Title:

Association of secondary metabolite gene clusters with host specialization in the cereal blast fungus *Pyricularia oryzae*

Running Title:

Fungal secondary metabolites in host specialization

Khyati Mehta[1], Jorge C. Navarro-Muñoz[2,#], Sayali Bakore[3], Jérôme Collemare[2,*] and Rajesh Patkar[1,3,*]

[1] *Bharat Chattoo Genome Research Centre, Department of Microbiology and Biotechnology Centre, The Maharaja Sayajirao University of Baroda, Vadodara, India*

[2] *Westerdijk Fungal Biodiversity Institute, Utrecht, the Netherlands*

[3] *Department of Biosciences & Bioengineering, Indian Institute of Technology Bombay, Mumbai, India*

[#] *Current address: Wageningen University and Research, Wageningen, the Netherlands*

[*] Author for correspondence – Rajesh Patkar - rajeshpatkar@iitb.ac.in and Jérôme Collemare - j.collemare@wi.knaw.nl

Abstract

Fungal plant pathogens constantly evolve and deploy novel peptide and metabolite effectors to break down plant resistance and adapt to new host plants. The blast fungal pathogen *Pyricularia oryzae* is a single species subdivided into multiple host-specific lineages. This host specialization is likely due to secreted effectors, including metabolite effectors. Here, we mined 68 genomes of *P. oryzae*, belonging to six host-specific lineages, to identify secondary metabolite (SM) biosynthetic gene clusters (BGCs) associated with host specialization. A similarity network analysis grouped a total of 4,501 BGCs into 283 gene cluster families (GCFs), based on the content and architecture of the BGCs. While most of the GCFs were present in all the *P. oryzae* lineages, two were found specifically in the *Oryza* lineage and one was found in the lineage specific to *Triticum*, *Lolium* and *Eleusine* hosts. Further analysis of the phylogenetic relationships between core biosynthetic genes confirmed that a BGC, comprising a reducing polyketide synthase (PKS) gene (MGG_08236) and four putative tailoring genes, was present only in the *Oryza* lineage. The predicted BGC was found expressed specifically during host penetration and colonization. We propose that this *Oryza* lineage-specific BGC is likely associated with a metabolite effector involved in specialization of *P. oryzae* to rice host. Our findings highlight the importance of further exploring the role of metabolite effectors in specialization of the blast fungus to different cereal hosts.

Keywords

Biosynthetic diversity, population genomics, *Magnaporthe oryzae*, gene cluster family, polyketide synthase, effector

Introduction

Many plant pathogenic fungi are adapted to infect specific host plants. The emergence of new pathogenic strains, as a result of host range expansion or specialization to a new host, can cause severe disease outbreaks (Giraud et al., 2010). It is crucial to understand the molecular basis of such host specialization to increase the durability of crop resistance to diseases (Giraud et al., 2010; Gladieux et al., 2014). The filamentous fungus *Pyricularia oryzae* (synonym *Magnaporthe oryzae*) is the causative agent of the economically important cereal blast disease worldwide, and is a model organism to study host-pathogen interactions (Ebbole, 2007; Talbot, 2003). Although capable of infecting several grasses, *P. oryzae* is classified as a single species, subdivided into several host-specialized lineages with limited primary host range (Gladieux et al., 2018).

Host selection pressure could be one of the factors that shape the virulence of the population of a pathogen, whereby specific virulence-related genes that encode secreted proteins and/or enzymes for secondary metabolite production - collectively known as effectors – are either gained or lost (Hartmann et al., 2018; Zhong et al., 2016). Effectors manipulate the host physiology or immunity to the pathogen's advantage, and thus are major determinants of specialization on a host (Sánchez-Vallet et al., 2018). Indeed, gain- or loss-of-function mutation in avirulence (AVR) genes of *P. oryzae* results in overcoming host resistance or non-host resistance in the case of a new host. For example, a widespread cultivation of wheat cultivar carrying the *RWT3* resistance gene led to a loss of function of the corresponding *PWT3* AVR gene in *P. oryzae* and subsequent emergence of wheat blast (Inoue et al., 2017). Similarly, although gain of the AVR gene *PWL1* led to a loss of pathogenesis on weeping

lovegrass in the *Eleusine*-lineage of *P. oryzae*, it conferred virulence on the new host, finger millet (Asuke et al., 2020).

Fungal secondary metabolites (SMs) have also been shown to exhibit important effector functions during plant-fungal interactions (Collemare et al., 2019). For example, ACE1, a secondary metabolite (SM) produced by *P. oryzae*, confers avirulence towards the rice cultivars carrying the corresponding resistance gene *Pi33* (Böhnert et al., 2004). Various fungal pathogens, including *P. oryzae*, are also known to produce analogues of plant hormones such as jasmonic acid, gibberellin and cytokinins, to manipulate plant growth or subvert plant hormone-based defense signalling pathways during invasion (Patkar & Naqvi, 2017; Shen et al., 2018). A derivative of jasmonic acid (12-hydroxyjasmonic acid) secreted by *P. oryzae* suppresses the host innate immunity specifically in rice but not in barley or wheat (Patkar et al., 2015). Many other fungal SMs such as host-selective toxins (HST) also act as virulence factors – for example, HC-toxin in *Cochliobous carbonum*, T-toxin in *C. heterostrophus*, AM-toxin and ACR-toxin in *Alternaria alternata* (Izumi et al., 2012; Lim & Hooker, 1971; Tsuge et al., 2013; Walton, 2006). These HSTs are essential for pathogenesis on specific hosts carrying the corresponding susceptible gene. The blast fungus produces SMs such as tenuazonic acid and pyriculol during pathogenesis (Chen & Qiang, 2017; Jacob et al., 2017).

The availability of an increasing number of fungal genomes and advances in computational biology facilitates the exploration of molecular determinants of host specialization or host range expansion. For instance, population genomics of the wheat fungal pathogen *Zymoseptoria tritici* has revealed a possible role of SMs in pathogen lineage diversification and adaptation to wheat host (Hartmann et al., 2018). In *P. oryzae* isolates, specialization to

rice host is likely associated with a small number of lineage-specific gene families, which could include genes involved in biosynthesis of SMs and secreted proteins of unknown functions (Chiapello et al., 2015). While most studies have focussed on the effector function of certain secreted proteins in the blast fungus, the link between diverse SM biosynthetic gene clusters (BGCs) and *P. oryzae* host-specific lineages has not been explored.

Here, we embarked on the exploration of genomes from the different lineages of the blast fungus, in search for SM BGCs associated with host specialization. Using similarity network, comparative genomics and phylogenetic analyses, we identified a candidate BGC that is likely involved in pathogenesis and specialization to rice and weeping lovegrass host plants.

Materials and methods

Fungal strains and culture conditions

We isolated 15 field strains of *P. oryzae* from different geographic locations in India and from three different host plants – rice (*Oryza sativa*), finger millet *(Eleusine coracana*) and foxtail millet (*Setaria italica*). Infected tissues with blast lesions were collected and used for single spore isolation as soon as possible. Leaf tissues with blast disease lesions were surface sterilized using 70% ethanol and 1% sodium hypochlorite, followed by three successive washes with sterile distilled water. Surface-sterilized leaf tissues were incubated under micro-humid conditions maintained in a Petri dish for 24-48 h to induce sporulation. Spores were suspended in sterile water and spread evenly on 2% water agar plate. Single germinating spores were picked up under a microscope, and inoculated on Prune Agar (PA: 1 g/L yeast extract, 2.5 g/L lactose, 2.5 g/L sucrose, 0.04% prune juice, pH 6.5; Soundararajan et al., 2004) plates and incubated at 28 °C. Vegetative growth of the fungus on PA plates was continued for 10 days at 28 °C, with initial 3-day incubation under dark conditions, followed by 7-day incubation under constant illumination for conidiation. Genomic DNA was extracted by grinding in liquid nitrogen the fungal biomass obtained from vegetative culture grown in a complete medium (CM: 6 g/L yeast extract, 6 g/L casein enzyme hydrolysate, 10 g/L sucrose) for ~3 days at 28 °C, followed by the standard protocols as described earlier (Dellaporta et al., 1983).

Genome sequencing and assemblies

Newly isolated strains were sequenced for whole genome using a paired-end sequencing approach at >50X depth on Illumina HiSeq2500 platform at AgriGenome Labs Pvt. Ltd.,

India. Short raw reads were processed using Trimmomatic v0.38 at a threshold for the minimum read length of 80 bp for paired-end reads (parameters PE, Leading:10, Trailing:10, Slidingwindow:4:20, Minlen:80; Bolger et al., 2014). *De novo* sequence assemblies were constructed using CLC Genomics Workbench v11.0 with default parameters (Word size:23, Bubble Size:50, Minimum contig length:500, Mismatch cost:2, Insertion cost:3, Deletion cost:3, Length fraction:0.5, Similarity fraction:0.8). Additionally, we used publicly available genome sequences of *P. oryzae* strains from NCBI (Supplementary table S1). The overall quality of each genome assembly was evaluated using BUSCO v5.2.2 (Simão et al., 2015) with the Sordariomycetes dataset. We carried out gene predictions on all the assemblies using Augustus v3.3.3 (Stanke et al., 2006) with Magnaporthe_grisea as a species model.

Construction of species tree

2655 BUSCO proteins conserved in 68 *P. oryzae* strains were retrieved from the BUSCO analysis and aligned using Mafft v7.475 (parameters -reorder; Katoh et al., 2002). TrimAL v1.4.1 (parameters -automated1; Capella-Gutiérrez et al., 2009) was used to remove poorly aligned regions (Supplementary dataset S1, S2). A maximum likelihood tree was generated using IQ-TREE v2.1.2 (Minh et al., 2020) with partition models (Chernomor et al., 2016; Lanfear et al., 2014), model finder (Kalyaanamoorthy et al., 2017) and ultrafast bootstrap (Hoang et al., 2018; parameters -m MF -p partition.nex -bb 1000) (Supplementary dataset S3, S4). Similarly, a species tree including 3 *Pyricularia grisea* and 68 *P. oryzae* strains was generated using a total of 2557 BUSCO proteins (Supplementary dataset S5-S7). Both trees were visualized and annotated using iTOL (Letunic & Bork, 2021).

Prediction of BGCs and similarity network analysis

BGCs were predicted using antiSMASH v6.0.1 (Blin et al., 2021; parameters -taxon fungi -minimal -cluster_hmmer -pfam2go). All the predicted BGCs and reference fungal BGCs from the MIBiG v2.1 (Kautsar et al., 2019) repository were included to perform a similarity network analysis using BiG-SCAPE v1.1.0 (Navarro-Muñoz et al., 2019) with glocal mode and with different cut-off values (parameters –include_singletons –mix –no_classify –clans-off –cutoffs 0.3 0.4 0.5 0.7). A cut-off value of 0.5 was found to be appropriate, as the relevant BGCs showed similarity with the reference BGCs from MIBiG known for metabolites such as DHN (melanin), epipyriculol, alternapyrone, squalestatin S1 and cytochalasans. The network files were visualized with Cytoscape v3.8.2 (Shannon et al., 2003). Visualization of BGC conservation was done with Clinker (Gilchrist & Chooi, 2021). Gene models belonging to BGC-O1 family were manually curated (Supplementary dataset S8).

Synteny analysis

To identify the syntenic genomic regions between any genome assembly and the reference 70-15 strain assembly, global alignments were generated using nucmer (parameters –c 100 --maxmatch) program from MUMMER v3.23 package (Kurtz et al., 2004). Alignments were further filtered using delta-filter utility with length >10 kb and percent identity >80% (parameters -l 10000 -i 80) to retrieve continuous syntenic blocks. The show-coords utility was used to extract alignment coordinates in tabular format and the plots were generated in R using KaryoploteR package (Gel & Serra, 2017).

