Comparative "Omics" of the *Fusarium fujikuroi* Species Complex Highlights Differences in Genetic Potential and Metabolite Synthesis

Eva-Maria Niehaus[1,†], Martin Münsterkötter[2,†], Robert H. Proctor[3], Daren W. Brown[3], Amir Sharon[4], Yifat Idan[4], Liat Oren-Young[4], Christian M. Sieber[5], Ondřej Novák[6], Aleš Pěnčík[6], Danuše Tarkowská[6], Kristýna Hromadová[6], Stanley Freeman[7], Marcel Maymon[7], Meirav Elazar[7], Sahar A. Youssef[8], El Said M. El-Shabrawy[8], Abdel Baset A. Shalaby[8], Petra Houterman[9], Nelson L. Brock[10], Immo Burkhardt[10], Elena A. Tsavkelova[11], Jeroen S. Dickschat[10], Petr Galuszka[6], Ulrich Güldener[12,*], and Bettina Tudzynski[1,*]

[1]Institut für Biologie und Biotechnologie der Pflanzen, Molecular Biology and Biotechnology of Fungi, Westfälische Wilhelms-Universität Münster, Münster, Germany

[2]Institute of Bioinformatics and Systems Biology, Helmholtz Zentrum München, German Research Center for Environmental Health (GmbH), Neuherberg, Germany

[3]United States Department of Agriculture, National Center for Agricultural Utilization Research, Peoria, Illinois

[4]Department of Molecular Biology and Ecology of Plants, Tel Aviv University, Tel Aviv, Israel

[5]Department of Energy Joint Genome Institute, University of California, Walnut Creek, Berkeley, California

[6]Centre of the Region Hana for Biotechnological and Agricultural Research, Palacky University, Olomouc, Czech Republic

[7]Department of Plant Pathology and Weed Research, Agricultural Research Organization (ARO), The Volcani Center, Bet-Dagan, Israel

[8]Plant Pathology Research Institute, Agricultural Research Center, Giza, Egypt

[9]University of Amsterdam, Swammerdam Institute for Life Sciences, Plant Pathology, Amsterdam, The Netherlands

[10]Rheinische Friedrich-Wilhelms-Universität Bonn, Kekulé-Institut für Organische Chemie und Biochemie, Germany

[11]Department of Microbiology Faculty of Biology, Lomonosov Moscow State University, Moscow, Russia

[12]Department of Genome-oriented Bioinformatics, Wissenschaftszentrum Weihenstephan, Technische Universität München, Maximus-von-Imhof-Forum 3, Freising, Germany

†These authors contributed equally to this work.

*Corresponding authors: E-mails: tudzynsb@uni-muenster.de; u.gueldener@tum.de.

Accepted: October 21, 2016

Data deposition: The genome and annotation data was submitted to the European Nucleotide Archive (ENA, *F. mangiferae*: http://www.ebi.ac.uk/ena/data/view/FCQH01000001-FCQH01000254; *F. proliferatum* ET1: http://www.ebi.ac.uk/ena/data/view/FJOF01000001-FJOF01000032; *F. proliferatum* 62905: http://www.ebi.ac.uk/ena/data/view/FCQG01000001-FCQG01000155).

Abstract

Species of the *Fusarium fujikuroi* species complex (FFC) cause a wide spectrum of often devastating diseases on diverse agricultural crops, including coffee, fig, mango, maize, rice, and sugarcane. Although species within the FFC are difficult to distinguish by morphology, and their genes often share 90% sequence similarity, they can differ in host plant specificity and life style. FFC species can also produce structurally diverse secondary metabolites (SMs), including the mycotoxins fumonisins, fusarins, fusaric acid, and beauvericin, and the phytohormones gibberellins, auxins, and cytokinins. The spectrum of SMs produced can differ among closely related species, suggesting that SMs might be determinants of host specificity. To date, genomes of only a limited number of FFC species have been sequenced. Here, we provide draft genome sequences of three more members of the FFC: a single isolate of *F. mangiferae*, the cause of mango malformation, and two isolates of *F. proliferatum*, one a pathogen of maize and the other an orchid endophyte. We compared these genomes to publicly available genome sequences of three other FFC species. The comparisons revealed species-specific and isolate-specific differences in the composition and expression (*in vitro* and *in planta*) of genes involved in

SM production including those for phytohormone biosynthesis. Such differences have the potential to impact host specificity and, as in the case of *F. proliferatum*, the pathogenic versus endophytic life style.

Key words: *Fusarium fujikuroi* species complex, genome sequencing, secondary metabolism, in planta expression, metabolomics, evolution.

Introduction

Fusarium is a species-rich genus of filamentous fungi that collectively can cause agriculturally significant diseases on virtually all crop plants. The plant diseases include numerous vascular wilts and rots of seedlings, tubers, bulbs, and corms as well as branch dieback and malformations of plant organs including roots, stalks, leaves, and flower- and seed-bearing structures (Leslie and Summerell 2006). During growth, these fungi produce secondary metabolites (SMs), including mycotoxins (e.g., fumonisins, moniliformin, fusaric acid, fusarins, trichothecenes, and zearalenone) that are of concern to food/feed safety and agricultural production. Ingestion of mycotoxin-contaminated food or feed is associated with a variety of animal diseases as well as cancer and neural tube defects in humans (Marasas et al. 2004; Leslie and Summerell 2006; Morgavi and Riley 2007). *Fusarium*-induced crop diseases as well as mycotoxin contamination problems result in significant economic losses to world agriculture every year (Wu 2007). A remarkable and unique aspect of *Fusarium* is that several species can synthesize multiple classes of phytohormones, including gibberellins (GAs) (Bömke and Tudzynski 2009; Troncoso et al. 2010) and auxins (Tsavkelova et al. 2012), that contribute to plant disease.

DNA-based phylogenetic analyses have resolved *Fusarium* species into a monophyletic lineage consisting of 20 species complexes that include almost 300 phylogenetically distinct species (Geiser et al. 2013; Ma et al. 2013; O'Donnell et al. 2013) (fig. 1*A*). The *Fusarium fujikuroi* species complex (FFC) includes at least 50 distinct species that group into three clades: the American, African, and Asian clades (O'Donnell et al. 1998a,b). FFC species occur on a wide range of plants and vary in host specificity. For example, the American-clade species *F. circinatum* and the Asian-clade species *F. mangiferae* are tree pathogens. *F. circinatum* causes pitch canker of multiple pine species and Douglas fir (Wingfield et al. 2008) whereas *F. mangiferae* causes malformation of flowers and shoots of mango trees (Ploetz and Freeman 2009; Freeman et al. 2014). Neither species is reported to occur widely or cause serious diseases on other plants (Britz et al. 2002; Leslie and Summerell 2006; Wingfield et al. 2008). In contrast, the Asian-clade species *F. fujikuroi* and African-clade species *F. verticillioides* are cereal pathogens (Leslie and Summerell 2006; Ma et al. 2010). *F. fujikuroi* causes *bakanae* disease of rice, which is characterized by seedling elongation resulting from production of GAs by the fungus; whereas *F. verticillioides* causes ear and stalk rot of maize. Both species can be recovered from other crops with some frequency, but neither typically causes severe disease on other crops. The Asian-clade species *F. proliferatum* is notable for its ability to cause significant levels of disease on a wide range of crops and other plants (Proctor et al. 2009). All five of these species have some mycotoxin biosynthetic genes and/or have been reported to produce some mycotoxins, however, production of fumonisin mycotoxins by *F. proliferatum* and *F. verticillioides* has received greater attention than other mycotoxins and species (Desjardins 2006; Leslie and Summerell 2006; Ma et al. 2010).

In fungi, SM biosynthetic genes are typically located adjacent to one another in gene clusters. Such biosynthetic gene clusters typically include a gene that encodes a core enzyme responsible for catalyzing the first committed step in an SM biosynthetic pathway: i.e., a polyketide synthase (PKS), nonribosomal peptide synthetase (NRPS), terpene cyclase (TC), or a dimethylallyl tryptophan synthase (DMATS) (Hoffmeister and Keller 2007). The clusters also usually include genes responsible for structural modifications of the molecule synthesized by the core enzyme (e.g., reduction, oxygenation, and acylation). In addition, the clusters often contain a gene encoding a transcription factor that regulates expression of genes within the same cluster, and a gene encoding a transport protein that facilitates export of the corresponding SM(s) out of the cell. Another characteristic of SM biosynthetic gene clusters is that genes within the same cluster exhibit similar patterns of expression (Hoffmeister and Keller 2007; Brakhage 2013). Thus, a region of a chromosome consisting of continguous, co-expressed genes that encode proteins commonly involved in SM biosynthesis (e.g., a core enzyme, modifying enzyme, transcription factor, and/or transporter) likely consistute a SM biosynthetic gene cluster.

At the start of the current study, genome sequences of only eight *Fusarium* species were publicly available: the FFC species *F. circinatum*, *F. fujikuroi* and *F. verticillioides*, and the nonFFC species *F. avenaceum, F. graminearum*, *F. oxysporum*, *F. pseudograminearum*, and *F. solani* (Cuomo et al. 2007; Coleman et al. 2009; Ma et al. 2010; Gardiner et al. 2012; Wingfield et al. 2012; Wiemann et al. 2013; Lysøe et al. 2014). *In silico* analyses indicated that each of these fusaria has the genetic potential to produce 30–60 structurally distinct families of SMs. The *F. fujikuroi* genome has been studied most comprehensively with respect to secondary metabolism. Expression studies by microarray and chromatin immunoprecipitation-sequencing ChIP-Seq coupled with extensive SM chemical analysis led to the identification of at least 46 SM gene clusters, each of which is likely required for synthesis of a

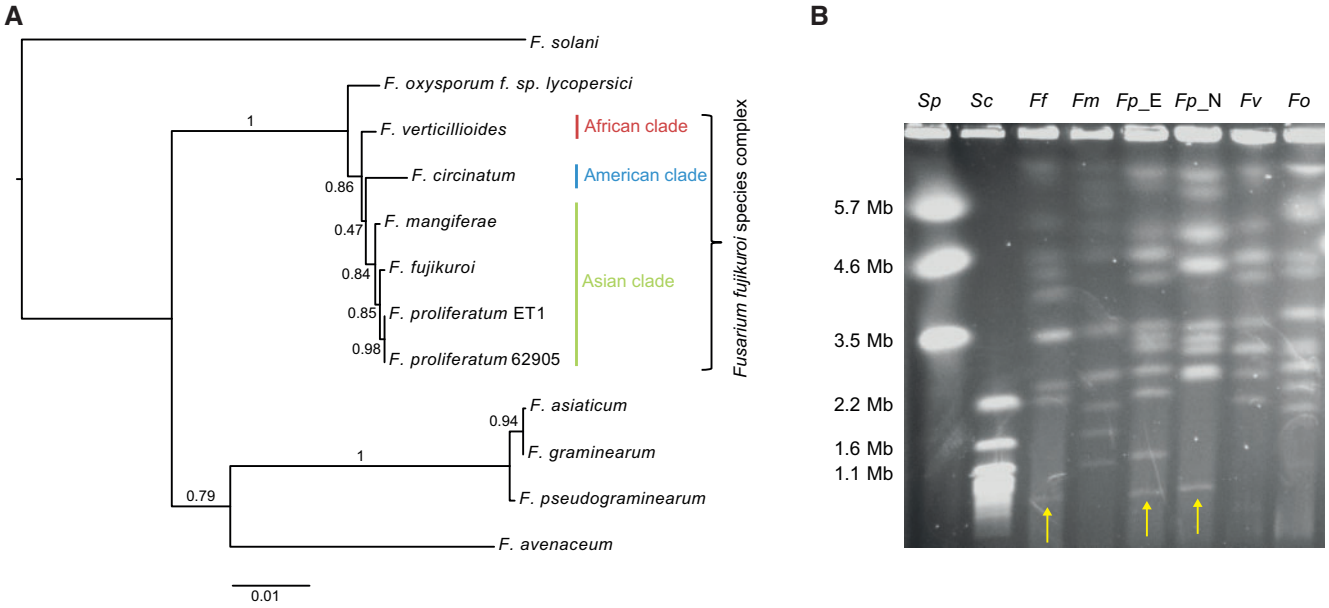

Fig. 1.—Phylogenetic analyses and separation of chromosomes of five *Fusarium* strains of the FFC and *F. oxysporum* as outgroup on a CHEF gel. (*A*) Phylogenetic analyses of different *Fusarium* strains. It was calculated based on the protein sequences of RNA polymerase II subunits RPB1 (FPRO_01340), RPB2 (FPRO_08166), and of the transcription elongation factor gene TEF1α. (FPRO_08841). (*B*) Chromosomes of *Schizzosaccharomyces pombe* (*Sp*) and *Saccharomyces cerevisiae* (*Sc*) were used as size standards; *Ff* = *F. fujikuroi*, *Fm* = *F. mangiferae*, *Fp*_E = *F. proliferatum* ET1, *Fp*_N = *F. proliferatum* 62905, *Fv* = *F. verticillioides*, *Fo* = *F. oxysporum*. Chromosome XII is marked with yellow arrows.

structurally distinct family of SMs. However, the SM product(s) of only 14 of the gene clusters have been reported thus far, probably due to no or minimal gene expression during growth in the lab (Wiemann et al. 2013). An additional PKS cluster has been recently identified in two other *F. fujikuroi* strains, KSU X-10626 and B14 which is not present in strain IMI58289 (Chiara et al. 2015). Although a majority of the putative SM gene clusters occur in other FFC species, some appear to be unique to a single species or even a single isolate of one species. For example, the NRPS31 cluster which is responsible for apicidin F biosynthesis (Wiemann et al. 2013; Niehaus et al. 2014a; von Bargen et al. 2015) is unique to *F. fujikuroi*. However, novel *F. fujikuroi* genome sequences showed that this cluster is not present in all strains of this species (Chiara et al. 2015).

The objective of the current study was to identify all potential SM biosynthetic gene clusters in a collection of phylogenetically closely related *Fusarium* species and to analyze their expression under *in vitro* and *in planta* conditions. To meet this objective, we conducted comparative genomic, transcriptomic and metabolomic analyses of five members of the FFC: one strain each of *F. fujikuroi*, *F. mangiferae*, and *F. verticillioides*, and two strains of *F. proliferatum* that were isolated from different plants. The analyses employed genome sequences for *F. mangiferae* strain MRC 7560 (Ploetz and Freeman 2009; Freeman et al. 2014) and *F. proliferatum* strains ET1 (Tsavkelova et al. 2003a) and NRRL 62905 (Munkvold et al. 1998), which were generated during the course of this study, as well as genome sequences for *F. fujikuroi* and *F. verticillioides*, which were generated previously (Ma et al. 2010; Wiemann et al. 2013). Multiple analyses also employed the previously generated genome sequence data from the closely related, but nonFFC species *F. oxysporum* (Leslie and Summerell 2006; Ma et al. 2010). We compared *in planta* expression profiles of the five FFC members to identify SM pathways that are commonly induced in all isolates and those that are expressed strain- or species-specific. For the first time, we provide a detailed analysis of the genetic potential of the FFC members to produce three classes of phytohormones, GAs, auxins, and cytokinins, using common or alternative biosynthetic pathways including two alternate biosynthetic pathways for auxin and one for cytokinin biosynthesis. These comprehensive analyses revealed both high levels of similarity as well as marked differences among and within species. These differences included variation in genome organization, SM (including phytohormones) production profiles, and expression of SM genes *in vitro* and *in planta*.

Results and Discussion

Three Novel FFC Genomes: *F. proliferatum* ET1, *F. proliferatum* 62905 and *F. mangiferae*

F. proliferatum strain ET1 was isolated from the roots of a tropical epiphytic orchid, *Dendrobium moschatum* (Buch.–Ham.) Swartz., cultivated in the Stock Greenhouse of the

Main Botanical Garden (Moscow, Russia) (Tsavkelova et al. 2003a,b). The plant from which ET1 was isolated appeared healthy, without any symptoms of blight, wilt, and root rot, discoloration, softening, or decay suggesting that *D. moschatum* and ET1 have an endophytic relationship. *F. proliferatum* strain 62905 was isolated from a kernel of a maize (*Zea mays*) ear with symptoms of ear rot and that was grown in Iowa, USA (Munkvold et al. 1998). *F. mangiferae* strain MRC 7560 was isolated from a malformed inflorescence of a mango (*Mangifera indica*, cultivar Kent) plant grown in Ginosar, Israel (Britz et al. 2002; Ploetz and Freeman 2009; Freeman et al. 2014).

Genome sequence data for *F. mangiferae* 7560 and *F. proliferatum* ET1 and 62905 was generated with a whole-genome shotgun approach using 454 pyrosequencing and Illumina HiSeq 2000 technologies (see "Material and Methods" section). The assembly processes resulted in 254 scaffolds for *F. mangiferae* (46.3 Mb, 104-fold coverage), 32 scaffolds for *F. proliferatum* ET1 (45.2 Mb, 100-fold coverage), and 155 scaffolds for *F. proliferatum* 62905 (43.2 Mb, 98-fold coverage). A total of 15,804 (*F. mangiferae*), 16,143 (*F. proliferatum* ET1), and 15,254 (*F. proliferatum* 62905) gene loci were annotated by a combination of prediction tools and manual validation. The most in-depth studied genome of a member of the FFC to date is the recently published genome of *F. fujikuroi* which was assembled into near-complete chromosome sequences (Wiemann et al. 2013). We therefore used the genome of *F. fujikuroi*, but also those of *F. verticillioides*, *F. oxysporum*, and *F. graminearum* (Cuomo et al. 2007; Ma et al. 2010; Wiemann et al. 2013) as a reference for subsequent comparative studies.

