Supplementary Information: ***Comparative genomics reveals mobile pathogenicity chromosomes in Fusarium oxysporum***

Outline:

A. Genome sequencing, assembling and mapping.

B. Annotation.

B.1. Repetitive sequences and transposable elements.

B.2. Predictions protein coding genes.

***B.3. Characterization of protein families*.**

C. Features common to *Fusarium* genomes.

***C.1. Transcription factors*.**

***C.2. CAZY enzymes*.**

C.3. Transporters.

C.4. Signaling.

D. Secondary metabolism gene clusters

E. Features of *Fusarium oxysporum* LS chromosomes.

E.1. Synteny and Fol LS chromosomes.

E.2. Variation in Fol LS regions among Fo strains.

***E. 3. Transposable elements*.**

E. 4. Segmental duplication.

E.5. Genes and gene categories expanded on LS chromosomes.

E.6. RIP and other silencing pathways.

F. Transfer of LS chromosomes between strains.

F.1. Transformation of* Fusarium *strains

F.2. Co-incubation of spores.

F.3. Bioassays.

F.4. PCR Amplications

F.5 . CHEF and Southern blotting.

A. Genome sequencing, assembling and mapping.

An 8X whole genome shotgun (WGS) assembly of *Fv* strain 7600 was generated combining genomic sequences from the Broad Institute and Syngenta Biotechnology Inc. (Syngenta). A 6.28X *Fo* assembly was generated at the Broad Institute also using WGS approach (Table S1). Both genomes were sequenced using Sanger technology and assembled using Arachne [1].

A total of 142 RFLP markers comprising the *F. verticillioides* (*Fv*) genetic linkage map [2] were sequenced in both directions at the Broad. Both forward and reverse sequences were treated as independent markers, resulting in 284 marker sequences. More than 99% of the assembly was anchored to the genetic map by locating the position of the sequence of 131 marker sequences within the current assembly using BlastN. The majority of unplaced markers on the map were caused by incomplete sequence data.

An optical map of *F. oxysporum f. sp. lycopersici* from the sequenced strain was constructed using the restriction enzyme *Bsi*WI [3], consisting of 15 single optical map contigs (~63 Mb) with an average of 55X coverage. Alignments were made between optical maps and the *in silico* maps of the sequence scaffolds based on the restriction recognition sequence and the lengths of the restriction fragments using map aligner software developed at the Broad Institute. The optical maps and sequence assemblies were highly congruent with more than 96% of the assembled scaffolds aligned to the optical map. The assembled sequence scaffolds were also ordered and oriented, and the gaps were estimated based on the alignment of the assembled sequence scaffolds to the optical maps (Table S2). Overall, more than 91% of the optical maps were covered by assembled sequence scaffolds.

The genetic linkage group maps for *Fv* and the optical linkage group maps for *Fol* can be accessed at:http://www.broad.mit.edu/annotation/genome/fusarium_group.3/maps/Index.html

B. Annotation.

B.1. Repetitive sequences and transposable elements.

Repeat sequences were detected by searching the genome sequence against itself using CrossMatch, filtering for alignments longer than 200 bp with greater than 60% sequence similarity. Full-length transposable elements were annotated using a combination of computational predictions based on BLAST analysis for transposase genes and manual

inspection. Their genome distribution was characterized using the sensitive mode of RepeatMasker version open-3.0.8, with cross_match version 0.990329, RepBase Update 9.04, RM database version 20040702.

B.2. Predictions of protein coding genes.

To improve annotation accuracy, the gene structures of the *Fol*, *Fv,* and *F. graminearum (Fg)* genomes were predicted with a combination of manual annotation, FGENESH, GENEID, and FindORFs. The manually curated gene sets were generated at the Broad Institute and used to train the *ab initio* gene prediction programs FGENESH and GENEID. The genome sequences of the three *Fusarium* species were then analyzed with both FGENESH and GENEID to provide an initial set of predicted genes. ESTs from *Fg* (68,000) and *Fv* (87,000) were analyzed with FindORFs, an EST-based gene finding program developed at the Broad Institute, to identify mis-annotated gene loci and improve gene structure predictions. In addition, the final gene set of *Fg* was mapped onto the *Fv* and *Fol* genomes by FindORFs to provide additional lines of evidence to improve the automated gene prediction in these two fungi. For *Fg*, the gene set predicted at MIPS [5] also was included to improve gene annotation. Furthermore, targeted manual annotation was carried out for loci in cases where automated gene predictions were inconsistent or where there was conflict with EST evidence or BlastX analysis. The re-annotation of *Fg* greatly improved the gene structures and greater than 90% accuracy was reached when comparing this annotation against a set of 4000 high confidence genes.

***B.3.* Characterization of protein families.**

Transcription factors were identified using the data analysis pipeline in the Fungal Transcription Factor Database (FTFD; http://ftfd.snu.ac.kr/) [4]. **Carbohydrate-active enzymes** were identified by searching against a library of catalytic and carbohydrate-binding modules of carbohydrate-active enzymes (see also www.cazy.org) [5] using blastp. T**ransporters** were identified by searching for the PFAM domains representing ABC transporters (PF00005, PF00664, PF01061) and assisted with manual inspection. The potential **secreted proteins** were predicted using SignalP 3.0 D score (http://www.cbs.dtu.dk/services/SignalP/). The trans-membrane proteins were removed using the TMHMM software (http://www.cbs.dtu.dk/services/TMHMM/) and the Phobius predicted transmembrane proteins (except in the first 50 amino acids). We further removed putative mitochondrial proteins using TargetP mitochondrial target peptide prediction RC score 1 or 2. **Small cysteine-rich proteins**

were identified among predicted secreted proteins that are less than 200 aa in length and contain at least 4% cysteine residues. **GPI-anchored proteins** were identified by the GPI-anchor attachment signal with a PERL script among the predicted secreted proteins. **Kinomes** were characterized using a multi-level hidden Markov model (HMM) library of the protein kinase superfamily under the HMMER software suite (v. 2.3.2, http://hmmer.janelia.org), correcting for database size with the '-Z' option. The automatically-retrieved sequences were individually inspected and protein kinase homologies were determined by building kinase group-specific phylogenetic trees with the annotated kinomes of *S. cerevisiae*, *S. pombe* and *Encephalitozoon cuniculi* [6]. **G proteins** were identified from the best hits using the corresponding *N. crassa* protein, using a cut-off score of 1e-12 during BLAST searches. For **G protein coupled receptors**, the proteins listed are the best hits (E-value cutoff: 1e-09) using the *N. crassa, M. grisea* or *F. graminearum* proteins as queries in BLASTp searches. For GprK homologues *A. nidulans* AN7795, *N. crassa* NCU09883 and *F. graminearum* FGSG_04628 were used for searches (E-value cut-off 1e-09) with BLASTp. All GPCR hits have 7-transmembrane (7-TM) helices, with the N-terminus outside and the C-terminus on the cytoplasmic side, as verified by HMTMM (http://www.cbs.dtu.dk/services/TMHMM-2.0/), TMPRED (http://www.ch.embnet.org/software/TMPRED_form.html), PHOBIUS (http://phobius.binf.ku.dk/) or TOPPRED (http://bioweb.pasteur.fr/ seqanal/interfaces/toppred.html), except as noted in parentheses (supplemental results). Deviations from the 7-TM structure may be due to mis-annotation or sequencing errors and many of the e values were much lower than 1e-09. **Pathogenicity and virulence factors** were identified using the pathogen-host interactions database [8] (http://phi-base.org) and through additional searches of the peer-reviewed literature.

C. Features common to *Fusarium* genomes.

Orthologous gene sets were determined based on pair-wise syntenic alignments that projected genes from one genome onto the other. All the data are available through the website: http://www.broad.mit.edu/annotation/genome/fusarium_group.3/OrthologSearch.html

The protein sequences for *Fusarium solani* [9] (telomorph: *Nectria haematococca*), used as outgroup for phylogenetic relationships among the *Fg, Fv,* and *Fol* genomes, was downloaded from the Joint Genome Institute (http://genome.jgi-psf.org/Necha2/Necha2.home.html). For a full list of annotated genes for each category, see web site: http://www.broad.mit.edu/annotation/genome/fusarium_group/SupplementaryPage.html.

Compared to other Ascomycete genomes, these three-species orthologs are enriched for predicted

transcription factors (p=2.6e-6), lytic enzymes (p=0.001), and transmembrane transporters (p=7e-9) (Tables S3-8).

C.1 Transcription factors.

We identified 2,280 transcription factors (TFs) in the three *Fusarium* genomes, which were classified into 46 TF families (Table S4). For a full list of annotated genes for each category, see: http://www.broad.mit.edu/annotation/genome/fusarium_group/SupplementaryPage.html. The most abundant TF family in *Fusarium* genomes is Zn2Cys6, a family of fungal specific transcriptional regulatory proteins that contain an N-terminal Cys-rich motif and play essential roles in both primary and secondary metabolism, drug resistance, and meiotic development [10]. The numbers of TFs in most families were similar across the three species. However, the numbers of TF families of bHLH, Zn2Cys6 and bZIP were significantly higher in *Fol* than in the other species. For example, in the bHLH TF family, *Fol* has 46 TFs, while the other species have only 16 (*Fg*) or 15 (*Fv*). Similarly, the number of TFs in the bZIP family was 55 in *Fol*, whereas those in *Fg* and *Fv* are 21, and 19, respectively. The higher numbers in these classes as well as higher numbers of homeodomain-like transcription factors can be explained by the addition of genes in *Fol* LS regions.

C.2. CAZY enzymes.

We conducted a comprehensive analysis of various carbohydrate-active enzymes (glycosidases, glycosyltransferases and polysaccharide lyases) encoded by the three *Fusarium* (Table S3) and other fungal genomes in a standardized manner to allow objective comparisons. For a full list of annotated genes for each category, see web site: http://www.broad.mit.edu/annotation/genome/fusarium_group/SupplementaryPage.html.

