

RESEARCH ARTICLE

# Biosynthesis of Antibiotic Leucinostatins in Bio-control Fungus *Purpureocillium lilacinum* and Their Inhibition on *Phytophthora* Revealed by Genome Mining

Gang Wang<sup>1</sup>, Zhiguo Liu<sup>2</sup>, Runmao Lin<sup>1,3</sup>, Erfeng Li<sup>1</sup>, Zhenchuan Mao<sup>1</sup>, Jian Ling<sup>1</sup>, Yuhong Yang<sup>1</sup>, Wen-Bing Yin<sup>2\*</sup>, Bingyan Xie<sup>1\*</sup>

**1** Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences, Beijing, PR China, **2** State Key Laboratory of Mycology, Institute of Microbiology, Chinese Academy of Sciences, Beijing, PR China, **3** College of Life Sciences, Beijing Normal University, Beijing, PR China

☉ These authors contributed equally to this work.

\* [yinwb@im.ac.cn](mailto:yinwb@im.ac.cn) (WBY); [xiebingyan@caas.cn](mailto:xiebingyan@caas.cn) (BX)



CrossMark  
click for updates

 OPEN ACCESS

**Citation:** Wang G, Liu Z, Lin R, Li E, Mao Z, Ling J, et al. (2016) Biosynthesis of Antibiotic Leucinostatins in Bio-control Fungus *Purpureocillium lilacinum* and Their Inhibition on *Phytophthora* Revealed by Genome Mining. *PLoS Pathog* 12(7): e1005685. doi:10.1371/journal.ppat.1005685

**Editor:** Xiaorong Lin, Texas A&M University, UNITED STATES

**Received:** February 21, 2016

**Accepted:** May 18, 2016

**Published:** July 14, 2016

**Copyright:** © 2016 Wang et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All genome sequences files are available from the GenBank database (accession numbers LSBH00000000, LSBI00000000 and LSBJ00000000).

**Funding:** This work was supported by Special Fund for Agro-scientific Research in the Public Interest (201103018) and the Earmarked fund for China Agriculture Research System (CARS-25-B-01). BX had received the fundings. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Abstract

*Purpureocillium lilacinum* of Ophiocordycipitaceae is one of the most promising and commercialized agents for controlling plant parasitic nematodes, as well as other insects and plant pathogens. However, how the fungus functions at the molecular level remains unknown. Here, we sequenced two isolates (PLBJ-1 and PLFJ-1) of *P. lilacinum* from different places Beijing and Fujian. Genomic analysis showed high synteny of the two isolates, and the phylogenetic analysis indicated they were most related to the insect pathogen *Tolypocladium inflatum*. A comparison with other species revealed that this fungus was enriched in carbohydrate-active enzymes (CAZymes), proteases and pathogenesis related genes. Whole genome search revealed a rich repertoire of secondary metabolites (SMs) encoding genes. The non-ribosomal peptide synthetase *LcsA*, which is comprised of ten C-A-PCP modules, was identified as the core biosynthetic gene of lipopeptide leucinostatins, which was specific to *P. lilacinum* and *T. ophioglossoides*, as confirmed by phylogenetic analysis. Furthermore, gene expression level was analyzed when PLBJ-1 was grown in leucinostatin-inducing and non-inducing medium, and 20 genes involved in the biosynthesis of leucinostatins were identified. Disruption mutants allowed us to propose a putative biosynthetic pathway of leucinostatin A. Moreover, overexpression of the transcription factor *lcsF* increased the production (1.5-fold) of leucinostatins A and B compared to wild type. Bioassays explored a new bioactivity of leucinostatins and *P. lilacinum*: inhibiting the growth of *Phytophthora infestans* and *P. capsici*. These results contribute to our understanding of the biosynthetic mechanism of leucinostatins and may allow us to utilize *P. lilacinum* better as bio-control agent.

**Competing Interests:** The authors have declared that no competing interests exist.

## Author Summary

*Purpureocillium lilacinum*, a well-known bio-control agent against various plant pathogens in agriculture, can produce antibiotic leucinostatins—peptaibiotic with extensive biological activities, including antimalarial, antiviral, antibacterial, antifungal, and antitumor activities, as well as phytotoxic. We have sequenced the genomes of two *P. lilacinum* isolates, and compared them with other fungi, focusing on their bio-control characteristics. We discovered a rich repertoire of CAZymes, proteases, SMs and pathogenesis related genes. We also identified a gene cluster containing 20 genes involved in the leucinostatins A and B biosynthesis by gene deletion, qRT-PCR and RNA-seq analyses. A transcription factor in the pathway was overexpressed, resulting in the upregulation of the related genes and a 1.5-fold increase in leucinostatins A and B. A new bioactivity of leucinostatins, inhibition of the growth of the notorious *Phytophthora*, was identified in this study by confronting incubation with *P. lilacinum*. These results provided new strategies for the agricultural development of leucinostatins and improving *P. lilacinum* strains.

