (Proctor et al., 2009). The role of thrichotecenes in virulence has been discussed earlier.

Like in other fungi, genes in *F. graminearum* that are involved in the production of secondary metabolites are often organized in clusters (Keller and Hohn, 1997). One representative secondary metabolite gene cluster in *F. graminearum* is *PKS12*. The *PKS12* gene cluster consists of 11 genes including those that encode polyketide synthase, the Zn2Cys6 transcription factor, FAD/FMN-containing dehydrogenases and flavin-containing mono-oxygenase. Knocking out one of the genes in this cluster caused reduced production of pigment, indicating that they function in concert (Frandsen et al., 2006). Another example of a cluster of genes is the ZEA cluster that includes *PKS4* and *PKS13* that are in close vicinity on the same chromosome (Gaffoor and Trail, 2006; Kim et al., 2005). Apart from these two polyketide synthases, this gene cluster also contains *ZEB1* and *ZEB2*, which encode isoamyl alcohol oxidase and a bZIP-DNA binding domain-containing protein, respectively. Deletion of each of these four genes showed reduced ZEA production (Kim et al., 2005). Also the *PKS9* gene cluster, involved in the production of fusarielin, has been analysed in detail. This cluster contains seven genes including those that encode polyketide synthase, esterase/lipase, aldose 1-epimerase, cytochrome P450 mono-oxygenase, enoyl reductase, AMP-binding protein and a transcription factor (Sorensen et al., 2012). In addition, *Tri* genes are also clustered, including one core cluster and two smaller ones (Proctor et al., 2009) (Figure 2).

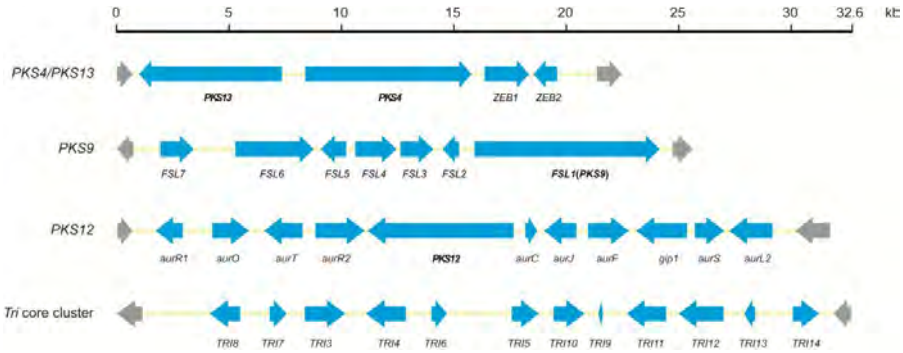


Figure 2. Four secondary metabolite gene clusters identified in *Fusarium graminearum*. The *PKS4* and *PKS13* gene clusters are required for the production of zearalenone; the *PKS9* gene cluster is required for the production of fusarielins; the *PKS12* gene cluster is essential for the production of aurofusarin; the *Tri* gene cluster is required for the production of trichothecenes.

Genome analysis of *F. graminearum*

The genome of *F. graminearum* has been sequenced and annotated (Cuomo et al., 2007). The total genome size of *F. graminearum* is 36.1 Mb, which is significantly less than that of the closely related species *F. verticillioides* (41.7 Mb) and *F. oxysporum* (59.9 Mb). The smaller genome size of *F. graminearum* could be explained by the fact that few high-identity duplicated sequences are present, which could be due to the presence of the phenomenon of repeat-induced point mutations (RIP) which inhibit genome proliferation. According to the MIPS (The Munich Information Center for Protein Sequences) *F. graminearum* Genome Database (http://mips.gsf.de/genre/proj/fusarium), its genome encodes 13826 genes, of which 1720 are very similar to known genes, 4058 are similar to known genes, 4915 show some similarity to unknown genes, and 3133 show no similarity to any genes in other organisms (Wong et al., 2011).

Compared to the closely related species *F. verticillioides* and *F. oxysporum*, *F. graminearum* has undergone complex chromosome rearrangements. *F. graminearum* only contains four chromosomes, whereas *F. verticillioides* contains 11 chromosomes, and *F. oxysporum* contains 11 core and four lineage-specific (LS) chromosomes (Ma et al., 2010). When comparing the genome of *F. graminearum* to *F. verticillioides*, it shows that *F. graminearum* chromosome (Fgchrom) 1 matches with *F. verticillioides* chromosome (Fvchrom) 1, 5 and 8; Fgchrom2 matches with Fvchrom 6, 9, 10 and 11; Fgchrom3 matches with Fvchrom 2, 4, 7; and Fgchrom 4 matches with

Fvchrom 2 and 3. This suggests that the current number of chromosomes of *F. graminearum* is the result of ancestral chromosome fusions. In particular, high SNP densities are mainly found in the telomere proximal regions. However, in addition to the telomeric regions, one or two large interstitial regions of high SNP densities were identified in the middle of three chromosomes of *F. graminearum* (Cuomo et al., 2007). These highly variable sequences within the chromosomes may reflect ancestral telomere locations, which suggest that current chromosomes of *F. graminearum* are the result of ancestral chromosome fusions. Other features identified in the genome of *F. graminearum* include (i) low numbers of duplicated sequences and transposons, (ii) genes specifically expressed *in planta* are enriched in the high-SNP-density regions (Cuomo et al., 2007).

Scope of this thesis

This thesis covers two aspects of the *F. graminearum*. In the first part we studied a transcription factor-encoding gene, *EBR1*, which is required for the radial growth and virulence of *F. graminearum*. The phenotype of mutant *ebr1* is described in detail. The second part of the thesis covers gene annotation and alternative splicing in *F. graminearum* by analyzing RNA-Seq data. Besides, variations in gene expression in different regions of chromosomes were investigated.

Chapter 1 introduces *F. graminearum* and FHB caused by this fungus. In addition, the current knowledge on virulence-related genes and secondary metabolite genes was briefly reviewed. Finally, aspects of the genome analysis of *F. graminearum* were discussed.

Chapter 2 describes mutant *Δebr1* (enhanced branching 1) that exhibits reduced radial growth and reduced virulence on wheat. By using a GFP-tagged isolate, it could be shown that infection of wheat by the *Δebr1* mutant was blocked at the point of rachis invasion. *EBR1* encodes a Zn_2Cys_6 transcription factor. GFP-tagged EBR1 protein was shown to constitutively localize in the nucleus. Complementing *Δebr1* with the orthologue from *F. oxysporum* f.sp. *lycopersici* restored radial growth and virulence of *Δebr1* to wild-type level, indicating that *EBR1* is conserved among different *Fusarium* species.

Chapter 3 describes new gene models and alternative splicing in *F. graminearum*. Using RNA-Seq data obtained from wild-type PH-1 and mutant *Δebr1*, 655 incorrectly annotated genes and 231 genes showing alternative splicing were identified in *F. graminearum*. In addition, 2459 novel transcriptionally active regions (nTARs) were identified.

Chapter 4 describes the variation in gene expression in different chromosome regions of *F. graminearum*. Genome-wide gene expression analysis showed that 13 regions on the four chromosomes of *F. graminearum* displayed relatively lower gene expression levels. Further investigation showed that these lowly expressed regions correlate with non-conserved chromosome regions in *F. graminearum*. Interestingly, weakly expressed regions are full of gene relocations, which probably led to the clustering of genes with similar expression patterns and also drove the evolution of new genes.

Chapter 5 discusses the results presented in the previous chapters in a broader context and highlights the contribution of our studies to the Fusarium research community. We also propose future experiments that need to be performed to further elucidate the role of EBR1 in virulence and the mechanism of genome evolution and genome-wide regulation of gene expression in *F. graminearum*.

Table 1. Genes that are involved in virulence of *Fusarium graminearum*.

Gene ID	Name	Type of protein	Effect on virulence	Effect on radial growth	Effect on conidiation	Effect on perithecium formation	Effect on DON production	Reference
FGSG_01665	FSR1	Multimodular protein	Reduced	Reduced	-	Failed	N.D.	**(Shim et al., 2006)**
FGSG_10114	RAS2	Ras GTPases	Reduced	Reduced	-	Failed	-	**(Bluhm et al., 2007)**
FGSG_05906	FGL1	Secreted lipase	Reduced	N.D.	N.D.	N.D.	N.D.	**(Voigt et al., 2005)**
FGSG_16251	TRI6	Transcription factor	Reduced	N.D.	N.D.	N.D.	Not detected	**(Seong et al., 2009)**
FGSG_03538	TRI10	Transcription factor	Reduced	N.D.	N.D.	N.D.	Not detected	**(Seong et al., 2009)**
FGSG_10825	MSY1	Methionine	Reduced	Reduced in	-	N.D.	N.D.	**(Seong et al., 2005)**
FGSG_01932	CBL1	Cystathionine beta-lyase	Reduced	Reduced in	-	N.D.	N.D.	**(Seong et al., 2005)**
FGSG_01555	ZIF1	bZIP transcription factor	Reduced	Slightly reduced	N.D.	Late stage of perithecium	Reduced	**(Seong et al., 2005; Wang et al., 2011b)**
FGSG_00376	NOS1	NADH dehydrogenase iron-sulfur protein	Reduced	N.D.	N.D.	N.D.	N.D.	**(Seong et al., 2005)**
FGSG_00332	TBL1	Transducin β-subunit-like protein	Reduced	N.D.	N.D.	N.D.	N.D.	**(Seong et al., 2005)**
FGSG_09280	FBP1	Fructose-1,6-bisphosphatase	Reduced	Reduced	N.D.	Failed	N.D.	**(Han et al., 2007)**
FGSG_09197	HMR1	Hydroxymethyl-glutaryl CoA reductase	Reduced	Reduced	-	N.D.	Reduced	**(Seong et al., 2006)**
FGSG_03747	NPS6	Nonribosomal Peptide Synthetase	Reduced	-	N.D.	N.D.	N.D.	**(Oide et al., 2006)**
FGSG_09614	GzGPA2	Ga subunit	Reduced	Slightly reduced	N.D.	-	-	**(Yu et al., 2008)**
FGSG_04104	GzGPB1	Gb subunit	Reduced	Reduced	N.D.	-	Increased	**(Yu et al., 2008)**

Gene ID	Name	Type of protein	Effect on virulence	Effect on radial growth	Effect on conidiation	Effect on perithecium formation	Effect on DON production	Reference
FGSG_06680	MES1	Novel fungal protein	Reduced	Reduced	Reduced	Failed	-	(Rittenour and Harris, 2008)
FGSG_09254	Npc1	Ergosterol transporter	Reduced	Reduced (temperature dependent)	N.D.	N.D.	N.D.	(Breakspear et al., 2011)
FGSG_10129	FgStuAp	Transcriptional regulator	Reduced	Reduced	Reduced	Failed	not detected	(Lysoe et al., 2011)
FGSG_01353	HDF1	Histone Deacetylase	Reduced	-	Reduced	Failed	Reduced	(Li et al., 2011)
FGSG_11955	FgVe1	Velvet	Reduced	-	Reduced	N.D.	Reduced	(Merhej et al., 2012)
FGSG_02519	FgATG15	Autophagy-like lipase	Reduced	Reduced	Reduced	N.D.	Reduced	(Nguyen et al., 2011)
FGSG_08948	FgRrg-1	Response regulator	Reduced	-	Reduced	N.D.	Reduced	(Jiang et al., 2011)
FGSG_04982	FgTep1p	Tensin-like phosphatase 1	Reduced	-	Reduced	N.D.	N.D.	(Zhang et al., 2010)
FGSG_04111	FgPtc1p	Type 2C protein phosphatase	Reduced	-	N.D.	N.D.	N.D.	(Jiang et al., 2010)
FGSG_00332	FTL1	Transducin Beta-Like protein	Reduced	Reduced	Reduced	N.D.	-	(Ding et al., 2009)
FGSG_08948	SSK1	Response regulator	Reduced	-	Reduced	Delayed	N.D.	(Oide et al., 2010)
FGSG_09612	HOG1	Kinase	Reduced	Reduced	Reduced	Delayed	N.D.	(Oide et al., 2010)
FGSG_04484		Kinase	Reduced	Reduced	Reduced	Failed	-	(Wang et al., 2011a)
FGSG_05547		Kinase	Reduced	Reduced	Reduced	Failed	-	(Wang et al., 2011a)
FGSG_08468		Kinase	Reduced	Reduced	increased	-	-	(Wang et al., 2011a)
FGSG_09903	STE7	kinase	Reduced	Slightly reduced	N.D.	N.D.	N.D.	(Ramamoorthy et al., 2007; Wang et al., 2011a)

Gene ID	Name	Type of protein	Effect on virulence	Effect on radial growth	Effect on conidiation	Effect on perithecium formation	Effect on DON production	Reference
FGSG_06939		Kinase	Reduced	Reduced	Reduced	-	Reduced	**(Wang et al., 2011a)**
FGSG_05734		Kinase	Reduced	Reduced	Reduced	-	Reduced	**(Wang et al., 2011a)**
FGSG_13318		Kinase	Reduced	Reduced	Reduced	-	-	**(Wang et al., 2011a)**
FGSG_06878		Kinase	Reduced	Reduced	Reduced	-	Reduced	**(Wang et al., 2011a)**
FGSG_10228		Kinase	Reduced	Reduced	Reduced	-	-	**(Wang et al., 2011a)**
FGSG_10381		Kinase	Reduced	Reduced	Reduced	Failed	-	**(Wang et al., 2011a)**
FGSG_08635		Kinase	Reduced	Reduced	Reduced	-	-	**(Wang et al., 2011a)**
FGSG_01641		Kinase	Reduced	Reduced	Reduced	-	-	**(Wang et al., 2011a)**
FGSG_06385	GPMK1	Kinase	Reduced	Reduced	Reduced	Failed	-	**(Wang et al., 2011a)**
FGSG_08691		Kinase	Reduced	Reduced	Reduced	Failed	-	**(Wang et al., 2011a)**
FGSG_05484	STE11	Kinase	Reduced	-	N.D.	N.D.	N.D.	**(Ramamoorthy et al., 2007; Wang et al., 2011a)**
FGSG_04947		Kinase	Reduced	Reduced	Reduced	-	Reduced	**(Wang et al., 2011a)**
FGSG_02795		Kinase	Reduced	Reduced	Reduced	Failed	-	**(Wang et al., 2011a)**
FGSG_07295		Kinase	Reduced	Reduced	Reduced	Failed	-	**(Wang et al., 2011a)**
FGSG_06326		Kinase	Reduced	Reduced	Reduced	Failed	-	**(Wang et al., 2011a)**
FGSG_10313	MGV1	Kinase	Reduced	Reduced	Reduced	Failed	-	**(Wang et al., 2011a)**
FGSG_09897		Kinase	Reduced	Reduced	Reduced	-	-	**(Wang et al., 2011a)**
FGSG_00408		Kinase	Reduced	Reduced	Reduced	Failed	-	**(Wang et al., 2011a)**
FGSG_07251		Kinase	Reduced	Reduced	Reduced	-	-	**(Wang et al., 2011a)**
FGSG_07329		Kinase	Reduced	Reduced	Reduced	Failed	-	**(Wang et al., 2011a)**
FGSG_10037		Kinase	Reduced	Reduced	Reduced	Failed	-	**(Wang et al., 2011a)**

Gene ID	Name	Type of protein	Effect on virulence	Effect on radial growth	Effect on conidiation	Effect on perithecium formation	Effect on DON production	Reference
FGSG_06957		Kinase	Reduced	Reduced	Reduced	-	-	(Wang et al., 2011a)
FGSG_10066		Kinase	Reduced	Reduced	No conidiation	Failed	-	(Wang et al., 2011a)
FGSG_01188		Kinase	Reduced	Reduced	No conidiation	Failed	-	(Wang et al., 2011a)
FGSG_00362		Kinase	Reduced	Reduced	No conidiation	Failed	-	(Wang et al., 2011a)
FGSG_04053		Kinase	Reduced	Reduced	Reduced	Failed	-	(Wang et al., 2011a)
FGSG_09274		Kinase	Reduced	-	Reduced	-	-	(Wang et al., 2011a)
FGSG_00472		Kinase	Reduced	-	Reduced	-	Reduced	(Wang et al., 2011a)
FGSG_04770		Kinase	Reduced	-	Slightly increased	-	Reduced	(Wang et al., 2011a)
FGSG_10095		Kinase	Reduced	-	Slightly reduced	-	Reduced	(Wang et al., 2011a)
FGSG_06793		Kinase	Reduced	-	Slightly reduced	-	Reduced	(Wang et al., 2011a)
FGSG_01312		Kinase	Reduced	-	Strongly reduced	-	Reduced	(Wang et al., 2011a)
FGSG_03284		Kinase	Reduced	-	-	-	-	(Wang et al., 2011a)
FGSG_05418		Kinase	Reduced	-	Reduced	-	-	(Wang et al., 2011a)
FGSG_08906		Kinase	Reduced	-	Slightly reduced	-	-	(Wang et al., 2011a)
FGSG_07344		Kinase	Reduced	-	Strongly reduced	-	-	(Wang et al., 2011a)
FGSG_11812		Kinase	Reduced	Reduced	-	-	-	(Wang et al., 2011a)
FGSG_10384		Transcription factor	Reduced	Reduced	-	Delayed	Reduced	(Son et al., 2011)

Gene ID	Name	Type of protein	Effect on virulence	Effect on radial growth	Effect on conidiation	Effect on perithecium formation	Effect on DON production	Reference
FGSG_06071		Transcription factor	Reduced	Reduced	Reduced	Delayed	-	**(Son et al., 2011)**
FGSG_01307		Transcription factor	Reduced	-	-	Delayed	-	**(Son et al., 2011)**
FGSG_06291		Transcription factor	Reduced	Reduced	Reduced	Delayed	Reduced	**(Son et al., 2011)**
FGSG_00515		Transcription factor	Reduced	Reduced	-	Delayed	-	**(Son et al., 2011)**
FGSG_01555		Transcription factor	Reduced	Reduced	-	Delayed	Reduced	**(Son et al., 2011)**
FGSG_05171		Transcription factor	Reduced	Reduced	-	Delayed	Reduced	**(Son et al., 2011)**
FGSG_06651		Transcription factor	Reduced	Reduced	-	-	Reduced	**(Son et al., 2011)**
FGSG_09832		Transcription factor	Reduced	Reduced	-	Delayed	-	**(Son et al., 2011)**
FGSG_10142		Transcription factor	Reduced	-	-	-	-	**(Son et al., 2011)**
FGSG_00477		Transcription factor	Reduced	Reduced	-	Delayed	-	**(Son et al., 2011)**
FGSG_01022		Transcription factor	Reduced	Reduced	-	Delayed	Reduced	**(Son et al., 2011)**
FGSG_01106		Transcription factor	Reduced	-	-	Delayed	Reduced	**(Son et al., 2011)**
FGSG_01341		Transcription factor	Reduced	-	-	Delayed	-	**(Son et al., 2011)**
FGSG_01350		Transcription factor	Reduced	Reduced	-	Delayed	Reduced	**(Son et al., 2011)**
FGSG_04083		Transcription factor	Reduced	-	-	Delayed	-	**(Son et al., 2011)**
FGSG_04134		Transcription factor	Reduced	Reduced	Reduced	Delayed	Reduced	**(Son et al., 2011)**
FGSG_06427		Transcription factor	Reduced	-	-	Delayed	-	**(Son et al., 2011)**
FGSG_06871		Transcription factor	Reduced	Reduced	-	Delayed	-	**(Son et al., 2011)**
FGSG_07052		Transcription factor	Reduced	Reduced	Reduced	-	-	**(Son et al., 2011)**
FGSG_07928		Transcription factor	Reduced	Reduced	-	Delayed	-	**(Son et al., 2011)**
FGSG_08617		Transcription factor	Reduced	Reduced	-	Delayed	Reduced	**(Son et al., 2011)**
FGSG_10517		Transcription factor	Reduced	Reduced	Reduced	Delayed	Reduced	**(Son et al., 2011)**

Gene ID	Name	Type of protein	Effect on virulence	Effect on radial growth	Effect on conidiation	Effect on perithecium formation	Effect on DON production	Reference
FGSG_11416		Transcription factor	Reduced	-	-	-	-	(Son et al., 2011)
FGSG_13711		Transcription factor	Reduced	Reduced	Reduced	Delayed	Reduced	(Son et al., 2011)
FGSG_1374		Transcription factor	Reduced	Reduced	Reduced	Delayed	Increased	(Son et al., 2011)
FGSG_02527		Transcription factor	Reduced	Reduced	Reduced	Delayed	Reduced	(Son et al., 2011)
FGSG_01182		Transcription factor	Reduced	Reduced	Reduced	Delayed	-	(Son et al., 2011)
FGSG_05304		Transcription factor	Reduced	Reduced	-	Delayed	Reduced	(Son et al., 2011)
FGSG_00385		Transcription factor	Reduced	Reduced	Reduced	Delayed	Reduced	(Son et al., 2011)
FGSG_00729		Transcription factor	Reduced	-	Reduced	Delayed	-	(Son et al., 2011)
FGSG_09868		Transcription factor	Reduced	Reduced	-	-	Reduced	(Son et al., 2011)
FGSG_09019		Transcription factor	Reduced	Reduced	-	Delayed	Reduced	(Son et al., 2011)
FGSG_06948		Transcription factor	Reduced	Reduced	-	Delayed	Reduced	(Son et al., 2011)
FGSG_10179		Transcription factor	Reduced	-	-	-	-	(Son et al., 2011)
FGSG_09339		Transcription factor	Reduced	-	-	Delayed	Reduced	(Son et al., 2011)
FGSG_00324		Transcription factor	Reduced	Reduced	Reduced	Delayed	Reduced	(Son et al., 2011)
FGSG_12781		Transcription factor	Reduced	Reduced	Reduced	Delayed	Reduced	(Son et al., 2011)
FGSG_09992		Transcription factor	Reduced	Reduced	Reduced	Delayed	Reduced	(Son et al., 2011)
FGSG_08737		Transcription factor	Reduced	-	-	-	-	(Son et al., 2011)
FGSG_09654		Transcription factor	Reduced	Reduced	Reduced	Delayed	-	(Son et al., 2011)
FGSG_13120		Transcription factor	Reduced	Reduced	Reduced	Delayed	Reduced	(Son et al., 2011)
FGSG_01665		Transcription factor	Reduced	Reduced	Reduced	Delayed	Reduced	(Son et al., 2011)
FGSG_06228		Transcription factor	Reduced	-	-	Delayed	Increased	(Son et al., 2011)
FGSG_06944		Transcription factor	Reduced	Reduced	-	Delayed	-	(Son et al., 2011)
FGSG_08481		Transcription factor	Reduced	Reduced	Reduced	Delayed	Reduced	(Son et al., 2011)

Gene ID	Name	Type of protein	Effect on virulence	Effect on radial growth	Effect on conidiation	Effect on perithecium formation	Effect on DON production	Reference
FGSG_08572		Transcription factor	Reduced	Reduced	Reduced	Delayed	Reduced	**(Son et al., 2011)**
FGSG_08719		Transcription factor	Reduced	Reduced	Reduced	Delayed	Reduced	**(Son et al., 2011)**
FGSG_10716		Transcription factor	Reduced	Reduced	Reduced	Delayed	Reduced	**(Son et al., 2011)**
FGSG_10069		Transcription factor	Reduced	Reduced	Reduced	Delayed	-	**(Son et al., 2011)**
FGSG_08769		Transcription factor	Reduced	Reduced	-	Delayed	Reduced	**(Son et al., 2011)**
FGSG_08182		Transcription factor	Reduced	-	-	-	-	**(Son et al., 2011)**
FGSG_08028		Transcription factor	Reduced	-	-	-	-	**(Son et al., 2011)**
FGSG_07133		Transcription factor	Reduced	-	Reduced	Delayed	Reduced	**(Son et al., 2011)**
FGSG_07067		Transcription factor	Reduced	-	-	Delayed	-	**(Son et al., 2011)**
FGSG_01293		Transcription factor	Reduced	-	-	-	-	**(Son et al., 2011)**
FGSG_01176		Transcription factor	Reduced	-	-	Delayed	Reduced	**(Son et al., 2011)**
FGSG_00719		Transcription factor	Reduced	-	-	-	-	**(Son et al., 2011)**
FGSG_00574		Transcription factor	Reduced	Reduced	-	Delayed	-	**(Son et al., 2011)**
FGSG_00147		Transcription factor	Reduced	-	-	-	-	**(Son et al., 2011)**
FGSG_04220		Transcription factor	Reduced	Reduced	-	Delayed	Reduced	**(Son et al., 2011)**

N.D. : Not determined.
"-" indicates that the phenotype of mutants is not affected.

Table 2. Secondary metabolite genes in *Fusarium graminearum*.

Gene	F. gram	F. vert[a]	F. oxy[b]	Production	Reference
PKS1	FGSG_17387	-	-	Unknown	-
PKS2	FGSG_04694	FVEG_00079	-	Unknown	-
PKS3	FGSG_17168	FVEG_03695	FOXG_05816	Perithecium pigment	(Gaffoor et al., 2005)
PKS4	FGSG_17745	-	-	Zearalenone	(Gaffoor and Trail, 2006; Kim et al., 2005; Lysoe et al., 2006)
PKS5	FGSG_17677	-	FOXG_10805	Unknown	-
PKS6	FGSG_08208	-	-	Unknown	-
PKS7	FGSG_08795	FVEG_01914	FOXG_03051	Unknown	-
PKS8	FGSG_03340	FVEG_10535	FOXG_11954	Unknown	-
PKS9	FGSG_10464	-	-	Fusarielins	(Sorensen et al., 2012)
PKS10	FGSG_07798	FVEG_11086	-	Fusarins	(Diaz-Sanchez et al., 2012)
PKS11	FGSG_01790	-	-	Unknown	-
PKS12	FGSG_02324	-	-	Aurofusarin	(Malz et al., 2005)
PKS13	FGSG_15980	-	-	Zearalenone	(Gaffoor and Trail, 2006; Kim et al., 2005; Lysoe et al., 2006)
PKS14	FGSG_03964	-	-	Unknown	-
PKS15	FGSG_04588	-	-	Unknown	-
NRPS1	FGSG_11026	FVEG_12503	FOXG_17422	Malonichrome	(Bushley et al., 2008)
NRPS2	FGSG_05372	FVEG_04296	FOXG_06448	Ferricrocin	(Tobiasen et al., 2007)
NRPS3	FGSG_10523	FVEG_03243	FOXG_04898	Unknown	-
NRPS4	FGSG_02315	FVEG_11762	FOXG_13024	Unknown	-
NRPS5	FGSG_17487	-	-	Unknown	-
NRPS6	FGSG_03747	FVEG_08697	FOXG_09785	Fusarinine	(Oide et al., 2006)
NRPS7	FGSG_08209	-	-	Unknown	-
NRPS8	FGSG_15673	-	-	Unknown	-
NRPS9	FGSG_10990	-	-	Unknown	-
NRPS10	FGSG_06507	FVEG_05643	FOXG_02458	Unknown	-
NRPS11	FGSG_03245	FVEG_08516	FOXG_09998	Unknown	-
NRPS12	FGSG_17574	FVEG_11841	FOXG_13405	Unknown	-
NRPS13	FGSG_13153	-	-	Unknown	-
NRPS14	FGSG_11395	-	-	Unknown	-
NRPS15	FGSG_02394	-	-	Unknown	-
NRPS16	FGSG_15872	-	-	Unknown	-
NRPS17	FGSG_10702	FVEG_14029	-	Unknown	-
NRPS18	FGSG_17386	-	-	Unknown	-
NRPS19	FGSG_15676	-	-	Unknown	-

a. Orthologous genes identified in *F. verticillioides*

b. Orthologous genes identified in *F. oxysporum* f.sp. *lycopersici*

References

Bai, G., and Shaner, G. (2004). Management and resistance in wheat and barley to fusarium head blight. Annu. Rev. Phytopathol. *42*, 135-161.

Bai, G.H., Desjardins, A.E., and Plattner, R.D. (2002). Deoxynivalenol-nonproducing *fusarium graminearum* causes initial infection, but does not cause disease spread in wheat spikes. Mycopathologia *153*, 91-98.

Bennett, J.W., and Klich, M. (2003). Mycotoxins. Clin. Microbiol. Rev. *16*, 497-516.

Bluhm, B.H., Zhao, X., Flaherty, J.E., Xu, J.R., and Dunkle, L.D. (2007). RAS2 regulates growth and pathogenesis in *Fusarium graminearum*. Mol. Plant Microbe Interact. *20*, 627-636.

Bottalico, A., and Perrone, G. (2002). Toxigenic *Fusarium* species and mycotoxins associated with head blight in small-grain cereals in Europe. Eur. J. Plant Pathol. *108*, 611-624.

Breakspear, A., Pasquali, M., Broz, K., Dong, Y., and Kistler, H.C. (2011). Npc1 is involved in sterol trafficking in the filamentous fungus *Fusarium graminearum*. Fungal Genet. Biol. *48*, 725-730.

Bushley, K.E., Ripoll, D.R., and Turgeon, B.G. (2008). Module evolution and substrate specificity of fungal nonribosomal peptide synthetases involved in siderophore biosynthesis. BMC evolutionary biology *8*, 328.

Calvo, A.M., Wilson, R.A., Bok, J.W., and Keller, N.P. (2002). Relationship between secondary metabolism and fungal development. Microbiol. Mol. Biol. Rev. *66*, 447-459.

Coccia, R., Foppoli, C., Blarzino, C., De Marco, C., and Rosei, M.A. (2001). Interaction of enkephalin derivatives with reactive oxygen species. Biochim. Biophys. Acta *1525*, 43-49.

Cuomo, C.A., Guldener, U., Xu, J.R., Trail, F., Turgeon, B.G., Di Pietro, A., Walton, J.D., Ma, L.J., Baker, S.E., Rep, M., *et al.* (2007). The *Fusarium graminearum* genome reveals a link between localized polymorphism and pathogen specialization. Science *317*, 1400-1402.

Dean, R., Van Kan, J.A.L., Pretorius, Z.A., Hammond-Kosack, K.E., Di Pietro, A., Spanu, P.D., Rudd, J.J., Dickman, M., Kahmann, R., Ellis, J., *et al.* (2012). The Top 10 fungal pathogens in molecular plant pathology. Mol. Plant Pathol. *13*, 414-430.

Desjardins, A.E., Bai, G., Plattner, R.D., and Proctor, R.H. (2000). Analysis of aberrant virulence of *Gibberella zeae* following transformation-mediated complementation of a trichothecene-deficient (Tri5) mutant. Microbiology *146*, 2059-2068.

Diaz-Sanchez, V., Avalos, J., and Limon, M.C. (2012). Identification and regulation of fusA, the polyketide synthase gene responsible for fusarin production in *Fusarium fujikuroi*. Appl. Environ.

Microbiol. *78*, 7258-7266.

Ding, S., Mehrabi, R., Koten, C., Kang, Z., Wei, Y., Seong, K., Kistler, H.C., and Xu, J.R. (2009). Transducin beta-like gene *FTL1* is essential for pathogenesis in *Fusarium graminearum*. Eukaryot. cell *8*, 867-876.

Dufresne, M., Hua-Van, A., El Wahab, H.A., Ben M'Barek, S., Vasnier, C., Teysset, L., Kema, G.H., and Daboussi, M.J. (2007). Transposition of a fungal miniature inverted-repeat transposable element through the action of a *Tc1*-like transposase. Genetics *175*, 441-452.

Dufresne, M., van der Lee, T., Ben M'barek, S., Xu, X., Zhang, X., Liu, T., Waalwijk, C., Zhang, W., Kema, G.H., and Daboussi, M.J. (2008). Transposon-tagging identifies novel pathogenicity genes in *Fusarium graminearum*. Fungal Genet. Biol. *45*, 1552-1561.

Fox, E.M., and Howlett, B.J. (2008). Secondary metabolism: regulation and role in fungal biology. Curr. Opin. Microbiol. *11*, 481-487.

Frandsen, R.J., Nielsen, N.J., Maolanon, N., Sorensen, J.C., Olsson, S., Nielsen, J., and Giese, H. (2006). The biosynthetic pathway for aurofusarin in *Fusarium graminearum* reveals a close link between the naphthoquinones and naphthopyrones. Mol. Microbiol. *61*, 1069-1080.

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

Gaffoor, I., and Trail, F. (2006). Characterization of two polyketide synthase genes involved in zearalenone biosynthesis in *Gibberella zeae*. Appl. Environ. Microbiol. *72*, 1793-1799.

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

Han, Y.K., Kim, M.D., Lee, S.H., Yun, S.H., and Lee, Y.W. (2007). A novel F-box protein involved in sexual development and pathogenesis in *Gibberella zeae*. Mol. Microbiol. *63*, 768-779.

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

Jansen, C., von Wettstein, D., Schafer, W., Kogel, K.H., Felk, A., and Maier, F.J. (2005). Infection patterns in barley and wheat spikes inoculated with wild-type and trichodiene synthase gene disrupted *Fusarium graminearum*. Proc. Natl. Acad. Sci. USA *102*, 16892-16897.

Jenczmionka, N.J., Maier, F.J., Losch, A.P., and 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.

Jiang, J., Yun, Y., Fu, J., Shim, W.B., and Ma, Z. (2011). Involvement of a putative response regulator FgRrg-1 in osmotic stress response, fungicide resistance and virulence in *Fusarium graminearum*. Mol. Plant Pathol. *12*, 425-436.

Jiang, L., Yang, J., Fan, F., Zhang, D., and Wang, X. (2010). The type 2C protein phosphatase FgPtc1p of the plant fungal pathogen *Fusarium graminearum* is involved in lithium toxicity and virulence. Mol. Plant Pathol. *11*, 277-282.

Keller, N.P., and Hohn, T.M. (1997). Metabolic pathway gene clusters in filamentous fungi. Fungal Genet. Biol. *21*, 17-29.

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

Li, Y., Wang, C., Liu, W., Wang, G., Kang, Z., Kistler, H.C., and Xu, J.R. (2011). The *HDF1* histone deacetylase gene is important for conidiation, sexual reproduction, and pathogenesis in *Fusarium graminearum*. Mol. Plant Microbe Interact. *24*, 487-496.

Lysoe, E., Klemsdal, S.S., Bone, K.R., Frandsen, R.J., Johansen, T., Thrane, U., and Giese, H. (2006). The *PKS4* gene of *Fusarium graminearum* is essential for zearalenone production. Appl. Environ. Microbiol. *72*, 3924-3932.

Lysoe, E., Pasquali, M., Breakspear, A., and Kistler, H.C. (2011). The transcription factor FgStuAp influences spore development, pathogenicity, and secondary metabolism in *Fusarium graminearum*. Mol. Plant Microbe Interact. *24*, 54-67.

Ma, L.J., van der Does, H.C., Borkovich, K.A., Coleman, J.J., Daboussi, M.J., Di Pietro, A., Dufresne, M., Freitag, M., Grabherr, M., Henrissat, B., *et al.* (2010). Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*. Nature *464*, 367-373.

Malz, S., Grell, M.N., Thrane, C., Maier, F.J., Rosager, P., Felk, A., Albertsen, K.S., Salomon, S., Bohn, L., Schafer, W., *et al.* (2005). Identification of a gene cluster responsible for the biosynthesis of aurofusarin in the *Fusarium graminearum* species complex. Fungal Genet. Biol. *42*, 420-433.

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

Merhej, J., Urban, M., Dufresne, M., Hammond-Kosack, K.E., Richard-Forget, F., and Barreau, C. (2012). The *velvet* gene, *FgVe1*, affects fungal development and positively regulates trichothecene biosynthesis and pathogenicity in *Fusarium graminearum*. Mol. Plant Pathol. *13*, 363-374.

Mesterhazy, A. (1995). Types and components of resistance to Fusarium head blight of wheat. Plant Breeding *114*, 377-386.

Nasmith, C.G., Walkowiak, S., Wang, L., Leung, W.W., Gong, Y., Johnston, A., Harris, L.J., Guttman, D.S., and Subramaniam, R. (2011). Tri6 is a global transcription regulator in the phytopathogen *Fusarium graminearum*. PLoS Pathog. *7*, e1002266.

Nguyen, L.N., Bormann, J., Le, G.T., Starkel, C., Olsson, S., Nosanchuk, J.D., Giese, H., and Schafer, W. (2011). Autophagy-related lipase FgATG15 of *Fusarium graminearum* is important for lipid turnover and plant infection. Fungal Genet. Biol. *48*, 217-224.

Oide, S., Liu, J., Yun, S.H., Wu, D., Michev, A., Choi, M.Y., Horwitz, B.A., and Turgeon, B.G. (2010). Histidine kinase two-component response regulator proteins regulate reproductive development, virulence, and stress responses of the fungal cereal pathogens *Cochliobolus heterostrophus* and *Gibberella zeae*. Eukaryot. cell *9*, 1867-1880.

Oide, S., Moeder, W., Krasnoff, S., Gibson, D., Haas, H., Yoshioka, K., and Turgeon, B.G. (2006). *NPS6*, encoding a nonribosomal peptide synthetase involved in siderophore-mediated iron metabolism, is a conserved virulence determinant of plant pathogenic ascomycetes. Plant cell *18*, 2836-2853.

Park, A.R., Son, H., Min, K., Park, J., Goo, J.H., Rhee, S., Chae, S.K., and Lee, Y.W. (2015). Autoregulation of *ZEB2* expression for zearalenone production in *Fusarium graminearum*. Mol. Microbiol. 97, 942-956.

Placinta, C.M., D'Mello, J.P.F., and Macdonald, A.M.C. (1999). A review of worldwide contamination of cereal grains and animal feed with *Fusarium* mycotoxins. Anim. Feed Sci. Tech. *78*, 21-37.

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

Proctor RH, McCormick SP, Alexander NJ, Desjardins AE (2009) Evidence that a secondary metabolic biosynthetic gene cluster has grown by gene relocation during evolution of the filamentous fungus Fusarium. Mol. Microbiol. 74,1128-1142.

Ramamoorthy, V., Zhao, X., Snyder, A.K., Xu, J.R., and Shah, D.M. (2007). Two mitogen-activated protein kinase signalling cascades mediate basal resistance to antifungal plant defensins in *Fusarium graminearum*. Cell. Microbiol. *9*, 1491-1506.

Rittenour, W.R., and Harris, S.D. (2008). Characterization of *Fusarium graminearum* Mes1 reveals roles in cell-surface organization and virulence. Fungal Genet. Biol. *45*, 933-946.

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

Seong, K., Li, L., Hou, Z., Tracy, M., Kistler, H.C., and Xu, J.R. (2006). Cryptic promoter activity in the coding region of the HMG-CoA reductase gene in *Fusarium graminearum*. Fungal Genet. Biol. *43*, 34-41.

Seong, K.Y., Pasquali, M., Zhou, X., Song, J., Hilburn, K., McCormick, S., Dong, Y., Xu, J.R., and Kistler, H.C. (2009). Global gene regulation by *Fusarium* transcription factors *Tri6* and *Tri10* reveals adaptations for toxin biosynthesis. Mol. Microbiol. *72*, 354-367.

Shim, W.B., Sagaram, U.S., Choi, Y.E., So, J., Wilkinson, H.H., and Lee, Y.W. (2006). *FSR1* is essential for virulence and female fertility in *Fusarium verticillioides* and *F. graminearum*. Mol. Plant Microbe Interact. *19*, 725-733.

Son, H., Seo, Y.S., Min, K., Park, A.R., Lee, J., Jin, J.M., Lin, Y., Cao, P., Hong, S.Y., Kim, E.K., *et al.* (2011). A phenome-based functional analysis of transcription factors in the cereal head blight fungus, *Fusarium graminearum*. PLoS Pathog. *7*, e1002310.

Sorensen, J.L., Hansen, F.T., Sondergaard, T.E., Staerk, D., Lee, T.V., Wimmer, R., Klitgaard, L.G., Purup, S., Giese, H., and Frandsen, R.J. (2012). Production of novel fusarielins by ectopic activation of the polyketide synthase 9 cluster in *Fusarium graminearum*. Environ. microbiol. *14*, 1159-1170.

Tobiasen, C., Aahman, J., Ravnholt, K.S., Bjerrum, M.J., Grell, M.N., and Giese, H. (2007). Nonribosomal peptide synthetase (*NPS*) genes in *Fusarium graminearum*, *F. culmorum* and *F. pseudograminearium* and identification of NPS2 as the producer of ferricrocin. Curr. Genet. *51*, 43-58.

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

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

Wang, C., Zhang, S., Hou, R., Zhao, Z., Zheng, Q., Xu, Q., Zheng, D., Wang, G., Liu, H., Gao, X., *et al.* (2011a). Functional analysis of the kinome of the wheat scab fungus *Fusarium graminearum*. PLoS Pathog. *7*, e1002460.

Wang, Y., Liu, W., Hou, Z., Wang, C., Zhou, X., Jonkers, W., Ding, S., Kistler, H.C., and Xu, J.R. (2011b). A novel transcriptional factor important for pathogenesis and ascosporogenesis in *Fusarium graminearum*. Mol. Plant Microbe Interact. *24*, 118-128.

Windels, C.E. (2000). Economic and social impacts of fusarium head blight: changing farms and rural communities in the northern great plains. Phytopathology *90*, 17-21.

Wong, P., Walter, M., Lee, W., Mannhaupt, G., Munsterkotter, M., Mewes, H.W., Adam, G., and Guldener, U. (2011). FGDB: revisiting the genome annotation of the plant pathogen *Fusarium*

graminearum. Nucleic Acids Res. *39*, D637-639.

Yu, H.Y., Seo, J.A., Kim, J.E., Han, K.H., Shim, W.B., Yun, S.H., and Lee, Y.W. (2008). Functional analyses of heterotrimeric G protein Gα and Gβ subunits in *Gibberella zeae*. Microbiology *154*, 392-401.

Yu, J.H., and Keller, N. (2005). Regulation of secondary metabolism in filamentous fungi. Annu. Rev. Phytopathol. *43*, 437-458.

Zhang, D., Fan, F., Yang, J., Wang, X., Qiu, D., and Jiang, L. (2010). FgTep1p is linked to the phosphatidylinositol-3 kinase signalling pathway and plays a role in the virulence of *Fusarium graminearum* on wheat. Mol. Plant Pathol. *11*, 495-502.

Chapter 2

EBR1, a novel Zn_2Cys_6 transcription factor, affects virulence and apical dominance of hyphal tip in *Fusarium graminearum*

Chunzhao Zhao[1,2,3,4], Cees Waalwijk[4], Pierre J.G.M. de Wit[5,6], Theo van der Lee[4] and Dingzhong Tang[1]

1. State Key Laboratory of Plant Cell and Chromosome Engineering, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing 100101, China
2. Graduate University of Chinese Academy of Sciences, Beijing 100049, China
3. Graduate School Experimental Plant Sciences, Wageningen, The Netherlands
4. Plant Research International, P.O. Box 6708 PB, Wageningen, The Netherlands
5. Wageningen University, Laboratory of Phytopathology, P.O. Box 6708 PB, Wageningen, The Netherlands
6. Department of Botany and Microbiology, King Saud University, Ryadh, Saudi Arabia

Published in Mol Plant Microbe Interact 24: 1407-1418 (2011)

Abstract

Zn_2Cys_6 transcription factors are unique to fungi and have been reported to be involved in different regulatory functions. Here we characterized EBR1 (Enhanced Branching 1), a novel Zn_2Cys_6 transcription factor of *Fusarium graminearum*. Knocking out *EBR1* in *F. graminearum* strain PH-1 caused reduction of both radial growth and virulence. The conidia of knock-out strain PH-1*Δebr1* germinated faster than those of wild-type PH-1, but the conidiation of the mutant was significantly reduced. Detailed analysis showed that the reduced radial growth might be due to reduced apical dominance of the hyphal tip leading to increased hyphal branching. Inoculation assays on wheat heads with a GFP-labeled PH-1*Δebr1* mutant showed that it was unable to penetrate the rachis of the spikelets. Protein fusion with GFP showed that EBR1 is localized in the nucleus of both conidia and hyphae. Knocking out the orthologous gene, FOXG_05408, in *F. oxysporum* f. sp. *lycopersici* caused a much weaker phenotype than the PH-1*Δebr1* mutant, which may be due to the presence of multiple orthologous genes in this fungus. Transformation of FOXG_05408 into PH-1*Δebr1* restored the mutant phenotype. Similar to EBR1, FOXG_05408 is localized in the nucleus of *F. oxysporum* f. sp. *lycopersici*. Possible functions of EBR1 and its relation with other fungal transcription factors are discussed.

Introduction

Gibberella zeae (anamorph *Fusarium graminearum*) is a worldwide occurring pathogen causing a wide range of cereal diseases including Fusarium head blight (FHB) on wheat and barley, and ear and stalk rot on corn (Bluhm et al. 2007; Lee et al. 2001). High humidity and temperatures between 20 and 25 °C at flowering time can result in serious FHB epidemics causing yield losses of up to 90% in wheat. *Fusarium* infection can result in the accumulation of mycotoxins such as deoxynivalenol (DON) and nivalenol (NIV) that are very harmful to humans and animals (Desjardins et al. 1993; McMullen et al. 1997; Proctor et al. 1995).

Several *F. graminearum* genes involved in virulence, growth, conidiation and sexual development have been studied extensively. Using a gene knock-out strategy, genes in the trichothecene cluster involved in the production of DON have been reported to be required for wheat ear infection. Strains that no longer produce these mycotoxins can infect the initial inoculated spikelet but are blocked to spread to neighboring spikelets (Bai et al. 2002; Desjardins et al. 1996; Desjardins et al. 2000). Also MAP kinases, such as *Gpmk1* have been shown to be involved in production of conidia, sexual development and virulence (Jenczmionka et al. 2003). A mutation in *Mgv1*, another MAP kinase, led to reduced virulence, reduced vegetative growth and increased sensitivity to the defensin MsDef1 (Hou et al. 2002; Ramamoorthy et al. 2007). Several additional genes involved in development and virulence have been identified. These include *FSR1*, encoding a putative multimodular protein, and essential for virulence and perithecia formation (Shim et al. 2006), and *RAS2*, a Ras GTPase-encoding gene essential for spore germination, vegetative growth and virulence (Bluhm et al. 2007). Also disruption of *NPS6*, encoding a non-ribosomal peptide synthetase, a conserved protein in plant pathogenic fungi such as *Cochliobolus miyabeanus*, *Fusarium graminearum* and *Alternaria brassicicola*, caused reduced virulence and reduced tolerance to H_2O_2 (Oide et al. 2006). Finally, disruption of *FGL1* encoding a secreted lipase results in reduced extracellular lipolytic activity and decreased virulence (Voigt et al. 2005).

However, the number of genes that have been functionally analyzed is still small compared to the 14000 genes predicted in the *F. graminearum* genome (Cuomo et al. 2007); many genes that are involved in virulence, vegetative growth and sexual development in *F.*

graminearum still need to be identified. To facilitate high-throughput functional analysis of *F. graminearum* genes, the transposable element *mimp1* was previously successfully employed to generate a transposon- inactivated gene pool (Dufresne et al. 2007; Dufresne et al. 2008). By this strategy we identified *rev112*, a mutant that showed reduced virulence on wheat head and reduced radial growth on potato dextrose agar. Tail-PCR showed that the transposable element *mimp1* was inserted in a putative Gal4-like Zn_2Cys_6 transcription factor, FGSG_10057 (Dufresne et al. 2008).