Among the 13,332 predicted genes for *Fg*, 14,179 for *Fv* and 17,735 for *Fol*, we detected, respectively, 243, 300, and 364 genes encoding glycoside hydrolases (GHs) and 102, 106, and 123 genes encoding glycosyltransferases (GTs). Overall, the three fungi have a complete set of plant cell wall degrading enzymes (including enzymes for the digestion of three major categories: cellulose, hemicelluloses and pectins). Notably, there are more pectin and pectate lyases from families PL1 and PL3 in *Fusarium* genomes than in other fungi. Several GH families have greatly

expanded in *Fol*. In comparison to the highly variable number of GHs, the number of GTs and the population of each GT family is much more constant. This observation is in line with the fact that while most GHs are secreted to the environment, and therefore co-evolved with the ecological niche of the fungi, GTs are dedicated to tasks that are less dependent on the outside medium, such as synthesis of the fungal cell wall polysaccharides, N- and O-glycosylation of proteins, and metabolism of storage carbohydrates.

The over-representation of these lytic enzymes is the combination of the increase of copy numbers of some gene family as well as the unique presence of some gene families. For instance, all three *Fusarium* genomes contain the highest number of pectate lyases (21,23, 24) comparing to less than 10 in other genomes. The PL family uniquely present in *Fusarium* genomes includes PL9, while the expansion is observed for families PL1, PL3, and PL4 (Table S5). The detailed comparison can be found at online supplementary data at: http://www.broadinstitute.org/annotation/genome/fusarium_group/SupplementaryPage.html

C.3 Transporters:

All three *Fusarium* species have a considerably higher number of ABC transporters than other fungi (Table S3). Typically, most fungi have 30–45 ABC transporters encoded in their genomes, regardless of whether they exist as a yeast, saprophyte, or plant/animal pathogen. However, the number of *Fusarium* ABC transporters ranges from 54 in *Fg* to 67 in *Fol* (Table S3). All five families of ABC transporters are represented in the *Fusarium* (Table S6). *Fol* has an 18% increase in the number of ABC transporters when compared to the number of transporters in *Fv* and an even larger increase when compared to *Fg*. For example, *Fol* contains four copies of a partial ABC transporter in the *Fol* LS regions (FOXG_16340.2, FOXG_07314.2, FOXG_06754.2, and FOXG_06510.2). Two of these copies are on the same chromosome (chromosome 3), while one is on chromosome 6, and the fourth is on the unpositioned supercontig 34. The larger number of ABC transporters may explain the increase in host range of *Fo* as a species as they may provide tolerance to antimicrobials generated by different hosts.

There are five ABC transporters present in all the sequenced *Fusarium* isolates, which are not present in any other sequenced fungal genome (Table S6, colored in red). The genes surrounding some of the ABC transporter genes are conserved and so may bring insight into their function. Since the genes are orthologous, only the *Fg* name will be used for this discussion. FGSG_11028.3 is similar to YOR1 transporters, residing in a cluster of genes, which are co-

expressed in plants [11] and contain the gene *NPS1* which encodes a siderophore that has yet to be chemically characterized [12.] The juxtaposition of the two genes suggests a scenario in which the ABC transporter may be partially responsible for efflux of the siderophore. Similarly, both FGSG_11028.3 and another Fusarium specific ABC transporter, FGSG_06141.3, are adjacent to a gene for a major facilitator transporter in all three *Fusarium* sp. The other two *Fusarium*-specific transporters (FGSG_03805.3 and FGSG_03032.3) are similar to heavy metal tolerance transporters (HMT). FGSG_10577.3 is a Pleiotropic Drug Resistance (PDR) protein transporter unique to *Fusarium*.

The family of PDR transporters (also referred to as the ABCG family) is largely responsible for tolerance to otherwise toxic compounds, including azoles, cyclohexamide, and plant secondary metabolites [13]. Moreover, the only pathogenicity related *Fusarium* ABC transporters, *F. sambucinum* GpABC1, which is responsible for tolerance to the potato antimicrobial compounds rishitin and lubimin [14], and *F. culmorum* FcABC1 [15] belong to this family. All three sequenced isolates have orthologs of this protein (FGSG_04580.3, FVEG_11089.3, and FOXG_13653.2). This family has more members in each of the *Fusarium* sp. than most other fungi (Table S3), such as *Neurospora crassa* (7 members), *Magnaporthe grisea* (8), *Aspergillus fumigatus* (15), and *A. nidulans* (16). In addition to the overall large number of PDR genes in all three *Fusarium* genomes, there is a 26% increase in these transporters in the broad host range *Fol* when compared to the *Fv* and *Fg*. The increased number of these proteins in all *Fusarium* species may explain why they are not typically inhibited *in vitro* with miconazole, itraconazole, or fluorocytosine [16].

C.4 Signaling.

Previous studies have demonstrated high conservation of G protein subunits and predicted G-protein coupled receptors (GPCRs) in filamentous fungi [17]. The three *Fusarium* species contain one protein corresponding to each of the three classes of Gα subunits of heterotrimeric G proteins found in other filamentous fungi (Table S7). However, *Fol* possesses an additional protein (FOXG_08807) with weak homology to the Group II Gα subunit. This protein is not similar to Gpα4 from *Ustilago maydis* [18], raising the possibility that certain filamentous fungi have evolved diverse Gα proteins. In *N. crassa*, the Group II Gα *GNA*-2 is required for mass accumulation on poor carbon sources and also appears to play a compensatory role in the absence of the Group I or Group III subunit [19]. Therefore, the additional Group II subunit in *Fol* may be required for growth on alternative carbon sources and may also provide additional genetic buffering, should

the Group I or Group III subunit genes become nonfunctional. *Fg*, *Fol* and *Fv* have two Gβ subunits; one homologous to canonical Gβ proteins identified in other filamentous fungi and another similar to *N. crassa CPC*-2, a RACK-1 homologue (Table S7). Like other filamentous fungi, the three *Fusarium* species also contain one Gg subunit. These three species do not possess homologues of the Gβ (*Krh1*/*Krh2*) or Gg (*Gpg1*) mimics identified in *S. cerevisiae*. RGS (Regulator of G protein signaling) proteins are negative regulators of Gα proteins in eukaryotes. *A. nidulans*, *N. crassa*, and *M. grisea* carry one copy of the RGS protein *FlbA*. In contrast, two copies of *FlbA* were identified in *Fg* and *Fv* genomes, and, notably *F. oxysporum* possesses eight *FlbA* -like genes.

The three *Fusarium* genomes all possess the nine classes of seven transmembrane-helix (7-TM), predicted G protein coupled receptors (GPCRs) present in filamentous Ascomycetes [17] (Table S8). The number of pheromone, carbon-sensing, *Stm*-1 related and rat growth hormone releasing factor-related GPCRs are similar in the three *Fusarium* species and other sequenced filamentous fungi. However, the *Fusarium* genomes contain more members of several other classes of predicted GPCRs. The *Fusarium* species possess a greater number of cAMP receptor-like GPCRs; *Fg* and *Fv* have four and *Fol* has five homologous proteins (Table S8). While *Fg* only has two homologues of *Homo sapiens* membrane progestin receptor (mPR)-like GPCRs, *Fv* has three and *Fol* has seven. The mPR receptors couple to G proteins on the plasma membrane and down-regulate adenylyl cyclase in response to progesterone exposure in reproductive tissues of animals [20]. The GprK/AtRGS1-like GPCR, which has 7-TMs and a RGS domain, has not been demonstrated to function as a GPCR in fungi [7,17]. Nevertheless, *Fg* has one GprK homologue and *Fol* and *Fv* each have two proteins. The *Fusarium* species possess a greater number of microbial opsins/opsin-related proteins [21] than other filamentous fungi: three in *Fg* and *Fv* and four in *Fol*. Finally, proteins homologous to the pathogenesis-related *PTH11* GPCR are abundantly present in all three *Fusarium* species, with 99 in *Fg*, 98 in *Fol* and 90 in *Fv*, far more than those in *Magnaporthe grisea* (60), *Aspergillus nidulans* (70) and *N. crassa* (25).

D. Secondary metabolism gene clusters

Fusarium produces numerous secondary metabolites that are diverse in biosynthetic origin, structure and biological activity. Some exhibit toxicity to mammals and as a result are considered mycotoxins. The two *Fusarium* mycotoxins of greatest concern are trichothecenes and fumonisins. Trichothecenes inhibit eukaryotic protein synthesis and induce hemorrhaging and

feed refusal in livestock [22]. Fumonisins disrupt sphingolipid metabolism, cause multiple animal diseases [23], and are epidemiologically associated with esophageal cancer and neural tube defects in humans [24,25]. Thus, the presence of these mycotoxins in cereal grains and grain-based food and feed poses a potential hazard to human and animal health. *Fg* and *Fv* are among the most important mycotoxin-producing *Fusarium* species because of their worldwide occurrence and because they produce trichothecenes and fumonisins, respectively.

In fungi, genes responsible for secondary metabolite biosynthesis are often clustered. The clusters typically include a terpene synthase (TS), polyketide synthase (PKS), or nonribosomal peptide synthetase (NRPS) gene that is responsible for a fundamental step in metabolite biosynthesis. Clusters also typically include the core genes responsible for structural modifications of initial metabolite, for metabolite transport, and for coordinated transcriptional regulation of cluster genes.

Few secondary metabolite biosynthetic gene clusters have been identified and functionally characterized in *Fusarium*. Given the structural diversity of secondary metabolites produced by the genus as a whole, many more *Fusarium* clusters are likely to be present in the genomes of these fungi. Here, genome sequences of *Fg*, *Fo*, and *Fv* were used to identify potential genes clusters associated with PKS and TS genes. Such analyses may provide insight into conservation of secondary metabolic pathways among species, into enzyme and cluster evolution, and into the functions of secondary metabolites in the ecology of producing species.