Table 1 summarizes physical genome features of the newly and previously sequenced FFC members, *F. oxysporum* and the distantly related *F. graminearum*. Most key genome

Table 1
Comparative Genome Statistics

	Asian-clade of the FFC				American-clade	African-clade	Outgroups	
	F. proliferatum 62905	*F. proliferatum* ET1	*F. mangiferae* MRC7560	*F. fujikuroi* IMI58289	*F. circinatum* FSP 34	*F. verticillioides* 7600	*F. oxysporum* 4287	*F. graminearum* PH-1
Genome size (Mb)	43.2	45.2	46.2	43.9	44.3	41.8	61.4	36.4
Chromosomes	12 (?)	12 (?)	11 (?)	12	n.d.	11	15	4
Scaffolds	155	32	254	12	4509	36	114	31
N_{50} scaffold (Mb)	0.9	3.3	2.5	4.2	0.025	2	2	5.4
GC-content (%)	48.7	48.1	47.6	47.4	47.3	48.6	47.3	48.0
Protein coding genes (splice variants)	15,254	16,143	15,804	14,817	15,022	15,869	20,925	13,826
Gene density (Number of genes per Mb)	353	358	343	337	339	380	341	380
Total coding exon length (Mb)	22.1	23.1	22.9	21.5	19.7	21.4	21.5	26.1
Total intron length (Mb)	1.8	1.9	1,8	1.9	1.8	1.8	1.8	2.6
Average distance between genes (kb)	1.3	1.3	1,4	1.4	1.5	1.2	2.1	1.6
Percent coding (%)	51.2	51.1	49,4	49.2	48.44	51.1	42.5	56.75
GC-content coding (%)	51.5	51.5	51.6	51.5	51.7	51.6	51.4	51.6
Average coding transcript length (kb)	1.5	1.4	1.4	1.5	1.3	1.3	1.2	1.4
Mean protein length (amino acids)	483	475.1	481.2	484.7	436.5	448.5	409.7	416.6
Coding exons	42,531	45,034	43,998	41,583	41,023	42,304	54,197	38,453
Average exon length (bp)	519.8	512.9	519.2	518.1	479.20	504.9	481.5	488.83
Exons/gene	2.79	2.78	2.78	2.81	2.73	2.66	2.6	2.78
Introns	27,277	28,831	28,175	26,769	26,001	26,435	33,272	24,627
Average intron length (bp)	64.8	66.1	64.9	69.2	68.52	69.6	79.0	76.62

features of the three newly sequenced FFC members are similar to those of *F. fujikuroi* (Wiemann et al. 2013). The completeness of the three draft genomes was documented comparing each predicted proteome to two different, highly conserved eukaryote protein sets by BLAST (Aguileta et al. 2008; Parra et al. 2009). To estimate protein conservation and functional coverage, the proteomes were compared against each other and the Swiss-Prot database by BLAST. Pair-wise BLAST analysis to the *F. fujikuroi* proteome revealed that 88.2% of *F. mangiferae*, 91.1% of *F. proliferatum* ET1, and 90.8% of *F. proliferatum* 62905 proteins are similar to proteins in *F. fujikuroi*, whereas ~10% of the proteins have very low (*e*-value $\leq 10^{-5}$) or no similarity and are thus likely strain specific. *F. verticillioides* was most dissimilar to *F. fujikuroi* with 18% of its proteins sharing low or no similarity to *F. fujikuroi* proteins. Blast analysis against the Swiss-Prot database of the proteomes revealed that only about 32% have a likely homolog in Swiss-Prot.

Despite the apparent close relationship of the three newly sequenced fusaria to the other sequenced members of the FFC (*F. fujikuroi*, *F. circinatum*, and *F. verticillioides*) and the more distantly related species *F. oxysporum* (fig. 1*A*), we expected some karyotypic diversity. Visualization of chromosome content of the six strains by pulse field gel electrophoresis (PFGE) combined with clamped homogeneous electric fields (CHEF) verified that all six strains contained multiple chromosomes of varying sizes. The exact number of chromosomes in each strain could not be determined. A comparison of chromosome size between strains indicated that most chromosomes differed substantially in length (fig. 1*B*). In particular, the small chromosome XII was found to be present in some but not all *F. fujikuroi* isolates and some other species of the FFC (Xu and Leslie 1996; Wiemann et al. 2013). Chromosome XII is also present in the two *F. proliferatum* strains but not in *F. mangiferae, F. verticillioides*, and *F. oxysporum*. The CHEF gel analysis also revealed chromosome size differences in addition to what we previously noted (e.g., chromosome IV of *F. fujikuroi* is missing 285 and 820 kb from the left and right arms respectively, as compared with chromosome IV of *F. verticillioides* (Wiemann et al. 2013).

Despite missing chromosomal segments relative to each other, the present orthologous genes are highly collinear compared with *F. fujikuroi*. The colinearity to *F. fujikuroi*, as estimated by OrthoCluster (Vergara and Chen 2009), is high for all four genomes: 0.936 for *F. mangiferae*; 0.995 for *F. proliferatum* ET1; 0.978 for *F. proliferatum* 62905, and 0.963 for *F. verticillioides*.

Genome-wide comparisons of gene families that might contribute to host specificity, such as transcription factors (TFs), transporters, cell-wall degrading enzymes, monooxygenases, and secreted and effector-like proteins, revealed differences in gene numbers per family among the FFC examined (i.e., *F. mangiferae*, *F. proliferatum* ET1 and *F. proliferatum* 62905, *F. fujikuroi*, and *F. verticillioides*) and the outgroup *F. oxysporum* (table 2).

Whereas the number of TFs is similar for all FFC members, the outgroup *F. oxysporum* encodes considerably more TFs (1645) consistent with its larger genome (Rep and Kistler 2010) (table 2). The number of genes encoding putative secreted proteins is very similar within the FFC with the greatest number present in the orchid-associated endophyte *F. proliferatum* ET1. This distribution is consistent in all subclasses of secreted protein-encoding genes, namely, classical, nonclassical which lack an N-terminal signal peptide (Bendtsen et al. 2004), and small secreted proteins (SSPs) containing $\leq$300 amino acids (Gohari et al. 2015; Lo Presti et al. 2015). The larger number of SSPs with a potential effector-like function seems to be a specific feature of *F. oxysporum* (Gawehns et al. 2014; Williams et al. 2016).

To investigate orthologous clusters of genes, deduced proteomes of the five FFC members were subjected to a QuartetS analysis (Yu et al. 2011). The genes group into 16,104 orthology clusters. *F. proliferatum* has genes in the most clusters (95.3% for ET1 and 92.5% for 62905), whereas *F. mangiferae* has genes in 91.4%, *F. fujikuroi* has genes in 89.3%, and *F. verticillioides* has genes in 85.9% of the clusters. Because *F. proliferatum* has genes in most of the clusters, it may represent the ancestral FFC genome of this group. The analysis also revealed that 2.9, 7.0, 5.2, 2.2, and 12.6% of the predicted genes in the *F. fujikuroi*, *F. mangiferae*, *F. proliferatum* ET1, *F. proliferatum* 62905, and *F. verticillioides* genome, respectively, have no ortholog in the genomes of the other FFC species examined, and therefore are unique to these species/strains.

We also determined the relative abundance of repetitive DNA by identifying *de novo* repeat elements using RepeatScout (Price et al. 2005) followed by an assessment of known repeat elements. According to RepBase (Jurka et al. 2005) transposable elements were classified into the four main transposable element categories using TEclass: DNA transposon, long interspersed nuclear element (LINE), short interspersed nuclear element (SINE), and retrotransposon with long terminal repeats (LTRs) (Abrusán et al. 2009). The FFC members examined had an overall coverage of repeat DNA between 3.74% and 8.16%. In contrast, the genome of the outgroup species *F. oxysporum* contains 21.46% repeats, which contribute to its larger genome (61.4 Mb) (Ma et al. 2010).

In summary, comparative analysis of the five FFC genomes revealed highly similar sets of genes and a high level of collinearity among the five *Fusarium* isolates. However, we also identified subsets of genes unique to each species that may be at least partially responsible for the differences among them, e.g., in host specificity or pathogenic versus endophytic life style.

Table 2

Distribution of Selected Gene Families and Other Genetic Elements in Genome Sequences of Six *Fusarium* Species

Gene family	Asian-clade				African-clade	Outgroup
	F. proliferatum 62905	*F. prolifertatum* ET1	*F. mangiferae* MRC7560	*F. fujikuroi* IMI58289	*F. verticillioides* 7600	*F. oxysporum* 4287
Secondary metabolite biosynthetic genes[a]						
PKS	17 (3|0)	17 (2|0)	15 (2|1)	14 (1|0)	13 (0|2)	13 (3|0)
PKS/NRPS	4 (1|0)	4 (0|0)	4 (1|0)	4 (0|0)	4 (1|0)	3 (0|0)
NRPS	16 (0|0)	20 (0|1)	17 (1|0)	16 (0|1)	18 (3|0)	17 (0|1)
DMATS	3 (1|0)	3 (0|0)	4 (0|1)	2 (0|0)	3 (0|0)	3 (0|0)
TC	10 (0|0)	12 (1|1)	12 (1|0)	12 (3|0)	10 (2|0)	10 (0|0)
Secreted protein-encoding genes						
Total secreted proteins (SP)	1,672	1,825	1,746	1,566	1,684	2,041
Small secreted proteins (< 300aa) (SSP)	547	634	569	516	557	716
Transporter genes	930	954	938	879	901	1,074
Transcription factors (TFs)	1,116	1,254	1,257	1,125	1,136	1,644
Zn(2)C6 fungal type TFs (largest group of TFs)	527	583	584	510	537	677
Other gene families						
Histone modifying enzymes	112	129	127	119	114	166
Cytochrome P450	171	185	159	143	145	180
Monooxygenases	86	89	78	69	70	84
Oxidoreductases	46	40	41	47	44	49
Protein kinases and phosphatases	253	293	294	251	253	346
Dehydrogenases	392	409	385	370	363	470
Glycoside hydrolases	331	350	343	314	335	410
Repetitive elements						
DNA Transposon	0.23%	0.31%	0.45%	1.06%	0.26%	3.94%
LINE	0.05%	0.05%	0.16%	0.10%	0.05%	4.05%
SINE	0.04%	0.05%	0.05%	0.05%	0.05%	0.12%
LTR Retrotransposon	0.30%	0.44%	0.41%	2.31%	0.52%	9.14%
Unclassified nonLTR-Retrotransposon	0.01%	0.01%	0.01%	0.00%	0.00%	0.04%
Unclassified Retrotransposon	0.16%	0.22%	0.21%	0.44%	0.15%	1.84%
Unclassified	0.04%	0.05%	0.07%	0.26%	0.05%	0.60%
Total TEclass	0.80%	1.09%	1.33%	4.18%	1.07%	19.25%
Simple sequence repeats	0.60%	0.62%	0.62%	0.65%	0.64%	0.42%
Tandem repeats	2.78%	3.43%	4.12%	5.60%	3.04%	2.76%
Total repeat coverage	3.74%	4.58%	5.29%	8.16%	4.22%	21.46%

[a]SM gene predictions are based on InterPro domains, manually validated and corrected based on reports in the literature and on comparative analysis of fusaria. Values are the number of genes per genome; values in brackets are (number of pseudogenes/remnants | genes unique to a genome among the species examined). PKS—polyketide synthase; NRPS—nonribosomal peptide synthetase; DMATS—dimethylallyl tryptophan synthase; TC—terpene cyclase.

Comparative Analysis of Secondary Metabolite Gene Clusters

Despite the close phylogenetic relationships of species of the FFC, they differ considerably in the presence and absence of SM genes in their genomes (supplementary table S1, Supplementary Material online). There are 21 NRPS-, 24 PKS-, 4 DMATS-, and 13 TC-encoding genes present in the five FFC members, i.e., 62 unique core-enzyme encoding genes that could give rise to 60 distinct families of SMs (the fusaric acid cluster encodes two core enzymes, PKS6 and NRPS34; two DTCs for GA biosynthesis in *F. proliferatum* ET1). Two of the 24 PKS genes (*PKS15* and *PKS39*) are only present in the genome of *F. verticillioides*, the member of the African clade. Four additional core enzyme-encoding genes (*PKS12a-1*, *PKS12a-2*, *PKS41*, and *NRPS32*) are not present in any genome of the FFC members, but only in the outgroup *F. oxysporum* (supplementary table S1, Supplementary Material online).

Differences in SM production potential are also apparent among the Asian clade members, *F. mangiferae*, *F. proliferatum*, and *F. fujikuroi* (supplementary table S1, Supplementary Material online). For instance, *NRPS17* is absent in *F. mangiferae*, but present in the other Asian clade members, and *NRPS36* is present only in the two *F. proliferatum* strains and the outgroup *F. oxysporum*. There are also differences within *F. proliferatum* as the two strains differ in NRPS, PKS, DMATS,

and TC gene content; *NRPS1*, *NRPS24*, *NRPS26*, *NRPS35*, and *STC7* are present in *F. proliferatum* ET1, but not in *F. proliferatum* 62905. In some cases, pseudogenes are present in strain *F. proliferatum* 62905 whereas *F. proliferatum* ET1 has a functional copy (*PKS10*, *PKS40*, and *DMATS3*) (supplementary table S1, Supplementary Material online). The most striking difference between the two *F. proliferatum* strains is the presence of two complete diterpene gene clusters for GA biosynthesis and three instead of two cytokinin gene clusters in the genome of *F. proliferatum* ET1.

In general, the number of NRPS genes per *Fusarium* genome analyzed (including *F. oxysporum*) ranges from 16 to 20, the number of PKS genes from 13 to 17, the number of hybrid (PKS/NRPS) genes from 3 to 4, the number of DMATSs from 2 to 4, and the number of TCs per genome from 10 to 12. In contrast to these *Fusarium* genomes, three recently sequenced *F. avenaceum* strains encode 75–80 SM core enzymes, including 25–27 PKSs, 25–28 NRPSs, four DMATSs, and 12–13 TCs (Lysøe et al. 2014).

SM products have been assigned only to 23 of the 60 predicted PKS, NRPS, DMATS, and TC-derived SM gene clusters in members of the FFC due likely to the fact that most are not or only minimally expressed under laboratory conditions (supplementary table S1, Supplementary Material online) (Brakhage 2013). More than half (36) of these predicted clusters are present in all FFC genomes and for 14 of them the products are known: four NRPS genes required for synthesis of the siderophores ferricrocin and fusarinine, the insecticidal metabolite beauvericin and fusaric acid (together with PKS6); six PKS genes required for synthesis of fusarubins, bikaverin, fusaric acid (together with NRPS8), fusarins, equisetin, and gibepyrone; three STCs required for synthesis of (–)-α-acorenol, eremophilene, and (+)-koraiol; one triterpene cyclase (TrTC) involved in lanosterol biosynthesis, and one tetraterpene cyclase (TeTC) required for synthesis of carotenoids (supplementary table S1, Supplementary Material online).

A total of 85 full-length, and apparently functional PKS genes, from the genomes of the four species of the FFC, *F. oxysporum* 4287 and three additional *F. oxysporum* strains were used to generate a phylogenetic tree. Phylogenetic analysis of these genes resolved them into 27 well-supported clades (supplementary fig. S1, Supplementary Material online). About 26 of the 27 clades correspond to PKS orthology groups, or clades, recently described by Brown and Proctor (2016) and Hansen et al. (2015). Based on the chemical structure of the polyketides produced by some of the PKSs and on PKS domain structures, the PKSs in each clade are likely required for synthesis of a structurally distinct polyketide (supplementary fig. S1, Supplementary Material online). The two *F. oxysporum* PKSs in clade 12, FOXG_14850 and FOXG_15886, are likely paralogues and thus likely synthesize the same polyketide (Brown et al. 2012a). Among the 27 orthology groups, some are unique to one of the species examined, others are shared by some but not all of the FFC species, and still others are conserved among the four FFC species (supplementary table S1, Supplementary Material online). Nine of the orthology groups consist of PKS genes that are conserved and potentially functional among all four FFC species exmined. Eight of these nine groups are also present in *F. oxysporum*, indicating distribution of the genes beyond the FFC (supplementary table S1, Supplementary Material online). In contrast, three orthology groups consist of a PKS gene that occurs only once among the four FFC species examined: *PKS43* (FMAN_09800) in *F. mangiferae*; and *PKS39* (FVEG_14911) and *PKS15* (FVEG_05537) in *F. verticillioides*. However, BLAST analysis against the NCBI GenBank database indicated that there are closely related homologs of these genes in more distantly related species; e.g., homologs of *PKS43* occur in the *F. avenaceum*, *F. graminearum* and *F. pseudograminearum* genomes (Brown and Proctor 2016).

PKS genes in 11 orthology groups exhibit a discontinuous distribution in that their occurrence is the same in some more distantly related species but different in some more closely related species. This distribution likely reflects gene loss events, resulting from pseudogenization and/or deletion, and gene gain events, e.g., via horizontal gene transfer (HGT) (Akagi et al. 2009). For example, a functional full-length homolog of *FUS1*, the PKS-NRPS gene (PKS10) required for synthesis of fusarin mycotoxins, is present in *F. fujikuroi*, *F. proliferatum* ET1, and *F. verticillioides*. However, a deletion event has likely resulted in the presence of only half of *FUS1* and none of the eight other *FUS* cluster genes in *F. proliferatum* 62905 and *F. mangiferae* (fig. 2*A*).

Phylogenetic analysis indicated that the partial *FUS1* homolog in *F. proliferatum* 62905 is more closely related to the partial homolog in *F. mangiferae* than to the homolog in *F. proliferatum* ET1 (supplementary fig. S2, Supplementary Material online). In addition, the four genes in the 5′ flanking regions of *FUS1* in 62905 and *F. mangiferae* are homologous, but not homologous to genes in the corresponding region of ET1 (fig. 2*A*). These findings indicate that partial loss of *FUS1*, as well as the rest of the *FUS* cluster did not occur independently in *F. mangiferae* and *F. proliferatum*. A possible explanation for this finding is that the partial deletion of the *FUS* cluster occurred in a common ancestor of *F. mangiferae* and *F. proliferatum* in such a way that it resulted in two alleles of the *FUS* cluster: an intact cluster allele and a partially deleted cluster allele. Subsequently, the two alleles would have been retained during at least one speciation event such that *F. proliferatum* acquired both alleles by vertical inheritance from the ancestral species. PCR analysis of a collection of 36 additional isolates of *F. proliferatum* from a variety of hosts and geographic locations (Proctor et al. 2009) indicated that 33 of the strains have an intact *FUS* cluster, and only three isolates have the partially deleted *FUS* cluster.