Core gene phylogeny

Homologs of MGG_08236 were identified in *Pezizomycotina* from the Joint Genome Institute MycoCosm repository (Grigoriev et al., 2014) using BLASTp. The resulting Blast hits were filtered for top 3 hits of each species. Amino acid sequences were aligned using MAFFT v7.475 (Katoh et al., 2002) and poorly aligned regions were removed with TrimAl v1.4.1 (Capella-Gutiérrez et al., 2009). A phylogenetic tree was constructed using Fasttree v2.1.11 (parameters -lg; Price et al., 2010) to identify orthologues and closest paralogues (Supplementary dataset S11). The identified orthologous protein sequences were used to build a phylogenetic tree using IQ-TREE v2.1.2 (Minh et al., 2020) with LG substitution model and ultrafast bootstrap (Hoang et al., 2018), and the SH-like approximate likelihood-ratio test (Guindon et al., 2010) for branch support (parameters -mset LG -bb 1000 -alrt 1000) (Supplementary dataset S12). The tree was visualized and annotated using iTOL (Letunic & Bork, 2021).

Whole-plant infection assay

Conidia were harvested as per the protocol described earlier (Patkar et al., 2012) from different strains collected in India. Whole-plant infection assay was carried out by spraying ~ 10^5 conidia/mL containing 0.05% w/v gelatin onto ~ 4-week-old host plants (Rice and finger millet) and ~3-week-old foxtail millet plants, followed by incubation at 28–30 °C under humid conditions, initially for 24 h in the dark and then for 4–6 days under 14 h light and 10 h dark cycles. Development of disease symptoms was monitored regularly during the entire incubation period and recorded around 5-8 days post inoculation depending on the appearance of lesions on the leaves. The infection assays were carried out with at least three biological replicates.

Gene expression analysis by semi-quantitative RT-PCR

The expression profiles of four genes (MGG_08236, MGG_15107, MGG_15108 and MGG_12496) were studied during different stages of infection. Briefly, 20 μL conidial suspension (~10^5 conidia/mL with 0.01% Tween-20) of B157 strain from *Oryza* lineage was drop-inoculated on to the surface-sterilized 2-week-old barley detached leaf blades placed on water agar plates containing 2 μg/mL Kinetin, followed by incubation under dark (10 h) and light (14 h) cycles at 25 °C. Samples were collected by excising inoculated portions on the leaf blades, at 12-, 24-, 48-, 72- and 96-hours post inoculation (hpi). Three-day-old vegetative mycelia, grown in liquid CM, was used as a control condition. A sample from mock-inoculation (leaf blades inoculated with only 0.01% Tween-20) was used as another experimental control to check any non-specific amplification during RT-PCR. Total RNA was extracted from all samples using TRIzol® reagent (Invitrogen, USA) as per the manufacturer's instructions. Two μg of total RNA each was used for cDNA synthesis using iScript™ cDNA Synthesis Kit (Bio-Rad, USA). Oligonucleotide primers used for RT-PCR for each gene are listed in Supplementary Table S4. Optimized thermal cycling conditions for RT-PCR were as follows: initial denaturation step at 95 °C for 5 min, followed by 30 cycling reactions each at 95 °C (denaturation step) for 30 s, 60 °C (annealing step) for 30 s and 72 °C (extension step); and the final extension step at 72 °C for 5 min. Expression of each gene was quantified based on the intensity of the band measured using ImageJ, and was assessed relative to that of β-tubulin (MGG_00604) as an endogenous control. The data on expression of PKS and tailoring genes are from three independent biological experiments and represented as mean ± standard deviation of mean (SDM) in the graph. The MGG_08236 ORF was amplified from the genomic DNA of *P. oryzae* strains from different lineages using the same thermal cycling conditions as above, except for the extension step, which was set at 1 min in each cycle.

Results

Genome sequencing of host-specific strains of *P. oryzae*

In this study, we used 68 *P. oryzae* genomes isolated from 6 different host plants, rice (*Oryza sativa*), finger millet *(Eleusine coracana*), foxtail millet *(Setaria sp.*), wheat (*Triticum aestivum*), perennial ryegrass *(Lolium sp.*) and weeping lovegrass (*Eragrostis curvula*) (Fig. 1A, Supplementary Table S1). These genomes included 15 field isolates that were collected from infected rice, finger millet and foxtail millet plants grown in different parts of India (Fig. 1A, Supplementary Table S1) and sequenced using Illumina technology (Supplementary Table S2). The remaining 53 genomes were obtained from publicly available genome sequences (Supplementary Table S1, Fig. 1A). Additionally, sequence data from three *P. grisea* strains previously isolated from crabgrass (*Digitaria sp.)* was used as control (Supplementary Table S1). Sixteen out of the publicly available assemblies were obtained from long-read sequencing technology and were highly-contiguous. We performed gene prediction for all 71 assemblies and analysis of BUSCO genes indicated high-quality gene prediction and completeness of the assemblies (>90% BUSCO-score; Supplementary Table S1).

We used 2655 BUSCO genes present in the 68 *P. oryzae* genomes to build a phylogenomic tree (Supplementary dataset S4). The topology of the obtained species tree confirms that *P. oryzae* consists of multiple genetic lineages specialized on different hosts, including *Oryza, Setaria, Eleusine, Eragrostis, Triticum* and *Lolium* (Fig. 1B). This tree is concordant with the previously reported population structure (Gladieux et al., 2018). A similar phylogenomic analysis using additional three genomes from *P. grisea* (*Digitaria* isolates) confirmed that *P.*

grisea is a separate species from *P. oryzae*, as previously reported (Supplementary Fig. S1; Supplementary dataset S7; Gladieux et al., 2018).

We performed infection assays to check the pathogenesis of 16 Indian isolates of *P. oryzae* on three hosts – rice, finger millet and foxtail millet. We found that all isolates but one induced clear blast lesions on their host-of-origin, while they triggered no or few symptoms on non-hosts (Supplementary Fig. S2). In several cases, small dark necrotic spots were observed, indicative of a hypersensitive response caused by recognition of the strain (Supplementary Fig. S2). Intriguingly, MOS1 and MOS4, two strains isolated from infected rice leaf tissue, were found to be genetically similar to that of the *Eleusine* lineage and not related to any lineage, respectively (Fig. 1B). The whole-plant infection assay using the MOS1 strain showed differential virulence pattern, where inoculated finger millet and rice plants developed highly-susceptible and moderately-resistant lesions, respectively, similar to other isolates from finger millet (Supplementary Fig. S2). This result indicates that MOS1 is indeed an *Eleusine* lineage strain which was isolated from infected rice tissue. This isolation could be due to isolates from the *Eleusine* lineage being able to initiate the infection process on rice leaves, resulting sometimes in a typical blast lesion (Supplementary Fig. S2). Alternatively, the strain MOS1 is likely evolving to adapt to rice host, or it randomly landed on the collected rice leaf and benefited from an *Oryza* lineage strain infection. Further studies are required to understand whether it was associated with any co-infection of the concerned rice plant tissue with two genetically distinct *P. oryzae* strains. The MOS4 strain also isolated from rice plant grown in field showed similar virulence pattern as MOS3 and MOS5, with clear interaction with rice. However, only few lesions were observed, and small necrotic spots could indicate recognition by the rice cultivar CO39. These results suggest that MOS4 strain defines a

second lineage that is able to infect rice, yet it may have another primary host that remains to be determined.

Genome mining identified lineage-specific biosynthetic gene clusters in *P. oryzae*

The identification of genomic regions that contain SM BGCs was performed in all 71 genomes using fungiSMASH (Blin et al., 2021). This analysis resulted in predicting a total of 4224 BGCs, with an average number of 59 BGCs per strain (Supplementary Table S3). The predicted BGCs were classified based on the core biosynthetic genes, such as genes encoding polyketide synthases (PKSs), non-ribosomal peptide synthetases (NRPSs) or terpene cyclases (TCs) (Supplementary Fig. S3, Supplementary Table S3). All the host-specific lineages contained similar number of BGCs from all different classes of BGCs, with type I PKSs being the most abundant (Supplementary Fig. S3). Overall, these analyses highlight the great potential of *P. oryzae* to produce SMs, some of which could be involved in virulence and/or host specialization.

To determine if any putative BGC is associated with the ability to infect specific host plants, we performed a network similarity analysis using BiG-SCAPE (Navarro-Muñoz et al., 2019). This analysis included the 4224 predicted BGCs and 277 characterized BGCs from the MIBiG database (Kautsar et al., 2019) used as a reference. The 4501 BGCs were clustered into 283 gene cluster families (GCFs)/subnetworks, of which 180 are singletons, including 160 from the reference characterized BGCs (Fig. 2). Twelve out of 91 GCFs present in *Pyricularia* strains comprised a characterized BGC, including largely conserved ones for DHN melanin, epipyriculol, alternapyrone, squalestatin and cytochalasans in *P. oryzae*, and betaneone in *P. grisea* (Fig. 2; Supplementary Fig. S4-S8). Forty-four GCFs are found in several *P. oryzae* lineages as well as in *P. grisea*, most of them remain uncharacterized.

Interestingly, 13 and 14 out of 91 GCFs appeared to be specific to *P. grisea* and *P. oryzae*, respectively, suggesting their likely role in pathogenesis on *Digitaria* and other relevant hosts (Fig. 2). Of the *P. oryzae*-specific GCFs, three could potentially be associated with host specialization. While BGC-O1 and BGC-O2 were found mostly specific to the *Oryza* lineage, BGC-TLE was found specific to *Tritici*, *Lolium*, *Eleusine* and *Eragrostis* lineages (Fig. 2), which all share a common origin and diverged from the *Oryza* and *Setaria* lineages (Fig. 1). BGC-O1 consists of genes encoding a reducing PKS and tailoring enzymes (see below), whereas BGC-O2 is restricted to a truncated PKS gene alone that is likely not functional (Supplementary Fig. S9). In the case of BGC-TLE, although the standalone core biosynthetic NRPS gene therein is present in all the strains, the gene structure appeared different (Supplementary Fig. S10). Manual curation of the NRPS gene showed the deletion of 557 bp in *Triticum* and *Lolium* lineages corresponding to the second exon in the *Eleusine* lineage (CD156 strain). Further, the presence of a stop codon within the adenylation domain of the NRPS gene in the *Triticum* and *Lolium* lineages suggests pseudogenization. BGC-TLE is thus likely not functional in those strains (Supplementary Fig. S11). Yet, it could be either important for virulence on *Eleusine* host plants only, or its product might have an avirulence effector-like role in *Triticum* and *Lolium* hosts. Thus, out of three candidate BGCs, we further investigated BGC-O1 as potentially involved in specialization on rice and weeping lovegrass hosts.

A novel polyketide biosynthetic gene cluster specifically found in rice/Eragrostis-infecting lineage of *P. oryzae*

BGC-O1 was found in 23 out of the 24 *Oryza* lineage strains used in the study, as well as in one *Eragrostis* lineage strain (Fig. 3). This GCF appeared to contain two different networks,

and comparison of the loci indeed clearly differentiated two putative BGCs - one conserved in all lineages, and BGC-O1 restricted to the *Oryza* and *Eragrostis* lineages (Fig. 3). BGC-O1 contains a novel reducing PKS gene (MGG_08236) as the core biosynthetic gene along with flanking tailoring genes, including one methyl transferase (MGG_15107), two cytochrome P450 monooxygenases (MGG_12496 and MGG_12497) and one Co-A transferase (MGG_15108).