To identify secondary metabolite biosynthetic (SMB) gene clusters in *Fusarium*, an automated prediction program was developed based on the presence of pfam domains described for PKSs and other proteins involved in fungal secondary metabolism. A significant initial effort was made to identify all PKS genes present in each of the four *Fusarium* genomes. This effort used a combination of BLAST and text search analyses of preliminary annotations at the Broad and JGI websites. The BLAST query sequences were predicted keto synthase domains from reducing and non-reducing (NR) PKSs, the two major classes of fungal PKSs. These analyses provided a preliminary list of putative KS-encoding DNA from the four species. Nucleotide sequences flanking the putative KS-encoding DNA were subjected to BLASTX analysis and manually annotated to identify the full-length PKSs. Since most secondary gene clusters contain genes with common domains related to Transporter; Cytochrome P450; oxidoreductase; Zn(2)-Cys(6) transcription factor; Oxidoreductases; FAD binding domain; Transporter; O-methyltransferase; Transferase family, we searched the putative SMB gene clusters by pfam domains analysis of 20

kb flanking either side of each PKS genes. PKS gene-flanking regions that included minimum of four pfam domains were considered putative SMB clusters and subjected to microarray analysis using available microarray data from both *Fg* and *Fv*. Conservation of the putative clusters among the four sequenced *Fusarium* genomes was used to further define the cluster boundaries. A similar approach was also taken to identify terpenoid biosynthetic gene clusters. Briefly, BLAST analysis was used to identify putative terpene synthetase (TS) genes, and the regions flanking these genes were then subjected to the pfam analysis described above.

Phylogenetic relationships of *Fusarium* PKSs within and among species were examined by maximum parsimony analysis of deduced amino acid sequences of the ketosynthase (KS) and acyl transferase (AT) domains of all the PKS genes identified in the *Fusarium* genomes. Because of a lack of conservation in 95 to160-amino acid region between the KS and AT domains, this inter-domain region was removed from the alignment. The aligned sequences were then used to construct a gene genealogy using parsimony in PAUP* 4.0b10 [26]. Bootstrap values were based on 1000 pseudoreplications. Phylogenetic trees generated from deduced amino acid sequences of other functional domains in the PKSs were very similar to each other and to the tree based on the combined KS and AT domains.

A total of 39 PKS and 7 TS related SMB gene clusters were identified (Table S9), including 18 from *Fg*, 16 from *Fv* and 12 from *Fol*. The average cluster contains eight genes and spans over 24 kb. Despite its increased genome size and large protein encoding gene set, *Fol* has the least SMB gene clusters identified through this approach.

Microarray analysis confirmed co-regulation of genes in 14 of 18 *Fg*, and 10 of 16 *Fv* clusters (online materials). Differential expression of *Fg* clusters was examined under multiple conditions, including during sexual reproduction, infection of multiple plant hosts, and growth in complete medium. *Fg* SMB clusters expressed *in planta* include the previously described trichothecene (FG3_21) and fusarin C clusters (FG3_19). Six SMB clusters (FG3_15, FG3_20, FG3_28, FG3_35, FG3_40 and FG3_45) were also preferentially expressed during perithecial formation and sexual reproduction. Cluster FG3_33 is responsible for biosynthesis of aurofusarin, a red pigment commonly observed in laboratory cultures of *Fg*. Indeed FG3_33 cluster genes were expressed exclusively in culture, and closely related orthologs of cluster genes were absent in the genomes of the other fusaria examined. Two clusters (FG3_12, FG3_13) are constitutively expressed and four, including the zearalenone cluster (FG3_34) were poorly expressed under all

conditions tested (online materials). The lack of expression of this known cluster is attributable to the fact that zearalenone inducing conditions were not included among microarray tests.

Among the identified PKS SMB gene clusters, only 2 (FG3_15, and FG3_28) appear to be conserved in all species examined (Table S9). This finding is consistent with the limited similarity shared by the PKSs (Figure S5). Surprisingly, both of the orthologous PKS clusters (FG3_15, and FG3_28) are preferentially expressed during perithecial formation. Cluster FG3_15 includes the PKS gene *PGL1* responsible for biosynthesis of the dark pigment in the perithecial wall in *Fg* and *Fv* [27]. It is interesting to note that *F. solani* has orthologs of *PGL1* (JGI_101778) and of cluster FG3_15. However, another non-reducing PKS gene, *PKSN* (JGI33672), is responsible for the red perithecial pigmentation in this species [28]. The conservation of the *PGL1* cluster in the asexual species *Fol* places in doubt the direct functional role of this cluster in the sexual reproduction processes. However, the *pgl1* mutants form un-pigmented perithecia with viable ascospores demonstrating that the dark pigments produced through such pathway is not fundamentally essential for sexual reproduction/perithecial formation, and instead may protect the organism from UV light, stress, and oxidative damage. We also noticed that one centrally located gene unique to the FG3_15 cluster, FGSG_09185 (adenylosuccinate lyase), is constitutively expressed and not regulated with the other genes of the cluster. This suggests that the cluster may have acquired this gene relatively recently, subsequent to the evolution of co-regulation of the rest of the cluster.

Interestingly, the non-reduced PKS encoding SMB gene clusters are uniquely expanded in the *Fg* genome (5, 2, 2 in *Fg*, *Fv*, and *Fol* respectively). This group of PKS gene clusters includes the previously characterized biosynthetic clusters for the mycotoxin zearalenone (FG3_34) and the red pigment aurofusarin (FG3_33) [27,29]. One *Fg* specific expanded gene cluster (FG3_26) include 17 genes, are expressed specific *in planta* (Figure S5 and online materials). In contrast, clade III of the reducing PKS encoding clusters including the known fumonisin biosynthetic gene cluster (FV3_21) is uniquely expanded in *Fv*, (1:5:3 in *Fg, Fv* and *Fol* respectively). The two *Fv* specific clusters have 8 and 13 genes identified and span 24 and 33 kb. The only *Fg* clade III cluster is expressed during sexual development [30]. The second clade of the reduced PKS gene clusters has expansion in the *Fv* and *Fol* lineages over *Fg* (2:4:4), including three orthologous clusters conserved between *Fv* and *Fol*. Orthologs of the putative fusarin C biosynthetic cluster [31] is present in *Fg*, *Fv*, and *Fs*, but absent in *Fol*.

There is a clear expansion of terpenoid related SMB gene clusters in *Fg* (5, 1, 1 in *Fg*, *Fv*, and *Fo* respectively), including the well-characterized trichothecene biosynthetic cluster FG3_21. The co-regulation of several of the *Fg* TS clusters are confirmed by expression analysis (Table S9 and online materials).

Species within genus *Fusarium* have evolved extraordinary chemical diversity and produce highly complex and biologically active secondary metabolites. The function of most SMB clusters described here are unknown. These novel SMB clusters can be used to discover new secondary metabolites and their potential impacts on the biological properties of the fungi. The co-expression of most of such clusters, especially the novel clusters, supports their potential functionality and will promote the study of these novel secondary metabolites. Only a very small fraction of the identified SMB clusters are shared among species, reflecting great genetic diversity within the *Fusarium*.

E. Features of *Fusarium oxysporum* LS chromosomes

E.1 Synteny and Fol LS chromosomes.

We performed multiple genome alignments to detect synteny and large-scale genomic re-arrangements in order to understand genome evolution among the three *Fusarium* genomes. PatternHunter [32] (1e-10) was used to compute local-alignment anchors. Pairs of anchors separated by fewer than 10,000 nucleotides on both genomes were merged progressively to form co-linear blocks, which were further filtered to ensure that no block overlaps another block by more than 90% of its length. The average sequence similarity between each pair of genomes was calculated based on the syntenic alignments in 100 base pair windows (Table S10).

Over 90% of the *Fv* genome can be unambiguously aligned to the orthologous regions in *Fol* with an average 90% sequence identity (Table S10, Figure 2). The pair-wise genome comparison reveals that all eleven chromosomes in *Fv* have corresponding chromosomes in *Fol* with minimal local rearrangements and only one major chromosomal translocation event (Figure 2). In contrast, four of the *Fol* chromosomes (chr3, chr6, chr14, and chr15), parts of chromosome 1 and 2 (scaffold 27, scaffold 31), and most of the unmapped scaffolds, lack significant orthologous sequence in the closely related genomes (Figure 2), defining the *Fol* LS regions.

The smaller number of chromosomes in *F. graminearum* appears to be a result of chromosome fusion. Whole genome alignments of *F. graminearum* to *F. verticillioides* reveals that the vast majority (86% overall) of each *F. verticillioides* chromosome maps to a single *F. graminearum*

chromosome (Figure S3). The alignments display end to end synteny in large blocks, with the exception of *F. verticillioides* chromosome ends. The previously described highly polymorphic regions of *F. graminearum* [33] correspond to *F. verticillioides* chromosome ends, including internal fusion sites. This suggests that the fused subtelomeric regions in *F.graminearum* continue to evolve similarly to conserved subtelomeres, and that a sequence or epigenetic modification continues to mark the fused sites similarly to subtelomeres.

E.2 the variation of Fol LS regions between strains within the F. oxysporum species complex.

