



Evolution and Diversity of Biosynthetic Gene Clusters in *Fusarium*

Koen Hoogendoorn^{1,2}, Lena Barra³, Cees Waalwijk², Jeroen S. Dickschat³,
Theo A. J. van der Lee^{2*} and Marnix H. Medema^{1*}

¹ Bioinformatics Group, Wageningen University, Wageningen, Netherlands, ² Biointeractions and Plant Health, Plant Research International, Wageningen University and Research, Wageningen, Netherlands, ³ Kekulé-Institut für Organische Chemie und Biochemie, Rheinische Friedrich-Wilhelms-Universität Bonn, Bonn, Germany

OPEN ACCESS

Edited by:

Mohamed Hijri,
Université de Montréal, Canada

Reviewed by:

Jinshui Zheng,
Huazhong Agricultural University,
China
Vania Aparecida Vicente Aparecida
Vicente,
Universidade Federal do Paraná, Brazil

*Correspondence:

Theo A. J. van der Lee
theo.vanderlee@wur.nl
Marnix H. Medema
marnix.medema@wur.nl

Specialty section:

This article was submitted to
Fungi and Their Interactions,
a section of the journal
Frontiers in Microbiology

Received: 20 December 2017

Accepted: 14 May 2018

Published: 05 June 2018

Citation:

Hoogendoorn K, Barra L, Waalwijk C,
Dickschat JS, van der Lee TAJ and
Medema MH (2018) Evolution and
Diversity of Biosynthetic Gene Clusters
in *Fusarium*. *Front. Microbiol.* 9:1158.
doi: 10.3389/fmicb.2018.01158

Plant pathogenic fungi in the *Fusarium* genus cause severe damage to crops, resulting in great financial losses and health hazards. Specialized metabolites synthesized by these fungi are known to play key roles in the infection process, and to provide survival advantages inside and outside the host. However, systematic studies of the evolution of specialized metabolite-coding potential across *Fusarium* have been scarce. Here, we apply a combination of bioinformatic approaches to identify biosynthetic gene clusters (BGCs) across publicly available genomes from *Fusarium*, to group them into annotated families and to study gain/loss events of BGC families throughout the history of the genus. Comparison with MIBiG reference BGCs allowed assignment of 29 gene cluster families (GCFs) to pathways responsible for the production of known compounds, while for 57 GCFs, the molecular products remain unknown. Comparative analysis of BGC repertoires using ancestral state reconstruction raised several new hypotheses on how BGCs contribute to *Fusarium* pathogenicity or host specificity, sometimes surprisingly so: for example, a gene cluster for the biosynthesis of hexadecydro-astechrome was identified in the genome of the biocontrol strain *Fusarium oxysporum* Fo47, while being absent in that of the tomato pathogen *F. oxysporum* f.sp. *lycopersici*. Several BGCs were also identified on supernumerary chromosomes; heterologous expression of genes for three terpene synthases encoded on the *Fusarium poae* supernumerary chromosome and subsequent GC/MS analysis showed that these genes are functional and encode enzymes that each are able to synthesize koraiol; this observed functional redundancy supports the hypothesis that localization of copies of BGCs on supernumerary chromosomes provides freedom for evolutionary innovations to occur, while the original function remains conserved. Altogether, this systematic overview of biosynthetic diversity in *Fusarium* paves the way for targeted natural product discovery based on automated identification of species-specific pathways as well as for connecting species ecology to the taxonomic distributions of BGCs.

Keywords: biosynthetic gene cluster, *Fusarium*, koraiol, ancestral state reconstruction (ASR), supernumerary chromosome

INTRODUCTION

The *Fusarium* genus is an extensive fungal genus of ascomycetes consisting of mostly saprotrophic soil-borne species. However, some *Fusarium* species are important plant pathogens. A survey, distributed to fungal pathologists, placed two *Fusarium* species in the top five of fungal plant pathogens based on scientific and economic importance (Dean et al., 2012). In addition, some *Fusarium* species may cause animal and human infections (Nucci and Anaissie, 2007). So while most of the species in the genus are harmless, the ones that actually show pathogenic abilities cause alarming levels of financial as well as clinical damage. Infection of wheat and barley by *Fusarium graminearum*, for example, causes fusarium head blight, FHB (Windels, 2000; McMullen and Stack, 2011). This disease eventually causes kernels to shrivel. When infected kernels are used as seeds for new generations of wheat plants, growth is severely hindered, and final production gravely decreased. Reports have estimated that financial damage from *F. graminearum* exceeds tens of millions of dollars *per annum* (Windels, 2000; Wu, 2007). Apart from the decreased kernel quality, these kernels often contain high amounts of mycotoxins. When consumed, these mycotoxins cause serious health hazards in humans and other animals, ranging from irritations to death.

A range of fungal specialized metabolites have been described in *Fusarium* (Marasas et al., 1979; Emery, 1980; Wiebe and Bjeldanes, 1981; Bacon et al., 1996; Logrieco et al., 1998; Linnemannstöns et al., 2002; Malz et al., 2005; Oide et al., 2006; Troncoso et al., 2010; Yoshida and Nakajima, 2010; Studt et al., 2012; Brock et al., 2013; Cruz et al., 2013; Kakule et al., 2013; Sørensen et al., 2013, 2014; Yin et al., 2013; Jørgensen et al., 2014; Niehaus et al., 2014; Arndt et al., 2015; Von Bargen et al., 2015; Burkhardt et al., 2016; Janevska et al., 2016, 2017; Johns et al., 2017; Reynolds et al., 2017; Wollenberg et al., 2017). Beside functions as mycotoxins, other fungal specialized metabolites play roles in host infection, interspecies competition and defense against predators. To achieve the production of these complex molecules, several enzymes cooperate in a pathway to generate the final product. Genes involved these biosynthetic pathways are often located in close proximity to each other, forming a biosynthetic gene cluster (BGC; see Keller et al., 2005). Along with genes responsible for the biosynthesis of the specialized metabolite, genes with regulatory and transport functions are usually also present in these clusters. Bioinformatic tools such as antiSMASH, CASSIS and SMURF make it possible to rapidly and automatically identify these BGCs in fungal genome sequences (Khaldi et al., 2010; Medema et al., 2011; Wolf et al., 2015). However, obtaining a systematic perspective on BGC diversity across multiple genomes has been more challenging. A few years ago, two key studies have appeared that cataloged the diversity of polyketide synthases (PKSs) and non-ribosomal peptide synthetases (NRPSs) in a number of *Fusarium* genomes (Hansen et al., 2012, 2015). Since then, several methods have been developed to compute all pairwise distances between BGCs in a dataset (Cimermanic et al., 2014; Doroghazi et al., 2014), which can be used to generate sequence similarity networks and group BGCs into gene cluster families (GCFs) based on similarities observed across the entire gene cluster. Moreover, the