We wondered if only the BGC-O1 cluster region was specific to the *Oryza* lineage, and we thus checked whether flanking genomic regions of BGC-O1 were conserved or variable in different lineages. We mapped representative high-quality assemblies from each lineage with the reference genome assembly of the 70-15 strain from *Oryza* lineage to assess synteny at global level. We found that BGC-O1 was located in the sub-telomeric region of chromosome 2 (NC_017850.1), and approximately 528 kb downstream of the *ACE1* gene cluster in the reference strain 70-15. The macrosynteny analysis between 70-15 and GUY11, both strains from *Oryza* lineage, showed conservation of ~134 kb upstream and ~117 kb downstream flanking regions, along with the BGC-O1 cluster (Fig. 4A). The ~10 kb-long immediate flanking regions were found to be swapped between the GUY11 and 70-15 genomes (Fig. 4A). Similarly, comparison between 70-15 and FR13 strains revealed a partial conservation (~14 kb) of BGC-O1, with an ~9.5 kb region from BGC-O1 in 70-15 strain translocated further downstream on the same contig in the FR13 strain (Fig. 4A). Further, while ~102 kb upstream flanking region was syntenic in 70-15 and FR13, the downstream flanking region showed a few structural rearrangements (Fig. 4A). In contrast, comparison with non-*Oryza* lineages, like *Eleusine* and *Setaria*, showed a significant loss of synteny in the sub-telomeric region of chromosome 2. While the upstream flanking region (~97 kb upstream of BGC-O1 in 70-15) was found to be syntenic to a corresponding region in MZ5-1-6, CD156 and US71, the

BGC-O1 cluster and downstream sequences were either absent or poorly conserved in the genomes of these lineages (Fig. 4B).

Importantly, BGC-O1 did not show any similarity with any of the reference BGCs in MIBiG, suggesting it is a unique BGC likely involved in the biosynthesis of a novel SM associated with specialization to rice host.

BGC-O1 likely originates from a duplication event in the common ancestor of *P. oryzae* and *Colletotrichum eremochloae*

To understand the evolutionary history of BGC-O1, we searched for orthologues and close paralogues of the MGG_08236 PKS protein within *Pezizomycotina* from the MycoCosm repository (Grigoriev et al., 2014). Only two close homologues could be retrieved from the fungus *Colletotrichum eremochloae*. A phylogenetic analysis showed that the PKS 670826 (scaffold_178) from *C. eremochloae* is an orthologue of MGG_08236 PKS, forming an outgroup to the *P. oryzae* lineage (Fig 5A). The other paralogue in *C. eremochloae*, 679399 (scaffold_5), belongs to a clade that is restricted to the *Colletotrichum* genus and forms a sister clade to the MGG_08236 PKS clade. Comparison of the genomic loci of both homologues in *C. eremochloae* and BGC-O1 in *P. oryzae* showed that these two clusters shared the PKS and two cytochrome P450 genes with BGC-O1 (Fig. 5B), while the methyltransferase and Co-A transferase genes were present elsewhere as single copies – 671186 and 576843 – on scaffold_185 and scaffold_10, respectively, in the *C. eremochloae* genome. The presence of the two similar BGCs in *C. eremochloae* suggests that BGC-O1 likely originated from an ancestral duplication event followed by many deletion events.

BGC-O1 genes are expressed specifically during host invasion

The presence of PKS MGG_08236 gene was checked by performing PCR on ORF region of the gene, using genomic DNA of *P. oryzae* strains from rice and finger millet host plants (Fig. 6A). The MGG_08236 gene was present in all the *Oryza* strains studied, except for MOS3, which was identified as the only *Oryza* lineage strain lacking the BGC-O1 based on the *in-silico* analysis. The MGG_08236 ORF was absent in all the *Eleusine* strains used as well as in the MOS1 and MOS4 strains, which are placed outside of *Oryza* lineage in the phylogenetic analysis (Fig. 1B and 6A).

We then studied the expression of the PKS (MGG_08236) and three tailoring genes – methyl transferase (MGG_15107), CoA-transferase (MGG_15108) and cytochrome P450 (MGG_12496) of the BGC-O1 locus during different stages of infection, using semi-quantitative RT-PCR. Total RNA was isolated from fungal vegetative mycelia grown in complete medium, barley leaves inoculated with *P. oryzae* and incubated for different time-points, and uninoculated barley leaves, to assess the levels of transcripts of the aforementioned genes relative to those of β-Tubulin as an internal control. While expression of the core PKS gene was undetected or barely seen in mycelium and at 12 hpi (pre-invasive stage), respectively, its transcripts started accumulating during pathogenic development, with significantly elevated expression at 24 hpi, followed by a stable, slightly increasing expression until 72 hpi (Fig. 6B and 6C). The expression pattern of MGG_15107 and MGG_15108 tailoring genes partly correlated with that of the PKS, albeit with a lower expression level (Fig. 6B and 6C). In contrast, MGG_12496 showed a different expression pattern with a highest expression in mycelium, and at 48 hpi during infection (Fig. 6B and 6C), suggesting it does not belong to the BGC. This result shows that the MGG_08236 PKS gene is specifically expressed during pathogenesis and potentially has a key role to play during host colonization. Altogether, our *in-silico* analyses identified a novel PKS gene

cluster in the *Oryza*-specific lineage, which likely played a key role in shaping specialization of the blast fungus to rice and weeping lovegrass host plants.

Discussion:

Population genomics identifies BGCs with potential roles in virulence and host specialization in *P. oryzae*

We identified three lineage-specific BGCs in the blast fungal pathogen *P. oryzae* using a population genomics approach. While two of them appeared to encode non-functional enzymes, the *Oryza* and *Eragrostis*-specific BGC, named BGC-O1, appears functional and is present in almost all the genomes from *Oryza* (23 out of 24) and *Eragrostis* (1 out of 1) lineages studied here. Our population-genomics-based finding of the *Oryza* lineage-specific BGC is supported by gene expression analyses. Indeed, we found that the *Oryza*-lineage-specific PKS (MGG_08236) transcript accumulated specifically during penetration and colonization of the host. Recent transcriptomics-based studies have shown that the expression of PKS from BGC-O1 (MGG_08236) and neighboring tailoring genes (methyl transferase, MGG_15107; Co-A transferase, MGG_15108; cytochrome P450, MGG_12496) was induced during the initial biotrophic colonization stage of blast disease (Jeon et al., 2020; Yan et al., 2023). These findings strongly suggest that the SM produced by BGC-O1 could have an effector-like function. Many plant fungal pathogens such as *Colletotrichum higginsianum, Fusarium graminearum, Zymoseptoria tritici*, and *P. oryzae* are known to produce a number of SMs during pathogenesis, especially during host penetration and colonization (Collemare, Billard, et al., 2008; Dallery et al., 2017; Harris et al., 2016; Jeon et al., 2020; Palma-Guerrero et al., 2017; Patkar et al., 2015; Rudd et al., 2015; Yan et al., 2023). To study if the BGC-O1 has any role in virulence in a host-specific manner, deletion of the *PKS* gene (MGG_08236) was attempted in an *Oryza*-specific isolate. Despite using different approaches and a number of attempts, generation of a deletion mutant has been unsuccessful so far, possibly due to its

presence in the sub-telomeric region of chromosome 2 of the reference strain 70-15. A similar difficulty was faced in generating deletion mutants at the *ACE1* locus (Collemare, Pianfetti, et al., 2008).

While only three BGCs were present in host-specific lineages, most of the other BGCs were found in all lineages. However, such apparently conserved BGCs can show variation in their tailoring gene content. Such differences could result in the production of distinct SMs with potential effector function in a host specific manner. Alternatively, conservation of BGCs in different lineages could indicate a key role in virulence on cereals. Thus, analyses investigating variations among conserved SM-producing BGCs across different *P. oryzae* lineages could be useful in identifying the likely diversity in the SMs produced and their association with specific hosts.

Conservation of BGC-O1 in an unstable genomic context

We found that the *Oryza*-specific cluster BGC-O1 is located in the sub-telomeric region on the chromosome 2 (NC_017850.1) of the reference strain 70-15. While the genomic regions covering the BGC-O1 and its flanking sequences were highly syntenic in strains from the *Oryza* lineage, only the upstream flanking region of this locus was found conserved in all the other lineages. The chromosome ends tend to have substantially higher levels of polymorphism because they frequently undergo rearrangements due to the presence of AT-rich repeats and transposable elements (Ahead, 2012; Farman & Kim, 2005; Rahnama et al., 2020). In the absence of sexual reproduction, asexually propagating pathogens likely utilize chromosomal rearrangements as one of the mechanisms of host-specialization as shown in *Verticillium dahliae* (Jonge et al., 2013). Similarly, the regions adjacent to a telomere in *P. oryzae* are reported to be highly polymorphic and enriched with genes involved in interaction

with its host plants (Rahnama et al., 2021). Gain and/or loss of effector genes, lineage-specific gene families and chromosomal rearrangements are likely the major evolutionary mechanisms involved in such host specificity and adaptation of *P. oryzae* (Chiapello et al., 2015; Chuma et al., 2011; Gómez Luciano et al., 2019; Jonge et al., 2013; Yoshida et al., 2016). Given that BGC-O1 is also flanked by transposons in most strains of the *Oryza* lineage of *P. oryzae*, we hypothesize that genomic alterations, most likely through deletions, may have resulted in the loss of the BGC in other host-specialized lineages. We further found that BGC-O1 is located on a mini chromosome (UEMA03000009.1) in the FR13 isolate, which also carries the *ACE1* BGC (Langner et al., 2021). Thus, retention of BGC-O1 in *Oryza* and *Eragrostis* lineages, despite its localization in unstable sub-telomeric regions, suggests a significant role of host selection pressure in its conservation therein. The most likely selection pressure to retain BGC-O1 relates to the ability to successfully infect a host, which is also indicated by the specific expression profile during infection. The presence of homologous BGCs in *C. eremochloae*, the causal agent of anthracnose disease on centipede-grass turf in southern United States (Crouch & Tomaso-Peterson, 2012), suggests a particular function for BGC-O1 in infecting certain cereals only (*Oryza*, *Eragrostis*, *Eremochloa*). Further investigations are needed to determine if the *C. eremochloae* BGC is expressed during infection and whether it produces the same metabolite as BGC-O1 in *P. oryzae*.

Altogether, our findings highlight the importance of population genomics-based studies in identifying secondary metabolite BGCs and the corresponding genomic rearrangements, likely driven by the host selection pressure, leading to a host-specialized lineage of the blast fungal pathogen.

Acknowledgement

We acknowledge late Prof. Bharat B. Chattoo for his guidance and providing sophisticated laboratory facility at the Bharat Chattoo Genome Research Centre (BCGRC) and acquiring the grant (BT/PR6949/AGIII/103/862/2012) from DBT, GoI. We also fondly remember late Dr. Johannes Manjrekar for the useful scientific discussions during this work. We thank Dr. Malali Gowda and Dr. T. R. Sharma for providing blast-infected tissues for some of the isolates used in this study. KM was supported by the DBT (GoI)-funded grant (BT/PR6949/AGIII/103/862/2012) and partly by the intramural University Research Scholarship from The M. S. University of Baroda. KM thanks EMBO for the Scientific Exchange Grant (SEG no. 8866) to conduct part of this work at the Westerdijk Institute, The Netherlands. We also acknowledge the funding support (EMR/2017/005303) from SERB, GoI, and the Seed Grant, IIT Bombay, awarded to RP. We thank the BCGRC group at MSU for useful discussions.