We have used both Illumina sequence of *F. oxysporum* strain Fo5176, a pathogen of *Arabidopsis* [34], and EST sequences from *F. oxysporum* f. sp. *vasinfectum* [35], a strain of *F. oxysporum* that infects cotton, to access the variation of the *Fol* LS regions within the *F. oxysporum* species complex. We have generated one run of Illumina reads for *F. oxysporum* strain Fo5176, a pathogen of *Arabidopsis*, using the Applied Biosystems and Illumina high-throughput sequencing platform Genome Analyzer. The raw data was processed by removing low quality reads using quality filter, resulting a total of 26,743,945 reads of 51 bp in length (1 Gb). The software MAQ [36] was used to map a total of 10,688,499 reads to the reference assembly of *Fol*, covering 46 Mb (60%) of the *Fol* assembly. In the aligned region the median SNP density was 14 SNPs per 1Kb, reflecting a short evolutionary distance between these two strains (Figure S13 A). Despite the low sequence divergence indicated from shared sequences, 40% of the reference genome was not represented in strain Fo5176. Measuring by the coverage of the Fo5176 reads in 10 Kb windows, we observed that the reference genome was not equally covered by the Illumina reads (Figure S13 B). By contrast, the majority of the genome (70% of the 10 Kb windows) has a coverage centering around 85% sequence alignment (ranging from 60%-100%). We found a clear enrichment of genomic regions that have a very low coverage (0-10%), most of which corresponded to *Fol*-LS regions in the reference assembly (Figure S13 B). While the genomic regions conserved between *Fol* and *Fv* had an greater than 80% coverage when compared to Fo5176, with a drop of coverage near telomeric regions of each conserved chromosomes, the *Fol*-LS regions had much lower sequence coverage (average 34.5%), indicating considerable sequence variation in the *Fol* LS regions within the *F. oxysporum* species complex. This provides direct support, analyzed at a genome wide level, for the existence of large LS-regions in genomes of members of the *F. oxysporum* species complex, as predicted by the *Fol-Fv* comparison and is consistent with studies where strain-specific sequences have been isolated in the *F. oxysporum* complex [37].

A total of 3,034 expressed sequence tags (ESTs) isolated from cotton root and hypocotyl tissues infected with *Fusarium oxysporum* f. sp. *vasinfectum* [35] were downloaded from NCBI and searched against *Fol* transcripts (blastn 1e-100). A total 198 fungal ESTs were uniquely matched to the *Fol* transcripts with high sequence similarity (an average of 99%). All these transcripts are located on orthologous chromosomes conserved between *Fol* and *Fv*. For a full list of annotated genes in each category see: http://www.broad.mit.edu/annotation/genome/fusarium_group/SupplementaryPage.html.

***E. 3 Transposable elements*.**

Many types of transposable elements [38], including both retroelements and DNA transposons, have been identified in the genome of *Fol*. Interestingly, DNA transposons were the most represented elements among the manually annotated TEs (65%), a unique situation as generally, in fungal genomes, the most abundant class of TEs are retroelements [39]. DNA transposons can be assigned to five distinct superfamilies: *Tc1-mariner* with the *pogo* and the *impala* families, *hAT* with 13 distinct families, *Mutator*-like with 5 families, MITEs with 6 distinct families and a *Helitron*-like family.

The *pogo* family, which is generally well-represented in fungal genomes [39,40] is uniquely amplified in the *Fol* genome with more than 500 copies belonging to eight different subfamilies, and representing approximately half of the DNA transposons. The copy number varies greatly from one *pogo* subfamily to another, from less than ten copies to half of the overall *Fot* copies. In the most numerous subfamilies, potentially active elements could be identified. Interestingly, a comparison of the representation of each *Fot* family in two different *F. oxysporum* strains, strain FOM24 (f. sp. *melonis*) and *Fol* suggested that amplification is strain-dependent. This is exemplified by *Fot1*, present in more than 100 copies in FOM24 [39] but only five copies in *Fol* (this work). Finally, on the basis of their small and homogeneous size, criteria commonly accepted for MITEs, we also identified 140 copies of a novel MITE, sharing TIRs with *Fot5*.

The *impala* family is characterized by a small number of copies diversified in three different subfamilies, none of which contain active elements. A family of MITEs, *mimp*, has been extensively characterized in the *Fol* genome [41] and demonstrated to be transactivated by the *impala* transposase both in *Fo* and heterologous genetic backgrounds.

The *hAT* superfamily is also well-represented in *Fol*: thirteen distinct families have been recognized, including the *Folyt* element which has already been demonstrated to be active in *Fol* [42]. The number of copies varies from four to 60 depending of the family. Four distinct families of small elements, three of which could clearly be associated with an identified full length *hAT*-like element were detected and classified as MITEs. *Hop* elements, previously characterized as *Mutator*-like TEs [43] were also identified in *Fol*. Although present in a much lower number of copies, they were classified into five distinct subfamilies.

A *Helitron* (Strider) of 6129 bp was identified in *Fol*. The ORF contains one intron and encodes an 1899 amino acid replicase/helicase/endonuclease. Four copies are disrupted by a *hAT* transposon and there are at least 15 nearly identical full length copies and another nine copies perhaps full length but truncated due to sequencing gaps or contig ends. Their high sequence similarities suggest a recent expansion of this element in the *Fol* genome.

Among retroelements, at least four different families of LTR elements were recognized or newly identified: the *Yaret2* family, a *gypsy*-like element (*Skippy*) and two *copia*-like retrotransposons (*Han*, *Gollum*), three of these composed of both full length and solo LTR elements. Non-LTR elements were also found in *Fol*: one family of LINE (*Yaret1*) and one family of SINE (*Foxy*). The SINE-like retrotransposon *Foxy* is the most abundant retroelement in *Fol*, which is in agreement with earlier AFLP analysis [44].

Transposable elements are over-represented in the LS regions of *Fol* (Table S11). More than 74% of all the identifiable TEs, including 95% of all the DNA transposable elements are present in *Fol* LS regions. This biased distribution of TEs is more significant for DNA transposons than for retro-elements, which are the main TEs present in *Fv* and *Fg* genomes. Our analysis suggests there has been a recent lineage-specific expansion of the DNA elements in association with the formation of *Fol* LS regions. A significant part of this expansion (approximately 50%) results from the presence of segmental duplications, some large genomic regions being found up to four times. For a full list of annotated genes for each category, see web site: http://www.broad.mit.edu/annotation/genome/fusarium_group/SupplementaryPage.html.

E. 4. *Segmental duplication.*

We aligned each of the 15 *Fol* optical contigs with all other optical contigs using Map Aligner [3] and discovered four large segmental duplications among *Fol* LS chromosomes, covering about 7 Mb (Figure 2C). The map alignment confidence score rewards matching restriction sites and

fragment length, and at the same time penalizes discrepancies including missing or extra restriction sites and differing length of the restriction fragments. A map alignment score of six with at least five matched restriction fragments is considered to be significant. The local map alignment confidence scores for all the four duplicated regions range from 49.99 to 269.34, and the matched restriction fragment numbers range from 58 to 210, reflecting high confidence for the aligned duplicated regions.

All of the segmental duplications detected by the optical mapping process are located in the *Fol* LS regions with high nucleotide sequence identity among the aligned regions (>99%), indicating that the observed duplications occurred very recently (Figure S7). To detect all potential segmental duplications, especially potential duplications across the conserved versus the *Fol* LS regions that may have accumulated more nucleotide changes, we searched for two genomic regions that contain a minimum of three genes that are in the same order and orientation. We identified three small duplications occurring between *FOL* LS regions (chromosome 3, scaffold 25) and telomere proximal regions of conserved chromosomes 4, 5 and 9, accounting for a total of 14 genes. These results clearly indicate that such limited exchange of genetic material between the conserved and *Fol* LS regions is not the major process involved in formation of the *Fol* LS regions.

E.5 ***Genes and gene categories expanded on LS chromosomes.***

Blast score ratio tests

Blast Score Ratio (BSR) analysis is designed for the rapid comparison of complete proteomes of three genomes, and enables a visual evaluation of the overall degree of similarity of these proteomes and their genomic structure [45]. The BLAST raw score for each *Fol* protein against itself is stored as the Reference score. Each *Fol* protein is then compared to each protein in *Fg* and *Fv* proteomes with each best BLAST raw score recorded as Query score respectively. The BSR for each reference protein is computed (Query score divided by the Reference score) with a normalization from 0 to 1. A score of 1 indicates a perfect match of the Reference to a Query sequence and score of 0 indicates no BLAST match of the Reference in the Query proteome. For each protein in the *Fol* genome, two numbers are generated from the best matches in *Fg* and *Fv*, respectively. The normalized pair of numbers were plotted using Matlab 7.0 software package. The *Fol* proteins encoded in the conserved regions versus the *Fol* LS regions exhibited very different profiles under this test (Figure 3).

Codon usage tests

Codon usage tables for each chromosome were computed using the EMBOSS tool 'cusp' (http://oryx.ulb.ac.be/embosshelp/cusp.html). The Pearson correlation coefficient (r) was calculated to determine the linear dependence of codon usage between each pair of chromosomes. While the r is close to 1 for both comparisons among the conserved chromosomes (0.999 ± 0.00167) and among the *Fol* LS chromosomes (0.998 ± 0.002), the r between conserved and *Fol* LS chromosomes is lower (0.990 ± 0.002), showing a weaker linear relationship between these two regions (Table S19).

To quantify this, we calculated the Pearson correlation coefficients of codon usage preferences between different chromosomes, and grouped them into three classes: A) between chromosomes within the conserved regions (top 10 rows in Table S19), B) between chromosomes within the LS regions (last 3 columns in Table S19), and C) between chromosomes in the conserved regions and the chromosomes in the LS regions (highlighted red in Table S19). The correlation coefficients in group A showed no difference from those in group B (p=0.30, two-sample t-test), indicating that the codon usage preferences are all highly similar within each group. However, the correlation coefficients between group A and C are significantly different (p=4.3e-38), and so are the correlation coefficients between B and C (p=1.9e-10), suggesting that the genes in the LS regions utilize a significantly different codon sets from those in the conserved regions (Table S19). The most significant differences were observed for amino acid Q, C, A, G, V, E, and T, with a preference for G and C over A and T among the genes encoded in the *Fol* LS regions (Table S20).

In addition, we computed the codon adaptation index (CAI) using the the EMBOSS tool 'cai' (http://oryx.ulb.ac.be/embosshelp/cai.html) as the geometric mean of the relative synonymous codon usage (RSCU). The CAI distribution for the genes encoded in the conserved region of *Fol* is almost identical to the CAI distribution of genes encoded in the *Fv* genome, however, the CAI distribution of the genes encoded in the *Fol* LS regions is shifted toward the lower end (Supplementary Figure S10).