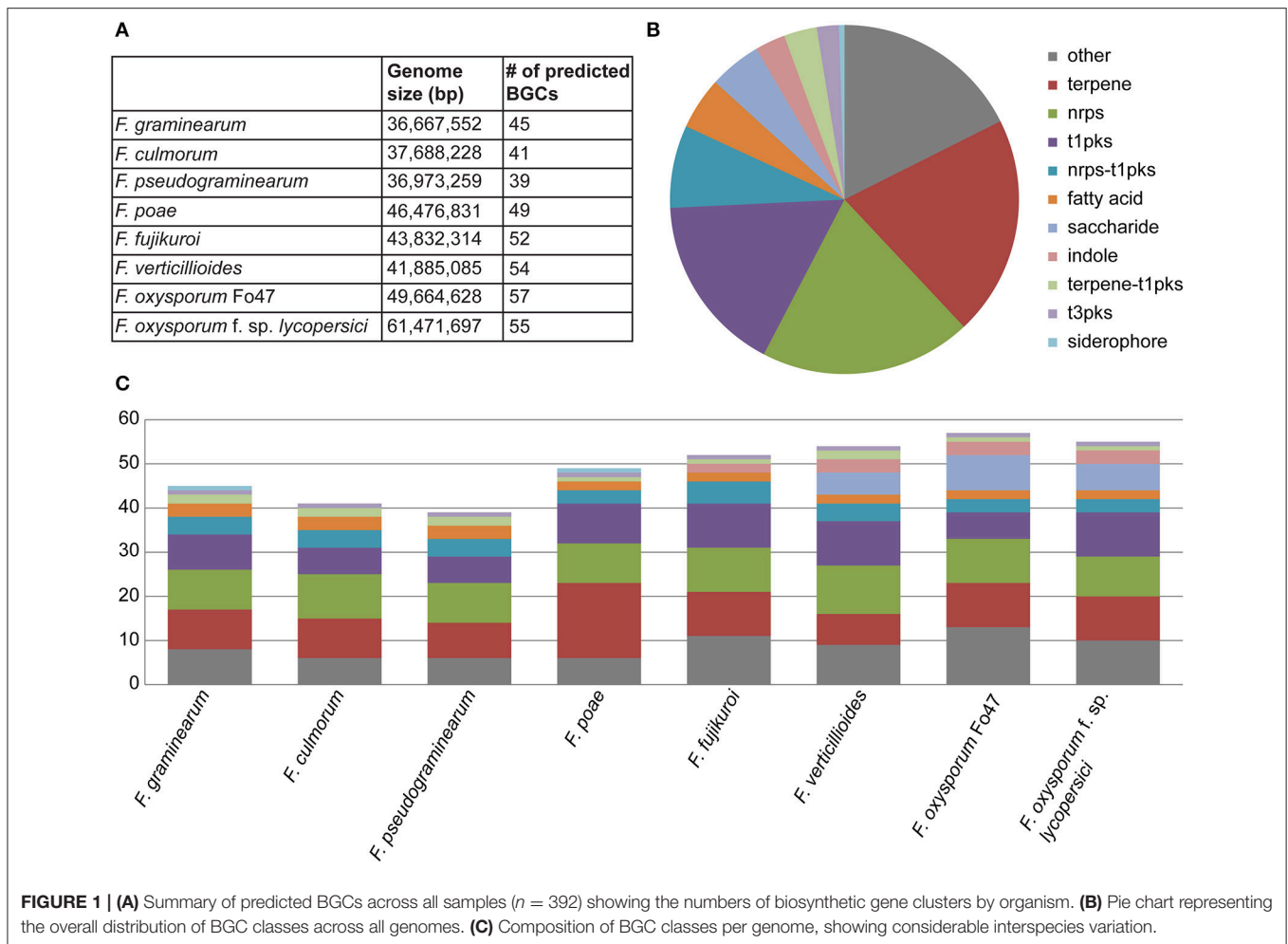
recently launched Minimum Information about a Biosynthetic Gene cluster (MIBiG) repository provides rich reference data to connect these GCFs to known products based on gene cluster homology, even with clusters from different fungal genera.

Here, we apply these techniques to provide a systematic overview of the taxonomic diversity of BGCs with known products as well as uncharacterized BGCs in eight high-quality genomes across the genus *Fusarium* (Niehaus et al., 2016). Based on this overview, we offer a reconstruction of the evolutionary history of BGC repertoires, and identify clues on ways in which BGCs evolve toward potential new functionalities. Specifically, we show that biosynthetic genes on supernumerary chromosomes are functional, and present a case in which at least three copies are found of a terpene synthase gene that is functionally redundant with the chromosomal copy; this indicates that these supernumerary chromosomes can host genes that are free to evolve toward novel functions without loss of the original function. Finally, we provide strategies to leverage the information on BGC taxonomic distribution for targeted natural product discovery. The network-guided prediction of BGC functions based on a combination of ancestral state reconstruction and transcriptomic analysis exploited in this study will deepen our understanding of the different life styles adopted by *Fusarium* species as well as pave the path toward discovery of bioactive compounds.

RESULTS AND DISCUSSION

Identification of a Diverse Set of High-Quality *Fusarium* Genomes

A comprehensive study on biosynthetic diversity critically depends on the completeness of gene content and contiguity within the genome sequences chosen for study, especially because absence of a BGC cannot be confidently established in low-quality genome sequences. For this reason, publicly available genomes were assessed for the covered gene space of the genomes and their corresponding structural annotations using BUSCO (Simão et al., 2015). Eight genomes were included in this study, which had BUSCO scores of at least 97% complete genes (Table S1; Cuomo et al., 2007; Ma et al., 2010; Al-Reedy et al., 2012; Gardiner et al., 2012; Moolhuijzen et al., 2013; Wiemann et al., 2013; Vanheule et al., 2016). Additionally, assembly statistics were evaluated; all genomes used in this study had a minimal N50 contig size of 40 kb and a minimal N50 scaffold size of 1.9 Mb (Table S1). A final indication of the completeness of the assemblies is the presence of certain gene clusters in a corresponding genome annotation; in *Fusarium* species, three gene clusters are considered to be conserved in all species of the genus (Wiemann et al., 2013). These clusters encode fusarubins, gibepyrone A and a yet unknown compound (PKS07). For all species passing the first two criteria, all three of these BGCs were identified on single contigs. The eight strains fall into two distinct clades: *Fusarium graminearum* PH-1, *Fusarium culmorum* CS7071, *Fusarium pseudograminearum* CS3096 and *Fusarium poae* fall into the *Fusarium sambucinum* species complex, while the *Fusarium verticillioides* 7600, *Fusarium*



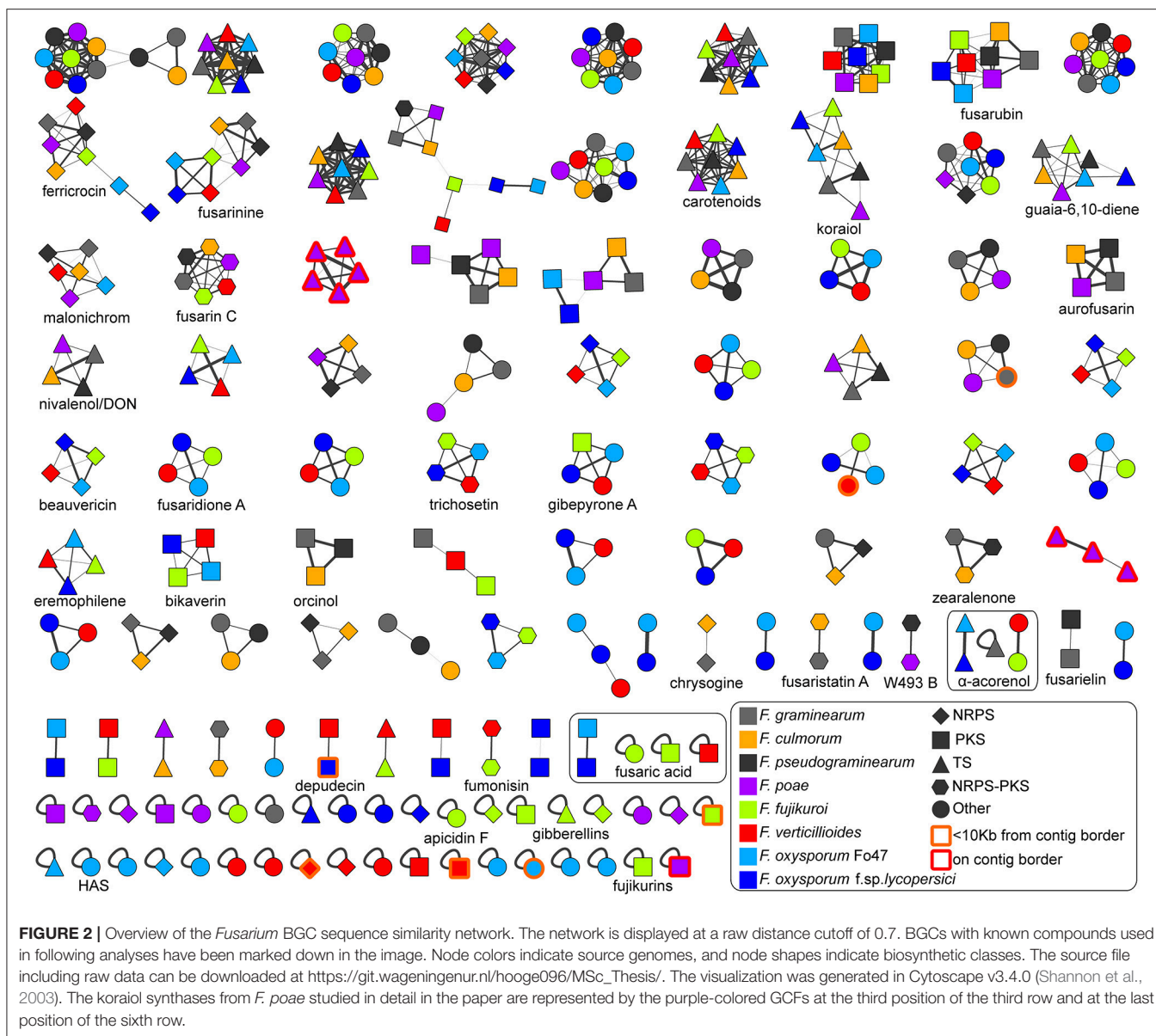
fujikuroi IMI58289, and the *Fusarium oxysporum* Fo47 and *lycopersici* 4287 strains fall into the closely related *F. fujikuroi* and *F. oxysporum* species complexes (O'Donnell et al., 2013). The two *F. oxysporum* strains included in this study were notably different: *F. oxysporum* f. sp. *lycopersici* infects tomato, while *F. oxysporum* Fo47 is a biological control strain that grows in association with tomato without causing symptoms and can protect against infection by other *Fusarium* species (Alabouvette et al., 1993).