References

Ahead, P. A. (2012). *Telomere-targeted retrotransposons in the rice blast fungus*. *191*(859), 389–406. https://doi.org/10.1534/genetics.111.137950

Asuke, S., Tanaka, M., Hyon, G.-S., Inoue, Y., Vy, T. T. P., Niwamoto, D., Nakayashiki, H., & Tosa, Y. (2020). Evolution of an *Eleusine* -Specific Subgroup of *Pyricularia oryzae* Through a Gain of an Avirulence Gene. *Molecular Plant-Microbe Interactions®*, *33*(2), 153–165. https://doi.org/10.1094/MPMI-03-19-0083-R

Blin, K., Shaw, S., Kloosterman, A. M., Charlop-Powers, Z., Van Wezel, G. P., Medema, M. H., & Weber, T. (2021). antiSMASH 6.0: Improving cluster detection and comparison capabilities. *Nucleic Acids Research*, *49*(W1), W29–W35. https://doi.org/10.1093/nar/gkab335

Böhnert, H. U., Fudal, I., Dioh, W., Tharreau, D., Notteghem, J. L., & Lebrun, M. H. (2004). A putative polyketide synthase/peptide synthetase from Magnaporthe grisea signals pathogen attack to resistant rice. *Plant Cell*, *16*(9), 2499–2513. https://doi.org/10.1105/tpc.104.022715

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics (Oxford, England)*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Capella-Gutiérrez, S., Silla-Martínez, J. M., & Gabaldón, T. (2009). trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, *25*(15), 1972–1973. https://doi.org/10.1093/bioinformatics/btp348

Chen, S., & Qiang, S. (2017). Recent advances in tenuazonic acid as a potential herbicide. *Pesticide Biochemistry and Physiology*, *143*, 252–257. https://doi.org/10.1016/j.pestbp.2017.01.003

Chernomor, O., Von Haeseler, A., & Minh, B. Q. (2016). Terrace Aware Data Structure for Phylogenomic Inference from Supermatrices. *Systematic Biology*, *65*(6), 997–1008. https://doi.org/10.1093/SYSBIO/SYW037

Chiapello, H., Mallet, L., Guérin, C., Aguileta, G., Amselem, J., Kroj, T., Ortega-Abboud, E., Lebrun, M.-H., Henrissat, B., Gendrault, A., Rodolphe, F., Tharreau, D., & Fournier, E. (2015). Deciphering Genome Content and Evolutionary Relationships of Isolates from the Fungus Magnaporthe oryzae Attacking Different Host Plants. *Genome Biology and Evolution*, *7*(10), 2896–2912. https://doi.org/10.1093/gbe/evv187

Chuma, I., Isobe, C., Hotta, Y., Ibaragi, K., Futamata, N., Kusaba, M., Yoshida, K., Terauchi, R., Fujita, Y., Nakayashiki, H., Valent, B., & Tosa, Y. (2011). Multiple translocation of the AVR-Pita effector gene among chromosomes of the rice blast fungus Magnaporthe oryzae and related species. *PLoS Pathogens*, *7*(7), e1002147. https://doi.org/10.1371/journal.ppat.1002147

Collemare, J., Billard, A., Böhnert, H. U., & Lebrun, M.-H. (2008). Biosynthesis of secondary metabolites in the rice blast fungus Magnaporthe grisea: The role of hybrid PKS-NRPS in pathogenicity. *Mycological Research*, *112*(2), 207–215. https://doi.org/10.1016/j.mycres.2007.08.003

Collemare, J., O'Connell, R., & Lebrun, M. H. (2019). Nonproteinaceous effectors: The terra incognita of plant–fungal interactions. *New Phytologist*, *223*(2), 590–596. https://doi.org/10.1111/nph.15785

Collemare, J., Pianfetti, M., Houlle, A.-E., Morin, D., Camborde, L., Gagey, M.-J., Barbisan, C., Fudal, I., Lebrun, M.-H., & Böhnert, H. U. (2008). *Magnaporthe grisea* avirulence gene *ACE1* belongs to an infection-specific gene cluster involved in secondary

metabolism. *New Phytologist*, *179*(1), 196–208. https://doi.org/10.1111/j.1469-8137.2008.02459.x

Crouch, J. A., & Tomaso-Peterson, M. (2012). Anthracnose disease of centipedegrass turf caused by Colletotrichum eremochloae, a new fungal species closely related to Colletotrichum sublineola. *Mycologia*, *104*(5), 1085–1096. https://doi.org/10.3852/11-317

Dallery, J.-F., Lapalu, N., Zampounis, A., Pigné, S., Luyten, I., Amselem, J., Wittenberg, A. H. J., Zhou, S., de Queiroz, M. V., Robin, G. P., Auger, A., Hainaut, M., Henrissat, B., Kim, K.-T., Lee, Y.-H., Lespinet, O., Schwartz, D. C., Thon, M. R., & O'Connell, R. J. (2017). Gapless genome assembly of Colletotrichum higginsianum reveals chromosome structure and association of transposable elements with secondary metabolite gene clusters. *BMC Genomics*, *18*(1), 667. https://doi.org/10.1186/s12864-017-4083-x

Dellaporta, S. L., Wood, J., & Hicks, J. B. (1983). A plant DNA minipreparation: Version II. *Plant Molecular Biology Reporter*, *1*(4), 19–21. https://doi.org/10.1007/BF02712670

Ebbole, D. J. (2007). Magnaporthe as a model for understanding host-pathogen interactions. *Annual Review of Phytopathology*, *45*, 437–456. https://doi.org/10.1146/annurev.phyto.45.062806.094346

Farman, M. L., & Kim, Y. S. (2005). Telomere hypervariability in Magnaporthe oryzae. *Molecular Plant Pathology*, *6*(3), 287–298. https://doi.org/10.1111/J.1364-3703.2005.00285.X

Gel, B., & Serra, E. (2017). karyoploteR: an R/Bioconductor package to plot customizable genomes displaying arbitrary data. *Bioinformatics*, *33*(19), 3088–3090. https://doi.org/10.1093/BIOINFORMATICS/BTX346

Gilchrist, C. L. M., & Chooi, Y. H. (2021). clinker & clustermap.js: Automatic generation of gene cluster comparison figures. *Bioinformatics*, *37*(16), 2473–2475. https://doi.org/10.1093/BIOINFORMATICS/BTAB007

Giraud, T., Gladieux, P., & Gavrilets, S. (2010). Linking the emergence of fungal plant diseases with ecological speciation. *Trends in Ecology & Evolution*, *25*(7), 387–395. https://doi.org/10.1016/J.TREE.2010.03.006

Gladieux, P., Condon, B., Ravel, S., Soanes, D., Maciel, J. L. N., Nhani, A., Chen, L., Terauchi, R., Lebrun, M., Tharreau, D., Mitchell, T., Pedley, K. F., Valent, B., Talbot, N. J., Farman, M., & Fournier, E. (2018). Gene Flow between Divergent Cereal- and Grass-Specific Lineages of the Rice Blast Fungus Magnaporthe oryzae. *mBio*, *9*(1), 1–19. https://doi.org/10.1128/mBio.01219-17

Gladieux, P., Ropars, J., Badouin, H., Branca, A., Aguileta, G., De Vienne, D. M., Rodríguez De La Vega, R. C., Branco, S., & Giraud, T. (2014). Fungal evolutionary genomics provides insight into the mechanisms of adaptive divergence in eukaryotes. *Molecular Ecology*, *23*(4), 753–773. https://doi.org/10.1111/mec.12631

Gómez Luciano, L. B., Tsai, I. J., Chuma, I., Tosa, Y., Chen, Y. H., Li, J. Y., Li, M. Y., Lu, M. Y. J., Nakayashiki, H., & Li, W. H. (2019). Blast Fungal Genomes Show Frequent Chromosomal Changes, Gene Gains and Losses, and Effector Gene Turnover. *Molecular Biology and Evolution*, *36*(6), 1148–1161. https://doi.org/10.1093/molbev/msz045

Grigoriev, I. V., Nikitin, R., Haridas, S., Kuo, A., Ohm, R., Otillar, R., Riley, R., Salamov, A., Zhao, X., Korzeniewski, F., Smirnova, T., Nordberg, H., Dubchak, I., & Shabalov, I. (2014). MycoCosm portal: Gearing up for 1000 fungal genomes. *Nucleic Acids Research*, *42*(D1), D699–D704. https://doi.org/10.1093/NAR/GKT1183

Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. (2010). New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Systematic Biology*, *59*(3), 307–321. https://doi.org/10.1093/SYSBIO/SYQ010

Harris, L. J., Balcerzak, M., Johnston, A., Schneiderman, D., & Ouellet, T. (2016). Host-preferential Fusarium graminearum gene expression during infection of wheat, barley, and maize. *Fungal Biology*, *120*(1), 111–123. https://doi.org/10.1016/j.funbio.2015.10.010

Hartmann, F. E., McDonald, B. A., & Croll, D. (2018). Genome-wide evidence for divergent selection between populations of a major agricultural pathogen. *Molecular Ecology*, *27*(12), 2725–2741. https://doi.org/10.1111/mec.14711

Hoang, D. T., Chernomor, O., Von Haeseler, A., Minh, B. Q., & Vinh, L. S. (2018). UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Molecular Biology and Evolution*, *35*(2), 518–522. https://doi.org/10.1093/MOLBEV/MSX281

Inoue, Y., Vy, T. T. P., Yoshida, K., Asano, H., Mitsuoka, C., Asuke, S., Anh, V. L., Cumagun, C. J. R., Chuma, I., Terauchi, R., Kato, K., Mitchell, T., Valent, B., Farman, M., & Tosa, Y. (2017). Evolution of the wheat blast fungus through functional losses in a host specificity determinant. *Science (New York, N.Y.)*, *357*(6346), 80–83. https://doi.org/10.1126/science.aam9654

Izumi, Y., Ohtani, K., Miyamoto, Y., Masunaka, A., Fukumoto, T., Gomi, K., Tada, Y., Ichimura, K., Peever, T. L., & Akimitsu, K. (2012). A Polyketide Synthase Gene, *ACRTS2* , Is Responsible for Biosynthesis of Host-Selective ACR-Toxin in the Rough Lemon Pathotype of *Alternaria alternata*. *Molecular Plant-Microbe Interactions®*, *25*(11), 1419–1429. https://doi.org/10.1094/MPMI-06-12-0155-R

Jacob, S., Grötsch, T., Foster, A. J., Schüffler, A., Rieger, P. H., Sandjo, L. P., Liermann, J. C., Opatz, T., & Thines, E. (2017). Unravelling the biosynthesis of pyriculol in the rice blast fungus magnaporthe oryzae. *Microbiology (United Kingdom)*, *163*(4), 541–553. https://doi.org/10.1099/mic.0.000396

Jeon, J., Lee, G. W., Kim, K. T., Park, S. Y., Kim, S., Kwon, S., Huh, A., Chung, H., Lee, D. Y., Kim, C. Y., & Lee, Y. H. (2020). Transcriptome profiling of the rice blast fungus magnaporthe oryzae and its host oryza sativa during infection. *Molecular Plant-Microbe Interactions*, *33*(2), 141–144. https://doi.org/10.1094/MPMI-07-19-0207-A

Jonge, R. de, Bolton, M., Kombrink, A., Berg, G. V. D. van den, Yadeta, K. A., & Thomma, B. (2013). Extensive chromosomal reshuffling drives evolution of virulence in an asexual pathogen. *Genome Research*. https://doi.org/10.1101/GR.152660.112

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., Von Haeseler, A., & Jermiin, L. S. (2017). ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nature Methods 2017 14:6*, *14*(6), 587–589. https://doi.org/10.1038/nmeth.4285

Katoh, K., Misawa, K., Kuma, K. I., & Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, *30*(14), 3059–3066. https://doi.org/10.1093/NAR/GKF436

Kautsar, S. A., Blin, K., Shaw, S., Navarro-Muñoz, J. C., Terlouw, B. R., van der Hooft, J. J. J., van Santen, J. A., Tracanna, V., Suarez Duran, H. G., Pascal Andreu, V., Selem-Mojica, N., Alanjary, M., Robinson, S. L., Lund, G., Epstein, S. C., Sisto, A. C., Charkoudian, L. K., Collemare, J., Linington, R. G., … Medema, M. H. (2019). MIBiG 2.0: A repository for biosynthetic gene clusters of known function. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkz882

Kurtz, S., Phillippy, A., Delcher, A. L., Smoot, M., Shumway, M., Antonescu, C., & Salzberg, S. L. (2004). Versatile and open software for comparing large genomes. *Genome Biology*, *5*(2), R12. https://doi.org/10.1186/gb-2004-5-2-r12

Lanfear, R., Calcott, B., Kainer, D., Mayer, C., & Stamatakis, A. (2014). Selecting optimal partitioning schemes for phylogenomic datasets. *BMC Evolutionary Biology*, *14*(1), 1–14. https://doi.org/10.1186/1471-2148-14-82

Langner, T., Harant, A., Gomez-Luciano, L. B., Shrestha, R. K., Malmgren, A., Latorre, S. M., Burbano, H. A., Win, J., & Kamoun, S. (2021). Genomic rearrangements generate hypervariable mini-chromosomes in host-specific isolates of the blast fungus. *PLOS Genetics*, *17*(2), e1009386. https://doi.org/10.1371/journal.pgen.1009386