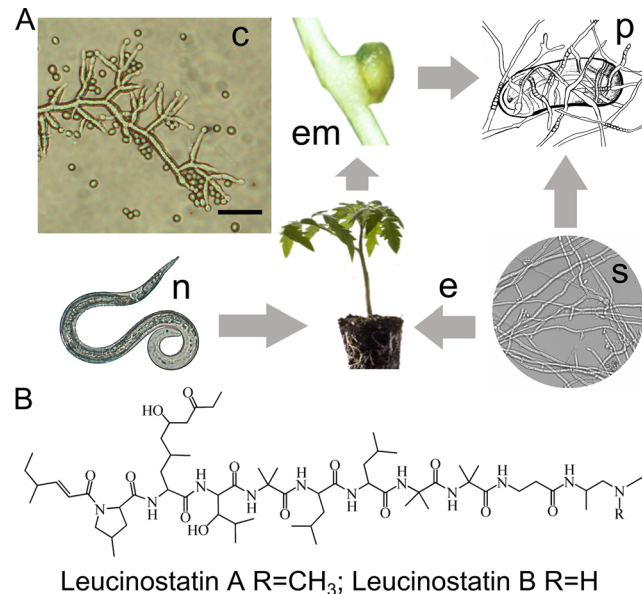
## Introduction

Plant parasitic nematodes with wide host ranges cause enormous crop and economic losses amounting to \$157 billion annually worldwide [1, 2]. Biological control by fungi has become increasingly popular due to nematicides' risks of environmental toxicity and adverse effects on human health [3]. One of the most promising and commercialized agents, *Purpureocillium lilacinum*, has been evaluated to assess its bio-control activity against plant nematodes in a number of studies [2, 4]. In particular, *P. lilacinum* has been reported to effectively control such species as the cotton aphid *Aphis gossypii* [5], the greenhouse whitefly *Trialeurodes vaporariorum*, the glasshouse red spider mite *Tetranychus urticae* [6], and the leaf-cutting ant *Acromyrmex lundii* [7].

The genus *Purpureocillium* was recently proposed for of Ophiocordycipitaceae, based on the internal transcribed spacer (ITS) and translation elongation factor 1- $\alpha$  (TEF) sequences of *P. lilacinum*, although it was originally classified in the genus *Paecilomyces* [8]. *P. lilacinum* is commonly isolated from soil, plant roots, nematodes and insects, and it occasionally infects people. This fungus employs flexible lifestyles, including soil-saprobies, plant-endophytes and nematode pathogens. Opportunistic infection occurs when nematode eggs encounter *P. lilacinum*; therefore, parasitism can be a mechanism for nematode bio-control (Fig 1A). It has now been confirmed that a serine protease [9], a cuticle-degrading protease [10] and chitinase [11] play important roles in infection by degrading nematode eggshells.

Recently, the production of SMs has been shown to be a mechanism that kills nematodes. For example, culture filtrates of *P. lilacinum*, in which leucinostatins were produced, caused strong mortality and inhibited nematode reproduction [12]. In addition to leucinostatins, a few other SMs have been isolated from *P. lilacinum*. The novel pyridone alkaloid paecilomide, an acetylcholinesterase inhibitor, was produced when this fungus was co-cultured with *Salmonella typhimurium* [13]. Two xanthone-anthraquinone heterodimers, acremoxanthone C and acremonidin A, were isolated in the course of a search for calmodulin ligands [14].

The leucinostatins (Fig 1B) are a family of lipopeptide antibiotics isolated from *P. lilacinum* [15], *Paecilomyces marquandii* [16–18] and *Acremonium* sp. [19]. Leucinostatin A contains nine amino acid residues, including the unusual amino acid 4-methyl-L-proline (MePro), 2-amino-6-hydroxy-4-methyl-8-oxodecanoic acid (AHyMeOA), hydroxyleucine (HyLeu),  $\alpha$ -aminoisobutyric acid (AIB),  $\beta$ -Ala, and a 4-methylhex-2-enoic acid at the N-terminus as well