Other examples of gene loss and pseudogenization events that likely contribute to biosynthetic diversity of the fusaria

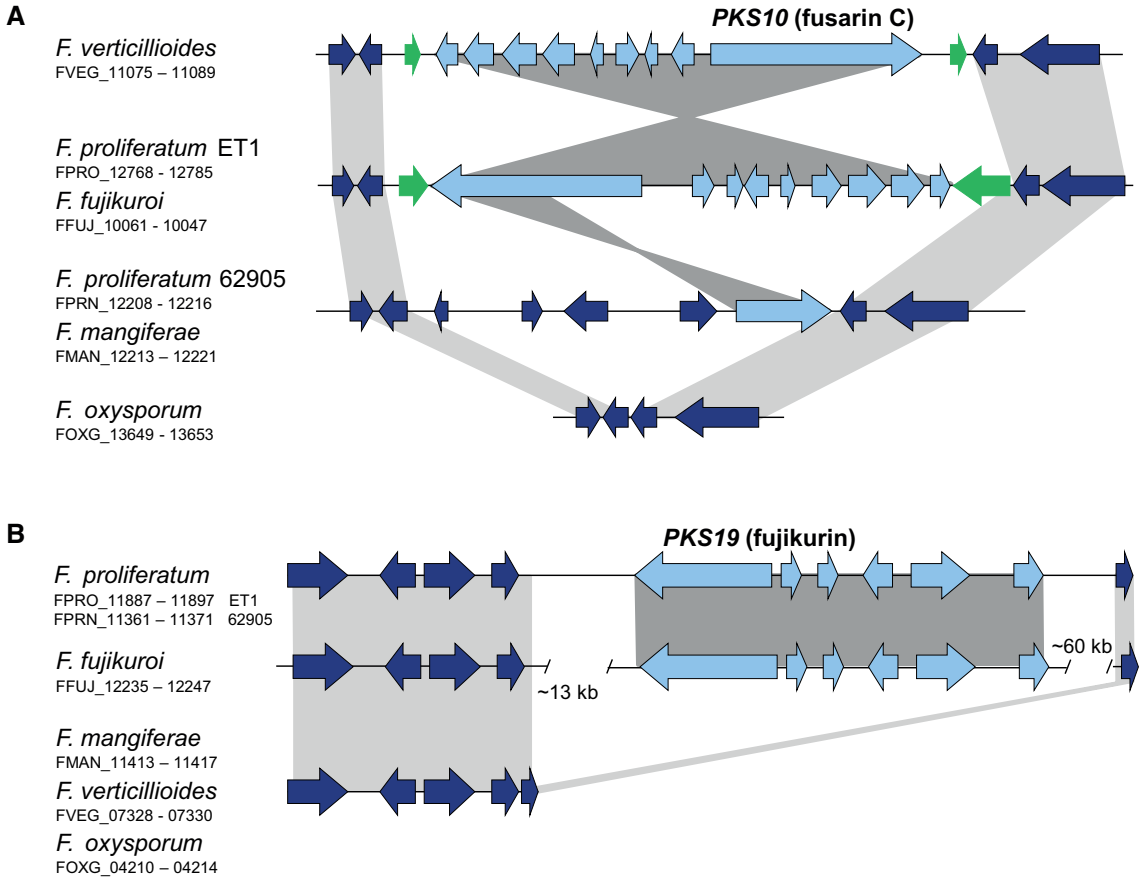

Fig. 2.—Examples of functional and pseudogenized PKS genes in *Fusarium* spp. (*A*) The fusarin C (*FUS1*; PKS10) gene cluster is intact only in *F. proliferatum* ET1, *F. fujikuroi* and *F. verticillioides*. (*B*) PKS19 cluster involved in fujikurin biosynthesis in *F. fujikuroi* is also present in *F. proliferatum*.

examined involved *PKS12*, *PKS15*, *PKS38*, and *PKS39*. *PKS12* is potentially functional in *F. mangiferae*, *F. fujikuroi*, and *F. verticillioides*, but is pseudogenized in both *F. proliferatum* isolates and *F. oxysporum* (supplementary fig. S3, Supplementary Material online). Furthermore, the *PKS39* cluster, which is most likely responsible for the synthesis of depudecin is pseudogenized in *F. proliferatum*, *F. mangiferae* and *F. fujikuroi* but intact in *F. verticillioides* and four of 12 *F. oxysporum* strains (Wight et al. 2009; Brown and Proctor 2013; Brown and Proctor 2016) (supplementary fig. S3, Supplementary Material online). *PKS38* and four genes immediately upstream of it are present in the two *F. proliferatum* strains, whereas *F. mangiferae* has only a truncated *PKS38* gene (supplementary fig. S3, Supplementary Material online). In contrast, the synteny of genes outside the region occupied by *PKS38* and its four flanking genes is largely conserved among the other FFC members and in *F. oxysporum*. These results and the predicted functions of the four *F. proliferatum* genes upstream of *PKS38* suggest that together the five genes constitute a SM biosynthetic gene cluster that is unique to *F. proliferatum*. Likewise, the *PKS15* gene (FVEG_05537) and seven genes upstream of it are unique to *F. verticillioides*, but there is conservation of gene synteny in the region flanking *PKS15* in the other FFC members and *F. oxysporum* (data not shown).

An example of a potential recent acquisition of a PKS gene is *PKS19*. Previous functional analyses in *F. fujikuroi* indicated that *PKS19* and five genes flanking it constitute an SM gene cluster responsible for synthesis of fujikurin (Wiemann et al. 2013; von Bargen et al. 2015). Comparative analyses in the current study indicate that the fujikurin cluster is present in *F. fujikuroi* and both *F. proliferatum* strains but absent in *F. mangiferae*, *F. verticillioides*, and *F. oxysporum* (fig. 2*B*). These findings are consistent with a previous analysis in which *PKS19* was detected in *F. fujikuroi* and *F. proliferatum* but not in another Asian-clade species *F. sacchari*, or in any of five American-clade or two African-clade species examined (Brown and Proctor 2016).

In addition to variation in distribution within the FFC, PKSs, and associated gene clusters can also be located at different positions in the genome of producing organisms. One example is the fumonisin (*PKS11*) gene cluster. This cluster occurs in

the same genomic context in *F. proliferatum* and *F. fujikuroi*, at another genomic position in *F. verticillioides* (Proctor et al. 2013), but is absent in the genomes of *F. mangiferae* (supplementary table S1, Supplementary Material online). *F. oxysporum* O-1819 has an intact cluster, whereas the cluster is absent in 12 other sequenced *F. oxysporum* strains (Proctor et al. 2013). Recently, it has been shown that *Aspergillus niger* has an intact fumonisin gene cluster (Frisvad et al. 2007). Differences between the *A. niger* and *F. verticillioides* clusters highlight how a cluster can diverge through addition or loss of genes, or by rearrangement of gene positions within a cluster (Khaldi and Wolfe 2011).

A total of 104 NRPS genes were identified in the genomes of the five FFC members examined and *F. oxysporum* 4287. In a phylogenetic analysis, the genes were resolved into 36 orthology groups. Twelve of the potentially functional NRPS genes were found in all five FFC strains as well as *F. oxysporum* (supplementary table S1, Supplementary Material online).

The presence or absence of some NRPSs likely responsible for the production of siderophores reflects an underlying redundancy in iron uptake systems in fungi (Bushley and Turgeon 2010). *NRPS1*, *NRPS2*, *NRPS6*, *NRPS17*, and *NRPS36* are likely responsible for the production of malonichrom, fericrocin, fusarinine, ferrichrome, and ferrirhodin siderophores, respectively, based on sequence homology (Bushley and Turgeon 2010; Munawar et al. 2013; Hansen et al. 2015; Oide et al. 2015). Only two of the corresponding NRPS genes, *NRPS2* and *NRPS6*, are present in all five FFC members examined (supplementary table S1, Supplementary Material online). *NRPS1* is absent in the genome of one of the two *F. proliferatum* isolates (62905) and in *F. fujikuroi*. *NRPS17* is absent in *F. mangiferae* and *F. oxysporum*, whereas *NRPS36* is present only in both *F. proliferatum* isolates and in some *F. oxysporum* isolates (supplementary table S1, Supplementary Material online). These findings indicate that closely related FFC species and even strains within a species can vary in the presence and absence of iron acquisition systems.

NRPS34 is a noncanonical NRPS with similarity to fungal α-aminoadipate reductases (AAR) involved in lysine biosynthesis in fungi (Bushley and Turgeon 2010), and was recently shown to be essential for fusaric acid biosynthesis along with PKS6 encoded by the *FUB1* gene (Brown et al. 2015; Studt et al. 2016). After the initial descriptions of the five-gene (*FUB1–FUB5*) fusaric acid gene cluster in *F. verticillioides* and *F. fujikuroi* (Brown et al. 2012b; Niehaus et al. 2014b), seven additional clustered fusaric acid biosynthetic genes were identified about 14 kb from *FUB1–FUB5* in both fungi. These additional genes included the gene *FUB8* encoding NRPS34.

Wiemann et al. (2013) identified NRPS31 (FFUJ_00003) in the *F. fujikuroi* genome sequence but not in other *Fusarium* genome sequences that were available at the time. Overexpression of the TF of the otherwise silent gene cluster led to the production of a novel apicidin analog, apicidin F (Niehaus et al. 2014a). The apicidin cluster is also absent in both *F. proliferatum* isolates and *F. mangiferae* (supplementary fig. S4A and table S1, Supplementary Material online) as well as two newly sequenced *F. fujikuroi* genome sequences (Chiara et al. 2015). An apicidin cluster homolog was originally identified in the distantly related species *Fusarium semitectum* (Jin et al. 2010) and more recently in the genome of another distantly related species, *F. avenaceum* (Lysøe et al. 2014). The cluster in *F. avenaceum* has undergone extensive rearrangements, including loss of three genes and gain of three genes relative to the *F. fujikuroi* apicidin cluster. These differences in gene content of the *F. fujikuroi* and *F. avenaceum* clusters suggest that the two clusters yield different products (Lysøe et al. 2014) (supplementary fig. S4A, Supplementary Material online). HGT provides one possible explanation for the presence of apicidin-like clusters in such distantly related species and its absence in all other fusaria examined thus far.

In contrast to the apicidin *NRPS* gene, *NRPS22* (FFUJ_09296) required for synthesis of the mycotoxin beauvericin is present and apparently functional in four members of the FFC and *F. oxysporum* (López-Berges et al. 2013; Niehaus et al. 2016), but likely nonfunctional in *F. verticillioides* due to a N-terminal truncation (supplementary fig. S4B and table S1, Supplementary Material online). *NRPS22* was probably acquired from *Beauveria bassiana* or another beauvericin-producing fungus (Xu et al. 2008). Recently, the beauvericin gene cluster (*BEA1–BEA4*) has been functionally characterized in *F. fujikuroi*: the cluster consists of two genes involved in biosynthesis, and two additional genes encoding an ABC transporter and a cluster-specific transcription factor, respectively (Niehaus et al. 2016).

Collectively, the five FFC and *F. oxysporum* genomes examined have 12 orthology groups of genes that encode putative TCs or related enzymes involved in terpene biosynthesis (supplementary table S1, Supplementary Material online). These enzymes include nine sesquiterpene cyclases (STC1–STC9), one diterpene cyclases (*ent*-kaurene synthase, DTC1), one triterpene cyclase (lanosterol synthase, TrTC1), and one tetraterpene cyclase (phytoene synthase/lycopene cyclase, TeTC1). STCs are type I terpene cyclases and consist of ca. 350 amino acids that contain: an aspartate-rich motif (DDXXD) for substrate and Mg^{2+} cofactor binding; a highly conserved arginine near amino acid 180, which is proposed to be a pyrophosphate sensor; an NSE triad (NDXXSXXXE), which is usually 46 amino acids upstream of the conserved arginine; and the RY dimer near amino acid 330 (Dickschat 2015). Homologs of the *STC1–STC9* genes occur in all FFC strains examined and *F. oxysporum*, except for the following: *STC1* is absent in *F. oxysporum*, *STC8* is absent in both *F. proliferatum* strains; *STC5* is absent in *F. verticillioides*; and *STC7* is absent in *F. proliferatum* 62905 (supplementary table S1, Supplementary Material online). Phylogenetic analysis of the deduced amino acid sequences of these genes along with 90 STC genes from distantly related *Fusarium* species resolved them into 11 clades, or orthology groups (labled

STC1–STC11 (supplementary fig. S5, Supplementary Material online).

Functional studies of STC-encoding genes in *F. fujikuroi* IMI58289 revealed that STC6 is an (–)-α-acorenol synthase and STC4 is a (+)-koraiol synthase (Brock et al. 2011, 2013). Both sesquiterpene alcohols are the main constituents of *F. fujikuroi* headspace extracts (supplementary fig. S6, Supplementary Material online). Recently, the products of STC3 and STC5 were identified by heterologous expression of *F. fujikuroi STC3* and *F. mangiferae STC5* in *Escherichia coli* as eremophilene and guaia-6,10(14)-diene, respectively (Burkhardt et al. 2016). The failure to detect the product of *F. fujikuroi STC5* was likely due to a point mutation in the conserved NSE triad motif (supplementary fig. S7, Supplementary Material online). Alternation of the *F. fujikuroi STC5* NSE triad motif by site directed mutagenesis led to the production of guaia-6,10(14)-diene (Burkhardt et al. 2016). Pseudogenization also likely played a role in the failure to detect products of the other *F. fujikuroi* STCs. The predicted ORF of *F. fujikuroi STC8* has a nucleotide insertion causing a frameshift and premature stop codon. In addition, *STC8* may be nonfunctional because all predicted STC8 proteins are missing the third aspartate residue of the aspartate-rich motif (supplementary fig. S7, Supplementary Material online). Several predicted STC3 and STC7 proteins also exhibit remarkable sequence deviations and thus may be nonfunctional (data not shown).

Together, the FFC *Fusarium* genomes examined in this study have four DMATSs genes (supplementary table S1, Supplementary Material online). *F. mangiferae* is the only FFC with all four, both *F. proliferatum* strains and *F. verticillioides* have the same three, althouth *DMATS3* is likely nonfunctional in *F. proliferatum* 62905, whereas *F. fujikuroi* has only *DMATS1* and *DMATS3*. A fifth DMATS gene, *DMATS5* is unique to *F. circinatum* and has been previously described (Wiemann et al. 2013; Baer et al. 2014). So far, no product of any *Fusarium* DMATS has been identified.

Comparative Analysis of Phytohormone Biosynthetic Genes

The presence of GA and auxin (indole-3-acetic acid, IAA) biosynthetic gene clusters in most FFC members that have been examined represents a potentially unique trait of these fungi (Bömke and Tudzynski 2009; Wiemann et al. 2013). Complete and partial homologs of the GA clusters also occur in some *F. oxysporum* isolates, although no GA or GA-like compounds have been reported in this species (Wiemann et al. 2013). A GA gene cluster with homologs of six of the seven GA biosynthetic genes of *F. fujikuroi* was also found recently in three genome sequences of *F. avenaceum* strains. All three homologs lacked a *P450-3* homolog encoding the 13-hydroxylase (Lysøe et al. 2014). In *F. fujikuroi*, this enzyme catalyzes the last step of the GA biosynthetic pathway, the conversion of GA_7 to GA_3 (Bömke and Tudzynski 2009). In addition to the absence of *P450-3*, the organization of the *F. avenaceum* GA cluster differs from the *F. fujikuroi* cluster in the position of the desaturase gene (*DES*) (fig. 3).

In the current study, we found that *F. proliferatum* ET1 has two homologs of the GA cluster, referred to as GAC1 and GAC2, whereas *F. proliferatum* 62905 and other members of the Asian clade have only one homolog of the cluster. The gene content in the flanking regions of *F. proliferatum* ET1 GAC1 is almost identical to the corresponding regions in other Asian-clade species. In contrast, there are marked differences in content and organization of genes flanking the GAC2 cluster in *F. proliferatum* ET1 (fig. 3). First, GAC2 is inverted relative to the flanking regions: genes flanking the *DES* side of GAC2 flank the *P450-3* side of GAC1, andgenes flanking the *P450-3* side of GAC2 flank the *DES* side of GAC1. Second, the region flanking the *DES* side of GAC2 is inverted relative to GAC1. The content and arrangement of genes flanking *F. proliferatum* ET1 GAC2 is much more similar to genes flanking the GA cluster in the American-clade species *F. circinatum* (fig. 3).

Phylogenetic analysis of individual or concatenated nucleotide sequences of coding regions of GA biosynthetic genes revealed that the *F. proliferatum* ET1 GAC1 is most closely related to the GA cluster in other Asian-clade species (fig. 4*A*). In contrast, GAC2 genes are most closely related to GA cluster genes from the American-clade species *F. anthophilum* and *F. circinatum.* Genes flanking GAC1 and GAC2 exhibit the same types of phylogenetic relationships as the GA genes (supplementary fig. S8, Supplementary Material online) suggesting that the origin of GAC2 (as well as the regions flanking it) in *F. proliferatum* ET1 is a *Fusarim* species that is either in or closely related to the American clade.

Among several auxin biosynthetic pathways known in microorganisms, the IAM (indole-3-acetamide) pathway is considered to be the most significant in plant-associated bacteria (Tudzynski and Sharon 2002). The two genes that constitute the IAM pathway in plant-associated bacteria are *IAAM* (encoding tryptophan monooxygenase) and *IAAH* (encoding indole-3-acetamide hydrolase). Homologs of the bacterial *IAAM* and *IAAH* genes were previously found in four *Fusarium* species: *F. proliferatum* ET1, *F. verticillioides*, *F. fujikuroi*, and *F. oxysporum* (Tsavkelova et al. 2012). In each species, *IAAM* and *IAAH* are located next to each other in a head to head orientation separated by a short noncoding region (fig. 5). In a previous study (Tsavkelova et al. 2012), considerable amounts of the auxins IAM and IAA were observed in culture extracts of *F. proliferatum* ET1, whereas no or relatively small amounts of these indolic compounds were observed in culture extracts of the other *Fusarium* species examined. Deletion of the 2-gene locus in *F. proliferatum* ET1 resulted in marked reduction but not complete loss of IAA, suggesting the presence of an alternative pathway for IAA synthesis in this strain (Tsavkelova et al. 2012).

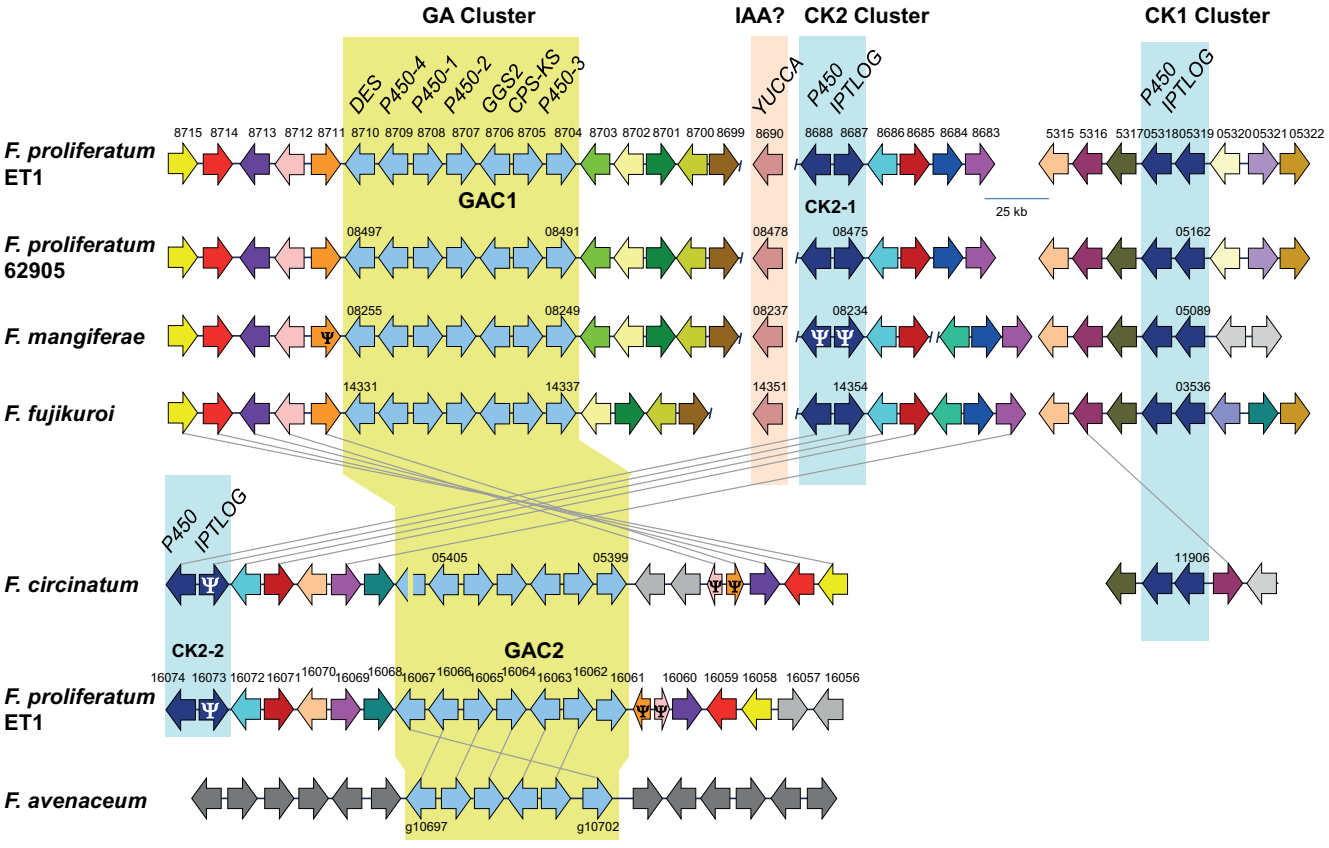

Fig. 3.—Content and arrangement of genes in the GA and CK cluster regions. Gene content of the GA and CK regions indicate that the clusters are in the same/similar genomic contexts in FFC members. However, *F. proliferatum* ET1 GAC2 and the GA cluster in the American-clade species *F. circinatum* are inverted relative to other GA cluster homologs examined. The genomic context of the GA cluster in the nonFFC species *F. avenaceum* differs from the context in the FFC. In different species, arrows that are the same color represent homologous genes. Exceptions to this are: the light blue arrows that represent the GA biosynthetic genes; the dark blue arrow that represent the two-gene cytokinin biosynthetic clusters CK1 and CK2. Grey arrows indicate genes that do not have a closely related homolog in the GA or CK cluster regions of the species examined. The Greek letter Ψ in an arrow indicates a pseudogene. Numbers above the genes correspond to gene model designations in the Pedant database (supplementary table S1, Supplementary Material online).