Letunic, I., & Bork, P. (2021). Interactive Tree Of Life (iTOL) v5: An online tool for phylogenetic tree display and annotation. *Nucleic Acids Research*, *49*(W1), W293–W296. https://doi.org/10.1093/NAR/GKAB301

Lim, S. M., & Hooker, A. L. (1971). Southern Corn Leaf Blight: Genetic Control of Pathogenicity and Toxin Production in Race T and Race O of *Cochliobolus heterostrophus*. *Genetics*, *69*(1), 115–117. https://doi.org/10.1093/genetics/69.1.115

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., Von Haeseler, A., Lanfear, R., & Teeling, E. (2020). IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Molecular Biology and Evolution*, *37*(5), 1530–1534. https://doi.org/10.1093/molbev/msaa015

Navarro-Muñoz, J. C., Selem-Mojica, N., Mullowney, M. W., Kautsar, S. A., Tryon, J. H., Parkinson, E. I., De Los Santos, E. L. C., Yeong, M., Cruz-Morales, P., Abubucker, S., Roeters, A., Lokhorst, W., Fernandez-Guerra, A., Cappelini, L. T. D., Goering, A. W., Thomson, R. J., Metcalf, W. W., Kelleher, N. L., Barona-Gomez, F., & Medema, M. H. (2019). A computational framework to explore large-scale biosynthetic diversity. *Nature Chemical Biology*, 1–9. https://doi.org/10.1038/s41589-019-0400-9

Palma-Guerrero, J., Ma, X., Torriani, S. F. F., Zala, M., Francisco, C. S., Hartmann, F. E., Croll, D., & McDonald, B. A. (2017). Comparative Transcriptome Analyses in Zymoseptoria tritici Reveal Significant Differences in Gene Expression Among Strains During Plant Infection. *Molecular Plant-Microbe Interactions: MPMI*, *30*(3), 231–244. https://doi.org/10.1094/MPMI-07-16-0146-R

Patkar, R. N., Benke, P. I., Qu, Z., Constance Chen, Y. Y., Yang, F., Swarup, S., & Naqvi, N. I. (2015). A fungal monooxygenase-derived jasmonate attenuates host innate immunity. *Nature Chemical Biology*, *11*(9), 733–740. https://doi.org/10.1038/nchembio.1885

Patkar, R. N., & Naqvi, N. I. (2017). Fungal manipulation of hormone-regulated plant defense. *PLoS Pathogens*, *13*(6), 10–14. https://doi.org/10.1371/journal.ppat.1006334

Patkar, R. N., Ramos-Pamplona, M., Gupta, A. P., Fan, Y., & Naqvi, N. I. (2012). Mitochondrial β-oxidation regulates organellar integrity and is necessary for conidial germination and invasive growth in *Magnaporthe oryzae*: Mitochondrial β-oxidation in *Magnaporthe*. *Molecular Microbiology*, *86*(6), 1345–1363. https://doi.org/10.1111/mmi.12060

Price, M. N., Dehal, P. S., & Arkin, A. P. (2010). FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. *PLoS ONE*, *5*(3), e9490. https://doi.org/10.1371/journal.pone.0009490

Rahnama, M., Novikova, O., Starnes, J. H., Zhang, S., Chen, L., & Farman, M. L. (2020). Transposon-mediated telomere destabilization: A driver of genome evolution in the blast fungus. *Nucleic Acids Research*, gkaa287. https://doi.org/10.1093/nar/gkaa287

Rahnama, M., Wang, B., Dostart, J., Novikova, O., Yackzan, D., Yackzan, A., Bruss, H., Baker, M., Jacob, H., Zhang, X., Lamb, A., Stewart, A., Heist, M., Hoover, J., Calie, P., Chen, L., Liu, J., & Farman, M. L. (2021). Telomere Roles in Fungal Genome Evolution and Adaptation. *Frontiers in Genetics*, *12*, 676751. https://doi.org/10.3389/fgene.2021.676751

Rudd, J. J., Kanyuka, K., Hassani-Pak, K., Derbyshire, M., Andongabo, A., Devonshire, J., Lysenko, A., Saqi, M., Desai, N. M., Powers, S. J., Hooper, J., Ambroso, L., Bharti, A., Farmer, A., Hammond-Kosack, K. E., Dietrich, R. A., & Courbot, M. (2015). Transcriptome and metabolite profiling of the infection cycle of Zymoseptoria tritici on wheat reveals a biphasic interaction with plant immunity involving differential pathogen chromosomal contributions and a variation on the hemibiotrophic lifestyle definition. *Plant Physiology*, *167*(3), 1158–1185. https://doi.org/10.1104/pp.114.255927

Sánchez-Vallet, A., Fouché, S., Fudal, I., Hartmann, F. E., Soyer, J. L., Tellier, A., & Croll, D. (2018). The Genome Biology of Effector Gene Evolution in Filamentous Plant Pathogens. *Annual Review of Phytopathology*, *56*(1), 21–40. https://doi.org/10.1146/annurev-phyto-080516-035303

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., & Ideker, T. (2003). Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Research*, *13*(11), 2498–2504. https://doi.org/10.1101/GR.1239303

Shen, Q., Liu, Y., & Naqvi, N. I. (2018). Fungal effectors at the crossroads of phytohormone signaling. *Current Opinion in Microbiology*, *46*, 1–6. https://doi.org/10.1016/J.MIB.2018.01.006

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., & Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, *31*(19), 3210–3212. https://doi.org/10.1093/BIOINFORMATICS/BTV351

Soundararajan, S., Jedd, G., Li, X., Ramos-pamplon, M., Chua, N. H., & Naqvi, N. I. (2004). Woronin body function in Magnaporthe grisea is essential for efficient pathogenesis and for survival during nitrogen starvation stress. *The Plant Cell*, *16*(June), 1564–1574. https://doi.org/10.1105/tpc.020677.2001

Stanke, M., Schöffmann, O., Morgenstern, B., & Waack, S. (2006). Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources. *BMC Bioinformatics*, *7*(1), 1–11. https://doi.org/10.1186/1471-2105-7-62/TABLES/2

Talbot, N. J. (2003). On the Trail of a Cereal Killer: Exploring the Biology of Magnaporthe grisea. *Annual Review of Microbiology*, *57*(1), 177–202. https://doi.org/10.1146/annurev.micro.57.030502.090957

Tsuge, T., Harimoto, Y., Akimitsu, K., Ohtani, K., Kodama, M., Akagi, Y., Egusa, M., Yamamoto, M., & Otani, H. (2013). Host-selective toxins produced by the plant pathogenic fungus *Alternaria alternata*. *FEMS Microbiology Reviews*, *37*(1), 44–66. https://doi.org/10.1111/j.1574-6976.2012.00350.x

Walton, J. D. (2006). HC-toxin. *Phytochemistry*, *67*(14), 1406–1413. https://doi.org/10.1016/j.phytochem.2006.05.033

Yan, X., Tang, B., Ryder, L. S., MacLean, D., Were, V. M., Eseola, A. B., Cruz-Mireles, N., Ma, W., Foster, A. J., Osés-Ruiz, M., & Talbot, N. J. (2023). The transcriptional landscape of plant infection by the rice blast fungus *Magnaporthe oryzae* reveals

distinct families of temporally co-regulated and structurally conserved effectors. *The Plant Cell*, koad036. https://doi.org/10.1093/plcell/koad036

Yoshida, K., Saunders, D. G. O., Mitsuoka, C., Natsume, S., Kosugi, S., Saitoh, H., Inoue, Y., Chuma, I., Tosa, Y., Cano, L. M., Kamoun, S., & Terauchi, R. (2016). Host specialization of the blast fungus Magnaporthe oryzae is associated with dynamic gain and loss of genes linked to transposable elements. *BMC Genomics*, *17*(1), 370. https://doi.org/10.1186/s12864-016-2690-6

Zhong, Z., Norvienyeku, J., Chen, M., Bao, J., Lin, L., Chen, L., Lin, Y., Wu, X., Cai, Z., Zhang, Q., Lin, X., Hong, Y., Huang, J., Xu, L., Zhang, H., Chen, L., Tang, W., Zheng, H., Chen, X., … Nielsen, R. (2016). Directional Selection from Host Plants Is a Major Force Driving Host Specificity in Magnaporthe Species. *Scientific Reports*, *6*, 25591. https://doi.org/10.1038/srep25591

Data accessibility Statement

Raw sequence data for the newly sequenced Indian strains under this study have been made accessible at SRA database under the BioProject ID PRJNA973860. All datasets generated for this study are included in the article/Supplementary Material.

Benefit-sharing section statement

A research collaboration established with scientists from Westerdijk institute, The Netherlands, was helpful in *in-silico* analyses of the secondary metabolism related gene clusters in the blast fungal pathogen. All collaborators are included as co-authors. The results of research, addressing a priority concern, will be shared with the broader scientific community.

Author contributions

KM collected samples and contributed to the isolation of the Indian strains of blast fungus and phenotyping experiments. KM conducted formal analyses on diversity in BGCs, JC and JCN-M helped with the investigation of the analyses. KM, RP and JC conducted image analyses and statistical analyses. JCN-M provided important in-house scripts for similarity network analysis. SB helped KM with the gene expression analysis. KM wrote the original draft of the manuscript. All authors contributed to the review and editing of the manuscript. KM, RP and JC designed the study. RP provided the resources for the study. RP and JC supervised the study.

Conflict of interest statement

The authors declare that there is no conflict of interest.

754

Figure legends:

Figure 1. *Pyricularia oryzae* worldwide populations are organized in host-specific lineages. A) Geographic distribution of sequenced *P. oryzae* strains. The color-coded dots denote the host plant of origin of the strains. B) Phylogenetic tree constructed based on concatenation of 2655 BUSCO genes present in all 68 *P. oryzae* genomes used in this study. Colored shapes in the background depict different host-specific genetic lineages.

Figure 2. Similarity network analysis of biosynthetic gene clusters (BGCs) from *Pyricularia oryzae* and *Pyricularia grisea*. A BiG-SCAPE analysis with a cutoff c0.5 depicts similarity of 4224 BGCs from *P. oryzae* or *P. grisea* with 277 reference BGCs from MIBiG database. Each dot represents a BGC and is color-coded according to the lineage. Gene cluster families (GCF; subnetworks) marked with green boxes share significant homology with reference MIBiG BGCs (grey-colored circle). GCFs marked with red boxes are found to be unique to host-specific lineages. The length of the gray lines is proportional to the genetic distance between BGCs. Singletons are shown as individual dots at the bottom.

Figure 3. BGC-O1 is predominantly present in *Oryza* lineage of *Pyricularia oryzae*. BiG-SCAPE network showing BGC-O1 present in 23 genomes from *Oryza* and one genome from *Eragrostis* lineages. Conservation of the BGC in selected strains is depicted using Clinker. Vertical bars besides the name of strain are colored according to the lineages. Red and gray dashed boxes separate BGC-O1 and BGC-O1-like clusters, respectively.

Figure 4. Genomic localization of BGC-O1 and synteny between the *Oryza* and non-*Oryza* lineages. A) Synteny analysis using pairwise comparison within *Oryza* lineage. Genomes of GUY11 and FR13 aligned individually with reference genome 70-15. The syntenic BGC-O1 locus (orange) and the flanking regions (blue) on chromosome 2 of 70-15 are depicted. The differential lengths of BGC-O1 in different isolates are marked with a bar and corresponding length in kilobases. The black arrowhead depicts genomic rearrangement (swapping) in the flanking ~10 Kb region. The red arrowhead marks the rearrangement of ~9 Kb region of the BGC-O1 in FR13. B) Synteny analyses of representative genome assemblies belonging to *Setaria* (US71), *Eleusine* (MZ5-1-6 and CD156) and *Oryza* (70-15) lineages. The BGC-O1 locus (orange) and flanking region (blue) sequences from US71, MZ5-1-6 and CD156 aligned with that of 70-15 are depicted.