Zn_2Cys_6 transcription factors are unique for fungi and have been extensively studied in *Saccharomyces cerevisiae*, *Aspergillus flavus* and *Aspergillus nidulans* (Chang et al. 1995; Johnston 1987; Shimizu et al. 2003; Vienken et al. 2005; Woloshuk et al. 1994; Yu et al. 1996). In *S. cerevisiae*, Gal4 is one of the best studied Zn_2Cys_6 transcription factors that regulates galactose metabolism (Johnston 1987). In *A. flavus*, *AflR* encodes a Zn_2Cys_6 transcription factor that is involved in the regulation of aflatoxin biosynthesis (Woloshuk et al. 1994). In *A. nidulans*, the Zn_2Cys_6 transcription factor RosA is a negative regulator of sexual development and over- expression of *RosA* causes reduced hyphal growth (Vienken et al. 2005).

In *Fusarium* species only few Zn_2Cys_6 transcription factors have been functionally characterized. In *F. graminearum*, GIP2 is a Zn_2Cys_6 transcription factor that regulates expression of genes involved in the biosynthesis of the toxin aurofusarin (Kim et al. 2006). Two Zn_2Cys_6 transcription factors, FUM and ZFR1, in *Fusarium verticillioides*, are involved in the regulation of fumonisin biosynthesis (Brown et al. 2007; Flaherty and Woloshuk 2004). Fow2, a Zn_2Cys_6 transcription factor in *F. oxysporum* f. sp. *melonis* was reported to be required for virulence but not for vegetative growth and conidiation (Imazaki et al. 2007).

In this study, we identified and characterized EBR1 (Enhanced Branching 1), a novel type of Zn_2Cys_6 transcription factor in *F. graminearum* that is required for the radial growth, virulence, germination rate and conidiation. In addition, we investigated the function of one of the orthologous genes in *F. oxysporum* f. sp. *lycopersici* named FOXG_05408. We found that the FOXG_05408 knock-out mutant showed only a slight reduction of radial growth, but this gene could fully complement mutant PH-1*Δebr1*, which suggests a conserved function of EBR1-like Zn_2Cys_6 transcription factors in different fungi.

Results

EBR1, a Gal4-like Zn_2Cys_6 transcription factor, is conserved in fungi

We identified a Gal4-like Zn_2Cys_6 domain in EBR1 in the region between amino acid (aa) 282 and 308 according to SMART analysis (Schultz et al. 1998). As shown in Supplemental Table 1, there are 35 annotated Zn_2Cys_6 transcription factors in the *F. graminearum* genome that vary both in size, location of the Zn_2Cys_6 domain and in function. From those Zn_2Cys_6 transcription factors, EBR1 is the largest and the only protein with the Zn_2Cys_6 domain located after residue 280. The aa sequence of EBR1 was blasted against the database in the Broad Institute (http://www.broadinstitute.org/scientific-community/data) and nine additional EBR1-like Zn_2Cys_6 transcription factors were identified from nine different fungi (Supplemental Table 2). By aligning these ten EBR1-like Zn_2Cys_6 transcription factors, we identified, apart from the six conserved cysteine residues, several other conserved amino acids (Figure 1A). The overall length of EBR1-like Zn_2Cys_6 transcription factors is around 1100 aa and the Zn_2Cys_6 domain is located at around aa 230 to 310. The phylogenetic relationship between the EBR1-like Zn_2Cys_6 transcription factors is presented in Figure 1B.

Construction of a knock-out of *EBR1* in *Fusarium graminearum*

To confirm that the initially observed phenotype in *rev112* was only due to transposon inactivation and not to mutations elsewhere in the genome, we constructed a null mutant of *EBR1* by replacing the entire *EBR1* gene by homologous recombination in *F. graminearum* strains Fg820 and PH-1 using protoplast-mediated transformation (Figure 2A). We selected 20 hygromycin-resistant colonies from both strains and replacement of *EBR1* by the hygromycin cassette was confirmed by PCR in three transformants derived from Fg820 and one derived from PH-1 (PH-1*Δebr1*) (Figure 2B). All these knock-out strains exhibited reduced radial growth on CMA medium (Figure 2C, D; Supplemental Figure 1A) and reduced virulence on wheat heads, as compared to the wild-type strain and ectopic transformants (Figure 2E; Supplemental Figure 1B). In addition, all knock-out strains produced more pigment than wild-type strain and ectopic transformants when grown on CMA medium for three days (Figure 2C; Supplemental Figure 1A). Mutant strain

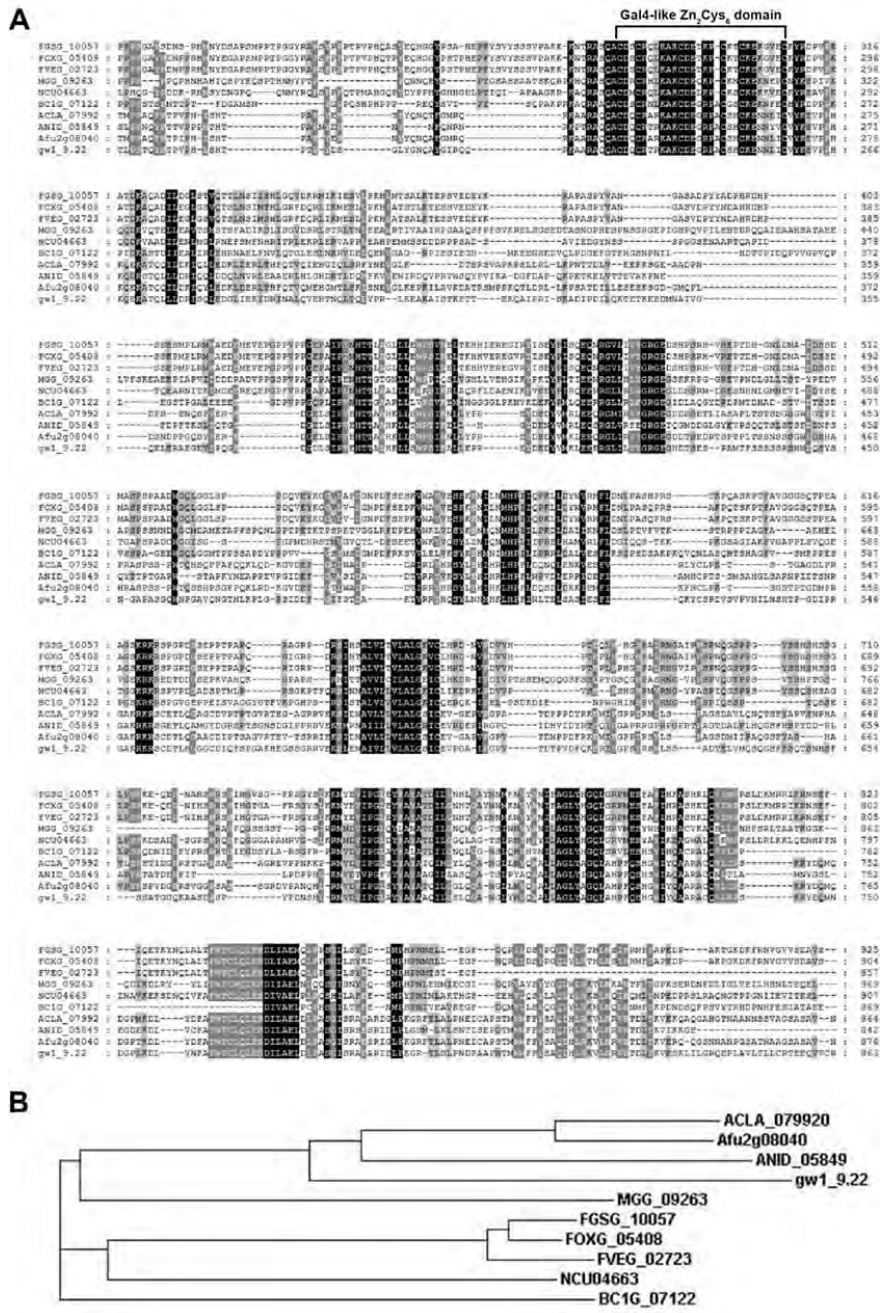


Figure 1. Phylogenetic analysis of EBR1-like Zn_2Cys_6 transcription factors. A. Alignment of EBR1-like Zn_2Cys_6 transcription factors from ten different fungi: *Aspergillus clavatus, A. fumigatus, A. nidulans, A. niger, Fusarium graminearum, F. oxysporum* f. sp. *lycopersici, F. verticillioides, Magnaporthe oryzae, Neurospora crassa* and *Botrytis cinerea.* The alignment

shows several blocks of conserved amino acids in addition to the six conserved cystein residues in the Gal4-like Zn_2Cys_6 domain. B. Phylogenetic tree constructed with the full length protein sequences of ten EBR1-like Zn_2Cys_6 transcription factors using the Clustal W program (Chenna et al. 2003).

PH-1*Δebr1* was used for further analysis.

PH-1*Δebr1* is defective in conidiation

The morphology of conidia from PH-1*Δebr1* was similar to that of wild-type and ectopic transformant. To investigate whether the mutant strain PH-1*Δebr1* was affected in conidiation, we used both mycelium and conidia as inoculum to determine the production of conidia. With mycelium as starting inoculum, same amounts of fresh mycelium of wild-type PH-1, PH-1*Δebr1* and ectopic transformant were picked from CMA medium and transferred to liquid mung bean medium for incubation. After three days of incubation, wild-type PH-1 and ectopic transformant produced about 5×10^5 conidia/ml, whereas the knock-out mutant PH-1*Δebr1* produced only 0.2×10^5 conidia/ml (Table 1). In contrast, when the same amount of conidia (10μl of a 10^5 conidia/ml suspension) of each strain was used as starting inoculum and incubated in liquid mung bean medium for three days, no significant differences in production of conidia were observed between wild-type and mutant (Table 1). These data indicate that the conidiation of mutant PH-1*Δebr1* is significantly affected when mycelium was used as starting inoculum, but not when conidia were used as starting inoculum.

Table 1. Conidiation and conidial germination rate of *Fusarium graminearum* strains.

Strain	Conidiation (10^5/ml) *		Germination rate (%) **
	Mycelium	Conidia	
PH-1	5.37 ± 0.67 [a]	4.26 ± 0.67	35.25 ± 8.98 [a]
PH-1*Δebr1*	0.17 ± 0.04 [b]	4.02 ± 1.55	68.50 ± 7.10 [b]
Ectopic	6.32 ± 0.80 [a]	4.02 ± 0.61	35.75 ± 7.19 [a]

*The mycelium or conidia of each strain were used as original inoculums to produce conidia. The conidia were counted three days after incubation in liquid mung bean medium. **100 random conidia of each strain were analyzed for germination 4 h after incubation. The different lower cases represent significant differences at $P \leq 0.01$. All the experiments were repeated at least 3 times.

The germination rate of conidia from PH-1*Δebr1* is increased as compared to wild-type

To determine whether the reduced radial growth of mutant PH-1*Δebr1* on CMA

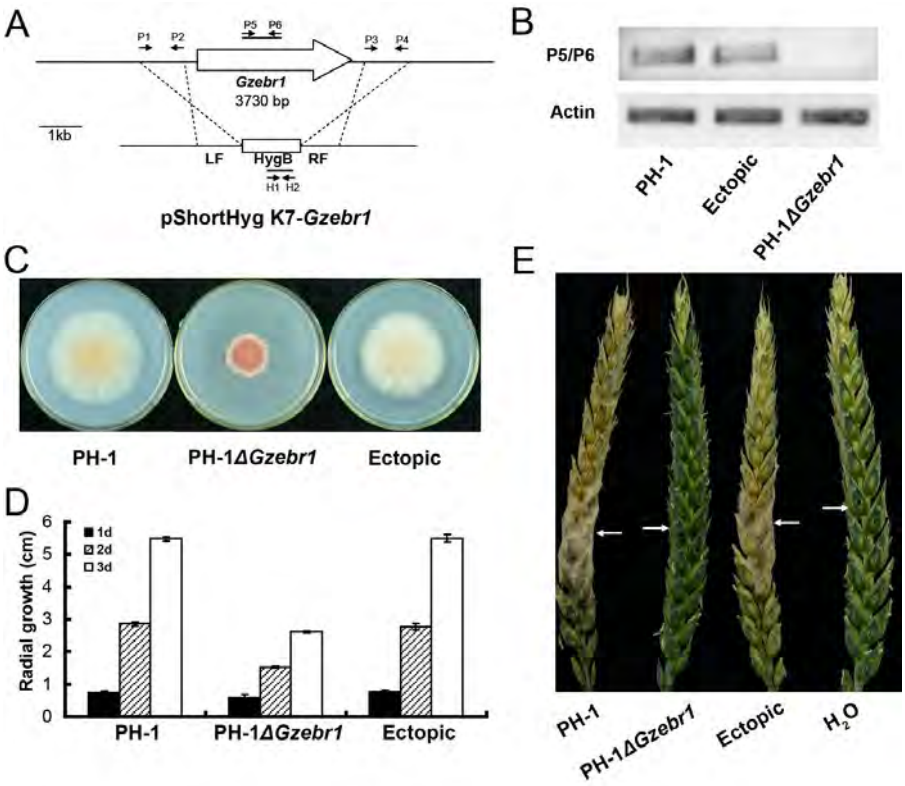


Figure 2. Properties of replacement mutant strain PH-1*Δebr1* of *F. graminearum*. A. The construct containing upstream (LF) and downstream (RF) fragments of the *EBR1* gene flanking HygB cassette was transformed into protoplasts of wild-type PH-1 strain to obtain replacement mutants. B. RT-PCR was performed to confirm replacement of *EBR1* by the HygB cassette in PH-1*Δebr1*. Actin expression was used as control. C. Strains PH-1, PH-1*Δebr1* and an ectopic transformant were incubated onto CMA medium and grew for three days. D. Radial growth of the various strains on CMA medium was measured at regular intervals until three days post inoculation (dpi). E. Wheat heads were inoculated with strains PH-1, PH-1*Δebr1* or ectopic transformant. Water-treated ears were used as control. Photographs were taken at 14 dpi. White arrows show the originally inoculated spikelets.

medium was due to a reduced germination rate of conidia, we analyzed the germination rate of conidia at 3, 4,, 5 and 6 hours after incubation on CMA medium. 100 randomly picked conidia of each strain were analyzed for germination. After 3 hours the germination rate of conidia of PH-1*Δebr1* (31%) appeared around three times as high as that of wild-type PH-1 and ectopic transformant (10 %) (Figure 3). After 4 and 5 hours, PH-1*Δebr1* still showed a higher germination rate than wild-type PH-1 and ectopic transformant. However, after 6 hours

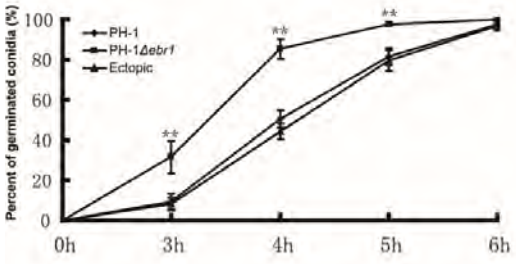


Figure 3. Germination rate of conidia on solid CMA medium. The germination of conidia was detected at 3, 4, 5 and 6h after incubation by microscopy. Before 6h, knock-out mutant PH-1*Δebr1* germinated faster than wild-type strain PH-1 and ectopic transformant. However, after 6h, nearly all conidia of all three strains had germinated. The experiment was repeated three times with similar results. * represents the significant differences at $P \leq 0.01$.

of incubation nearly all conidia of wild-type strain PH-1, mutant PH-1*Δebr1* and ectopic transformant had germinated (Figure 3). These data suggest that EBR1 might be a negative regulator of conidial germination.

Mycelium of PH-1*Δebr1* shows increased branching compared to wild-type PH-1

To determine which stage of growth was affected in PH-1*Δebr1*, we monitored the growth of hyphae after the transfer of spores on CMA medium for 24 h. After 8 hours no significant differences between PH-1*Δebr1*, wild-type PH-1 and ectopic transformant were observed. The conidia of all these strains had germinated and grew over the surface showing no significant differences in hyphal length. In contrast, at 24 h, PH-1*Δebr1* showed more branched hyphae than wild-type PH-1 and ectopic transformant (Figure 4A, B). This suggests that EBR1 has a positive effect on the apical dominance of the hyphal tip and a negative effect on hyphal branching.

PH-1*Δebr1* is impaired in infecting the rachis of wheat spikelets

To uncover which stage of the infection process was affected in PH-1*Δebr1*, we retransformed construct PSC001-*EBR1* into the wild-type PH-1 strain to obtain GFP-tagged EBR1 knock-out strain PSC*Δebr1* (Supplemental Figure 2A). PSC*Δebr1* showed reduced

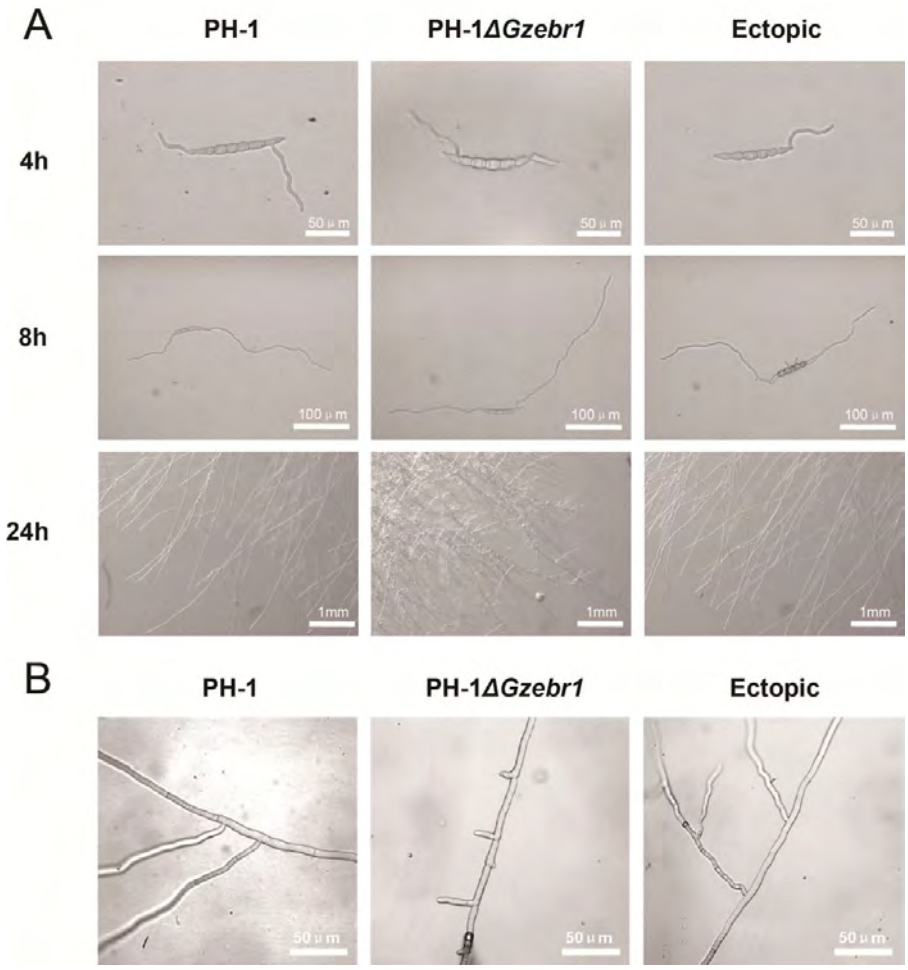


Figure 4. Growth of wild-type PH-1, replacement strain PH-1*Δebr1* and ectopic transformant on CMA medium. A. Conidia of strains PH-1, PH-1*Δebr1* and ectopic transformant were incubated on CMA medium. The growth of hyphae was monitored microscopically. Until eight hours of incubation no difference in growth pattern was observed, but at 24h the radial growth of PH-1*Δebr1* was reduced compared to the PH-1 wild-type with much more short branch hyphae. This experiment was repeated three times with consistent differences between PH-1*Δebr1* and PH-1. B. Mutant PH-1*Δebr1* develops many short branches when compared with wild-type PH-1 and ectopic transformant.

radial growth on CMA medium similar to PH-1*Δebr1* (Supplemental Figure 2B). As a control for the GFP-tagged wild-type strain, we transformed empty vector PSC001 into the wild-type strain to obtain the transformant GFPX19. Both PSC*Δebr1* and GFPX19 strains showed a clear GFP signal in their hyphae (Supplemental Figure 2C). In the wheat head virulence assay,

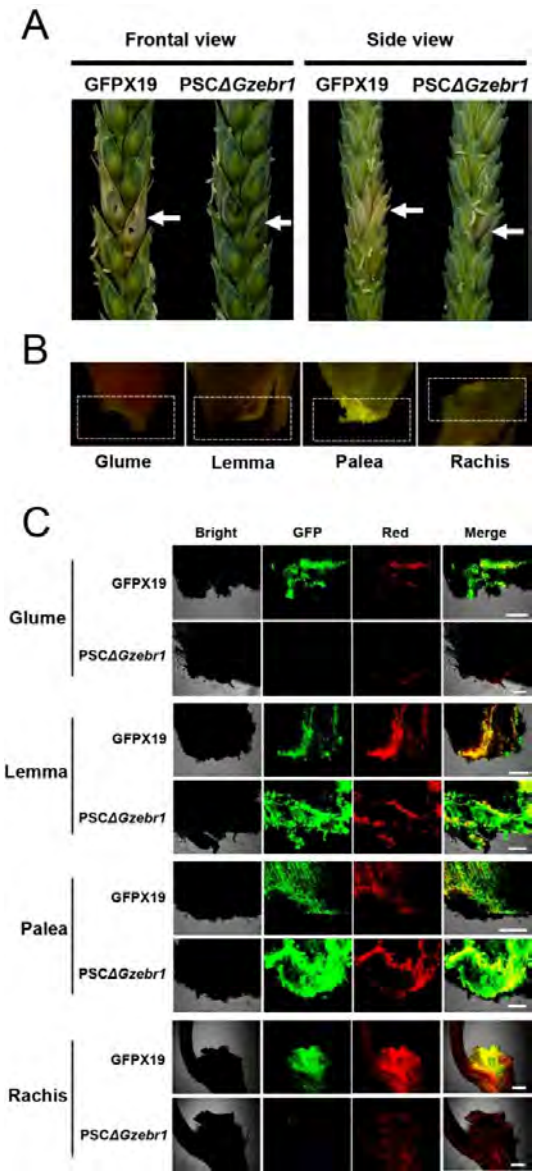


Figure 5. Wheat infection assay with GFP-tagged wild-type strain GFPX19 and replacement strain PSC*Δebr1*. A. Wheat spikelets were inoculated with strains GFPX19 and PSC*Δebr1* for seven days. Photographs show a frontal view (left) and a side view (right) of infected heads. The arrows point to the originally inoculated spikelets. B. Photographs show the morphology of glume, lemma, palea and spikelet rachis. The spikelet rachis and lower part of the glume, lemma, palea (white rectangle) were observed using a confocal microscopy. C. Many hyphae were detected at the lower part of lemma and palea for both strains. However, hyphae were detected only in the glume when inoculated with wild-type strain GFPX19 but not with replacement strain PSC*Δebr1*. In addition, GFPX19 had penetrated the rachis at seven dpi. In contrast, nearly no GFP signal was detected in the rachis inoculated with PSC*Δebr1*. GFP shows the signal of hyphae. Red shows the autofluorescence of chloroplast (Bar=100μm).

PSC*Δebr1* displayed reduced virulence when compared with strain GFPX19 (Figure 5A). The spikelets that were inoculated with either the PSC*Δebr1* or the GFPX19 strain were split into glume, lemma, palea and rachis at seven days post inoculation (Figure 5B). All tissues were analyzed by confocal microscopy to observe the presence of GFP signals. In total, 48

inoculated wheat spikelets for each isolate were used for analysis. At 7dpi, wild-type strain GFPX19 generated GFP signals in the glume, lemma, palea and rachis of nearly all inoculated spikelets, whereas PSC*Δebr1* had only penetrated lemma and palea and had partially penetrated glume, but nearly no GFP signal could be detected in rachis (Figure 5C, Supplemental Figure 3). These results indicate that mutant PSC*Δebr1* is able to infect lemma and palea and partially glume, but not rachis.

EBR1 is exclusively localized in the nucleus

EBR1 is responsible for conidial germination and hyphal growth on solid CMA medium. To further understand the function of EBR1 in hyphal growth, we studied the cellular location of EBR1 in conidia and hyphae. For this purpose, we tagged the *EBR1* genomic sequence with GFP under control of its native promoter and transformed it into wild-type strain PH-1. Several transformants were obtained and a clear GFP signal was observed by confocal microscopy. EBR1 appeared to be localized exclusively in the nucleus in both conidia and hyphae (Figure 6A). Interestingly, during the process of conidial germination, we observed two nuclei in each conidial cell (Figure 6B), with EBR1 located in both nuclei. To confirm that GFP- tagged EBR1 has the proper function in the cell, we transformed the GFP-tagged full genomic sequence into the mutant PH-1*Δebr1*. The transformants were selected on CMA medium containing geneticin (300µg/ml). In total 21 transformants were obtained and all of them did fully restore radial growth of mutant PH-1*Δebr1* to the level of wild-type strain PH-1 (Supplemental Figure 4). These data confirm that EBR1 indeed localizes in the nucleus and plays an important role in regulating hyphal growth.

Disruption of FOXG_05408 only slightly affects radial growth of *F. oxysporum* f. sp. *lycopersici*

To determine whether the orthologous gene FOXG_05408 in *F. oxysporum* f. sp. *lycopersici* (Fol) is also required for radial growth on CMA, it was replaced by the HygB cassette in Fol strain 4287 using homologous recombination (Figure 7A). Forty

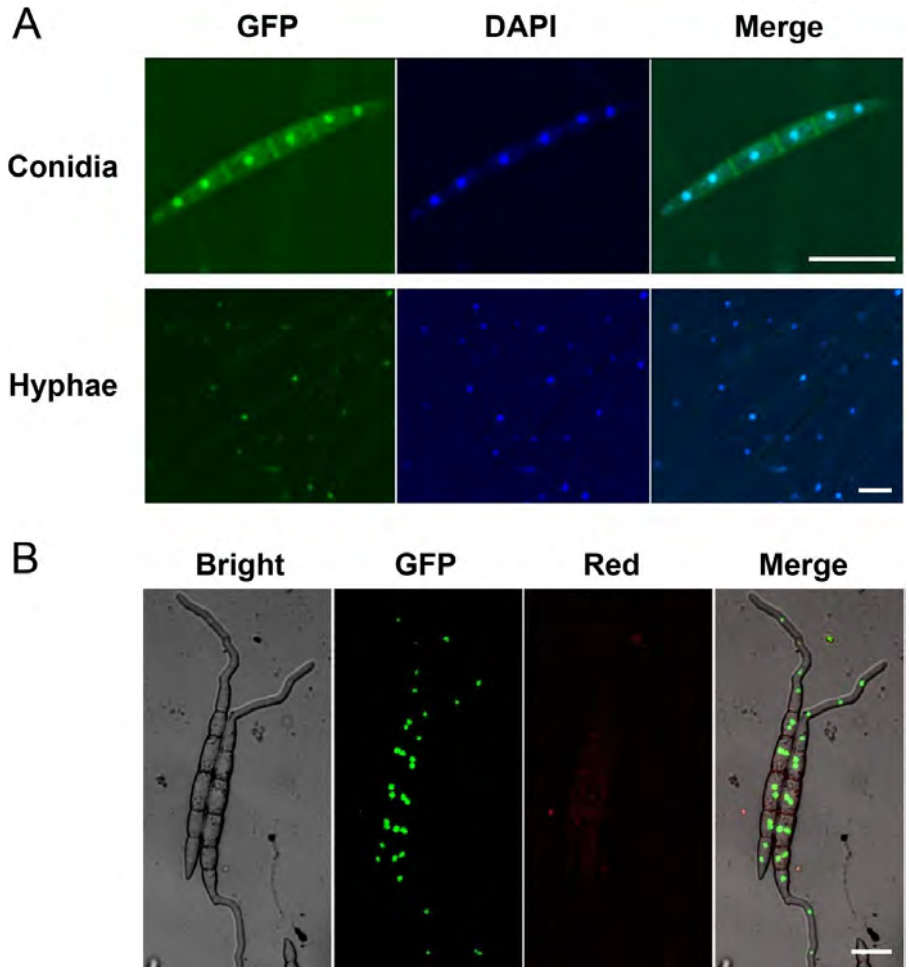


Figure 6. Expression of GFP-tagged EBR1. A. Localization of EBR1 was examined by confocal microscopy. EBR1 protein was found to localize in the nucleus of both conidia and hyphae. DAPI staining was used as control (Bar=10μm). B. In germinating conidia, two nuclei were observed in each cell of the conidium and EBR1 was localized in the nucleus of the germinating conidia (Bar=10μm).

hygromycin-resistant transformants were obtained as was confirmed by PCR. We continued further analysis with a transformant designated 4287*ΔFO05408*, in which the FOXG_05408 gene was replaced by the HygB cassette. Subsequently RT-PCR confirmed the absence of FOXG_05408 in strain 4287*ΔFO05408* (Figure 7B). The radial growth rate of 4287*ΔFO05408* on CMA medium was only slightly affected as compared to wild-type and ectopic transformant (Figure 7C, Supplemental Figure 5). This could be caused by functional

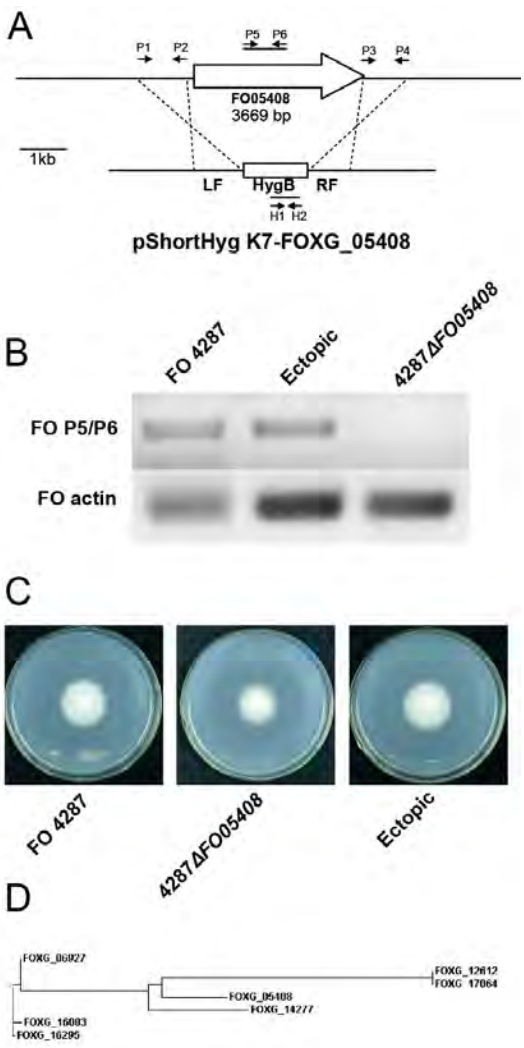


Figure 7. Replacement of FOXG_05408 in *F. oxysporum* f. sp. *lycopersici* by the HygB cassette. A. Strategy to construct the replacement strain 4287*ΔFO05408*. B. RT-PCR analysis to confirm the absence of FOXG_05408 in the replacement mutant strain 4287*ΔFO05408*. Actin was used as a control. C. Strains were incubated on CMA medium for determination of radial growth. Photographs were taken three days post incubation (dpi). Replacement mutant 4287*ΔFO05408* shows only a very slight reduction in radial growth when compared with wild-type FO 4287 and ectopic transformant. D. Paralogous genes of FOXG_05408 occurring in *F. oxysporum* f. sp. *lycopersici*.

redundancy as the *F. oxysporum* f. sp. *lycopersici* genome contains six potential paralogous genes of the FOXG_05408 gene (Figure 7D), whereas no potential paralogue of *EBR1* was identified in the *F. graminearum* genome.

The FOXG_05408 gene can complement the mutant strain PH-1*Δebr1*

FOXG_05408 is one of seven homologous genes present in the genome of *F. oxysporum* f. sp. *lycopersici*. To examine whether FOXG_05408 could complement mutant

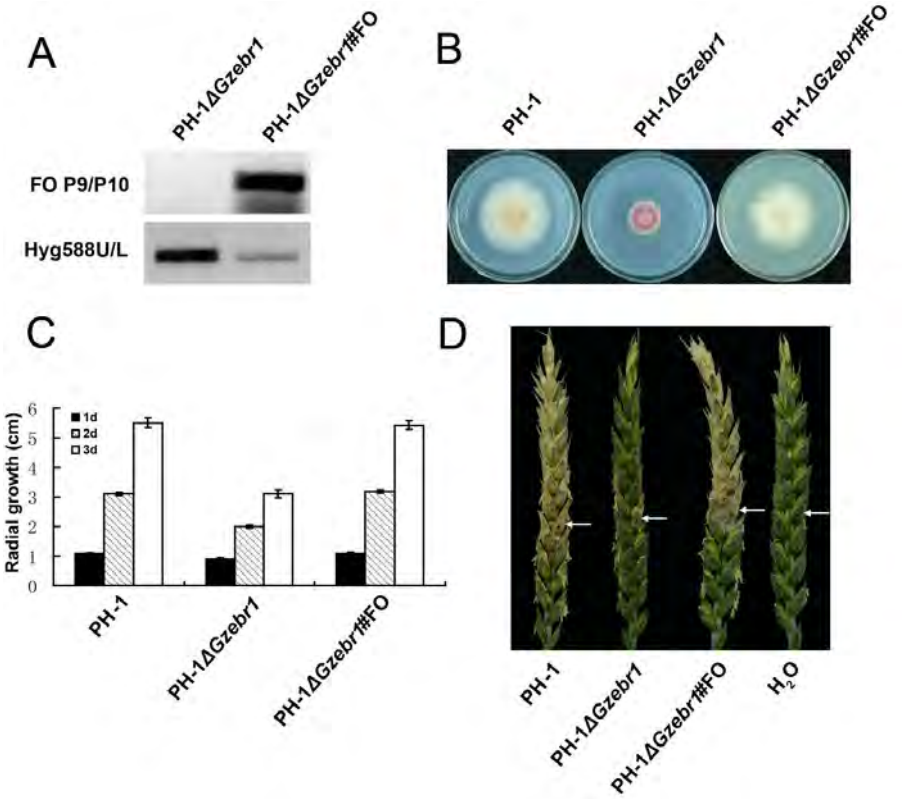


Figure 8. FOXG_05408 can complement the replacement strain PH-1*Δebr1*. A. The genomic sequence of FOXG_05408 was transformed into replacement mutant strain PH-1*Δebr1*. Geneticin resistant transformants were selected for PCR confirmation using FOXG_05408 specific primers and using the HygB fragment as a control. B-C. FOXG_05408 restored radial growth of replacement mutant PH-1*Δebr1* on CMA medium to wild-type level. D. FOXG_05408 restored the virulence of replacement mutant PH-1*Δebr1* strain to wild-type level in the wheat head assay. White arrows show the originally inoculated spikelets.

strain PH-1*Δebr1*, we retransformed PH-1*Δebr1* with FOXG_05408 under control of its own promoter by protoplast-mediated transformation. Geneticin-resistant colonies were selected to confirm the presence of the FOXG_05408 gene using FOXG_05408-specific primers (Figure 8A). Transformant PH-1*Δebr1*#FO could fully restore radial growth and virulence of PH-1*Δebr1* to wild-type (Figure 8B, C, D). Also PH-1*Δebr1*#FO showed restoration of conidiation and conidial germination of PH-1*Δebr1* to wild-type levels (Supplemental Table 3). These data prove that FOXG_05408 and EBR1 exhibit similar functions in regulation of hyphal growth, conidiation, conidial germination and virulence.

EBR1 is localized in the nucleus of both conidia and hyphae. To investigate whether FOXG_05408 is also localized in the nucleus, we transformed the GFP-tagged FOXG_05408 genomic sequence into *F. oxysporum* f. sp. *lycopersici* strain 4287 by protoplast-mediated transformation. As expected, FOXG_05408 is also localized in the nucleus of both conidia and hyphae (Supplemental Figure 6), which confirms a similar location of EBR1-like Zn_2Cys_6 transcription factors in different fungi.

Discussion

To sustain the continuous race with its host, timely development and a fast targeted growth are crucial for a fungal plant pathogen. *F. graminearum* is one of the fastest growing fungi, with a radial growth rate on PDA medium of 16-20 mm/day at 25 ^{o}C in the dark (Leslie and Summerell, 2006). During infection of wheat spikelets, fast directional growth allows the fungus to reach the rachis in less than three days after a conidium has landed on a spikelet. In this way, it will fully surround the developing seed and get access to the rachis, a key factor in colonization of the remaining part of the wheat head and possibly other parts of the wheat plant (Jansen et al. 2005). The Zn_2Cys_6 transcription factor EBR1 appears to play a key role for *F. graminearum* in regulation of the infection process.

Zn_2Cys_6 transcription factors have been widely studied in various fungi, especially in *Saccharomyces cerevisiae*, *Aspergillus nidulans* and *A. flavus* (Shimizu et al. 2003; Johnston 1987; Pan and Coleman 1990; Yu et al. 1996; Anderson et al. 1995), but much less so in fungal plant pathogens. Previous studies have shown that Zn_2Cys_6 transcription factors are involved, among others, in the regulation of the production of mycotoxins, secondary metabolites, meiosis and nitrate metabolism (Burger et al. 1991; Brown et al. 2007; Woloshuk et al. 1994; Anderson et al. 1995). In this study we found that the Zn_2Cys_6 transcription factor EBR1 in *F. graminearum* is involved in the regulation of hyphal growth and virulence.

Bioinformatic analysis showed the presence of orthologous *EBR1* genes in other fungi including *Aspergillus clavatus*, *A. fumigatus*, *A. nidulans*, *A. niger*, *F. oxysporum* f. sp. *lycopersici*, *F. verticillioides*, *Magnaporthe oryzae*, *Neurospora crassa* and *Botrytis cinerea*. In each fungus, except *F. oxysporum*, the orthologous gene was identified as a single copy

gene, encoding EBR1-like Zn_2Cys_6 transcription factors. Phylogenetic analysis showed that four EBR1-like Zn_2Cys_6 transcription factors from *Aspergillus* and three EBR1-like Zn_2Cys_6 transcription factors from *Fusarium* are categorized into one group. This suggests that EBR1-like Zn_2Cys_6 transcription factors could be candidates to study the evolutionary relationship between different fungi species.

EBR1-like Zn_2Cys_6 proteins represent a novel type of Zn_2Cys_6 transcription factors that are clearly different from those that have been described previously. Gal4 is a typical Zn_2Cys_6 transcription factor that contains a Cys-X_2-Cys-X_6-Cys-X_6 -Cys-X_2-Cys-X_6-Cys motif, which is reported to bind two Zn (Ⅱ) or Cd (Ⅱ) ions (Pan and Coleman 1990). Another well studied transcription factor is AflR of *A. nidulans*, a protein of 444 aa with the same motif as Gal4 (Shimizu et al. 2003). EBR1-like Zn_2Cys_6 transcription factors also contain a Zn_2Cys_6 DNA binding domain with the conserved Cys-X_2-Cys-X_6-Cys–$X_{5/6}$-Cys-X_2-Cys-X_6-Cys motif. However, unlike the Gal4 and AflR proteins of which the Zn_2Cys_6 DNA binding domain is located close to the N-terminus, the Zn_2Cys_6 DNA binding domain of EBR1-like Zn_2Cys_6 transcription factors is located further downstream between aa residues 230 and 310. In addition to the Zn_2Cys_6 domain, Gal4-like proteins contain a fungus-specific domain. In EBR1-like Zn_2Cys_6 transcription factors, besides the Zn_2Cys_6 domain, additional domains consisting of groups of highly conserved amino acid residues are present. However, the possible function(s) of these domains are not yet known.

In the mutant strain PH-1*Δebr1*, the production of conidia was significantly reduced when mycelium was used as inoculum. However, there was no difference between wild-type and mutant when conidia were used as inoculum. We hypothesize that increased hyphal branching affects conidiation. Unexpectedly, the initial germination rate of conidia of mutant strain PH-1*Δebr1* was higher than that of the wild-type within four hours after incubation on CMA medium. Until now, in *F. graminearum*, only a few genes have been reported to be associated with germination of conidia. A mutant disrupted in the Ras GTPase also showed delayed spore germination, but did not cause a significant phenotypic change in emergence of germ tubes from conidia and rate of conidial germination (Bluhm et al. 2007). The reasons why conidial germination is faster in mutant strain PH-1*Δebr1* than in wild-type is not clear,

but it might be that EBR1 represents a downstream component of the Ras GTPase signal transduction pathway that negatively regulates germination of conidia.

Detailed observation of hyphal development of mutant strain PH-1*Δebr1* on solid CMA medium showed that its radial growth rate is not affected until eight hours after incubation, but is significantly inhibited at 24 hours when increased branching of hyphae is clearly visible. We speculate that the reduced radial growth rate of mutant strain PH-1*Δebr1* is caused by reduction of the apical dominance of the hyphal tip leading to increased hyphal branching. The apical dominance of hyphal tips is a complex process which is regulated by various intracellular signals, including reactive oxygen species (Schmid and Harold, 1988) and calcium ions (Silverman-Gavrila and Lew, 2002, 2003). To date, no transcription factors have been reported involved in the apical dominance of hyphal tip and the Zn_2Cys_6 transcription factor EBR1 we identified represents the first example. The reason for the reduced hyphal tip growth of mutant PH-1*Δebr1* is not clear. One possibility is some genes that are directly regulated by EBR1 play an important role for the hyphal tip growth. In addition, it appears that mutant strain PH-1*Δebr1* produces more pigment than wild-type when grown on CMA medium. Previous studies have shown that the growth rate of a pigment-defective mutant (*ΔGIP2*) was increased approximately 30% relative to wild-type when grown on PDA medium. In contrast, constitutive expression of GIP2 in *Gibberella zeae* leads to high accumulation of pigment and reduced radial growth on PDA medium (Kim et al. 2006). EBR1 seems to have the opposite effect, and we conclude that it may have a negative regulatory effect on the expression of *GIP2.*

Also virulence of mutant strain PH-1*Δebr1* is significantly reduced when assayed on wheat heads. After 14 dpi, wild-type PH-1 has colonized the whole wheat head showing the typical symptoms of head blight, whereas mutant strain PH-1*Δebr1* only colonized the spikelets that were initially inoculated with conidia. The GFP-tagged PSC*Δebr1* strain clearly showed absence of penetration of the rachis at seven dpi, a phenotype that is reminiscent of naturally occurring resistance “type two” of wheat, where colonization of the spikelets neighboring the initially inoculated one is blocked. Similarly, previous studies have reported that infection of wheat heads by mutant *top1* (Topoisomerase I) was restricted to the

inoculated florets while wild-type strain had spread from inoculated spikelets to neighboring non-inoculated spikelets after 20 days (Baldwin et al. 2010). Also a novel transcription factor, ZIF1, was reported to be important for *F. graminearum* to penetrate the rachis (Wang et al. 2011). There could be different causes leading to the failure to penetrate the rachis. Here we speculate that reduced hyphal tip growth of PH-1*Δebr1* might allow the host to mobilize effective defense responses in time to block further colonization of the fungus.

Replacing the orthologous gene FOXG_05408 in *F. oxysporum* f. sp. *lycopersici* only slightly affected radial growth on CMA medium. This might be due to redundancy as six homologous genes are present in *F. oxysporum* f. sp. *lycopersici*, whereas in *F. graminearum* *EBR1* is a single copy gene. Retransforming PH-1*Δebr1* with FOXG_05408 fully restored the mutant to wild-type confirming conserved functions between EBR1 and FOXG_05408 in *F. graminearum* and *F. oxysporum* f. sp. *lycopersici*, respectively.

GFP-tagged-EBR1 showed that the protein is exclusively localized in the nucleus of both conidia and hyphae. Also the orthologous protein FOXG_05408 is localized in the nucleus of *F. oxysporum* f. sp. *lycopersici*. This suggests that EBR1-like Zn_2Cys_6 transcription factors are conserved proteins that are constitutively expressed. Previous reports have shown that the localization of Zn_2Cys_6 transcription factors is regulated by a G protein signaling pathway or by extracellular signals. For example, in *A. nidulans*, the Zn_2Cys_6 transcription factor, AflR, is located in the cytoplasm when protein kinase A (*pkaA*) is overexpressed, while it resides in the nucleus when an RGS (regulator of G-protein signaling) gene *FlbA* is inactivated (Shimizu et al. 2003). AmyR is another Zn_2Cys_6 transcription factor that regulates the expression of amylolytic genes in *A. nidulans*. The location of AmyR is dependent on isomaltose. In the absence of isomaltose, AmyR locates in the nucleus and cytoplasm, but in the presence of isomaltose, AmyR locates exclusively in the nucleus (Makita et al. 2009). The mechanism of regulation varies between different types of Zn_2Cys_6 transcription factors. It is possible that both the germination of conidia and hyphal tip growth require constitutive expression of *EBR1* in the nucleus. However, it remains to be investigated which domain of EBR1 is required for its nuclear localization.

Materials and Methods

Plants

The spring wheat cultivar Thassos was used in this study. Plants were grown in the greenhouse at 18 °C and a 16 h light/8 h dark regime. At anthesis, spikelets were inoculated with 5 µl of a conidial suspension (10^5 conidia/ml) of *F. graminearum*.

Fungal strains and culture conditions

The *F. graminearum* strains Fg820 and PH-1 were used in this study. Fg820 was isolated from infected wheat plants in the Netherlands and PH-1 is the sequenced strain that originates from a wheat plot in Michigan, USA (Cuomo et al. 2007). The sequenced *F. oxysporum* f. sp. *lycopersici* strain 4287 was also used in parallel experiments. All strains were cultured on solid PDA plates (25 °C) and stored in 2ml Eppendorf tubes covered with 15% DMSO in liquid nitrogen.

For production of conidia, strains were grown on complete medium (CM) (Leach et al. 1982) amended with 2% agar (CMA) for three days, after which a small piece of CMA with mycelium (25mm^2) was transferred to 3 ml of liquid mung bean medium (Bai et al. 1996) and incubated on a rotary shaker for three to five days (25 °C, 200 rpm). The number of conidia was determined microscopically using a GLASSTIC® SLIDE 10 with GRIDS (Hycor Biomedical Inc. California, USA).

Growth, conidiation and germination assays

For growth assay of *F. graminearum*, conidia were diluted to a concentration of 10^5 conidia/ml and 500 conidia (5 µl) were transferred to 9 cm Petri plates containing CMA. The Petri plates were placed in an incubator (25 °C) in the dark and the radial growth of *F. graminearum* was measured at 24 h, 48 h and 72 h after incubation. To determine the radial growth of *F. oxysporum* f. sp. *lycopersici*, conidia were diluted to 10^6 conidia/ml and 5 µl conidial suspension was spotted in the center of CMA plates and growth rate measurements were conducted after 3 days.

For conidiation assays using mycelium as initial inoculum, an equal volume (3 mm^2

from CMA medium) of fresh mycelium of each strain was transferred into 3 ml liquid mung bean medium and incubated for three days in a rotary shaker (25 °C, 200 rpm). For conidiation assays using conidia as initial inoculum, a 10 µl conidial suspension (10^5 conidia/ml) of each strain was transferred to 3 ml liquid mung bean medium and incubated for three days in a rotary shaker (25 °C, 200 rpm). The number of conidia was determined microscopically using a GLASSTIC® SLIDE 10 with GRIDS (Hycor Biomedical Inc. California, USA).