In the current study, we identified likely functional homologs of *IAAM* and *IAAH* in *F. proliferatum* 62905 as well as in *F. mangiferae*. Phylogenetic analysis of coding regions revealed that the IAA cluster genes from the Asian-clade species are more closely related, whereas the genes from *F. verticillioides* and *F. oxysporum* are more distantly related (fig. 4*B*).

Recently, Hinsch et al. (2015) described for the first time a functional cytokinin (CK) biosynthetic gene cluster in the ergot fungus *Claviceps purpurea*. The cluster consisted of two genes: *IPTLOG* and *P450*. *IPTLOG* encodes a bifunctional enzyme containing the isopentenyltransferase and lonely guy domains necessary for *de novo* isopentenyladenine production. *P450* encodes a cytochrome P450 monooxygenase. Comparative genomic analysis revealed that most members of the FFC have two copies of the CK cluster, designated CK1 and CK2 (figs. 3 and 4*C*). However, in *F. mangiferae* and *F. oxysporum*, one or both genes in one cluster are nonfunctional due to mutations that lead to premature termination of the predicted ORF.

F. proliferatum ET1 is unique in that it has three CK clusters, although one of the *IPTLOG* homologs is likely nonfunctional due to mutations that lead to premature termination of the ORF relative to other homologs. Phylogenetic analysis indicated that one of the *F. proliferatum* ET1 CK clusters is most closely related to CK1, whereas the other two clusters are more closely related to CK2. As a result, the latter two CK clusters in *F. proliferatum* ET1 have been designated CK2-1 and CK2-2. Interestingly, CK2-1 is most similar to CK2 from other Asian-clade *Fusarium*, whereas CK2-2 is most similar to the CK2 from the American-clade species *F. circinatum* (fig. 4*C*). In FFC genomes examined, the CK2 cluster is located on the same chromosome as the GA cluster (fig. 3). In *F. proliferatum* ET1, the CK2-1 and GAC1 clusters are 37 kb apart and most of the 15 genes in between them are

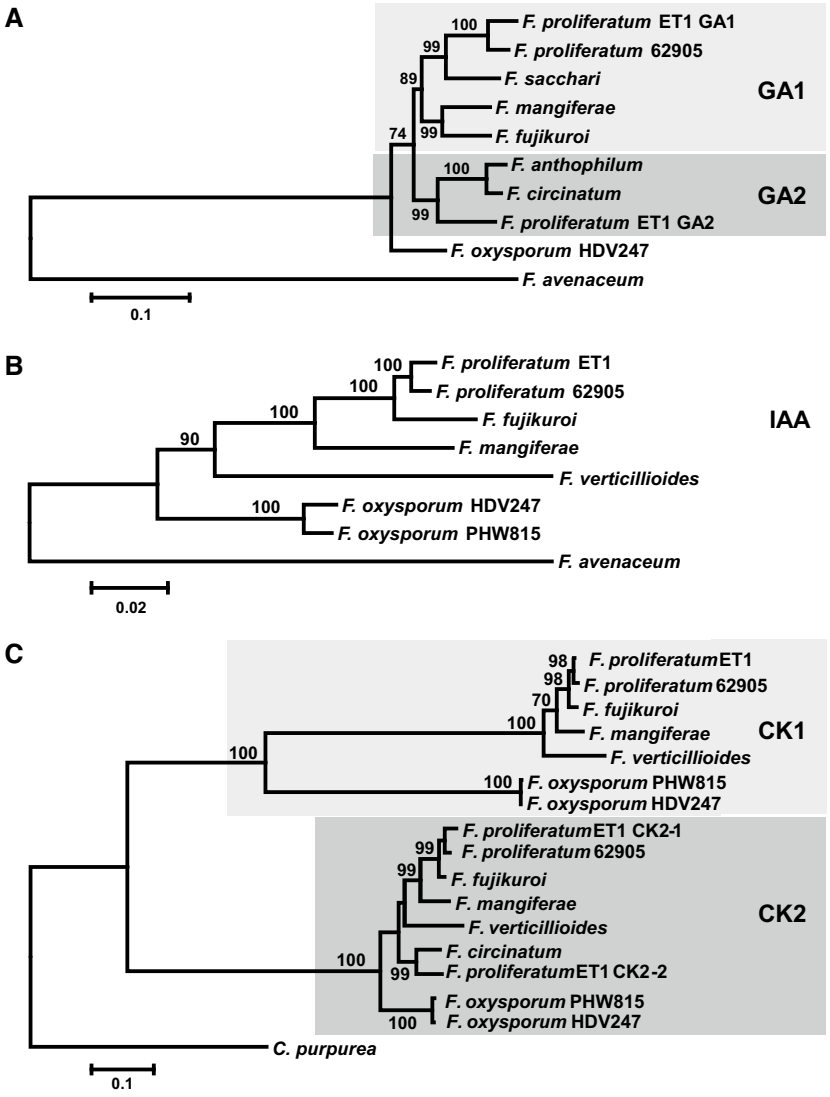

Fig. 4.—Maximum likelihood trees inferred from concatenated nucleotide sequences of coding regions of phytohormone biosynthetic genes. (*A*) GA biosynthetic genes *DES*, *P450-4*, *P450-1*, *P450-2*, *GGS2*, and *CPS-KS*. *P450-3* was excluded from the analysis because it is absent in *F. avenaceum*. (*B*) IAA biosynthetic genes *IAAM* and *IAAH*. (C) Cytokinin biosynthetic genes *IPTLOG-1* and *P450-1*(CK1 cluster), and *IPTLOG-2* and *P450-2* (CK2 cluster). Numbers near branches are bootstrap values based on 500 pseudoreplicates. Gene numbers are shown in supplementary table S1, Supplementary Material online.

homologs of the 15 genes located between the CK2 and GAC clusters in *F. fujikuroi* and *F. proliferatum* 62905. The second CK2 cluster (CK2-2) in *F. proliferatum* and GAC2 clusters are 13.6 kb apart, and the five genes located between them are homologs of the five genes located between the CK2 and GA clusters in *F. circinatum*. However, the relative position of the CK2-2 and GAC2 clusters is reversed compared with the CK2-1 and GAC1 clusters. Based on phylogenetic analysis of genes in the GA cluster and CK2 region as well as the organization of genes flanking the two clusters, we hypothesize that the presence of GAC2 and CK2-2 in *F. proliferatum* ET1 might be the result of the same HGT event (figs. 3 and 4C).

The absence of GAC2 in *F. proliferatum* 62905 suggests that the putative HGT event of GAC2 occurred relatively recently, after *F. proliferatum* ET1 and 62905 diverged from their common ancestor. Alternatively, the transfer could have occurred prior to divergence of these two strains if it was followed by loss of the GAC2-CK2-2 region from the *F. proliferatum* lineage leading to 62905. To test these alternatives, we used PCR to assess the presence of GAC2 in 52 isolates of *F. proliferatum* that were obtained from different plant hosts and geographical regions (Proctor et al. 2009). GAC2 was not detected in any other isolate of this species. This supports the hypothesis that HGT of the GAC2-CK2-2 region occurred recently.

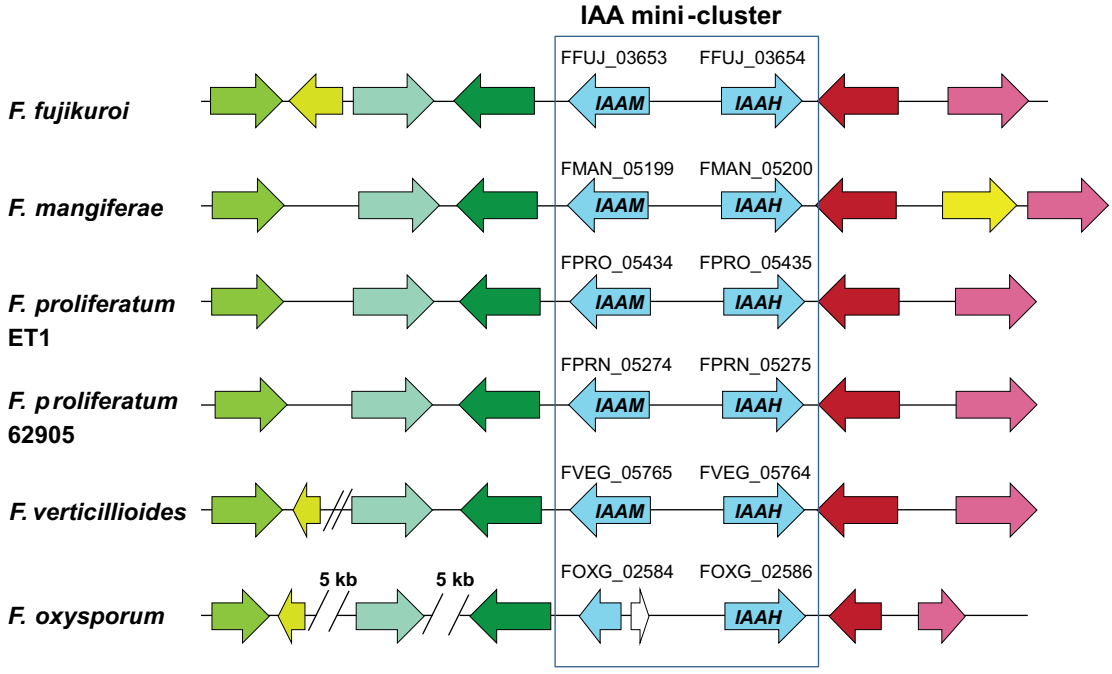

Fig. 5.—Arrangement of genes in the IAA cluster region. The two-gene IAA cluster: *IAAM* (encoding tryptophan monooxygenase; FFUJ_03653) and *IAAH* (encoding indole-3-acetamide hydrolase; FFUJ_03654). Homologous genes are indicated by identical color.

Does the Presence of SM Gene Clusters Correlate with Product Formation?

The production of known PKS-, NRPS-, and STC-derived products was investigated in the five FFC strains and in *F. oxysporum* 4287 under low and high nitrogen conditions as previously described for *F. fujikuroi* (Wiemann et al. 2013). Despite the presence of most of the gene clusters in most of the strains, there were remarkable differences in SM production (table 3). For example, although the PKS3 gene cluster for the perithecia pigment fusarubin (Studt et al. 2012) was present in all *Fusarium* strains analyzed, fusarubin was not detected in culture extracts of *F. proliferatum* 62905 and *F. oxysporum*.

Second, although *STC4* and *STC6* are present in all genomes analyzed, production of (–)-α-acorenol (STC6) was only detected in headspace extracts from *F. fujikuroi, F. proliferatum and F. verticillioides*, and (+)-koraiol (STC4) was only produced by *F. fujikuroi* and *F. oxysporum* (table 3; supplementary fig. S6, Supplementary Material online).

Third, although the GA cluster is present in *F. fujikuroi*, both *F. proliferatum* ET1 and 62905 as well as *F. mangiferae*, *ent*-kaurene (supplementary fig. S6, Supplementary Material online) (Brock et al. 2011) was only produced by *F. fujikuroi* and *F. proliferatum* ET1 whereas large amounts of GAs were produced by *F. fujikuroi*, and only small amounts were observed in *F. proliferatum* ET1 (fig. 6*A*).

In a fourth example, the high levels of IAA and IAM produced by the orchid isolate *F. proliferatum* ET1 correlate with the presence of functional pathway genes *IAAM* and *IAAH* (Tsavkelova et al. 2012) as it has been previously shown for *Colletotrichum* sp. (Maor et al. 2004). Non-functional homologs were found in other *Fusarium* (Tsavkelova et al. 2012) as well as in *F. proliferatum* 62905, although low levels of IAA, but almost no IAM, were detected in the latter strain (supplementary fig. S9, Supplementary Material online).

In minimal media with no tryptophan, relatively high levels of IAA were detected in both the mycelium and filtrates of all WT strains, but only *F. proliferatum* ET1 and to a smaller extent *F. verticillioides*, produce substantial amounts of IAM, the intermediate of the IAM pathway, in mycelium (fig. 6*B*; supplementary fig. S9, Supplementary Material online). The enormous level of IAA exclusively produced by *F. proliferatum* ET1 in media supplemented with tryptophan (supplementary fig. S10, Supplementary Material online) indicates that the IAM pathway is inducible by tryptophan as shown before (Tsavkelova et al. 2012), and that this pathway seems not to be active in any other WT strain.

Overexpression of the *F. proliferatum* ET1 *IAM* genes, in *F. proliferatum* ET1 (homologous) or in FFC members and *F. oxysporum* (heterologous), increased IAM and IAA production in all mutants compared with WT progenitor strains (fig. 6*B*; supplementary figs. S9 and S10, Supplementary Material online).

The contribution of other metabolic pathway(s) to IAA production is supported by the relative high auxin levels in *F. mangiferae*, and *F. fujikuroi* and *F. verticillioides* WT strains

Table 3

Presence of Secondary Metabolite Gene Clusters and Production of the Concomitant Product under Standard Laboratory Conditions

	Asian-clade of the FFC				African-clade	Outgroup
Metabolite	***F .proliferatum* 62905**	***F. proliferatum* ET1**	***F. mangiferae***	***F. fujikuroi***	***F. verticillioides***	***F. oxysporum***
Gibberellins*[a]	–	+	–	+	No cluster	No cluster
Bikaverin*[a]	+	+	+	+	+	+
Fusarubins*[c]	–	+	+	+	+	–
Fusarins*[b]	No cluster	+	No cluster	+	+	No cluster
Fusaric acid*[b]	+	+	+	+	+	+
Apicidin F*[b]	No cluster	No cluster	No cluster	+	No cluster	No cluster
ent-Kaurene**	–	+	–	+	No cluster	No cluster
α-Acorenol**	+	+	–	+	+	–
(+)-Koraiol**	–	–	–	+	–	+

NOTE.—The strains were grown for 7 days in ICI media supplemented with different nitrogen sources under inducing conditions for the respective secondary metabolite. Analyses were done by

*HPLC-DAD and

**GS-MS.

[a]6 mM glutamine;

[b]60 mM glutamine;

[c]6 mM $NaNO_3$.

(fig. 6*B*), and by the detection of precursors, other than IAM, that are associated with alternative IAA biosynthetic pathways (supplementary fig. S9, Supplementary Material online). The presence of indole-3-pyruvic acid (IPyA) in mycelia and culture filtrates of *F. mangiferae* suggests an active plant-like IAA pathway that could include a tryptophan aminotransferase and a flavin monooxygenase-like protein (e.g., YUCCA enzyme) (Kasahara 2016). In support of this possibility, a putative YUCCA-like gene is located in the ~37 kb region between the CK2-1 and GAC1 clusters in *F. proliferatum* isolates ET1 and 62905, *F. mangiferae* and *F. fujikuroi* (fig. 3).

Another alternative IAA biosynthetic pathway, independent of YUCCA and *IAAM* and *IAAH*, includes the intermediate indole-3-acetaldehyde (IAAld) and was described recently by Mashiguchi et al. (2011). IAAld was the most abundant IAA precursor detected in filtrates of all *Fusarium* except *F. mangiferae* and in mycelium of three out of six *Fusarium* strains (supplementary fig. S9, Supplementary Material online). This pathway utilizes a putative tryptophan decarboxylase, aminoxidases or indole-3-acetaldehyde dehydrogenase. Putative orthologs for the encoding genes have been identified in *Fusarium* genome annotation sets.

In a fifth example, all six *Fusarium* isolates produce four types of isoprenoid CKs under laboratory conditions: isopentenyladenine (iP), *trans*-zeatin (tZ), *cis*-zeatin (cZ), and dihydrozeatin (DHZ) although substantial differences in the production yields were observed among *Fusarium* species as well as among biological replicates of one species (fig. 6*C*).

Two independent pathways for CKs were recently proposed for the ascomycete *C. purpurea* (Hinsch et al. 2015). In the *de novo* CK pathway, the product of *IPTLOG* catalyzes formation of iP which is then hydroxylated by the product of *P450* yielding tZ. In the second pathway, the tRNA degradation pathway, the product of the tRNA-isopentenyltransferase (*tRNA-IPT*) catalyzes the formation of an intermediate that is then converted by the cytokinin-activating enzyme called "lonely guy" (*LOG*) yielding either iP or cZ. When the deletion of the *de novo* pathway did not lead to a complete loss of CKs, a double deletion mutant of *tRNA-IPT* and *IPTLOG* in *C. purpurea* was generated (Hinsch et al. 2016). The marked reduction in pathogenicity in the "CK-free" double deletion mutant clearly established for the first time the importance of CKs in host–fungus interaction (Hinsch et al. 2016). The importance of CKs for virulence has recently been demonstrated for the blast fungus *Magnaporthe oryzae* (Chanclud et al. 2016). Mutants lacking *tRNA-IPT* exhibited reduced virulence on rice. The absence of the *IPTLOG* and *P450* genes in the *M. oryzae* genome suggests that the tRNA degradation pathway is the only source of CKs in this fungus.

After *Claviceps*, *Fusarium* is the second ascomycete in which two independent CK biosynthetic pathways have been identified. qPCR analysis revealed that that *IPTLOG* and *P450* expression of both CK clusters were strongly repressed in axenic culture in all fusaria examined (data not shown), suggesting that the observed levels of CK production can be attributed to the tRNA degradation pathway. Accordingly, iP- or cZ-derivatives are the predominate CKs detected in the majority of analyzed samples (supplementary fig. S11, Supplementary Material online). Interestingly, culture filtrates of *F. mangifereae* and *F. proliferatum* 62905 contain higher levels of iP- and cZ-derivatives than other fusaria suggesting a highly active tRNA degradation pathway in these fungi.