Cross-kingdom proteome comparison suggests that the *Fol* LS regions are of fungal origin. Among the predicted proteins sharing homologous sequences to the protein set at the NCBI database (excluding *Fusarium* proteins), less than 2.6% of the LS-encoded proteins have bacterial proteins as their best hit, comparable to proteins encoded in the conserved regions (2.3%). Significantly, 93% have their best BLAST hit to other Ascomycete fungi (Figure S12). Using present/absent as a measure of closeness among these LS chromosomes and other fungal

genomes within Ascomycetes, the closest hit is *Fs*, where HT is also reported [9]. A distance matrix to show the distance between the genes encoded in LS regions and other selected fungal genomes are shown (Table S21)

E.5. Functional enrichment of genes encoded in Fol LS regions

Gene ontology (GO) terms [46] were assigned to 7,709 out of the total 17,735 genes predicted for *Fol* (43.5%) based on BLASTP (1e-20) using the program Blast2GO [47]. A total of 595 genes out of the 3,261 (18.3%) encoded on the *Fol* LS chromosomes were assigned with 318 GO terms. Among these, only 24 GO terms are enriched (FDR <0.01) (Table S12), while 98 of them are under-represented (also see detailed results from the GO-enrichment test at http://www.broad.mit.edu/annotation/genome/fusarium_group/SupplementaryPage.html). The *Fol* LS enriched functional categories include the proteins secreted to the extracellular space, genes involved in the fungal cell wall degrading process, signal transduction (PLC, *FlbA*, and kinases), and gluconeogenesis (Figure S9). Detailed annotation of the genes encoded on chromosome 14, a *Fol* LS chromosome capable of mobilization between vegetatively incompatible strains and encoding host range determinants, also reveals enrichment for secreted proteins (9.5% of total proteins), TFs (10% of proteins with functional assignment) and function related with DNA modifications (20%). Conversely, house-keeping genes are nearly absent from this chromosome (See detailed results from chromosome 14 at http://www.broad.mit.edu/annotation/genome/fusarium_group/SupplementaryPage.html).

Secreted proteins

Among the potentially secreted proteins (See all the secreted proteins at http://www.broad.mit.edu/annotation/genome/fusarium_group/SupplementaryPage.html), we identified 253 that are encoded on *Fol* LS regions (Table S13) that comprise small cysteine-rich proteins (potential plant effectors), necrosis and ethylene-inducing proteins (NEPs) and carbohydrate active enzymes (CAZYs), all considered likely to play a role in pathogenesis. In contrast, the *Fol* LS regions contain none of the 137 predicted to be wall-anchored proteins. Small, cysteine-rich proteins produced by fungi may be of ecological relevance by functioning as toxins, inhibitors of hydrolytic enzymes secreted by other organisms (e.g. proteases and chitinases) or by suppressing defense responses of host plants. Importantly, peptides secreted by the fungus within a host may be recognized by the plant innate immune system, triggering disease resistance. Indeed, almost all avirulence factors of fungi identified to date are small, cysteine-rich

proteins [48,49]. In fact, the three avirulence factors of *Fol*, corresponding to the three known resistance genes in tomato, are small proteins secreted in xylem ("SIX" proteins), each having between two and eight cysteine residues [50]. At least two AVR genes (probably all three) are encoded on LS chromosomes [37,50]. Seven additional SIX proteins also have between two and eight cysteines and are encoded by LS chromosomes; several also have roles in virulence and/or enhancement or suppression of disease resistance response.

Other secreted proteins uniquely encoded on the LS chromosomes include plant necrosis–inducing proteins (NEPs) that can act as virulence factors during infection of plants [51]. NEPs activate host defense-associated responses, including accumulation of pathogenesis-related (PR) proteins and reactive oxygen species [49], that result in cell death. Remarkably, three NEP genes from *Fol* located on LS regions (FOXG_14409.2, FOXG_15072.2 and FOXG_17014.2) are closely related to "wilt-specific" gene lineage containing only genes from the vascular pathogens *F. oxysporum* and *Verticillium dahliae* (AAS45247.1). Possibly, the *F. oxysporum*- and *Verticillium* specific NEP genes contribute to the development of the characteristic wilt symptoms of infected plants.

Secreted carbohydrate active enzymes (CAZYs) have multiple functions in virulence, including host penetration, nutrient acquisition and pathogen cell wall remodeling. The *Fg* LS chromosomes are significantly enriched for glycoside hydrolase (GH) genes, resulting from expansion of several families of GHs that are important for degrading plant cell wall components (Table S3 and S14. Also see detailed annotation of all CAZY at http://www.broad.mit.edu/annotation/genome/fusarium_group/SupplementaryPage.html). The expansion of secreted GHs and their enrichment on the LS chromosomes suggests that the acquisition of LS chromosomes by *Fol* has increased the diversity of its emzyme arsenal against major plant cell wall components, such as pectins or arabinogalactans. This may well be related to the broad host range of *Fol*.

To test whether these GH genes are expressed during early stages of plant infection, reverse transcriptase PCR was performed on total RNA isolated from infected tomato roots at three and seven days after inoculation (Figure S8 and Table S15). The analysis included GH families previously implicated in fungal virulence such as β-glucuronidase (GH79), β-1,3-glucanase (GH64), polygalacturonase (GH28) or chitinase (GH18), among others (Table S15). For most enzyme families, two classes of members were included: 1) *Fol* specific genes located on LS chromosomes, most of them present in multiple copies (e.g. FOXG_06568, FOXG_12407,

FOXG_12535, FOXG_15491), and 2) single orthologous copy genes present in all three *Fusarium* species (e.g. FOXG_02657, FOXG_09503, FOXG_13051, FOXG_00921). We detected transcripts of all the orthologous GH genes and approximately half of the *Fol* LS genes tested. By contrast, we failed to detect transcripts of any of the four GT genes included in the analysis (Figure S8 and Table S15).

Among the transcripts detected *in planta* are two *Fol* LS specific chitinases (FOXG_14329, FOXG_15151) and two *Fol* LS specific genes encoding a chitin-binding protein (CBM12) that shares homology only with bacterial genes (FOXG_15172, FOXG_14308). Interestingly, the LS chromosomes contain four copies (FOXG_14840, FOXG_16128, FOXG_06883, FOXG_16251) of an endochitinase gene orthologous to *ech-42* from the fungal mycoparasite *Trichoderma harzianum* [52]. Expansion of chitinases and other fungal wall degrading enzymes on LS chromosomes may reflect selection for competitiveness of *Fg* with other cohabitant fungi in its main habitat, the plant rhizosphere.

Expression of genes for secreted CAZYs in planta. Two week-old tomato plants were inoculated with microconidia of *F. oxysporum* f.sp. *lycopersici* strain 4287 as described [52]. Total RNA was isolated from roots of non-infected and infected plants three days after inoculation, using the RNeasy Plant Mini Kit (Qiagen, Germany) according to the manufacturer's instructions. Quality and quantity of extracted RNA were verified by running aliquots in ethidium bromide stained agarose gels and by spectrophotometric analysis in a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, USA), respectively. The isolated RNA was treated with deoxyribonuclease I (DNase I, Fermentas, USA) and reverse transcribed into first strand cDNA with ribonuclease inhibitor RNasin Plus RNase inhibitor (Promega, USA) and M-MLV reverse transcriptase (Invitrogen S.A., Spain) using a poly-dT antisense primer. Gene-specific primers were used for PCR amplification (Table S16). The *F. oxysporum* actin 1 gene was used as a positive control for gene expression, and *F. oxysporum* genomic DNA was used as a control template for size comparison with bands amplified from cDNA of intron-containing genes. PCR products were electrophoretically separated in 2% agarose gels, stained with ethidium bromide and photographed.

Peroxidases.

Three genes on LS-chromosomes encode catalase-peroxidases (EC 1.11.1.6) (FOXG_14234, FGXG_17106, and FOXG_17130), a subgroup of heme-dependent peroxidases. This group comprises secretory peroxidases that exhibit both peroxidase and catalase activities and provide

protection to cells under oxidative stress [53]. Plant cells in contact or close proximity to the pathogen usually respond defensively with an oxidative burst. This may lead to direct damage to the pathogen, plant cell wall reinforcement and/or programmed cell death. The catalase-peroxidases of *Fol* may have important roles in reducing the effects of oxidative stress. The three copies of these genes on LS chromosomes in *Fol* are paralogs of a syntenic copy (FOXG_17180) found in all species indicating a specific expansion within *Fol*.

Lipid metabolism and signaling.

Species-specific expansions are observed in *Fol* among gene families that are important in lipid metabolism and generation of lipid-derived second messengers. Particularly, we observed a large expansion of genes for the perilipin-like Cap20 protein that regulates the release of non-esterified free fatty acids (NEFA), as well as phospholipase C (PLC) genes that generate the signaling molecules inositol 1,4,5-triphosphate (IP3) and diacyl glycerol (Table S17). These lipid-derived molecules play crucial roles in intra- and inter-cellular signaling [54] and regulate pathogenic traits required for the development of fungal diseases [55].

There is an increasing body of knowledge about how perilipin-like proteins regulate lipid homeostasis in mammals [56,57], however, little is known about the factors that mediate lipid storage and control lipid homeostasis in fungi. In mammalian systems, Perilipin A [25] is a lipid droplet-associated phosphoprotein that acts dually as a suppressor of basal (constitutive) lipolysis and as an enhancer of cyclic AMP-dependent protein kinase (PKA)-stimulated lipolysis through controlling the accessibility of lipase to the hydrophobic triglyceride core [57]. Such control is mediated through a G protein-coupled receptor (GPCR) that activates adenylyl cyclase and, consequently, protein kinase A (PKA) [58]. Adipose triglyceride lipase and Hormone-sensitive lipase are the major enzymes involved in adipose tissue triacylglycerol catabolism [59] and promote the release of non-esterified free fatty acid (NEFA). The resulting NEFA can be used for energy generation in beta-oxidation, membrane phospholipid synthesis, signaling, and in regulation of transcription factors such as the peroxisome proliferator-activated receptors (PPARs) [56,57].