**Fig 1. Lifestyles of nematophagous *P. lilacinum* and the structures of leucinostatins.** (A) Microscopic conidiophores and conidia (c) of *P. lilacinum*. Scale bar = 10  $\mu$ m. The soil saprophyte (s) *P. lilacinum* colonizes plant roots as an endophyte (e), and the parasite (p) can occur in nematode eggs in the egg mass (em) generated after the infection with the plant nematode (n). (B) Chemical structure of leucinostatins A and B.

doi:10.1371/journal.ppat.1005685.g001

as an N1,N1-dimethylpropane-1,2-diamine (DPD) at the C-terminus. Twenty-four different structures have been described in the leucinostatin series [20]. Leucinostatin A significantly suppressed prostate cancer growth in a coculture system in which prostate stromal cells stimulated the growth of DU-145 human prostate cancer cells through insulin-like growth factor I [21]. When screening for antitrypanosomal compounds among several peptide antibiotics, leucinostatins showed the most potent activity against trypanosomes. Trypanosome infection causes human African trypanosomiasis, which is one of the world's most neglected diseases lacking satisfactory drugs [22]. Furthermore, leucinostatins have displayed broad bioactivity against bacteria and fungi. These antibiotics' functions are based on their ability to inhibit ATP synthesis in the mitochondria as well as different phosphorylation pathways [23]. These findings drew our attention to the relationships between the bio-control function of *P. lilacinum* and leucinostatins. Furthermore, genetic and molecular information regarding the biosynthesis of this family of lipopeptide antibiotics, of which little was known to date, could contribute to increasing its production and screening for more efficient derivative compounds.

Genome sequences have shed light on the mechanism of the endoparasitic lifestyle or nematode control beyond biological research. During the preparation of our manuscript, the genome sequence of *P. lilacinum* was published [24]. Two other plant nematode endoparasitic fungi, *Pochonia chlamydosporia* [25] and *Hirsutella minnesotensis* [26], were recently sequenced. Genome sequencing revealed that *P. chlamydosporia* encoded a wide array of hydrolytic enzymes and transporters expressed at the mRNA level, which supported its multitrophic lifestyle, and *H. minnesotensis*, which mainly invades juvenile stage cyst nematodes, putatively conducted its parasitic process through lectins, secreted proteases and SMs. Thus, the genome sequence of *P. lilacinum* provides an opportunity to better understand its mechanism in controlling plant nematodes, and it would be useful to enhance its capabilities as a bio-control agent. At the same time, the genome sequence has the potential to solve the biosynthetic puzzle

of leucinostatins as well as to detect novel genes and metabolites that might be of value in agriculture and medicine.

Here, we present the results of genome sequencing of the PLBJ-1 and PLFJ-1 strains of the bio-control agent *P. lilacinum*, and we increased our knowledge of its bio-control capabilities by comparing the sequences of *P. lilacinum* with those of other fungi. The genome revealed a repertoire of SM-encoding genes that illustrated the potential for using this fungus to discover natural products. Furthermore, we identified the leucinostatin gene cluster (*lcs* cluster) and proposed a hypothetical pathway for biosynthesis through genetic manipulation. In the course of screening for new activities of leucinostatins, we found that they inhibited the most notorious oomycetes *P. infestans*, which causes potato late blight and results in global yield losses of 16% [27].