To examine the functionality of the two *de novo* CK gene clusters in *Fusarium*, *IPTLOG* and *P450* from both CK1 and CK2 clusters of *F. fujikuroi* were constitutively overexpressed in

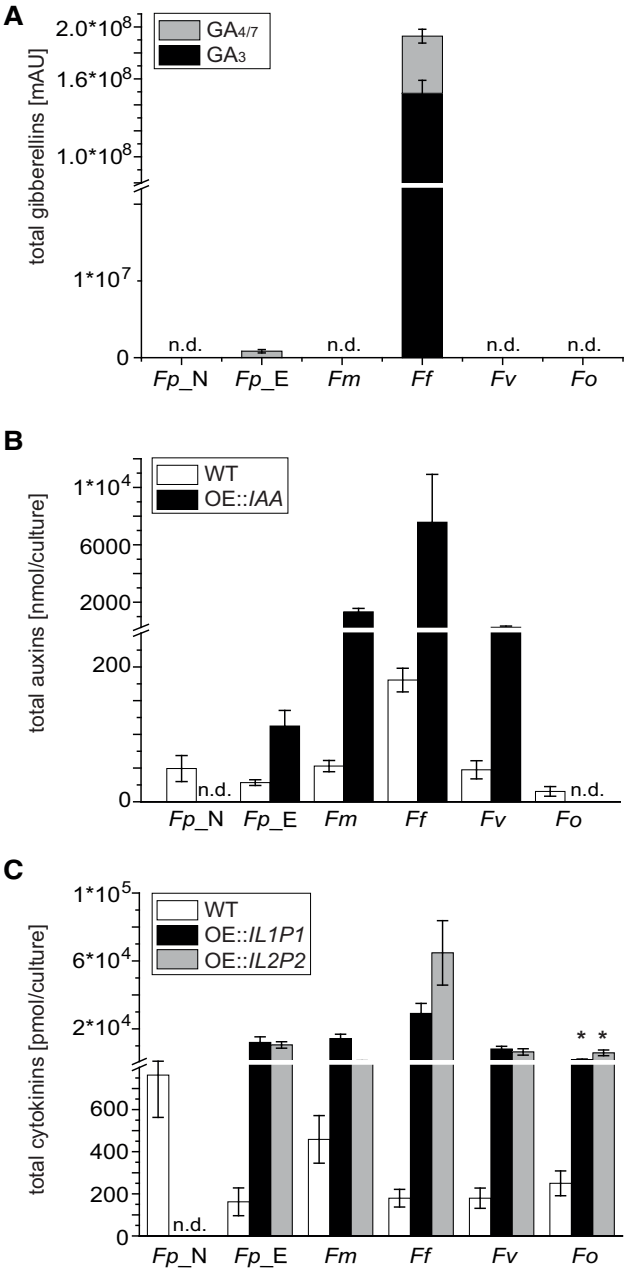

Fig. 6.—*In vitro* production of phytohormones by the six analyzed *Fusarium* strains. (*A*) Total concentration of GAs in the WT strains analyzed via HPLC-DAD (λ = 210 nm). (*B*) Total concentration of auxins in the WT and overexpression mutant strains (OE::*IAA*). (*C*) Total concentration of cytokinins in the WT and the double overexpression mutant strains OE::*IL1P1* and OE::*IL2P2*; * in *F. oxysporum* only *IL1* and *IL2* genes, respectively, were overexpressed. The data represent the average of three independent biological repetitions, the bars show standard deviations. n.d. - not detected.

five different *Fusarium* species. Analyses of extracts of transformants after growth in axenic culture revealed the presence of high levels of multiple CK analogs and indicate that all four genes encode functional enzymes (fig. 6*C*; supplementary fig. S11, Supplementary Material online). Whereas overexpression of only *IPTLOG* (*IL1* or *IL2*) led to higher accumulation of iP than tZ in most cases, simultaneous overexpression of both *IPTLOG* and *P450* (*IL1P1* or *IL2P2*) resulted almost exclusively in accumulation of tZ. These data suggest that the putative monooxygenases encoded by the *P450* genes in both CK clusters likely add a hydroxyl group to iP to form tZ, as shown for *C. purpurea* (Hinsch et al. 2015).

Summarizing, all studied FFC isolates have the genetic capacity to produce GAs (except for *F. verticillioides*), auxins and CKs. However, pseudogenization likely prevent their synthesis in some of the fungi. In the culture conditions examined, GAs are only produced by *F. fujikuroi* and to a lower extent by *F. proliferatum* ET1 whereas all isolates produce IAA. However, IAM, the intermediate of the IAM pathway, was only observed in *F. proliferatum* ET1 and to a lower extent in *F. verticillioides*. The high levels of IAA in cultures of the other fusaria indicate that alternative IAA biosynthetic pathways are present. In contrast, although all *Fusarium* have at least one intact CK cluster, minimal gene expression is likely responsible for the very low levels of tZ, whereas the rest of produced CKs originate from evolutionary highly conserved tRNA dependent pathway. This pathway is activated upon axenic growth mainly in *F. mangiferae* and *F. proliferatum* strain 62905.

Infection of Maize

The observation that more than half of the SM gene clusters are transcriptionally silent under the growth conditions tested has stymied efforts to characterize their products and determine their biological functions. Expression profiling *in planta* could lead to the chemical identification of these orphan SM clusters as well as provide insight into a potential role some SMs play during the fungus–plant interaction.

With the exception of *F. mangiferae*, which causes fruit malformation in mango, the other *Fusarium* species have been previously shown to invade the roots, crown, or flowers of cereals, including barley, maize, rice, and wheat (Maor et al. 2004; Desjardins 2006; Leslie and Summerell 2006; Bajguz and Piotrowska 2009). Before analyzing the *in planta* expression of SM gene clusters, we first studied the capability of the five FFC strains to penetrate and invade maize roots and to cause symptoms.

In the maize seedling assay, we found that the mean germination was often lower, for all species, except *F. fujikuroi* (supplementary table S2, Supplementary Material online). We also found that mean seedling height was also often lower, for all species, except *F. fujikuroi* which caused strong elongation of maize seedlings compared with the control (supplementary fig. S12 and table S2, Supplementary Material online). However, although mean germination and seedling height were often lower for all species, except *F. fujikuroi*, than for the uninoculated control, none of the differences

were statistically significant (supplementary table S3, Supplementary Material online). In contrast, *F. fujikuroi* caused significant elongation of maize (supplementary fig. S12, Supplementary Material online). This elongation was reminiscent of *bakanae* disease; i.e., elongation of rice seedlings resulting from infection with *F. fujikuroi* and associated with GA production (Wiemann et al. 2013).

The increased height of maize seedlings induced by *F. fujikuroi* and reductions, albeit inconsistent reductions, in seed germination and seedling height induced by the other fusaria suggest that all the strains were capable of infecting the seeds/seedlings (supplementary fig. S12 and table S2, Supplementary Material online). Therefore, in addition to the maize seedling assay, we tracked growth of the fungi inside the seedling tissue using GFP-tagged strains and confocal microscopy. Despite their variation in ability to induce disease symptoms and their different host origins, all six *Fusarium* strains examined were consistently able to infect and grow within roots (supplementary fig. S13, Supplementary Material online). Both roots and crown tissue contained considerable quantities of fungal biomass, mainly mycelia with occasional pockets of conidia. There were differences in the amounts of hyphae between 10 and 20 dpi, and between roots and crowns. There were also differences in growth of the two fusaria, *F. verticillioides* and *F. proliferatum* 62905, that originated from maize compared with the other fusaria. *F. proliferatum* 62905 produced dense and congested mycelia in roots and crown at 10 dpi. At 20 dpi, the amount of biomass in the crowns of *F. verticillioides*-infected plants was greater compared with 10 dpi, and reached similar levels to those observed for *F. proliferatum* 62905, which was more abundant and more uniformly distributed in the crown tissue compared with all the other species. These results are in agreement with previous reports showing that the crown is the preferred tissue for early stages of maize infection by *F. verticillioides* (Oren et al. 2003). Interestingly and in contrast to the maize isolate (62905) of *F. proliferatum*, biomass of the orchid isolate (ET1) was abundant in roots but sparce in the crown.

In summary, the maize seedling assay and the microscopic analysis revealed that all species, including the mango pathogen *F. mangiferae*, were capable of entering the roots. However, the strains originating from maize proliferated faster and produced more biomass in the roots as well as in the crown (and possibly in the upper parts of seedlings).

In Planta Expression of SM Gene Clusters

In the present study, SM production in liquid media differed among the *Fusarium* strains examined even when the corresponding gene clusters was shared by all the strains (table 3). Previous genome-wide expression studies in *F. fujikuroi* revealed that many SM gene clusters are not expressed in culture (Wiemann et al. 2013). To gain insight into how expression of gene clusters contributes to variation in SM production among the *Fusarium* strains, and whether there are different expression pattern between *in vitro* and *in planta* conditions, we used RNA-seq to compare expression of SM cluster genes during growth on CM agar (*in vitro*) and in maize roots at 4, 10, and 20 dpi (*in planta*) (fig. 7; supplementary table S4, Supplementary Material online).

The analysis revealed differences in SM cluster expression: (1) for the same strain growing in CM culture versus maize roots; (2) for the same strain at different time points in maize roots; and (3) for different strains growing under the same condition. In contrast, some SM clusters exhibited similar levels of expression in all strains and/or all conditions (fig. 7; supplementary table S4, Supplementary Material online). Examples of the latter clusters are those involved in synthesis of the siderophores ferricrocin (core gene *NRPS2*) and fusarinines (core gene *NRPS6*), and those involved in synthesis of gibepyrone and neurosporaxanthin, as well as several clusters for which the corresponding SM products are not known: e.g., clusters with the core genes *DMATS1*, *NRPS3*, *NRPS10*, *NRPS11*, *NRPS13*, *TrTC1* (fig. 7).

In addition, although the fujikurin (*PKS19*) cluster was present in only three of the five fusaria, they all expressed the cluster under all conditions examined. Genes responsible for synthesis of the two volatile compounds (+)-koraiol (*STC4*) and (−)-α-acorenol (*STC6*) are highly expressed in all species, either in roots only or in both roots and culture (fig. 7; supplementary table S4, Supplementary Material online). In maize roots, the high levels of expression of terpene cyclase genes required for synthesis of these volatiles indicate that these compounds may have a role in plant–fungal interactions, as has recently shown for other volatile terpenes in the mutualistic association of plant roots with ectomycorrhizal fungi (Ditengou et al. 2015). In general, expression of SM gene clusters in culture as well as in maize roots suggests that the corresponding SMs play a role supporting fungal growth and/or fungus–plant interaction that is independent of host specificity.

The majority of SM clusters exhibited different patterns of expression in culture versus in maize roots and/or among the different strains. Some clusters were highly expressed in only one or two species and not or only minimally expressed in the others. For example, the fusaric acid cluster genes *FUB1* (*PKS6*) and *FUB8* (*NRPS34*) were highly and coordinately expressed in *F. mangiferae*, both on CM and at early stages in maize roots. In *F. verticillioides* and in *F. fujikuroi*, these genes were specifically and strongly up-regulated in maize roots at 4 and 10 dpi. Surprisingly no expression of fusaric acid genes was observed by *F. proliferatum* 62905 (supplementary table S4, Supplementary Material online). Another example is the fumonisin cluster, which includes *FUM1* (*PKS11*). This gene exhibited high levels of expression in *F. verticillioides* and *F. proliferatum* ET1 on CM and in maize roots, but it exhibited low-level expression in *F. fujikuroi* and surprisingly no expression in *F. proliferatum* 62905. The fumonisin cluster is not present in the *F. mangiferae* genome (fig. 7). Low-level

Fp _N				*Fp* _E				*Fm*				*Ff*				*Fv*				Key-enzyme	Product
CM	4	10	20	CM	4	10	20	CM	4	10	20	CM	4	10	20	CM	4	10	20		
4.7	4.96	2	4.2	2.5	3.6	4.3	3.9	2.2	1.6	3.1	2.1	3.5	3.8	4.7	4.4	3.3	2.4	3.5	2.6	DMATS1	
0.7	3.28	0.2	0.6	4.8	3.3	2.1	5	4.5	1.2	1.1	0	3.6	3.4	0.8	2.5	4.8	4.1	4.8	6.2	DMATS3	
2.6	0	1.2	1.9	3	2.8	1.6	3	3	1.6	2.4	1.7	n.d.	n.d.	n.d.	n.d.	4.8	3	4.1	4.8	DMATS4	
0	0	0.1	0.5	0	2.2	4.1	2.1	0	0	0.3	0	7.9	8.5	11	9.1	n.d.	n.d.	n.d.	n.d.	DTC1-1	Gibberellins
n.d.	n.d.	n.d.	n.d.	1.7	1.5	5.7	1.3	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	DTC1-2	Gibberellins
0	0.65	0	0	3.7	5.8	6.8	5.9	0.1	0.5	0	0	0	0	0	0	0.8	6.3	5.4	5.3	iaaM	Auxins (IAA)
4	1.05	0.2	2.1	4.4	2.9	3.8	4.3	2.4	2.4	2	1.6	0.5	0.5	0	0.4	0.2	0	0	0	YUCCA	Auxins
0	0	0	0	0.4	0	0	0	0	0	0	0	0	0	0	0.5	0	0	0	0	IPT_LOG1	Cytokinins
0	0	0	0	0	0	0.9	0	0	0	0	0	2.1	2.2	0.4	3.2	1.5	2.2	4.1	0	IPT_LOG2	Cytokinins
3.7	3.73	3.6	2.4	3.3	4.4	3.8	3.9	3.7	2.9	4.3	4.8	4.2	3.9	4.1	3.6	3.1	3	3.3	3	tRNA-IPT	Cytokinins
n.d.	n.d.	n.d.	n.d.	0	1.7	1.4	2	0	0	0.3	0.4	n.d.	n.d.	n.d.	n.d.	0.2	1	0.4	0	NRPS1	Malonichrom
5.3	3.03	3.8	3.4	5.1	4.8	6	5.3	5	4.2	4.9	4.8	4.5	4	6.2	5.7	4.9	4.4	4.1	4.4	NRPS2	Ferricrocin
2.7	1.74	1.3	1.1	3.2	3.2	2.1	2.7	2.8	2.9	2.2	2.1	3.5	3.1	2.3	2.3	2.8	2.9	2.2	1.3	NRPS3	
0.4	1.17	0	0.4	4.2	4.2	2.9	3.8	4.9	2.3	1.2	1.1	0.8	0.6	0.2	0.3	5.2	3.7	2.9	3.9	NRPS4	
1.6	3.14	1.1	1.4	3.5	5.5	3.8	4.1	4.3	4	2.7	3	3.6	3.5	2.4	3.5	2.7	3.3	3.7	2.8	NRPS6	Fusarinine
3.9	3.74	5.6	2.6	5.2	5.1	3.9	3.5	4.9	3.9	3.2	4.5	5.5	5.4	3.6	4.7	4.5	3.4	2.7	3.8	NRPS10	
3	4.62	0.9	3.1	1.8	6.2	5.6	7.2	3.2	6.9	8	7	4.7	4.7	3.1	3	5	7.4	6	6.2	NRPS11	
0	3.43	0.2	1.6	0.4	6.5	5	7.2	0.7	5.9	5.1	7.1	0.1	0	0.8	1.3	0.2	6.1	5.4	3.7	NRPS12	
2.5	1.44	2.5	0.9	3.1	3.2	3.2	3.2	2.7	2.6	2.9	2.7	1.5	1.3	2.4	2.4	4.2	3.9	3.7	2.7	NRPS13	
0	0	0	0.3	0.1	0	0	0	n.d.	n.d.	n.d.	n.d.	0	0	0	0	2.2	3.6	1.5	1.9	NRPS17	Ferrichrom
0	0	0	0	0	0	0	0	0.1	0	0.1	0.5	0.4	0.3	0	0.2	0.1	0	0	0	NRPS20	
0	0.17	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0.1	0	0	0.5	NRPS21	
0	0.49	0	0	5.9	4.2	2.3	2.8	7.6	1.6	0.3	2.6	0	0	0	0	0	0.6	1.2	0.1	NRPS22	Beauvericin
0.1	2.88	0.3	0.4	2.1	4	4.2	5	0.1	0	0.5	0	0.4	0.3	0	0	0.3	5.1	4.7	0.5	NRPS23	
n.d.	n.d.	n.d.	n.d.	3.4	4.6	4.2	3.5	0.3	1.4	0.7	0	n.d.	n.d.	n.d.	n.d.	3.8	5.4	3.7	1.9	NRPS24	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.3	0	1.1	0	NRPS25	
n.d.	n.d.	n.d.	n.d.	2	2.9	2.6	2.1	0.1	1.2	0.4	0.7	n.d.	n.d.	n.d.	n.d.	2.4	4.6	2.6	0.8	NRPS26	
n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	0	0.7	0.4	1.6	n.d.	n.d.	n.d.	n.d.	NRPS31	Apicidin F
0	0	0	0	3.8	1.4	0.6	0.7	7.2	6.9	2.1	0	0	2.6	0.4	0	1.1	3.7	4.8	0.9	NRPS34	Fusaric Acid
0	0	0.1	0	0	0	0	0	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	NRPS36	Ferrirhodin-type
0.1	0.31	0.3	0	0.2	0	0.5	0.6	1.1	0.3	0.5	0.3	0	0	0	0	0.3	0.5	0.1	0.7	PKS1	Equisetin
0	1.54	0.1	0	3	1.7	2.7	1.9	0	0.3	0.2	0	0.7	0.4	0	0.6	3.5	2.4	2.9	0.8	PKS2	
0.5	0.51	0.1	0.3	1.2	1	1.1	0.4	0.6	0.4	0.9	0	0.2	0.1	0.2	0.5	0	0.6	1.1	2	PKS3	Fusarubins
0.4	3.31	0.2	0.8	0.5	0	0	0	3.5	4.7	2.4	0.7	6.3	6.1	4.6	5.3	6.8	6	1.3	1.6	PKS4	Bikaverin
0.4	2.5	0.7	1.6	0.7	1.2	1	2.6	0.2	3.2	3	2.5	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	PKS5	
0.1	0	0.1	0.2	3.1	1.7	1.1	1.2	7.8	6.9	1.7	0.4	0.1	2.9	0.4	0.1	1.2	3.9	4.4	0.5	PKS6	Fusaric Acid
0.2	0	0	0.1	1.4	1.2	0.3	0.1	0.7	0.7	1.2	0.4	0.3	0.2	0	0.3	1.3	0.9	1.1	0.4	PKS7	
1.3	2.36	0.4	1.3	4.7	2.6	2.5	2.8	3.7	2	2	1.8	4.3	3.8	4	2.8	0.9	1.1	0.8	0.3	PKS9	
1.6	0.65	0.2	0.7	1.9	0.3	1.8	0.5	0.1	0	0.2	0	2.5	2.1	3.4	5.8	2.5	0.9	1.8	0.3	PKS10	Fusarin C
0.3	0	0	0.6	8.5	6.5	4.9	2.9	n.d.	n.d.	n.d.	n.d.	2	2	0.1	0.3	8.9	6.9	2.1	6.4	PKS11	Fumonisin
5.9	5.22	2.3	5.1	2.5	2.7	2.4	4	2.7	5.8	4.9	5.7	4.5	4.2	4.2	3.8	3.6	2.6	3.8	1.4	PKS13	Gibepyron
0	0.62	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.7	0	PKS14	
n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	4.6	2.5	1	0.6	PKS15	
0.6	1.15	0.2	0	2.3	0.7	0	0.3	n.d.	n.d.	n.d.	n.d.	0	0.1	0.1	0.1	n.d.	n.d.	n.d.	n.d.	PKS16	
0	0.17	0	0	0	0	1.8	0.4	0	0	0	0	0	0	0	0.1	n.d.	n.d.	n.d.	n.d.	PKS17	
0	0	0	0	0	0.8	2.2	0.9	0	0	0	0	0	0	0	0	n.d.	n.d.	n.d.	n.d.	PKS18	
5.9	2.21	0.6	1.4	2.2	2.2	2.3	2.5	n.d.	n.d.	n.d.	n.d.	2.1	1.9	2.6	2.5	n.d.	n.d.	n.d.	n.d.	PKS19	Fujikurin
0.1	0.49	0	0.2	5	2.6	2	2.5	0.9	3	1.8	1.6	0.7	0.4	0.2	0.5	0.4	0	0	0	PKS20	
0.3	1.34	0.1	0.9	6.8	3.9	1.4	5.1	1	0.7	0.5	0.9	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	PKS38	
n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	0	0	0.2	0.2	PKS39	Depudecin
0.2	1.51	0.3	0.5	0	0	0.1	0	0.1	0	0.5	0	n.d.	n.d.	n.d.	n.d.	0	0.3	0.9	0.1	PKS40	
n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	0	0	0	0	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	PKS43	
1.6	0	0.2	1.8	0.9	0	0.5	0.3	0.3	1.3	3.6	3.3	2.6	2.7	0	1.6	1	4	4.5	4.6	STC1	
2	1.05	1.1	3.4	1.6	1.9	2.4	0.3	2.1	2.6	2.7	0	2.3	3.2	3.4	2.6	1.8	0	1.5	3	STC2	
0.1	0	0	0	0.3	0	0	0	1.1	0.9	0.4	0	0	0	0	0	0	0	0	0	STC3	Eremophilene
0.1	6.13	3.1	4.1	4.8	2.6	5.8	5.2	0.6	3.9	3.1	3.8	7.4	7.4	4.5	6.1	6.4	3.5	4.8	2.9	STC4	(+)-Koraiol
0	0	0	0	2.3	0	0	3.1	0.2	0	0.7	0	0.3	0.6	0.5	0	n.d.	n.d.	n.d.	n.d.	STC5	Guaia-6,10(14)-diene
1	4.14	1	1.9	1	3.8	2.8	2.4	0.7	4.8	4.2	4.3	3.3	3.8	4.3	5	0.7	3.5	4.2	2.8	STC6	(-)-α-Acorenol
n.d.	n.d.	n.d.	n.d.	0.8	0	0	0	0	0	0	0	0.1	0.4	0	0	1.6	2.1	0.9	0	STC7	
n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	5.2	2.2	0.4	0	0.2	0	0	0	4.5	0	1	3.1	STC8	
1	3.03	0.4	1.5	1.6	0	0.9	0	2.1	2.1	1.7	1.2	3.7	3.5	0.5	1.4	1.1	3	1.9	1.3	STC9	
3.1	3.58	5.4	3.5	4.9	5.9	4.9	5.2	5	4.9	4	4.6	6	5.7	5.2	4.9	4.5	4.8	4.9	3.8	TrTC1	
3	0.68	3.4	3.4	5.4	4.9	2	2.4	8.4	1.9	1.7	2	0.5	2.9	4.1	2.9	7.4	3.9	2.3	6	TeTC1	Carotenoids