As pathogenicity determinants, perilipin-like *CAP*20 (first identified in the phytopathogenic fungus *Colletotrichum gloeosporioides)* and its homologous gene *MPL1* (*Metarhizium anisopliae* perilipin-like *MPL1*) were uniquely induced by a host signal during appressorium formation and localized at the infection front and within infected host tissue [60]. A single copy of this gene is detected in pezizomycotinal fungi of the phylum Ascomycota, but is absent in yeasts and other

fungi. As anticipated, three *Fusarium* genomes share an orthologous *CAP20* gene. However, the *Fol* genome contains eight additional copies of *CAP20* homologs in the LS genomic regions. In addition to the expansion of the *CAP20* homologues in *Fol* LS genomic regions, we also observed the expansions of lipases that act on carboxylic esters, a family of GPCRs, and a group of Regulator of G protein Signaling (RGS) proteins among these regions as well (Table S17). The expanded GPCRs share sequence similarity to human membrane Progestin Receptor (mPR). In addition to two ancestral copies conserved among other fungi, *Fol* encodes another four mPR-like GPCRs in the LS regions. The mPR-GPCRs transfer signals to G proteins at the plasma membrane and down-regulate adenylyl cyclase [20]. Even though homologues of human mPR-like proteins have not been functionally characterized in any fungus [17], based on mammalian systems, we predict that they could regulate lipid homeostasis through *Cap20*. In connection with regulating Gα proteins and presumably acting upstream of PKA, we observed the expansion of one of the RGS, a *FlbA*-like protein. *FlbA*-like gene products may act as GTPase-activating proteins to negatively control Gα protein signaling [61]. While most sequenced fungal genomes contain a single *FlbA* gene, two orthologous copies were identified in the three *Fusarium* genomes, while the *Fol* LS regions encode six additional *FlbA*-like genes. Another key element of lipolysis, the lipase gene family, is also expanded in the LS regions, parti cularly the genes for lipases that act on carboxylic esters.

The PLC gene (YPL268W in yeast) encodes a calcium-dependent phospholipase that hydrolyzes phosphatidylinositol 4,5-biphosphate (PIP2) to generate the signaling molecules diacylglycerol DAG and inositol 1,4,5-triphosphate (IP3). Both DAG and IP3 serve as secondary messengers that regulate cellular processes such as transcription, mRNA export, DNA repair, chromatin remodeling, glucose-induced calcium signaling, telomere elongation, vacuolar biogenesis, stress responses and cell wall synthesis [62]. DAG also been implicated as a second message in the plant pathogenic fungus *M. grisea* where it promotes formation of appressoria, structures required for infection, on normally non-conducive hydrophilic surfaces [63]. In addition to one orthologous PLC1 gene encoded in each *Fusarium* genome, there are six copies in the *Fol* LS regions. It is noteworthy that four Fol PLC genes, FOXG_07102, FOXG_07611, FOXG_15032 and FOXG_14448, have a Pleckstrin homology (PH)-like domain at the amino terminus, which is lacking from other fungal PLC genes, such as those in *Neurospora* and yeast. PH-like sequences, thought to target PLC to the plasma membrane through binding to PIP3, the phospholipid product of PI-3 kinase, are also found at the amino terminus of human b and d isozymes. The presence of

this domain might be important for targeting the *Fusarium* PLCs to the plasma membrane in certain cell types.

E.6. RIP and other silencing pathways.

One clear difference among the three genomes is Repeat Induced Point mutation (RIP), an Ascomycete-specific defense mechanism that recognizes and inactivates duplicated segments during the sexual cycle by C:G to T:A mutations [62,64]. RIP was recently shown to operate on introduced repeated sequences in the homothallic *Fg* [33], and a severe form of RIP was observed in a group of skippy-like elements in *Fv* (Figure S18). Four almost full-length SLREs are found on chromosomes 8, 9, 11 and on an unmapped contig. None of these elements are predicted to be functional, but based on the mutational profile, the *Fol* and *Fv* SLRE elements shared a common ancestor. The *Fol* SLRE, with nine full-length or almost full-length copies, shows no evidence of RIP. The predicted proteins are relatives of "Maggy" from *M. grisea* [65]. No C:G to T:A mutations are found in *Fol* within nine individual elements that are located on chromosomes 1, 3, 4, 6, 10, 14, 15, and in unmapped contigs (Figure S19a). The ~240-bp 5'- and 3'-LTRs are retained in all nine copies (Figure S19 b, c). Because this element is found on both LS and non-LS *Fol* chromosomes, these elements either were introduced to *Fol* with the LS chromosomes or invaded even more recently.

If not all *Fusarium* species can make use of RIP to inactivate transposable elements, how do they control the movement of such "selfish DNA" or at least dampen their effects? Two additional well-studied silencing phenomena, DNA methylation and RNAi-related mechanisms akin to "quelling" in *Neurospora* [66] may help to inactivate TEs during asexual reproduction and vegetative growth (Table S18). All *Fusarium* species have the potential to express a *Neurospora* DIM-2 homologue but to date ESTs for this gene has only been found in *Fv*. Two additional genes that encode proteins required for DNA methylation in *Neurospora*, the histone H3 K9 methyltransferase DIM-5 [67] and the heterochromatin proteins HP1 [68] are present in all three *Fusarium* species (Table S18). ESTs from five of the six genes have been found. The one gene without ESTs, *Fg Hpo*, has been expressed as a fusion protein with both GFP and RFP and its localization is in heterochromatic foci as observed for *Neurospora* HP1 (L. Connolly, P. Phatale and M. Freitag, in preparation). While the sequenced *Fg* strain and another frequently used strain (Z-3639) [69] have very little DNA methylation, both *Fol* and *Fv* strains show different banding patterns when examined with a pair of isoschizomers that are insensitive (*Dpn* II) or sensitive (*Sau3A* I) to cytosine DNA methylation (K. Pomraning, J. Whalen and M. Freitag, in

preparation). Thus methylation may be one way to control expression of TEs in *Fol* and *Fv* but further experiments to evaluate the distribution of methylated DNA in *Fusarium* species are required.

While the DNA methylation and small RNA silencing machineries appear not to be connected in *Neurospora* [68], RNA interference or quelling can act on its own to inhibit expression of aberrant RNAs or TEs [70]. This type of silencing occurs in many organisms, including those that do not methylate their DNA. All three *Fusarium* genomes contain homologues for genes previously shown to be important for quelling or meiotic silencing in *Neurospora* [71], i.e., two dicer homologues, two argonaute homologues and three homologues of RNA-dependent RNA polymerases (Table S18). Interestingly, three additional *QDE*-2 homologues are found in *Fol*, all of them on LS chromosomes (Figure S20 and Table S18). As with heterochromatin silencing pathways that rely on histone or DNA methylation, little is known about RNAi or meiotic silencing in *Fusarium* species.

F. Transfer of LS chromosomes between strains.

F.1 Transformation of Fusarium strains

To select for transfer of chromosome 14 from into Fom, Foc, Fo-47 or Fol, the receiving strains were transformed with a hygromycin (*HYG*) resistance gene, inserted randomly into the genome - three independent hygromycin resistant transformants were selected for the experiments. The hygromycin resistance gene was present in the vector pPK2-HPH-GFP which has the *GFP* gene cloned in frame behind *HPH* [72] and was transformed to Fom (NRRL 26406), Fo-47 and Foc (NRRL 25603) using *Agrobacterium* mediated transformation [73]. For transformation of Fol (Fol007) with HYG the same strategy was applied, except that a vector was used with *CFP* cloned behind *HPH* (pPK2-HPH-CFP). The donor strain (Fol007) has a phleomycin resistance gene (*BLE*, confers resistance against zeocin) inserted in chromosome 14 (B chr. 14) 180 bp downstream of the *SIX1* ORF and was obtained through homologous recombination at the *SIX1* locus, as described in [74]. All media containing zeocin were buffered with 100mM Tris pH 8.0.

F.2. Co-incubation of spores

To select for transfer of chromosome 14 from the donor strain (Fol007) into Fom, Foc, Fo-47 or Fol007, one donor and one receiving strain were co-incubated. All *F. oxysporum* strains are grown at 25 °C in the dark. For co-incubation, microconidia of the different strains were isolated

after three days of growth (175 rpm) in synthetic medium (1% KNO_3, 3% sucrose, 0.17% Yeast Nitrogen Base w/o amino acids and ammonium), mixed in a 1:1 ratio (10^5 conidia of each strain) and plated out on rich medium (potato dextrose agar-PDA). After six days of growth new microconidia had formed. Five milliliter sterile MQ water was pipetted on the co-incubated plates and microconidia were scraped off. For selection of double resistant colonies 20, 100 and 500 µl of the microconidia suspension was plated out on agar plates (PDA) containing both zeocin and hygromycin (buffered with 100 mM Tris pH 8.0). To determine the total number of plated conidia 20, 100 and 500 µl of a 1:1,000,000 dilution was plated on PDA plates without antibiotic. After two days, the number of double drug resistant colonies resistant were scored.

F.3. Bioassays

To test pathogenicity the root dip method was used [75]. Briefly, microconidia were collected from 5-day-old cultures in Czapek Dox broth (CDB; Difco) and used for root inoculation of 10-day-old tomato plants (Moneymaker C32) at a spore density of 10^7 spores per ml. The seedlings were then potted individually and grown at 25 °C in the greenhouse. Three weeks after inoculation, the overall phenotype and the extent of browning of vessels was scored on a scale of 0–4: 0, no symptoms; 1, slightly swollen and/or bent hypocotyl; 2, one or two brown vascular bundles in hypocotyl; 3, at least two brown vascular bundles and growth distortion (strong bending of the stem and asymmetric development); 4, all vascular bundles are brown, plant either dead or stunted and wilted. Reisolation of isolates from diseased plants was done by placing slices of hypocotyls on agar plates (PDA) containing both zeocin and hygromycin. From mycelium growing out of the slices, single-spore colonies were derived and used for infection of tomato to satisfy Koch's postulates.