Systematic Mapping of *Fusarium* Biosynthetic Diversity

In the eight genomes, a total of 392 biosynthetic gene clusters were found using antiSMASH, averaging 49 clusters per genome, with total numbers ranging from 39 to 57 (Figure 1A). As expected, the most prevalent classes of BGCs were those encoding PKSs, NRPSs and terpene synthases (TSs) (Figures 1B,C). The ClusterFinder algorithm also identified several putative BGCs of unknown function across the eight genomes (Cimermanic et al., 2014). Interestingly, many of these were strain-specific, reinforcing the hypothesis that they carry out specialized metabolic functions specific to certain ecological niches. While most BGCs identified were found on

the main chromosomes, several clusters were detected on the supernumerary chromosomes of *Fusarium poae*. Because of the high repeat content of these supernumerary chromosomes, resulting in a fragmentary assembly of these chromosomes, only partial sequences were obtained for these BGCs.

To identify orthology relationships between BGCs across genomes, we reconstructed a BGC sequence similarity network and visualized it using Cytoscape (Figure 2). At a raw distance threshold of 0.7 (see Methods for details on the distance metric), analysis of BGCs with known product shows that the connected components corresponded very well with GCFs for 29 natural products (Marasas et al., 1979; Emery, 1980; Wiebe and Bjeldanes, 1981; Bacon et al., 1996; Logrieco et al., 1998; Linnemannstöns et al., 2002; Malz et al., 2005; Oide et al., 2006; Troncoso et al., 2010; Yoshida and Nakajima, 2010; Studt et al., 2012; Brock et al., 2013; Cruz et al., 2013; Kakule et al., 2013; Sørensen et al., 2013, 2014; Yin et al., 2013; Jørgensen et al., 2014; Niehaus et al., 2014; Arndt et al., 2015; Von Bargen et al., 2015; Burkhardt et al., 2016; Janevska et al., 2016, 2017; Johns et al., 2017; Reynolds et al., 2017; Wollenberg et al., 2017); only the fusaric acid and α -acorenonol BGCs were split over a few separate connected components, and comparison with MIBiG reference

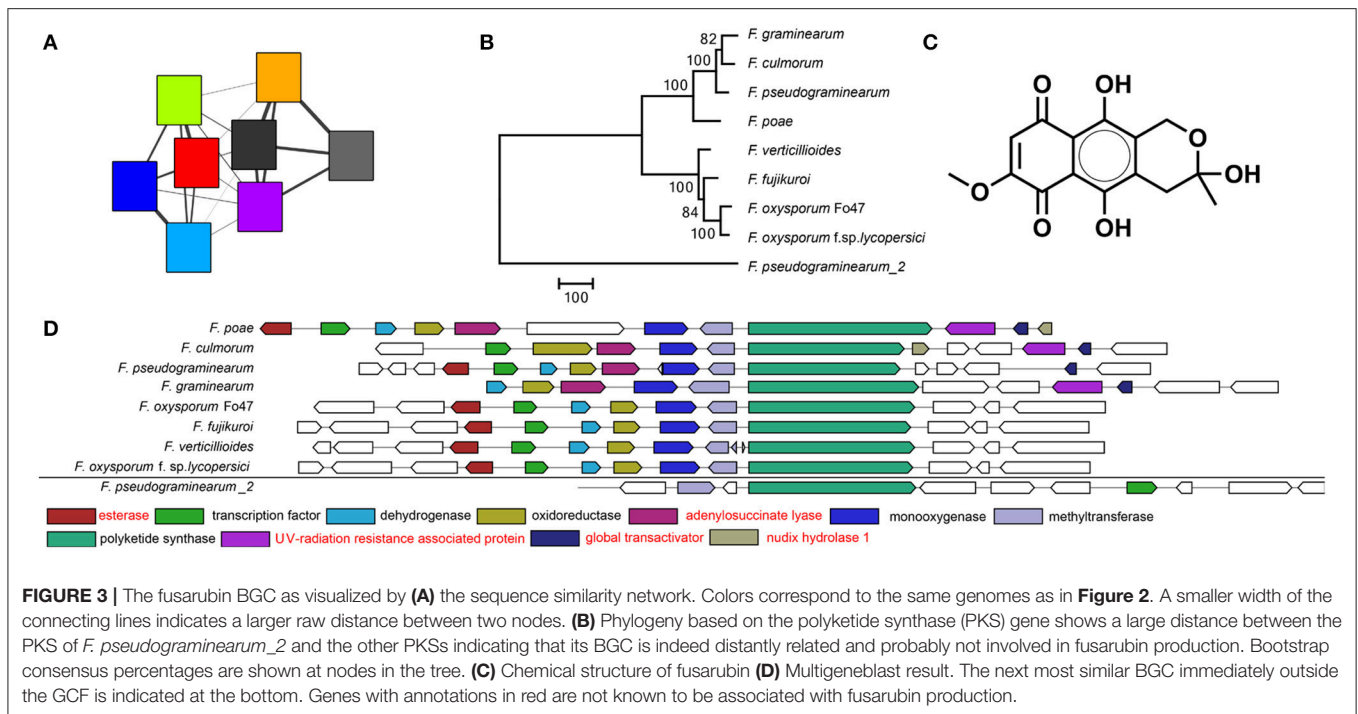


BGCs (Medema et al., 2015) allowed assignment of 30 GCFs to pathways responsible for the production of known compounds. Per GCF, this assignment was validated by manual analysis based on multiple-BGC alignments generated using MultiGeneBlast (Medema et al., 2013). In complex cases, phylogenetic analysis was performed to validate the GCF assignment (see **Figure 3** for an example). Finally, absence of BGCs was re-evaluated by manual BlastP and tBlastN searches with the scaffold biosynthesis enzymes as query, to check for cases in which misannotations, pseudogenization or gene cluster content diversification lead to clusters being missed; this *post-hoc* analysis revealed additional seemingly intact copies of koraiol synthase in *F. verticillioides* and malonichrome synthetase in *F. oxysporum* f. sp. *lycopersici*. It should also be noted that in some cases, BLAST hits were observed with closely related strains, but not with the genome

under study; e.g., a malonichrome synthetase homolog was found in *F. fujikuroi* UK99, but not in *F. fujikuroi* IMI58289.

Most of the smaller connected components in the network can be divided into two main groups that correspond to the two major clades in the species phylogeny: the ones that are present in *F. graminearum*, *F. pseudograminearum*, *F. culmorum*, *F. poae* but absent in *F. verticillioides*, *F. fujikuroi*, *F. oxysporum* f. sp. *lycopersici*, *F. oxysporum* Fo47, and vice versa. Increasing the raw distance threshold did not combine these types of clusters, indicating that there is a considerable diversity in BGC repertoires between these two clades.

In all except two cases, GCFs contained maximally one copy per species, indicating that recent whole cluster duplications within genomes were generally absent. One of these exceptions corresponds to terpene BGCs on the



supernumerary chromosomes of *F. poae*; a sequence analysis and phylogeny of the terpene synthase sequences in these BGCs showed that these represent paralogous copies of the koraol sesquiterpene synthase present on the core chromosome (Figure S1). An additional “GCF” was present that contained three small fragments of an additional single koraol synthase-encoding gene, which had not been fully assembled.