For germination assays, 5 µl conidial suspension (10^5 conidia/ml) were pipetted onto a glass slide containing a thin layer of CMA and incubated in a 25 °C incubator. The germination rate of 100 randomly selected conidia was counted microscopically (Olympus CX31) 4 h after incubation. The same samples were also used to monitor the growth of hyphae at 8 h and 24 h after incubation. The experiments were repeated twice.

Virulence assays

When the first anthers extruded from the wheat heads, 5 µl conidial suspension (10^5 conidia/ml) was pipetted into the floral cavity between the lemma and palea in one single floret of a spikelet. For each head, two spikelets (the ninth and tenth spikelet from the bottom of wheat head) were inoculated and the plants were kept in greenhouse at 20 °C and a 16 h light/8 h dark regime. The symptoms were observed at 14 days post inoculation (dpi).

To follow the infection process in wheat heads by GFP-tagged strains, the infected wheat heads were cut at seven dpi and the spikelets were carefully split into glume, lemma, palea and rachis. For glume, lemma and palea, the lower part of tissue was analyzed, respectively. For rachis, the tissue was first cut vertically into a thin piece. All tissues were placed onto a microscopic slide and the GFP signal was detected by confocal microscopy.

Bioinformatic analysis

Nine orthologous genes of *EBR1* were identified by BLAST analysis of the complete protein sequence of EBR1 against nine known fungal genomes available in the BROAD database. The sequences of ten proteins were aligned using Megalign software (DNASTAR,

Inc,) and the alignment was edited with Genedoc software (Nicholas et al. 1997). To construct a phylogenetic tree, ten EBR1-like transcription factors were aligned using the Clustal W program (Thompson et al. 1994).

DNA and RNA isolation

Genomic DNA was extracted from fresh mycelium by the CTAB method as described previously (Saghai-Maroof et al. 1984). For RNA isolation, fresh hyphae were collected from three-day-old cultures grown on CMA medium. RNA was extracted using TRIzol reagent (Invitrogen, Cat. No. 15596-018) according to the manufacturer's instructions.

Generation of transformation constructs

To prepare a construct for knocking out the *EBR1* gene, two fragments upstream and downstream of the *EBR1* gene were amplified from genomic DNA using primers Fg10057 P1/P2 and Fg10057 P3/P4, respectively (Primers are listed in Supplemental Table 4). PCR was conducted according to the instruction of Phusion High-Fidelity DNA polymerase (Finnzymes, Espoo, Finland). The amplified fragments were isolated from agarose gel and ligated into the vector pShortHyg K7 (Supplemental Figure 7A) using the restriction sites *Bam*HI/*Xba*I and *Kpn*I/*Xho*I, respectively. The resulting construct was named pShortHyg K7-*EBR1* (Figure 2A). The knock-out construct for the orthologous gene FOXG_05408 of *F. oxysporum* f. sp. *lycopersici* was prepared in a similar way. Primers FO05408 P1/P2 and FO05408 P3/P4 were used to amplify an upstream and downstream fragment of FOXG_05408, respectively. The same PCR protocol was used for amplification and both fragments were ligated into pShortHyg K7 to obtain the knock-out construct pShortHyg K7-FOXG_05408 (Figure 7A). To construct the GFP-tagged *EBR1* gene, the *EBR1* genomic sequence together with its 1613 bp promoter region was amplified from genomic DNA using primers Fg10057 P11/P12. The PCR product was digested with *Spe*I and ligated to the *Spe*I site of vector pShortHyg K7-GFP-Tnos (Supplemental Figure 7B). The obtained construct pShortHyg K7-*EBR1* -GFP-Tnos was used for transformation. For GFP-tagged *EBR1* knock-out strain, vector PSC001 (Supplemental Figure 7C) was used to obtain the final

construct PSC001-*EBR1* using the same procedure described for the construct pShortHygK7-*EBR1*. For the complementation assay, genomic sequence of FOXG_05408 under control of its native promoter (1570bp) was amplified using primers FO05408 P7/P8. The PCR product was ligated to the *Spe*I site of vector pShortGen K7-Tnos which contains the geneticin cassette (Supplemental Figure 7D).

Protoplast-mediated transformation

Transformation of *F. graminearum* and *F. oxysporum* f. sp. *lycopersici* was performed as described previously (Turgeon et al. 1987) with some minor modifications. Approximately 10^5 conidia grown in liquid mung bean medium were transferred to 50ml liquid CM medium. The conidia were incubated for two days on a rotary shaker (25 °C, 150 rpm) to produce mycelium. The mycelium was homogenized by vortexing and 10ml of the homogenate was transferred to 200ml fresh CM liquid medium and incubated overnight (25 °C, 150 rpm). The mycelium was collected by filtration on sterile cheese cloth and washed twice with sterile water. The mycelium was transferred to 20 ml Driselase/Glucanase (2.5%/1.5% in 700mM NaCl, PH 5.6) in a 50ml tube and incubated for 2.5 h-3h (22 °C, 43.5 rpm). When the number of protoplasts was higher than 10^7/ml, the protoplasts were filtered (48µm) and 20 ml 700mM NaCl was added to the protoplast filtrate. The protoplasts were collected by centrifugation at 1300g for 5 min (Centrifuge 5810 R, Eppendorf, Hamburg, Germany) and resuspended in 10ml ice-cold 700mM NaCl. The protoplasts were washed and collected by centrifugation at 830g for 5 min. The washing step was repeated. Finally, the protoplasts were collected by centrifugation (830g for 5 min) and resuspended in STC (0.8M sorbitol, 50mM Tris HCl PH 8.0, 50mM $CaCl_2$) with SPTC (4:1) (0.8M sorbitol, 40% polyethylene glycol 4000, 50mM Tris-HCl pH 8.0, 50mM CaCl2) at a final protoplast concentration of 0.5-2×10^8/ml. To 100µl protoplast suspension, 30 µg plasmid and 5µl heparin (5mg/ml in STC) were added in 2 ml Eppendorf tubes. The mixture was incubated on ice for 30 min, after which 1 ml SPTC was added to the mixture and incubated for 20 min at 22 °C. The mixture was added gently to 200ml regeneration medium (yeast extract (0.1%), casein hydrolysate (0.1%), sucrose (34.2%), granulated agar (1.6%)) at 43 °C. The regeneration medium was poured in 20 ml

portions in 94 mm Petri dishes. The protoplasts were incubated for 12-24 h at 28 °C after which the selection top layer was added (1.2% granulated agar in water containing 300µg/ml hygromycin B or 300µg/ml geneticin) to the plates. The plates were incubated at 28 °C for three days. Colonies that grew through the selection medium were transferred to CMA medium containing 100 µg/ml hygromycin B or 100 µg/ml geneticin and incubated at 28°C.

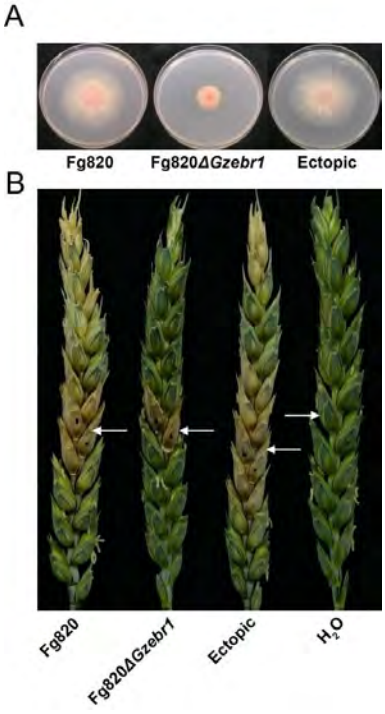
Validation of the knock-out transformants by PCR

DNA was isolated from the mycelium of transformants as described above. The DNA was used as template for the PCR amplification using the primers from the HygB cassette to confirm the presence of the *HygB* gene and the primers from the ORF region of the genes to confirm whether the targeted gene has been knocked out. Positive knock-out strains were selected for RNA isolation and reverse transcription (RT)-PCR was performed to further confirm the absence of mRNA.

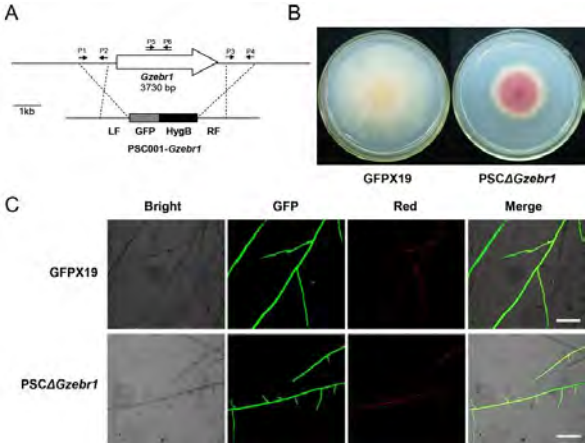
Acknowledgements

We thank Dr. M. Rep for kindly providing *F. oxysporum* f. sp. *lycopersici* strain 4287. This work was supported by grant from National Basic Research Program of China (2011CB100700). P.J.G.M. is supported by grants from the Royal Netherlands Academy of Arts and Sciences and the Centre for BioSystems Genomics.

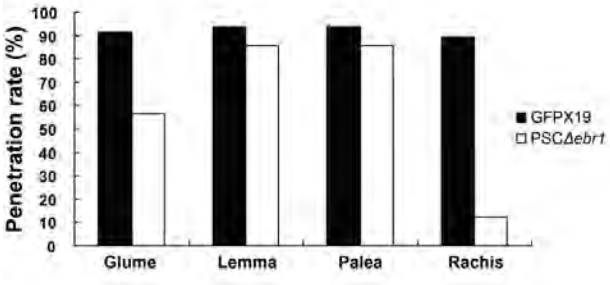
Supplementary data



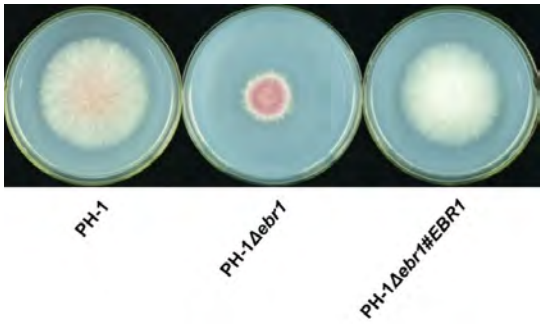
Supplemental Figure 1. Knock-out mutant of *EBR1* in *Fusarium graminearum* Fg820. A. Fg820*Δebr1* shows reduced radial growth on CMA medium. Photographs were taken three days after incubation. B. Fg820*Δebr1* shows reduced virulence compared with wt Fg820 and ectopic strains. Photographs were taken14 days post inoculation. White arrows show the originally inoculated spikelets.



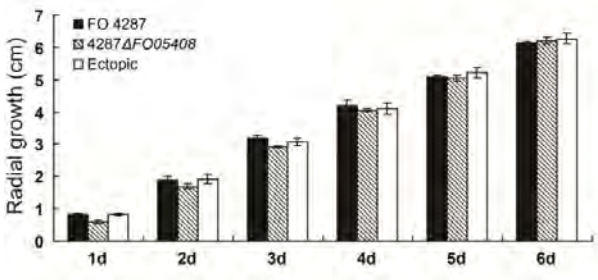
Supplemental Figure 2. Phenotype of knock-out mutant of *EBR1* in *Fusarium graminearum* strain PH-1 (PSC*Δebr1*). A. Knock-out construct PSC001-*EBR1* with upstream fragment (LF) and downstream fragment (RF) of *EBR1* flanking the Hygromycin and GFP cassettes. B. Transformant PSC*Δebr1* shows reduced radial growth on CMA medium compared with GFP-tagged wild-type strain GFPX19 three days after incubation (Bar=10mm). C. Strains GFPX19 and PSC*Δebr1* show a strong GFP signal. PSC*Δebr1* develops more short branched hyphae compared with the GFPX19 control strain (Bar=50μm).



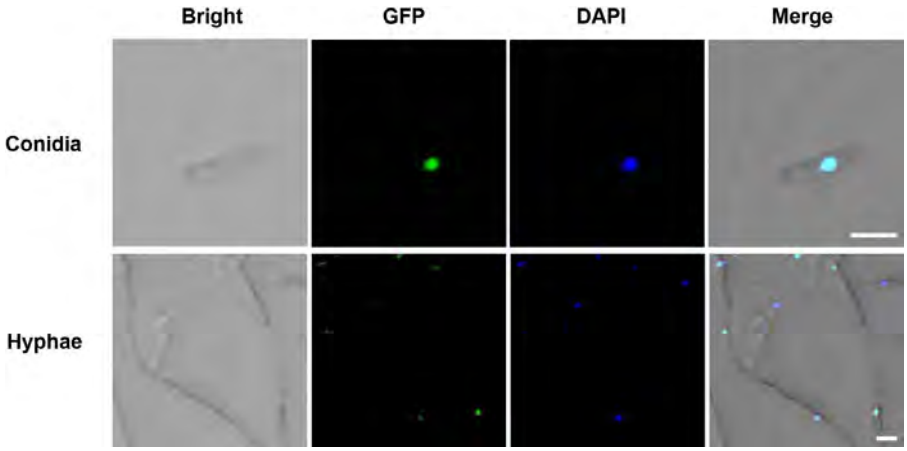
Supplemental Figure 3. Mutant PSC*Δebr1* is blocked in penetration of the rachis of wheat head. In total 48 inoculated spikeletes for each isolate were analyzed for penetration of glume, lemma, palea and rachis at 7 days post incubation (dpi). Similar to wild-type strain, mutant PSC*Δebr1* was able to penetrate the lemma and palea. However, the ability of mutant PSC*Δebr1* to penetrate the glume was reduced compared to wild-type strain. Mutant PSC*Δebr1* can hardly penetrate the rachis, while wild-type strain can fully penetrate and infect the rachis.



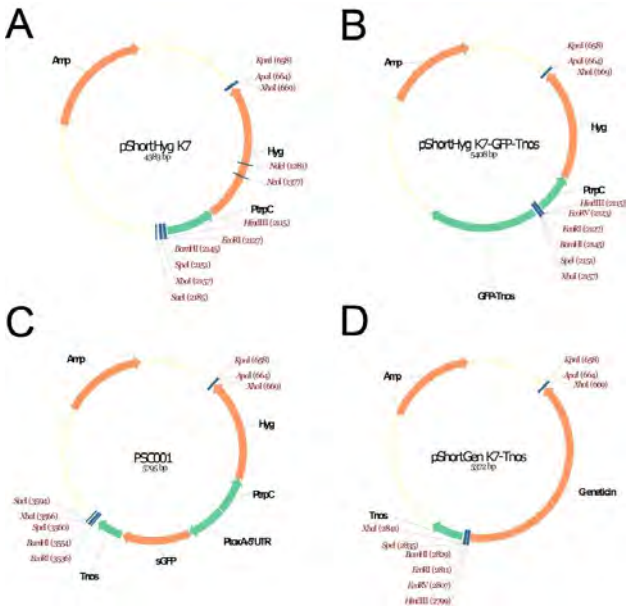
Supplemental Figure 4. Complementation of mutant PH-1*Δebr1*. GFP-tagged *EBR1* genomic sequence was transformed into mutant PH-1*Δebr1*. The transformant PH-1*Δebr1#EBR1* could fully restore radial growth of mutant PH-1*Δebr1* to wild-type level.



Supplemental Figure 5. Quantitative measurements of radial growth of *F. oxysporum* isolates on CMA medium. The conidia of wild-type stain FO 4287, knock-out mutant 4287*ΔFO05408* and ectopic transformant were dropped onto CMA medium and incubated. The radial growth was measured at regular intervals until six days post inoculation (dpi). Compared to mutant PH-1*Δebr1*, the radial growth of mutant 4287*ΔFO05408* was only slightly affected.



Supplemental Figure 6. Localization of FOXG_05408 in *F. oxysporum* f. sp. *lycopersici*. GFP-tagged FOXG_05408 genomic sequence under its own promoter was transformed into *F. oxysporum* f. sp. *lycopersici* strain 4287. FOXG_05408 localizes in the nucleus of both conidia and hyphae (Bar=5μm).



Supplemental Figure 7. Vectors used in this study. A. Vector pShortHyg K7 was used to knock out genes *EBR1* and FOXG_05408. B. Vector pShortHyg K7-GFP- Tnos used to tag *EBR1* with GFP for visualization of EBR1 in *F. graminearum*. C. Vector PSC001 used to knock out *EBR1* in the GFP-tagged strain. D. Vector p*ShortGen K7-Tnos containing the geneticin cassette used to complement mutant strain PH-1Δebr1* with genomic sequence of FOXG_05408.

Supplemental Table 1. Comparison of Zn_2Cys_6 domain-containing proteins in *Fusarium graminearum*.

Gene	Protein length	Location of domain (from ATG)	Length of domain	Other domain	Predicted function
FGSG_12134	148 aa	aa 52 – aa 78	27 aa	No	Transcriptional activator xlnR
FGSG_12714	490 aa	aa 65 – aa 91	27 aa	No	Transcriptional activator xlnR
FGSG_04311	505 aa	aa 40 – aa 67	28 aa	No	Hypothetical protein
FGSG_12631	515 aa	aa 166 – aa 192	27 aa	No	Hypothetical protein
FGSG_07591	527 aa	aa 23 – aa 51	29 aa	No	Hypothetical protein
FGSG_03857	558 aa	aa 46 – aa 73	28 aa	No	Hypothetical protein
FGSG_13386	563 aa	aa 28 – aa 54	27 aa	Fungal specific transcription factor domain	Hypothetical protein
FGSG_04191	590 aa	aa 85 – aa 112	28 aa	No	Hypothetical protein
FGSG_05068	611 aa	aa 37 – aa 64	28 aa	No	Hypothetical protein
FGSG_00069	613 aa	aa 13 – aa 39	27 aa	Fungal specific transcription factor domain	Hypothetical protein
FGSG_03786	624 aa	aa 12 – aa 38	27 aa	No	Hypothetical protein
FGSG_12385	644 aa	aa 51 – aa 79	29 aa	Fungal specific transcription factor domain	Hypothetical protein
FGSG_11271	654 aa	aa 23 – aa 49	27 aa	Fungal specific transcription factor domain	Hypothetical protein
FGSG_02083	656 aa	aa 10 – aa 36	27 aa	Fungal specific transcription factor domain	Hypothetical protein
FGSG_13098	674 aa	aa 18 – aa 44	27 aa	Fungal specific transcription factor domain	Hypothetical protein
FGSG_10639	684 aa	aa 28 – aa 54	27 aa	Fungal specific transcription factor domain	Hypothetical protein
FGSG_00404	691 aa	aa 32 – aa 59	28 aa	Fungal specific transcription factor domain	Hypothetical protein
FGSG_03138	702 aa	aa 107 – aa 137	31 aa	Fungal specific transcription factor domain	Hypothetical protein
FGSG_03159	702 aa	aa 51 – aa 77	27 aa	Fungal specific transcription factor domain	Hypothetical protein
FGSG_05958	704 aa	aa 12 – aa 38	27 aa	Fungal specific transcription factor domain	Hypothetical protein
FGSG_03727	706 aa	aa 33 – aa 60	28 aa	Fungal specific transcription factor domain	Hypothetical protein
FGSG_07638	718 aa	aa 27 – aa 53	27 aa	No	Hypothetical protein
FGSG_08392	740 aa	aa 194 – aa 225	32 aa	Fungal specific transcription factor domain	Hypothetical protein
FGSG_06750	741 aa	aa 27 – aa 53	27 aa	Fungal specific transcription factor domain	Hypothetical protein
FGSG_02215	743 aa	aa 39 – aa 68	30 aa	Fungal specific transcription factor domain	Hypothetical protein

FGSG_07924	743 aa	aa 119 – aa 146	28 aa	No	Hypothetical protein
Gene	Protein length	Location of domain (from ATG)	Length of domain	Other domain	Predicted function
FGSG_06436	773 aa	aa 29 – aa 55	27 aa	Fungal specific transcription factor domain	Hypothetical protein
FGSG_09594	773 aa	aa 35 – aa 61	27 aa	Fungal specific transcription factor domain	Hypothetical protein
FGSG_02799	777 aa	aa 77 – aa 104	28 aa	Fungal specific transcription factor domain	Hypothetical protein
FGSG_01564	847 aa	aa 124 – aa 152	29 aa	Fungal specific transcription factor domain	Hypothetical protein
FGSG_09921	886 aa	aa 24 – aa 51	28 aa	Fungal specific transcription factor domain	Transcriptional activator protein acu-15
FGSG_06810	893 aa	aa 64 – aa 91	28 aa	Fungal specific transcription factor domain	Hypothetical protein
FGSG_08791	903 aa	aa 178 – aa 204	27 aa	Zinc finger, C2H2 type	Hypothetical protein
FGSG_02531	969 aa	aa 59 – aa 85	27 aa	Fungal specific transcription factor domain	Hypothetical protein
FGSG_10057	1136 aa	aa 282 – aa 308	27 aa	No	Hypothetical protein

Supplemental Table 2. Comparison of EBR1-like Zn_2Cys_6 transcription factors in fungi.

Gene	Protein length	Location of domain (from ATG)	Length of domain	Predicted function	*Organism*
ACLA_079920	1139 aa	aa 240 - aa 267	28 aa	n.d.	*Aspergillus clavatus*
Afu2g08040	1148 aa	aa 243 – aa 270	28 aa	C6 finger domain protein	*Aspergillus fumigatus*
ANID_05849	1077 aa	aa 236 – aa 263	28 aa	Conserved hypothetical protein	*Aspergillus nidulans*
gw1_9.22	1102 aa	aa 231 – aa 258	28 aa	n.d.	*Aspergillus niger*
BC1G_07122	1099 aa	aa 237 - aa 264	28 aa	Conserved hypothetical protein	*Botrytis Cinerea*
FGSG_10057	1172 aa	aa 282 – aa 308	27 aa	Hypothetical protein	*Fusarium graminearum*
FOXG_05408	1152 aa	aa 262 – aa 288	27 aa	Conserved hypothetical protein	*Fusarium oxysporum*
FVEG_02723	916 aa	aa 264 – aa 290	27 aa	Conserved hypothetical protein	*Fusarium verticillioides*
NCU04663	1183 aa	aa 258 – aa 284	27 aa	Hypothetical protein	*Neurospora crassa*
MGG_09263	1226 aa	aa 288 – aa 314	27 aa	C6 zinc finger domain- containing protein	*Magnaporthe oryzae*

Supplemental Table 3. Conidiation and conidial germination rate of *Fusarium graminearum* strains.

Strain	Conidiation (10^5/ml) * (Mycelium)	Germination rate (%)**
PH-1	3.48 ± 0.79 [a]	50.75 ± 4.78 [a]
PH-1*Δebr1*	0.24 ± 0.10 [b]	78.25 ± 2.36 [b]
PH-1*Δebr1*#FO	3.96 ± 0.66 [a]	46.50 ± 1.73 [a]

*The mycelium of each strain was used as original inoculum to produce conidia. The conidia were counted three days after incubation in liquid mung bean medium. **100 random conidia of each strain were analyzed for germination 4 h after incubation. The different lower cases represent the significant differences at $P \leq 0.01$.

Supplemental Table 4. Primers used in this study.

Primer	Sequence (5' to 3')
Fg10057-P1	TGCTCTAGATTATCTGCTTACCTGCTCTG
Fg10057-P2	CGCGGATCCGTCTGAAACGACTGGGTATG
Fg10057-P3	TGCGGGCCCTGACGGATAACGCAACAATAC
Fg10057-P4	CGGGGTACCTAGGGAACGAATCTGGGAGTC
Fg10057-P5	ATCGAGAGTACCCGTGCTTT
Fg10057-P6	GCAGTTCCTCATCCACATAG
Fg10057-P7	AACTGGGCGGACTCAGTCCT
Fg10057-P8	GGGATAGCTCCGTTGCGAAT
Fg10057-P9	CGACTGCTGCCAATACTACG
Fg10057-P10	CACTTTGCCTTTACTAACCTCC
Fg10057-P11	CGGACTAGTATAATCTCCCTTGTTTACCAC
Fg10057-P12	CGGACTAGTCTGCCCGTGCGATGTCATCTG
Hyg588U-H1	AGCTGCGCCGATGGTTTCTACAA
Hyg588L-H2	GCGCGTCTGCTGCTCCATACAA
Fg-Actin-F	ATCCACGTCACCACTTTCAA
Fg-Actin-R	TGCTTGGAGATCCACATTTG
FO05408 P1	TGCTCTAGATATTCGCTCAGCTAAAGAAA
FO05408 P2	CGCGGATCCGAACAAGGCGAAGAGAAAAG
FO05408 P3	CCGCTCGAGTAGAGTCGACGATATTTGCC
FO05408 P4	CGGGGTACCGAAGACTTTTCTTTTGGCGG
FO05408 P5	GATCCAGCACCTCCAGATAG
FO05408 P6	AGGGTAATGTTGATACGGGC
FO05408 P7	CGGACTAGTTCAAGCGCACCACCACAATC
FO05408 P8	CGGACTAGTTCACTGCCCGTGTGATGAGTTC
FO05408 P9	TCCTCCCCATCCTCCTCATCC
FO05408 P10	GTGGATGTTGATGCTGCTGTT
FO-Actin-F	CCCAAGACCCAAACTACATT
FO-Actin-R	TCCATATCATCCCAGTTGGT

References

Anderson, S.F., Steber, C.M., Esposito, R.E., and Coleman, J.E. (1995). Ume6, a negative regulator of meiosis in *Saccharomyces cerevisiae*, contains a C-Terminal Zn_2Cys_6 binuclear cluster that binds the URS1 DNA sequence in a zinc-dependent manner. Prot. Sci. *4*, 1832-1843.

Bai, G.H., Desjardins, A.E., and Plattner, R.D. (2002). Deoxynivalenol-nonproducing *Fusarium graminearum* causes initial infection, but does not cause disease spread in wheat spikes. Mycopathologia *153*, 91-98.

Baldwin, T.K., Urban, M., Brown, N., and Hammond-Kosack, K.E. (2010). A role for topoisomerase I in *Fusarium graminearum* and *F. culmorum* pathogenesis and sporulation. Mol. Plant Microbe. Interact. *23*, 566-577.

Bluhm, B.H., Zhao, X., Flaherty, J.E., Xu, J.R., and Dunkle, L.D. (2007). *RAS2* regulates growth and pathogenesis in *Fusarium graminearum*. Mol. Plant Microbe. Interact. *20*, 627-636.

Brown, D.W., Butchko, R.A.E., Busman, M., and Proctor, R.H. (2007). The *Fusarium verticillioides FUM* gene cluster encodes a Zn(II)2Cys6 protein that affects *FUM* gene expression and fumonisin production. Eukaryot. Cell *6*, 1210-1218.

Burger, G., Strauss, J., Scazzocchio, C., and Lang, B.F. (1991). *niaA*, the pathway-specific regulatory gene of nitrate assimilation in *Aspergillus nidulans*, encodes a putative GAL4-type zinc finger protein and contains 4 introns in highly conserved regions. Mol. Cell Biol. *11*, 5746-5755.

Chang, P.K., Ehrlich, K.C., Yu, J.J., Bhatnagar, D., and Cleveland, T.E. (1995). Increased expression of *Aspergillus parasiticus aflR*, encoding a sequence-specific DNA-binding protein, relieves nitrate inhibition of aflatoxin biosynthesis. Appl. Environ. Microb. *61*, 2372-2377.

Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T.J., Higgins, D.G., and Thompson, J.D. (2003). Multiple sequence alignment with the Clustal series of programs. Nucleic Acids Res. *31*, 3497-3500.

Cuomo, C.A., Gueldener, U., Xu, J.R., Trail, F., Turgeon, B.G., Di Pietro, A., Walton, J.D., Ma, L.J., Baker, S.E., Rep, M., *et al.* (2007). The *Fusarium graminearum* genome reveals a link between localized polymorphism and pathogen specialization. Science *317*, 1400-1402.

Desjardins, A.E., Bai, G.H., Plattner, R.D., and Proctor, R.H. (2000). Analysis of aberrant virulence of *Gibberella zeae* following transformation-mediated complementation of a trichothecene-deficient (Tri5) mutant. Microbiology *146*, 2059-2068.

Desjardins, A.E., Hohn, T.M., and Mccormick, S.P. (1993). Trichothecene biosynthesis in *Fusarium* species: chemistry, genetics, and significance. Microbiol. Rev. *57*, 595-604.

Desjardins, A.E., Proctor, R.H., Bai, G.H., McCormick, S.P., Shaner, G., Buechley, G., and Hohn, T.M. (1996). Reduced virulence of trichothecene-nonproducing mutants of *Gibberella zeae* in wheat field tests. Mol. Plant Microbe. Interact. *9*, 775-781.

Dufresne, M., Hua-Vant, A., el Wahab, H.A., Ben M'Barek, S., Vasnier, C., Teysset, L., Kema, G.H.J., and Daboussi, M.J. (2007). Transposition of a fungal miniature inverted-repeat transposable element through the action of a *Tc1*-like transposase. Genetics *175*, 441-452.

Dufresne, M., van der Lee, T., Ben M'Barek, S., Xu, X.D., Zhang, X., Liu, T.G., Waalwijk, C., Zhang, W., Kema, G.H.J., and Daboussi, M.J. (2008). Transposon-tagging identifies novel pathogenicity genes in *Fusarium graminearum*. Fungal. Genet. Biol. *45*, 1552-1561.

Flaherty, J.E., and Woloshuk, C.P. (2004). Regulation of fumonisin biosynthesis in *Fusarium verticillioides* by a zinc binuclear cluster-type gene, *ZFR1*. Appl. Environ. Microb. *70*, 2653-2659.

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

Imazaki, I., Kurahashi, M., Iida, Y., and Tsuge, T. (2007). Fow2, a Zn(II)2Cys6-type transcription regulator, controls plant infection of the vascular wilt fungus *Fusarium oxysporum*. Mol. Microbiol. *63*, 737-753.

Jansen, C., von Wettstein, D., Schafer, W., Kogel, K.H., Felk, A., and Maier, F.J. (2005). Infection patterns in barley and wheat spikes inoculated with wild-type and trichodiene synthase gene disrupted *Fusarium graminearum*. Proc. Natl. Acad. Sci. USA *102*, 16892-16897.

Jenczmionka, N.J., Maier, F.J., Losch, A.P., and 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.

Johnston, M. (1987). A model fungal gene regulatory mechanism: the GAL genes of *Saccharomyces cerevisiae*. Microbiol. Rev. *51*, 458-476.

Kim, J.E., Jin, J.M., Kim, H., Kim, J.C., Yun, S.H., and Lee, Y.W. (2006). GIP2, a putative transcription factor that regulates the aurofusarin biosynthetic gene cluster in *Gibberella zeae*. Appl. Environ. Microb. *72*, 1645-1652.

Leach, J., Lang, B.R., and Yoder, O.C. (1982). Methods for selection of mutants and in vitro culture of *Cochliobolus heterostrophus*. J. Gen. Microbiol. *128*, 1719-1729.

Lee, T., Oh, D.W., Kim, H.S., Lee, J., Kim, Y.H., Yun, S.H., and Lee, Y.W. (2001). Identification of deoxynivalenol- and nivalenol-producing chemotypes of *Gibberella zeae* by using PCR. Appl. Environ. Microb. *67*, 2966-2972.

Leslie J., Summerell B. (2006). The *Fusarium* Laboratory Manual. Blackwell Publishing, Ames, IA, U.S.A.

Makita, T., Katsuyama, Y., Tani, S., Suzuki, H., Kato, N., Todd, R.B., Hynes, M.J., Tsukagoshi, N., Kato, M., and Kobayashi, T. (2009). Inducer-dependent nuclear localization of a Zn(II)(2)Cys(6) transcriptional activator, AmyR, in *Aspergillus nidulans*. Biosci. Biotechnol. Biochem. *73*, 391-399.

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

Nicholas, K.B., Nicholas, H.B.J., and Deerfield, D.W.I. (1997). Gene-Doc: Analysis and visualization of genetic variation. EMBNEW. News *4*, 14.

Oide, S., Moeder, W., Krasnoff, S., Gibson, D., Haas, H., Yoshioka, K., and Turgeon, B.G. (2006). *NPS6*, encoding a nonribosomal peptide synthetase involved in siderophore-mediated iron metabolism, is a conserved virulence determinant of plant pathogenic ascomycetes. Plant Cell *18*, 2836-2853.

Pan, T., and Coleman, J.E. (1990). GAL4 transcription factor is not a "zinc finger" but forms a Zn(II)2Cys6 binuclear cluster. Proc. Natl. Acad. Sci. USA *87*, 2077-2081.

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

Ramamoorthy, V., Zhao, X., Snyder, A.K., Xu, J.R., and Shah, D.M. (2007). Two mitogen-activated protein kinase signalling cascades mediate basal resistance to antifungal plant defensins in *Fusarium graminearum*. Cell. Microbiol. *9*, 1491-1506.

Saghai-Maroof, M.A., Soliman, K.M., Jorgensen, R.A., and Allard, R.W. (1984). Ribosomal DNA spacer-length polymorphisms in barley: mendelian inheritance, chromosomal location, and population dynamics. Proc. Natl. Acad. Sci. USA *81*, 8014-8018.

Schmid, J., and Harold, F.M. (1988). Dual roles for calcium ions in apical growth of *Neurospora crassa*. J. Gen. Microbiol. *134*, 2623-2631.

Schultz, J., Milpetz, F., Bork, P., and Ponting, C.P. (1998). SMART, a simple modular architecture research tool: identification of signaling domains. Proc. Natl. Acad. Sci. USA *95*, 5857-5864.

Semighini, C.P., and Harris, S.D. (2008). Regulation of apical dominance in *Aspergillus nidulans* hyphae by reactive oxygen species. Genetics *179*, 1919-1932.

Shim, W.B., Sagaram, U.S., Choi, Y.E., So, J., Wilkinson, H.H., and Lee, Y.W. (2006). *FSR1* is

essential for virulence and female fertility in *Fusarium verticillioides* and *F. graminearum*. Mol. Plant Microbe. Interact. *19*, 725-733.

Shimizu, K., Hicks, J.K., Huang, T.P., and Keller, N.P. (2003). Pka, Ras and RGS protein interactions regulate activity of AflR, a Zn(II)2Cys6 transcription factor in *Aspergillus nidulans*. Genetics *165*, 1095-1104.

Silverman-Gavrila, L.B., and Lew, R.R. (2002). An IP_3-activated Ca^{2+} channel regulates fungal tip growth. J. cell sci. *115*, 5013-5025.

Silverman-Gavrila, L.B., and Lew, R.R. (2003). Calcium gradient dependence of *Neurospora crassa* hyphal growth. Microbiology *149*, 2475-2485.

Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. *22*, 4673-4680.

Turgeon, B.G., Garber, R.C., and Yoder, O.C. (1987). Development of a fungal transformation system based on selection of sequences with promoter activity. Mol. Cell Biol. *7*, 3297-3305.

Vienken, K., Scherer, M., and Fischer, R. (2005). The Zn(II)2Cys6 putative *Aspergillus nidulans* transcription factor repressor of sexual development inhibits sexual development under low-carbon conditions and in submersed culture. Genetics *169*, 619-630.

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

Wang, Y., Liu, W., Hou, Z., Wang, C., Zhou, X., Jonkers, W., Ding, S., Kistler, H.C., and Xu, J.R. (2011). A novel transcriptional factor important for pathogenesis and ascosporogenesis in *Fusarium graminearum*. Mol. Plant Microbe. Interact. *24*, 118-128.

Woloshuk, C.P., Foutz, K.R., Brewer, J.F., Bhatnagar, D., Cleveland, T.E., and Payne, G.A. (1994). Molecular characterization of aflR, a regulatory locus for aflatoxin biosynthesis. Appl. Environ. Microbiol. *60*, 2408-2414.

Yu, J.H., Butchko, R.A., Fernandes, M., Keller, N.P., Leonard, T.J., and Adams, T.H. (1996). Conservation of structure and function of the aflatoxin regulatory gene *aflR* from *Aspergillus nidulans* and *A. flavus*. Curr. Genet. *29*, 549-555.

Chapter 3

RNA-Seq analysis reveals new gene models and alternative splicing in the fungal pathogen *Fusarium graminearum*

Chunzhao Zhao,[1,2,3,4] Cees Waalwijk,[1,2] Pierre J. G. M. de Wit,[2,5] Dingzhong Tang[3] and Theo van der Lee[1,2]

1. Plant Research International, P.O. Box 6708 PB, Wageningen, The Netherlands
2. Graduate School Experimental Plant Sciences, Wageningen, The Netherlands
3. State Key Laboratory of Plant Cell and Chromosome Engineering, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing 100101, China
4. Graduate University of Chinese Academy of Sciences, Beijing 100049, China
5. Wageningen University, Laboratory of Phytopathology, P.O. Box 6708 PB, Wageningen, The Netherlands

Published in BMC Genomics 14:21 (2013)

Abstract

The genome of *Fusarium graminearum* has been sequenced and annotated previously, but correct gene annotation remains a challenge. In addition, posttranscriptional regulations, such as alternative splicing and RNA editing, are poorly understood in *F. graminearum*. Here we took advantage of RNA-Seq to improve gene annotations and to identify alternative splicing and RNA editing in *F. graminearum*. We identified and revised 655 incorrectly predicted gene models (5% of all predicted genes), including revisions of intron predictions, intron splice sites and prediction of novel introns. 231 genes were identified with two or more alternative splice variants, mostly due to intron retention. Interestingly, the expression ratios between different transcript isoforms appeared to be developmentally regulated. Surprisingly, no RNA editing was identified in *F. graminearum*. Moreover, 2459 novel transcriptionally active regions (nTARs) were identified and our analysis indicates that many of these could be missed genes. Finally, we identified the 5' UTR and/or 3' UTR sequences of 7666 genes. A number of representative novel gene models and alternatively spliced genes were validated by reverse transcription polymerase chain reaction and sequencing of the generated amplicons.

Introduction

Fusarium graminearum is an ascomycete that can cause diseases in a variety of agronomically important crops, including Fusarium Head Blight (FHB) on wheat, barley and oat, and stalk rot on corn (Bluhm et al., 2007; Kazan et al., 2011). Infection by *F. graminearum* not only causes severe yield losses but also contaminates seeds with mycotoxins, such as deoxynivalenol (DON) and nivalenol (NIV) (Lee et al., 2001; Lysoe et al., 2006), which are very harmful to humans and animals (Desjardins et al., 1993; Proctor et al., 1995). The infection mechanism of wheat and other cereals by *F. graminearum* is still poorly understood, but genome and transcriptome research will enable us to identify genes that are required for pathogenicity and improve our understanding of infection mechanism of *F. graminearum* on its host plants. The genome of *F. graminearum* has been sequenced and currently two different annotations of the same genome assembly are available. One was generated by the Broad Institute (Cuomo et al., 2007), and a second one by MIPS (Guldener et al., 2006; Wong et al., 2011).

The correctness of predicted gene models is extremely important for further comparative and functional genome studies. Gene model predictions performed at the Broad Institute were mainly generated by machine annotation based on a combination of the Calhoun annotation system and the FGENESH program (Cuomo et al., 2007). The MIPS *F. graminearum* database was constructed based on Broad gene calls by integrating several sources and programs, including (i) integration of different gene prediction programs, (ii) comparison of current *F. graminearum* gene models with related *Fusarium* species (*F. oxysporum, F. verticillioides* and *F. solani*) and other *Ascomycetes* including *Neurospora crassa*, and (iii) inclusion of expression sequence tag (EST) data (Wong et al., 2011). Compared to the Broad gene set, 1770 gene models were revised and 691 new gene calls were added to MIPS gene set [9]. Although many gene models have been improved by these different approaches, most of them lack experimental support and for species-specific and non-conserved genes the gene model predictions are often incorrect or partially incorrect. In addition, it is difficult to identify novel genes and delineate untranslated regions (UTRs) using traditional bioinformatics tools. To further improve gene model predictions, large-scale transcript information is required.

Genome sequencing and annotation have provided a global view of the genes present in *F. graminearum*, but little is known about their transcriptional and post-transcriptional regulation. In

Homo sapiens, *Mus musculus*, *Drosophila melanogaster* and *Arabidopsis thaliana*, alternative splicing has been reported to occur in many genes, which enables these organisms to enlarge their proteome diversity by increasing transcript variations in their genome (Filichkin et al., 2009; Pan et al., 2008; Reddy, 2007; Sultan et al., 2008; Venables et al., 2011; Yeo et al., 2005). A striking example of alternative splicing is the *Dscam* gene of *D. melanogaster*, which potentially generates more than 38,000 different transcripts (Schmucker et al., 2000). In mammals, alternative splicing plays an important role in developmental processes, such as stem cell self-renewal and differentiation (Mayshar et al., 2008; Salomonis et al., 2010; Yeo et al., 2007), development of heart and brain (Gehman et al., 2011; Kalsotra et al., 2008; Xu et al., 2005), and in the response to extracellular stimuli, such as immune cell activation and neuronal depolarization (Heyd and Lynch, 2011; Li et al., 2007). In *A. thaliana*, alternative splicing has been shown to play an important role in its development (Ali et al., 2007) and in the response to environmental cues, such as high light treatment, cold and heat treatment (Filichkin et al., 2009). Alternative splicing has also been reported in fungi, including *Cryptococcus neoformans*, *Ustilago maydis*, *Magnaporthe grisea*, *Aspergillus nidulans*, and *F. verticillioides* (Brown et al., 2008; Galagan et al., 2005; Loftus et al., 2005; McGuire et al., 2008). However, so far, alternative splicing has not been reported to occur in *F. graminearum*.

Recently, next-generation sequencing technology (RNA-Seq) has become available as a powerful tool to investigate the transcriptional profiles in many organisms, such as *H. sapiens*, *Saccharomyces cerevisiae*, *A. thaliana*, *Candida albicans* and *C. parapsilosis* (Bruno et al., 2010; Filichkin et al., 2009; Guida et al., 2011; Mortazavi et al., 2008; Nagalakshmi et al., 2008). It has been demonstrated that RNA-Seq data can be efficiently used to improve gene model prediction and to identify novel transcripts (Gan et al., 2011; Li et al., 2011; Wang et al., 2009). In addition, RNA-Seq technology is much more sensitive and efficient than previously used dedicated microarrays to compare gene expression profiles (Sultan et al., 2008). RNA-Seq data also have been successfully used to identify alternative splicing in genes of different species (Filichkin et al., 2009; Pan et al., 2008; Ramani et al., 2010). Moreover, RNA-Seq technology has recently been used to identify RNA editing in *H. sapiens* (Peng et al., 2012).

Previously, we have identified and phenotypically characterized knock out mutant *ebr1* (*Enhanced branch 1*) that shows reduced radial growth and reduced pathogenicity (Zhao et al.,

2011). EBR1 encodes a Gal4-like Zn_2Cys_6 transcription factor that is localized in the nucleus during vegetative growth. In order to further unravel the regulatory role of EBR1 in radial growth, we have performed RNA-Seq on wild-type isolate PH-1 (PH-1) and mutant *ebr1* (*ebr1*) to identify differentially expressed genes. In this study, we focused on the use of RNA-Seq data from both PH-1 and *ebr1* to improve gene model predictions, identify novel genes, and search for alternative spicing and RNA editing in *F. graminearum*. The obtained results were validated using RT-PCR and sequencing of the generated products. These analyses have greatly improved numerous gene models and provided a comprehensive insight of RNA splicing in *F. graminearum*.

Results

Quality analysis of the RNA-Seq data from *F. graminearum*

To perform RNA-Seq analysis, RNA was isolated from mycelia of PH-1 and *ebr1* grown in liquid CM medium for 30 h. The isolated RNA was prepared to be sequenced by next generation sequencing technology (Illumina). Of each isolate, two technical replicates were analyzed. In total 12,791,946 reads (90 nucleotides for each read) from PH-1 and 12,928,704 reads from *ebr1* were obtained. Using the genome annotation in Broad database, transcripts were detected for 76.9% of the 13321 predicted genes in PH-1 and for 81.2% of the genes in *ebr1*. By combing the RNA-Seq data from both PH-1 and *ebr1* we detected the expression of over 84% of the predicted genes. In addition, the combined RNA-Seq data showed that 74.8% of the reads matched to exonic regions, 10.6% to untranslated regions (UTRs), 12.9% to intergenic regions and only 1.7% to intronic regions (Figure 1A). Among all matched reads, 84% matched to unique locations, and 16% to multiple locations in the genome, of which 2% matched to between 2 to 10 different locations, and 14% to more than 10 different locations (Figure 1B). Of all reads that matched to multiple locations, 69.5% matched to intergenic regions, 30.4% to UTRs, and only 0.1% to coding regions (Figure 1C).

To evaluate the quality of RNA-Seq data, several quality control analyses were performed. Firstly, the total coverage of reads from 5' to 3' end of genes was examined. For both PH-1 and *ebr1* RNA-Seq data the reads were evenly distributed over genes with exception of the very 5' and 3' ends (Figure 1D). In addition, for 54% of the genes in PH-1 and 60% of the genes in *ebr1* the read coverage was more than 90% (Supplemental Figure 1). Finally, comparison of the two

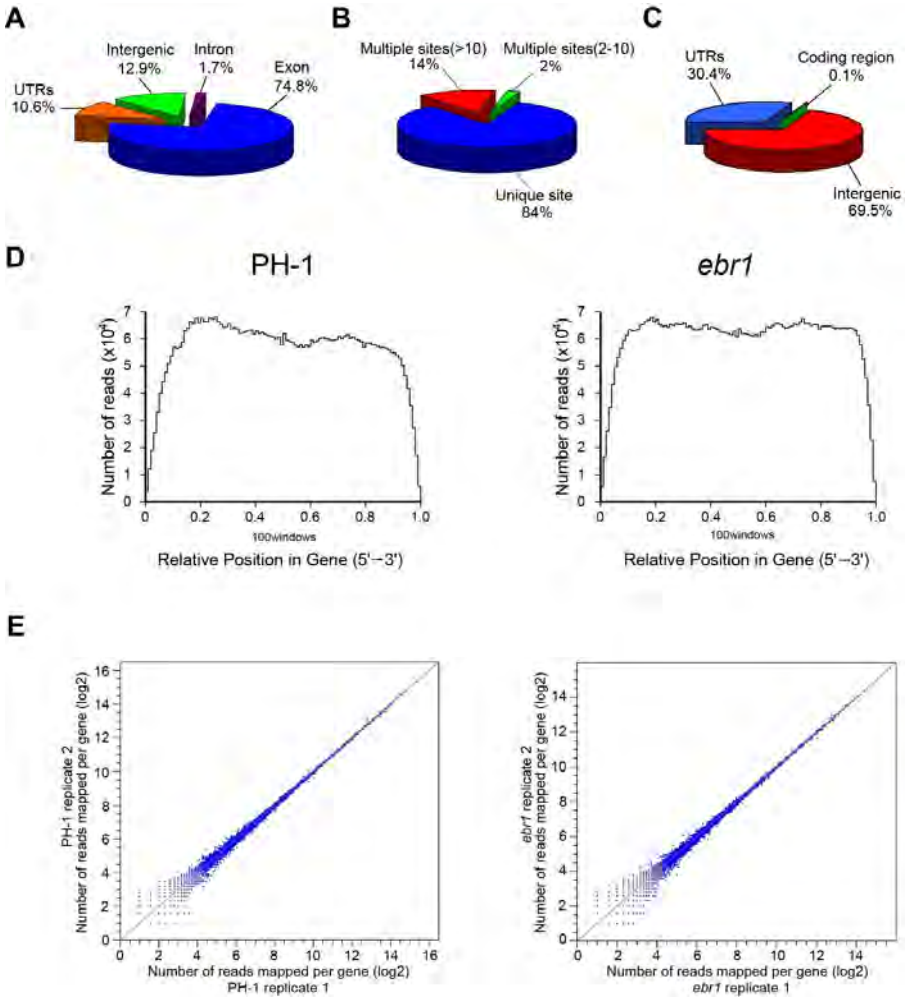


Figure 1. Statistical and quality control analysis of RNA-Seq data. A. Reads distribution over exons, introns, untranslated regions (UTRs) and intergenic regions. B. 84% of all reads matched to unique locations and 16% of the reads to multiple locations in the genome of *F. graminearum*. Of the reads matching to multiple locations 2% matched to 2-10 sites and 14% to more than 10 different locations. C. The distribution of all reads matching to multiple locations in intergenic regions, UTRs and coding regions. D. The total read coverage along the gene body from 5' to 3' end in wt PH-1 and mutant *ebr1*. The genes were divided into 100 equal windows. E. Scatter plot analysis of two technical replicates from both wt PH-1 and mutant *ebr1*. Log2 transformed reads number of all predicted genes was used for comparison.

technical replicates of both PH-1 and *ebr1* clearly showed that the RNA-Seq data are highly reproducible (Figure 1E).