Figure 5. Evolutionary origin of the reducing polyketide synthase (PKS) MGG_08236 gene. A) Maximum-likelihood phylogenetic tree using protein sequences of reducing PKS genes from BGC-O1 and related gene cluster families, as well as homologues retrieved from MycoCosm repository. Tip labels depicting sequences from *Oryza* and *Eragrostis*-specific clade are marked with blue and pink background, respectively. The *Colletotrichum eremocloae* ortholog (jgi.p_Coler1_670826), closer to MGG_08236, is denoted with pink label; whereas the more distant paralogue (jgi.p_Coler1_679399) is shown in orange. Gray shaded triangles denote collapsed clades with distant sequences. Branches were supported by > 95% Bootstrap values indicated with gray circles at the nodes. The tree is rooted at

midpoint. B) Comparative analysis, using Clinker tool, of the BGC-O1 or BGC-O1-like loci from three *Oryza*-specific and *C. eremochloae* genomes.

Figure 6. BGC-O1 genes are specifically expressed during pathogenic development. A) Gel electrophoresis of PCR products displaying presence or absence of ORF region of rPKS MGG_08236 from genomic DNA of the indicated *P. oryzae* strains. B157, MOS2, MOS3 and MOS5 belong to the *Oryza* lineage; whereas, the MOS1, MEC1, MEC4, MEC5, MEC6 and MEA1 belong to the *Eleusine* lineage. * - the only *Oryza* strain that lacked the BGC-O1 in *in-silico* analysis. ** - the *Oryza* strains placed outside *Oryza* lineage. Arrowheads corresponds to the 1 Kb size of band from Marker. B) RT-PCR gel depicting expression of genes MGG_08236 (Polyketide synthase), MGG_15107 (Methyl transferase), MGG_15108 (Co-A transferase) and MGG_12496 (Cytochrome P450) in *Oryza*-specific strain B157 at different stages of barley infection (12 to 96 hours post inoculation) and during vegetative growth (mycelium) in complete medium. C) Quantification of the expression of BGC-O1 genes. The intensity of each band was measured using ImageJ, and relative gene expression was calculated relative to that of β-tubulin (MGG_00604) as an endogenous control. The data on expression of all the genes represents mean ± standard deviation of mean (SDM) from three independent biological experiments.

Supplementary Information

Supplementary tables:

Supplementary Table S1: *Pyricularia oryzae* and *Pyricularia grisea* genomes used in this study.
Supplementary Table S2: *De novo* assembly statistics and gene prediction content of *Pyricularia oryzae* strains sequenced in this study.
Supplementary Table S3: Number of BGCs predicted using fungiSMASH in this study.
Supplementary Table S4: Oligonucleotide primers used to study the relative expression of BGC-O1 genes.

Supplementary figures:

Supplementary Figure S1: *Pyricularia oryzae* and *Pyricularia grisea* are evolutionarily distinct species adapted to different hosts.
Supplementary Figure S2: Pathogenicity tests of *P. oryzae* strains isolated from different hosts and geographic locations in India.
Supplementary Figure S3: Occurrence of BGCs associated with specific classes of SM in individual genomes of *P. oryzae* and *P. grisea*.
Supplementary Figure S4: Homology of melanin-associated *P. oryzae* and *P. grisea* gene cluster family with reference BGCs in MIBiG database.
Supplementary Figure S5: Analysis of homology between *ACE1* gene cluster family in *P. oryzae* and *P. grisea* and reference BGCs in MIBiG database.
Supplementary Figure S6: Analysis of homology between putative epipyriculol-associated gene cluster family in *P. oryzae* and *P. grisea* and reference BGCs in MIBiG database.
Supplementary Figure S7: Homology of putative Squalestatin S1-associated *P. oryzae* and *P. grisea* gene cluster family with reference BGCs in MIBiG database.
Supplementary Figure S8: Analysis of homology between Tenuazonic acid-associated gene cluster families in host-adapted *P. oryzae* and *P. grisea* isolates.
Supplementary Figure S9: BGC-O2 is predominantly present in *Oryza*-specific lineage of *P. oryzae*.
Supplementary Figure S10: BGC-TLE is predominantly present in *Triticum*, *Lolium* and *Eleusine*-specific lineages of *P. oryzae*.
Supplementary Figure S11: Comparison of core biosynthetic gene region of BGC-TLE shows the pseudogenization in NRPS gene belonging to *Triticum* lineage.

Supplementary datasets:

Supplementary dataset S1: Amino acid sequences of 2655 BUSCOs present in 68 strains of *P. oryzae*.
Supplementary dataset S2: Amino acid sequence alignments of 2655 BUSCOs present in 68 strains of *P. oryzae*.
Supplementary dataset S3: Best partitioning scheme used to construct Maximum-likelihood tree for *P. oryzae* strains.

Supplementary dataset S4: Maximum-likelihood tree based on 2655 BUSCOs present in 68 strains of *P. oryzae*.

Supplementary dataset S5: Amino acid sequences of 2557 BUSCOs present in 71 strains of *P. oryzae* and *P. grisea*.

Supplementary dataset S6: Amino acid sequence alignments of 2557 BUSCOs present in 71 strains of *P. oryzae* and *P. grisea*.

Supplementary dataset S7: Maximum-likelihood tree based on 2557 BUSCOs present in 71 strains of *P. oryzae* and *P. grisea*.

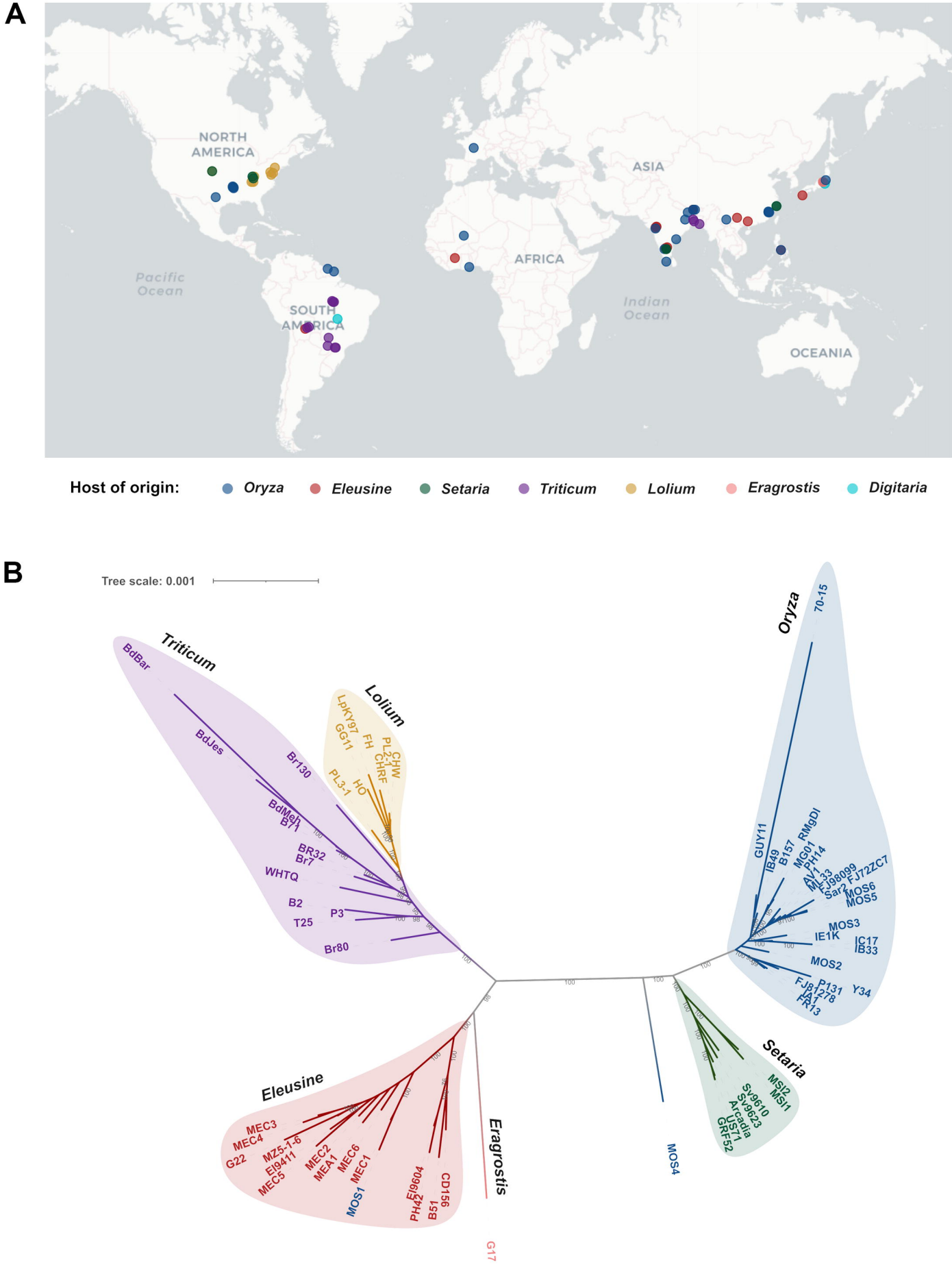
Supplementary dataset S8: Manually curated gene models from *Oryza*-lineage-specific BGC-O1.

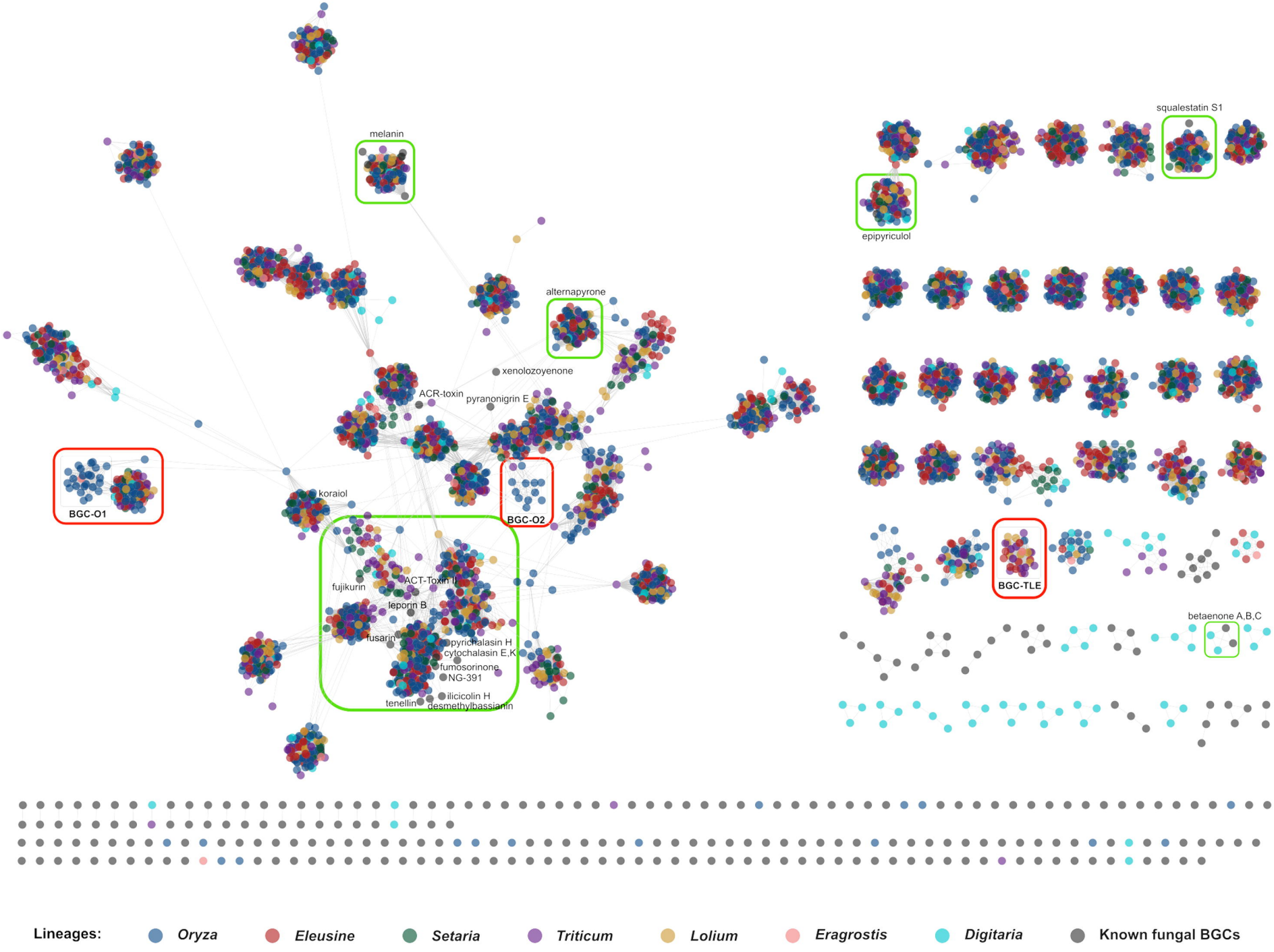
Supplementary dataset S9: Amino acid sequences of rPKS (MGG_08236) homologs in *Pezizomycotina* group.