Identification of Functionally Redundant Sesquiterpene Synthases on the Supernumerary Chromosome(s) of *Fusarium Poae*

The GCFs containing the terpene synthases found on the supernumerary chromosome(s) in *F. poae* were subjected to a more detailed analysis to identify their relation to each other and potentially discover the structure of their products and their functions. All of the identified BGCs on these chromosomes contained the full sequence of a predicted sesquiterpene synthase, sharing 83% identity with the ortholog of the previously identified koraol synthase from *F. fujikuroi* on the main chromosome of *Fusarium poae* (Brock et al., 2013). Sequence alignment of all predicted terpene synthases revealed that multiple copies of the same terpene synthase are present on the supernumerary chromosomes; however, slight sequence variation between the copies was observed. Sequencing errors or assembly errors were excluded as the cause of these differences by mapping back the original sequence reads to each of the newly identified terpene synthases with a very high stringency (Length overlap $\geq 95\%$, similarity $\geq 80\%$). The

frequency with which the differing bases occurred within reads showed that the differences were genuine, and validated the antiSMASH prediction and sequence similarity clustering of the BGCs. From the six valid variants that were discovered, four were synthesized and cloned into an expression vector, which was then transformed into *Escherichia coli* BL21 (Figure S2, Table S2). Three of the four enzymes showed high activity and converted farnesyl diphosphate (FPP) into koraol as a single product (Figures S3–S6 and Table S3). The incubation with geranyl diphosphate (GPP) yielded a series of monoterpenes for these three enzymes, while geranylgeranyl diphosphate (GGPP) was converted into an unidentified diterpene. In contrast, the fourth enzyme exhibited low activity toward FPP and GPP, and did not accept GGPP. This enzyme has an additional arginine in position 260, located directly in front of the highly conserved NSE triad that is important for functionality of terpene cyclases, which may be the reason for the disturbed enzyme activity. The additional arginine residue could also be responsible for a misfolding of the protein, leading to the observed low expression levels of soluble protein. The presence of at least three functionally redundant copies of koraol synthase on the supernumerary chromosome reinforces the hypothesis that these genomic elements provide opportunities for evolution to experiment with little cost associated, as the original function remains preserved in the chromosomal copy (Vanheule et al., 2016). The absence of repeat-induced point mutation may also increase the evolutionary rate of protein-coding genes here (Galagan and Selker, 2004). Sometimes this may lead to loss of function (as perhaps in the case of the fourth enzyme above, although evolution toward an unknown function cannot be fully excluded), but it may also lead to neofunctionalization. Further

studies on a wider range of *Fusarium* species and BGCs will be required to establish the latter.

Ancestral State Reconstruction Sheds Light on the Evolution of *Fusarium* Biosynthetic Repertoires

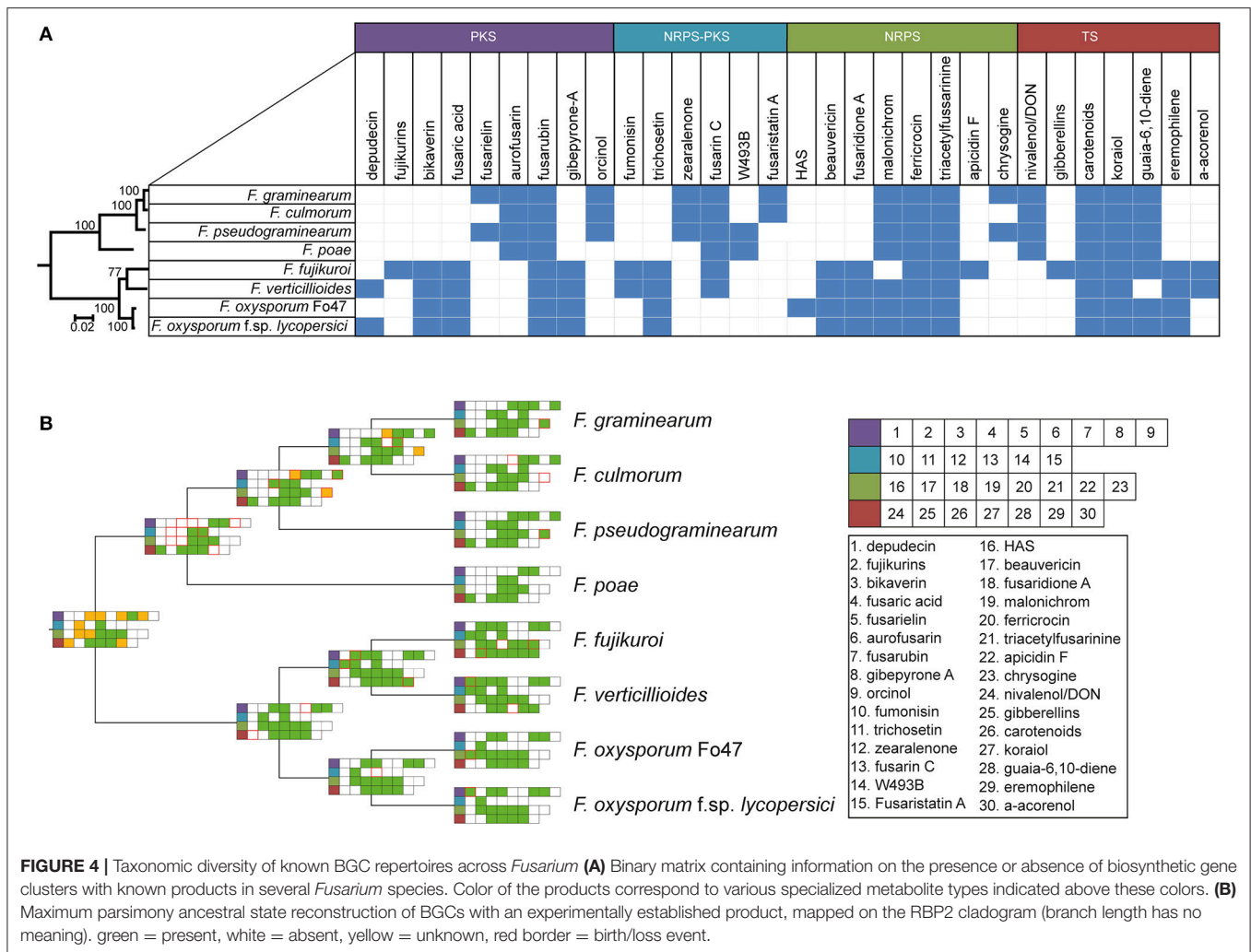
Ancestral state reconstruction, the inference of mutations, genome duplications, BGC births and deaths (Proctor et al., 2013) in the past from genomic data obtained from organisms in the present has been an essential technique to increase our understanding of the evolutionary history of a species. Reconstructing the ancestral states of GCFs gives insight into the development of BGC repertoires through time. Focusing on the BGCs for the production of known molecules, we used the GCF assignments to construct a presence/absence matrix of the 30 GCFs to which products were assigned (see details of absence/presence of specific BGCs in Table S4). Subsequently, gains and losses of clusters were inferred using a maximum parsimony approach. To construct a reference phylogeny, the RPB2 gene sequence was used, which corroborated the topology of a previously built, more extensive, phylogeny (O'Donnell et al., 2013). This allowed for reconstruction of the ancestral states of the gene clusters (Figure 4, Figure S7).

The absence/presence matrix and ancestral state reconstruction reveal a number of interesting hallmarks of the evolution of biosynthetic diversity in *Fusarium* species. For example, four GCFs are conserved throughout all genomes, whereas 22 GCFs seem limited to either the genomes of either the *F. sambucinum* species complex lineage or to the *F. fujikuroi* / *F. oxysporum* species complexes; these two groups are also the basal monophyletic groups in species phylogeny. In the case of bikaverin and aurofusarin, both of which are pigments (Malz et al., 2005; Schumacher et al., 2013), this bifurcation is clearly observed (Figure 4A). While their exact roles are not known, these pigments often play a role in protection against environmental conditions (Durán et al., 2002); the two pigments may either be functionally equivalent or provide specific advantages in the varying ecological context where the two groups of fusaria abide. Still other GCFs are limited to just one or two species, and are otherwise only known from different genera, which is suggestive of recent horizontal gene transfer. For example, the BGC for depudecin, which is known to be a virulence factor of *Alternaria brassicicola* in the infection of green cabbage (*Brassica oleracea*) (Wight et al., 2009), follows an interesting pattern: only two species, *F. oxysporum* f. sp. *lycopersici* and *F. verticillioides*, contain this cluster. Interestingly, detailed analysis of the genome of the biocontrol strain *F. oxysporum* Fo47 showed that the cluster appears there in a truncated, most likely non-functional, form. This indicates that the cluster was most likely acquired recently during evolution and partially lost again in some *F. oxysporum* strains, corroborating earlier work by Reynolds et al. (2017). Hence, it may well be (partially) responsible for the difference in lifestyle between these two strains, as *F. oxysporum* f. sp. *lycopersici* is pathogenic to tomato while the biocontrol strain *F. oxysporum* Fo47 is not.

The core enzymes of GCFs seem to follow the *Fusarium* phylogeny derived from RPB2 (see tested examples in Figure S8) or only slightly deviate from it. This indicates that these GCFs were present in the common ancestor of these species and were maintained during evolution. Our study suggests that horizontal gene transfer events occur at a relatively low frequency; still, when they occur, these events may have a large impact. Therefore, correlating these events to species phenotypes may open up new ways for targeted natural product discovery. For example, there are 17 GCFs of unknown function that occur only in one of the two *F. oxysporum* strains; these intriguing candidates may explain adaptations to a specific ecological niche or lifestyle.