Matching of the reads to the *F. graminearum* gene database showed that the 98.1% of the

matched reads supported the gene models present in the Broad *F. graminearum* database, implying that reads matched to exonic, but not to intronic regions of genes. For example, all 168 reads from PH-1 and all 216 reads from *ebr1* uniquely matched to the exonic regions of gene FGSG_04412, and no reads were found matching to the intronic regions (Figure 2A). Comparing the reads matching to gene FGSG_04412 from PH-1 and *ebr1* showed that the distribution pattern of reads along the gene was very similar (exemplified in Figure 2A). Similar distribution patterns of reads in PH-1 and *ebr1* were also found in other genes. To evaluate the background of non-specific transcripts, we compared the expression of *EBR1* in PH-1 and *ebr1*. In PH-1, 75 reads were found that matched to the coding region of *EBR1*, whereas no reads were found matching to *EBR1* in *ebr1* (Figure 2B). Altogether these results clearly show that the obtained RNA-Seq data are of high quality and form a firm basis for further analysis.

Strategies to identify incorrect gene models and alternative splicing

We combined the RNA-Seq data of PH-1 and *ebr1* and employed three different strategies to identify incorrect gene models and alternative splicing in *F. graminearum* (Supplemental Figure 2A). The first strategy was to identify reads that matched to intronic regions. Reads matching intronic regions originate from either incorrectly annotated or alternatively spliced genes. The second strategy was aimed at predicting transcripts with non-matched or mismatched regions. Of highly expressed genes, the transcripts should be well covered by reads. However, of some transcripts regions not matched by reads or not perfectly matched by reads were identified which points to novel introns or incorrectly predicted introns in these genes. Two examples of this type of transcripts are shown in Supplemental Figure 3. In total, 436 possibly incorrectly annotated or alternatively spliced genes were identified by the first strategy and 343 by the second. To further refine incorrect gene models and identify genes with alternative splicing, the TopHat program was applied. This program identifies intron splice sites and has been widely applied to identify incorrect gene annotations and alternative splicing (Trapnell et al., 2012). By applying this program, we obtained 228 putatively new genes. Comparing all these three strategies, we identified 287 genes that were exclusively identified by the first strategy, 243 genes by the second, and 153 genes by the TopHat program. Only 6 genes were identified by all three strategies (Supplemental Figure 2B).

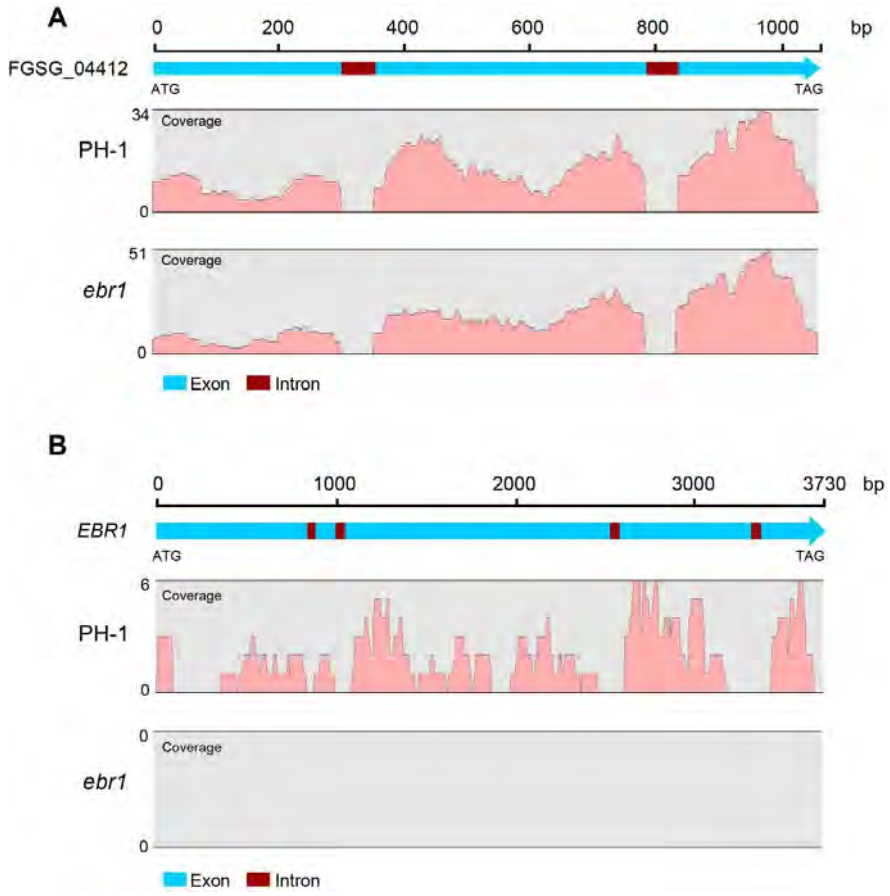


Figure 2. Visualization of RNA-Seq data in the CLC genomic workbench software. A. Reads from the RNA-Seq data of wt PH-1 and mutant *ebr1* were mapped to the gene FGSG_04412. 168 reads in PH-1 and 216 reads in *ebr1* uniquely matched to the exon of FGSG_04412. In contrast, no reads matched to the intronic regions. Y-axis represents the coverage number of reads of each nucleotide. B. Analysis of the reads matching to the *EBR1* gene in wt PH-1 and mutant *ebr1*, respectively. 75 reads from wt PH-1 RNA-Seq data matched to *EBR1*, whereas no reads from mutant *ebr1* RNA-Seq data matched to *EBR1*.

Using these three strategies, 842 genes with possibly incorrect gene models or alternative splicing were identified when compared with the Broad *F. graminearum* annotation. We further examined these genes in the MIPS *F. graminearum* database and found that 278 of the identified genes had already been revised. Subsequently, we manually examined the remaining 564 genes in the CLC software package and classified them into two distinct groups: incorrect gene models and alternatively spliced genes. To distinguish between these two options, we carefully examined reads for the presence of splice sites. Genes that matched reads showing both reference splice site

and additional splice site were considered to be the result of alternative splicing; genes that matched reads only showing additional splice site but not reference splice site were grouped into incorrect gene models.

Identification of incorrect gene models

377 genes that were incorrectly annotated in the Broad *F. graminearum* database and had not yet been revised in the MIPS *F. graminearum* database were further analyzed. They were divided into four classes: (i) gene models with incorrectly predicted introns, (ii) gene models with incorrect intron splice sites, (iii) gene models with novel introns and (iv) gene models with other incorrect annotations (Figure 3A). In total 119 genes with incorrect intron predictions were identified. For example, according to the annotation in the Broad *F. graminearum* database, there are two introns in gene FGSG_01636. However, several RNA-Seq reads that matched to the second intron indicated that the proposed second intron does not exist (Figure 3B). For confirmation of this result, primers flanking the supposed second intron were designed and RT-PCR was performed. Genomic DNA isolated from PH-1 and two cDNAs isolated from PH-1 and *ebr1* were used as templates, respectively. RT-PCR confirmed that the proposed second intron in this gene is absent (Figure 3C). Open reading frame inspection showed that the newly proposed gene model is translated into a functional protein without a premature stop. Furthermore, two additional randomly chosen genes, FGSG_04300 and FGSG_08487, were inspected using the same strategy as described for FGSG_01636, and both lacked the predicted introns in the amplified fragments. In contrast, gene FGSG_10264 that was selected as positive control confirmed the predicted intron (Figure 3C).

In addition to incorrect intron predictions, we identified novel introns in 40 genes with an average length of around 50 bp. Supplemental Figure 4A shows an example of a novel intron identified in gene FGSG_06363. To validate the presence of novel introns, flanking primers for five randomly selected introns were designed and for all of them the presence of the introns was confirmed by RT-PCR (Supplemental Figure 4B).

In 164 genes incorrectly predicted splice sites were identified, including incorrect donor and acceptor sites or both; they were manually revised according to our RNA-Seq data. Supplemental Figure 5A shows an example of an incorrectly predicted splice site. Three genes with incorrectly

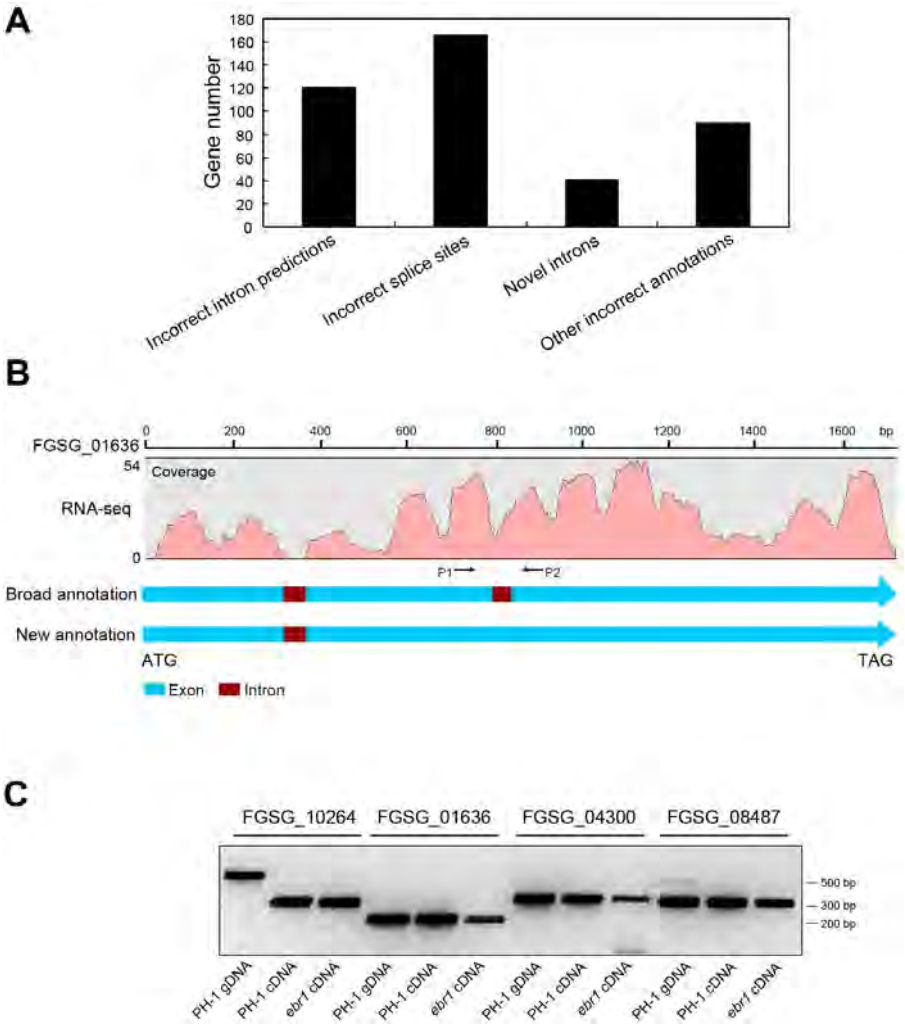


Figure 3. Identification of incorrect gene models in the *F. graminearum* database. A. Statistical analysis of the incorrect gene models identified from RNA-Seq data, including incorrect predictions of introns, incorrect intron splice sites, novel introns and other incorrect annotations. B. One example of a gene with incorrect intron predictions is shown. Two introns were annotated in gene FGSG_01636 in the Broad *F. graminearum* database, but RNA-Seq data clearly showed that the second intron is absent. C. Three genes with incorrect intron predictions were selected for confirmation by RT-PCR. For each gene, one genomic DNA template, two cDNA templates (one is from wt PH-1, one is from mutant *ebr1*) were applied to perform RT-PCR. Primers were designed flanking the intronic regions, as shown in B. Bands with identical size were amplified from both genomic DNA and cDNA template for all three genes, indicating that predicted introns do not exist. The correctly annotated gene FGSG_10264 was used as a control.

predicted splice sites were randomly selected and were all confirmed by RT-PCR (Supplemental Figure 5B). In addition, 88 genes were identified with incorrect gene models of which the correct

splice sites could not be assigned yet due to low read coverage or other reasons.

Gene expression analysis showed that for 15% of the predicted genes transcripts were absent in the RNA-Seq data. To determine whether these genes result from incorrect gene calls in databases or were not expressed under the condition tested, we performed a homology search of the predicted proteins using blastP against the NCBI database. As orthologous genes could be identified for 86.5% of these genes (E-value<1E-10), we conclude that these genes are correctly annotated but not or very lowly expressed in liquid CM medium.

Identification of alternatively spliced genes in *F. graminearum*

From our RNA-Seq data, 231 genes were identified with alternative splicing, including exon skipping, intron retention, or alternative 5' and 3' splice sites (Figure 4A). Most of the alternatively spliced genes involved intron retentions, of which one example is shown in Figure 4B. In gene FGSG_05122, there are four reads that confirm reference intron splice sites, whereas several reads uniquely matched to a presumed intronic region. To confirm retention of these introns, RT-PCR was performed for three randomly selected genes, and in all cases the predicted intron retentions were confirmed (Figure 4C). Open reading frame analysis of genes with intron retention showed that most of them lead to premature termination of transcription.

In addition to intron retention, we identified 28 genes with alternative 5' or 3' splice sites or both. Similar to genes with intron retention, most genes with alternative 5'and 3' splice sites led to premature termination of transcription. Among all these alternatively spliced introns, we identified three introns with two alternative 3' splice sites and three with two or three alternative 5' splice sites. For instance, gene FGSG_06760 that encodes a HMG-box protein with a coiled coil and HMG domain contains two alternative 3' splice sites according to RNA-Seq data (Figure 5A). The alternative 3' splice sites were confirmed by RT-PCR (Figure 5B). Compared to the predicted transcript based on the gene model present in the Broad database, the two alternative 3' splice sites do not lead to premature termination of the transcript. However, the two alternative transcripts lead to proteins lacking 16 amino and 17 amino acids, respectively, that are located between the coiled coil and HMG domain (Figure 5C). Another example is gene FGSG_06124 for which there are four different 5' splice sites in its second intron (Supplemental Figure 6A) that were all confirmed by RT-PCR (Supplemental Figure 6B). All alternative transcripts are not prematurely

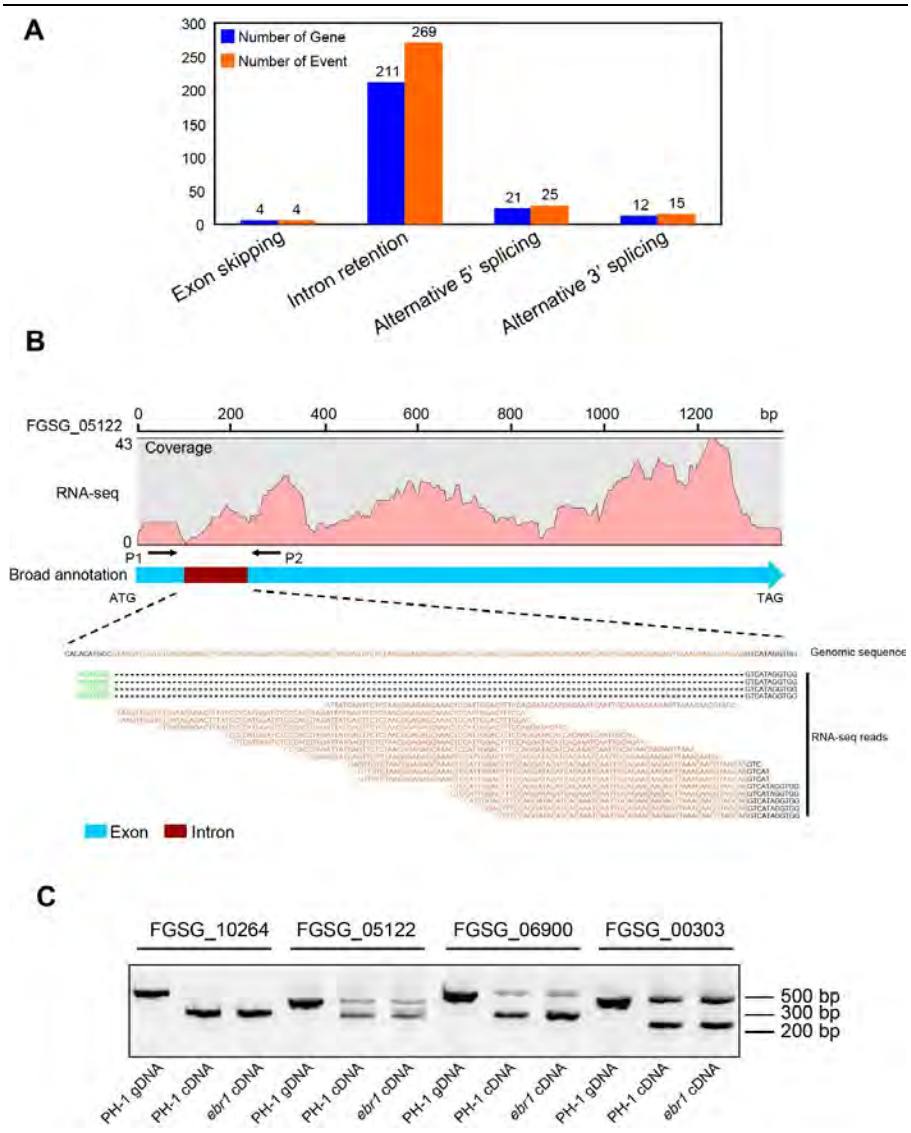


Figure 4. Identification of alternative splicing. A. Statistical analysis of alternative splicing identified from RNA-Seq data. Four different types of alternative splicing were identified. Blue bars represent the number of genes that are alternatively spliced. Orange bars represent the total number of alternative splicing events. B. FGSG_05122 is an example of a gene showing intron retention. There are four reads representing the intron splice sites (green letters). Meanwhile, several reads uniquely match to the intronic region. Black letters represent exonic region; orange letters represent intronic region. C. Three genes with intron retention were selected for confirmation by RT-PCT. Primers were designed flanking the intronic regions. Gene FGSG_10264 was included as a control.

terminated (Supplemental Figure 6C). FGSG_06124 encodes a hypothetical protein, with a putative transmembrane and prolipoprotein diacylglyceryl transferase domain. Both domains are

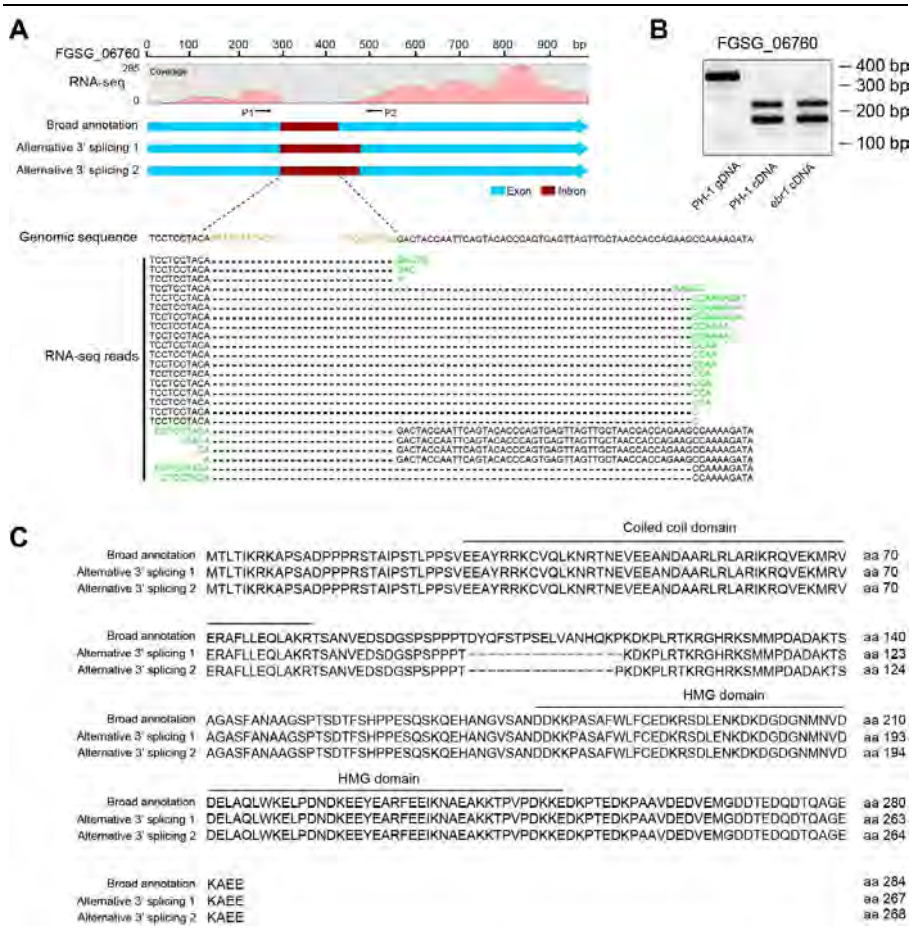


Figure 5. Example of alternative 3' splicing. A. Three different 3' intron splice sites were identified in the intron of gene FGSG_06760. There are seven reads showing the reference 3' intron splice site, 16 reads showing alternative 3' intron splice site 1, and one read showing alternative 3' intron splice site 2. Black letters represent the exonic region; orange letters represent the intronic region; green letters represent different 3' intron splice sites. B. RT-PCR confirmed the alternative 3' splicing events in the intron of FGSG_06760. C. Protein alignment shows that 16 or 17 amino acids located between coiled coil and HMG domain are lacking in the proteins encoded by the two alternatively spliced transcripts.

present in all four predicted proteins.

Finally, we identified four cases of exon skipping. FGSG_00786 is an example of a gene with alternative exon skipping (Supplemental Figure 7A) that encodes a serine/threonine-protein kinase srk1 with an S_TKc domain between amino acid (aa) residues 101 and 405 (Supplemental Figure 7B). The third exon in FGSG_00786 is sometimes lacking in transcripts as was confirmed by RT-PCR (Supplemental Figure 7C), leading to the loss of 17 aa residues in the S_TKc domain.

From above, six genes with alternative splicing were confirmed in both PH-1 and *ebr1* by RT-PCR. We further analyzed all remaining alternatively spliced genes by using RNA-Seq data from PH-1 and *ebr1*, respectively, in the CLC software package. Nearly all of the alternative splicing events can be identified in both PH-1 and *ebr1*. This indicates that disruption of *EBR1* in *F. graminearum* does not affect alternative splicing. To further understand possible roles of all alternatively spliced genes, we functionally categorized them by using the MIPS FunCatDB database. The alternatively spliced genes did not belong to one specific functional class of genes, but were classified in many different categories, of which “proteins with binding function or cofactor requirement” (P value=1.91E-06) and “Protein synthesis” (P value=2.61E-04) prevailed.

Alternative splicing is developmentally regulated

To test whether alternative splicing in *F. graminearum* is developmentally regulated, we performed RT-PCR on four alternatively spliced genes (FGSG_00303, FGSG_06760, FGSG_05122 and FGSG_04141) on PH-1 RNA samples isolated at five different time points (0 h, 2 h, 8 h, 24 h and 36 h after incubation of conidia in liquid CM medium) (Figure 6). In this medium, the macroconidia of *F. graminearum* swell within 2 hours, germinate after 3 hours and the hyphae elongate and develop into mycelium after 8 hours (Seong et al., 2008). For gene FGSG_00303 (encoding a transcriptional elongation regulator), the expression ratio between the two transcript isoforms is independent of the growth stage, whereas this ratio (reference isoform/alternative isoform) increased for gene FGSG_06760 (encoding a HMG box protein) after 8 hours. The expression of genes FGSG_05122 (encoding a FAD dependent oxidoreductase) and FGSG_04141 (encoding a DNA repair protein) show strong developmental regulation and the expression ratio between the two transcript isoforms changes strongly at different vegetative growth stages. The RT-PCRs for these genes were biologically repeated with similar result. These data suggest that for some of the alternatively spliced genes the expression levels and the ratio between transcripts change during vegetative growth.

Non-canonical splice sites

Manually inspection of the 842 genes with putative incorrect gene models or alternative splicing resulted in the identification of 28 genes with non-canonical splice sites. In total the 842

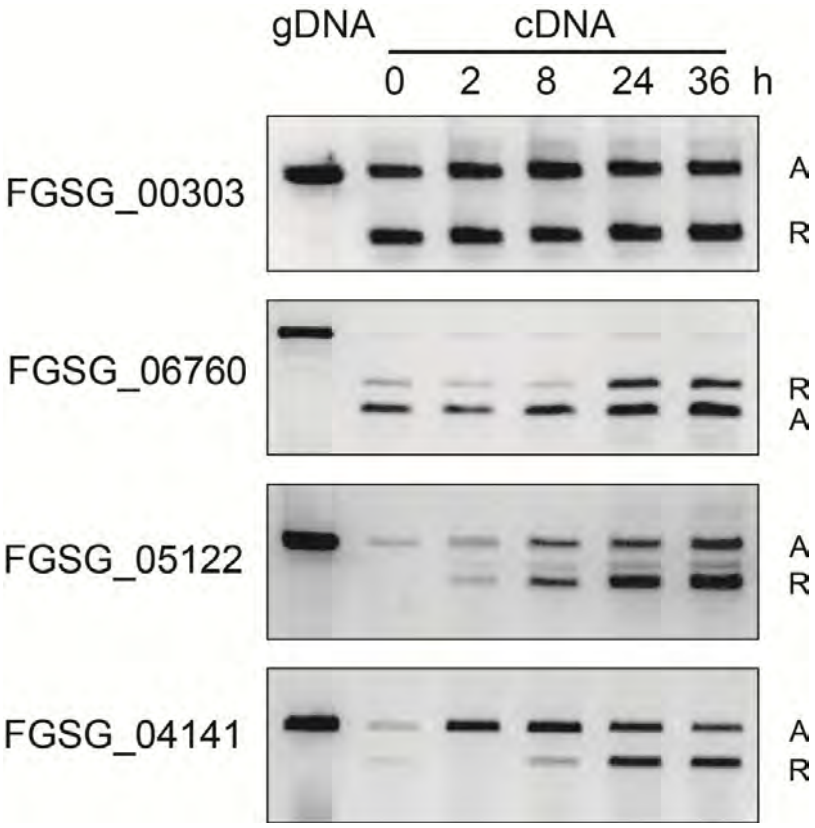


Figure 6. Alternative splicing is developmentally regulated. RNA samples were isolated from wt PH-1 isolate at five different time points (0 h, 2 h, 8 h, 24 h, 36 h after incubation of conidia in liquid CM medium-containing shake cultures). RT-PCRs were performed at the different time points for four alternatively spliced genes (FGSG_00303, FGSG_06760, FGSG_05122 and FGSG_04141). Genomic DNA template was used as control. Capital letter R represents reference transcript isoform. Capital letter A represents alternatively spliced transcript isoform. The experiment was independently performed twice with similar results.

genes contain 2604 introns, of which 98.92% carry canonical GT-AG donor-acceptor sites, 0.77% introns carry GC-AG donor-acceptor sites, and the remaining 0.31% carry other non-canonical donor-acceptor sites, including GT-GG, AC-AC, GG-TA, TA-AG, AT-AC and GA-AG. Seven genes with non-canonical splice sites are shown in Figure 7A and the correctness of all these splice sites was confirmed by Sanger sequencing showing that they were not caused by sequencing errors.

The 20 introns with GC-AG splice sites were analyzed for the presence of conserved flanking nucleotides by using motif comparison tool (Bailey et al., 2009). AG nucleotides

predominantly flank the GC donor site, whereas in the intronic region, AAGT occurs more frequently. The nucleotides flanking the AG acceptor site are less conserved. However, a C or T prevails in the intronic region flanking AG (Figure 7B).

Identification of novel transcriptionally active and untranslated regions

By mapping RNA-Seq reads against Broad *F. graminearum* database, 12.9% of the reads matched to intergenic regions, from which 2459 novel transcriptionally active regions (nTARs) were obtained. To determine whether these nTARs encode proteins, they were blasted against the MIPS *F. graminearum* and Broad *Fusarium* databases. Of these 2459 nTARs, 355 had already been predicted as novel genes in the MIPS *F. graminearum* database, 118 of which show orthologs in either the *F. oxysporum*, *F. verticillioides* or both. In addition, we identified 74 nTARs that had not yet been annotated in the MIPS *F. graminearum* database but are putatively derived from genes as orthologs were identified in either *F. oxysporum*, *F. verticillioides* or both. In addition, we found 123 nTARs (5%) have introns, indicating that they could be real genes. Supplemental Figure A shows an example of an nTAR, TU358, which contains three introns. To confirm that the identified nTARs are real, five were selected and confirmed by RT-PCR (Supplemental Figure 8B).

The RNA-Seq data also allowed identification of the boundaries of 5' and 3' UTRs of genes. For 5951 genes 5' UTRs and for 6405 genes 3' UTRs were identified. Comparing UTRs identified by RNA-Seq analysis with those present in the annotated genome in the Broad *F. graminearum* database showed some genes with incorrectly predicted UTRs. One example is shown in Supplemental Figure 9B where the 3' UTR prediction in gene FGSG_01403 is different from that predicted by RNA-Seq analysis.

Screening for RNA editing in *F. graminearum*

In total 695 single nucleotide polymorphisms (SNPs) were identified when comparing RNA-Seq data with the genome sequences by using the CLC software package. All SNPs were manually examined and a large number was identified in stretches of multiple cytosine residues. In addition, many SNPs were identified near intron splice sites and appeared to be caused by misalignment of cDNA to the genomic DNA sequence. Twelve representative SNPs were selected

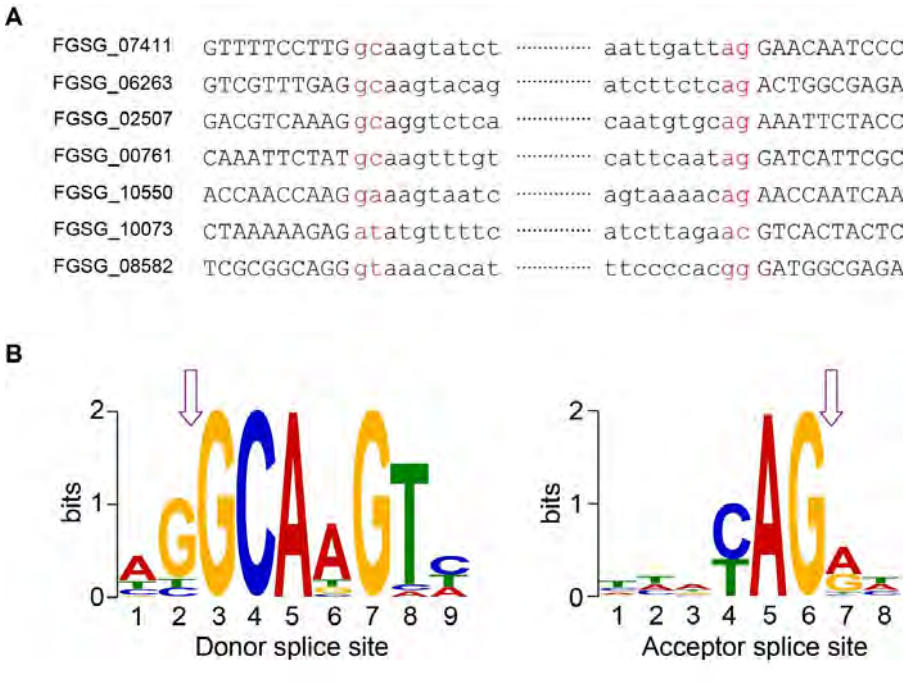


Figure 7. Non-canonical splice sites identified in *F. graminearum*. A. Genes with non-canonical splice sites are shown. Capital letters represent exonic sequences, small letters represent intronic sequences. B. Analysis of nucleotides preference flanking GC donor splice site and AG acceptor splice site by using the MEME program (Bailey et al., 2009). Y-axis indicates sequence conservation at each position. The height of symbols in each position represents the relative frequency of each nucleic acid. Arrows indicate splice sites.

for confirmation by Sanger sequencing of the PCR amplicons obtained from both genomic DNA and cDNA. In four cases the SNPs were not real and due to sequencing errors present in the genomic DNA sequence of PH-1. For the remaining eight SNPs, no differences were observed between cDNA and genomic sequences after re-sequencing suggesting that in the latter cases discrepancies between the RNA-Seq data and the genome sequence could be explained by sequencing errors in the initial RNA-Seq data set. These results suggest that no RNA editing occurs in *F. graminearum* according to our analysis.

Discussion

In this study we analyzed the transcriptome of *F. graminearum* grown in liquid CM medium by Illumina sequencing to investigate the correctness of predicted gene models present in the

annotated Broad *F. graminearum* genome database and to identify the occurrence of alternative splicing, RNA editing, non-canonical splice sites, novel transcripts and the sequences of the 5' UTR and 3' UTR regions. The total coverage of reads along the genes was evenly distributed except for the ultimate 5' and 3' ends indicating that overall our RNA-Seq data are of high quality (Wang et al., 2009). Although overall the read coverage was evenly distributed over genes, for individual genes the coverage was not evenly distributed; this phenomenon has also been reported in other RNA-Seq studies (Guida et al., 2011; Nagalakshmi et al., 2008). Interestingly, for nearly all genes the read coverage pattern between wt PH-1 and mutant *ebr1* is very similar. This suggests that each gene has a characteristic RNA-Seq profile, which could be related to secondary structure of particular domains of RNA molecules that interfere with RNA shearing and subsequent sequencing.

The background reads in RNA-Seq data sets have been reported to be low. For example in RNA-Seq data obtained from yeast, no reads matching a 3.5-kb deleted region were obtained, and very few reads matching to nontranscribed centromeres were identified (Nagalakshmi et al., 2008). The similar result was found in our RNA-Seq analysis; a comparison of the *EBR1* expression level between PH-1 and *ebr1* showed hardly any transcription of the *EBR1* gene in the *ebr1* deletion mutant. In addition, we found that RNA-Seq data analysis for both PH-1 and *ebr1* RNA-Seq is very reproducible.

Analysis of the read distribution suggests that 12.9% of reads matched to intergenic regions, which is relatively high in comparison to 3% and 5% found in *H. sapiens* and *A. thaliana,* respectively (Filichkin et al., 2009; Mortazavi et al., 2008). This high percentage may at least partly reflect the lower quality of gene model prediction in *F. graminearum* compared to *H. sapiens* and *A. thaliana*. In the latter two genomes, several rounds of gene annotation have been performed and more experimental evidence has been provided to support the gene models. In addition, 16% of the reads could not be matched to a single location in the genome, a finding that was also reported in other species (Filichkin et al., 2009; Mortazavi et al., 2008). For instance, in *H. sapiens*, 20% of the intergenic reads match to multiple locations in the genome, of which 6% match to 2-10 locations and 14% to more than 10 locations (Mortazavi et al., 2008). Furthermore, we identified that most of the reads mapping to multiple locations originate from intergenic regions and UTRs, whereas only very few reads matched to coding regions, which suggests that

the reads matching to each transcript are very specific and the read coverage of each transcript is a reliable reflection of the gene expression level.

RNA-Seq has been widely used to identify incorrect gene models and alternative splicing in different organisms (Bruno et al., 2010; Filichkin et al., 2009; Guida et al., 2011; Sultan et al., 2008). However, to distinguish incorrect gene models from alternative splicing is a challenging and laborious task. In this study, all selected genes were manually examined in the CLC software package to identify reads showing splice sites. RT-PCR analysis on the selected genes confirmed that identification of incorrectly annotated gene models and alternative splicing appears reliable. In total 655 genes were identified with incorrect gene models in the Broad *F. graminearum* database. Excluding genes with no detectable expression or with low read coverage (less than 50 reads), the fraction of incorrect gene models in the published annotation of the Broad *F. graminearum* database is 10.3%. Gene model predictions in the MIPS *F. graminearum* database were considered to be of higher quality than those in the Broad *F. graminearum* database (Guldener et al., 2006), which was confirmed by our RNA-Seq analysis. Nonetheless we could still improve many gene models predicted in the MIPS *F. graminearum* database. Even some of the manually revised gene models in the MIPS *F. graminearum* database appeared to be incorrect, indicating that gene annotations in *F. graminearum* database still need to be improved and RNA-Seq analysis can significantly improve the published gene models. In this study, RNA-Seq data were generated from mycelia growing in nutrient-rich medium. To investigate whether the incorrectly annotated genes are caused by alternative splicing, we also analyzed the available EST data generated from other conditions, such as carbon- and nitrogen- starved media and cultures of maturing perithecia (Pan et al., 2008; Wang et al., 2008). These EST data support our discoveries that genes are incorrectly annotated, but six genes were identified that have two different transcripts, indicating that they are alternatively spliced. This indicates that few of genes classified as incorrectly annotated genes in this study could in fact be alternative spliced genes. In future, high-throughput transcript data of *F. graminearum* generated from different growth conditions and different development stage could be required to identify all possible intron splice sites in these incorrectly annotated genes.

Alternative splicing has been investigated in many organisms including *H. sapiens*, *Caenorhabditis elegans*, *A. thaliana* and *C. neoformans* (Filichkin et al., 2009; Ramani et al., 2010;

Wang et al., 2008). In *H. sapiens*, 95% of the genes undergo alternative splicing (Pan et al., 2008; Wang et al., 2008); in *A. thaliana*, alternative splicing is estimated to be 42% (Filichkin et al., 2009). In fungi much lower percentages of alternative splicing have been predicted, including 4.3% (277 genes) in *C. neoformans*, 1.3 % (162 genes) in *A. flavus* and 1.4% (151 genes) in *M. grisea* (Loftus et al., 2005; McGuire et al., 2008). We found alternative splicing in 231 genes (1.7%) in *F. graminearum*, but it should be noted that we have only analyzed expression in one growth condition and as fungi can adapt to many different environmental conditions we expect that this percentage will increase when transcription profiles under more different growth conditions are analyzed. At least 4 different types of alternative splicing exist in *F. graminearum*, of which intron retention appeared most prevalent, which is also the case in *A. thaliana* (Filichkin et al., 2009; Ner-Gaon et al., 2004; Wang and Brendel, 2006), whereas in *H. sapiens*, exon skipping is most prevalent (Sultan et al., 2008).

In-frame analysis showed that the majority of the alternatively spliced transcripts identified in *F. graminearum* cause premature termination codons (PTCs), of which most are located in intronic regions. Also in *H. sapiens* and *A. thaliana*, a high percentage of alternatively spliced transcripts contain PTCs (Filichkin et al., 2009; Saltzman et al., 2008). In *A. thaliana*, 77.9% of the alternatively spliced genes introduce PTCs and most of them are considered as potential targets of the nonsense mediated mRNA decay (NMD) (Filichkin et al., 2009; Wang and Brendel, 2006). NMD was initially identified in *S. cerevisiae* and later widely studied in higher eukaryotes (Cui et al., 1995; Lee and Culbertson, 1995; Leeds et al., 1992), but so far, only a few studies on NMD are reported in filamentous fungi (Morozov et al., 2006) and whether the PTCs identified in *F. graminearum* are also associated with NMD needs to be further investigated. Apart from PTC isoforms, some alternatively spliced transcripts encoding proteins with diverse length were identified. The effects of the diversity in length on the biological function of proteins are still unknown, but several functions including binding properties, intracellular localization, enzymatic activity or stability might be affected (Stamm et al., 2005).

Alternative splicing appears widespread in eukaryotes, but the biological function of alternative splicing is still poorly understood. Some studies have shown that alternative splicing events are developmentally regulated or associated with the response to different environmental conditions (Kalsotra and Cooper, 2011; Palusa et al., 2007). For instance, in *A. thaliana*, the

CIRCADIAN CLOCK ASSOCIATED 1 (CCA1) gene produces two different transcripts and their expression ratio is dependent of light and temperature (Filichkin et al., 2009). Similarly, in *A. thaliana* splicing of serine/arginine-rich protein-encoding genes is altered in response to hormones or abiotic stresses (Kalsotra and Cooper, 2011). In *H. sapiens*, a number of genes involved in apoptosis (Schwerk and Schulze-Osthoff, 2005) and differentiation of embryonic stem cells is regulated by alternative splicing (Pritsker et al., 2005; Salomonis et al., 2010). In our study, we have also demonstrated that for some genes the alternative splicing events are regulated at different vegetative growth stages in *F. graminearum*; their biological implications are not yet understood, but they might be important in adaptation of *F. graminearum* to changing external environmental conditions that occur during different growth stages.

As reported previously in other species, in addition to the canonical GT donor and AG acceptor sites in introns there are several non-canonical donor and acceptor sites, of which GC occurs most frequently as an alternative donor site (Burset et al., 2000; Filichkin et al., 2009; Sheth et al., 2006). Our discovery of non-canonical splice sites in *F. graminearum* also showed that the GC donor, AG acceptor combination is prevalent, of which the proportion is consistent with what has been found in other organisms (Filichkin et al., 2009; Sheth et al., 2006). In addition, the nucleotide preferences flanking the GC donor splice site and AG acceptor splice site identified in *F. graminearum* are consistent with previous reports in other organisms (Sheth et al., 2006).

Conclusions

We have analyzed the transcriptome of *F. graminearum* during growth in nutrient-rich medium by RNA-Seq and identified transcripts of 84 % of the predicted genes, which allowed us to not only significantly revise existing gene models present in the Broad *Fusarium* database but also to get preliminary information on the presence of alternative splicing in this fungus. This is one of the most comprehensive reports on alternative splicing in filamentous fungi. Our analyses indicate that the frequency of alternative splicing in *F. graminearum* is lower than *H. sapiens* and *A. thaliana*. Nevertheless, the expression of alternatively spliced genes appeared tightly regulated in different growth stage and can change from spore to mycelium within a few hours. This is the first indication that alternative splicing may be important in the developmental regulation in filamentous fungi. In the future, the biological functions of the different transcript isoforms and

their encoded proteins need to be studied in more detail.

Methods

Fungal strains and culture conditions

F. graminearum isolates wt PH-1 (PH-1) and the mutant *ebr1* (*ebr1*) were used in this study. PH-1 is the sequenced strain (Cuomo et al., 2007) and *ebr1* is a knock out mutant derived from PH-1 and its phenotype was recently described (Zhao et al., 2011). To prepare the mycelium for RNA-Seq, both PH-1 isolate and *ebr1* were grown as shake cultures in liquid mung bean medium for 3 days to produce conidia (25 °C, 200 rpm). Then 10^5 conidia of PH-1 and *ebr1* were transferred to 400 ml liquid complete medium (CM) (Leach et al., 1982) and grown as shake culture for 30 h to produce mycelium (25 °C, 200 rpm).

RNA isolation and RT-PCR

Mycelium harvested from PH-1 and *ebr1* was collected from liquid CM medium by filtration and ground in liquid nitrogen using a mortar and pestle. Ground mycelium was used for RNA extraction by TRIzol reagent (Invitrogen, Cat. No. 15596-018) according to the manufacturer's instructions. The quality of RNA was evaluated by Agilent 2100.

For reverse transcription (RT)-PCR, isolated RNA was treated with DNase I (Fermentas, #EN0521) according to the manufacturer's manual. The DNase I-treated RNA was reversely transcribed into cDNA by using M-MLV Reverse Transcriptase (Promega) according to the protocol described in the manual. cDNA was used as template to perform RT-PCR according to the following procedures: 20 μl reaction mixture including 2 μl 10 x reaction buffer, 0.8 μl dNTP (5mM), 0.5 μl forward primer (10μM), 0.5 μl reverse primer (10μM), 1 μl template, 0.3 μl Taq DNA polymerase (Roche) and 14.9 μl ddH_2O; reaction conditions including step 1 (94 °C 4 min), step 2 (94 °C 30s; 56 °C 30s; 72 °C 60s; this step was repeated 34 times) and step 3 (72 °C 10 min).

RNA-Seq analysis

Isolated RNA was enriched for mRNA by using oligo dT Dynabeads (Invitrogen) according to the manufacturer's instructions and fragmented into fragments of 200-700 nucleotides by

incubating at 70 °C for 15 min in fragmentation buffer (Ambion). Fragmentation of mRNA was terminated by adding stop solution (Ambion) and used as template to synthesize the first strand cDNA by using random hexamers (Invitrogen). Subsequently dNTPs, RNase and DNA polymerase were added to the reaction solution to synthesize the second strand cDNA. The synthesized cDNA was purified by Qiaquick PCR kits and blunted by an End Repair reaction. Subsequently a single "A" base was added to the 3' end of cDNA by using dATP and Klenow Exo Fragment. Later Illumina adaptors were linked to the cDNA ends. The adapted cDNA was run on agrose gel and ~200 bp cDNA fragments were selected. Finally, the cDNA was amplified and the obtained cDNA pool was subjected to high-throughput sequencing by Illumina HiSeq^{TN} 2000.

Reads mapping

The gene database, the transcript database, the supercontig database and the UTR database of *F. graminearum* were downloaded from the Broad Institute (http://www.broadinstitute.org/annotation/genome/fusarium_group/MultiDownloads.html). All these databases along with RNA-Seq raw data were imported into the CLC genomic workbench software according to the method described in the manual. The "RNA-Seq analysis" option was used to map reads to each database at the following settings: minimum length fraction 0.9, minimum similarity fraction 0.8, and maximum number of hits for a read 30. The matched reads were visualized in the CLC interface.

Identification of incorrect gene models and alternative splicing

Three strategies were employed to identify incorrect gene models and alternative splicing. In the first strategy, we mapped all reads from PH-1 and *ebr1* RNA-Seq data (25,720,650 reads in total) against complete transcript database (only exonic regions). After this round of mapping, 13,073,825 unmapped reads were obtained that were subsequently aligned against the 5' UTR (1000 bp) and 3' UTR (1000 bp) databases for the second round of mapping, after which 6,995,901 unmapped reads were obtained. This set of unmapped reads can be divided in four fractions: (i) reads matching to intergenic regions, (ii) reads matching to intronic regions, (iii) reads matching to the border of coding regions and UTRs, and (iv) non-mapped reads. Finally, the 6,995,901 unmapped reads were aligned against the gene database (including exons, introns and

UTRs). From this round of mapping, 732,254 reads were identified matching to genes, from which we could collect all genes with introns matched by reads. In the second strategy, all reads were mapped against the transcript database and the matched reads for each transcript were visualized in the CLC interface. We browsed all transcripts that contained more than 200 matched reads, from which we collected transcripts with non-matched or mismatched regions. For the non-matched regions, there must be at least one read flanking this region showing a splice site. In the third strategy, we employed the TopHat program to identify incorrect gene models and alternative splicing according to a previously described protocol (Trapnell et al., 2009; Trapnell et al., 2012).