Fig. 7.—Expression of SM cluster genes encoding core enzymes during growth of five strains of *Fusarium* in culture and in maize roots. Expression levels are represented as a heat map scaled from 0 to 11. 0 is represented by a white background and indicates the lowest level of expression, and 11 is represented by a deep red background and indicates the highest level of expression. Treatments: *Fp*_N and *Fp*_E indicate the *F. proliferatum* strains 62905 and ET1 respectively; *Fm*, *Ff*, and *Fv* indicate *F. mangiferae*, *F. fujikuroi*, and *F. verticillioides* respectively; CM indicates gene expression in 5-day-old CM agar cultures; and 4, 10, and 20 indicate gene expression in maize tissue at 4, 10, and 20 days after inoculation. See Methods for details on RNA isolation, sequencing, mapping, and quantification.

expression of the fumonisin genes in *F. fujikuroi* compared with *F. verticillioides* was also demonstrated recently in liquid cultures (Rösler et al. 2016). A third example is the beauvericin biosynthetic gene *NRPS22*, which was highly expressed in *F. mangiferae* on CM and in *F. proliferatum* ET1 on CM and in maize roots. The bikaverin gene cluster, which include *PKS4/BIK4*, exhibited no or minimal expression in *F. proliferatum* ET1 and 62905, but exhibited high-level expression in *F. mangiferae*, *F. fujikuroi*, and *F. verticillioides* under some of the conditions. The fusarin gene cluster, which includes *PKS10/FUS1*, exhibits relatively high levels of expression only in *F. fujikuroi* with a maximum expression on day 20 in maize (fig. 7; supplementary table S3, Supplementary Material online). *STC5*, which is responsible for production of guaia-6,10(14)-diene (Burkhardt et al. 2016), was expressed only in *F. proliferatum* ET1 on CM and at 20 dpi.

The expression of several gene clusters with unknown SM products was also specifically up-regulated in maize roots in one or two strains. For example, although the putative *PKS17* and *PKS18* clusters are present in all Asian-clade strains, they were expressed, albeit at low levels, only in *F. proliferatum* ET1 in maize roots at 10 dpi. The *NRPS23* and *NRPS24* clusters were expressed only by *F. proliferatum* ET1 and *F. verticillioides*, whereas the *PKS20* and *PKS38* genes were only expressed in *F. proliferatum* ET1 (supplementary table S4, Supplementary Material online).

The phytohomormone (GA, IAA, and CK) biosynthetic genes also exhibited variation in gene expression. The highest expression levels for any SM biosynthetic gene cluster were those exhibited by the GA biosynthetic genes in *F. fujikuroi*, both on CM and *in planta* (fig. 7). This indicates that in this species GA production is not restricted to its predominant host, rice. These data correspond well with the *bakanae*-like symptoms observed in the maize seedling assay (supplementary fig. S12, Supplementary Material online). In *F. proliferatum* ET1, both GA gene clusters were also induced in maize roots. However, based on the symptoms in the maize seedling assays, the amount of GAs produced were insufficient to induce elongation. *F. proliferatum* ET1, and to a lesser extent *F. verticillioides*, were the only strains with an active IAM auxin biosynthetic pathway (supplementary fig. S9, Supplementary Material online). The IAA production levels are consistent with the high-level expression of *IAA* genes in these two strains, especially in maize roots (fig. 7). From the two CK gene clusters present in all studied isolates, only the CK2 cluster was expressed, and only in *F. fujikuroi* and *F. verticillioides* (supplementary table S4, Supplementary Material online).

In summary, although the different species share many of the same SM gene clusters, their expression varies markedly among strains as well as within strains under different conditions. The expression of some SM genes only in maize suggests that they may play a role in plant–fungus interactions. Most striking are the expression pattern for the endophyte *F. proliferatum* ET1: (1) it is the only strain with an highly active IAM biosynthetic pathway, (2) it is the only strain with two GA clusters, both of which are expressed in maize, and (3) it has an unexpectedly high portion of highly expressed SM gene clusters for known SMs (e.g., fumonisin, beauvericin, and fusarin) as well as gene clusters for which the SM products are not known (e.g., *NRPS23*, *NRPS24*, and *PKS38*). The markedly different expression patterns of these genes in the other *F. proliferatum* strain could reflect differences in adaptations to the different plants from which they originated (maize versus orchids) and/or life style (endophytic *versus* pathogenic). Future comparative analysis of *in planta*-expressed genes of pathogenicity-related gene families, e.g., for cell wall-degrading enzymes, small secreted proteins, or transcription factors will provide a further insight into inter- and intraspecific differences and the probable impact of host specificity and life style on gene expression profiles.

Conclusions

Our study presents high-quality genome sequences of three additional members of the Asian clade of the FFC, namely, *F. mangiferae*, and two strains of *F. proliferatum*. We compared the genomes of these strains to those of *F. fujikuroi* (Wiemann et al. 2013) and *F. verticillioides* (Ma et al. 2010), which belong to the Asian and African-clades of the FFC, respectively, and to the closely related but nonFFC species *F. oxysporum*. We explored the genetic potential of these fusaria to produce SMs, including phytohormones, and tested their ability to produce SMs both *in vitro* and *in planta*. Despite the high level of sequence conservation among species, there are several differences in SM gene content and production that have the potential to contribute to host specificity. Transcriptomic analysis of maize tissue individually infected with the five *Fusarium* isolates revealed SM biosynthetic genes that are upregulated in all the strains and others that were upregulated only in one strain. Among the most remarkable findings are the following:

1. Although most of SM gene clusters are common to the five *Fusarium* strains, the *in vitro* expression of most clusters differs considerably among the strains. Therefore, the presence of a gene cluster does not always correlate with the ability to produce the respective SM.
2. Despite the high level of similarity among the genomes of the sequenced FFC strains, there are SM biosynthetic genes that are unique to each species and may be at least partially responsible for the differences in host specificity or pathogenic *versus* endophytic life style.
3. Members of the FFC may be unique among ascomycetes in their genetic potential to produce three classes of phytohormones: GAs, CKs and auxins. However, the ability to produce these hormones is species- and even strain-specific. In particular, there are fundamental differences between the two *F. proliferatum* strains in their ability to produce GAs and auxins. These differences in

phytohormone biosynthesis may be at least partially responsible for the pathogenic *versus* endophytic life style.
4. In members of the FFC other than *F. proliferatum* ET1, the major auxin, indole-3-acetic acid, is formed not by previously described bacterial IAM pathway but presumably by plant-like pathways via indole-3-acetaldehyde and especially in *F. mangiferae* via indole-3-pyruvic acid using YUCCA enzyme encoded together with CK and GA biosynthesis genes in close proximity on the same chromosome.
5. In axenic culture, CKs are produced mainly by the evolutionary conserved tRNA degradation pathway, which is upregulated in *F. mangifereae* and *F. proliferatum* strain 62905.
6. Although most of the gene clusters are commonly present in the genomes of the FFC members, their expression during the growth in maize may significantly differ between the species and even between two isolates of one species.

Material and Methods

Fungal and Bacterial Strains

Strain IMI58289 (Commonwealth Mycological Institute, Kew, United Kingdom) is a GA-producing WT strain of *F. fujikuroi* used unless otherwise indicated. The mango pathogen *F. mangiferae* MRC7560, originating in Israel, is deposited in the MRC culture collection (National Research Institute for Nutritional Diseases, Tygerberg, South Africa). *F. proliferatum* ET1, *F. proliferatum* 62905, *F. verticillioides* M-3125 (FGSC 7600), and *F. oxysporum* 4287 were provided by Elena Tsavkelova, Moscow State University, Russia, Robert H. Proctor and Daren W. Brown, U. S. Department of Agriculture, USA, and Antonio Di Pietro, Universidad de Córdoba, Spain, respectively. *Escherichia coli* strain Top10 F' (Invitrogen, Groningen, The Netherlands) was used for plasmid propagation. The uracil-auxotrophic *Saccharomyces cerevisiae* FGSC 9721 (FY 834) was provided by the Fungal Genetics Stock Center (Kansas State University) and used for yeast recombination cloning.

Genome Sequencing and Assembly

F. mangiferae

Sequencing was carried out by shot gun sequencing of 350–450 bp and an 8 kb library with paired end 100 bp read length using Illumina HiSeq 2000 by Eurofins MWG Operon, Germany. The assembly was performed by ALLPATHS-LG (Gnerre et al. 2011) using 130-fold pair coverage of the 350 bp library and 5-fold coverage of the 8 kb-LJD mate pairs library resulting in 254 scaffolds (545 contigs) and a genome size of 46.3 Mb. 58 scaffolds were >10 kb and covered 45.9 Mb whereas 124 scaffolds were >2 kb and covered 46.1 Mb.

F. proliferatum ET1

Sequencing and assembly were performed by Eurofins MWG Operon, Germany using the Roche 454 GS-FLX system with the Titanium series chemistry and Illumina HiSeq 2000 (two libraries, 350–450 bp and 8 kb). A part of the Illumina shotgun paired end data (100-fold coverage) was assembled by velvet (v1.2.01) and the script VelvetOptimiser (v2.2.0). The kmer size used was 63. The final assembly was performed using Newbler assembler (v2.6). (input data: the velvet-scaffolds fragmented into pieces of 1998 bases length with an overlap of 1000 bases, a 20-fold coverage of 8 kb-LJD mate pairs, a 10-fold coverage of 8 kb-LJD unpaired shotgun sequences and the 454 shotgun library). The 55 scaffolds were error corrected by mapping all Illumina shotgun paired-end data and further scaffolded using SSPACE (Boetzer et al. 2011) resulting in 32 final scaffolds. About 24 scaffolds were >10 kb and covered 42.2 Mb whereas 32 scaffolds were >2 kb and covered 45.2 Mb.

F. proliferatum 62905

Sequencing process was carried out by shot gun library sequencing (500 bp) with paired end 90 bp read length using Illumina HiSeq 2000 by BGI Tech Solutions (Hong Kong) Co., Limited. Using SOAPde novo, Version 1.05 (Li et al. 2009), the assembly resulted in 155 scaffolds (197 contigs) with a total size of 43.16 Mb. 89 scaffolds were >10 kb and covered 42.97 Mb whereas 122 scaffolds were >2 kb and covered 43.12 Mb).

Structural Annotation

Draft gene models for *F. mangiferae*, *F. proliferatum* ET1, and *F. proliferatum* 62905 were generated by three *de novo* prediction programs: (1) Fgenesh (Salamov and Solovyev 2000) with different matrices (trained on *Aspergillus nidulans*, *Neurospora crassa* and a mixed matrix based on different species); (2) GeneMark-ES (Ter-Hovhannisyan et al. 2008) and (3) Augustus (Stanke et al. 2006) with *Fusarium* ESTs and RNA-seq based transcripts as training sets. Annotation was aided by exonerate (Slater and Birney 2005) hits of protein sequences from *F. verticillioides*, *F. oxysporum*, *F. fujikuroi, F.circinatum, F. proliferatum*, and *F. mangiferae*, respectively to uncover gene annotation gaps. Transcripts were assembled on the RNA-seq data sets using Trinity (Grabherr et al. 2011). The different gene structures and evidence (exonerate mapping, RNA-seq reads and transcripts) were visualized in GBrowse (Donlin 2009), allowing manual validation of coding sequences. The best fitting model per locus was selected manually and gene structures were adjusted by splitting or fusion of gene models and redefining exon–intron boundaries if necessary. tRNAs were predicted using tRNAscan-SE (Lowe and Eddy 1997). The predicted protein sets were searched for highly conserved single (low) copy genes to assess the completeness of the

genomic sequences and gene predictions. Orthologous genes to all 246 single copy genes were identified for all three proteomes by blastp comparisons (eVal: 10^{-3}) against the single-copy families from all 21 species available from the FunyBASE (Aguileta et al 2008). Additionally, all of the 248 core-genes commonly present in higher eukaryotes (CEGs) could be identified by Blastp comparisons (eVal: 10^{-3}) for all three proteomes (Parra et al. 2009).

Functional Annotation of Predicted Open Reading Frames and Data Repositories

The protein coding genes were analyzed and functionally annotated using the PEDANT system (Walter et al. 2009). Data sets are accessible at http://pedant.helmholtz-muenchen.de/genomes.jsp?category=fungal. The genome and annotation data was submitted to the European Nucleotide Archive (ENA, *F. mangiferae*: http://www.ebi.ac.uk/ena/data/view/FCQH01000001-FCQH01000254; *F. proliferatum* ET1: http://www.ebi.ac.uk/ena/data/view/FJOF01000001-FJOF01000032; *F. proliferatum* 62905: http://www.ebi.ac.uk/ena/data/view/FCQG01000001-FCQG01000155).

Genome Comparison

The genomes of *F. verticillioides* and *F. oxysporum* were retrieved from http://www.broadinstitute.org/on October 28, 2013 (Ma et al. 2010). The genome of *F. graminearum*, annotation version 3.2, was retrieved from http://www.helmholtz-muenchen.de/en/ibis/institute/groups/fungal-microbial-genomics/resources/index.html (Wong et al. 2011). Brenda Wingfield, University of Pretoria, South Africa provided access to genome sequence of *F. circinatum* FSP 34 (Wingfield et al. 2012). The genome of *F. fujikuroi* IMI58289 was retrieved from the PEDANT analysis (Wiemann et al. 2013). All genomes were analyzed using the PEDANT system (Walter et al. 2009). To avoid misleading ortholog information based on similarity and bi-directional best hits, QuartetS (Yu et al. 2011) was applied to retrieve a reliable ortholog matrix which was used for all comparative representations. The ortholog data was also used to determine the level of colinearity among species using the tool Orthocluster (Zeng et al. 2008) with standard parameters.

Phylogenetic Tree Analyses of the 12 *Fusarium* Species

The phylogenetic tree of 12 *Fusarium* species was calculated based on the protein sequences of RNA polymerase II subunits RPB1, RPB2 and of the transcription elongation factor gene TEF1α. Orthologs of the sequences were aligned separately using MUSCLE (Edgar 2004). After that, we concatenated the alignments and removed columns with gaps using Gblocks (Talavera and Castresana 2007). To infer the tree, we used the maximum likelihood method PhyML (Guindon et al. 2010) with the LG (Le and Gascuel 2008) substitution model and 100 bootstraps.

Contour-Clamped Homogeneous Electric Field (CHEF)-Gel Analysis

Protoplasts from *Fusarium* strains were prepared as described previously (Tudzynski et al. 1999). The protoplast suspension was mixed with 1.2% InCert agarose (Lonza Group AG, Basel, Switzerland) and then loaded on a CHEF (contour-clamped homogeneous electric field) gel as described in Crawford et al. (2008). Chromosomes of *Schizosaccharomyces pombe* and *Saccharomyces cerevisiae* were used as a molecular size marker (Bio-Rad, Munich, Germany).

Identification of Repetitive DNA Elements

Determination of repeat sequences or the interspersed repeat library for each genome involved first identifying all interspersed repeat elements or families de-novo using RepeatScout (Price et al. 2005) followed by comparing the *de novo* set to known families. New families were comprised of 10 or more repeats and had a consensus sequence length longer than 50 bp. Low complexity and simple sequence repeat families were determined with NSEG (Wootton and Federhen 1996) and Tandem-Repeat-Finder (Benson 1999), in RepeatScout, and then removed from the interspersed repeat library.