F.4. PCR amplifications

To test whether just the *BLE* resistance gene or the entire chromosome 14 was present in double drug resistant strains derived from the co-incubation experiments, three double drug resistant colonies of each successful combination (*Fol*007 with three independent *HYG*-transformants of *Fom* and Fo-47) were tested for the presence of genes located on different parts of chromosome 14, including *SIX1*, *SIX2, SIX3*, *SIX5, SIX6*, *SIX7* and *ORX1* (Figure S15). All seven genes were present in each of the nine double drug resistant strains (18 strains in total) indicating the presence of most or all of chromosome 14 in every case (Figure 4A – *SIX6* and *ORX1* are also present in the *Fom* strain used and could therefore not be used to distinguish between *Fol* and *Fom*-derived material). PCRs were performed on genomic DNA as described in [37]. For *SIX1-7* and *ORX1* the following primers were used:

SIX1-F1:	CCCCGAATTGAGGTGAAG
SIX1-R1:	AATAGAGCCTGCAAAGCATG
SIX2-F2:	TCTATCCGCTTTCTTCTCTC
SIX2-R2:	CAACGCCGTTTGAATAAGCA
SIX3-F1:	CCAGCCAGAAGGCCAGTTT
SIX3-R2:	GGCAATTAACCACTCTGCC
SIX5-F1:	ACACGCTCTACTACTCTTCA
SIX5-R1:	GAAAACCTCAACGCGGCAAA
SIX6-F1:	CTCTCCTGAACCATCAACTT
SIX6-R1:	CAAGACCAGGTGTAGGCATT
SIX7-F1:	CATCTTTTCGCCGACTTGGT
SIX7-R1:	CTTAGCACCCTTGAGTAACT
ORX1-F1:	CCAGGCCATCAAGTTACTC
ORX1-R1:	TCTCCAATATGGCAGATTGTG

To track individual chromosomes, a marker system based on differences in locations of the Foxy transposon was set up. The screen uses specific primers based on sequences ca. 500 bp upstream of 48 Foxy insertions based on the Fol genome sequence as released by the Broad Institute (isolate Fol4287). Each specific primer was used in combination with a reverse primer in Foxy and together the primer combinations covered all 15 chromosomes. (Table S25). Each was used with a reverse primer in Foxy (GAGAGAATTCTGGTCGGTG). First the primers were tested on Fol007 and Fo-47. From the primers that were positive on Fol007 and negative on Fo-47, twenty-nine primers were selected for testing the double drug-resistant strains for the presence of Fol007-derived chromosomes.

The additional Foxy insertion in strain 1C and 2A is located in scaffold 18 (1.3 Mb) on chromosome 3 and scaffold 21 (1.0 Mb) on chromosome 6. The presence of most or all of the sequence of scaffold 18/21 in strains 1C and 2A was confirmed with an additional 9 primers for Foxy insertions scattered over this region (Table S25). Given the differences in karyotype between *Fol*4287 and *Fol*007, it is highly likely that this region (scaffold 18/21 in *Fol*4287) is present in the smallest chromosome in *Fol*007 (Figure 4D).

F.5. CHEF & Southern blotting

Protoplasts (10^8/ml) from Fol4287, Fol007, three independent HYG transformants of Fo-47 and nine double drug-resistant strains were loaded on a CHEF gel as described in [44]. The gel (1% Seakem Gold agarose [FMC] in 0.5x TBE) ran for 255 hours, using switch times between 1200-4800 s at 1.5 V/cm. Chromosomes of *S. pombe* were used as a molecular size marker (BioRad). Chromosomes separated in the gel were blotted to Hybond-N+ (Amersham Pharmacia). PCR with primers SIX6-F1 and SIX6-R1 was used to generate a ~800 bp fragment containing the entire *SIX6* open reading frame. This fragment was radioactively labelled with γ^{32}P dATP using the DecaLabel™ DNA labeling kit from MBI Fermentas (Vilnius, Lithuania). Hybridization was done overnight at 65°C in 0.5M phosphate buffer pH 7.2 containing 7% SDS and 1 mM EDTA. Blots were washed at 65°C with 0.5 X SSC, 0.1% SDS. The position of chromosomes hybridizing to the *SIX6* sequence was visualized by phosphoimaging (Molecular Dynamics).

Reference cited

1 Jaffe, D. B. et al., Whole-genome sequence assembly for mammalian genomes: Arachne 2. *Genome Res* **13** (1), 91 (2003).

2 Xu, J. R. and Leslie, J. F., A genetic map of *Gibberella fujikuroi* mating population A (*Fusarium moniliforme*). *Genetics* **143** (1), 175 (1996).

3 Zhou, S. et al., Single-molecule approach to bacterial genomic comparisons via optical mapping. *J Bacteriol* **186** (22), 7773 (2004).

4 Park, J. et al., FTFD: an informatics pipeline supporting phylogenomic analysis of fungal transcription factors. *Bioinformatics* **24** (7), 1024 (2008).

5 Cantarel, B. L. et al., The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res* **37** (Database issue), D233 (2009).

6 Miranda-Saavedra, D. and Barton, G. J., Classification and functional annotation of eukaryotic protein kinases. *Proteins* **68** (4), 893 (2007).

7 Lafon, A. et al., G-protein and cAMP-mediated signaling in aspergilli: a genomic perspective. *Fungal Genet Biol* **43** (7), 490 (2006).

8 Winnenburg, R. et al., PHI-base update: additions to the pathogen host interaction database. *Nucleic Acids Res* **36** (Database issue), D572 (2008).

9 Coleman, J. J. et al., The genome of *Nectria haematococca*: contribution of supernumerary chromosomes to gene expansion. *PLoS Genet* **5** (8), e1000618 (2009).

10 Todd, R. B. and Andrianopoulos, A., Evolution of a fungal regulatory gene family: the Zn(II)2Cys6 binuclear cluster DNA binding motif. *Fungal Genet Biol* **21** (3), 388 (1997).

11 Guldener, U. et al., Development of a *Fusarium graminearum* Affymetrix GeneChip for profiling fungal gene expression *in vitro* and *in planta*. *Fungal Genet Biol* **43** (5), 316 (2006).

12 Haas, H., Eisendle, M., and Turgeon, B. G., Siderophores in fungal physiology and virulence. *Annual Review of Phytopathology* **46**, 149 (2008).

13 Stergiopoulous, I., Zwiers, L.-H. , and De Waard, M.A., Secretion of natural and synthetic toxic compounds from filamentous fungi by membrane transports of the ATP-binding cassette and major facilitator subfamily. *European Journal of Plant Pathology* **108**, 719 (2002).

14 Fleissner, A., Sopalla, C., and Weltring, K. M., An ATP-binding cassette multidrug-resistance transporter is necessary for tolerance of *Gibberella pulicaris* to phytoalexins and virulence on potato tubers. *Mol Plant Microbe Interact* **15** (2), 102 (2002).

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

16 Reuben, A. et al., Antifungal susceptibility of 44 clinical isolates of *Fusarium* species determined by using a broth microdilution method. *Antimicrob Agents Chemother* **33** (9), 1647 (1989).

17 Li, L. et al., Heterotrimeric G protein signaling in filamentous fungi. *Annual Review of Microbiology* **61**, 423 (2007).

18 Regenfelder, E. et al., G proteins in *Ustilago maydis*: transmission of multiple signals? *The EMBO Journal* **16** (8), 1934 (1997).

19 Kays, A. M. and Borkovich, K. A., Severe impairment of growth and differentiation in a *Neurospora crassa* mutant lacking all heterotrimeric G alpha proteins. *Genetics* **166** (3), 1229 (2004).

20 Thomas, P., Characteristics of membrane progestin receptor alpha (mPRalpha) and progesterone membrane receptor component 1 (PGMRC1) and their roles in mediating rapid progestin actions. *Frontiers in Neuroendocrinology* **29** (2), 292 (2008).

21 Baasiri, R. A. et al., Overlapping functions for two G protein alpha subunits in *Neurospora crassa*. *Genetics* **147** (1), 137 (1997).

22 Murphy, M. and Armstrong, D., Fusariosis in patients with neoplastic disease. . *Infect.Med* **12**, 66 (1995).

23 Seefelder, W. et al., Induction of apoptosis in cultured human proximal tubule cells by fumonisins and fumonisin metabolites. *Toxicol Appl Pharmacol* **192** (2), 146 (2003).

24 Hendricks, K., Fumonisins and neural tube defects in South Texas. *Epidemiology* **10** (2), 198 (1999).

25 Marasas, W. F. et al., Fumonisins disrupt sphingolipid metabolism, folate transport, and neural tube development in embryo culture and in vivo: a potential risk factor for human neural tube defects among populations consuming fumonisin-contaminated maize. *J Nutr* **134** (4), 711 (2004).

26 Swofford, D. L., *PAUP. Phylogenetic Analysis Using Parsimony*. (Sinauer Associates, Sunderland, Massachusetts., 2003).

27 Gaffoor, I. and Trail, F., Characterization of two polyketide synthase genes involved in zearalenone biosynthesis in *Gibberella zeae*. *Applied and environmental microbiology* **72** (3), 1793 (2006).

28 Gaffoor, I. et al., Functional analysis of the polyketide synthase genes in the filamentous fungus *Gibberella zeae* (anamorph *Fusarium graminearum*). *Eukaryot Cell* **4** (11), 1926 (2005).

29 Proctor, R. H., Butchko, R. A., Brown, D. W., and Moretti, A., Functional characterization, sequence comparisons and distribution of a polyketide synthase gene required for perithecial pigmentation in some Fusarium species. *Food Addit Contam* **24** (10), 1076 (2007).

30 Frandsen, R. J. et al., The biosynthetic pathway for aurofusarin in *Fusarium graminearum* reveals a close link between the naphthoquinones and naphthopyrones. *Molecular Microbiology* **61** (4), 1069 (2006).