All the genes collected by these three strategies were first examined in the MIPS *F. graminearum* database to exclude genes that had already been revised manually. The remaining genes were manually examined in the CLC software package by comparing RNA-Seq reads with the predicted gene models to identify incorrectly annotated genes or alternatively spliced genes. A number of genes from each category was selected for confirmation by RT-PCR.

Identification of nTARs

We aligned all reads against the supercontigs of *F. graminearum* and collected all reads matched regions (more than two read coverage on average and more than 150 bp in length) that located in the intergenic regions 200 bp away from flanking gene models. To analyze whether these nTARs encode mRNAs, we collected their sequences and blasted them against the MIPS *F. graminearum* database to identify novel genes that had already been annotated and against the Broad *Fusarium* database to identify orthologous genes in *F. oxysporum* and *F. verticillioides.*

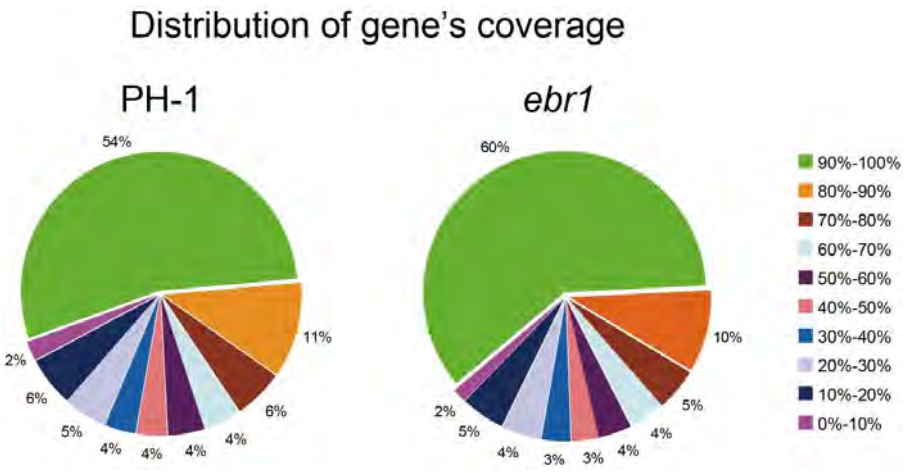
RNA editing analysis

All reads from PH-1 and *ebr1* were first aligned against the gene database of *F. graminearum* in CLC and the “SNP analysis” module was used to analyze putative SNPs between the RNA-Seq and the genome data. To confirm SNPs, primers were designed in flanking regions and PCRs were performed by using genomic DNA and cDNA, respectively, as templates. The amplicons were sequenced and the obtained sequences were aligned to the annotated genomic sequence to identify putative RNA editing.

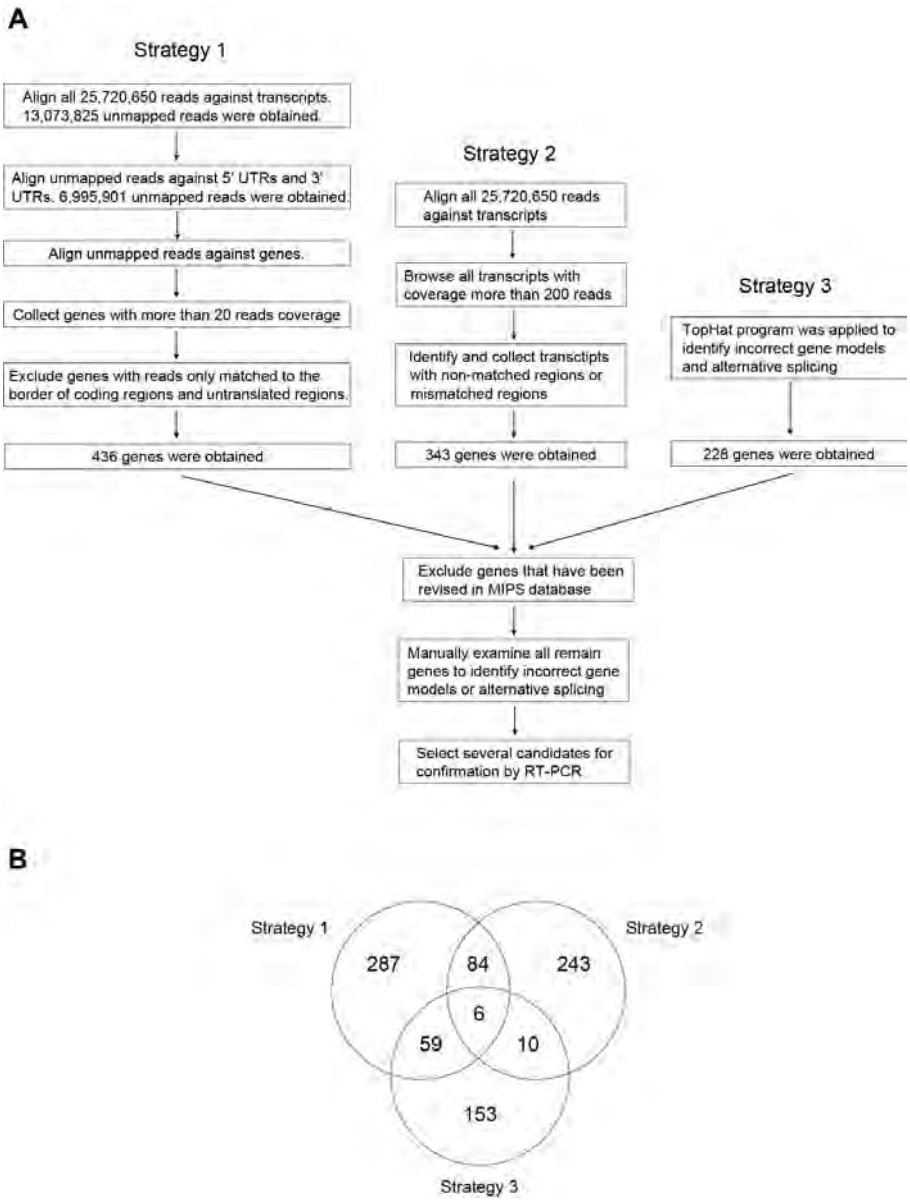
Acknowledgments

This work was supported by a grant from the National Basic Research Program of China (2011CB100700). P. J. G. M. de Wit is supported by grants from the Royal Netherlands Academy of Arts and Sciences, the Centre for BioSystems Genomics.

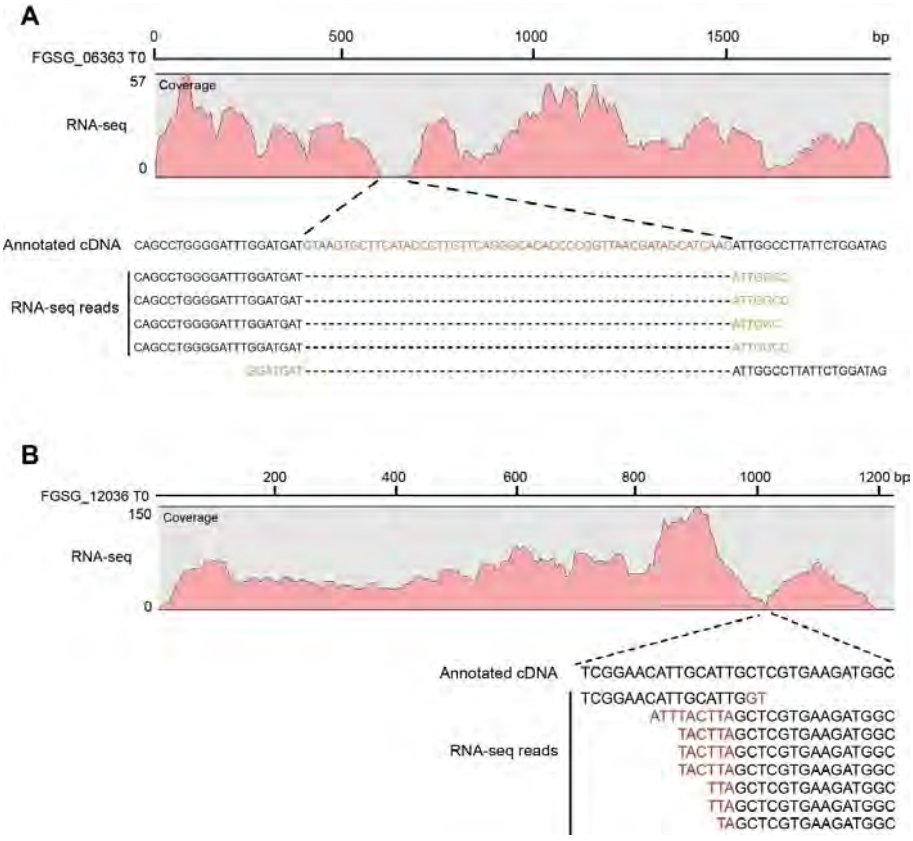
Supplementary data



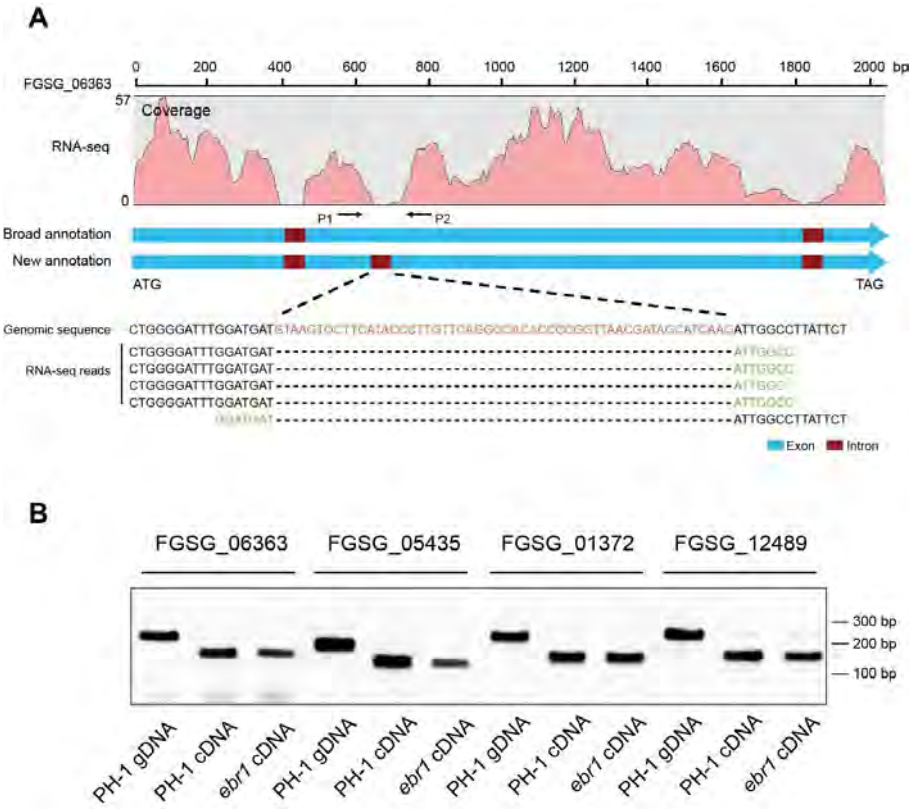
Supplemental Figure 1. Statistical analysis of gene coverage by RNA-seq reads in wt PH-1 and mutant *ebr1*. Graphs show that 54% of the genes in PH-1 and 60% of the genes in *ebr1* are over 90% covered by reads.



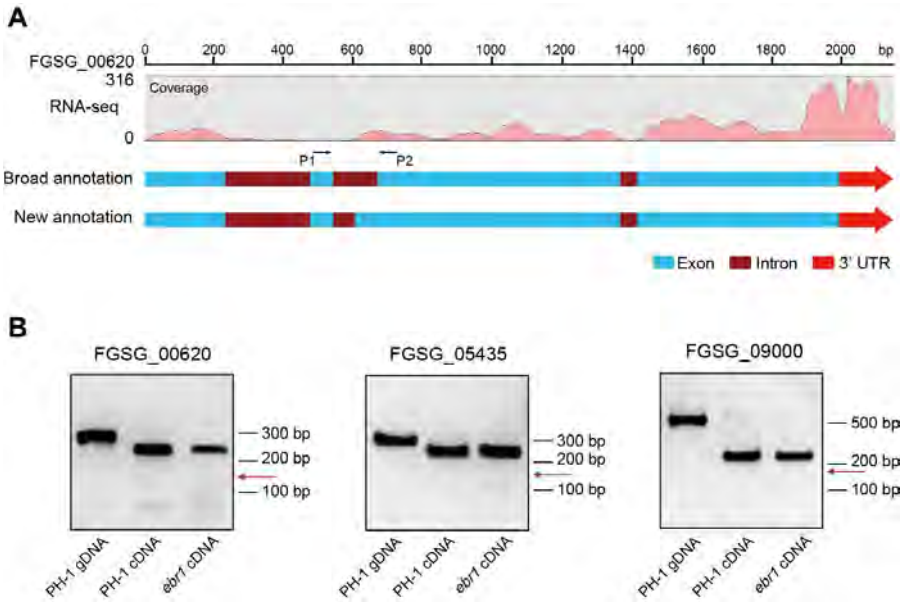
Supplemental Figure 2. Strategies used to identify incorrect gene models and alternative splicing. A. Three complementary strategies were employed to identify genes with incorrect gene models or alternative splicing. B. Comparison of genes identified by three different strategies.



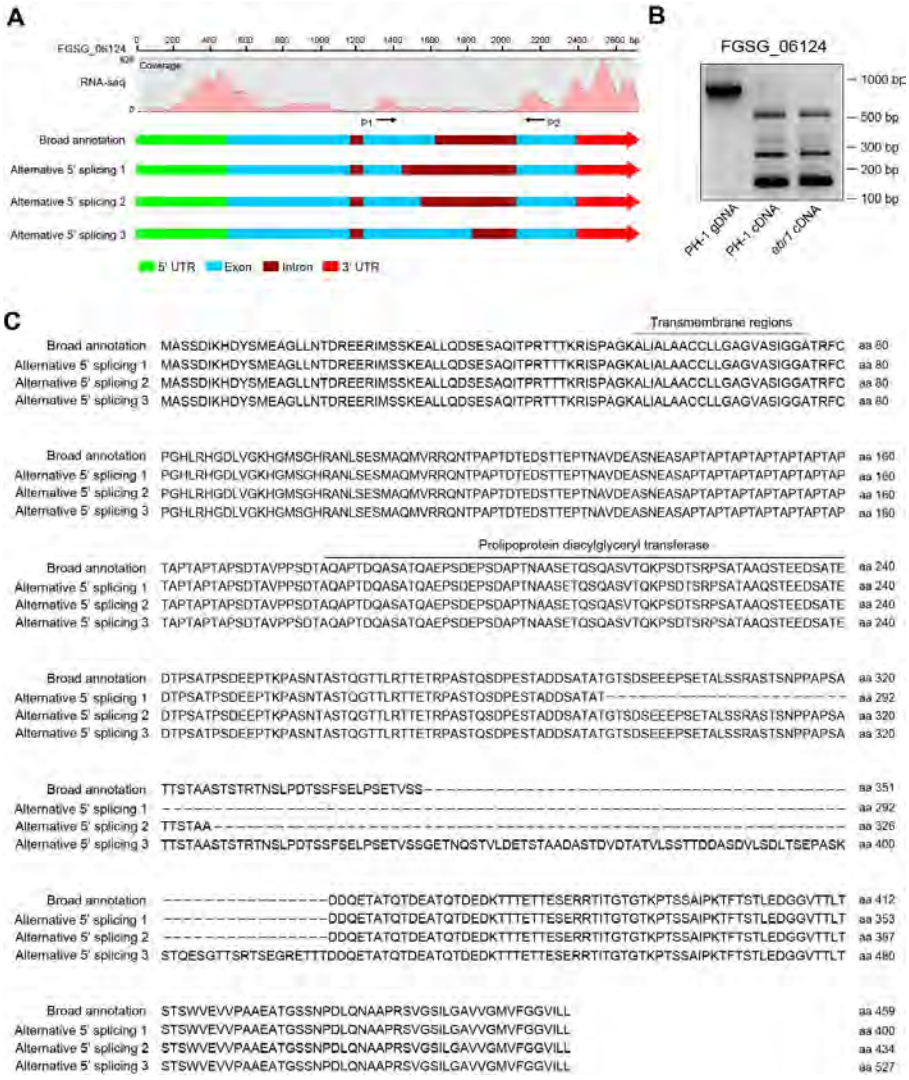
Supplemental Figure 3. Two examples of transcripts derived from incorrect gene models. A. Transcript with non-matched region. When aligning reads with the annotated transcript of gene FGSG_06363, a non-matched region was identified. Five reads flanking this region show splice sites (green letters), indicative of a novel intron in gene FGSG_06363. B. Transcript with mismatched region. One position with mismatched reads was identified when aligning reads with the annotated transcript of gene FGSG_12036. The nucleotides (red letter) at the end of eight reads do not match to the transcript of FGSG_12036, indicative of an incorrectly predicted intron in gene FGSG_12036. Black letters represent exonic region; orange letters represent intronic region.



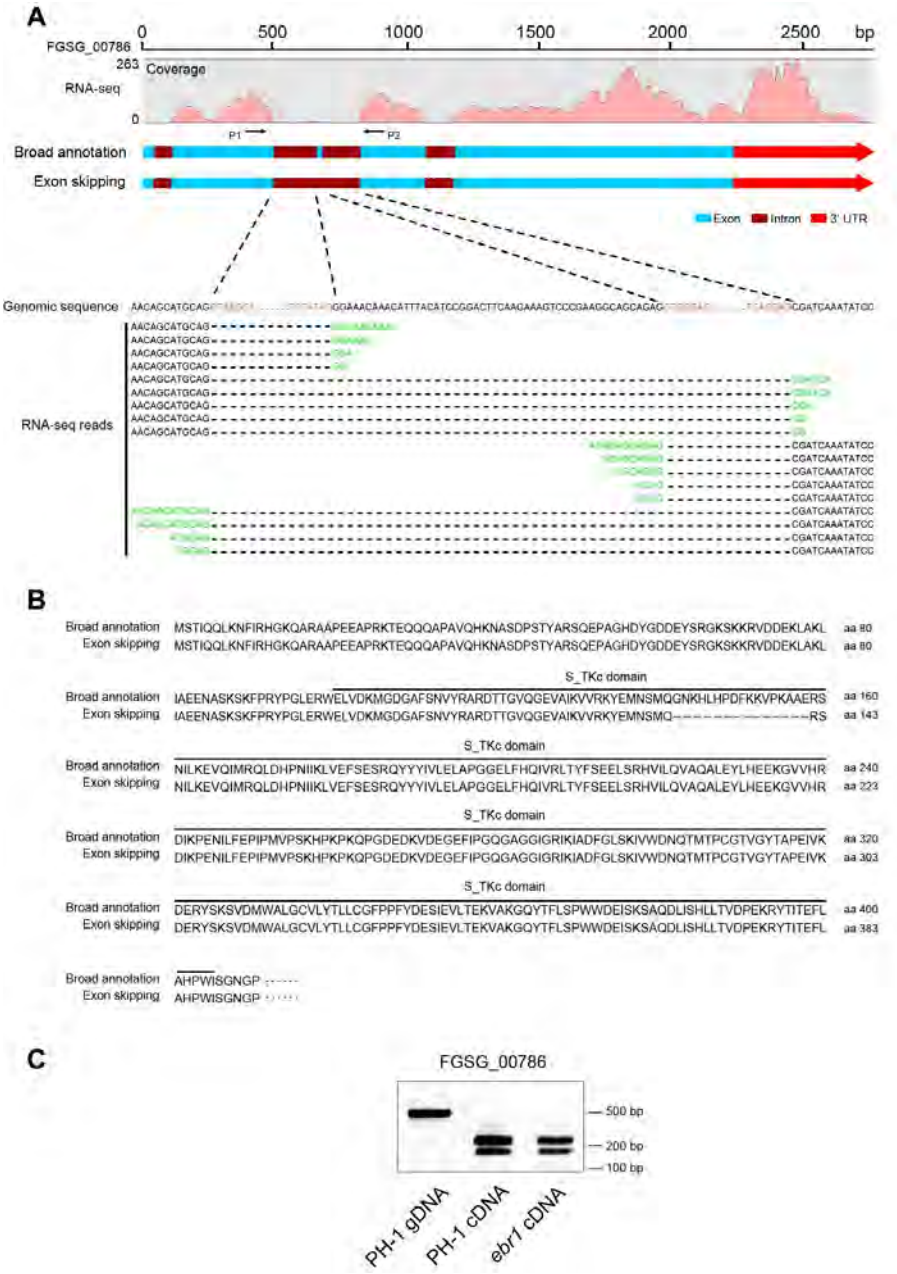
Supplemental Figure 4. Identification of novel introns. A. A novel intron was identified in gene FGSG_06363. The novel intron was supported by the reads that show the intron splice sites (green letters). Black letters represent exonic region; orange letters represent the intronic region. B. Four genes were randomly selected for the confirmation of novel introns. Primers were designed flanking the intronic regions as shown in A.



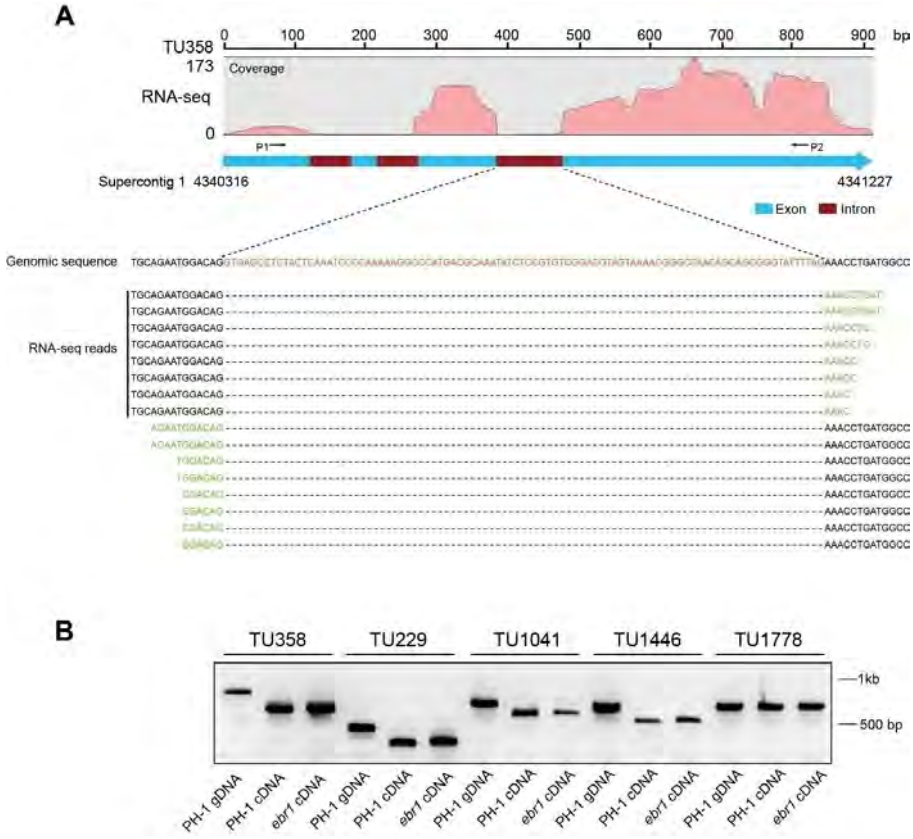
Supplemental Figure 5. Identification of incorrect intron splice sites. A. RNA-Seq data show that the 3' splice site of the second intron of gene FGSG_00620 is different from the annotated gene present in the Broad *F. graminearum* database. B. Three genes were selected to confirm the incorrect intron splice sites. Red arrows indicate the size of the amplified products by PCR that should be obtained according to gene models present in the Broad *F. graminearum* database. Primers were designed flanking the intronic regions.



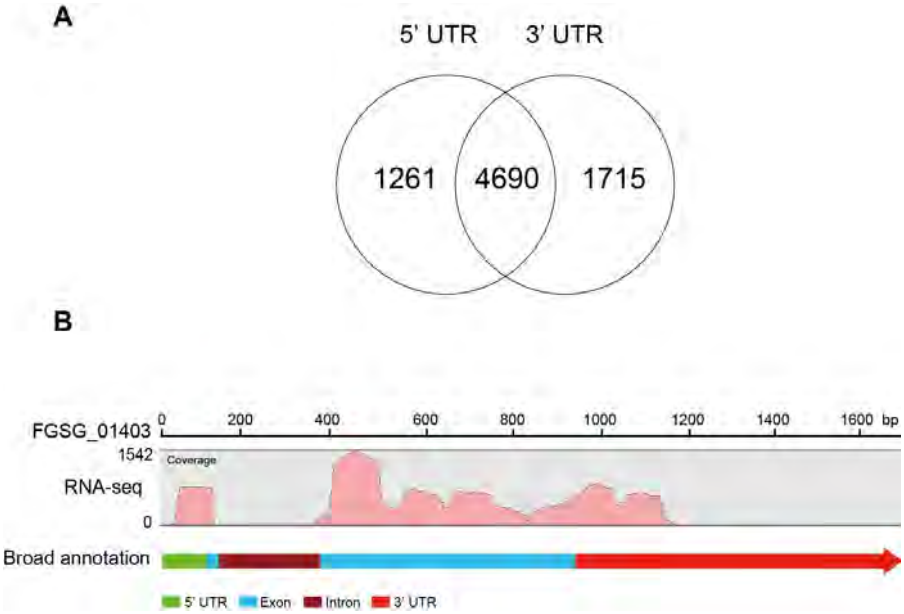
Supplemental Figure 6. Example of alternative 5' splicing. A. RNA-seq reads show that there are four different 5' splice sites of the second intron in gene FGSG_06124. B. All four transcripts with different sizes were confirmed by RT-PCR. Primers were designed flanking the intronic region as shown in A. C. Protein alignment shows the consequence of the alternative 5' splicing in gene FGSG_06124. Several amino acids are present or absent in the alternative proteins compared to the annotated protein.



Supplemental Figure 7. Example of exon skipping. A. RNA-seq data show that the third exon is skipped in gene FGSG_00786. From the reads on the left side, some intron splice sites (green letters) match to the third exon and some to the fourth exon. Similarly, the splice sites from the reads on the right side match to either the third or the second exon. B. Protein alignment shows that 17 amino acids are absent in the S_TKc domain in case of exon skipping. C. RT-PCR confirmed exon skipping of the third exon by detection of two bands with the expected sizes.



Supplemental Figure 8. Identification of novel transcriptional active regions (nTARs). A. TU358 is shown as one of the nTARs encoding a putative mRNA. Three introns in TU358 were identified and one of them is displayed. B. Four nTARs containing intronic regions and one nTARs without intronic region were selected for confirmation by RT-PCR.



Supplemental Figure 9. Determination of 5' and 3' UTR regions. A. For 4690 genes, both 5' UTRs and 3'UTRs were identified. For 1261 genes only 5'UTRs and for 1715 genes only 3' UTRs were identified. B. Comparison of UTRs predicted in the Broad *F. graminearum* database to RNA-seq data identified some incorrectly predicted UTRs. The graph shows one example where the length of the 3'UTR in gene FGSG_01403 is shorter than the length annotated in the Broad *F. graminearum* database.

References

Ali, G.S., Palusa, S.G., Golovkin, M., Prasad, J., Manley, J.L., and Reddy, A.S. (2007). Regulation of plant developmental processes by a novel splicing factor. PLoS One *2*, e471.

Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., Ren, J., Li, W.W., and Noble, W.S. (2009). MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res. *37*, W202-208.

Bluhm, B.H., Zhao, X., Flaherty, J.E., Xu, J.R., and Dunkle, L.D. (2007). *RAS2* regulates growth and pathogenesis in *Fusarium graminearum*. Mol. Plant Microbe Interact. *20*, 627-636.

Brown, D.W., Butchko, R.A., and Proctor, R.H. (2008). Genomic analysis of *Fusarium verticillioides*. Food Addit. Contam. Part A *25*, 1158-1165.

Bruno, V.M., Wang, Z., Marjani, S.L., Euskirchen, G.M., Martin, J., Sherlock, G., and Snyder, M. (2010). Comprehensive annotation of the transcriptome of the human fungal pathogen *Candida albicans* using RNA-seq. Genome Res. *20*, 1451-1458.

Burset, M., Seledtsov, I.A., and Solovyev, V.V. (2000). Analysis of canonical and non-canonical splice sites in mammalian genomes. Nucleic Acids Res. *28*, 4364-4375.

Cui, Y., Hagan, K.W., Zhang, S., and Peltz, S.W. (1995). Identification and characterization of genes that are required for the accelerated degradation of mRNAs containing a premature translational termination codon. Genes Dev. *9*, 423-436.

Cuomo, C.A., Guldener, U., Xu, J.R., Trail, F., Turgeon, B.G., Di Pietro, A., Walton, J.D., Ma, L.J., Baker, S.E., Rep, M., *et al.* (2007). The *Fusarium graminearum* genome reveals a link between localized polymorphism and pathogen specialization. Science *317*, 1400-1402.

Desjardins, A.E., Hohn, T.M., and McCormick, S.P. (1993). Trichothecene biosynthesis in *Fusarium* species: chemistry, genetics, and significance. Microbiol. Rev. *57*, 595-604.

Filichkin, S.A., Priest, H.D., Givan, S.A., Shen, R., Bryant, D.W., Fox, S.E., Wong, W.K., and Mockler, T.C. (2009). Genome-wide mapping of alternative splicing in *Arabidopsis thaliana*. Genome Res. *20*, 45-58.

Galagan, J.E., Henn, M.R., Ma, L.J., Cuomo, C.A., and Birren, B. (2005). Genomics of the fungal kingdom: insights into eukaryotic biology. Genome Res. *15*, 1620-1631.

Gan, X., Stegle, O., Behr, J., Steffen, J.G., Drewe, P., Hildebrand, K.L., Lyngsoe, R., Schultheiss, S.J., Osborne, E.J., Sreedharan, V.T., *et al.* (2011). Multiple reference genomes and transcriptomes for *Arabidopsis thaliana*. Nature *477*, 419-423.

Gehman, L.T., Stoilov, P., Maguire, J., Damianov, A., Lin, C.H., Shiue, L., Ares, M., Jr., Mody, I., and Black, D.L. (2011). The splicing regulator Rbfox1 (A2BP1) controls neuronal excitation in the mammalian brain. Nat. Genet. *43*, 706-711.

Guida, A., Lindstadt, C., Maguire, S.L., Ding, C., Higgins, D.G., Corton, N.J., Berriman, M., and Butler, G. (2011). Using RNA-seq to determine the transcriptional landscape and the hypoxic response of the pathogenic yeast *Candida parapsilosis*. BMC Genomics *12*, 628.

Guldener, U., Mannhaupt, G., Munsterkotter, M., Haase, D., Oesterheld, M., Stumpflen, V., Mewes, H.W., and Adam, G. (2006). FGDB: a comprehensive fungal genome resource on the plant pathogen *Fusarium graminearum*. Nucleic Acids Res. *34*, D456-458.

Heyd, F., and Lynch, K.W. (2011). Degrade, move, regroup: signaling control of splicing proteins. Trends Biochem. Sci. *36*, 397-404.

Kalsotra, A., and Cooper, T.A. (2011). Functional consequences of developmentally regulated alternative splicing. Nat. Rev. Genet. *12*, 715-729.

Kalsotra, A., Xiao, X., Ward, A.J., Castle, J.C., Johnson, J.M., Burge, C.B., and Cooper, T.A. (2008). A postnatal switch of CELF and MBNL proteins reprograms alternative splicing in the developing heart. Proc. Natl. Acad. Sci. USA *105*, 20333-20338.

Kazan, K., Gardiner, D.M., and Manners, J.M. (2011). On the trail of a cereal killer: recent advances in *Fusarium graminearum* pathogenomics and host resistance. Mol. Plant Pathol. *13*, 399-413.

Leach, J., Lang, B.R., and Yoder, O.C. (1982). Methods for selection of mutants and in vitro culture of *Cochliobolus heterostrophus*. Microbiology *128*, 1719-1729.

Lee, B.S., and Culbertson, M.R. (1995). Identification of an additional gene required for eukaryotic nonsense mRNA turnover. Proc. Natl. Acad. Sci. USA *92*, 10354-10358.

Lee, T., Oh, D.W., Kim, H.S., Lee, J., Kim, Y.H., Yun, S.H., and Lee, Y.W. (2001). Identification of deoxynivalenol- and nivalenol-producing chemotypes of *Gibberella zeae* by using PCR. Appl. Environ. Microbiol. *67*, 2966-2972.

Leeds, P., Wood, J.M., Lee, B.S., and Culbertson, M.R. (1992). Gene products that promote mRNA turnover in *Saccharomyces cerevisiae*. Mol. Cell. Biol. *12*, 2165-2177.

Li, Q., Lee, J.A., and Black, D.L. (2007). Neuronal regulation of alternative pre-mRNA splicing. Nat. Rev. Neurosci. *8*, 819-831.

Li, Z., Zhang, Z., Yan, P., Huang, S., Fei, Z., and Lin, K. (2011). RNA-Seq improves annotation of protein-coding genes in the cucumber genome. BMC Genomics *12*, 540.

Loftus, B.J., Fung, E., Roncaglia, P., Rowley, D., Amedeo, P., Bruno, D., Vamathevan, J., Miranda, M., Anderson, I.J., Fraser, J.A., *et al.* (2005). The genome of the basidiomycetous yeast and human pathogen *Cryptococcus neoformans*. Science *307*, 1321-1324.

Lysoe, E., Klemsdal, S.S., Bone, K.R., Frandsen, R.J., Johansen, T., Thrane, U., and Giese, H. (2006). The *PKS4* gene of *Fusarium graminearum* is essential for zearalenone production. Appl. Environ. Microbiol. *72*, 3924-3932.

Mayshar, Y., Rom, E., Chumakov, I., Kronman, A., Yayon, A., and Benvenisty, N. (2008). Fibroblast growth factor 4 and its novel splice isoform have opposing effects on the maintenance of human embryonic stem cell self-renewal. Stem cells *26*, 767-774.

McGuire, A.M., Pearson, M.D., Neafsey, D.E., and Galagan, J.E. (2008). Cross-kingdom patterns of alternative splicing and splice recognition. Genome Biol. *9*, R50.

Morozov, I.Y., Negrete-Urtasun, S., Tilburn, J., Jansen, C.A., Caddick, M.X., and Arst, H.N., Jr. (2006). Nonsense-mediated mRNA decay mutation in *Aspergillus nidulans*. Eukaryot. Cell *5*, 1838-1846.

Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L., and Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat. Methods *5*, 621-628.

Nagalakshmi, U., Wang, Z., Waern, K., Shou, C., Raha, D., Gerstein, M., and Snyder, M. (2008). The transcriptional landscape of the yeast genome defined by RNA sequencing. Science *320*, 1344-1349.

Ner-Gaon, H., Halachmi, R., Savaldi-Goldstein, S., Rubin, E., Ophir, R., and Fluhr, R. (2004). Intron retention is a major phenomenon in alternative splicing in *Arabidopsis*. Plant J. *39*, 877-885.

Palusa, S.G., Ali, G.S., and Reddy, A.S. (2007). Alternative splicing of pre-mRNAs of *Arabidopsis* serine/arginine-rich proteins: regulation by hormones and stresses. Plant J. *49*, 1091-1107.

Pan, Q., Shai, O., Lee, L.J., Frey, B.J., and Blencowe, B.J. (2008). Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. Nat. Genet. *40*, 1413-1415.

Peng, Z., Cheng, Y., Tan, B.C., Kang, L., Tian, Z., Zhu, Y., Zhang, W., Liang, Y., Hu, X., Tan, X., *et al.* (2012). Comprehensive analysis of RNA-Seq data reveals extensive RNA editing in a human transcriptome. Nat. Biotechnol. *30*, 253-260.

Pritsker, M., Doniger, T.T., Kramer, L.C., Westcot, S.E., and Lemischka, I.R. (2005). Diversification of stem cell molecular repertoire by alternative splicing. Proc. Natl. Acad. Sci. USA *102*, 14290-14295.

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

Ramani, A.K., Calarco, J.A., Pan, Q., Mavandadi, S., Wang, Y., Nelson, A.C., Lee, L.J., Morris, Q.,

Blencowe, B.J., Zhen, M., *et al.* (2010). Genome-wide analysis of alternative splicing in *Caenorhabditis elegans*. Genome Res. *21*, 342-348.

Reddy, A.S. (2007). Alternative splicing of pre-messenger RNAs in plants in the genomic era. Annu. Rev. plant Biol. *58*, 267-294.

Salomonis, N., Schlieve, C.R., Pereira, L., Wahlquist, C., Colas, A., Zambon, A.C., Vranizan, K., Spindler, M.J., Pico, A.R., Cline, M.S., *et al.* (2010). Alternative splicing regulates mouse embryonic stem cell pluripotency and differentiation. Proc. Natl. Acad. Sci. USA *107*, 10514-10519.

Saltzman, A.L., Kim, Y.K., Pan, Q., Fagnani, M.M., Maquat, L.E., and Blencowe, B.J. (2008). Regulation of multiple core spliceosomal proteins by alternative splicing-coupled nonsense-mediated mRNA decay. Mol. Cell. Biol. *28*, 4320-4330.

Schmucker, D., Clemens, J.C., Shu, H., Worby, C.A., Xiao, J., Muda, M., Dixon, J.E., and Zipursky, S.L. (2000). *Drosophila* Dscam is an axon guidance receptor exhibiting extraordinary molecular diversity. Cell *101*, 671-684.

Schwerk, C., and Schulze-Osthoff, K. (2005). Regulation of apoptosis by alternative pre-mRNA splicing. Mol. cell *19*, 1-13.

Seong, K.Y., Zhao, X., Xu, J.R., Guldener, U., and Kistler, H.C. (2008). Conidial germination in the filamentous fungus *Fusarium graminearum*. Fungal Genet. Biol. *45*, 389-399.

Sheth, N., Roca, X., Hastings, M.L., Roeder, T., Krainer, A.R., and Sachidanandam, R. (2006). Comprehensive splice-site analysis using comparative genomics. Nucleic Acids Res. *34*, 3955-3967.

Stamm, S., Ben-Ari, S., Rafalska, I., Tang, Y., Zhang, Z., Toiber, D., Thanaraj, T.A., and Soreq, H. (2005). Function of alternative splicing. Gene *344*, 1-20.

Sultan, M., Schulz, M.H., Richard, H., Magen, A., Klingenhoff, A., Scherf, M., Seifert, M., Borodina, T., Soldatov, A., Parkhomchuk, D., *et al.* (2008). A global view of gene activity and alternative splicing by deep sequencing of the human transcriptome. Science *321*, 956-960.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics *25*, 1105-1111.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., and Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat. protoc. *7*, 562-578.

Venables, J.P., Tazi, J., and Juge, F. (2011). Regulated functional alternative splicing in *Drosophila*. Nucleic Acids Res. *40*, 1-10.

Wang, B.B., and Brendel, V. (2006). Genomewide comparative analysis of alternative splicing in plants. Proc. Natl. Acad. Sci. USA *103*, 7175-7180.

Wang, E.T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., Kingsmore, S.F., Schroth, G.P., and Burge, C.B. (2008). Alternative isoform regulation in human tissue transcriptomes. Nature *456*, 470-476.

Wang, Z., Gerstein, M., and Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. Nat. Rev. Genet. *10*, 57-63.

Wong, P., Walter, M., Lee, W., Mannhaupt, G., Munsterkotter, M., Mewes, H.W., Adam, G., and Guldener, U. (2011). FGDB: revisiting the genome annotation of the plant pathogen *Fusarium graminearum*. Nucleic acids Res. *39*, D637-639.

Xu, X., Yang, D., Ding, J.H., Wang, W., Chu, P.H., Dalton, N.D., Wang, H.Y., Bermingham, J.R., Jr., Ye, Z., Liu, F., *et al.* (2005). ASF/SF2-regulated CaMKIIdelta alternative splicing temporally reprograms excitation-contraction coupling in cardiac muscle. Cell *120*, 59-72.

Yeo, G.W., Van Nostrand, E., Holste, D., Poggio, T., and Burge, C.B. (2005). Identification and analysis of alternative splicing events conserved in human and mouse. Proc. Natl. Acad. Sci. USA *102*, 2850-2855.

Yeo, G.W., Xu, X., Liang, T.Y., Muotri, A.R., Carson, C.T., Coufal, N.G., and Gage, F.H. (2007). Alternative splicing events identified in human embryonic stem cells and neural progenitors. PLoS Comput. Biol. *3*, 1951-1967.

Zhao, C., Waalwijk, C., de Wit, P.J., van der Lee, T., and Tang, D. (2011). EBR1, a novel Zn(2)Cys(6) transcription factor, affects virulence and apical dominance of the hyphal tip in *Fusarium graminearum*. Mol. Plant Microbe Interact. *24*, 1407-1418.

Chapter 4

Relocation of genes generates non-conserved chromosomal segments in *Fusarium graminearum* that show distinct and co-regulated gene expression patterns

Chunzhao Zhao,[1,2,3,4] Cees Waalwijk,[1,2] Pierre J. G. M. de Wit,[2,5] Dingzhong Tang[3] and Theo van der Lee[1,2]

1. Plant Research International, P.O. Box 6708 PB, Wageningen, The Netherlands
2. Graduate School Experimental Plant Sciences, Wageningen, The Netherlands
3. State Key Laboratory of Plant Cell and Chromosome Engineering, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing 100101, China
4. Graduate University of Chinese Academy of Sciences, Beijing 100049, China
5. Wageningen University, Laboratory of Phytopathology, P.O. Box 6708 PB, Wageningen, The Netherlands

Published in BMC Genomics 15:191 (2014)

Abstract

Genome comparisons between closely related species often show non-conserved regions across chromosomes. Some of them are located in specific regions of chromosomes and some are even confined to one or more entire chromosomes. The origin and biological relevance of these non-conserved regions are still largely unknown. Here we used the genome of *Fusarium graminearum* to elucidate the significance of non-conserved regions. The genome of *F. graminearum* harbours thirteen non-conserved regions dispersed over all of the four chromosomes. Using RNA-Seq data from the mycelium of *F. graminearum*, we found weakly expressed regions on all of the four chromosomes that exactly matched with non-conserved regions. Comparison of gene expression between two different developmental stages (conidia and mycelium) showed that the expression of genes in conserved regions is stable, while gene expression in non-conserved regions is much more influenced by developmental stage. In addition, genes involved in the production of secondary metabolites and secreted proteins are enriched in non-conserved regions, suggesting that these regions could also be important for adaptations to new environments, including adaptation to new hosts. Finally, we found evidence that non-conserved regions are generated by sequestration of genes from multiple locations. Gene relocations may lead to clustering of genes with similar expression patterns or similar biological functions, which was clearly exemplified by the PKS2 gene cluster.

Introduction

The genus *Fusarium* includes a large group of phytopathogenic fungi, each having a different or partially overlapping host range. *F. graminearum* is an important pathogen on wheat, barley and maize, causing Fusarium head blight (Bai and Shaner, 2004; Parry et al., 1995), while *F. verticillioides* mainly infects maize, causing rot and wilting (Oren et al., 2003). In contrast, *F. oxysporum* causes disease on more than 100 different plant species, but specific strains usually infect only a single host species (Kistler et al., 1998). This has led to the introduction of the *forma specialis* concept in *F. oxysporum*, where strains are categorized according to the host plant they infect, such as tomato or banana [5]. In addition, some strains of *F. oxysporum* may also cause infections in humans (Gordon and Martyn, 1997; Ortoneda et al., 2004). Genomic sequences of several *Fusarium* species have been generated, including *F. graminearum* (strain PH-1), *F. verticillioides* (strain FGSC 7600), *F. pseudograminearum* (strain CS3096), *F. solani* (strain FGSC 9596), *F. fujikuroi* (strain IMI58289) and *F. oxysporum* f. sp. *lycopersici* strain 4287. In addition, 11 different *F. oxysporum* strains, including pathogens from different hosts as well as a biocontrol strain have been sequenced (Coleman et al., 2009; Cuomo et al., 2007; Gardiner et al., 2012; Ma et al., 2010; Wiemann et al., 2013). Genome comparisons showed that chromosome XII of *F. fujikuroi* is absent in the genome sequence of related species *F. verticillioides*, while 285 and 820 kb genome sequences at both ends of chromosome IV of *F. verticillioides* are absent in *F. fujikuroi* [11]. Comparing the genome of *F. pseudograminearum* with *F. graminearum* showed that 89.8% of genomic sequence could be aligned at >70% nucleotide identity [7]. Strikingly, *F. oxysporum* 4287 includes 11 core chromosomes and four lineage-specific (LS) chromosomes. LS chromosomes are specific to *F. oxysporum* strain 4287 and have no collinear chromosomes in either *F. graminearum, F. verticillioides* or other *F. oxysporum* strains. The LS chromosomes in *F. oxysporum* 4287 are suggested to have resulted from horizontal transfer from unknown sources (Ma et al., 2010). On LS chromosome 14, several effector-encoding genes have been identified that facilitate infection of its host plant (Houterman et al., 2009; Rep et al., 2004; van der Does et al., 2008).

In contrast to the high number of chromosomes present in *F. verticillioides* (11), *F. oxysporum* (15) and *F. fujikuroi* (12), *F. graminearum* has only four chromosomes, which

probably resulted from fusions of ancestral chromosomes (Cuomo et al., 2007). Comparing the genome of *F. graminearum* isolate PH-1 with *F. graminearum* isolate GZ3639 revealed several regions with high SNP density (Cuomo et al., 2007). In addition, comparison of the genome of *F. graminearum* with the closely related species *F. verticillioides* and *F. oxysporum* revealed several non-conserved regions (Ma et al., 2010). Further analysis showed that high SNP density regions match with non-conserved regions. Although many of the genes specifically expressed during plant infection are enriched in non-conserved regions (Cuomo et al., 2007), the origin and the biological relevance for these non-conserved regions are still largely unknown. In this study, we explored RNA-Seq data from both conidia and mycelium of *F. graminearum* to investigate the putative effect of gene locations (in conserved or non-conserved regions) on their expression patterns. In addition, by comparing the genome of *F. graminearum* with those of *F. verticillioides* and *F. oxysporum*, we show a possible mechanism for the generation of non-conserved regions.