BGC Networking Allows Rapid Screening for Bioactive Compound-Synthesizing Capacity in *Fusarium* Genomes

Fungal BGCs have a wide variety of ecological roles, most of which are not yet fully understood. From the 30 BGCs with known products, only a few are known to be directly involved in pathogenicity. Several mycotoxins are produced, and some of these mycotoxins are involved in pathogenicity toward the host. Depudecin, fumonisin, gibepyrone and fusaric acid have been shown to be virulence factors (Wight et al., 2009; Cruz et al., 2013; Janevska et al., 2016; Studt et al., 2016). Fumonisin, a strong virulence factor for maize, produced by *F. verticillioides*, is not directly related to disease in rice when infected by *F. fujikuroi* (Cruz et al., 2013). This shows that much about its role in virulence is still unknown. By applying sequence similarity network analysis to genome sequences, a quick overview can be acquired concerning the potential for the production of toxins and other important bioactive compounds in a species. For example, *F. oxysporum* Fo47, which is widely used as a biological control strain to prevent other *Fusaria* from infecting a plant, still contains a BGC similar to the one that encodes production of beauvericin in *Aspergillus fumigatus* (Figure S9A). This is a worrisome discovery, since beauvericin production has been found in other *F. oxysporum* strains, and is a mycotoxin capable of triggering apoptosis in several human cell lines (Logrieco et al., 1998). While this compound is not necessarily synthesized by *F. oxysporum* Fo47 under normal conditions, it is possible that through alteration of gene expression through either mutations or changing ecological conditions, production of beauvericin is resumed. We also (for the first time, to our knowledge) detected a hexadecahydro-astechrome (HAS) BGC in this *F. oxysporum* strain, showing $\pm 78\%$ sequence identity at the amino acid level with the NRPS identified in *Aspergillus fumigatus* (Figure S9B). Overexpression of this compound is associated with increased mortality in mice when infected by *A. fumigatus* (Yin et al., 2013). While it is thought to only be produced by a small group of opportunistic human pathogenic fungi, the genomic information indicates that a homologous BGC is actually present in *F. oxysporum* Fo47. This should be reason for concern regarding the use of this strain for biological control purposes. Many more of these cases might arise in the future, and preliminary screening of genomes for toxin production capabilities using BGC diversity analyses such as provided here



can assist with making decisions concerning widespread use of an organism for either agricultural, biotechnological or medical purposes. Alternatively, they can guide the identification of natural mutants (or the engineering of mutants) that lack such unfavorable traits.

Network Guided Re-analysis of RNA-Seq Data Aids Identification of Specialized Metabolites Putatively Involved in Pathogenesis

As shown for the cases of HAS and beauvericin, BGC sequence similarity networks facilitate the identification of BGCs for known molecules in well-studied strains in which these had not been reported before. This also means that previous RNA-seq data can be reevaluated based on this information; moreover, any information from function arising from this may be translated across species. To demonstrate this, an RNA-seq dataset from *F. graminearum* was downloaded from NCBI (Boedi et al., 2016). This dataset was originally used to identify genes involved in pathogenicity, and contained reads

extracted from *F. graminearum* growing on living wheat (pth), dead wheat (spr) and a secondary metabolite inducing medium (smi). Additionally, a dataset containing reads from a wild type strain of *F. graminearum* growing on rich medium (wt), as well as a mutant thought to induce trichothecene production (tri) grown on the same rich medium (Zhao et al., 2013) were added to produce the final dataset. Based on analysis of these data (see Methods), we identified genes that were differentially expressed at the statistical threshold of $\alpha = 0.01$ and were present in the list of anchor genes from the BGCs found in *F. graminearum* (Figure S10). A heatmap of the expression values of the remaining genes was generated and clustered on both axes to reveal patterns of interest (Figure S11). A particular group of interest consisted of the set of anchor genes predominantly active during pathogenic growth: three BGCs from *Fusarium graminearum* producing yet unknown products, along with the BGCs encoding the biosynthetic enzymes for malonichrome and fusaristatin A, as also noted in the previous publication (Boedi et al., 2016). While malonichrome has already been described to be of some importance during infection by regulation of the iron homeostasis in the fungus, the *in vivo* function of

fusaristatin A is poorly studied. Interestingly, the BGC for fusaristatin A is also present in *Fusarium culmorum*, prompting further investigation of its possible roles in pathogenicity in this species. The three BGCs producing unknown compounds each have homologs in *F. pseudograminearum* and/or *F. culmorum* according to the network analysis, and have previously been found to be under control by the same transcription factor; *FGPI* (Jonkers et al., 2012; Oide et al., 2015). Mutation studies showed that the *FGPI* gene is essential for virulence, and that $\Delta fgp1$ mutants are no longer able to infect their hosts. It is likely that the BGCs controlled by *FGPI* play a role in pathogenicity, and therefore are interesting to study in more detail.

CONCLUSIONS

The present work showcases how computational analysis of BGC repertoires using sequence similarity networks and GCF reconstruction allows a systematic analysis of biosynthetic diversity in a genus with a complex specialized metabolome. Moreover, the combination of GCF taxonomic distributions, species phenotypes and ancestral state reconstruction analysis potentially allows for intelligent prioritization strategies to characterize the function of unknown and novel GCFs through synthesis and/or heterologous expression. In this way, the metabolic capacity of a genus can be systematically explored and directly linked to the specific ecologies of its constituent species and strains. This is highlighted by the case of the two *F. oxysporum* strains, one being a pathogen and one a biocontrol strain on tomato, which differ in the presence/absence of depudecin and HAS BGCs, as well as several BGCs of unknown function, which may (partially) explain their large phenotypic differences. The next step will be to extend this kind of analysis to other fungal genera and larger sets of genomes, accelerated through the latest bioinformatic advances. Moreover, the hypotheses generated by these analyses offer promising new avenues for experimental research. Such future studies may lead to new ways to aid in the combat against pathogens, reducing health risks and improving crop yields.

MATERIALS AND METHODS

Genome Sequences

The genomic dataset consisted of eight *Fusarium* genomes (Table S1). From the GenBank file, FASTA files as well as GFF3 files with gene features were created using custom-made Python scripts to be used as input for BUSCO. To assess genome completeness, BUSCO v3.0.1 was used along with its sordariomyceta_odb9 database. Default parameters were used for all input files, and no additional parameters were used.

Phylogeny Construction

The RBP2 gene was used as the sole gene to base the phylogeny of this *Fusarium* subset on. Using MEGA6, a bootstrapped neighbor-joining tree was constructed

using the multiple sequence alignment of the BLAST results from the RBP2 gene in *F. oxysporum* f. sp. *lycopersici* (chromosome 7, location: 3840192-3844693, locus: FOXG_10639). Default settings were used, with 1,000 bootstrap replicates.

Computational BGC Identification

For identification of gene clusters, antiSMASH was used. A development version was used that utilizes ClusterFinder to trim borders of BGCs, in order to enhance sequence similarity networking. Official releases are available at <http://antismash.secondarymetabolites.org> and development versions are available at <https://bitbucket.org/antismash/antismash>.

Sequence Similarity Network and BGC Family Reconstruction

AntiSMASH output was used as input to generate a sequence similarity network according to previously published methods (Cimermanic et al., 2014), with a 0.6 weight for the domain duplication index (including sequence identity, according to the same methods as by Cimermanic et al. (2014) but using hmalign instead of Muscle for multiple-sequence alignment), a 0.2 weight for the Jaccard index and a 0.2 weight for the Goodman-Kruskal gamma index. Singletons were included to ensure that unique BGCs would not be left out in the final network. In order to visualize the networks, Cytoscape v.3.4.0 was used (Shannon et al., 2003). Edge width was set to decrease with an increasing raw distance, to better visualize internodal relations.

Multigeneblast

To visualize similarity between BGCs in different species, the MultiGeneBlast algorithm was used (Medema et al., 2013). AntiSMASH cluster output files were modified slightly to contain the species name and cluster number in the version/accession feature to make results easier to interpret. The MultiGeneBlast database used for the analysis comprised all predicted clusters across all samples. Predicted clusters were blasted one by one against this database to acquire visual similarity results that could be used to validate GCFs assigned using the sequence similarity network. Default settings were used when running MultiGeneBlast.

Ancestral State Reconstruction

Ancestral state reconstruction of BGCs with known compounds was conducted using Mesquite (Maddison and Maddison, 2016). Since the dataset was small, and maximum likelihood trees can be unreliable and biased in certain cases, a maximum parsimony approach was taken while reconstructing ancestral states (Kolaczkowski and Thornton, 2004). With the use of a phylogenetic tree based on the RBP2 gene, emergence, disappearance and eventual duplication of gene clusters can be traced back through evolutionary history. A binary matrix containing species and

gene clusters with known compounds was created to be used in Mesquite.