31 Hallen, H. E. and Trail, F., The L-type calcium ion channel cch1 affects ascospore discharge and mycelial growth in the filamentous fungus *Gibberella zeae* (anamorph *Fusarium graminearum*). *Eukaryot Cell* **7** (2), 415 (2008).

32 Li, M., Ma, B., Kisman, D., and Tromp, J., Patternhunter II: highly sensitive and fast homology search. *J Bioinform Comput Biol* **2** (3), 417 (2004).

33 Cuomo, C. A. et al., The *Fusarium graminearum* genome reveals a link between localized polymorphism and pathogen specialization. *Science* **317** (5843), 1400 (2007).

34 Thatcher, L. F., Manners, J. M., and Kazan, K., *Fusarium oxysporum* hijacks COI1-mediated jasmonate signaling to promote disease development in *Arabidopsis*. *Plant J* **58** (6), 927 (2009).

35 Dowd, C., Wilson, I. W., and McFadden, H., Gene expression profile changes in cotton root and hypocotyl tissues in response to infection with *Fusarium oxysporum* f. sp. *vasinfectum*. *Mol Plant Microbe Interact* **17** (6), 654 (2004).

36 Li, H., Ruan, J., and Durbin, R., Mapping short DNA sequencing reads and calling variants using mapping quality scores. *Genome Res* **18** (11), 1851 (2008).

37 van der Does, H. C. et al., The presence of a virulence locus discriminates *Fusarium oxysporum* isolates causing tomato wilt from other isolates. *Environmental Microbiology* (2008).

38 Daboussi, M. J. and Capy, P., Transposable elements in filamentous fungi. *Annual Review of Microbiology* **57**, 275 (2003).

39 Daboussi, M. J., Daviere, J. M., Graziani, S., and Langin, T., Evolution of the *Fot1* transposons in the genus *Fusarium*: discontinuous distribution and epigenetic inactivation. *Mol Biol Evol* **19** (4), 510 (2002).

40 Daboussi, M. J., Langin, T., and Brygoo, Y., *Fot1*, a new family of fungal transposable elements. *Mol Gen Genet* **232** (1), 12 (1992).

41 Bergemann, M. et al., Genome-wide analysis of the *Fusarium oxysporum* mimp family of MITEs and mobilization of both native and *de novo* created mimps. *J Mol Evol* (2008).

42 Gomez-Gomez, E., Anaya, N., Roncero, M. I., and Hera, C., *Folyt1*, a new member of the hAT family, is active in the genome of the plant pathogen *Fusarium oxysporum*. *Fungal Genet Biol* **27** (1), 67 (1999).

43 Chalvet, F. et al., *Hop*, an active Mutator-like element in the genome of the fungus *Fusarium oxysporum*. *Mol Biol Evol* **20** (8), 1362 (2003).

44 Teunissen, H. A. et al., Construction of a mitotic linkage map of *Fusarium oxysporum* based on Foxy-AFLPs. *Mol Genet Genomics* **269** (2), 215 (2003).

45 Rasko, D. A., Myers, G. S., and Ravel, J., Visualization of comparative genomic analyses by BLAST score ratio. *BMC Bioinformatics* **6**, 2 (2005).

46 Ashburner, M. et al., Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nature Genetics* **25** (1), 25 (2000).

47 Conesa, A. et al., Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **21** (18), 3674 (2005).

48 Catanzariti, A. M., Dodds, P. N., and Ellis, J. G., Avirulence proteins from haustoria-forming pathogens. *FEMS Microbiology Letters* **269** (2), 181 (2007).

49 Fudal, I. et al., Heterochromatin-like regions as ecological niches for avirulence genes in the *Leptosphaeria maculans* genome: map-based cloning of *AvrLm6*. *Mol Plant Microbe Interact* **20** (4), 459 (2007).

50 Houterman, P. M. et al., The effector protein Avr2 of the xylem colonizing fungus *Fusarium oxysporum* activates the tomato resistance protein I-2 intracellularly. *Plant J* (2009).

51 Qutob, D. et al., Phytotoxicity and innate immune responses induced by *Nep1*-like proteins. *The Plant Cell* **18** (12), 3721 (2006).

52 Di Pietro, A. and Roncero, M. I., Cloning, expression, and role in pathogenicity of pg1 encoding the major extracellular endopolygalacturonase of the vascular wilt pathogen *Fusarium oxysporum*. *Mol Plant Microbe Interact* **11** (2), 91 (1998).

53 Welinder, K. G., Bacterial catalase-peroxidases are gene duplicated members of the plant peroxidase superfamily. *Biochimica et Biophysica Acta* **1080** (3), 215 (1991).

54 Wang, X., Lipid signaling. *Current Opinion in Plant Biology* **7** (3), 329 (2004).

55 Shea, J. M. and Del Poeta, M., Lipid signaling in pathogenic fungi. *Current Opinion in Microbiology* **9** (4), 352 (2006).

56 Beller M, Sztalryd C, Southall N, Bell M, Jäckle H, et al., COPI complex is a regulator of lipid homeostasis. *PLoS Biology* **6** (11), e292 doi:10.1371/journal.pbio.0060292 (2008).

57 Brasaemle, D. L., Thematic review series: adipocyte biology. The perilipin family of structural lipid droplet proteins: stabilization of lipid droplets and control of lipolysis. *Journal of Lipid Research* **48** (12), 2547 (2007).

58 Patel, R. T., Soulages, J. L., Hariharasundaram, B., and Arrese, E. L., Activation of the lipid droplet controls the rate of lipolysis of triglycerides in the insect fat body. *The Journal of Biological Chemistry* **280** (24), 22624 (2005).

59 Schweiger, M. et al., Adipose triglyceride lipase and hormone-sensitive lipase are the major enzymes in adipose tissue triacylglycerol catabolism. *The Journal of Biological Chemistry* **281** (52), 40236 (2006).

60 Hwang, C. S., Flaishman, M. A., and Kolattukudy, P. E., Cloning of a gene expressed during appressorium formation by *Colletotrichum gloeosporioides* and a marked decrease in virulence by disruption of this gene. *The Plant Cell* **7** (2), 183 (1995); Wang, C. and St Leger, R. J., The *Metarhizium anisopliae* perilipin homolog MPL1 regulates lipid

metabolism, appressorial turgor pressure, and virulence. *The Journal of Biological Chemistry* **282** (29), 21110 (2007).

61 Yu, J. H., Wieser, J., and Adams, T. H., The *Aspergillus FlbA* RGS domain protein antagonizes G protein signaling to block proliferation and allow development. *The EMBO Journal* **15** (19), 5184 (1996).

62 Cambareri, E. B., Jensen, B. C., Schabtach, E., and Selker, E. U., Repeat-induced G-C to A-T mutations in *Neurospora*. *Science* **244** (4912), 1571 (1989).

63 Thines, E., Weber, R. W., and Talbot, N. J., Signal transduction leading to appressorium formation in germinating conidia of *Magnaporthe grisea*: Effects of second messengers discylglycerols, ceramindes and sphingomycelin. *FEMS Microbiol. Lett.* **156**, 91 (1997).

64 Nolan, T. et al., The post-transcriptional gene silencing machinery functions independently of DNA methylation to repress a LINE1-like retrotransposon in *Neurospora crassa*. *Nucleic Acids Res* **33** (5), 1564 (2005).

65 Farman, M. L., Tosa, Y., Nitta, N., and Leong, S. A., MAGGY, a retrotransposon in the genome of the rice blast fungus *Magnaporthe grisea*. *Mol Gen Genet* **251** (6), 665 (1996).

66 Selker, E. U., Repeat-induced gene silencing in fungi. *Adv Genet* **46**, 439 (2002).

67 Tamaru, H. and Selker, E. U., A histone H3 methyltransferase controls DNA methylation in *Neurospora crassa*. *Nature* **414** (6861), 277 (2001).

68 Freitag, M. et al., HP1 is essential for DNA methylation in *Neurospora*. *Mol Cell* **13** (3), 427 (2004).

69 Lee, J., Jurgenson, J. E., Leslie, J. F., and Blowden, R. L., Alignment of genetic and physical maps of *Gibberella zeae*. *Applied and Environmental Microbiology* **74** (8), 2349 (2008).

70 Chicas, A. et al., Small interfering RNAs that trigger posttranscriptional gene silencing are not required for the histone H3 Lys9 methylation necessary for transgenic tandem repeat stabilization in *Neurospora crassa*. *Mol Cell Biol* **25** (9), 3793 (2005).

71 Borkovich, K. A. et al., Lessons from the genome sequence of *Neurospora crassa*: tracing the path from genomic blueprint to multicellular organism. *Microbiol Mol Biol Rev* **68** (1), 1 (2004).

72 Michielse, C. B. et al., Insight into the molecular requirements for pathogenicity of *Fusarium oxysporum* f. sp. *lycopersici* through large-scale insertional mutagenesis. *Genome Biol* **10** (1), R4 (2009).

73 Takken, F. L. et al., A one-step method to convert vectors into binary vectors suited for Agrobacterium-mediated transformation. *Current genetics* **45** (4), 242 (2004).

74 van der Does, H. C. et al., Expression of effector gene *SIX1* of *Fusarium oxysporum* requires living plant cells. *Fungal Genet Biol* **45** (9), 1257 (2008).

75 Wellman, F.L., A technique for studying host resistance and pathogenicity in tomato *Fusarium* wilt. *Phytopathology* **29**, 945 (1939).

76 Houterman, P. M. et al., The mixed xylem sap proteome of *Fusarium oxysporum* infected tomato plants. *Molecular Plant Pathology* **8** (2), 215 (2007).

77 Rep, M. et al., A small, cysteine-rich protein secreted by Fusarium oxysporum during colonization of xylem vessels is required for I-3-mediated resistance in tomato. *Molecular Microbiology* **53** (5), 1373 (2004).