Results

Synteny block analysis between the genomes of *Fusarium graminearum*, *F. verticillioides* and *F. oxysporum*

To identify collinear regions on the chromosomes of *F. graminearum* (Fg) and *F. verticillioides* (Fv), we performed synteny block analysis presented in Figure 1. This analysis revealed that *F. graminearum* chromosome 1 (Fgchr 1) is largely collinear with *F. verticillioides* chromosomes (Fvchrs) 1, 5 and 8; Fgchr 2 is collinear with Fvchrs 6, 9, 10 and 11; Fgchr 3 is collinear with Fvchrs 2, 4 and 7; while Fgchr 4 is largely collinear with Fvchrs 2 and 3 (Figure 1A). These major collinear regions between chromosomes indicate that the four *F. graminearum* chromosomes were likely formed primarily through end-to-end joining of two to four smaller ancestral chromosomes, descendants of which still exist in *F. verticillioides*. Similar collinear patterns were observed when comparing the genome of *F. graminearum* with those of *F. oxysporum* (Figure 1B). In addition to these major collinear regions, translocations and inversions of many chromosomal segments are observed in the chromosomes of *F. graminearum*. For instance, the first half of Fgchr 1 is homologous to Fvchr 1, but the order of genomic sequences on Fgchr 1 is different from Fvchr 1. Some

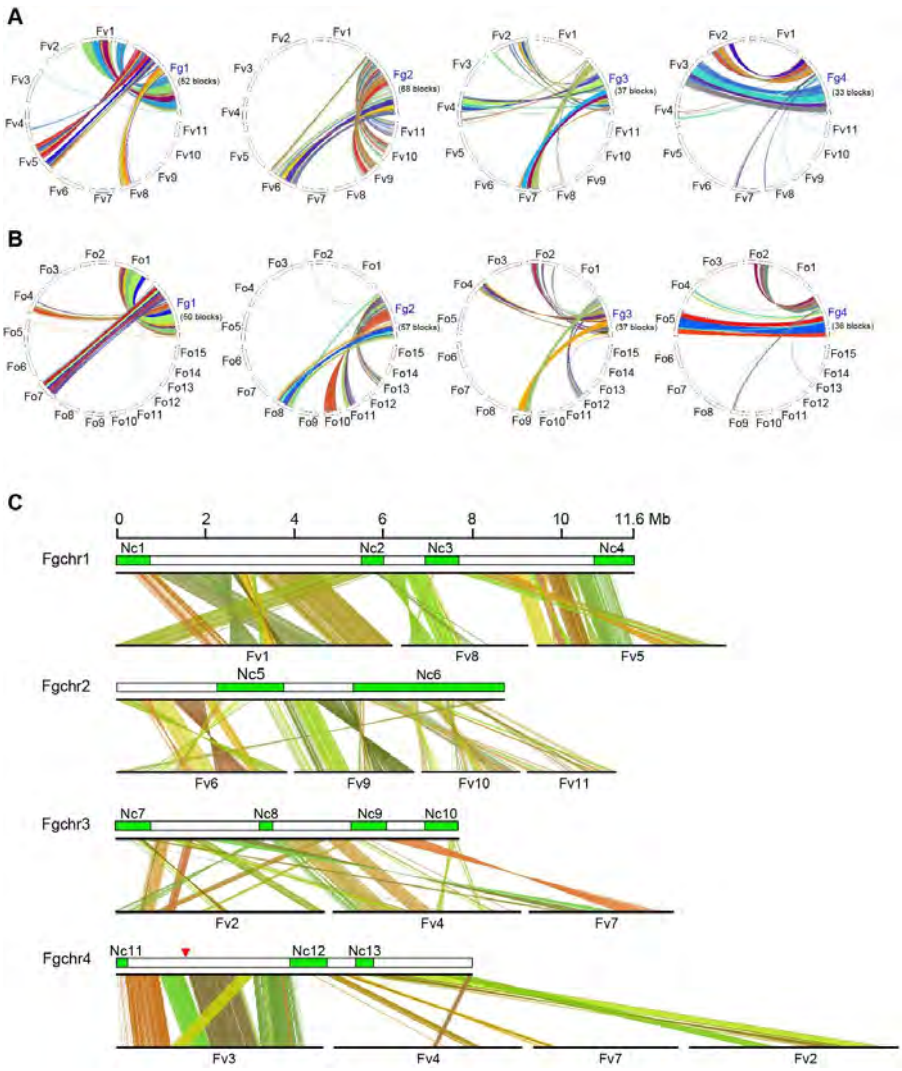


Figure 1. Comparison of synteny blocks in *Fusarium graminearum* with *F. verticillioides* and *F. oxysporum*. The genes on each chromosome of *F. graminearum* were compared with the genes of *F. verticillioides* (A) and *F. oxysporum* (B) and synteny blocks were analyzed by *MCScanX* (Wang et al., 2012). The results are shown in a cycle plot. *F. graminearum* contains four chromosomes, *F. verticillioides* 11, and *F. oxysporum* 15, respectively. Each colour represents an independent synteny block, the number of which is shown for each chromosome. (C) Detailed analysis of synteny blocks of each chromosome of *F. graminearum* with their collinear chromosomes of *F. verticillioides* by using *MCScanX*. The results are shown as a dual synteny plot. Thirteen non-conserved regions were identified on the chromosomes of *F. graminearum*, which were designated Nc1 to Nc13. Red triangle represents the region of *F. graminearum* that matches unassembled genomic regions of *F. verticillioides* (not shown).

chromosomal regions are translocated to new positions, while some are inversed (Figure 1A, C). In addition, thirteen non-conserved regions, which were designated Nc1-Nc13, were found on the four chromosomes of *F. graminearum*. These non-conserved regions exhibit almost no synteny with the chromosomes of *F. verticillioides* (Figure 1C). Notably, although the second half of Fgchr 2 shows collinearity with Fvchrs 10 and 11, a large number of small synteny blocks was identified, indicating multiple rearrangements in this part of chromosome 2 (Figure 1A, C). Similar chromosome rearrangement patterns were observed when the chromosomes of *F. graminearum* were compared with those of *F. oxysporum* 4287 (Figure 1B). We also performed synteny block analysis between *F. oxysporum* and *F. verticillioides*. Consistent with the notion that *F. oxysporum* and *F. verticillioides* are more closely related to each other than to *F. graminearum*, the collinearity between *F. oxysporum* and *F. verticillioides* was much higher (Supplemental Figure 1).

Gene expression analysis on individual chromosomes of *F. graminearum*

To investigate a possible correlation between chromosome rearrangements and gene expression patterns on each chromosome of *F. graminearum*, we performed RNA-Seq analysis on mycelium of wild-type isolate PH-1 grown in nutrient-rich medium. Three biologically independent replicates were conducted. In total, 35,103,060 RNA-Seq reads were mapped against the four chromosomes and the expression level of each chromosome was evaluated by using RPKM (reads per kilobase per million mapped reads) values. The averaged expression levels of genes on chromosome 1 are significantly higher than other chromosomes, while the expression levels of genes on chromosome 2 are significantly lower (Figure 2A). Subsequently, we evaluated the expression levels of all predicted genes of *F. graminearum* and divided them into five categories based on their RPKM values: RPKM 0, RPKM 0-1, RPKM 1-10, RPKM 10-100 and RPKM >100. The proportion of each category on each chromosome was calculated. As expected, the proportion of weakly expressed genes (e.g. RPKM <1) is higher on Fgchr 2 (45.1%) than on the other chromosomes (30.2-35.3%) (Figure 2B), while the proportion of highly expressed genes (e.g. RPKM >10) on Fgchr 2 (32.7%) is lower than that on the other chromosomes (44-50%). Box plot analysis also

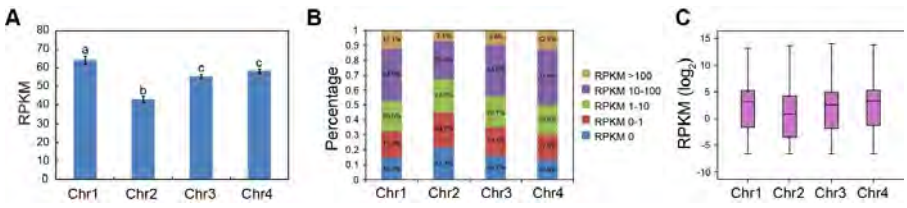


Figure 2. Evaluation of gene expression levels on the four chromosomes of *Fusarium graminearum*. (A) The expression levels of genes on each chromosome were quantified by the number of reads per kilobase per million reads mapped. Genes on chromosome 2 showed relative low expression levels. Different letters represent statistically significant differences (P<0.01, one-way ANOVA). (B) All genes of *F. graminearum* were grouped into five classes according to their RPKM values (0, 0-1, 1-10, 10-100, >100), and the proportion of each class was calculated for each chromosome. (C) Box plot analysis of gene expression on each chromosome. The expression of each gene was quantified by its RPKM value. Log2-transformed RPKM values were used to draw a box plot graph for each chromosome.

showed that relatively more weakly expressed genes (RPKM<1) are located on Fgchr 2 (*p* value=4.04e-39, hypergeometric test) (Figure 2C).

Detailed analysis of chromosome regions with weakly expressed genes in *F. graminearum*

To identify positional differences in gene expression, we divided each chromosome into 20 kb windows and for each window the log2-transformed read coverage was used to quantify the expression level. This analysis demonstrated that on Fgchr 2, two large regions, Nc5 (2.37 Mb - 3.82 Mb) and Nc6 (5.44 Mb - 8.89 Mb) exhibited significantly lower expression levels (Lines B in Figure 3) (*p* value<0.01, Student's *t*-test) when compared with other parts of the same chromosome or with other chromosomes. In addition, four regions on Fgchr 1, three regions on Fgchr 3 and three regions on Fgchr 4 showed significantly lower gene expression levels. These gene expression patterns were observed for each of three biologically independent replicates (Supplemental Figure 2). Subsequently, we determined the number of genes in each of the 20 kb windows. Interestingly the regions with weakly expressed genes showed slightly higher gene density (Lines A in Figure 3). These data demonstrate that the expression of genes is not equally distributed across the chromosomes of *F. graminearum* and that the lower expression is not caused by a lower gene density in these regions.

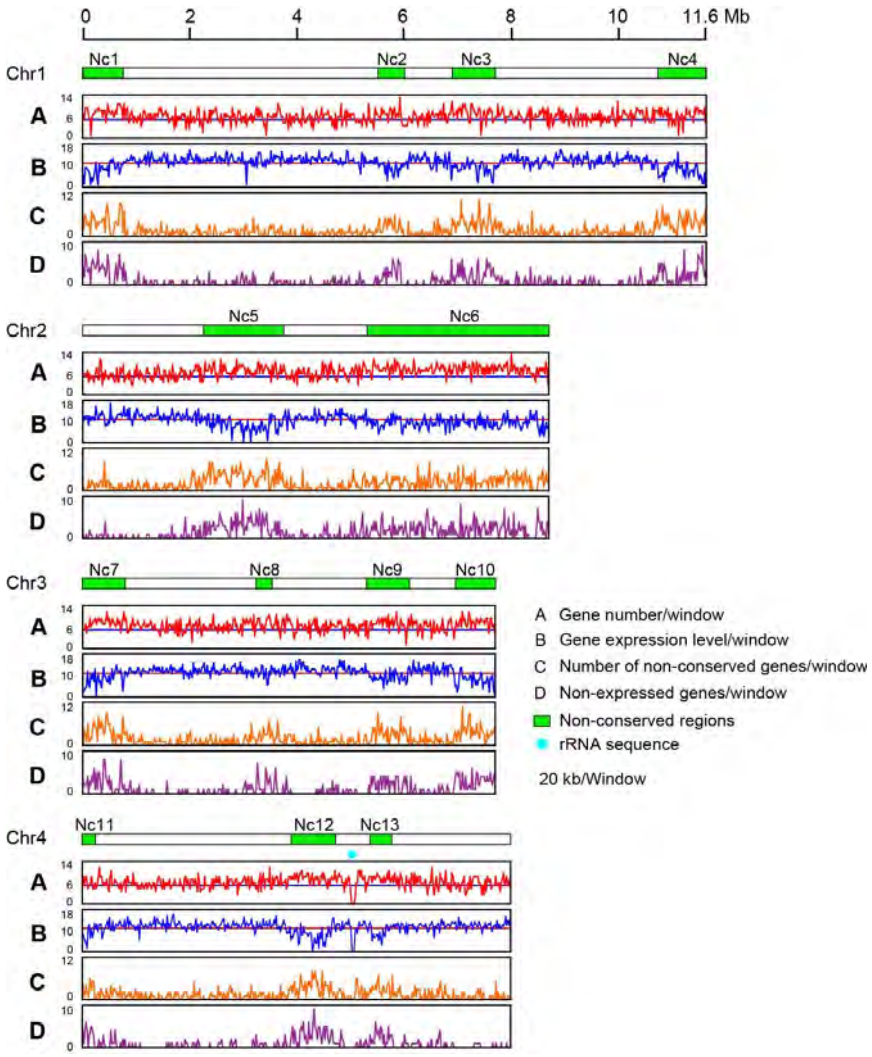


Figure 3. Gene expression patterns across each of the four chromosomes of *Fusarium graminearum*. Each chromosome was divided into 20 kb windows. (A) Number of genes per window. The horizontal blue lines represent gene number of 6. (B) Gene expression levels per window. Log2-transformed read number per window was used to quantify gene expression levels. Twelve regions with low expression levels were identified on the four chromosomes, which matched non-conserved regions (except Nc8). The horizontal red lines represent an expression level of 10. (C) Number of non-conserved genes per window. All genes predicted in *F. graminearum* were matched with the genes of *F. verticillioides* using BLASTn tool and genes without homologs (*p* value>1E-5) were selected. The number of these non-conserved genes per 20 kb window is shown in each chromosome of *F. graminearum*. (D) The number of non-expressed genes in each window of each chromosome is shown. Green boxes represent non-conserved regions in *F. graminearum*. The rDNA region on chromosome 4 is indicated by a blue dot.

Low-expression regions coincide with non-conserved regions

Detailed inspection showed that regions with low levels of expression correspond to non-conserved regions of *F. graminearum* (confer lines B and C in Figure 3). This was further analyzed by comparing genes predicted in the genome of *F. graminearum* to all the genes in *F. verticillioides* using BLASTn. For 9297 genes in *F. graminearum*, conserved genes were found in *F. verticillioides* (*p* value≤1E-5), while for 4024 genes in *F. graminearum*, no conserved ortholog in *F. verticillioides* was identified (*p* value>1E-5). When calculating the number of non-conserved genes in each 20 kb window, we found that non-conserved genes are more abundant in chromosomal regions, where genes are weakly expressed (Lines C in Figure 3). This indicates that the regions with weakly expressed genes coincide with the non-conserved regions in *F. graminearum*. In addition, genes that are not expressed in nutrient-rich medium are enriched in non-conserved regions (Lines D in Figure 3).

To investigate whether non-conserved genes show lower levels of expression, the expression levels of all conserved and non-conserved genes were categorized according to their RPKM values. On each chromosome, the conserved genes showed much higher gene expression levels than non-conserved genes (Figure 4A). Furthermore, we categorized the genes of *F. graminearum* into five groups according to the *p* value of the predicted genes in the BLAST analysis using the predicted genes of *F. verticillioides* as a reference (*p* value 0, 0 - 1E-100, 1E-100 - 1E-50, 1E-50 - 1E-5, >1E-5). Compared to the other three chromosomes, Fgchr 2 contains a higher proportion of non-conserved genes (*p* value>1E-5) (Figure 4B). In addition, a strong correlation between the degree of gene conservation and gene expression level was observed (Figure 4C), as genes with high similarity to their homologs showed relatively higher expression levels, while genes with less similarity to their homologs showed relatively lower expression levels. To specifically demonstrate this phenomenon the expression levels of all 710 transcription factors, previously identified by Son *et al.* (Son et al., 2011) were analyzed. As observed for the expression levels of all genes, the expression levels of genes encoding transcription factors with high similarity to their putative homologs in *F. verticillioides* were higher than genes encoding non-conserved transcription factors (Figure 4D). Furthermore, when each category of genes, grouped by the degree of similarity, was

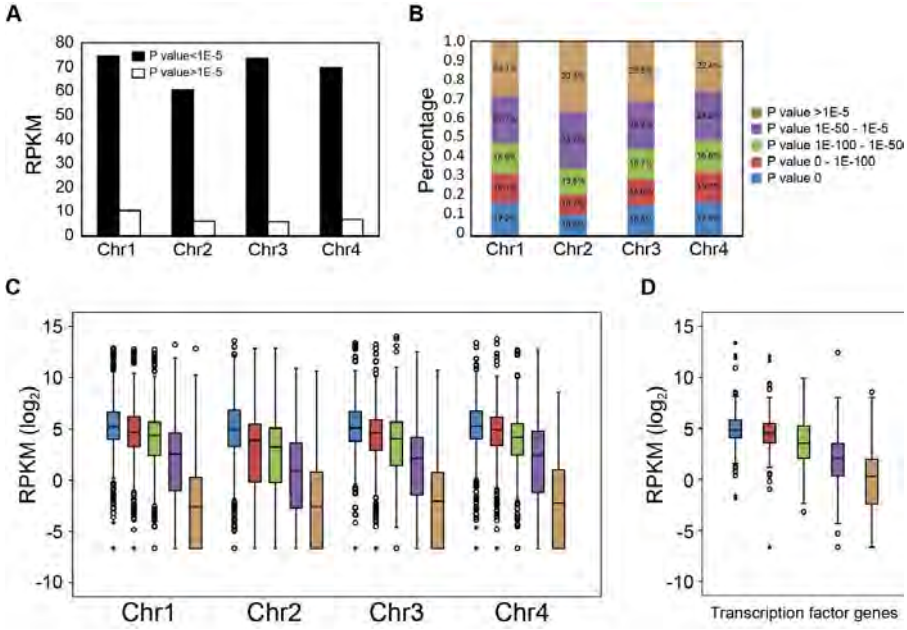


Figure 4. Gene expression levels in relation to their degree of similarity. (A) The expression levels of genes with orthologs (*p* value<1E-5) or genes without orthologs (*p* value>1E-5) in *F. verticillioides* were quantified by their RPKM values. (B) Genes were categorized into five groups according to their degree of similarity with those of *F. verticillioides* (*p* value 0, 0 - 1E-100, 1E-100 - 1E-50, 1E-50 - 1E-5 and >1E-5). The proportion of each category of genes on each chromosome was calculated. (C) Box plot analysis of the expression of genes on each chromosome according to their degree of similarity with those of *F. verticillioides*. Log2-transformed RPKM values were used to evaluate gene expression levels. (D) Box plot analysis of the expression of 710 transcription factor genes (as identified by Son *et al*., 2011) according to the degree of similarity with those of *F. verticillioides*. Log2-transformed RPKM values were used to evaluate gene expression levels.

mapped on the four chromosomes of *F. graminearum*, a clear differential gene distribution pattern was observed. Conserved genes tend to cluster together as well as less conserved genes (Supplemental Figure 3). Strikingly, no conserved genes (*p* value=0) are located at one of the terminal regions of Fgchr 3.

Expression of genes in non-conserved regions is highly variable in different developmental stages

To investigate whether gene expression in non-conserved regions is invariably low, or whether it is related to a specific growth stage, RNA-Seq analysis was performed on conidia

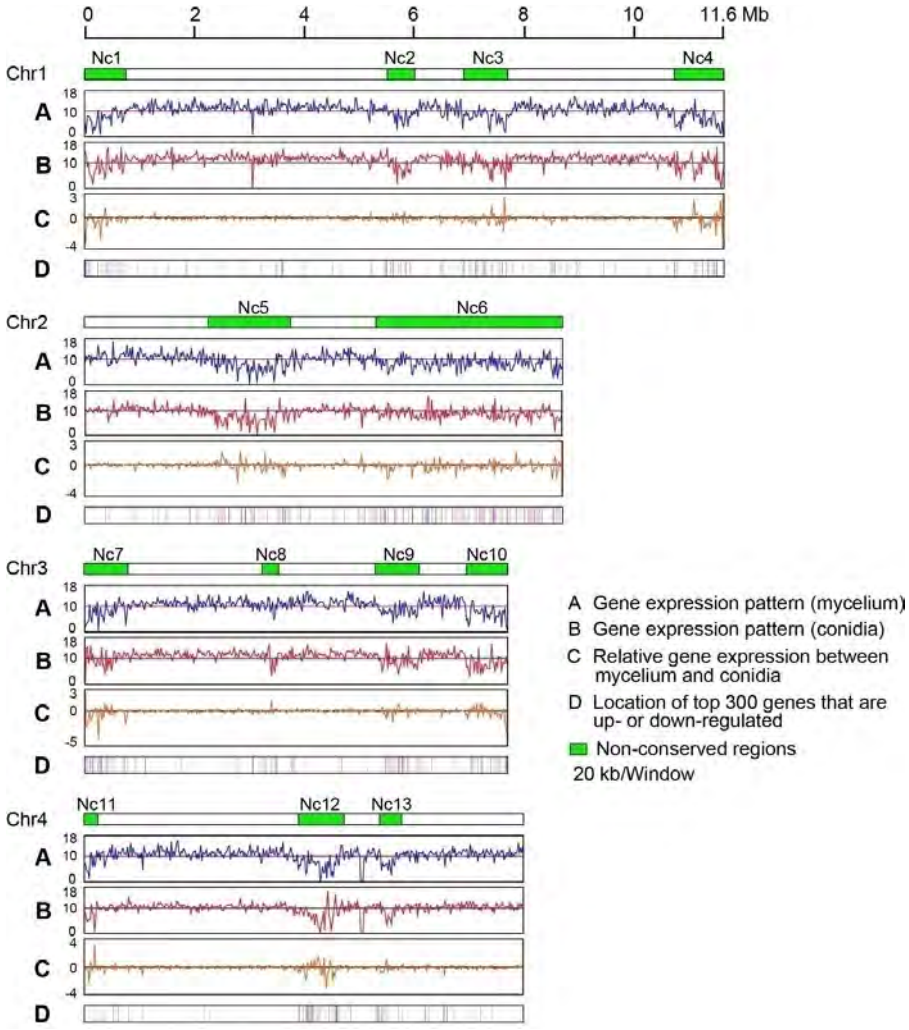


Figure 5. Comparison of gene expression patterns between mycelium and conidia on the chromosomes of *Fusarium graminearum*. (A) Each chromosome was divided into 20 kb windows. Log2-transformed read coverage per window was used to quantify gene expression levels in mycelium of *F. graminearum*. (B) Similarly, the Log2-transformed read coverage per window was used to quantify gene expression levels in conidia of *F. graminearum*. (C) The ratio of the log2-transformed RPKM values per window between mycelium and conidia. (D) The locations of the 300 genes that show the strongest up-regulation comparing mycelium to conidia (red lines) and the 300 genes that show the strongest down-regulated (blue lines) on each chromosome of *F. graminearum*. Green boxes represent non-conserved regions in *F. graminearum*.

of *F. graminearum* isolate PH-1 produced in mung bean medium. Three biologically independent replicates were analyzed and in total 40,615,406 reads were obtained. Gene

expression patterns on each chromosome of *F. graminearum* were analyzed with the obtained RNA-Seq data. Firstly, similar gene expression patterns on each chromosome of *F. graminearum* were observed for the three biological replicates (Supplemental Figure 4). Secondly, also in conidia nearly all non-conserved regions exhibited significantly lower gene expression levels than conserved regions (p value<0.01, Student's t-test). However, some parts of non-conserved regions, especially in regions Nc4 and Nc7 showed expression levels similar to conserved regions (lines B in Figure 5). In addition, we compared the gene expression levels between conidia and mycelium in each 20 kb window on each chromosome. Strikingly, the relative expression (e.g. expression in mycelium/expression in conidia) of genes located in conserved regions is relatively constant, while the relative expression of genes in non-conserved regions is sometimes highly variable (lines C in Figure 5). Furthermore, we selected the 300 genes that exhibit the strongest up- and down-regulation between the two different developmental stages and mapped their locations on each chromosome of *F. graminearum*. We found that these highly induced or repressed genes are significantly abundant in non-conserved regions (p value=2.57e-49, hypergeometric test) (lines D in Figure 5). These data suggest that gene expression in non-conserved regions might be important for *F. graminearum* to respond to external stimuli or during specific phases of its life cycle, while gene expression in conserved regions are more stable and are more likely to be involved in housekeeping functions.

Non-conserved regions are enriched for genes encoding secreted proteins and enzymes required for the production of secondary metabolites

To further understand the biological relevance of non-conserved regions, we examined which categories of genes are enriched in the non-conserved regions. Firstly, we studied the distribution of all genes encoding secreted proteins in *F. graminearum*. Previously, Brown *et al*. [16] identified 574 genes encoding secreted proteins in the *F. graminearum* genome (Brown et al., 2012). Distribution of these genes on each chromosome of *F. graminearum* showed that 384 locate in non-conserved regions and 190 in conserved regions (Figure 6A), suggesting that they are significantly enriched in non-conserved regions (p value=4.32e-69, hypergeometric test). Moreover, a higher proportion of genes encoding secreted proteins were

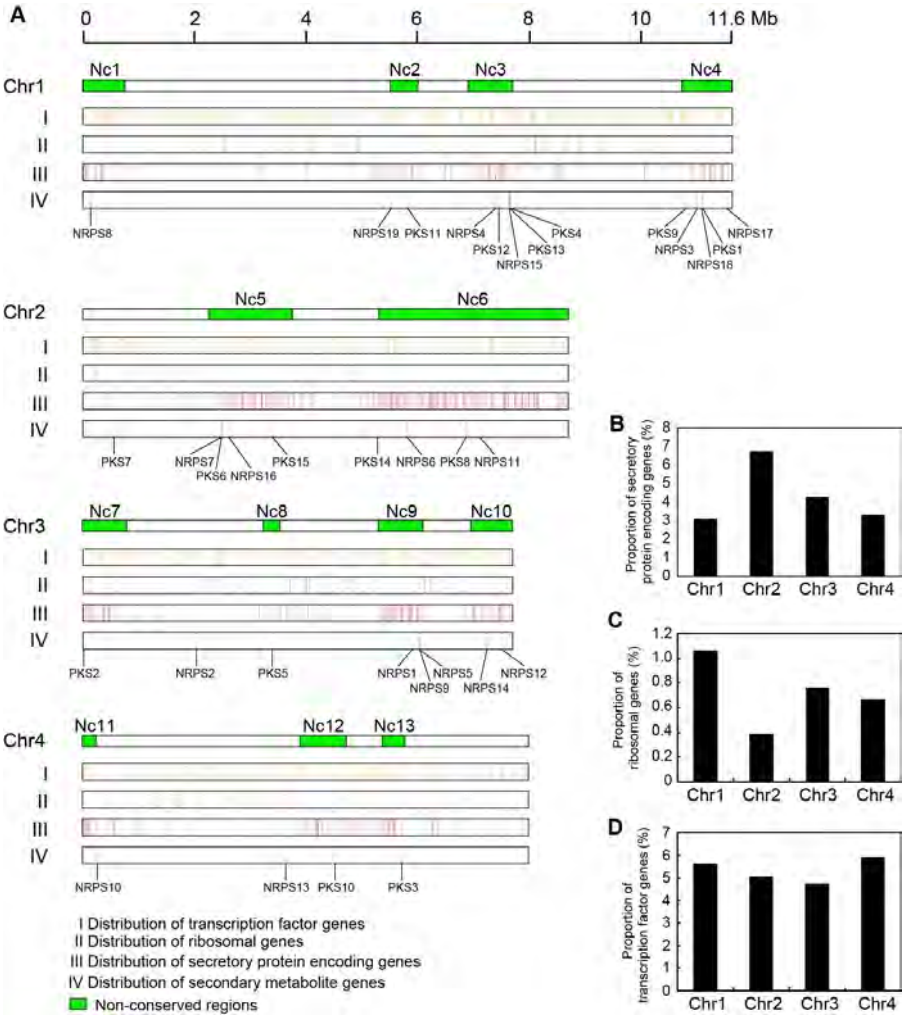


Figure 6. Distribution of different categories of genes on each chromosome of *F. graminearum*. (A) 710 transcription factor genes, 547 genes encoding secreted proteins, 98 ribosomal genes and 34 genes encoding enzymes putatively involved in the production of secondary metabolites are mapped on the chromosomes. Genes encoding secreted proteins and genes encoding enzymes for production of secondary metabolites are enriched in weakly expressed regions, ribosomal genes are predominantly found in conserved regions, while transcription factor genes are randomly distributed. (B) Proportion of genes encoding secreted proteins on each chromosome of *F. graminearum*. (C) Proportion of ribosomal protein genes on each chromosome of *F. graminearum*. (D) Proportion of transcription factor encoding genes on each chromosome of *F. graminearum*. Green boxes represent non-conserved regions identified in *F. graminearum*.

observed on chromosome 2 (*p* value=5.77e-15, hypergeometric test) (Figure 6B). Secondly,

the distribution of secondary metabolite genes, including 15 polyketide synthase (PKS) genes and 19 non-ribosomal peptide (NRPS) genes (Hansen et al., 2012) also appeared to be unevenly distributed: 14 PKS genes and 17 NRPS genes locate in non-conserved regions, while only one PKS gene and two NRPS genes were found in conserved regions. This indicates that genes involved in secondary metabolism are also enriched in non-conserved regions. (*p* value=3.09e-9, hypergeometric test). In contrast, ribosomal genes that are essential for growth are enriched in conserved regions (*p* value=2.56e-14, hypergeometric test) and fewer ribosomal genes are identified on chromosome 2 of *F. graminearum* (Figure 6A, C). In contrast, transcription factor-encoding genes were randomly distributed on each chromosome of *F. graminearum* (Figure 6A, D) (*p* value=0.03, hypergeometric test).

Non-conserved regions are enriched for relocated genes

As mentioned above, 9297 genes in *F. graminearum* have homologs in *F. verticillioides*. To further demonstrate that genes with their putative homologs in *F. verticillioides* are likely to be true orthologs, the homologs in *F. verticillioides* were aligned against the genes in *F. graminearum* using BLASTn and we found that 9208 gene pairs are the reciprocal best blast hits, suggesting that they are orthologs. These genes can be further divided into two groups based on their collinearity with their orthologs. Most *F. graminearum* genes have their putative orthologs on one of the collinear chromosomes of *F. verticillioides*, while some genes have their putative orthologs located on non-collinear chromosomes of *F. verticillioides*. The number of genes matching to putative orthologs on each chromosome of *F. verticillioides* is presented in Table 1. For example, on Fgchr 1, 2794 genes have a putative ortholog on the collinear chromosomes (1, 5, 8) of *F. verticillioides*, whereas for 187 genes, putative orthologs were shown on non-collinear chromosomes of *F. verticillioides*. We then mapped all these genes on the chromosomes of *F. graminearum* and found that non-conserved regions are highly enriched for these genes (Figure 7A). This indicates that non-conserved regions in *F. graminearum* are enriched in genes that have undergone relocation. A similar pattern was observed when comparing *F. graminearum* with *F. oxysporum* (Figure 7B). Further evidence for these relocations comes from the fact that in 355 (80.5%) of these relocations a single gene was relocated , but in 46 (10.4%) cases two neighbouring genes and in 40 (9.1%) cases

Table 1 Number of genes on each chromosome of *Fusarium graminearum* that have orthologs on each chromosome of *F. verticillioides*

	Fvchr1	Fvchr2	Fvchr3	Fvchr4	Fvchr5	Fvchr6	Fvchr7	Fvchr8	Fvchr9	Fvchr10	Fvchr11
Fgchr1	**1478**	15	24	74	**958**	17	13	**358**	12	19	13
Fgchr2	31	26	17	35	17	**772**	12	27	**568**	**333**	**281**
Fgchr3	24	**485**	14	**681**	18	7	**623**	72	17	15	20
Fgchr4	8	**632**	**1034**	**152**	8	11	103	68	20	14	18

Values in bold type are the numbers of genes that have orthologs on collinear chromosomes of *F. verticillioides*, while values in regular type are the numbers of genes that have orthologs on non-collinear chromosomes of *F. verticillioides*.

three or more neighbouring genes were relocated together. Figure 7C shows one example of a relocation of four neighbouring genes in *F. graminearum*. On Fvchr1 four genes were found between genes FVEG_01045 and FVEG_01040, whereas no genes were identified between orthologs FGSG_00567 and FGSG_00568 in *F. graminearum*, indicating that these four genes have been relocated to a new location in *F. graminearum*. BLAST analysis of the four *F. verticillioides* genes showed that their orthologs in *F. graminearum* are relocated to the non-conserved region Nc9 of Fgchr 3.

To understand whether genes that are relocated to non-conserved regions are also affected in expression, we compared the expression levels of relocated genes with non-relocated genes. This analysis indicated that expression of relocated genes was significantly lower than non-relocated genes under the conditions examined (confer A and B in Supplemental Figure 5).

Secondary metabolite gene clusters are assembled in non-conserved regions by gene relocations

As shown above, most of secondary metabolite genes are located in non-conserved regions, and previous studies have shown that genes producing secondary metabolites are often clustered [17], but whether these secondary metabolite gene clusters are assembled by gene relocations is still unknown. Here we found that gene clusters, such as *PKS1*, *PKS2*, *PKS6*, *PKS8* and *PKS10*, are probably formed by gene relocations, as genes in these clusters have orthologs that are located on different chromosomes of *F. verticillioides*. This is exemplified by the PKS2 gene cluster of which most genes have high similarity to their orthologs (Table 2). The ortholog of PKS2 gene is located on chromosome 1 of *F.*

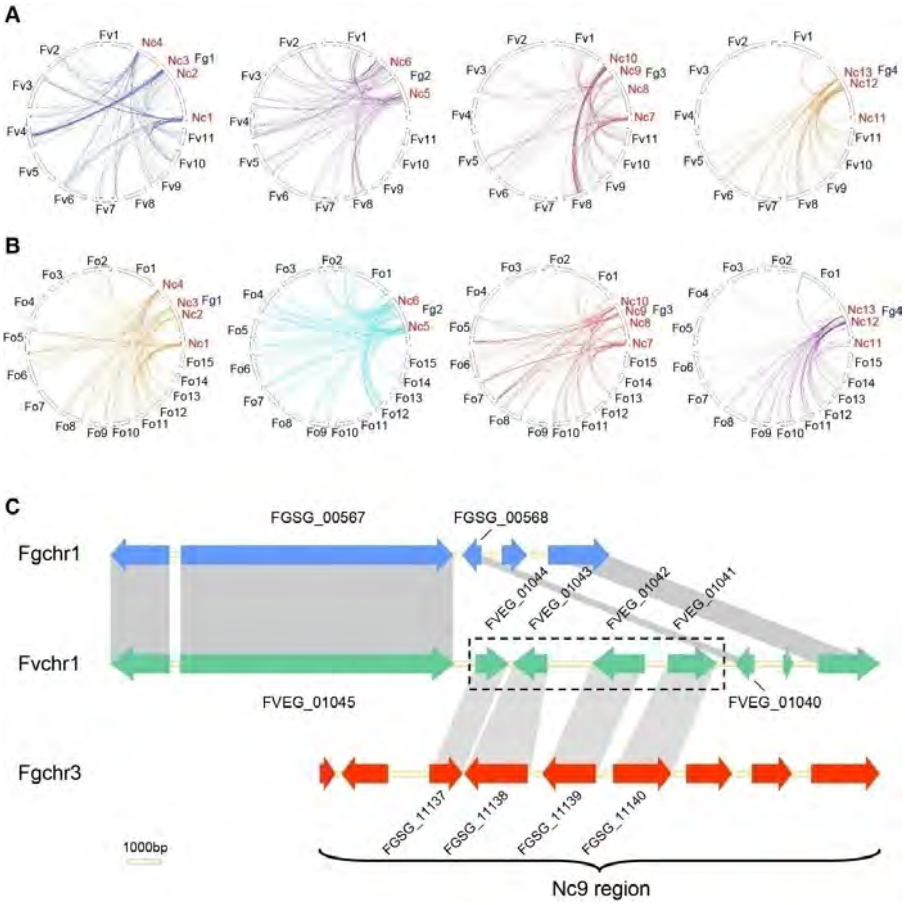


Figure 7. Frequent gene relocations in non-conserved regions of the four chromosomes of *Fusarium graminearum*. (A-B) Genes matching their orthologs on non-collinear chromosomes of *F. verticillioides* (A) or *F. oxysporum* (B) are present on each of the four chromosomes of *F. graminearum*. (C) One representative example showing relocation of four genes to region Nc9. Orthologs of four *F. verticillioides* genes (FVEG_01044, FVEG_01043, FVEG_01042 and FVEG_01041) were not identified in the collinear chromosomal region, but were found in region Nc9 of *F. graminearum*.

verticillioides. Other genes that are typically associated with secondary metabolite gene clusters, such as MFS monosaccharide transporter, NAD binding oxidoreductase, alcohol dehydrogenase, cytochrome P450 family protein and C6 transcription factor, have their orthologs located on the chromosomes 8, 9, 10 and 11 of *F. verticillioides*, respectively (Figure 8 and Table 2). This result suggests that secondary metabolite gene clusters could be formed by gene relocations as was found for other genes in the non-conserved regions.

Table 2 Analysis of PKS2 secondary metabolite gene cluster

Num	Gene ID	Putative function	Orthologs in *Fusarium verticillioides*		
			Gene ID	Chromosome	*p* value
1	FGSG_12582	Sterol 3-beta-glucosyltransferase	FVEG_00073	Fvchr 1	0
2	FGSG_12583	Putative bile acid 7-alpha protein	FVEG_00076	Fvchr 1	5.51E-155
3	FGSG_04692	Acyltransferase	FVEG_00077	Fvchr 1	0
4	FGSG_04693	Integral membrane protein PTH11	FVEG_00078	Fvchr 1	0
5	FGSG_04694 (PKS2)	Polyketide synthase	FVEG_00079	Fvchr 1	0
6	FGSG_04695	Hypothetical protein	-	-	-
7	FGSG_04696	Hypothetical protein	-	-	-
8	FGSG_04697	Hypothetical protein	FVEG_13993	Sc26	0
9	FGSG_04698	Hypothetical protein	-	-	-
10	FGSG_04699	Hypothetical protein	-	-	-
11	FGSG_04700	MFS monosaccharide transporter	FVEG_10988	Fvchr 9	9.16E-83
12	FGSG_04701	ARCA-like protein	-	-	-
13	FGSG_04702	NAD binding oxidoreductase	FVEG_10989	Fvchr 9	1.58E-55
14	FGSG_04703	Alcohol dehydrogenase	FVEG_13775	Fvchr 10	7.36E-61
15	FGSG_04704	Cytochrome P450 family protein	FVEG_07686	Fvchr 8	0
16	FGSG_04705	RTA1 like protein	-	-	-
17	FGSG_04706	RTA1 domain-containing protein	-	-	-
18	FGSG_12584	C6 transcription factor	FVEG_10599	Fvchr 11	6.25E-52

"-" represents that no ortholog was identified in *F. verticillioides*.

The origin of lineage-specific chromosomes in *F. oxysporum*

F. oxysporum f. sp. *lycopercisi* isolate 4287 contains eleven core chromosomes and four lineage-specific (LS) chromosomes (Ma et al., 2010). Comparison of core chromosomes of *F. oxysporum* with those of *F. verticillioides* through synteny analysis revealed several non-conserved regions (Supplemental Figure 1), especially at the end of chromosomes 1 and 2. To determine whether the four LS chromosomes and non-conserved regions on core chromosomes of *F. oxysporum* are enriched for gene relocations, we selected all genes of *F. oxysporum* with homologs on a non-collinear chromosome of *F. verticillioides* and assigned these genes to each chromosome of *F. oxysporum*. The four LS chromosomes appear to be enriched for genes that have relocated from core chromosomes (Supplemental Figure 6). Gene relocations were also supported by the observation that multiple groups of adjacent genes on LS chromosomes showed collinearity with their homologs on the chromosomes of *F. verticillioides* (Supplemental Table 1). In addition, relocation of genes was also observed in non-conserved regions of core chromosomes of *F. oxysporum*, especially in one telomere

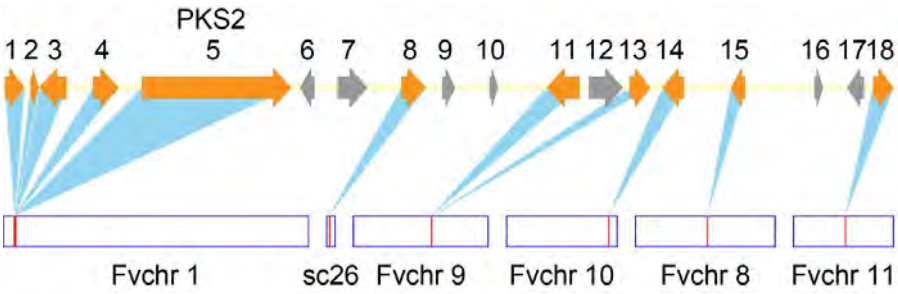


Figure 8. *PKS2* gene cluster is assembled in a non-conserved region by gene relocations. The *PKS2* gene and its flanking genes are matched with their orthologs in *F. verticillioides*. The numbers represent the genes of *F. graminearum* that are listed in Table 2. Orange colour represents genes that have orthologs, while grey colour indicates genes that have no clear orthologs in *F. verticillioides*. The boxes below represent the chromosomes of *F. verticillioides* and also the locations of orthologs on each chromosome. Sc26 is an unmapped supercontig in *F. verticillioides*.

proximal region of chromosomes 1 and 2. In addition, synteny block analysis of chromosomes in *F. oxysporum* was performed and we found that most of the genes present on LS chromosomes 3 and 6 are duplicated, which is consistent with previous findings [9]. Some genes on LS chromosome 14 might originate from LS chromosomes 3 or 6. Interestingly, genes on LS chromosome 15 are duplicated from core chromosome 1 (Supplemental Figure 7), suggesting that LS chromosome 15 originally arose from core chromosome.

Discussion

Genome comparison between closely related species may help us to understand the mechanism of genome evolution and apprehend how species evolve to adapt to new environments (Chimpanzee and Analysis, 2005; Paterson et al., 2010; Scally et al., 2012). Comparing the genome of *Homo sapiens* with the closely related species chimpanzee showed that 98.76% of the genomic sequences are similar (Navarro and Barton, 2003), but 1,576 putative inversions were identified on the chromosomes of *H. sapiens* (Feuk et al., 2005) that are likely to be involved in genome evolution. Large-scale translocations and inversions appear to have also occurred in the ascomycete *Saccharomyces cerevisiae* (Gordon et al., 2011; Wei et al., 2007). Here, we compared the genome of *F. graminearum* with that of two closely related species *F. verticillioides* and *F. oxysporum*. Again, a large number of translocations and inversions were identified. Next to these translocations and inversions of

large chromosomal segments, non-conserved regions are commonly discovered in closely related species. For instance, synteny analysis of the genome of *Aspergillus nidulans* with related species *A. fumigatus* and *A. oryzae* showed that around 78% of the genome could be mapped to conserved syntenic blocks, while the remaining genomic sequences lack significant syntenic blocks (Galagan et al., 2005). In addition, syntenic analysis of 12 sequenced *Drosophila* species showed that on average 66% of each genome was covered by syntenic blocks (Drosophila 12 Genomes et al., 2007), indicating that the location of the remaining 34% of the genome is not conserved. Comparing the genome of *F. graminearum* with those of *F. verticillioides* and *F. oxysporum*, thirteen non-conserved regions were identified. The presence of these non-conserved regions suggests that genome evolution has occurred unevenly across the chromosomes. Although non-conserved regions frequently occur in the genomes of many species, their origin and biological relevance are still largely unknown.

The development of RNA-Seq technology (Wang et al., 2009) allowed us to evaluate the global gene expression patterns along chromosomes (Filichkin et al., 2010; Nagalakshmi et al., 2008). In this study, we used RNA-Seq data obtained from the mycelium of the sequenced isolate PH-1 of *F. graminearum* grown in nutrient-rich medium to investigate the gene expression pattern along whole chromosomes. Interestingly, we found two striking features in *F. graminearum*: (i) there is a strong correlation between the degree of gene conservation and gene expression level and (ii) genes in the non-conserved regions showed lower expression levels than genes in the conserved regions. In addition, comparing gene expression levels between conidia and mycelium, we found that the expression of genes in non-conserved regions is highly variable, while the expressions of genes in conserved regions is surprisingly stable. This indicates the expression of genes in conserved regions is less dependent on fungal development (e.g. in conidia *vs.* mycelium), while the expression of genes in non-conserved regions seems more developmentally regulated. Furthermore, house-keeping genes are more abundant in conserved regions, while genes required in specific developmental stages or environmental conditions are found more often in non-conserved regions. This conclusion was supported by the fact that genes encoding secreted proteins or involved in the production of secondary metabolites, which are induced

under specific conditions (Cuomo et al., 2007; Gaffoor et al., 2005; Kim et al., 2005; Voigt et al., 2005), are more abundant in non-conserved regions, while ribosomal genes are mainly located in conserved regions.

Gene enrichment in specific chromosomal regions has also been studied in other species. For instance, secondary metabolite genes are enriched in subtelomeric regions of *Aspergillus* species. The rapid rearrangement of subtelomeric regions may promote the rapid evolution of these genes to become species-specific attributes (Galagan et al., 2005; Perrin et al., 2007). In the plant pathogenic fungus *Verticillium dahliae*, *in planta*-expressed genes are enriched in lineage-specific genomic regions that have developed by extensive chromosomal reshuffling, which was suggested to drive evolution of virulence (de Jonge et al., 2013). In *F. oxysporum*, genes encoding secreted effectors and virulence factors are more abundant in LS chromosomes, while few house-keeping genes are identified on LS chromosomes (Ma et al., 2010). The drivers and biological relevance of gene enrichment in specific chromosome regions are still not fully understood. One possible reason could be that clustering of genes with similar expression patterns may facilitate co-regulation of specific sets of genes under specific conditions or developmental stages. This phenomenon has been observed in many species, such as yeast, mouse and human (Cohen et al., 2000; Hurst et al., 2004; Su et al., 2004; Woo et al., 2010). Secondly, changes in gene expression levels have also been shown to be important in adaptive evolution (Brawand et al., 2011; Fraser et al., 2010), so perhaps enrichment of genes in non-conserved regions could facilitate them to rapidly change their expression pattern.

Our data indicate that genes in non-conserved regions are weakly expressed, but how the expression of these genes is suppressed is still unknown. Previous studies have shown that gene position in the nucleus is associated with their transcriptional regulation, for instance, the nuclear periphery was considered as a zone for transcriptionally repressed genes (Sexton et al., 2007). This type of organization could also occur in *F. graminearum* and non-conserved regions could form specific sub-compartments of the nucleus with repressed gene expression.

The non-conserved regions occur widely, but how they are generated is still unclear. In this study, many gene relocations were identified in non-conserved regions of *F. graminearum*, in contrast to conserved regions. Gene relocations have been described previously in *S.*

cerevisiae where the *DAL* gene cluster, including six genes, was formed by gene relocations from six different loci (Wong and Wolfe, 2005). In *F. graminearum* and *F. sporotrichioides*, trichothecene biosynthesis requires genes at three loci: the 12-gene *TRI* cluster, a second locus with two genes (*TR1* and *TRI16*), and a third locus with one gene (*TRI101*). However, in the more distantly related species *F. equiseti*, both *TRI1* and *TRI101* are located within the *TRI* core cluster suggesting that the latter two genes have been relocated during evolution of *F. equiseti* (Proctor et al., 2009). We also showed that four successive genes were relocated to a non-conserved region in *F. graminearum*. However, how these genes are marked for relocation is still obscure. Based on our findings that genes with low similarity to their homologs have low expression levels (refer to Figure 4C), we hypothesize that the relocated genes in non-conserved chromosome regions already had a low expression level before they were relocated. Possibly, in the three dimensional organization of chromosomes in the interphase these weakly expressed genes are in close proximity with the likewise weakly expressed non-conserved regions. Such a close proximity might facilitate targeted relocation of weakly expressed genes to one of these non-conserved regions.

In *F. oxysporum*, four LS chromosomes and 11 core chromosomes were identified. It was suggested that these four LS chromosomes were generated by horizontal transfer from a yet unknown fungal source (Ma et al., 2010). Also for the dispensable chromosomes in *Zymoseptoria tritici* horizontal gene transfer was hypothesized (Goodwin et al., 2011). Although horizontal gene transfer is one of the reasonable explanations for the origin of LS chromosomes in *F. oxysporum* (Ma et al., 2010), other mechanisms or events could also contribute to the formation of LS regions. In this study, the frequent gene relocations identified in non-conserved regions of *F. graminearum* drove us to hypothesize that the generation of LS chromosomes in *F. oxysporum* could also have included extensive gene relocations. This hypothesis was supported by the fact that around 700 genes on LS chromosomes have a homolog in *F. verticillioides*, and in 22 cases, two or more adjacent genes are collinear with their homologs. In addition, large part of LS chromosome 15 represents a duplication of the telomere proximal region of core chromosome 1, suggesting that it might originate from the core chromosome. Based on the studies by us and others [9], we proposed that most likely horizontal gene transfer from other fungal species and gene

relocations within the species both occurred in *F. oxysporum*.