Transcriptome Analysis

F. graminearum RNA-seq reads were mapped back onto the reference genome with the CLC genomics workbench. Preprocessing of reads was also done within CLC genomics workbench, where the last 5% of lowest quality bases were trimmed. If read lengths were shorter than 15 nucleotides after trimming, they were discarded entirely. Reads were mapped using two parameters: %identity and %length overlap. To create the differential expression heatmap, cutoffs used for these were 60 and 80% respectively. Cufflinks was then used to assemble transcripts from the mapped reads to determine expression levels per gene (Trapnell et al., 2010). A final transcriptome was assembled using cuffmerge, which was slightly altered after creation to allow the use of locus tags in following steps instead of the arbitrary IDs that cufflinks assigns to the assembled transcripts. To ensure that replicates of the same sample were consistent with each other, an MDS plot was created (Figure S12). Replicates were expected to end up in close proximity to each other in the graph, while samples were expected to have a larger spacing between them. The plot shows these exact patterns, which leads to the conclusion that the replicates within a sample are consistent with each other, and can be used for a differential expression analysis. Using the original mapped reads with cuffdiff, a differential expression analysis was conducted, which was visualized in R using cummRbund.

Script Availability

All scripts that were created during this project were developed in Python 2.7. Scripts, along with their documentation and data files have been made publicly available on http://git.wageningenur.nl/hooge096/MSc_Thesis.git.

Gene Cloning

All four targeted synthetic genes were amplified from a pUCIDT-AMP plasmid using Q5 polymerase and the primer pairs LB019f_538_48_FPOA_13441 and LB019r_538_48_FPOA_13441 (Table S2). The initial PCR product was elongated by PCR using Q5 polymerase and primer pairs LB020f_538_48_FPOA_13441 and LB020r_538_48_FPOA_13441 to obtain the desired gene carrying homology arms for homologous recombination in *Saccharomyces cerevisiae*. Transformation of *S. cerevisiae* with the elongated PCR products and linearized (XhoI and PvuII digestion) vector pYE-Express was conducted by using the LiOAc/SS carrier DNA protocol. The transformed cells were inoculated on SM-URA agar plates for 3 days at 28°C and plasmid DNA was isolated from grown cells using the kit Zymoprep Yeast Plasmid Miniprep II (Zymo Research). The obtained plasmids were transferred into *E. coli* BL 21 cells by electroporation. The transformed cells were plated on 2YT agar (tryptone 16 g, yeast extract 10 g, NaCl 5 g, agar 20 g, water 1 L, pH 7.2) containing kanamycin (50 mg/L) and grown at 37°C overnight. Single colony cells were picked

and inoculated in 2YT liquid medium (10 mL) and grown for 4 h at 37 °C, followed by plasmid isolation using the kit Plasmid Miniprep kit (Promega). The plasmid containing the targeted genes were analyzed by sequencing with t7 promoter primers.

Gene Expression and Protein Purification

Transformed *E. coli* BL21 mutants were precultured overnight in 2YT liquid medium containing kanamycin (50 mg/L) at 37°C and used for inoculation of 250 mL liquid 2YT medium with kanamycin (50 mg/L). Cells were grown to an OD₆₀₀ = 0.5 at 37°C and 160 rpm. After cooling to 18°C, IPTG (0.4 mM) was added and the cultures were incubated over night at 18°C and 160 rpm. After centrifugation at 4°C and 3,600 rpm for 30 min, the supernatant was discarded and the cell pellet was resuspended in 2 mL binding buffer (20 mM Na₂HPO₄, 0.5 M NaCl, 20 mM imidazole, 1 mM MgCl₂, pH 7.4) and cells were disrupted by ultra-sonication on ice for 5 × 30 s. The cell debris was removed by centrifugation (2 × 10 min) at 4°C and 11,000 rpm to yield the soluble enzyme fractions. The protein was purified by Ni²⁺-NTA affinity chromatography with Ni²⁺-NTA superflow (Qiagen) with binding buffer (4 × 1 mL; 20 mM Na₂HPO₄, 0.5 M NaCl, 20 mM imidazole, 1 mM MgCl₂, pH 7.4) and elution buffer (3 × 1 mL; 20 mM Na₂HPO₄, 0.5 M NaCl, 500 mM imidazole, 1 mM MgCl₂, pH 7.4). The obtained fractions were collected and analyzed by SDS-PAGE (Figure S2). Fractions containing protein of the correct size (elution fraction 1 and 2) were pooled and used for enzyme reactions.

Enzyme Reactions and Product Analysis

The obtained pooled elution fractions containing proteins of the expected size were aliquoted and mixed with freshly prepared solutions of GPP, FPP and GGPP (0.2 mL, ca. 1 mg/mL in water) and incubated at 28°C for 14 h. The reaction mixture was extracted with hexane (0.2 mL) and the organic layer was separated, dried with MgSO₄ and directly analyzed by GC/MS. The total ion chromatograms are shown in Figures S3–S5.

GC/MS Analysis

The crude extracts were analyzed by use of an Agilent HP7890B gas chromatograph, fitted with a HP-5MS silica capillary column (30 m, 0.25 mm i. d., 0.50 μm film), connected to a HP5977A mass detector. The GC-MS conditions were as follows: (1) inlet pressure: 77.1 kPa, He flow 23.3 mL min⁻¹; (2) injection volume: 1 μL; (3) injection mode: splitless, valve time 60 s; (4) oven temperature ramp: 5 min at 50°C increasing at 5°C min⁻¹ to 320°C; (5) carrier gas He at 1 mL min⁻¹; (6) transfer line: 250°C; (7) electron energy: 70 eV. Retention indices (*I*) were determined from a homologous series of *n*-alkanes (C₈-C₄₀).

AUTHOR CONTRIBUTIONS

MHM and TL conceived the project. KH, LB, CW, JSD, TL, and MHM participated in designing the experiments and analyses.

KH carried out bioinformatic analysis, under supervision of MHM and TL. LB performed cloning and chemical analysis, under supervision of JSD. KH wrote the first draft of the paper with significant input from LB and JSD. All authors contributed to data interpretation and editing of the paper.

ACKNOWLEDGMENTS

We thank Marley Yeong for aid in reconstructing the sequence similarity networks, and Dick de Ridder for valuable comments and suggestions. JSD is supported

by the Deutsche Forschungsgemeinschaft DFG by grant DI1536/9-1. TL and CW are supported by the H2020 project MycoKey (GA67878). MHM is supported by VENI grant 863.15.002 from The Netherlands Organization for Scientific Research (NWO).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2018.01158/full#supplementary-material>