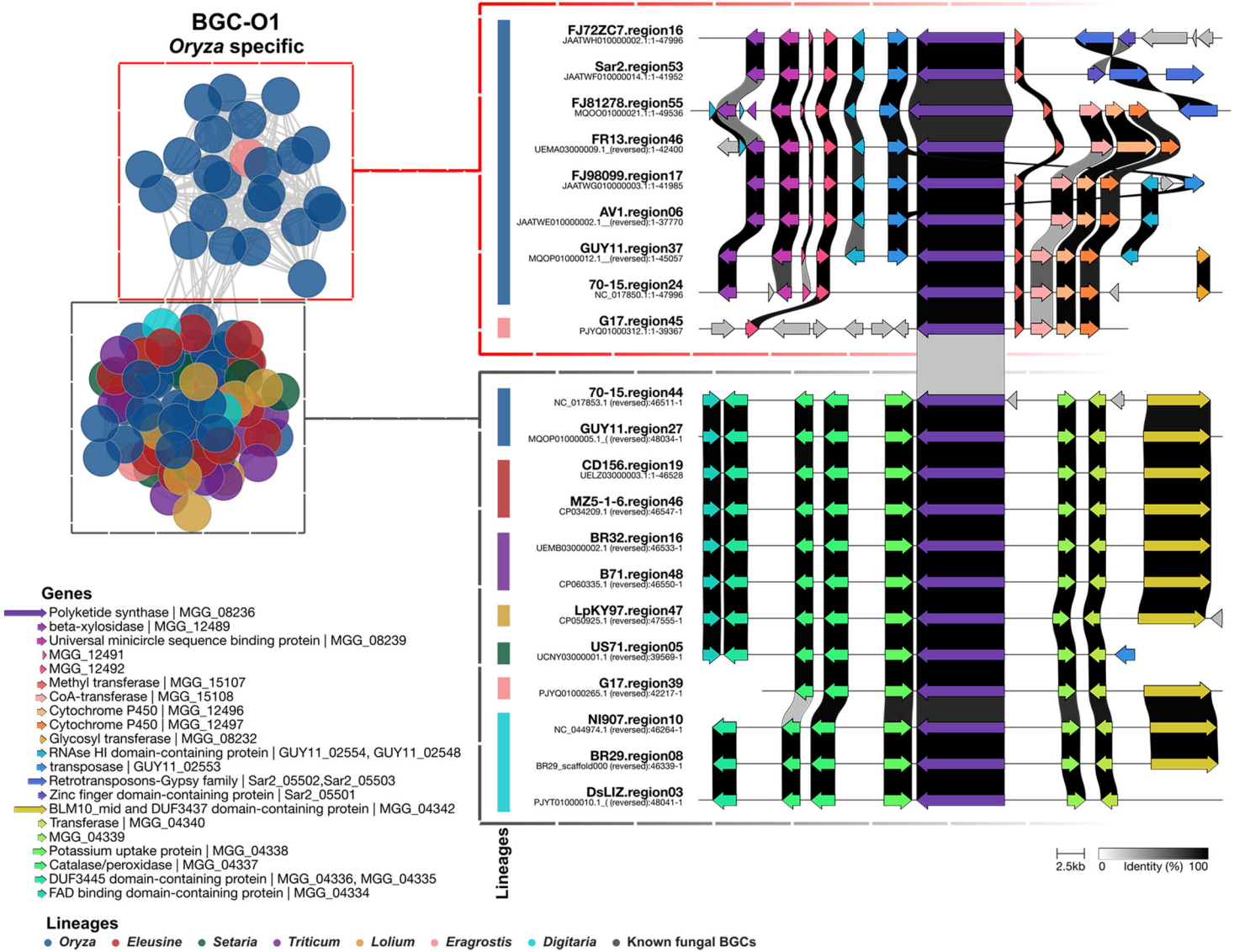
Supplementary dataset S10: Amino acid sequence alignments of rPKS (MGG_08236) homologs in *Pezizomycotina* group.

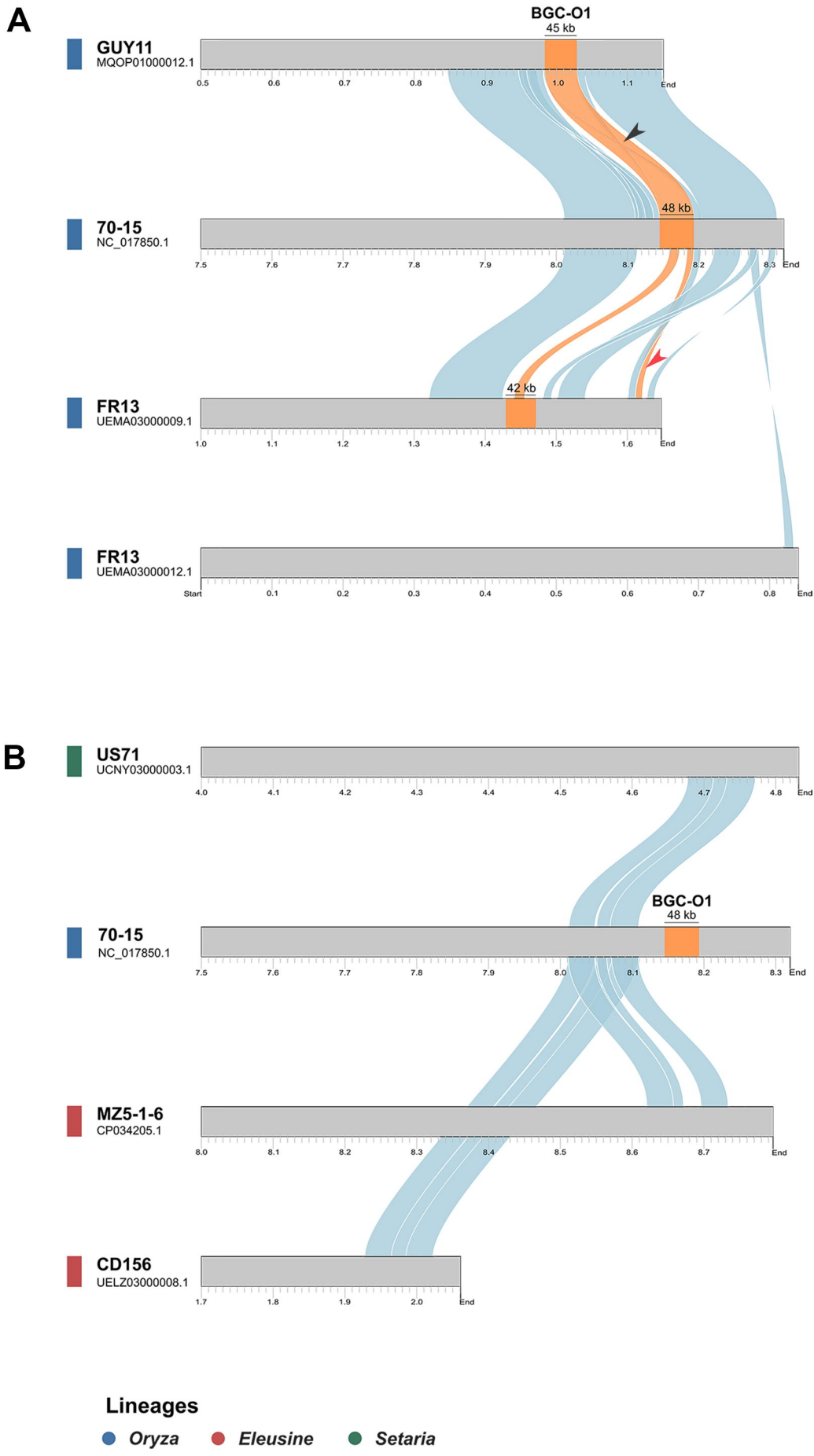
Supplementary dataset S11: Maximum-likelihood tree based on rPKS (MGG_08236) homologs in *Pezizomycotina* group.

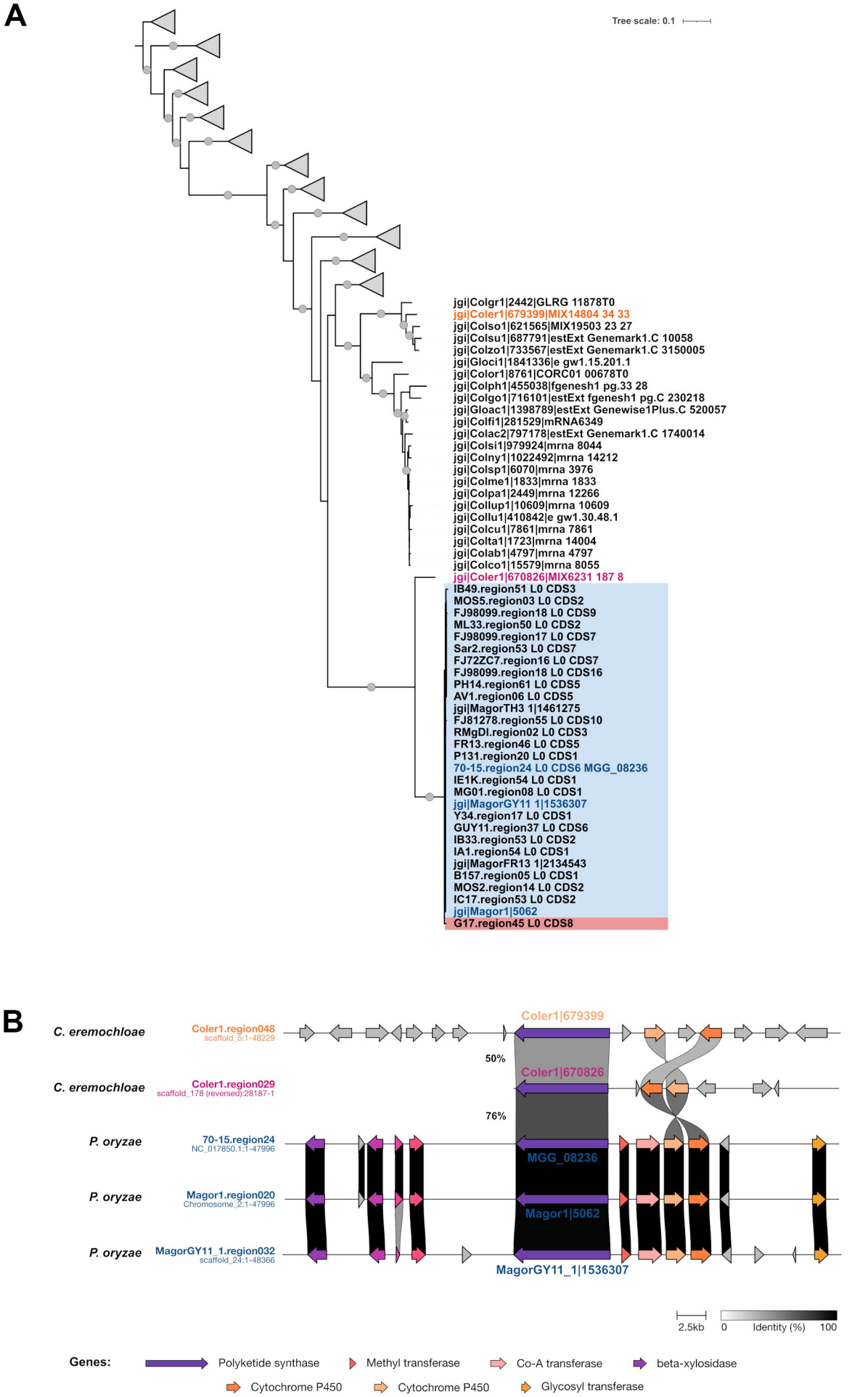
Supplementary dataset S12: Maximum-likelihood tree based on the subset of rPKS (MGG_08236) homologs in *Pezizomycotina* group and closely related cluster of BGC-O1.