Conclusions

Taken together, our data demonstrate that chromosomes of *F. graminearum* show distinct conserved and non-conserved regions. Subsequent gene expression analysis showed that genes in these regions exhibit different expression patterns. Genes showing high and stable expression levels are more abundant in conserved regions, while genes that are induced or repressed in specific developmental stages or under different environmental conditions (mung bean medium versus liquid CM) are significantly abundant in non-conserved regions. This type of genome arrangement may not only facilitate the co-regulation of specific sets of genes, but could also enable fungi to maintain on the one hand the required conservation of house-keeping genes and on the other to accelerate the evolution of species-specific genes to rapidly adapt to new environments or new hosts. Moreover, due to the selective transcription of genes in non-conserved regions, this could prevent organisms spending too much energy in transcription and translation of evolving genes that do not have acquired full functionality yet.

Methods

RNA isolation and RNA-Seq

Fusarium graminearum wild-type isolate PH-1 was used in this study. To prepare conidia and mycelium for RNA isolation, PH-1 was grown in 400 ml liquid mung bean medium for 3 days to produce conidia (25 °C, 200 rpm). The conidia were collected by centrifugation. 10e5 conidia of PH-1 were transferred to 400 ml liquid complete medium and subsequently incubated for 30 h to produce mycelium (25 °C, 200 rpm). Mycelium was harvested from liquid CM medium by filtration and grounded in liquid nitrogen using a mortar and pestle. The conidia and mycelium were used for RNA extraction using TRIzol reagent (Invitrogen, Cat. No. 15596-018) according to manufacturer's instructions. The quality of RNA was analyzed by Agilent 2100. RNA-Seq was performed according to protocols described previously (Zhao et al., 2013).

Analysis of gene homology

Gene homology was evaluated in CLC genomic workbench. The gene database of *F. graminearum*, *F. verticillioides* and *F. oxysporum* were downloaded from Broad Fusarium Comparative Database. The gene database (fasta file format) of each *Fusarium* species were imported into CLC, and "BLASTn" option was used to align the genes of *F. graminearum* against the genes of *F. verticillioides* or *F. oxysporum* using default settings. Genes of *F. graminearum* were grouped according to their *p* value after aligning with the genes of *F. verticillioides* by using BLASTn.

The identification of orthologs between *Fusarium* species was based on two criteria: (i) a cutoff of *p* value of 1E-5, and (ii) reciprocal best blast hits. For the genes in conserved chromosomal regions, synteny block analysis was also explored for the definition of orthologous relationship.

Synteny block analysis

Synteny block analysis was performed according to the program *MCScanX* with a small modification (Wang et al., 2012). The gene data sets (fasta format file) of *F. graminearum*, *F. verticillioides* and *F. oxysporum* were downloaded from the Broad *Fusarium* database. The local blast database pools of *F. verticillioides* and *F. oxysporum* were created by using program formatdb. All genes in *F. graminearum* were analyzed against the gene database pool of *F. verticillioides* and *F. oxysporum*, respectively, by using BLAST tool. The BLAST results were exported in m8 format. Besides, gff file containing the information of the chromosome number (e.g Fg1), gene name (e.g FGSG_00001), start position and stop position of each gene on each chromosome of *Fusarium* species was prepared. The BLAST file and gff file were imported for synteny block analysis according to the procedure described in the manual of *MCScanX*. The criteria for the synteny analysis are as follow: match score 50, match size > 5, gap_penalty -1, overlap_window 5, max gaps 25. Finally, two types of output (dual synteny plot and circle plot) were obtained by using two downstream programs.

Analysis of relocated genes

According to the BLASTn result, all *F. graminearum* genes with high similarity (*p* value≤1E-5) in *F. verticillioides* or *F. oxysporum* were collected, from which we selected

genes that have their putative orthologs on non-collinear chromosomes of *F. verticillioides* or *F. oxysporum*. The map of these genes to their orthologs was performed by using *MCScanX* software (Wang et al., 2012). Furthermore, the sequences of all these putative orthologs from either *F. verticillioides* or *F. oxysporum* were collected and matched with all predicated genes in *F. graminearum* using BLASTn tool to identify the best hits to show that they are the bidirectional best hits.

Gene expression analysis

To evaluate gene expression along chromosomes, RNA-Seq reads were mapped to chromosome sequences of *F. graminearum* using software available in the CLC Genomics Workbench. The RNA-Seq reads were mapped to each chromosome by using "RNA-Seq analysis" option with default settings. The number of reads matched to each chromosome was calculated and subsequently the expression level of each chromosome was evaluated by using RPKM (reads per kilobase per million mapped reads) values. Similarly, to evaluate the expression of each gene, the transcript database of *F. graminearum* were imported in CLC and the expression level of each gene was evaluated by RPKM value.

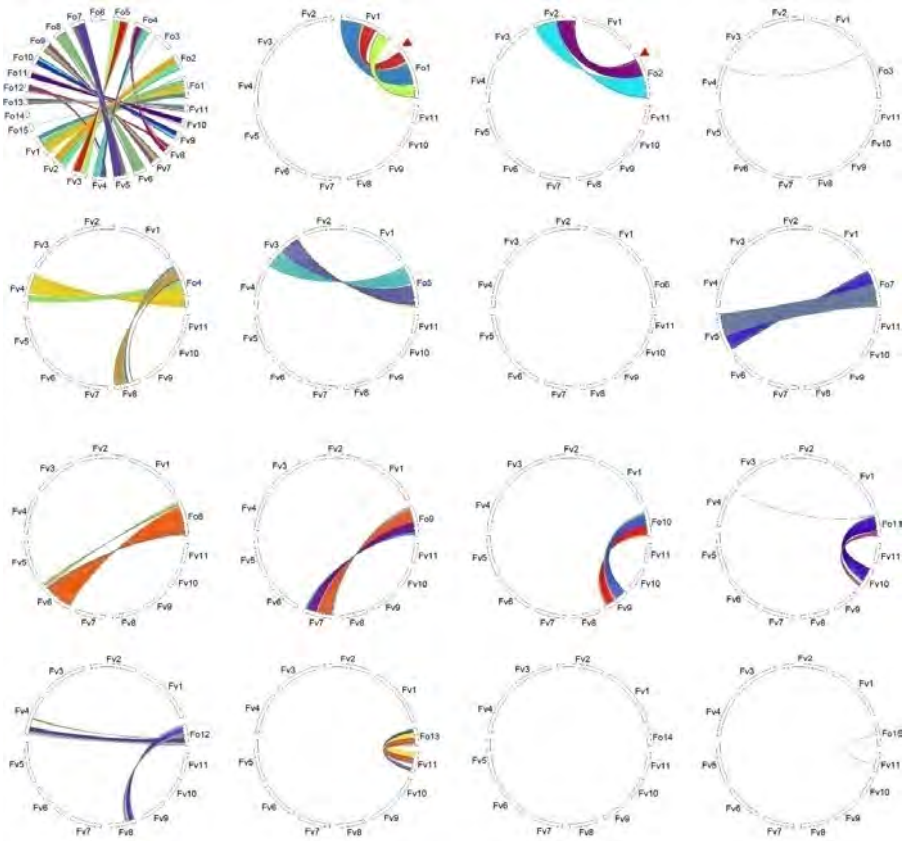
To draw the gene expression level along each chromosome, we divided the chromosomes into portions of 20 kb. The read coverage of each 20 kb window was calculated and log2-transformed reads coverage in each window was used to compare gene expression levels. To compare the gene expression level between conidia and mycelium, reads coverage of each window was compared and log2-transformed reads coverage fold change was used to evaluate gene expression differences. The total gene number and the number of non-conserved genes in each 20 kb window were calculated manually based on the criteria described above.

To analyze the expression levels of conserved and non-conserved genes, the transcript sequences of conserved and non-conserved genes were collected and assembled, respectively. Also in this case the PRKM value was used to evaluate the expression levels of conserved and non-conserved genes. Box plot analysis of gene expression was performed by using SPSS software.

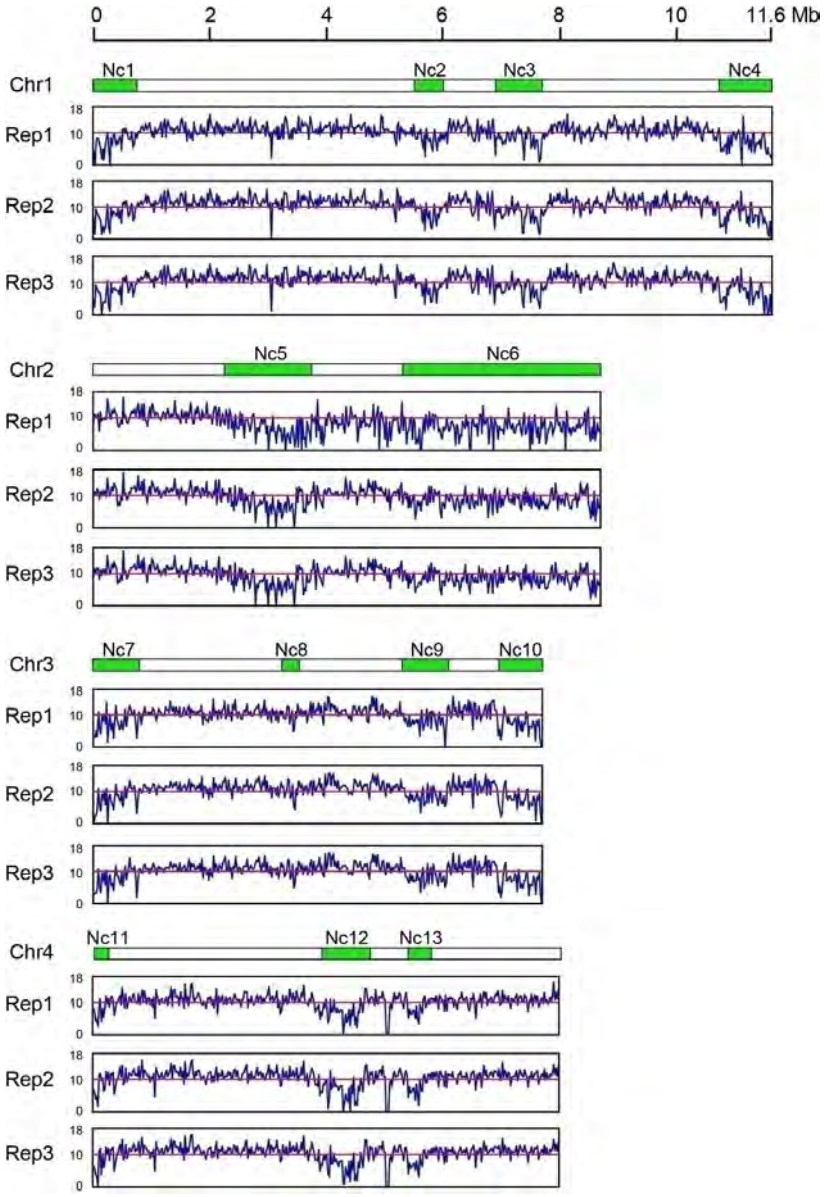
Acknowledgements

This work was supported by grants from the National Basic Research Program of China (2011CB100700) and Chinese Academy of Sciences (SAJC201305). P. J. G. M. de Wit is supported by grants from the Royal Netherlands Academy of Arts and Sciences, the Centre for BioSystems Genomics. We would like to thank Dr. Jan-Peter Nap (Applied Bioinformatics PRT WUR) for stimulating discussion.

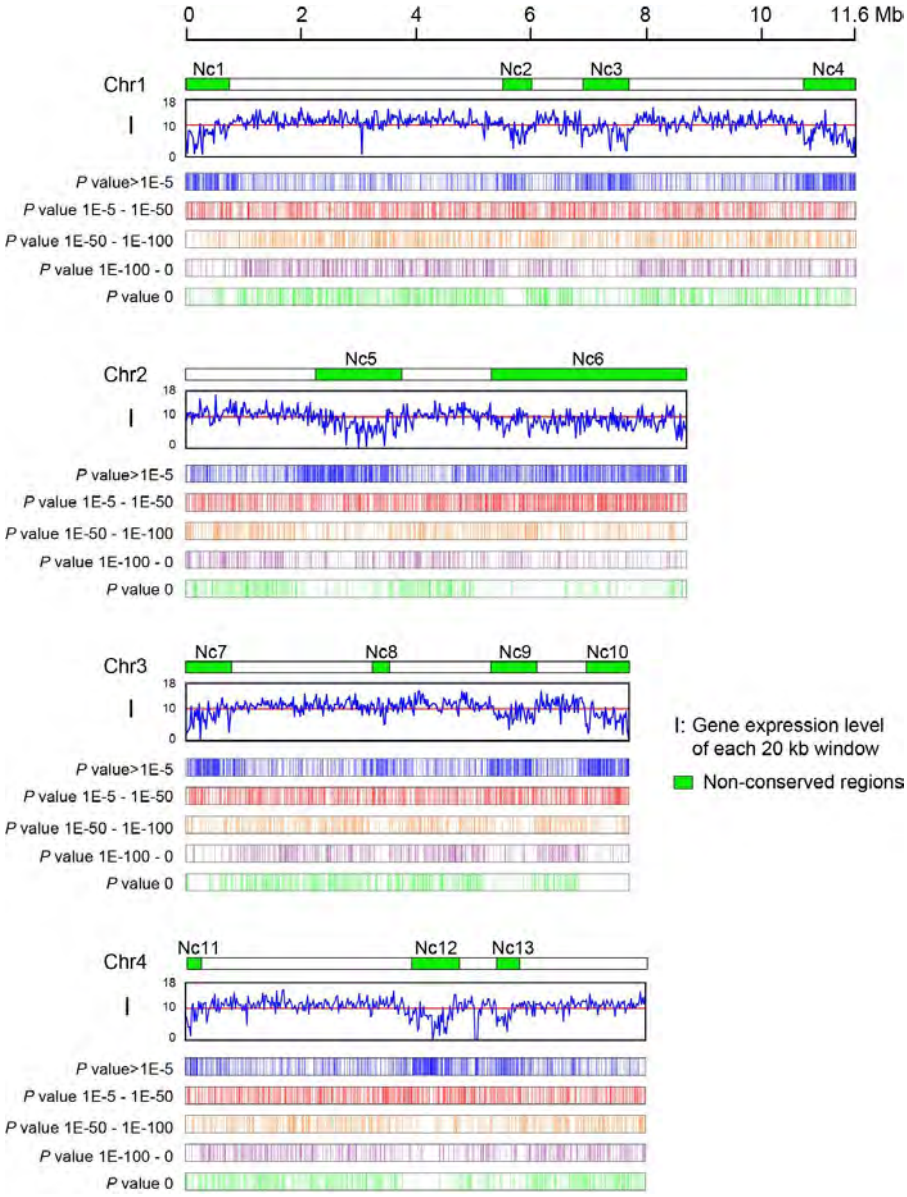
Supplementary data



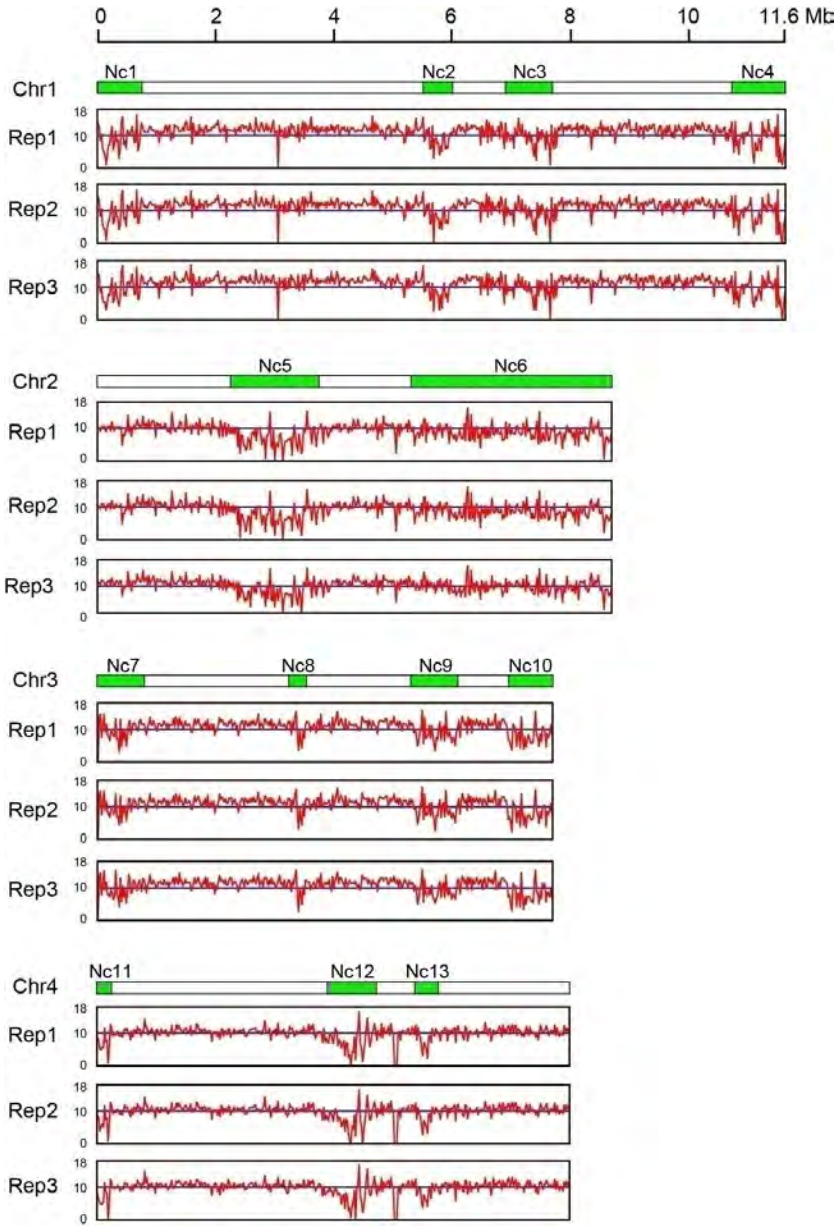
Supplemental Figure 1. Synteny block analysis of *Fusarium oxysporum* with the genomic sequence of *F. verticillioides*. The genomic sequence of *F. oxysporum* strain 4287 was used to compare the genomic sequence of *F. verticillioides* by using program *MCScanX*. Eleven core chromosomes of *F. oxysporum* contain collinear chromosomes in *F. verticillioides*, while four LS chromosomes do not contain collinear chromosomes in *F. verticillioides*. Red triangles represent non-conserved regions identified on the chromosome 1 and 2 of *F. oxysporum*.



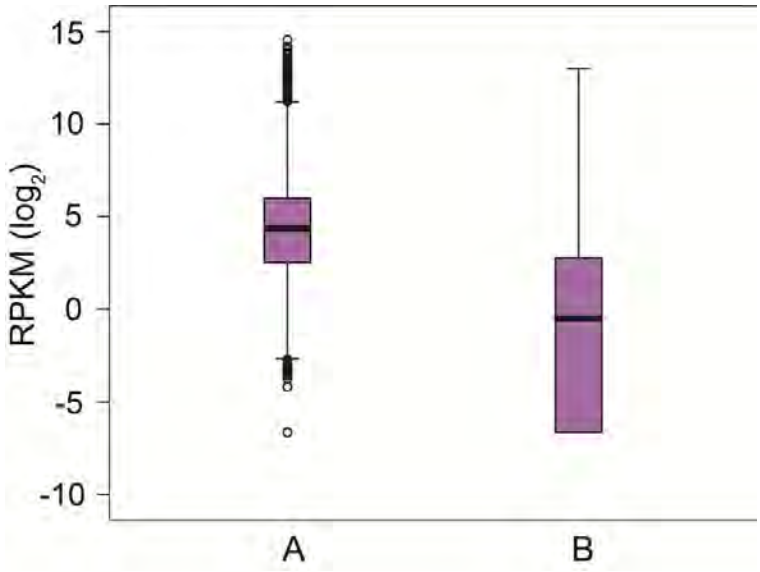
Supplemental Figure 2. Gene expression pattern analysis using RNA-Seq data obtained from mycelium of *F. graminearum*. Each chromosome of *F. graminearum* was divided into 20 kb windows. For each window, the log2-transformed reads coverage was drawn to show gene expression patterns along each chromosome of *F. graminearum*. The gene expression patterns analyzed by three biologically independent RNA-Seq data obtained from mycelium of *F. graminearum* are shown. Green boxes represent non-conserved regions identified in *F. graminearum*.



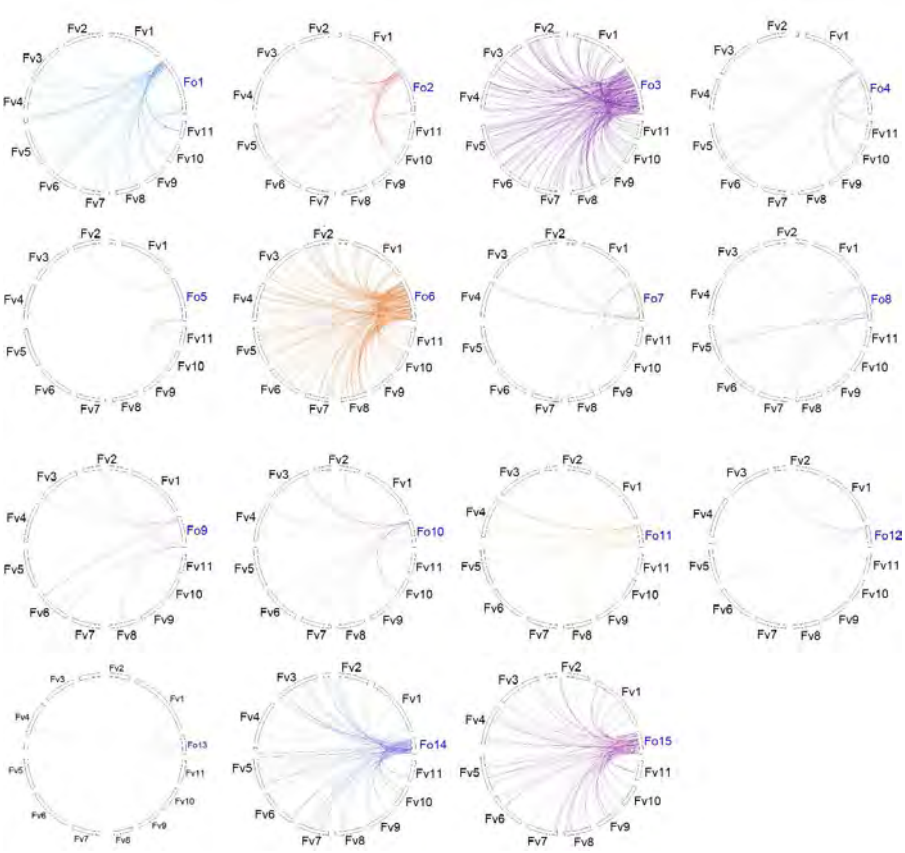
Supplemental Figure 3. Distribution of genes on each chromosome of *F. graminearum* according to their similarity to genes in *F. verticillioides*. Genes on each chromosome were divided into five groups according to their degree of similarity. Genes with a low degree of similarity to their orthologs are enriched in weakly expressed regions, while genes with a high degree of similarity to their orthologs are enriched in highly expressed regions. Green boxes represent non-conserved regions identified in *F. graminearum*. Note the absence of conserved genes (*p* value=0) at the telomeric region of Chr3, corresponding to Nc10.



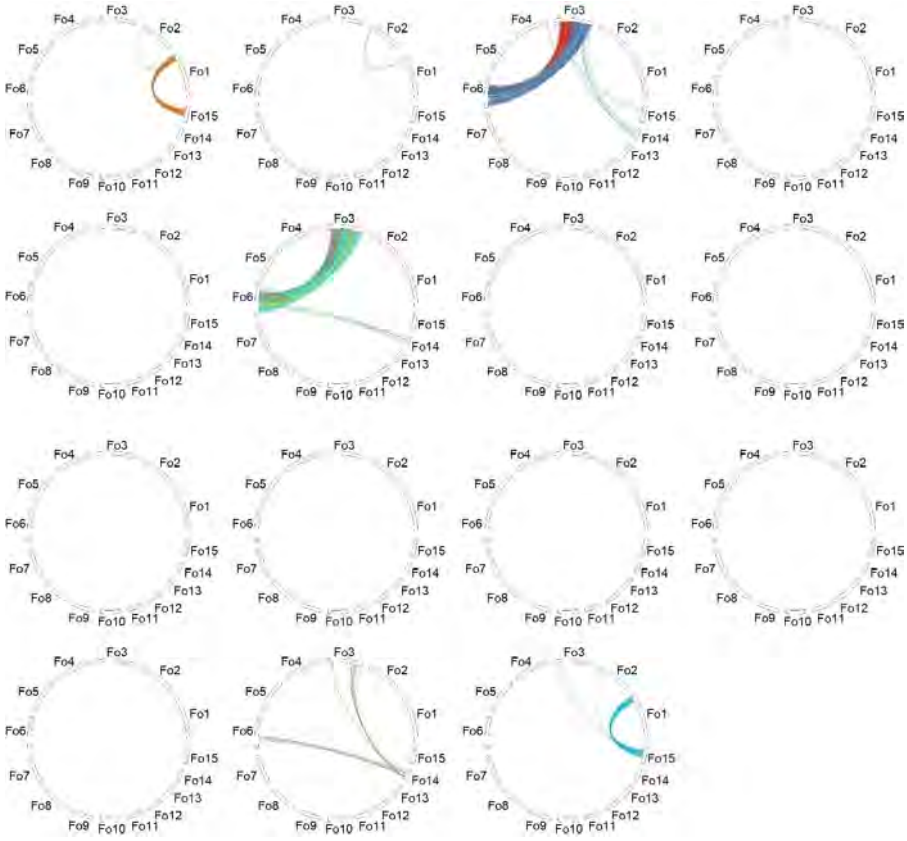
Supplemental Figure 4. Gene expression pattern analysis using RNA-Seq data obtained from conidia of *F. graminearum*. Each chromosome of *F. graminearum* was divided into 20 kb windows. For each window, the log2-transformed reads coverage was drawn to show gene expression patterns along each chromosome of *F. graminearum*. The gene expression patterns analyzed by three biologically independent RNA-Seq data obtained from the conidia of *F. graminearum* are shown. Green boxes represent non-conserved regions identified in *F. graminearum*.



Supplemental Figure 5. Comparison of the expression of relocated and non-relocated genes. All genes that have orthologs in *F. verticillioides* were divided into two groups: genes that are relocated to non-conserved regions and genes that are not relocated. Box plot analysis shows that the expression levels of relocated genes in non-conserved regions (B) are lower than non-relocated genes (A). Asterisk indicates significant difference (p value<0.01, Student's t-test).



Supplemental Figure 6. LS chromosomes in *F. oxysporum* are enriched for gene relocations. Genes that match their orthologs on non-collinear chromosomes of *F. verticillioides* are distributed on the chromosomes of *F. oxysporum*. Four LS chromosomes, Fo3, Fo6, Fo14 and Fo15, show multiple gene relocations. In addition, the telomere proximal regions of the core chromosomes, especially chromosome 1 and 2, show multiple gene relocations.



Supplemental Figure 7. Synteny block analysis of each chromosome against other chromosomes in *F. oxysporum*. The genomic sequence of each chromosome was used to compare other chromosomes by using program *MCScanX*. Genomic sequence duplications were identified between four LS chromosomes. Remarkably, LS chromosome 15 is duplicated from the telomere proximal region of core chromosome 1.

Supplemental Table 1 Group of neighboring genes on LS chromosomes of *Fusarium oxysporum* that are collinear with their homologs in *F. verticillioides*.

Group	Gene ID	Chromosome	Homologs in *Fusarium verticillioides*		
			GeneID	Chromosome	*p* value
1	FOXG_12459	Fochr3	FVEG_06980	Fvchr7	8E-09
	FOXG_12461	Fochr3	FVEG_06981	Fvchr7	1E-08
2	FOXG_12413	Fochr3	FVEG_08664	Fvchr10	2E-21
	FOXG_12412	Fochr3	FVEG_08665	Fvchr10	2E-65
3	FOXG_14888	Fochr3	FVEG_10773	Fvchr11	6E-16
	FOXG_14887	Fochr3	FVEG_10774	Fvchr11	6E-16
4	FOXG_06503	Fochr3	FVEG_13461	Fvchr8	0
	FOXG_06502	Fochr3	FVEG_13462	Fvchr8	3E-14
	FOXG_06501	Fochr3	FVEG_13464	Fvchr8	9E-56
	FOXG_06500	Fochr3	FVEG_13465	Fvchr8	1E-12
5	FOXG_06759	Fochr3	FVEG_13461	Fvchr8	0
	FOXG_06760	Fochr3	FVEG_13462	Fvchr8	4E-47
	FOXG_06761	Fochr3	FVEG_13464	Fvchr8	9E-56
	FOXG_06762	Fochr3	FVEG_13465	Fvchr8	1E-12
6	FOXG_06866	Fochr3	FVEG_13465	Fvchr8	0
	FOXG_06870	Fochr3	FVEG_13464	Fvchr8	2E-48
	FOXG_06871	Fochr3	FVEG_13462	Fvchr8	4E-36
7	FOXG_16139	Fochr3	FVEG_13462	Fvchr8	4E-36
	FOXG_16140	Fochr3	FVEG_13464	Fvchr8	2E-33
	FOXG_16144	Fochr3	FVEG_13465	Fvchr8	0
8	FOXG_07112	Fochr6	FVEG_02865	Fvchr5	8E-71
	FOXG_07100	Fochr6	FVEG_02866	Fvchr5	9E-07
9	FOXG_06967	Fochr6	FVEG_05732	Fvchr3	5E-18
	FOXG_06971	Fochr6	FVEG_05733	Fvchr3	9E-12
	FOXG_06970	Fochr6	FVEG_05734	Fvchr3	2E-21
	FOXG_06969	Fochr6	FVEG_05737	Fvchr3	8E-09
10	FOXG_14078	Fochr6	FVEG_06980	Fvchr7	8E-09
	FOXG_14076	Fochr6	FVEG_06981	Fvchr7	1E-08
11	FOXG_14124	Fochr6	FVEG_08664	Fvchr10	2E-21
	FOXG_14125	Fochr6	FVEG_08665	Fvchr10	1E-60
12	FOXG_06957	Fochr6	FVEG_13464	Fvchr8	3E-36
	FOXG_06958	Fochr6	FVEG_13462	Fvchr8	2E-49
	FOXG_06959	Fochr6	FVEG_13461	Fvchr8	0
13	FOXG_07172	Fochr6	FVEG_13461	Fvchr8	0
	FOXG_07173	Fochr6	FVEG_13462	Fvchr8	1E-51
	FOXG_07174	Fochr6	FVEG_13464	Fvchr8	2E-50
	FOXG_07175	Fochr6	FVEG_13465	Fvchr8	1E-12

14	FOXG_07321	Fochr6	FVEG_13461	Fvchr8	0
	FOXG_07322	Fochr6	FVEG_13462	Fvchr8	4E-47
	FOXG_07323	Fochr6	FVEG_13464	Fvchr8	2E-53
	FOXG_07324	Fochr6	FVEG_13465	Fvchr8	1E-12
15	FOXG_16235	Fochr6	FVEG_13465	Fvchr8	0
	FOXG_16239	Fochr6	FVEG_13464	Fvchr8	2E-33
	FOXG_16240	Fochr6	FVEG_13462	Fvchr8	4E-36
16	FOXG_14290	Fochr14	FVEG_02567	Fvchr6	0
	FOXG_14289	Fochr14	FVEG_02568	Fvchr6	0
17	FOXG_16453	Fochr14	FVEG_05746	Fvchr3	1E-19
	FOXG_16452	Fochr14	FVEG_05747	Fvchr3	9E-16
18	FOXG_14382	Fochr15	FVEG_10080	Fvchr9	6E-72
	FOXG_14384	Fochr15	FVEG_10081	Fvchr9	4E-47
19	FOXG_14360	Fochr15	FVEG_10478	Fvchr11	0
	FOXG_14359	Fochr15	FVEG_10479	Fvchr11	0
	FOXG_14358	Fochr15	FVEG_10480	Fvchr11	0
	FOXG_14357	Fochr15	FVEG_10482	Fvchr11	0
	FOXG_14356	Fochr15	FVEG_10483	Fvchr11	0
	FOXG_14355	Fochr15	FVEG_10484	Fvchr11	0
	FOXG_14354	Fochr15	FVEG_10485	Fvchr11	0
	FOXG_14352	Fochr15	FVEG_10486	Fvchr11	6E-116
	FOXG_14351	Fochr15	FVEG_10487	Fvchr11	0
20	FOXG_16778	Fochr15	FVEG_11801	Fvchr7	1E-94
	FOXG_16781	Fochr15	FVEG_11802	Fvchr7	2E-07
21	FOXG_14336	Fochr15	FVEG_12257	Fvchr4	3E-42
	FOXG_14335	Fochr15	FVEG_12258	Fvchr4	1E-57
22	FOXG_14327	Fochr15	FVEG_13769	Fvchr10	0
	FOXG_14330	Fochr15	FVEG_13770	Fvchr10	1E-08

References

Bai, G., and Shaner, G. (2004). Management and resistance in wheat and barley to fusarium head blight. Annu. Rev. Phytopathol. *42*, 135-161.

Brawand, D., Soumillon, M., Necsulea, A., Julien, P., Csardi, G., Harrigan, P., Weier, M., Liechti, A., Aximu-Petri, A., Kircher, M., *et al.* (2011). The evolution of gene expression levels in mammalian organs. Nature *478*, 343-348.

Brown, N.A., Antoniw, J., and Hammond-Kosack, K.E. (2012). The predicted secretome of the plant pathogenic fungus *Fusarium graminearum*: a refined comparative analysis. PLoS One *7*, e33731.

Chimpanzee, S., and Analysis, C. (2005). Initial sequence of the chimpanzee genome and comparison with the human genome. Nature *437*, 69-87.

Cohen, B.A., Mitra, R.D., Hughes, J.D., and Church, G.M. (2000). A computational analysis of whole-genome expression data reveals chromosomal domains of gene expression. Nat. Genet. *26*, 183-186.

Coleman, J.J., Rounsley, S.D., Rodriguez-Carres, M., Kuo, A., Wasmann, C.C., Grimwood, J., Schmutz, J., Taga, M., White, G.J., Zhou, S., *et al.* (2009). The genome of *Nectria haematococca*: contribution of supernumerary chromosomes to gene expansion. PLoS Genet. *5*, e1000618.

Cuomo, C.A., Gueldener, U., Xu, J.R., Trail, F., Turgeon, B.G., Di Pietro, A., Walton, J.D., Ma, L.J., Baker, S.E., Rep, M., *et al.* (2007). The *Fusarium graminearum* genome reveals a link between localized polymorphism and pathogen specialization. Science *317*, 1400-1402.

de Jonge, R., Bolton, M.D., Kombrink, A., van den Berg, G.C., Yadeta, K.A., and Thomma, B.P. (2013). Extensive chromosomal reshuffling drives evolution of virulence in an asexual pathogen. Genome Res. 23, 1271-1282.

Drosophila 12 Genomes, C., Clark, A.G., Eisen, M.B., Smith, D.R., Bergman, C.M., Oliver, B., Markow, T.A., Kaufman, T.C., Kellis, M., Gelbart, W., *et al.* (2007). Evolution of genes and genomes on the *Drosophila* phylogeny. Nature *450*, 203-218.

Feuk, L., MacDonald, J.R., Tang, T., Carson, A.R., Li, M., Rao, G., Khaja, R., and Scherer, S.W. (2005). Discovery of human inversion polymorphisms by comparative analysis of human and chimpanzee DNA sequence assemblies. PLoS Genet. *1*, e56.

Filichkin, S.A., Priest, H.D., Givan, S.A., Shen, R., Bryant, D.W., Fox, S.E., Wong, W.K., and Mockler, T.C. (2010). Genome-wide mapping of alternative splicing in *Arabidopsis thaliana*. Genome Res. *20*, 45-58.

Fraser, H.B., Moses, A.M., and Schadt, E.E. (2010). Evidence for widespread adaptive evolution of gene expression in budding yeast. Proc. Natl. Acad. Sci. USA *107*, 2977-2982.

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

Galagan, J.E., Calvo, S.E., Cuomo, C., Ma, L.J., Wortman, J.R., Batzoglou, S., Lee, S.I., Basturkmen, M., Spevak, C.C., Clutterbuck, J., *et al.* (2005). Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae*. Nature *438*, 1105-1115.

Gardiner, D.M., McDonald, M.C., Covarelli, L., Solomon, P.S., Rusu, A.G., Marshall, M., Kazan, K., Chakraborty, S., McDonald, B.A., and Manners, J.M. (2012). Comparative pathogenomics reveals horizontally acquired novel virulence genes in fungi infecting cereal hosts. PLoS Pathog. *8*, e1002952.

Goodwin, S.B., M'Barek S, B., Dhillon, B., Wittenberg, A.H., Crane, C.F., Hane, J.K., Foster, A.J., Van der Lee, T.A., Grimwood, J., Aerts, A., *et al.* (2011). Finished genome of the fungal wheat pathogen *Mycosphaerella graminicola* reveals dispensome structure, chromosome plasticity, and stealth pathogenesis. PLoS Genet. *7*, e1002070.

Gordon, J.L., Byrne, K.P., and Wolfe, K.H. (2011). Mechanisms of chromosome number evolution in yeast. PLoS Genet. *7*, e1002190.

Gordon, T.R., and Martyn, R.D. (1997). The evolutionary biology of *Fusarium oxysporum*. Annu. Rev. Phytopathol. *35*, 111-128.

Hansen, F.T., Sorensen, J.L., Giese, H., Sondergaard, T.E., and Frandsen, R.J. (2012). Quick guide to polyketide synthase and nonribosomal synthetase genes in *Fusarium*. Int. J. Food Microbiol. *155*, 128-136.

Houterman, P.M., Ma, L., van Ooijen, G., de Vroomen, M.J., Cornelissen, B.J., Takken, F.L., and Rep, M. (2009). The effector protein Avr2 of the xylem-colonizing fungus *Fusarium oxysporum* activates the tomato resistance protein I-2 intracellularly. Plant J. *58*, 970-978.

Hurst, L.D., Pal, C., and Lercher, M.J. (2004). The evolutionary dynamics of eukaryotic gene order. Nat. Rev. Genet. *5*, 299-310.

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

Kistler, H.C., Alabouvette, C., Baayen, R.P., Bentley, S., Brayford, D., Coddington, A., Correll, J., Daboussi, M.J., Elias, K., Fernandez, D., *et al.* (1998). Systematic numbering of vegetative

compatibility groups in the plant pathogenic fungus *Fusarium oxysporum*. Phytopathology *88*, 30-32.

Ma, L.J., van der Does, H.C., Borkovich, K.A., Coleman, J.J., Daboussi, M.J., Di Pietro, A., Dufresne, M., Freitag, M., Grabherr, M., Henrissat, B., *et al.* (2010). Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*. Nature *464*, 367-373.

Nagalakshmi, U., Wang, Z., Waern, K., Shou, C., Raha, D., Gerstein, M., and Snyder, M. (2008). The transcriptional landscape of the yeast genome defined by RNA sequencing. Science *320*, 1344-1349.

Navarro, A., and Barton, N.H. (2003). Chromosomal speciation and molecular divergence--accelerated evolution in rearranged chromosomes. Science *300*, 321-324.

Oren, L., Ezrati, S., Cohen, D., and Sharon, A. (2003). Early events in the *Fusarium verticillioides*-maize interaction characterized by using a green fluorescent protein-expressing transgenic isolate. Appl. Environ. Microbiol. *69*, 1695-1701.

Ortoneda, M., Guarro, J., Madrid, M.P., Caracuel, Z., Roncero, M.I., Mayayo, E., and Di Pietro, A. (2004). *Fusarium oxysporum* as a multihost model for the genetic dissection of fungal virulence in plants and mammals. Infect. Immun. *72*, 1760-1766.

Parry, D.W., Jenkinson, P., and Mcleod, L. (1995). *Fusarium* Ear Blight (Scab) in Small-Grain Cereals - a Review. Plant Pathol. *44*, 207-238.

Paterson, A.H., Freeling, M., Tang, H., and Wang, X. (2010). Insights from the comparison of plant genome sequences. Annu. Rev. Plant Biol. *61*, 349-372.

Perrin, R.M., Fedorova, N.D., Bok, J.W., Cramer, R.A., Wortman, J.R., Kim, H.S., Nierman, W.C., and Keller, N.P. (2007). Transcriptional regulation of chemical diversity in *Aspergillus fumigatus* by LaeA. PLoS Pathog. *3*, e50.

Proctor, R.H., McCormick, S.P., Alexander, N.J., and Desjardins, A.E. (2009). Evidence that a secondary metabolic biosynthetic gene cluster has grown by gene relocation during evolution of the filamentous fungus *Fusarium*. Mol. Microbiol. *74*, 1128-1142.

Rep, M., van der Does, H.C., Meijer, M., van Wijk, R., Houterman, P.M., Dekker, H.L., de Koster, C.G., and Cornelissen, B.J. (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. *53*, 1373-1383.

Scally, A., Dutheil, J.Y., Hillier, L.W., Jordan, G.E., Goodhead, I., Herrero, J., Hobolth, A., Lappalainen, T., Mailund, T., Marques-Bonet, T., *et al.* (2012). Insights into hominid evolution from the gorilla genome sequence. Nature *483*, 169-175.

Sexton, T., Schober, H., Fraser, P., and Gasser, S.M. (2007). Gene regulation through nuclear organization. Nat. Struct. Mol. Biol. *14*, 1049-1055.

Son, H., Seo, Y.S., Min, K., Park, A.R., Lee, J., Jin, J.M., Lin, Y., Cao, P., Hong, S.Y., Kim, E.K., *et al.* (2011). A phenome-based functional analysis of transcription factors in the cereal head blight fungus, *Fusarium graminearum*. PLoS Pathog. *7*, e1002310.

Su, A.I., Wiltshire, T., Batalov, S., Lapp, H., Ching, K.A., Block, D., Zhang, J., Soden, R., Hayakawa, M., Kreiman, G., *et al.* (2004). A gene atlas of the mouse and human protein-encoding transcriptomes. Proc. Natl. Acad. Sci. USA *101*, 6062-6067.

van der Does, H.C., Lievens, B., Claes, L., Houterman, P.M., Cornelissen, B.J., and Rep, M. (2008). The presence of a virulence locus discriminates *Fusarium oxysporum* isolates causing tomato wilt from other isolates. Environ. Microbiol. *10*, 1475-1485.

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

Wang, Y., Tang, H., Debarry, J.D., Tan, X., Li, J., Wang, X., Lee, T.H., Jin, H., Marler, B., Guo, H., *et al.* (2012). MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic acids Res. *40*, e49.

Wang, Z., Gerstein, M., and Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. Nat. Rev. Genet. *10*, 57-63.

Wei, W., McCusker, J.H., Hyman, R.W., Jones, T., Ning, Y., Cao, Z., Gu, Z., Bruno, D., Miranda, M., Nguyen, M., *et al.* (2007). Genome sequencing and comparative analysis of *Saccharomyces cerevisiae* strain YJM789. Proc. Natl. Acad. Sci. USA *104*, 12825-12830.

Wiemann, P., Sieber, C.M., von Bargen, K.W., Studt, L., Niehaus, E.M., Espino, J.J., Huss, K., Michielse, C.B., Albermann, S., Wagner, D., *et al.* (2013). Deciphering the cryptic genome: genome-wide analyses of the rice pathogen *Fusarium fujikuroi* reveal complex regulation of secondary metabolism and novel metabolites. PLoS Pathog. *9*, e1003475.

Wong, S., and Wolfe, K.H. (2005). Birth of a metabolic gene cluster in yeast by adaptive gene relocation. Nat. Genet. *37*, 777-782.

Woo, Y.H., Walker, M., and Churchill, G.A. (2010). Coordinated expression domains in mammalian genomes. PLoS One *5*, e12158.

Zhao, C., Waalwijk, C., de Wit, P.J., Tang, D., and van der Lee, T. (2013). RNA-Seq analysis reveals new gene models and alternative splicing in the fungal pathogen *Fusarium graminearum*. BMC Genomics *14*, 21.

Chapter 5

General discussion

Cereal crops are continuously threatened by a variety of different fungal pathogens in the field. Infection by some of them can cause severe diseases on crops, which leads to dramatic yield reduction. In addition, mycotoxins produced by these fungi can contaminate grains, which severely threatens the health of human beings upon cereal consumption. Considering the adverse effects of these fungal pathogens, it is extremely important to study the interactions with their hosts and find strategies to control them.

According to their lifestyle, pathogens can be divided into three categories: biotrophic, hemi-biotrophic and necrotrophic pathogens. Biotrophic pathogens are those that do not kill host cells during the infection process. They feed themselves by absorbing nutrients from living host cells, either through intercellular hyphae or through specific feeding structures known as haustoria. Well-known biotrophic plant pathogens include *Blumeria graminis*, *Puccinia* spp., *Ustilago maydis* and *Melampsora lini* (Dean et al. 2012; Duplessis et al. 2011; Giraldo and Valent 2013; Kamper et al. 2006; Nowara et al. 2010). In contrast, necrotrophic pathogens, including *Botrytis cinerea* and *Alternaria brassicicola* (Amselem et al. 2011; Oh et al. 2005), retrieve nutrients from dead or dying cells. During infection, necrotrophic pathogens secret various phytotoxic compounds and cell wall-degrading enzymes (CWDEs) to induce cell death and cause leakage of nutrients (Mengiste, 2012). Hemi-biotrophic pathogens explore both strategies. During the initial biotrophic stage, they suppress the host's immune system and cell death, which allows them to colonize the living host and spread throughout the plant tissues. At a late stage of infection, the increasing fungal biomass becomes a burden to the host, and toxins are secreted to induce cell death of host tissues initiating a necrotrophic lifestyle. Plant pathogens that belong to this group include *Magnaporthe oryzae*, *Fusarium graminearum*, *Fusarium oxysporum*, *Mycosphaerella graminicola* and *Colletotrichum* spp. (Ding et al., 2011; Goodwin et al., 2011; Liu et al., 2007; Lyons et al., 2015; Marcel et al., 2010; Trail, 2009).

During evolution, plants have developed two mechanisms of innate immunity: PAMP (pathogen-associated molecular pattern)-triggered immunity (PTI) or MAMP (microbe-associated molecular pattern)-triggered immunity (MTI) and effector-triggered immunity (ETI) (Boller and Felix, 2009). PTI or MTI are basal defense responses, which are based on PRR (pattern recognition receptor)-mediated recognition of either PAMPs or MAMPs. Well-studied PPR-PAMP/MAMP models in *Arabidopsis* include FLS2 (flagellin sensing 2)-mediated

recognition of flg22, a conserved part of the N terminus of bacterial flagellin (Gomez-Gomez and Boller, 2000); EFR (EF-Tu receptor)-mediated recognition of EF-Tu (elongation factor Tu) (Zipfel et al., 2006); and CERK1 (chitin elicitor receptor kinase 1)-mediated recognition of chitin, a main component of fungal cell walls (Miya et al., 2007). In order to suppress MTI or PTI, pathogens have evolved a variety of effectors to suppress defense proteins and enable them to successfully infect hosts. However, plants have also evolved certain resistance (R) proteins that specifically recognize effectors, triggering ETI. ETI is often accompanied by a hypersensitive response (HR), a type of programmed cell death (Cui et al., 2015).