The RepBase database (Jurka et al. 2005) was used to detect previously published families of transposable elements, pseudogenes and retroviruses. In order to determine the exact locations of the repetitive elements on the genome we used the RepBase library and the calculated library of interspersed repeat families as input for RepeatMasker (Smit et al. 1996). RepeatMasker was also used to find and mask genomic regions of low complexity. We applied the automated classification tool TEclass (Abrusán et al. 2009) to categorize the predicted repeat sequences into the four main transposable element categories DNA transposon, long interspersed nuclear element (LINE), short interspersed nuclear element (SINE), and retrotransposon with long terminal repeats (LTRs).

In Silico Identification of Secondary Metabolite Clusters

To identify SM clusters in each genome, the InterPro scan results of the PEDANT analysis were used as described (Sieber et al. 2014). Essentially, predicted proteins with homology to a domain of a signature SM core enzyme (e.g., PKS, NRPS, TC, or DMATS) were considered a marker for a gene cluster. A cluster was verified if any neighboring genes with homology to typical SM-cluster enzymes, like P450 monooxygenases, oxidases, methyltransferases, MSF, or ABC transporters or transcription factors, were identified. The extent of each putative gene cluster was then adjusted by comparison to previously published data and to homologous clusters in other *Fusarium* species.

Prediction of Secreted Proteins

We identified putative secreted proteins using a number of approaches. First we used SignalP 4.1 (Petersen et al. 2011), with a *D*-cutoff ≥ 0.45 (for SignalP-noTM networks); ≥ 0.5 (for SignalP-TM networks), and a TargetP RC score less than four (Emanuelsson et al. 2000) and SecretomeP hit, with a cutoff score of 0.65 (Bendtsen et al. 2004), to identify putative protein secretion signals. To add nonclassically secreted proteins from the remaining proteome we selected on predicted secreted proteins using SecretomeP with a cutoff score of 0.65 but excluded membrane proteins by utilizing TMHMM (Krogh et al. 2001). The set of nonclassically secreted proteins were further filtered for those which are preferentially predicted as extracellular or plasma membrane by WoLF PSORT (Tamura and Akutsu 2007).

Phylogenetic Analyses

Phylogenetic analyses of the coding regions of phytohormone biosynthetic and PKS genes were done using the computer program MEGA6 (Tamura et al. 2013). With the program, coding region sequences were translated, and the deduced amino acid sequences were then aligned using the ClustalW or Muscle alignment methods. Following alignment, the amino acid sequences were converted back to nucleotide sequences before subsequent analyses. The most appropriate nucleotide substitution model for each alignment was determined using the Find Best DNA/Protein Models function in MEGA6. The model with the lowest score for the Bayesian Information Criterion was used in subsequent phylogenetic analyses. Trees were generated with both maximum likelihood and maximum parsimony methods. However, because trees resulting from the two analyses were similar, the results from only one method or the other are presented. In trees, statistical support for branches was determined by bootstrap analysis with 500 pseudoreplicates. Alignments and phylogenetic analyses of sesquiterpene cyclase homologs (STCs) were performed using the Tree builder option in the program Geneious 7.0.6 software.

Culture Conditions

For SM production experiments, strains were first grown for 3 days in 300-mL Erlenmeyer flasks with 100 ml Darken medium (Darken et al. 1959) on a rotary shaker at 180 rpm at 28 °C. About 500 µL of this culture were then used to inoculate 100 mL of ICI (Imperial Chemical Industries, UK) medium (Krogh et al. 2001) containing either 6 mM glutamine, 60 mM glutamine or 6 mM sodium nitrate. For auxin and cytokinin analysis, the WT and mutant strains were grown in ICI medium with reduced (5 g/L) glucose concentration and 6 mM $NaNO_3$ as a nitrogen source. For secondary metabolite and phytohormone analysis, the strains were grown for 7 days on a rotary shaker at 28 °C in the dark. The mycelia were harvested, washed with deionized water and lyophilized.

For total indole analysis by thin layer chromatography (TLC) the wild type and *IAAM* and *IAAH* OE mutants were grown in ICI medium with the addition of 5 mM tryptophan (precursor of the IAM pathway) for 48 h. The culture filtrates were used for extraction of indoles.

For RNA preparation under *in vitro* conditions, strains were grown on CM agar covered with cellophane sheets for five days. The mycelia were flash frozen with liquid nitrogen prior to lyophilization. For *in planta* expression studies by RNA-seq, lyophilized roots of infected maize plants (4, 10, and 20 days postinfection) were used for RNA preparation.

Plasmid Construction

The vectors were generated using the yeast recombinational cloning system (Schumacher 2012). The pOE-iaaM and pOE:iaaH vectors were generated by first PCR amplification of the *IAAM* and *IAAH* genes from genomic DNA of *F. proliferatum* ET1 using the primer pairs OE_GPD_IAAM_FW/RV and OE_GPD_IAAH_FW/RV (supplementary table S5, Supplementary Material online). *S. cerevisiae* FY834 was transformed with a *Hin*dIII restricted, modified pRS426 plasmid (Colot et al. 2006), containing either the nourseothricin and a hygromycin resistance cassette, and the PCR product of either the *IAAM* or the *IAAH* gene. *IAAH* expression was directed by the *F. fujikuroi glnA* promoter whereas *IAAM* expression was directed by the *A. nidulans gpd* promoter. The *IL1*, *IL2*, *P4501*, and *P450* overexpression mutants were created by first amplifying each respective gene from *F. fujikuroi* (for primers see supplementary table S5, Supplementary Material online). *S. cerevisiae* FY834 was transformed with the obtained fragments together with the *Not*I- and *Nco*I-restricted, plasmid pNAN-OGG for pOE:IL1 and plasmid pNDN-OCT (Schumacher 2012) for pOE:IL2 which both include the nourseothricin resistance cassette. For the construction of pOE:P450-1 and pOE:P4502 a *Hin*dIII digested vector driven by the *gpd* promoter was used. The plasmids contain a hygromycin resistance cassette.

Transformations

Transformation of *Fusarium* spp. protoplasts was carried out according to (Wiemann et al. 2013). Regeneration of transformed protoplasts was performed over 4–5 days at 28 °C in a regeneration medium (0.7 M sucrose, 0.05% yeast extract) containing either 100 µg/mL nourseothricin (Werner-Bioagents, Jena, Germany) or 100 µg/mL hygromycin (Calbiochem, Darmstadt, Germany). Transformants were purified by single spore isolation (Wiemann et al. 2012) to homokaryons. Vector integration events were confirmed by diagnostic PCR using specific primers as indicated (supplementary table S5, Supplementary Material online).

Chemical Analysis of Secondary Metabolites

For analyses of the secondary metabolites, strains were grown in submerged cultures as described above. After seven days, mycelia were removed from the culture by filtration through Mirachloth (Calbiochem, Merck KGaA, Darmstadt, Germany). Small particulates were removed from the culture filtrates using 0.45 μm syringe filters (BGB®, Schloßböckelheim, Germany) which were then analyzed by high-pressure liquid chromatography with a diode array detector (HPLC-DAD; Hitachi Chromaster LC equipped with a 250 mm × 4.60 mm i.d., 5 μm, Gemini® C_{18} with a 4 mm × 3 mm Gemini® C_{18} guard column (Phenomenex, Aschaffenburg, Germany), a with 5160 pump, 5260 autosampler, 5310 column oven and 5430 DAD (VWR International GmbH, Darmstadt, Germany). For data analyses, the software EZChrome Elite (VWR, Darmstadt, Germany) was used. The filtrates were separated using a previously described method to identify bikaverin, fusarubins, fusarins, and fusaric acid (Wiemann et al. 2012). Prior to gibberellin analyses, nonpolar compounds were removed from 20 mL of the supernatant with a Speed SPE Octadecyl C_{18} cartridge (Applied Separation) as described previously (Wiemann et al. 2012). The samples were separated on an Ascentis Express C_{18} column, 10 cm × 3 mm, 2.7 μm (Sigma-Aldrich, Merck KGaA, Darmstadt, Germany). A gradient system with 0.05% H_3PO_4 in water (v/v) (eluent A) and acetonitrile (eluent B) was applied, beginning with 85% A for 8 min. In 10 min, eluent A was decreased to 60%. After holding these conditions for 2 min, the concentrations of eluent A and B were brought back to the starting conditions over 1 min and equilibrated for 3 min.

For analyses of volatile secondary metabolites, solid CM medium were inoculated with one piece of a preculture grown on the same medium. Cultures were grown for 5 days at 28 °C in the dark. The volatile compounds were collected by trapping on a charcoal filter using the CLSA technique as described before (Citron et al. 2012). The trapped compounds were extracted with dichloromethane (50 μL) and the resulting extract was subjected to GC-MS analysis using an HP 7890B gas chromatograph connected to an HP 5977A inert mass detector fitted with an HP5-MS (30 m, 0.25 mm i. d., 0.50 μm film) fused silica capillary column. Instrumental parameters were (1) inlet pressure, 77.1 kPa, He 23.3 mL min^{-1}, (2) injection volume, 2 μL, (3) transfer line, 250 °C, and (4) electron energy 70 eV. The GC program used was as follows: 5 min at 50 °C increasing at 5 °C min^{-1} to 320 °C, and operated in split mode (50:1, 60 s valve time). Helium was used as carrier gas at 1 mL min^{-1}. Retention indices (*I*) were determined from retention times of *n*-alkane standards (C_8–C_{40}). Compound identification was based on comparison of mass spectra to data base spectra and of retention indices to tabulated literature data, and by comparison to authentic standards.

Phytohormone Analysis

Stable isotope-labeled cytokinins (CK) internal standards (OlChemIm, Czech Republic) were added, each at 1 ppmol per sample, to check the recovery during the purification and to validate the determination. Precise quantification of CKs was done by high-pressure liquid chromatography (Acquity UPLC™; Waters, USA) coupled to a triple quadrupole mass detector (Xevo TQ MSTM; Waters) equipped with and electro-spray interface directly as described in (Hinsch et al. 2015). Lyophilized media or mycelia of overexpression mutants were extracted without addition of internal standards and purified on immuno-affinity columns as described (Hinsch et al. 2015). Samples were analyzed for gibberellins using the method described in Urbanova et al. (2013). Relative indole content was estimated by the Salkowski method (Gordon and Weber 1951) essentially as described (Tsavkelova et al. 2012). Briefly, 1 mL of culture filtrate was mixed with 1 ml of Salkowski reagent (1.015 g $FeCl_3 \cdot 6H_2O$, 150 mL H_2SO_4, 250 mL ddH_2O). The pH was adjusted to 2.8 with 0.1 N HCl, the tubes were incubated for 15 min at 30 °C for color development and then the absorbance at OD_{540} was determined. Standard curves were prepared from serial dilutions of 100 mM IAA stock solution and the relative levels of indoles were calculated accordingly. TLC analysis was used for specific detection of IAA and IAM in extracts of total indoles. Aliquots of total indoles were removed before staining with Salkowski reagent and 0.8 mL samples were extracted with ethyl acetate:water (1:2, v/v) by vigorous shaking for 10 min. The ethyl acetate fraction was removed to a light protected tube, evaporated to dryness and the solid residue was dissolved in 30 μL methanol. Samples were spotted onto a silica gel 60 F254 plates (Merck) and the indoles were developed with ethyl acetate:isopropanol:ammonia (45:35:20 v/v) as the mobile phase. After separation, the plates were dried, sprayed with van Ehmann's reagent (a 3:1 mixture of Salkowski's and Ehrich's reagents) and heated to 90 °C until spots were developed. The Rfs and colors of the spots were compared with 1 mM each of IAA and IAM standard compounds.

RNA Isolation

Total RNA was extracted from mycelia grown for 5 days on CM agar and from infected maize seedling roots after 4, 10, or 20 dpi using TRIzol Reagent (Life Technology, Karlsruhe, Germany) and purified using an RNeasy Plant MinElute Cleanup Kit (Qiagen, Hilden, Germany). The quality of DNase-treated RNA (28S:18S > 1.0; RIN≥ 6.5; OD260/280 ≥1.8; OD260/230 ≥ 1.8) was determined using an Agilent Bioanalyzer. The high quality RNA was sent to BGI Tech Solutions (Hong Kong) for library construction and sequencing.

PCR and Quantitative Real-Time PCR

PCR mixtures contained 25 ng of template DNA, 50 ng of each primer (supplementary table S5, Supplementary

Material online), 0.2 mM deoxynucleoside triphosphates, and 1 U of Biotherm Taq polymerase (Genecraft, Lüdinghausen, Germany). Reverse transcription-PCR (RT-PCR) was performed using the Superscript II (Invitrogen, Groningen, The Netherlands) and 1.5 µg of total RNA as the template, according to the manufacturer's instructions. Quantitative Real-Time PCR (qPCR) was performed using iTaq Universal SYBR Green Supermix (BioRad) and Superscript II cDNA as template, in a Biorad thermocycler iTaq. In all cases the qPCR efficiency was between 90% and110% and the annealing temperature was 60 °C. Every sample was run twice and the relative expression levels were calculated using the delta-delta-Ct method (Pfaffl 2001). The expression of an actin gene (primers FRAC_RTPCR_FW and FRAC_RTPCR_RV), a GDP-mannose transporter gene (primers FGMT_RTPCR_FW and FGMT_RTPCR_RV) and an ubiquitin gene (primers FUB_RTPCR_FW and FUB_RTPCR_RV) were used as references. The following primers were used for amplification of the indicated genes: for *IAAM*: CHECK_IAAM_FW and CHECK_I AAM_RV, and for *IAAH*: CHECK_IAAH_FW and CHECK_IAAH_RV (supplementary table S5, Supplementary Material online).

RNA Sequencing, Mapping and Quantification

Total RNA was sequenced using Illumina HiSeq2000 technology by BGI Tech Solutions (Hong Kong) Co., Limited. RNA-seq reads were mapped on the reference genome using tophat2 (v2.0.8). The interval for allowed intron lengths was set to min 20 nt and max 1 kb (Trapnell et al. 2009; Kim et al. 2013). We used cufflinks to determine the abundance of transcripts in FPKM (Fragments Per Kilobase of exon per Million fragments mapped) and calculated differentially expressed genes using cuffdiff (Trapnell et al. 2009, 2012). The gene models were included as raw junctions. Genes with a minimum of fourfold increase or decrease in expression ($|\log2$ of the FPKM values $+1| \geq 2$) between the two experimental conditions were considered as regulated. The RNA-seq data has been deposited in NCBI's Gene Expression Omnibus (Edgar et al. 2002) and are accessible through GEO Series accession number GSE77595.

Maize Seedling Assay

Fungicide-free seeds of the hybrid white sweet corn variety Silver Queen (Johnny's Selected Seeds, USA) were soaked in 0.82% sodium hypochlorite for one minute and then rinsed twice in sterile water for one minute each time. Seeds were then soaked in 30 mL of a culture of individual *Fusarium* strains grown for 3 days in liquid mung bean medium (Bai and Shaner 1996). After 2 days of soaking, 10 seeds were sown onto the surface of a water-saturated soil mixture consisting of sphagnum peat moss, vermiculite and dolomite lime (Sunshine Redi-Earth Professional Growing Mix), contained in a 10 cm^2 plastic pot. The seeds were then covered with approximately 1 cm of additional soil. The pots were incubated in a growth chamber with a light dark cycle consisting of 14 h light at 30 °C and 10 h dark at 20 °C. Disease severity was assessed as follows at seven days after sowing: (i) percent germination was determined by counting the number of seedlings per 10 seeds sown in each pot; and (ii) seedling height was determined by measuring the length of each seedling from the soil line to the top of the longest leaf. The resulting data were subjected to analysis of variance (ANOVA), and statistically significant differences between means were determined with a least squares means test using a Bonferroni adjustment. Statistical analyses were done using SAS Statistical Software (SAS Institute Inc.).

Microscopic Observation of Fungal Development in Maize

Maize seeds were inoculated as follows. Fungal biomass was produced by culturing for three days in potato dextrose broth (PDB) (Sigma-Aldrich). The mycelia and conidia were harvested by centrifugation and dispensed in water at 10 g/L. Barley seeds were added to the suspension and the mix was stirred for 15 min. The water was removed and the seeds coated with fungal biomas were spread on a tray and left in a fume hood overnight to dry. Soil was inoculated by mixing fungus-coated seeds with soil at the desired amount (5 g/L unless otherwise mentioned). Maize seeds were inoculated as follows: seeds were first incubated for three days in darkness at 25 °C, in humid Petri dishes and then planted in 1 L pots, 3 seeds per pot. Prior to sowing, the soil was inoculated with barley seeds coated with fungal biomass at 15 g/kg. Plants were sampled at 10 and 20 dpi. Roots and crown were collected, washed thoroughly in water to remove soil and external debris, the tissues were stained with propidium iodide (PI) by incubation in 10 µg/mL PI for 1 min, and then washed again with water. Hand sections of tissue samples were visualized with a Zeiss LSM 780 confocal microscope using excitation and emission wavelengths of 490/526 nm for visualization of GFP and 597/664 nm for visualization of PI.

Supplementary Material

Supplementary tables S1–S5 and figures S1–S13 are available at *Genome Biology and Evolution* online (http://www.gbe.oxfordjournals.org/).

Acknowledgments

This study was supported by the Deutsche Forschungsgemeinschaft (DFG), projects TU101/17-2 and GU1205/2-2, by the Austrian Science Fund FWF (special research project Fusarium, F3705/DFG ME1682/6-1) and the National Science Foundation, Czech Republic (grant number 16-10602S). Mention of trade names or commercial products in this publication is solely for the purpose of providing specific

information and does not imply recommendation or endorsement by the US Department of Agriculture. USDA is an equal opportunity provider and employer. We are grateful for the technical assistance of Stephanie Folmar, Marcie Moore and Crystal Probyn. We thank Martijn Rep (University of Amsterdam, The Netherlands) for support in performing CHEF gel analysis. We thank also José J. Espino for transforming *Fusarium* strains.