REFERENCES

- Alabouvette, C., Lemanceau, P., and Steinberg, C. (1993). Recent advances in the biological control of *Fusarium* wilts. *Pestic. Sci.* 37, 365–373. doi: 10.1002/ps.2780370409
- Al-Reedy, R. M., Malireddy, R., Dillman, C. B., and Kennell, J. C. (2012). Comparative analysis of *Fusarium* mitochondrial genomes reveals a highly variable region that encodes an exceptionally large open reading frame. *Fungal Genet. Biol.* 49, 2–14. doi: 10.1016/j.fgb.2011.11.008
- Arndt, B., Studt, L., Wiemann, P., Osmanov, H., Kleigrewe, K., Köhler, J., et al. (2015). Genetic engineering, high resolution mass spectrometry and nuclear magnetic resonance spectroscopy elucidate the bikaverin biosynthetic pathway in *Fusarium fujikuroi*. *Fungal Genet. Biol.* 84, 26–36. doi: 10.1016/j.fgb.2015.09.006
- Bacon, C. W., Porter, J. K., Norred, W. P., and Leslie, J. F. (1996). Production of fusaric acid by *Fusarium* species. *Appl. Environ. Microbiol.* 62, 4039–4043.
- Boedi, S., Berger, H., Sieber, C., Münsterkötter, M., Maloku, I., Warth, B., et al. (2016). Comparison of *Fusarium graminearum* transcriptomes on living or dead wheat differentiates substrate-responsive and defense-responsive genes. *Front. Microbiol.* 7:1113. doi: 10.3389/fmicb.2016.01113
- Brock, N. L., Huss, K., Tudzynski, B., and Dickschat, J. S. (2013). Genetic dissection of sesquiterpene biosynthesis by *Fusarium fujikuroi*. *ChemBioChem* 14, 311–315. doi: 10.1002/cbic.201200695
- Burkhardt, I., Siemon, T., Henrot, M., Studt, L., Rösler, S., Tudzynski, B., et al. (2016). Mechanistic characterisation of two sesquiterpene cyclases from the plant pathogenic fungus *Fusarium fujikuroi*. *Angew. Chem. Int. Ed.* 55, 8748–8751. doi: 10.1002/anie.201603782
- Cimermancic, P., Medema, M., Claesen, J., Kurita, K., Wieland Brown, L., Mavrommatis, K., et al. (2014). Insights into secondary metabolism from a global analysis of prokaryotic biosynthetic gene clusters. *Cell* 158, 412–421. doi: 10.1016/j.cell.2014.06.034
- Cruz, A., Marin, P., González-Jaén, M. T., Aguilar, K. G. I., and Cumagun, C. J. R. (2013). Phylogenetic analysis, fumonisin production and pathogenicity of *Fusarium fujikuroi* strains isolated from rice in the Philippines. *J. Sci. Food Agric.* 93, 3032–3039. doi: 10.1002/jsfa.6136
- Cuomo, C. A., Güldener, U., Xu, J.-R., Trail, F., Turgeon, B. G., Di Pietro, A., et al. (2007). The *Fusarium graminearum* genome reveals a link between localized polymorphism and pathogen specialization. *Science* 317, 1400–1402. doi: 10.1126/science.1143708
- Dean, R., Van Kan, J. A. L., Pretorius, Z. A., Hammond-Kosack, K. E., Di Pietro, A., Spanu, P. D., et al. (2012). The Top 10 fungal pathogens in molecular plant pathology. *Mol. Plant Pathol.* 13, 414–430. doi: 10.1111/j.1364-3703.2011.00783.x
- Doroghazi, J. R., Albright, J. C., Goering, A. W., Ju, K.-S., Haines, R. R., Tchalukov, K. A., et al. (2014). A roadmap for natural product discovery based on large-scale genomics and metabolomics. *Nat. Chem. Biol.* 10, 963–968. doi: 10.1038/nchembio.1659
- Durán, N., Teixeira, M. F. S., De Conti, R., and Esposito, E. (2002). Ecological-friendly pigments from fungi. *Crit. Rev. Food Sci. Nutr.* 42, 53–66. doi: 10.1080/10408690290825457
- Emery, T. (1980). Malonichrome, a new iron chelate from *Fusarium roseum*. *Biochim. Biophys. Acta* 629, 382–390. doi: 10.1016/0304-4165(80)90110-5
- Galagan, J. E., and Selker, E. U. (2004). RIP: the evolutionary cost of genome defense. *Trends Genet.* 20, 417–423. doi: 10.1016/j.tig.2004.07.007
- Gardiner, D. M., McDonald, M. C., Covarelli, L., Solomon, P. S., Rusu, A. G., Marshall, M., et al. (2012). Comparative pathogenomics reveals horizontally acquired novel virulence genes in fungi infecting cereal hosts. *PLoS Pathog.* 8:e1002952. doi: 10.1371/journal.ppat.1002952
- Hansen, F. T., Gardiner, D. M., Lysøe, E., Fuertes, P. R., Tudzynski, B., Wiemann, P., et al. (2015). An update to polyketide synthase and non-ribosomal synthetase genes and nomenclature in *Fusarium*. *Fungal Genet. Biol.* 75, 20–29. doi: 10.1016/j.fgb.2014.12.004
- Hansen, F. T., Sørensen, J. L., Giese, H., Sondergaard, T. E., and Frandsen, R. J. N. (2012). Quick guide to polyketide synthase and nonribosomal synthetase genes in *Fusarium*. *Int. J. Food Microbiol.* 155, 128–136. doi: 10.1016/j.ijfoodmicro.2012.01.018
- Janevska, S., Arndt, B., Baumann, L., Apken, L. H., Marques, L. M. M., Humpf, H. U., et al. (2017). Establishment of the inducible Tet-on system for the activation of the silent trichostin gene cluster in *Fusarium fujikuroi*. *Toxins (Basel)* 9:126. doi: 10.3390/toxins9040126
- Janevska, S., Arndt, B., Niehaus, E. M., Burkhardt, I., Rösler, S. M., Brock, N. L., et al. (2016). Gibepyrone biosynthesis in the rice pathogen *Fusarium fujikuroi* is facilitated by a small polyketide synthase gene cluster. *J. Biol. Chem.* 291, 27403–27420. doi: 10.1074/jbc.M116.753053
- Johns, A., Scharf, D. H., Gsaller, F., Schmidt, H., Heinekamp, T., Straßburger, M., et al. (2017). A nonredundant phosphopantetheinyl transferase, PptA, is a novel antifungal target that directs secondary metabolite, siderophore, and lysine biosynthesis in *Aspergillus fumigatus* and is critical for pathogenicity. *MBio* 8, e01504-16. doi: 10.1128/mBio.01504-16
- Jonkers, W., Dong, Y., Broz, K., and Kistler, H. C. (2012). The Wor1-like protein Fgp1 regulates pathogenicity, toxin synthesis and reproduction in the phytopathogenic fungus *Fusarium graminearum*. *PLoS Pathog.* 8:e1002724. doi: 10.1371/journal.ppat.1002724
- Jørgensen, S. H., Frandsen, R. J. N., Nielsen, K. F., Lysøe, E., Sondergaard, T. E., Wimmer, R., et al. (2014). *Fusarium graminearum* PKS14 is involved in orsellinic acid and orcinol synthesis. *Fungal Genet. Biol.* 70, 24–31. doi: 10.1016/j.fgb.2014.06.008
- Kakule, T. B., Sardar, D., Lin, Z., and Schmidt, E. W. (2013). Two related pyrrolidinedione synthetase loci in *Fusarium heterosporum* ATCC 74349 produce divergent metabolites. *ACS Chem. Biol.* 8, 1549–1557. doi: 10.1021/cb400159f
- Keller, N. P., Turner, G., and Bennett, J. W. (2005). Fungal secondary metabolism — from biochemistry to genomics. *Nat. Rev. Microbiol.* 3, 937–947. doi: 10.1038/nrmicro1286
- Khalidi, N., Seifuddin, F. T., Turner, G., Haft, D., Nierman, W. C., Wolfe, K. H., et al. (2010). SMURF: genomic mapping of fungal secondary metabolite clusters. *Fungal Genet. Biol.* 47, 736–741. doi: 10.1016/j.fgb.2010.06.003
- Kolaczowski, B., and Thornton, J. W. (2004). Performance of maximum parsimony and likelihood phylogenetics when evolution is heterogeneous. *Nature* 431, 980–984. doi: 10.1038/nature02917