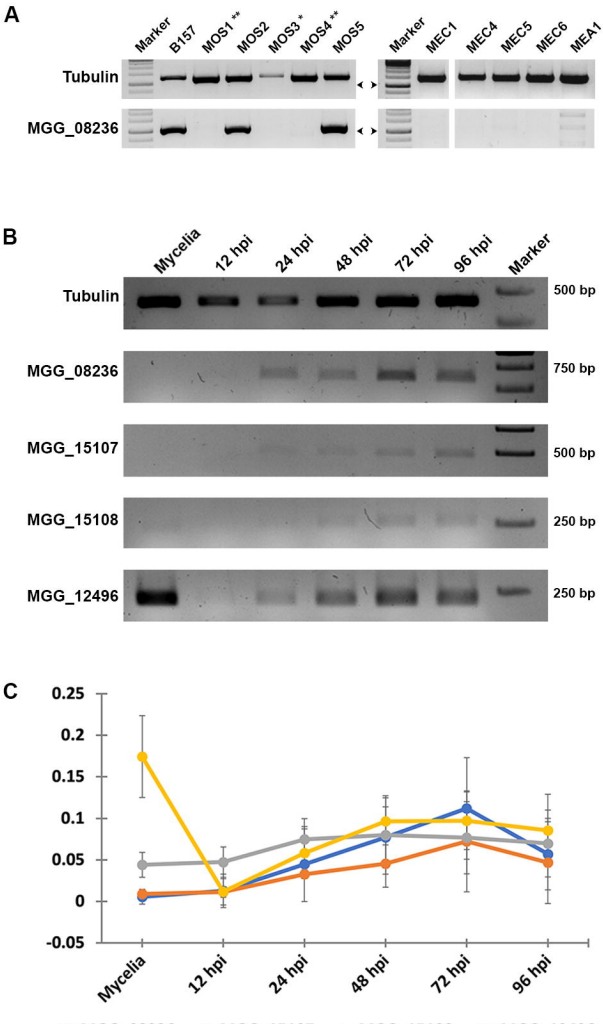

A
NORTH AMERICA
SOUTH AMERICA
AFRICA
ASIA
OCEANIA
Pacific Ocean
Indian Ocean
Host of origin:
Oryza
Eleusine
Setaria
Triticum
Lolium
Eragrostis
Digitaria
B
Tree scale: 0.001
Triticum
BdBar
BdJes
Br130
BdMeh
B71
BR32
Br7
WHTQ
B2
P3
T25
Br80
Lolium
LpKY97
GG11
FH
PL2-1
CHW
CHRF
HO
PL3-1
Oryza
70-15
GUY11
IB49
B157
RMgDI
MG01
PH14
AV1
ML33
FJ98099
Sar2
FJ72ZC7
MOS6
MOS5
MOS3
IE1K
IC17
IB33
MOS2
P131
Y34
FJ81278
IA1
FR13
Setaria
MSI2
MSI1
Sv9610
Sv9623
Arcadia
US71
GRF52
MOS4
Eragrostis
G17
Eleusine
MEC3
MEC4
G22
MZ5-1-6
EI9411
MEC5
MEC2
MEA1
MEC6
MEC1
MOS1
EI9604
PH42
B51
CD156


melanin
alternapyrone
xenolozoyenone
ACR-toxin
pyranonigrin E
BGC-O1
koraiol
BGC-O2
fujikurin
ACT-Toxin II
leporin B
fusarin
pyrichalasin H
cytochalasin E,K
fumosorinone
NG-391
tenellin
ilicicolin H
desmethylbassianin
squalestatin S1
epipyriculol
BGC-TLE
betaenone A,B,C
Lineages:
Oryza
Eleusine
Setaria
Triticum
Lolium
Eragrostis
Digitaria
Known fungal BGCs


BGC-O1
Oryza specific
FJ72ZC7.region16
JAATWH010000002.1:1-47996
Sar2.region53
JAATWF010000014.1:1-41952
FJ81278.region55
MQOO01000021.1:1-49536
FR13.region46
UEMA03000009.1_(reversed):1-42400
FJ98099.region17
JAATWG010000003.1:1-41985
AV1.region06
JAATWE010000002.1__(reversed):1-37770
GUY11.region37
MQOP01000012.1__(reversed):1-45057
70-15.region24
NC_017850.1:1-47996
G17.region45
PJYQ01000312.1:1-39367
70-15.region44
NC_017853.1 (reversed):46511-1
GUY11.region27
MQOP01000005.1_((reversed):48034-1
CD156.region19
UELZ03000003.1:1-46528
MZ5-1-6.region46
CP034209.1 (reversed):46547-1
BR32.region16
UEMB03000002.1 (reversed):46533-1
B71.region48
CP060335.1 (reversed):46550-1
LpKY97.region47
CP050925.1 (reversed):47555-1
US71.region05
UCNY03000001.1 (reversed):39569-1
G17.region39
PJYQ01000265.1 (reversed):42217-1
NI907.region10
NC_044974.1 (reversed):46264-1
BR29.region08
BR29_scaffold000 (reversed):46339-1
DsLIZ.region03
PJYT01000010.1_((reversed):48041-1
Lineages
2.5kb
0
Identity (%)
100
Genes
Polyketide synthase | MGG_08236
beta-xylosidase | MGG_12489
Universal minicircle sequence binding protein | MGG_08239
MGG_12491
MGG_12492
Methyl transferase | MGG_15107
CoA-transferase | MGG_15108
Cytochrome P450 | MGG_12496
Cytochrome P450 | MGG_12497
Glycosyl transferase | MGG_08232
RNAse HI domain-containing protein | GUY11_02554, GUY11_02548
transposase | GUY11_02553
Retrotransposons-Gypsy family | Sar2_05502,Sar2_05503
Zinc finger domain-containing protein | Sar2_05501
BLM10_mid and DUF3437 domain-containing protein | MGG_04342
Transferase | MGG_04340
MGG_04339
Potassium uptake protein | MGG_04338
Catalase/peroxidase | MGG_04337
DUF3445 domain-containing protein | MGG_04336, MGG_04335
FAD binding domain-containing protein | MGG_04334
Lineages
Oryza
Eleusine
Setaria
Triticum
Lolium
Eragrostis
Digitaria
Known fungal BGCs


A
BGC-O1
45 kb
GUY11
MQOP01000012.1
0.5
0.6
0.7
0.8
0.9
1.0
1.1
End
48 kb
70-15
NC_017850.1
7.5
7.6
7.7
7.8
7.9
8.0
8.1
8.2
8.3
End
42 kb
FR13
UEMA03000009.1
1.0
1.1
1.2
1.3
1.4
1.5
1.6
End
FR13
UEMA03000012.1
Start
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
End
B
US71
UCNY03000003.1
4.0
4.1
4.2
4.3
4.4
4.5
4.6
4.7
4.8
End
BGC-O1
48 kb
70-15
NC_017850.1
7.5
7.6
7.7
7.8
7.9
8.0
8.1
8.2
8.3
End
MZ5-1-6
CP034205.1
8.0
8.1
8.2
8.3
8.4
8.5
8.6
8.7
End
CD156
UELZ03000008.1
1.7
1.8
1.9
2.0
End
Lineages
Oryza
Eleusine
Setaria


A
Tree scale: 0.1
jgi|Colgr1|2442|GLRG 11878T0
jgi|Coler1|679399|MIX14804 34 33
jgi|Colso1|621565|MIX19503 23 27
jgi|Colsu1|687791|estExt Genemark1.C 10058
jgi|Colzo1|733567|estExt Genemark1.C 3150005
jgi|Gloci1|1841336|e gw1.15.201.1
jgi|Color1|8761|CORC01 00678T0
jgi|Colph1|455038|fgenesh1 pg.33 28
jgi|Colgo1|716101|estExt fgenesh1 pg.C 230218
jgi|Gloac1|1398789|estExt Genewise1Plus.C 520057
jgi|Colfi1|281529|mRNA6349
jgi|Colac2|797178|estExt Genemark1.C 1740014
jgi|Colsi1|979924|mrna 8044
jgi|Colny1|1022492|mrna 14212
jgi|Colsp1|6070|mrna 3976
jgi|Colme1|1833|mrna 1833
jgi|Colpa1|2449|mrna 12266
jgi|Collup1|10609|mrna 10609
jgi|Collu1|410842|e gw1.30.48.1
jgi|Colcu1|7861|mrna 7861
jgi|Colta1|1723|mrna 14004
jgi|Colab1|4797|mrna 4797
jgi|Colco1|15579|mrna 8055
jgi|Coler1|670826|MIX6231 187 8
IB49.region51 L0 CDS3
MOS5.region03 L0 CDS2
FJ98099.region18 L0 CDS9
ML33.region50 L0 CDS2
FJ98099.region17 L0 CDS7
Sar2.region53 L0 CDS7
FJ72ZC7.region16 L0 CDS7
FJ98099.region18 L0 CDS16
PH14.region61 L0 CDS5
AV1.region06 L0 CDS5
jgi|MagorTH3 1|1461275
FJ81278.region55 L0 CDS10
RMgDI.region02 L0 CDS3
FR13.region46 L0 CDS5
P131.region20 L0 CDS1
70-15.region24 L0 CDS6 MGG_08236
IE1K.region54 L0 CDS1
MG01.region08 L0 CDS1
jgi|MagorGY11 1|1536307
Y34.region17 L0 CDS1
GUY11.region37 L0 CDS6
IB33.region53 L0 CDS2
IA1.region54 L0 CDS1
jgi|MagorFR13 1|2134543
B157.region05 L0 CDS1
MOS2.region14 L0 CDS2
IC17.region53 L0 CDS2
jgi|Magor1|5062
G17.region45 L0 CDS8
B
C. eremochloae
Coler1.region048
scaffold_5:1-48229
Coler1|679399
50%
C. eremochloae
Coler1.region029
scaffold_178 (reversed):28187-1
Coler1|670826
76%
P. oryzae
70-15.region24
NC_017850.1:1-47996
MGG_08236
P. oryzae
Magor1.region020
Chromosome_2:1-47996
Magor1|5062
P. oryzae
MagorGY11_1.region032
scaffold_24:1-48366
MagorGY11_1|1536307
2.5kb
0
Identity (%)
100
Genes:
Polyketide synthase
Methyl transferase
Co-A transferase
beta-xylosidase
Cytochrome P450
Cytochrome P450
Glycosyl transferase


A
Marker
B157
MOS1**
MOS2
MOS3*
MOS4**
MOS5
Marker
MEC1
MEC4
MEC5
MEC6
MEA1
Tubulin
MGG_08236
B
Mycelia
12 hpi
24 hpi
48 hpi
72 hpi
96 hpi
Marker
Tubulin
500 bp
MGG_08236
750 bp
MGG_15107
500 bp
MGG_15108
250 bp
MGG_12496
250 bp
C
0.25
0.2
0.15
0.1
0.05
0
-0.05
Mycelia
12 hpi
24 hpi
48 hpi
72 hpi
96 hpi
MGG_08236
MGG_15107
MGG_15108
MGG_12496