Literature Cited

Abrusán G, Grundmann N, Demester L, Makalowski W. 2009. TEclass—a tool for automated classification of unknown eukaryotic transposable elements. Bioinformatics 25:1329–1330.
Aguileta G, et al. 2008. Assessing the performance of single-copy genes for recovering robust phylogenies. Syst Biol. 57:613–627.
Akagi Y, Akamatsu H, Otani H, Kodama M. 2009. Horizontal chromosome transfer, a mechanism for the evolution and differentiation of a plant-pathogenic fungus. Eukaryot Cell 8:1732–1738.
Baer P, et al. 2014. Induced-fit mechanism in class I terpene cyclases. Angew Chem Int Ed Engl. 53:7652–7656.
Bai G, Shaner G. 1996. Variation in *Fusarium graminearum* and cultivar resistance to wheat scab. Plant Dis. 80:975–979.
Bajguz A, Piotrowska A. 2009. Conjugates of auxin and cytokinin. Phytochemistry 70:957–969.
Bendtsen JD, Jensen LJ, Blom N, Von Heijne G, Brunak S. 2004. Feature-based prediction of non-classical and leaderless protein secretion. Protein Eng Des Sel. 17:349–356.
Benson G. 1999. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. 27:573–580.
Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W. 2011. Scaffolding pre-assembled contigs using SSPACE. Bioinformatics 27:578–579.
Bömke C, Tudzynski B. 2009. Diversity, regulation, and evolution of the gibberellin biosynthetic pathway in fungi compared to plants and bacteria. Phytochemistry 70:1876–1893.
Brakhage AA. 2013. Regulation of fungal secondary metabolism. Nat Rev Microbiol. 11:21–32.
Britz H, et al. 2002. Two new species of *Fusarium* section *Liseola* associated with mango malformation. Mycologia 94:722–730.
Brock NL, Huss K, Tudzynski B, Dickschat JS. 2013. Genetic dissection of sesquiterpene biosynthesis by *Fusarium fujikuroi*. Chembiochem 14:311–315.
Brock NL, Tudzynski B, Dickschat JS. 2011. Biosynthesis of sesqui- and diterpenes by the gibberellin producer *Fusarium fujikuroi*. Chembiochem 12:2667–2676.
Brown DW, Butchko RA, Baker SE, Proctor RH. 2012a. Phylogenomic and functional domain analysis of polyketide synthases in *Fusarium*. Fungal Biol. 116:318–331.
Brown DW, Butchko RA, Busman M, Proctor RH. 2012b. Identification of gene clusters associated with fusaric acid, fusarin, and perithecial pigment production in *Fusarium verticillioides*. Fungal Genet Biol. 49:521–532.
Brown DW, et al. 2015. Identification of a 12-gene fusaric acid biosynthetic gene cluster in *Fusarium* species through comparative and functional genomics. Mol Plant Microbe Interact. 28:319–332.
Brown DW, Proctor RH. 2016. Insights into natural products biosynthesis from analysis of 490 polyketide synthases from *Fusarium*. Fungal Genet Biol. 89:37–51.
Brown DW, Proctor RH. 2013. Diversity of polyketide synthases in *Fusarium*. In: Brown DW, Proctor RH, editors. Fusarium: genomics, molecular and cellular biology. Lewiston (NY): Caister Academic Press. p. 143–164.
Burkhardt I, et al. 2016. Mechanistic characterization of two sesquiterpene cyclases from the plant pathogen *Fusarium fujikuroi*. Angew Chem. 55:8748–8751.
Bushley KE, Turgeon BG. 2010. Phylogenomics reveals subfamilies of fungal nonribosomal peptide synthetases and their evolutionary relationships. BMC Evol Biol. 10:26.
Chanclud E, et al. 2016. Cytokinin production by the rice blast fungus is a pivotal requirement for full virulence. PLoS Pathog. 12:e1005457.
Chiara M, et al. 2015. Genome sequencing of multiple isolates highlights subtelomeric genomic diversity within *Fusarium fujikuroi*. Genome Biol Evol. 7:3062–3069.
Citron CA, Rabe P, Dickschat JS. 2012. The scent of bacteria: headspace analysis for the discovery of natural products. J Nat Prod. 75:1765–1776.
Coleman JJ, et al. 2009. The genome of *Nectria haematococca*: contribution of supernumerary chromosomes to gene expansion. PLoS Genet. 5:e1000618.
Colot HV, et al. 2006. A high-throughput gene knockout procedure for *Neurospora* reveals functions for multiple transcription factors. Proc Natl Acad Sci U S A. 103:10352–10357.
Crawford JM, Vagstad AL, Ehrlich KC, Townsend CA. 2008. Starter unit specificity directs genome mining of polyketide synthase pathways in fungi. Bioorg Chem. 36:16–22.
Cuomo CA, et al. 2007. The *Fusarium graminearum* genome reveals a link between localized polymorphism and pathogen specialization. Science 317:1400–1402.
Darken MA, Jensen AL, Shu P. 1959. Production of gibberellic acid by fermentation. Appl Microbiol. 7:301–303.
Desjardins AE. 2006. Fusarium mycotoxins: chemistry, genetics, and biology. St. Paul (MN): American Phytopathological Society (APS) Press.
Dickschat JS. 2015. Bacterial terpene cyclases. Nat Prod Rep. 33:87–110.
Ditengou FA, et al. 2015. Volatile signalling by sesquiterpenes from ectomycorrhizal fungi reprogrammes root architecture. Nat Commun. 6:6279.
Donlin MJ. 2009. Using the Generic Genome Browser (GBrowse). Curr Protoc Bioinformatics Chapter 9: Unit 9.9.
Edgar R, Domrachev M, Lash AE. 2002. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res. 30:207–210.
Edgar RC. 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics 5:1.
Emanuelsson O, Nielsen H, Brunak S, von Heijne G. 2000. Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. J Mol Biol. 300:1005–1016.
Freeman S, Shtienberg D, Maymon M, Levin AG, Ploetz RC. 2014. New insights into mango malformation disease epidemiology lead to a new integrated management strategy for subtropical environments. Plant Dis. 98:1456–1466.
Frisvad JC, Smedsgaard J, Samson RA, Larsen TO, Thrane U. 2007. Fumonisin B2 production by *Aspergillus niger*. J Agric Food Chem. 55:9727–9732.
Gardiner DM, et al. 2012. Comparative pathogenomics reveals horizontally acquired novel virulence genes in fungi infecting cereal hosts. PLoS Pathog. 8:e1002952.
Gawehns F, et al. 2014. The *Fusarium oxysporum* effector Six6 contributes to virulence and suppresses I-2-mediated cell death. Mol Plant Microbe Interact. 27:336–348.
Geiser DM, et al. 2013. One fungus, one name: defining the genus *Fusarium* in a scientifically robust way that preserves longstanding use. Phytopathology 103:400–408.
Gnerre S, et al. 2011. High-quality draft assemblies of mammalian genomes from massively parallel sequence data. Proc Natl Acad Sci U S A. 108:1513–1518.

Gohari AM, et al. 2015. Effector discovery in the fungal wheat pathogen *Zymoseptoria tritici*. Mol Plant Pathol. 16:931–945.
Gordon SA, Weber RP. 1951. Colorimetric estimation of indoleacetic acid. Plant Physiol. 26:192–195.
Grabherr MG, et al. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol. 29:644–652.
Guindon S, et al. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol. 59:307–321.
Hansen FT, et al. 2015. An update to polyketide synthase and non-ribosomal synthetase genes and nomenclature in *Fusarium*. Fungal Genet Biol. 75:20–29.
Hinsch J, Galuszka P, Tudzynski P. 2016. Functional characterization of the first filamentous fungal tRNA-isopentenyltransferase and its role in the virulence of *Claviceps purpurea*. New Phytol. 211:980–992.
Hinsch J, et al. 2015. De novo biosynthesis of cytokinins in the biotrophic fungus *Claviceps purpurea*. Environ Microbiol. 17:2935–2951.
Hoffmeister D, Keller NP. 2007. Natural products of filamentous fungi: enzymes, genes, and their regulation. Nat Prod Rep. 24:393–416.
Jin JM, et al. 2010. Functional characterization and manipulation of the apicidin biosynthetic pathway in *Fusarium semitectum*. Mol Microbiol. 76:456–466.
Jurka J, et al. 2005. Repbase Update, a database of eukaryotic repetitive elements. Cytogenet Genome Res. 110:462–467.
Kasahara H. 2016. Current aspects of auxin biosynthesis in plants. Biosci Biotechnol Biochem. 80:34–42.
Khaldi N, Wolfe KH. 2011. Evolutionary origins of the fumonisin secondary metabolite gene cluster in *Fusarium verticillioides* and *Aspergillus niger*. Int J Evol Biol. 2011:423821.
Kim D, et al. 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. 14:R36.
Krogh A, Larsson B, von Heijne G, Sonnhammer EL. 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol. 305:567–580.
Le SQ, Gascuel O. 2008. An improved general amino acid replacement matrix. Mol Biol Evol. 25:1307–1320.
Leslie JF, Summerell BA. 2006. The Fusarium laboratory manual, Oxford, UK: Blackwell Publishing.
Li R, et al. 2009. SOAP2: an improved ultrafast tool for short read alignment. Bioinformatics 25:1966–1967.
Lo Presti L, et al. 2015. Fungal effectors and plant susceptibility. Annu Rev Plant Biol. 66:513–545.
López-Berges, et al. 2013. The velvet complex governs mycotoxin production and virulence of Fusarium oxysporum on plant and mammalian hosts. Mol Microbiol. 87:49–65.
Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 25:955–964.
Lysøe E, et al. 2014. The genome of the generalist plant pathogen *Fusarium avenaceum* is enriched with genes involved in redox, signaling and secondary metabolism. PLoS One 9:e112703.
Ma LJ, et al. 2013. *Fusarium* pathogenomics. Annu Rev Microbiol. 67:399–416.
Ma LJ, et al. 2010. Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*. Nature 464:367–373.
Maor R, Haskin S, Levi-Kedmi H, Sharon A. 2004. In planta production of indole-3-acetic acid by *Colletotrichum gloeosporioides* f. sp. *aeschynomene*. Appl Environ Microbiol. 70:1852–1854.
Marasas WF, et al. 2004. 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. 134:711–716.
Mashiguchi K, et al. 2011. The main auxin biosynthesis pathway in *Arabidopsis*. Proc Natl Acad Sci U S A. 108:18512–18517.
Morgavi DP, Riley RT. 2007. An historical overview of field disease outbreaks known or suspected to be caused by consumption of feeds contaminated with *Fusarium* toxins. Anim Feed Sci Technol. 137:201–212.
Munawar A, Marshall JW, Cox RJ, Bailey AM, Lazarus CM. 2013. Isolation and characterisation of a ferrirhodin synthetase gene from the sugarcane pathogen *Fusarium sacchari*. Chembiochem 14:388–394.
Munkvold G, Stahr HM, Logrieco A, Moretti A, Ritieni A. 1998. Occurrence of fusaproliferin and beauvericin in *Fusarium*-contaminated livestock feed in Iowa. Appl Environ Microbiol. 64:3923–3926.
Niehaus E-M, et al. 2014a. Apicidin F: characterization and genetic manipulation of a new secondary metabolite gene cluster in the rice pathogen *Fusarium fujikuroi*. PLoS One 9:e103336.
Niehaus E-M, et al. 2014b. Characterization of the fusaric acid gene cluster in *Fusarium fujikuroi*. Appl Microbiol Biotechnol. 98:1749–1762.
Niehaus E-M, et al. 2016. Sound of silence: the beauvericin cluster in *Fusarium fujikuroi* is controlled by cluster-specific and global regulators mediated by H3K27 modification. Environ Microbiol. doi: 10.1111/1462-2920.13576.
O'Donnell K, Cigelnik E, Nirenberg HI. 1998a. Molecular systematics and phylogeography of the *Gibberella fujikuroi* species complex. Mycologia 90:465–493.
O'Donnell K, Kistler HC, Cigelnik E, Ploetz RC. 1998b. Multiple evolutionary origins of the fungus causing Panama disease of banana: concordant evidence from nuclear and mitochondrial gene genealogies. Proc Natl Acad Sci U S A. 95:2044–2049.
O'Donnell K, et al. 2013. Phylogenetic analyses of RPB1 and RPB2 support a middle Cretaceous origin for a clade comprising all agriculturally and medically important fusaria. Fungal Genet Biol. 52:20–31.
Oide S, Berthiller F, Wiesenberger G, Adam G, Turgeon BG. 2015. Individual and combined roles of malonichrome, ferricrocin, and TAFC siderophores in *Fusarium graminearum* pathogenic and sexual development. Front Microbiol. 5:759.
Oren L, Ezrati S, Cohen D, 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.
Parra G, Bradnam K, Ning Z, Keane T, Korf I. 2009. Assessing the gene space in draft genomes. Nucleic Acids Res. 37:289–297.
Petersen TN, Brunak S, von Heijne G, Nielsen H. 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods 8:785–786.
Pfaffl MW. 2001. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. 29:e45.
Ploetz RC, Freeman S. 2009. Foliar, floral and soilborne diseases, Chapter 8. In: Mukherjee SK, Litz RE, editors. The Mango: Botany, Production and Uses. Bodmin, UK: MPG Books Group. p. 231.
Price AL, Jones NC, Pevzner PA. 2005. *De novo* identification of repeat families in large genomes. Bioinformatics 21 Suppl 1:i351–i358.
Proctor RH, Desjardins AE, Moretti A. 2009. Biological and chemical complexity of *Fusarium proliferatum*. In: Strange RN, Lodovica Gullino M, editors. The role of plant pathology in food safety and food security. Berlin: Springer, pp. 97–111.
Proctor RH, et al. 2013. Birth, death and horizontal transfer of the fumonisin biosynthetic gene cluster during the evolutionary diversification of *Fusarium*. Mol Microbiol. 90:290–306.
Rep M, Kistler HC. 2010. The genomic organization of plant pathogenicity in *Fusarium* species. Curr Opin Plant Biol. 13:420–426.
Rösler SM, Sieber CMK, Humpf H-U, Tudzynski B. 2016. Interplay between pathway-specific and global regulation of the fumonisin gene cluster

in the rice pathogen *Fusarium fujikuroi*. Appl Microbiol Biotechnol. 100:5869–5882.
Salamov AA, Solovyev VV. 2000. Ab initio gene finding in *Drosophila* genomic DNA. Genome Res. 10:516–522.
Schumacher J. 2012. Tools for *Botrytis cinerea*: new expression vectors make the gray mold fungus more accessible to cell biology approaches. Fungal Genet Biol. 49:483–497.
Sieber CMK, et al. 2014. The *Fusarium graminearum* genome reveals more secondary metabolite gene clusters and hints of horizontal gene transfer. PLoS One 9:e110311.
Slater GS, Birney E. 2005. Automated generation of heuristics for biological sequence comparison. BMC Bioinformatics 6:31.
Smit AFA, Hubley R, Green P. 1996. RepeatMasker Open-3.0. Seattle (WA): Institute for Systems Biology.
Stanke M, et al. 2006. AUGUSTUS: ab initio prediction of alternative transcripts. Nucleic Acids Res. 34:W435–W439.
Studt L, et al. 2016. Two separate key enzymes and two pathway-specific transcription factors are involved in fusaric acid biosynthesis in *Fusarium fujikuroi*. Environ Microbiol. 18:936–956.
Studt L, Wiemann P, Kleigrewe K, Humpf H-U, Tudzynski B. 2012. Biosynthesis of fusarubins accounts for pigmentation of *Fusarium fujikuroi* perithecia. Appl Environ Microbiol. 78: 4468–4480.
Talavera G, Castresana J. 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Syst Biol. 56:564–577.
Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol. 30:2725–2729.
Tamura T, Akutsu T. 2007. Subcellular location prediction of proteins using support vector machines with alignment of block sequences utilizing amino acid composition. BMC Bioinformatics 8:466.
Ter-Hovhannisyan V, Lomsadze A, Chernoff YO, Borodovsky M. 2008. Gene prediction in novel fungal genomes using an ab initio algorithm with unsupervised training. Genome Res. 18:1979–1990.
Trapnell C, Pachter L, Salzberg SL. 2009. TopHat: discovering splice junctions with RNA-Seq. Bioinformatics 25:1105–1111.
Trapnell C, et al. 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc. 7:562–578.
Troncoso C, et al. 2010. Gibberellin biosynthesis and gibberellin oxidase activities in *Fusarium sacchari*, *Fusarium konzum* and *Fusarium subglutinans* strains. Phytochemistry 71:1322–1331.
Tsavkelova EA, et al. 2003a. Fungi associated with orchid roots in greenhouse conditions. Mikologiia I Fitopatologiia 37:57–63.
Tsavkelova EA, Lobakova ES, Kolomeitseva GL, Cherdyntseva TA, Netrusov AI. 2003b. Localization of associative cyanobacteria on the roots of epiphytic orchids. Microbiology 72:86–91.
Tsavkelova EA, et al. 2012. Identification and functional characterization of indole-3-acetamide-mediated IAA biosynthesis in plant-associated *Fusarium* species. Fungal Genet Biol. 49:48–57.
Tudzynski B, Homann V, Feng B, Marzluf GA. 1999. Isolation, characterization and disruption of the *areA* nitrogen regulatory gene of *Gibberella fujikuroi*. Mol Gen Genet. 261:106–114.
Tudzynski B, Sharon A. 2002. Biosynthesis, biological role and application of fungal phytohormones. In: Osiewacz HD, editor. Industrial applications. Berlin: Springer. pp. 183–211.
Urbanova T, Tarkowska D, Novak O, Hedden P, Strnad M. 2013. Analysis of gibberellins as free acids by ultra performance liquid chromatography-tandem mass spectrometry. Talanta 112:85–94.
Vergara IA, Chen N. 2009. Using OrthoCluster for the detection of synteny blocks among multiple genomes. Curr Protoc Bioinformatics Chapter 6: Unit 6.10 6.10.1–18.
von Bargen KW, et al. 2015. Isolation and structure elucidation of fujikurins A-D: products of the PKS19 gene cluster in *Fusarium fujikuroi*. J Nat Prod. 78:1809–1815.
Walter MC, et al. 2009. PEDANT covers all complete RefSeq genomes. Nucleic Acids Res. 37:D408–D411.
Wiemann P, et al. 2012. The Sfp-type 4′-phosphopantetheinyl transferase Ppt1 of *Fusarium fujikuroi* controls development, secondary metabolism and pathogenicity. PLoS One 7:e37519.
Wiemann P, 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.
Wight WD, Kim KH, Lawrence CB, Walton JD. 2009. Biosynthesis and role in virulence of the histone deacetylase inhibitor depudecin from *Alternaria brassicicola*. Mol Plant Microbe Interact. 22:1258–1267.
Williams AH, et al. 2016. Comparative genomics and prediction of conditionally dispensable sequences in legume–infecting Fusarium oxysporum formae speciales facilitates identification of candidate effectors. BMC Genomics 17:1.
Wingfield BD, et al. 2012. First fungal genome sequence from Africa: a preliminary analysis. S Afr J Sci. 108:01–09.
Wingfield MJ, et al. 2008. Pitch canker caused by *Fusarium circinatum*—a growing threat to pine plantations and forests worldwide. Australas Plant Pathol. 37:319–334.
Wong P, et al. 2011. FGDB: revisiting the genome annotation of the plant pathogen *Fusarium graminearum*. Nucleic Acids Res. 39:D637–D639.
Wootton JC, Federhen S. 1996. Analysis of compositionally biased regions in sequence databases. Meth Enzymol. 266:554–571.
Wu F. 2007. Measuring the economic impacts of *Fusarium* toxins in animal feeds. Anim Feed Sci Technol. 137:363–374.
Xu JR, Leslie JF. 1996. A genetic map of Gibberella fujikuroi mating population A (*Fusarium moniliforme*). Genetics 143:175–189.
Xu Y, et al. 2008. Biosynthesis of the cyclooligomer depsipeptide beauvericin, a virulence factor of the entomopathogenic fungus *Beauveria bassiana*. Chem Biol. 15:898–907.
Yu C, Zavaljevski N, Desai V, Reifman J. 2011. QuartetS: a fast and accurate algorithm for large-scale orthology detection. Nucleic Acids Res. 39:e88.
Zeng X, et al. 2008. OrthoCluster: a new tool for mining synteny blocks and applications in comparative genomics. 11th International Conference on Extending Technology (EDBT), March 25–30, 2008, Nantes, France. pp. 656–667.

Associate editor: Rebecca Zufall