- Linnemannstöns, P., Prado, M., Fernández-Martín, R., Tudzynski, B., and Avalos, J. (2002). A carotenoid biosynthesis gene cluster in *Fusarium fujikuroi*: the genes *carB* and *carRA*. *Mol. Genet. Genomics* 267, 593–602. doi: 10.1007/s00438-002-0690-5
- Logrieco, A., Moretti, A., Castella, G., Kostecki, M., Golinski, P., Ritieni, A., et al. (1998). Beauvericin production by *Fusarium* species. *Appl. Environ. Microbiol.* 64, 3084–3088.
- Ma, L.-J., van der Does, H. C., Borkovich, K. A., Coleman, J. J., Daboussi, M.-J., Di Pietro, A., et al. (2010). Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*. *Nature* 464, 367–373. doi: 10.1038/nature08850
- Maddison, W. P., and Maddison, D. R. (2016). *Mesquite: A Modular System for Evolutionary Analysis*. Available online at: <http://mesquiteproject.org>.
- Malz, S., Grell, M. N., Thrane, C., Maier, F. J., Rosager, P., Felk, A., et al. (2005). Identification of a gene cluster responsible for the biosynthesis of aurofusarin in the *Fusarium graminearum* species complex. *Fungal Genet. Biol.* 42, 420–433. doi: 10.1016/j.fgb.2005.01.010
- Marasas, W. F. O., van Rensburg, S. J., and Mirocha, C. J. (1979). Incidence of *Fusarium* Species and the mycotoxins, deoxynivalenol and zearalenone, in corn produced in esophageal cancer areas in Transkei. *J. Agric. Food Chem.* 27, 1108–1112. doi: 10.1021/jf60225a013
- McMullen, M., and Stack, R. (2011). *Fusarium* head blight (scab) of small grains. *NDSU Ext. Serv.* 804:6. doi: 10.1094/PHI-1-2003-0612-01
- Medema, M. H., Blin, K., Cimermancic, P., de Jager, V., Zakrzewski, P., Fischbach, M. A., et al. (2011). antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Res.* 39, W339–W346. doi: 10.1093/nar/gkr466
- Medema, M. H., Kottmann, R., Yilmaz, P., Cummings, M., Biggins, J. B., Blin, K., et al. (2015). Minimum Information about a Biosynthetic Gene cluster. *Nat. Chem. Biol.* 11, 625–631. doi: 10.1038/nchembio.1890
- Medema, M. H., Takano, E., and Breitling, R. (2013). Detecting sequence homology at the gene cluster level with MultiGeneBlast. *Mol. Biol. Evol.* 30, 1218–1223. doi: 10.1093/molbev/mst025
- Moolhuijzen, P. M., Manners, J. M., Wilcox, S. A., Bellgard, M. I., and Gardiner, D. M. (2013). Genome sequences of six wheat-infecting *Fusarium* species isolates. *Genome Announc.* 1:e00670-13. doi: 10.1128/genomeA.00670-13
- Niehaus, E. M., Janevska, S., Von Bargen, K. W., Sieber, C. M. K., Harrer, H., Humpf, H. U., et al. (2014). Apicidin F: characterization and genetic manipulation of a new secondary metabolite gene cluster in the rice pathogen *Fusarium fujikuroi*. *PLoS ONE* 9:e103336. doi: 10.1371/journal.pone.0103336
- Niehaus, E.-M., Münsterkötter, M., Proctor, R. H., Brown, D. W., Sharon, A., Idan, Y., et al. (2016). Comparative “omics” of the *Fusarium fujikuroi* species complex highlights differences in genetic potential and metabolite synthesis. *Genome Biol. Evol.* 8, 3574–3599. doi: 10.1093/gbe/evw259
- Nucci, M., and Anaissie, E. (2007). *Fusarium* infections in immunocompromised patients. *Clin. Microbiol. Rev.* 20, 695–704. doi: 10.1128/CMR.00014-07
- O'Donnell, K., Rooney, A. P., Proctor, R. H., Brown, D. W., McCormick, S. P., Ward, T. J., 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. doi: 10.1016/j.fgb.2012.12.004
- Oide, S., Berthiller, F., Wiesenberger, G., Adam, G., and Turgeon, B. G. (2015). Individual and combined roles of malonichrome, ferricrocin, and TAFc siderophores in *Fusarium graminearum* pathogenic and sexual development. *Front. Microbiol.* 6:759. doi: 10.3389/fmicb.2014.00759
- Oide, S., Moeder, W., Krasnoff, S., Gibson, D., Haas, H., Yoshioka, K., et al. (2006). NPS6, encoding a nonribosomal peptide synthetase involved in siderophore-mediated iron metabolism, is a conserved virulence determinant of plant pathogenic ascomycetes. *Plant Cell* 18, 2836–2853. doi: 10.1105/tpc.106.045633
- Proctor, R. H., Van Hove, F., Susca, A., Stea, G., Busman, M., van der Lee, T., 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. doi: 10.1111/mmi.12362
- Reynolds, H. T., Slot, J. C., Divon, H. H., Lysøe, E., Proctor, R. H., and Brown, D. W. (2017). Differential retention of gene functions in a secondary metabolite cluster. *Mol. Biol. Evol.* 34, 2002–2015. doi: 10.1093/molbev/msx145
- Schumacher, J., Gautier, A., Morgant, G., Studt, L., Ducrot, P. H., Le Pêcheur, P., et al. (2013). A functional bikaverin biosynthesis gene cluster in rare strains of *Botrytis cinerea* is positively controlled by VELVET. *PLoS ONE* 8:e53729. doi: 10.1371/journal.pone.0053729
- Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504. doi: 10.1101/gr.1239303
- Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., and Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31, 3210–3212. doi: 10.1093/bioinformatics/btv351
- Sørensen, J. L., Akk, E., Thrane, U., Giese, H., and Sondergaard, T. E. (2013). Production of fusarielins by *Fusarium*. *Int. J. Food Microbiol.* 160, 206–211. doi: 10.1016/j.ijfoodmicro.2012.10.016
- Sørensen, J. L., Sondergaard, T. E., Covarelli, L., Fuertes, P. R., Hansen, F. T., Frandsen, R. J. N., et al. (2014). Identification of the biosynthetic gene clusters for the lipopeptides fusaristatin A and W493 B in *Fusarium graminearum* and *F. pseudograminearum*. *J. Nat. Prod.* 77, 2619–2625. doi: 10.1021/np500436r
- Studt, L., Janevska, S., Niehaus, E.-M., Burkhardt, I., Arndt, B., Sieber, C. M. K., 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. doi: 10.1111/1462-2920.13150
- Studt, L., Wiemann, P., Kleigrew, K., Humpf, H. U., and Tudzynski, B. (2012). Biosynthesis of fusarubins accounts for pigmentation of *Fusarium fujikuroi* perithecia. *Appl. Environ. Microbiol.* 78, 4468–4480. doi: 10.1128/AEM.00823-12
- Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M. J., et al. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* 28, 511–515. doi: 10.1038/nbt.1621
- Troncoso, C., González, X., Bömke, C., Tudzynski, B., Gong, F., Hedden, P., et al. (2010). Gibberellin biosynthesis and gibberellin oxidase activities in *Fusarium sacchari*, *Fusarium konzum* and *Fusarium subglutinans* strains. *Phytochemistry* 71, 1322–1331. doi: 10.1016/j.phytochem.2010.05.006
- Vanheule, A., Audenaert, K., Warris, S., van de Geest, H., Schijlen, E., Höfte, M., et al. (2016). Living apart together: crosstalk between the core and supernumerary genomes in a fungal plant pathogen. *BMC Genomics* 17:670. doi: 10.1186/s12864-016-2941-6
- Von Bargen, K. W., Niehaus, E. M., Krug, I., Bergander, K., Würthwein, E. U., Tudzynski, B., 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. doi: 10.1021/np5008137
- Wiebe, L. A., and Bjeldanes, L. F. (1981). Fusarin C, a mutagen from *Fusarium moniliforme* grown on corn. *J. Food Sci.* 46, 1424–1426. doi: 10.1111/j.1365-2621.1981.tb04189.x
- Wiemann, P., Sieber, C. M. K., von Bargen, K. W., Studt, L., Niehaus, E.-M., Espino, J. J., 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. doi: 10.1371/journal.ppat.1003475
- Wight, W. D., Kim, K.-H., Lawrence, C. B., and Walton, J. D. (2009). Biosynthesis and role of virulence of the histone deacetylase inhibitor depudecin from *Alternaria brassicicola*. *Mol. Plant Microbe Interact.* 22, 1258–1267. doi: 10.1094/MPMI-22-10-1258
- Windels, C. E. (2000). Economic and social impacts of *Fusarium* head blight: changing farms and rural communities in the northern great plains. *Phytopathology* 90, 17–21. doi: 10.1094/PHYTO.2000.90.1.17
- Wolf, T., Shelest, V., Nath, N., and Shelest, E. (2015). CASSIS and SMIPS: promoter-based prediction of secondary metabolite gene clusters in eukaryotic genomes. *Bioinformatics* 32, 1138–1143. doi: 10.1093/bioinformatics/btv713
- Wollenberg, R. D., Saei, W., Westphal, K. R., Klitgaard, C. S., Nielsen, K. L., Lysøe, E., et al. (2017). Chrysoyone biosynthesis is mediated by a two-module nonribosomal peptide synthetase. *J. Nat. Prod.* 80, 2131–2135. doi: 10.1021/acs.jnatprod.6b00822

- Wu, F. (2007). Measuring the economic impacts of *Fusarium* toxins in animal feeds. *Anim. Feed Sci. Technol.* 137, 363–374. doi: 10.1016/j.anifeeds.2007.06.010
- Yin, W. B., Baccile, J. A., Bok, J. W., Chen, Y., Keller, N. P., and Schroeder, F. C. (2013). A nonribosomal peptide synthetase-derived iron(III) complex from the pathogenic fungus *Aspergillus fumigatus*. *J. Am. Chem. Soc.* 135, 2064–2067. doi: 10.1021/ja311145n
- Yoshida, M., and Nakajima, T. (2010). Deoxynivalenol and nivalenol accumulation in wheat infected with *Fusarium graminearum* during grain development. *Phytopathology* 100, 763–773. doi: 10.1094/PHYTO-100-8-0763
- Zhao, C., Waalwijk, C., de Wit, P. J. G. M., Tang, D., and van der Lee, T. (2013). RNA-Seq analysis reveals new gene models and alternative splicing in the fungal pathogen *Fusarium graminearum*. *BMC Genomics* 14:21. doi: 10.1186/1471-2164-14-21

Conflict of Interest Statement: MHM is a member of the Scientific Advisory Board of Hexagon Bio.

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Hoogendoorn, Barra, Waalwijk, Dickschat, van der Lee and Medema. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.