For plants that are infected by the obligate biotrophic fungal pathogens, such as rust fungi and powdery mildew fungi, many resistance (*R)* genes encoding R proteins have been identified that specifically recognize effectors and activate ETI (Giraldo and Valent, 2013). Similarly, several R proteins have been identified in tomato that specifically recognize the effectors secreted by *F. oxysporum* f. sp. *lycopersici* (*Fol*) (Houterman et al., 2009; Rep et al., 2004). So far, the identified effectors include Avr1, Avr2, Avr3 and 11 other secreted-in-xylem (Six) proteins (Gawehns et al., 2014; Houterman et al., 2007; Ma et al., 2010). Avr2 and Avr3, recognized by tomato plants carrying the *R* genes *I-2* and *I-3*, respectively, are required as effectors causing virulence on susceptible tomato lines (Houterman et al., 2009; Rep et al., 2004). Avr1, recognized by *I* and *I-1*, is not required as a virulence factor on susceptible tomato lines, but plays an important role in the suppression of *I-2*-and *I-3*-mediated resistance (Houterman et al., 2008). Although most effectors are secreted and recognized in xylem vessels of tomato, it has also been shown that Avr2 can move to the nuclei of the host cells and trigger an I-2-dependent HR (Houterman et al., 2009; Ma et al., 2013), which suggests that effectors secreted by *F. oxysporum* can be recognized both extracellularly and intracellularly.

However, so far, no *R* genes were identified in wheat or other crops that can recognize and mediate a resistance response against *F. graminearum*. Although many secreted proteins are predicted in the genome of *F. graminearum*, to date no effector genes were identified that are required for the infection of its host plant by this fungus (Brown et al., 2012; Jonkers et al., 2012). One possible reason could be that the function of the secreted protein were not well studied yet. Another possibility is that, during infection of spikes of wheat, the phytotoxins such as the trichothecenes, synthesized and secreted by *F. graminearum*, may disrupt the ETI response.

Although no *R* genes have been identified in wheat against *F. graminearum*, it is still important to identify other types of genes that might provide resistance to this fungus. So far, several plant and microbial (defense-related) genes, including those encoding pectin methylesterase, lactoferrin, radish defensin, a truncated form of yeast ribosomal protein L3, polygalacturonase-inhibiting proteins and NPR1 (Di et al. 2010; Han et al. 2012; Li et al. 2011b; Makandar et al. 2006; Volpi et al. 2011), have been transformed to wheat to increase resistance to FHB. In future, more of those genes need to be identified and introduced to obtain more resistance crops by natural breeding or transgenesis.

To effectively control outbreaks and spread of FHB, and to identify other sources of resistance, it is essential to investigate infection mechanisms of *F. graminearum* on its hosts. In order to achieve this, identification of virulence genes in *F. graminearum* is critical. So far, more than 100 genes have been identified that are involved in virulence. These genes include kinases, transcription factors, RasGTPase, a secreted lipase and metabolite-associated enzymes (Bluhm et al. 2007; Han et al. 2007; Oide et al. 2006; Seong et al. 2006; Seong et al. 2009; Son et al. 2011; Voigt et al. 2005; Wang et al. 2011). From these studies in might be concluded that successful infection of its hosts by *F. graminearum* depends on tight regulation of many different biological processes. So far, most of the research programmes have focused on studying a single gene. The interactions between the identified virulence genes are still largely unknown, which restricts our view of the infection mechanism of *F. graminearum*. One of the possible solutions to extend our knowledge is to identify more virulence genes by using high-throughput screening methods. The identification of multiple virulence genes and investigation of the intrinsic functions of the encoded products will give us further insight of the pathogenicity network of *F. graminearum*, which in turn will enable us to find the most appropriate strategy to control this fungal cereal pathogen in the field.

In search for the identification of additional virulence genes in *F. graminearum*, in this thesis, I performed high-throughput screenings for mutants that showed reduced virulence. Several candidate virulence genes were identified using this strategy, and one of these candidate genes, *EBR1* (*enhanced branching 1*), which encodes a Zn_2Cys_6 transcription factor, was selected to be studied in more detail. As shown in chapter 2, *EBR1* is required for the growth and virulence of *F. graminearum*. Using a GFP-tagged strain, we demonstreated that PH-1*Δebr1* mutants lost the

ability to penetrate the rachis of spikes, and showed a significantly reduced virulence when compared with the parental strain PH-1 of *F. graminearum*. Recently, Jonkers and colleagues (Jonkers et al. 2014) studied the intrinsic function of the *EBR1* gene in *Fol* strains. Unlike the single copy *EBR1* gene in *F. graminearum* and *F. verticillioides*, nine paralogues of *EBR1* gene were identified in *Fol* strains. These nine genes could be divided into four groups, named *EBR1* (one copy), *EBR2* (one copy), *EBR3* (three copies) and *EBR4* (four copies), based on the similarity and their location on the chromosomes of *Fol*. Interestingly, except for the *EBR1* gene, all other eight *EBR* genes locate on supernumerary chromosomes of *Fol*. Deletion of *EBR1* in different *F. oxysporum* species all cause reduced radical growth on both rich and poor media, suggesting that EBR1 is required for growth regulation. In addition, loss of *EBR1* in *F. oxysporum* species also leads to reduced virulence. So far, the evolutionary advantage of expansion of *EBR* genes in *F. oxysporum* is still not yet understood, but it may contribute to virulence or niche specialization. Apart from *Fusarium* species, one orthologous gene of *EBR1* was also identified in *M. oryzae* (Chung et al. 2013), which is named MGG_09263. Inactivation of this gene by T-DNA insertion resulted in impaired appressorium formation and reduced virulence. These results suggest that *EBR1* is essential for full virulence in multiple plant pathogens.

The roles of EBR1 in the growth and pathogenicity have been confirmed in both *F. graminearum* and *F. oxysporum*. However, as a transcription factor, the targets of *EBR1* are still not known. Jonkers and colleagues (Jonkers et al. 2014) performed a microarray assay on wild types and *ebr1* mutants in both *F. graminearum* and *F. oxysporum*. Although a similar number of genes are up-regulated and down-regulated in *ebr1* mutant in both *Fusarium* species, the overlap of the affected genes between these two *Fusarium* species is very limited. Specifically, genes that are down-regulated in *Fg*Δ*ebr1* are enriched in the categories of metabolism and energy, while genes that are up-regulated are enriched in the categories of cell rescue, defense and virulence. In contrast, the genes down-regulated in *Fo*Δ*ebr1* are enriched in cell rescue, defense and virulence, while genes that are up-regulated are enriched in the category of metabolism. The reason why the EBR1-mediated gene expression in *F. graminearum* and *F. oxysporum* is in opposite directions is still unknown. One possible explanation is that EBR1 may play different or even opposite roles in different hyphal growth stages. It is well known that *F. graminearum* grows faster than *F. oxysporum*, so even though RNA samples of *F. graminearum* and *F. oxysporum* were collected at a

similar time point (48 h) after incubation in the medium, the two species are likely in different developmental stages, explaining why genome-wide gene expression would be distinct. To answer this question, different time course experiments could be performed for analysis of gene expression in both *Fusarium* species. Using microarray data, the cellular pathways that are affected in the *ebr1* mutant have been identified, but the intrinsic function of EBR1 remains elusive. In future, Chip-Seq assays could be performed to identify the direct targets of EBR1, which might help to further understand the intrinsic function of EBR1 and explain why mutating *EBR1* caused reduced radial growth and reduced virulence.

In the second part of this thesis, we studied the genome annotation, post-transcriptional regulation and genome-wide gene expression in *F. graminearum* using RNA-Seq data. The genome of *F. graminearum* has been sequenced in 2006 and subsequently genome annotation was completed based on the bioinformatics methods and EST databases available at that time (Cuomo et al. 2007). However, due to the limitation of the bioinformatics methods and limited number of ESTs available in the public database, many genes have not been annotated correctly. Fortunately, the high-throughput RNA-Seq method can be used to analyse different populations of RNA, including total RNA, small RNA, such as miRNA, tRNA, and ribosomal profiling. RNA-Seq can also be used to more accurately determine exon/intron boundaries and verify or amend previously annotated 5' and 3' gene boundaries, which altogether improves the annotation of gene models. So far, RNA-Seq technology has been applied in many organisms, ranging from animals, plants, fungi to bacteria (Bruno et al. 2010; Li et al. 2011a; Lu et al. 2010; Mortazavi et al. 2008; Wang et al. 2013). In Chapter 3, we used RNA-Seq data from the *F. graminearum* PH-1 strain to comprehensively re-inspect the gene models. In total, 655 incorrectly predicted gene models were identified and revised. Furthermore, 2459 novel transcriptionally active regions (nTARs) were identified. In our RNA-Seq data, less than 60% of predicted genes were over 90% covered by reads, so probably many incorrectly annotated genes are missing in our study. In future, more RNA-Seq data could be incorporated to further improve the gene annotation of *F. graminearum*.

Another important application of RNA-Seq technology is to study post-transcriptional regulation, including alternative splicing and RNA-editing (Filichkin et al. 2010; Peng et al. 2012; Sultan et al. 2008). Alternative splicing has been widely identified in animals, plants and fungi (Brown et al., 2008; Filichkin et al., 2010; Loftus et al., 2005; McGuire et al., 2008; Sultan et al.,

2008). RNA-editing was mainly identified in animals and plants (Bentolila et al., 2012; Peng et al., 2012), and so far little is known about this phenomenon in fungi. In our study, around 220 genes that show alternative splicing were identified. There are four different forms of alternative splicing, including exon skipping, intron retention, alternative 5' splicing and alternative 3' splicing. In *F. graminearum*, the most prevalent form of alternative splicing is intron retention. Our studies showed that the fraction of alternatively spliced genes in *F. graminearum* is 1.7%, which is much lower than that in *Homo sapiens* (95%) and *Arabidopsis thaliana* (42%) (Filichkin et al. 2010; Pan et al. 2008), and also lower than that in other fungi, such as *Aspergillus oryzae* (8.6%) and *M. oryzae* (2.7%) (Ebbole et al. 2004; Wang et al. 2010). These results suggest that alternative splicing is less prevalent in fungi than in plants and animals. Interestingly, for some alternative spliced genes we found that the ratio of different transcripts altered throughout development, implying that different types of transcripts could have special functions in different developmental growth stages. We also searched for evidence of RNA editing in *F. graminearum*, but so far we did not find this. This suggests that RNA editing events mainly occur in animals and plants, or alternatively, insufficient cases in fungi have been studied so far.

RNA-Seq data were mostly used to analyze individual gene expression. In chapter 4, we explored RNA-Seq data to analyze gene expression patterns at the whole chromosome level. Interestingly, some lowly expression regions were identified on each chromosome of *F. graminearum*, suggesting that the level of gene expression is not evenly distributed on the different chromosomes. Besides, other features were found in the lowly expression regions, including low conservation and a large number of gene relocations. These analyses indicate that the non-conserved chromosomal regions of *F. graminearum* were assembled by frequent gene relocations that are expected to accelerate the evolution of species-specific genes. Furthermore, we found that genes that are induced in specific developmental stages or different environmental conditions are enriched in non-conserved genes, while house-keeping genes that are stably expression during different developmental stages are enriched in conserved regions. For example, genes encoding secreted proteins or secondary metabolite-related enzymes (Cuomo et al. 2007; Gaffoor et al. 2005; Kim et al. 2005; Voigt et al. 2005), which are induced under specific conditions, are abundant in non-conserved regions.

In conclusion, our studies pointed to possible mechanisms of the evolution of genome and

gene expression in *F. graminearum*. So far, the correlation between genome evolution and genome-wide gene expression regulation is still largely unknown. Our studies provided insight in how these two processes are interconnected and could contribute to the successful adaptation of fungi to complex ecological niches. In future, it will be interesting to investigate whether gene relocations and their effects on the evolution of genome and gene expression can also be detected in other fungi species.

References

Amselem, J., Cuomo, C.A., van Kan, J.A., Viaud, M., Benito, E.P., Couloux, A., Coutinho, P.M., de Vries, R.P., Dyer, P.S., Fillinger, S., *et al.* (2011). Genomic analysis of the necrotrophic fungal pathogens *Sclerotinia sclerotiorum* and *Botrytis cinerea*. PLoS Genet. *7*, e1002230.

Bentolila, S., Heller, W.P., Sun, T., Babina, A.M., Friso, G., van Wijk, K.J., and Hanson, M.R. (2012). RIP1, a member of an *Arabidopsis* protein family, interacts with the protein RARE1 and broadly affects RNA editing. Proc. Natl. Acad. Sci. USA *109*, E1453-1461.

Bluhm, B.H., Zhao, X., Flaherty, J.E., Xu, J.R., and Dunkle, L.D. (2007). RAS2 regulates growth and pathogenesis in *Fusarium graminearum*. Mol. Plant Microbe Interact. *20*, 627-636.

Boller, T., and Felix, G. (2009). A renaissance of elicitors: perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors. Annu. Rev. Plant Biol. *60*, 379-406.

Brown, D.W., Butchko, R.A., and Proctor, R.H. (2008). Genomic analysis of *Fusarium verticillioides*. Food Addit. Contam. Part A *25*, 1158-1165.

Brown, N.A., Antoniw, J., and Hammond-Kosack, K.E. (2012). The predicted secretome of the plant pathogenic fungus *Fusarium graminearum*: A Refined Comparative Analysis. PLoS One *7*, e33731.

Bruno, V.M., Wang, Z., Marjani, S.L., Euskirchen, G.M., Martin, J., Sherlock, G., and Snyder, M. (2010). Comprehensive annotation of the transcriptome of the human fungal pathogen *Candida albicans* using RNA-seq. Genome Res. *20*, 1451-1458.

Chung, H., Choi, J., Park, S.Y., Jeon, J., and Lee, Y.H. (2013). Two conidiation-related Zn(II)(2)Cys(6) transcription factor genes in the rice blast fungus. Fungal Genet. Biol. *61*, 133-141.

Cui, H.T., Tsuda, K., and Parker, J.E. (2015). Effector-triggered immunity: from pathogen perception to robust defense. Annu. Rev. Plant Biol. *66*, 487-511.

Cuomo, C.A., Gueldener, U., Xu, J.R., Trail, F., Turgeon, B.G., Di Pietro, A., Walton, J.D., Ma, L.J., Baker, S.E., Rep, M., *et al.* (2007). The *Fusarium graminearum* genome reveals a link between localized polymorphism and pathogen specialization. Science *317*, 1400-1402.

Dean, R., Van Kan, J.A., Pretorius, Z.A., Hammond-Kosack, K.E., Di Pietro, A., Spanu, P.D., Rudd, J.J., Dickman, M., Kahmann, R., Ellis, J., *et al.* (2012). The Top 10 fungal pathogens in molecular plant pathology. Mol. Plant Pathol. *13*, 414-430.

Di, R., Blechl, A., Dill-Macky, R., Tortora, A., and Tumer, N.E. (2010). Expression of a truncated form of yeast ribosomal protein L3 in transgenic wheat improves resistance to Fusarium head blight. Plant Sci. *178*, 374-380.

Ding, L., Xu, H., Yi, H., Yang, L., Kong, Z., Zhang, L., Xue, S., Jia, H., and Ma, Z. (2011). Resistance

to hemi-biotrophic *F. graminearum* infection is associated with coordinated and ordered expression of diverse defense signaling pathways. PLoS One *6*, e19008.

Duplessis, S., Cuomo, C.A., Lin, Y.C., Aerts, A., Tisserant, E., Veneault-Fourrey, C., Joly, D.L., Hacquard, S., Amselem, J., Cantarel, B.L., *et al.* (2011). Obligate biotrophy features unraveled by the genomic analysis of rust fungi. Proc. Natl. Acad. Sci. USA *108*, 9166-9171.

Ebbole, D.J., Jin, Y., Thon, M., Pan, H.Q., Bhattarai, E., Thomas, T., and Dean, R. (2004). Gene discovery and gene expression in the rice blast fungus, *Magnaporthe grisea*: Analysis of expressed sequence tags. Mol. Plant Microbe Interact. *17*, 1337-1347.

Filichkin, S.A., Priest, H.D., Givan, S.A., Shen, R.K., Bryant, D.W., Fox, S.E., Wong, W.K., and Mockler, T.C. (2010). Genome-wide mapping of alternative splicing in *Arabidopsis thaliana*. Genome Res. *20*, 45-58.

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

Gawehns, F., Houterman, P.M., Ichou, F.A., Michielse, C.B., Hijdra, M., Cornelissen, B.J.C., Rep, M., and Takken, F.L.W. (2014). The *Fusarium oxysporum*effector Six6 contributes to virulence and suppresses I-2-mediated cell death. Mol. Plant Microbe Interact. *27*, 336-348.

Giraldo, M.C., and Valent, B. (2013). Filamentous plant pathogen effectors in action. Nat. Rev. Microbiol.*11*, 800-814.

Gomez-Gomez, L., and Boller, T. (2000). FLS2: an LRR receptor-like kinase involved in the perception of the bacterial elicitor flagellin in *Arabidopsis*. Mol. Cell *5*, 1003-1011.

Goodwin, S.B., M'Barek S, B., Dhillon, B., Wittenberg, A.H., Crane, C.F., Hane, J.K., Foster, A.J., Van der Lee, T.A., Grimwood, J., Aerts, A., *et al.* (2011). Finished genome of the fungal wheat pathogen *Mycosphaerella graminicola* reveals dispensome structure, chromosome plasticity, and stealth pathogenesis. PLoS Genet.*7*, e1002070.

Han, J.G., Lakshman, D.K., Galvez, L.C., Mitra, S., Baenziger, P.S., and Mitra, A. (2012). Transgenic expression of lactoferrin imparts enhanced resistance to head blight of wheat caused by *Fusarium graminearum.* BMC Plant Biol. *12*, 1471-2229.

Han, Y.K., Kim, M.D., Lee, S.H., Yun, S.H., and Lee, Y.W. (2007). A novel F-box protein involved in sexual development and pathogenesis in *Gibberella zeae*. Mol. Microbiol. *63*, 768-779.

Houterman, P.M., Cornelissen, B.J.C., and Rep, M. (2008). Suppression of plant resistance gene-based immunity by a fungal effector. PLoS Pathog. *4*, e1000061.

Houterman, P.M., Ma, L., van Ooijen, G., de Vroomen, M.J., Cornelissen, B.J., Takken, F.L., and Rep, M. (2009). The effector protein Avr2 of the xylem-colonizing fungus *Fusarium oxysporum* activates the tomato resistance protein I-2 intracellularly. Plant J. *58*, 970-978.

Houterman, P.M., Speijer, D., Dekker, H.L., de Koster, C.G., Cornelissen, B.J.C., and Rep, M. (2007). The mixed xylem sap proteome of *Fusarium oxysporum*-infected tomato plants. Mol. Plant Pathol. *8*, 215-221.

Jonkers, W., Dong, Y.H., Broz, K., and Kistler, H.C. (2012). The Wor1-like protein Fgp1 regulates pathogenicity, toxin synthesis and reproduction in the phytopathogenic fungus *Fusarium graminearum*. PLoS Pathog. *8*, e1002724.

Jonkers, W., Xayamongkhon, H., Haas, M., Olivain, C., van der Does, H.C., Broz, K., Rep, M., Alabouvette, C., Steinberg, C., and Kistler, H.C. (2014). *EBR1* genomic expansion and its role in virulence of *Fusarium* species. Environ. Microbiol. *16*, 1982-2003.

Kamper, J., Kahmann, R., Bolker, M., Ma, L.J., Brefort, T., Saville, B.J., Banuett, F., Kronstad, J.W., Gold, S.E., Muller, O., *et al*.(2006). Insights from the genome of the biotrophic fungal plant pathogen *Ustilago maydis*. Nature *444*, 97-101.

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

Li, Z., Zhang, Z.H., Yan, P.C., Huang, S.W., Fei, Z.J., and Lin, K. (2011a). RNA-Seq improves annotation of protein-coding genes in the cucumber genome. BMC Genomics *12*, 540.

Li, Z., Zhou, M.P., Zhang, Z.Y., Ren, L.J., Du, L.P., Zhang, B.Q., Xu, H.J., and Xin, Z.Y. (2011b). Expression of a radish defensin in transgenic wheat confers increased resistance to *Fusarium graminearum* and *Rhizoctonia cerealis*. Funct. Integr. Genomics *11*, 63-70.

Liu, G., Kennedy, R., Greenshields, D.L., Peng, G., Forseille, L., Selvaraj, G., and Wei, Y. (2007). Detached and attached *Arabidopsis* leaf assays reveal distinctive defense responses against hemibiotrophic *Colletotrichum* spp. Mol. Plant Microbe Interact. *20*, 1308-1319.

Loftus, B.J., Fung, E., Roncaglia, P., Rowley, D., Amedeo, P., Bruno, D., Vamathevan, J., Miranda, M., Anderson, I.J., Fraser, J.A., *et al.* (2005). The genome of the basidiomycetous yeast and human pathogen *Cryptococcus neoformans*. Science *307*, 1321-1324.

Lu, T.T., Lu, G.J., Fan, D.L., Zhu, C.R., Li, W., Zhao, Q.A., Feng, Q., Zhao, Y., Guo, Y.L., Li, W.J., *et al.* (2010). Function annotation of the rice transcriptome at single-nucleotide resolution by RNA-seq. Genome Res. *20*, 1238-1249.

Lyons, R., Stiller, J., Powell, J., Rusu, A., Manners, J.M., and Kazan, K. (2015). *Fusarium oxysporum*

triggers tissue-specific transcriptional reprogramming in *Arabidopsis thaliana*. PLoS One *10*, e0121902.

Ma, L.J., van der Does, H.C., Borkovich, K.A., Coleman, J.J., Daboussi, M.J., Di Pietro, A., Dufresne, M., Freitag, M., Grabherr, M., Henrissat, B., *et al.* (2010). Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*. Nature *464*, 367-373.

Ma, L.S., Cornelissen, B.J.C., and Takken, F.L.W. (2013). A nuclear localization for Avr2 from *Fusarium oxysporum* is required to activate the tomato resistance protein I-2. Front. Plant Sci. *4*, 94.

Makandar, R., Essig, J.S., Schapaugh, M.A., Trick, H.N., and Shah, J. (2006). Genetically engineered resistance to Fusarium head blight in wheat by expression of *Arabidopsis NPR1*. Mol. Plant Microbe Interact. *19*, 123-129.

Marcel, S., Sawers, R., Oakeley, E., Angliker, H., and Paszkowski, U. (2010). Tissue-adapted invasion strategies of the rice blast fungus *Magnaporthe oryzae*. Plant Cell *22*, 3177-3187.

McGuire, A.M., Pearson, M.D., Neafsey, D.E., and Galagan, J.E. (2008). Cross-kingdom patterns of alternative splicing and splice recognition. Genome Biol.*9*, R50.

Mengiste, T. (2012). Plant immunity to necrotrophs. Annu. Rev. Phytopathol. *50*, 267-294.

Miya, A., Albert, P., Shinya, T., Desaki, Y., Ichimura, K., Shirasu, K., Narusaka, Y., Kawakami, N., Kaku, H., and Shibuya, N. (2007). CERK1, a LysM receptor kinase, is essential for chitin elicitor signaling in *Arabidopsis*. Proc. Natl. Acad. Sci. USA *104*, 19613-19618.

Mortazavi, A., Williams, B.A., Mccue, K., Schaeffer, L., and Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat. Methods *5*, 621-628.

Nowara, D., Gay, A., Lacomme, C., Shaw, J., Ridout, C., Douchkov, D., Hensel, G., Kumlehn, J., and Schweizer, P. (2010). HIGS: host-induced gene silencing in the obligate biotrophic fungal pathogen *Blumeria graminis*. Plant Cell *22*, 3130-3141.

Oh, I.S., Park, A.R., Bae, M.S., Kwon, S.J., Kim, Y.S., Lee, J.E., Kang, N.Y., Lee, S., Cheong, H., and Park, O.K. (2005). Secretome analysis reveals an *Arabidopsis* lipase involved in defense against *Alternaria brassicicola*. Plant Cell *17*, 2832-2847.

Oide, S., Moeder, W., Krasnoff, S., Gibson, D., Haas, H., Yoshioka, K., and Turgeon, B.G. (2006). *NPS6*, encoding a nonribosomal peptide synthetase involved in siderophore-mediated iron metabolism, is a conserved virulence determinant of plant pathogenic ascomycetes. Plant Cell *18*, 2836-2853.

Pan, Q., Shai, O., Lee, L.J., Frey, J., and Blencowe, B.J. (2008). Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. Nat. Genet. *40*, 1413-1415.

Peng, Z., Cheng, Y., Tan, B.C., Kang, L., Tian, Z., Zhu, Y., Zhang, W., Liang, Y., Hu, X., Tan, X., *et al.*(2012). Comprehensive analysis of RNA-Seq data reveals extensive RNA editing in a human transcriptome. Nat. Biotechnol. *30*, 253-260.

Rep, M., van der Does, H.C., Meijer, M., van Wijk, R., Houterman, P.M., Dekker, H.L., de Koster, C.G., and Cornelissen, B.J. (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. *53*, 1373-1383.

Seong, K., Li, L., Hou, Z.M., Tracy, M., Kistler, H.C., and Xu, J.R. (2006). Cryptic promoter activity in the coding region of the HMG-CoA rediactase gene in *Fusarium graminearum*. Fungal Genet. Biol. *43*, 34-41.

Seong, K.Y., Pasquali, M., Zhou, X.Y., Song, J., Hilburn, K., McCormick, S., Dong, Y.H., Xu, J.R., and Kistler, H.C. (2009). Global gene regulation by *Fusarium* transcription factors *Tri6* and *Tri10* reveals adaptations for toxin biosynthesis. Mol. Microbiol. *72*, 354-367.

Son, H., Seo, Y.S., Min, K., Park, A.R., Lee, J., Jin, J.M., Lin, Y., Cao, P.J., Hong, S.Y., Kim, E.K., *et al.* (2011). A phenome-based functional analysis of transcription factors in the cereal head blight fungus, *Fusarium graminearum*. PLoS Pathog. *7*, e1002310.

Sultan, M., Schulz, M.H., Richard, H., Magen, A., Klingenhoff, A., Scherf, M., Seifert, M., Borodina, T., Soldatov, A., Parkhomchuk, D., *et al.*(2008). A global view of gene activity and alternative splicing by deep sequencing of the human transcriptome. Science *321*, 956-960.

Trail, F. (2009). For blighted waves of grain: *Fusarium graminearum* in the postgenomics era. Plant physiol.*149*, 103-110.

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

Volpi, C., Janni, M., Lionetti, V., Bellincampi, D., Favaron, F., and D'Ovidio, R. (2011). The ectopic expression of a pectin methyl esterase inhibitor increases pectin methyl esterification and limits fungal diseases in wheat. Mol. Plant Microbe Interact. *24*, 1012-1019.

Wang, B., Guo, G.W., Wang, C., Lin, Y., Wang, X.N., Zhao, M.M., Guo, Y., He, M.H., Zhang, Y., and Pan, L. (2010). Survey of the transcriptome of *Aspergillus oryzae* via massively parallel mRNA sequencing. Nucleic Acids Res. *38*, 5075-5087.

Wang, C.F., Zhang, S.J., Hou, R., Zhao, Z.T., Zheng, Q., Xu, Q.J., Zheng, D.W., Wang, G.H., Liu, H.Q., Gao, X.L., *et al.* (2011). Functional analysis of the kinome of the wheat scab fungus *Fusarium graminearum*. PLoS Pathog. *7*, e1002460.

Wang, L., Zhao, S., Gu, C., Zhou, Y., Zhou, H., Ma, J.J., Cheng, J., and Han, Y.P. (2013). Deep

RNA-Seq uncovers the peach transcriptome landscape. Plant Mol. Biol. *83*, 365-377.

Zipfel, C., Kunze, G., Chinchilla, D., Caniard, A., Jones, J.D.G., Boller, T., and Felix, G. (2006). Perception of the bacterial PAMP EF-Tu by the receptor EFR restricts *Agrobacterium*-mediated transformation. Cell *125*, 749-760.

Summary

Fusarium graminearum is a destructive plant fungal pathogen that can infect a variety of important cereal crops, such as wheat, barley and oat, causing Fusarium head blight (FHB), which leads to yield reduction and contamination of kernels by mycotoxins produced by this fungus. In order to control this disease, a better understanding of the interaction of this fungus with its host plants is critical. The research described in this thesis aims to identify the virulence genes that are required for *F. graminearum* to infect its hosts. Also the post-transcriptional regulation and whole genome-wide gene expression of *F. graminearum* were studied to enable us to understand how this fungus adapts to different ecological niches through the regulation of gene expression.

Chapter 1 provides an overview of the current status of research on *F. graminearum*. So far, more than 100 genes have been identified that are required for the virulence of *F. graminearum*. The biological functions of some these virulence genes are described and discussed. The roles of secondary metabolites and gene clusters that are involved in their biogenesis are also discussed. Finally, the current genome annotation information and the comparison of the whole genome of *F. graminearum* with its closely related species *Fusarium verticillioides* and *Fusarium oxysporum* are described.

In **Chapter 2** the biological function of a Zn_2Cys_6 transcription factor, EBR1 (enhanced branching 1), was studied in detail. Mutant *Δebr1*, which showed reduced radial growth and virulence, was initially identified by screening a *F. graminearum* mutant pool generated by transposon *mimp1*-mediated mutagenesis. To further confirm the role of *EBR1*, the knock-out mutant strain PH-1*Δebr1* was generated through homologous recombination. The PH-1*Δebr1* strain also showed reduced radial growth and virulence. Although showing reduced radial growth, the PH-1*Δebr1* strain germinated faster than wild-type strain PH-1 on solid complete medium. Detailed microscopic observations showed that PH-1*Δebr1* developed more branched hyphae than the wild-type PH-1 after growth for 24 hrs on solid medium, suggesting that EBR1 probably plays a positive role on apical dominance of the hyphal tip, but a negative role in hyphal branching. Using GFP-tagged *EBR1* knock-out mutant strain PSC*Δebr1*, we found that the *Δebr1* mutant was unable to penetrate into the rachis of spikes, thereby

resulting in reduced virulence. Subcellular localization analysis indicated that EBR1 is exclusively localized in the nucleus in both conidia and hyphae, which is consistent with the transcriptional role of *EBR1* in cells. Finally, the role of orthologous gene FOXG_05408 was investigated in *F. oxysporum* f. sp.*lycopersici* (*Fol*). Disruption of FOXG_05408 in the *Fol* strain only slightly affected radial growth, which could be due to the redundancy of six potential paralogous genes of the FOXG_05408 gene in the *Fol* strain. However, transforming the FOXG_05408 gene into the mutant strain PH-1*Δebr1* fully restored to wild type phenotype, indicating that FOXG_05408 has similar biological function as *EBR1*.

The genome of *F. graminearum* has been sequenced and annotated previously, but correct gene annotation remains a challenge. In **Chapter 3**, using RNA-seq data, we comprehensively re-analyzed the current gene models described in the Broad database. In total, 655 incorrectly predicted gene models were identified and revised, including incorrect intron predictions, incorrect intron splice sites and prediction of novel introns. Besides, the post-transcriptional regulation, alternative splicing and RNA editing were also analyzed in *F. graminearum*. In total, 231 genes were identified with two or more alternative splice variants, 86% of which are due to intron retention. Interestingly, for some of the alternatively spiced genes, the expression ratio between different transcript isoforms changed during vegetative growth, suggesting that alternative splicing events are developmentally regulated. No RNA editing events were identified in our study. Finally, 2459 novel transcriptionally active regions (nTARs) were identified in our RNA-seq data, and detailed analysis showed that many of these could be missed genes.

Through the analysis of RNA-seq data, we found that transcriptional activity along the chromosomes of *F. graminearum* is not evenly distributed, and several regions showed significantly lower levels of expression. In **Chapter 4**, we analyzed the features of these lowly expression regions in detail. Firstly, the weakly expression regions on all four chromosomes of *F. graminearum* are exactly matched with the non-conserved regions that were identified by comparing to closely related species *F. verticillioides* and *F. oxysporum*. Secondly, genes that are involved in the production of secondary metabolites and secreted proteins are enriched in the non-conserved regions, whereas house-keeping genes, such as ribosome encoding genes, are enriched in the conserved regions. Thirdly, comparing gene

expression levels between conidia and mycelium showed that expression levels in the non-conserved regions are much more influenced by developmental stage than in the conserved regions. Interestingly, the top 300 genes that displayed the strongest up- or down-regulation between two developmental stages are highly over-represented in the non-conserved regions. Finally, we found that the weakly expressed regions are full of gene relocations, which probably has led to the clustering of genes with similar gene expression patterns or similar biological functions.

Chapter 5 provides a general discussion of our results in a broader context and suggests directions for future studies.

Acknowledgements

I was very lucky to get the opportunity to study at Wageningen University. Even though it took a bit longer to run this road, I finally reached my destination. When I recall the path I have passed, there are so many beautiful memories that I cannot forget and so many people that I should appreciate.

First of all, I am sincerely grateful to my supervisor Dr. Theo van der Lee. It was enjoyable to discuss with you my projects and always you gave me some good ideas to improve my research. I am also grateful to you for revising my presentation slides one by one, and encouraging me to give presentations once there were opportunities. Now I feel much more confident than before when I give presentations. Thank you for your efforts on the published papers and the thesis. Without your support, my research could not have been completed so smoothly. I am also very grateful for your help during my stay in Wageningen. You were the person who helped me shuttle between airport and Wageningen. You were the person who helped me find the suitable accommodation in Wageningen. You were also the person who organized impressive farewells for me when I left Wageningen.

Many thanks to my promotor Prof. Piere de Wit. I was always impressed by your professional help and kindness. I have learned a lot of skills from you, especially in scientific writing. Every time when I sent a manuscript to you, you always carefully revised it and sent it back to me very quickly. Thank you very much for your extensive efforts on my final pieces of PhD thesis, which made it much greater. Also I would like to mention that I had a wonderful stay with you when you visited Beijing.

I also like to thank very much my daily supervisor Dr. Cees Waalwijk. It felt very comfortable to stay with you. You were the person who interviewed me by phone before I came to Wageningen. Thank you for your excellent pronunciation, so I could easily understand your questions and pass the interview favorably. I am also grateful to you for all kinds of help when I stayed in Wageningen.

I am also especially grateful to my supervisor Dr. Dingzhong Tang in China. Thank you very much for giving me the opportunity to study at Wageningen University as a sandwich student. Overseas experience gave me the chance to learn a lot of skills that I probably could not have

learned in China, which enabled me to advance much faster. Thank you again for your supervision and support during my study in China.

I am also grateful to the persons who work at Plant Research International and Wageningen University. Firstly, I would like to thank to Kema Gert for the organization of lab meetings and discussion of my project. I am also grateful to Ineke de Vries, who has retired in 2010, but helped me a lot during my first year stay in Wageningen, with organizing the laboratory and ordering of enzymes and chemicals. Thank you Els Verstappen for your help to perform assays related to genetically modified organisms. Thank you Marga van Gent for your help with ordering experimental items. I would also like to thank other colleagues working at Plant Research International, especially to Sarrah Ben M'Barek, Caucasella Diaz Trujillo, Amir Mirzadi Gohari and Gilvan Ferreira Da Silva. Thank you for all your help in the lab.

I also would like to express my gratitude to my Chinese friends studying at Wageningen University. Getting to know all of you is one of my precious treasures in my life. Specifically, I would like to thank Wei Liu for all of the help and advice I received from you. Thank you Wei Song for all your help during my trip to London. It was a wonderful trip for me in my life. Many thanks to Hui Li and Lemeng Dong for your kindness. Thanks you Huachen Lei and Lisha Zhang for the nice discussions about projects. I also would like to give my thanks to other friends in the Netherlands. Cheng Xu, Fang Jie, Liu Qing, Ma Lisong, Song Chunxu, Ting Hieng Ming, Lin Ya-Fen, Wang Guodong, Xiao Tingting, Zhang Ying, Zhang Zhao and Zhao Yue.

I am very grateful to my parents, who always tried to give me the best education. Thank you for your hard work to bring me up. Thank you for all kinds of supports I received from you.

Finally, I am especially grateful to my wife, Haozhen Nie. We fell in love in 2010 and got married in 2014. Thank you very much for your support during my PhD period. Thank you for taking care of me in my daily life. You are the best gift I have ever received in my life.

Curriculum vitae

Chunzhao Zhao (赵春钊) was born in Wenzhou, Zhejiang Province, China, on February 8, 1985. From 2002-2006, he studied at the Department of Life Science and Biotechnology at Ningbo University, where he received his BSc degree in Biotechnology (Major) and International Economy and Trade (Minor). After graduation in 2006, he joined the laboratory of prof. Dingzhong Tang at the Institute of Genetics and Developmental Biology (IGDB), Chinese Academy of Sciences (CAS), Beijing, for a combined MSc and PhD program. In prof. Tang's laboratory, he studied the role of *Arabidopsis thaliana* EDR1 (enhanced disease resistance 1) in plant innate immunity. In January 2014, he obtained his PhD degree in Genetics at the IGDB. In 2009, he was also selected to enroll the joint PhD sandwich fellowship program between the Chinese Academy of Science and the Royal Netherlands Academy of Arts and Sciences (CAS-KNAW) to perform a PhD study at Wageningen University, The Netherlands. In Wageningen, he was supervised by Dr. Theo van der Lee, at Plant Research International and prof. Pierre de Wit at Wageningen University, and by prof. Tang at CAS, where he studied virulence genes and genome-wide gene expression in *Fusarium graminearum*, which are presented in this thesis. Currently he works in Dr. Jian-Kang Zhu's lab at Purdue University as a post-doctoral researcher studying abiotic stress in *A. thaliana.*

List of Publications

Publications in Wageningen:

Zhao, C., Waalwijk, C., de Wit, P.J.G.M, van der Lee, T.A., and Tang, D. (2011). EBR1, a novel Zn(2)Cys(6) transcription factor, affects virulence and apical dominance of the hyphal tip in *Fusarium graminearum*. Mol. Plant Microbe Interact. *24*, 1407-1418.

van der Fels-Klerx, H.J., de Rijk, T.C., Booij, C.J., Goedhart, P.W., Boers, E.A., **Zhao, C.,** Waalwijk, C., Mol, H.G., and van der Lee, T.A. (2012). Occurrence of *Fusarium* Head Blight species and *Fusarium* mycotoxins in winter wheat in the Netherlands in 2009. Food Addit. Contam. Part A: Chem. Anal. Control Expo. Risk Assess. *29*, 1716-1726.

Zhao, C., Waalwijk, C., de Wit, P.J.G.M, Tang, D., van der Lee, T.A. (2013) RNA-Seq analysis reveals new gene models and alternative splicing in the fungal pathogen *Fusarium graminearum*. BMC Genomics *14*, 21.

Zhao, C., Waalwijk, C., de Wit, P.J.G.M., Tang, D., and van der Lee, T.A. (2014) Relocation of genes generates non-conserved chromosomal segments in *Fusarium graminearum* that show distinct and co-regulated gene expression patterns. BMC Genomics *15*, 191.

Other publications:

Nie, H., **Zhao, C.,** Wu, G., Wu, Y., Chen, Y., and Tang, D. (2012). SR1, a calmodulin-binding transcription factor, modulates plant defense and ethylene-induced senescence by directly regulating *NDR1* and *EIN3*. Plant Physiology 158, 1847-1859.

Zhao, C., Nie, H., Shen, Q., Zhang, S., Lukowitz, W., and Tang, D. (2014) EDR1 physically interacts with MKK4/MKK5 and negatively regulates a MAP kinase cascade to modulate plant innate immunity. PLoS Genet. 10, e1004389.

Zhao, C., Lang, Z., and Zhu, J.K. (2015) Cold responsive gene transcription becomes more complex. Trends Plant Sci. *20*, 466-468.

Education Statement of the Graduate School

Experimental Plant Sciences

The Graduate School EXPERIMENTAL PLANT SCIENCES

Issued to: Chunzhao Zhao
Date: 8 December 2015
Group: PRI - Biointeractions & Plant Health / Laboratory of Phytopathology
University: Wageningen University & Research Centre

1) Start-up phase	*date*
► **First presentation of your project**	
Functional genomics in Fusarium	Apr 02, 2009
► **Writing or rewriting a project proposal**	
Isolation and Characterization of Fusarium Mutants altered in virulence	Apr 2009
► **Writing a review or book chapter**	
► **MSc courses**	
► **Laboratory use of isotopes**	
Subtotal Start-up Phase	*7.5 credits**

2) Scientific Exposure	*date*
► **EPS PhD student days**	
EPS PhD student day, University of Amsterdam	Nov 30, 2012
4th European Retreat for PhD Students in Plant Sciences, Norwich, UK	Aug 15-17, 2012
EPS theme symposia	
EPS Theme 2 symposium & Willie Commelin Scholten Day 'Interactions between Plants and Biotic Agents, Utrecht University	Jan 22, 2009
EPS Theme 4 symposium 'Genome Biology', Radboud University	Dec 07, 2012
► **NWO Lunteren days and other National Platforms**	
► **Seminars (series), workshops and symposia**	
Seminar series 'Phytopathology'(estimated 3X/year)	2009, 2012
Magnaporthe grisea in rice and other cereals	Apr 14, 2009
Cellular Signaling During Pollen Tube Growth	Jan 05, 2010
Prospects for engineering nitrogen fixation in crops	29 Jan 2010
Every step from a receptor kinase to a thousand target genes- the brassinosteroid signaling pathway	Feb 01, 2010
Functional Dissection of LRR-Kinase Receptors and Plant Defense Signaling.	Mar 11, 2010
The role of ABA and bacterial effectors in suppression plant defense responses	Mar 16, 2010
Grass is greener on the other side-biotechnology and gene discovery for genetic improvement of turf species	Apr 08, 2010
AgGTL1 regulates transcription and water-use efficiency by controlling stomatal number through transcriptional repression of SDD1	Apr 13, 2010
Keeping nodules sweet: sugar supplies for nitrogen fixation in the legume, Lotus japonicus	Apr 14, 2010
Wild emmer as a source of valuable alleles for wheat	Apr 14, 2010
Auxin regulation of its efflux carrier PIN2 trafficking and turnover in Arabidopsis	Apr 23, 2010
Epigenetic reprogramming during plant reproduction	May 13, 2010
Rice genomics - from sequencing to gene function	May 20,2010
A calcium sensor/protein kinase network for decoding calcium signals in plants.	Jun 30, 2010
Interception of fungal feeding structures by plant resistance proteins at the host-pathogen interface	Jul 15,2010
MAC: an evolutionarily conserved protein complex essential for plant immunity	Aug 05, 2010
Genome Evolution in Wheat and Related Species	Oct 14, 2010
A novel SCF complex which interacts with the COP9 signalosome and regulates light response in Arabidopsis	Oct 20, 2010
Signaling during pollen tube growth	Nov 02, 2010
Association mapping for enhancing crop genetics improvement	Nov 05, 2010
Rapid origination of essential functions: The role of new genes in development and network evolution	Nov 11, 2010
Quantitative molecular methods and High Through Put Fluorescent Imaging to Phenotype Wheat Resistance to Fusarium Head Blight	Nov 16, 2010
Dare to be different: Polarity In Control of Asymmetric Division	Dec 13, 2010
Recent studies of potassium channels and calcium-activated chloride channels	Apr 20, 2011
New Gene Origination: Mechanism, Pattern and Function	Apr 25, 2011
Biochemical, Molecular and Genetic Regulation of Sorghum Aluminum Tolerance Involving the Root Citrate Transporter, SbMATE	Apr 27, 2011
Polar Auxin Transport and Plant Growth Responses	May 09, 2011
Unraveling Nitrate Signaling Via Genetics and Genomics	May 11, 2011
Arabidopsis MPK3/MPK6 in plant disease resistance: Role of ethylene in the process	Jul 04, 2011
Mechanisms in the plant circadian clock	Jul 07, 2011
Understanding the interplay between plant immune responses and other cellular processes	Jul 15,2011
RNA silencing and stem cell development in Arabidopsis	Aug 05, 2011
Small RNA and RNAi machinery in plant immunity	Aug 08, 2011
Molecular mechanisms underlying plant-bacteria interactions	Aug 09, 2011
Applications of protein microarray in construction of networks and pathways	Oct 25, 2011
Applications of whole-genome sequencing	Feb 07, 2012
► **Seminar plus**	
► **International symposia and congresses**	
Spring meeting KNPV, Wageningen, NL	May 25, 2009
22nd New Phytologist Symposium: Effectors in plant-microbe interactions, Versailles, France	Sep 13-16, 2009
The 2nd International Conference on biotic Plant Interactions, Kunming China	Nov 12-16, 2011
11th European Conference On Fungal Genetics, Marburg, Germany	Mar 30-Apr 02, 2012
KNPV FUSARIUM-workinggroup, CBS Utrecht, NL	Oct 31, 2012
The 3rd International Conference on biotic Plant Interactions, Yangling China	Aug 18-22, 2013

►	**Presentations**	
	2nd International Conference on biotic Plant Interactions, Kunming China (Poster)	Nov 12-16, 2011
	Fusarium workshop in Marburg (Talk)	Mar 30, 2012
	11th European Conference On Fungal Genetics, Marburg, Germany (Poster)	Mar 30-Apr 02, 2012
	4th European Retreat for PhD Students in Plant Sciences, Norwich, UK (Poster)	Aug 15-17, 2012
	Fusarium workshop in Utrecht (Talk)	Oct 31, 2012
	The 3rd International Conference on biotic Plant Interactions, Yangling China (Poster)	Aug 18-22, 2013
►	**IAB interview**	
	Meeting with a member of the International Advisory Board of EPS	Nov 14, 2012
►	**Excursions**	
	Paris-sud	Jan 28, 2009
	PPO lisse	Mar 09, 2009
	Denmark	May 19-20, 2009
	Martijn Rep's lab in Amsterdam University	Sep 28, 2012
	Paris-sud	Dec 2012
	Subtotal Scientific Exposure	*20.8 credits**

3) In-Depth Studies		*date*
►	**EPS courses or other PhD courses**	
	5th Utrecht PhD Summer school on Environmental Signaling	Aug 24-26, 2009
	Systems Biology: Statistical Analysis of ¬Omics Data, Wageningen, NL	Dec 10-14, 2012
►	**Journal club**	
	Literature discussion at PRI	2009
	Literature discussion at IGDB, China	2010-2011
►	**Individual research training**	
	Subtotal In-Depth Studies	*3.9 credits**

4) Personal development		*date*
►	**Skill training courses**	
	Mobilising your - scientific - network	May 12 & 26, 2009
	Project- and Time Management	Sep 08, 22 and Oct 20, 2009
	NWO Talent Days	Oct 06, 2009
	Techniques for Writing and Presenting a Scientific Paper	Dec 15-19, 2009
►	**Organisation of PhD students day, course or conference**	
►	**Membership of Board, Committee or PhD council**	
	Subtotal Personal Development	*4.0 credits**

TOTAL NUMBER OF CREDIT POINTS*	**36.2**

Herewith the Graduate School declares that the PhD candidate has complied with the educational requirements set by the Educational Committee of EPS which comprises of a minimum total of 30 ECTS credits

** A credit represents a normative study load of 28 hours of study.*