## Results

### General structure of the *P. lilacinum* genome

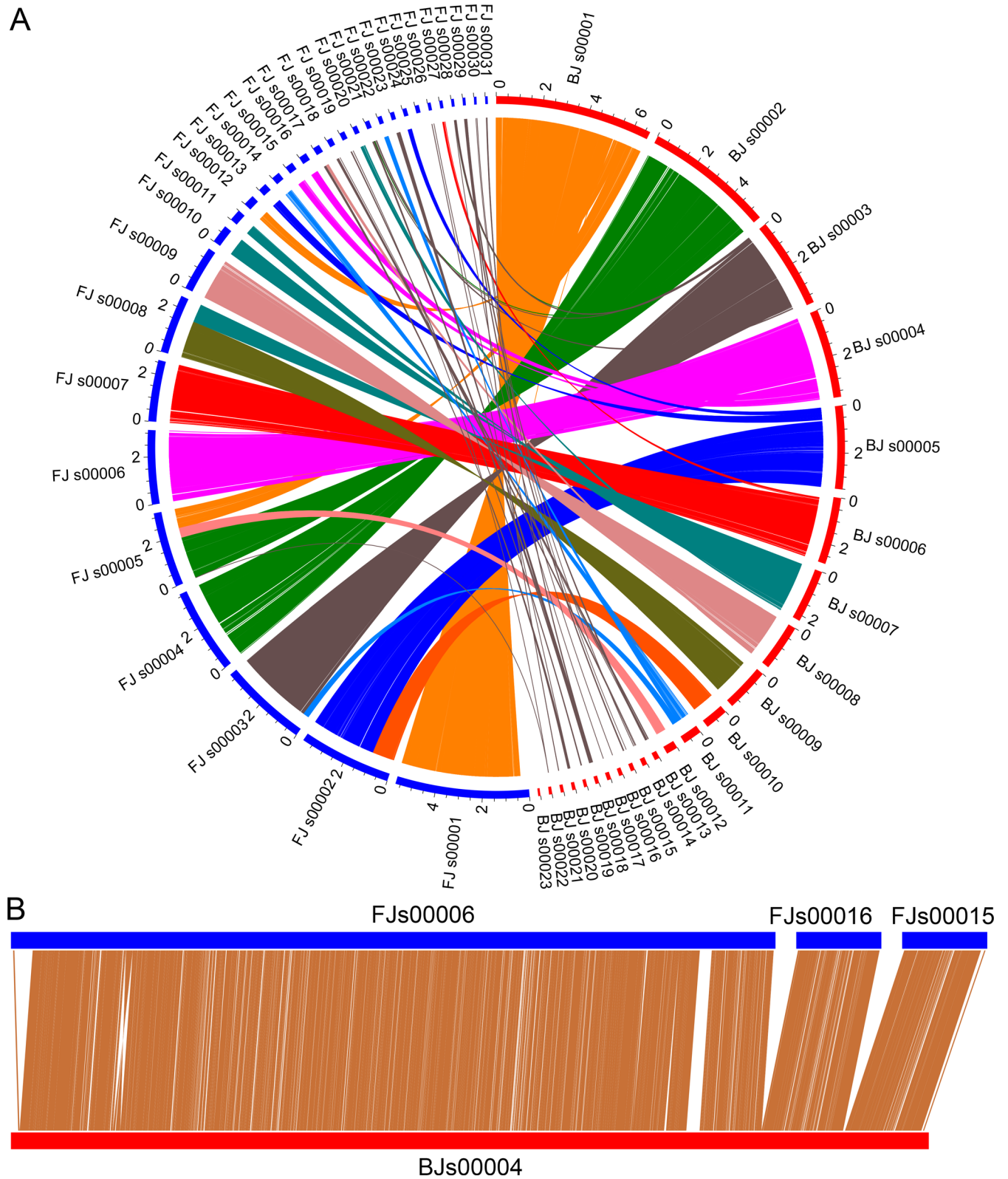
Two *P. lilacinum* isolates, PLBJ-1 and PLFJ-1, were sequenced to ensure the accuracy of the genome information and the subsequent analysis. PLBJ-1 and PLFJ-1 were assembled into 144 and 163 scaffolds, respectively, with total sizes of 38.14 and 38.53 Mb, while the published TERIBC 1 was assembled into 301 scaffolds with a total size of 38.82 Mb (Table 1). The comparative genome sizes of related fungi species are listed in S1 Table. A total of 11,773 and 11,763 gene models were predicted in both genomes, respectively, parallel to other ascomycetes fungi (S1 Table). BLASTN analysis was performed between the two genomes and demonstrated that 83.56% of the PLBJ-1 genome and 82.79% of the PLFJ-1 genome shared high synteny (Fig 2A). According to the syntenic relationship of PLBJ-1 and PLFJ-1, we reconstructed 10 super-scaffolds (S2 Table), which illustrated the physical ubiquities of the assembled scaffolds; e.g., scaffold 00006, scaffold 00016 and scaffold 00015 in PLFJ-1 were combined into a super-scaffold (Fig 2B). The overall syntenic relationship of PLBJ-1 and TERIBC 1 showed that 76.52% of the PLBJ-1 genome and 75.12% of the TERIBC 1 genome shared high synteny (S1 Fig). Approximately 6.07% of the repeat sequences that included transposon elements (TEs) (~4.37%) and tandem repeats (~1.70%) were identified in PLBJ-1. The Class I

**Table 1. Genome feature of the three *P. lilacinum* isolates.**

Feature	PLBJ-1	PLFJ-1	TERIBC 1
Accession number	LSBH00000000	LSBI00000000	LOFA00000000.1
Length (Mb)	38.14	38.53	38.82
Scaffold number	144	163	301
Contig number	596	818	301
Max scaffold length (Mb)	6.69	5.70	3.7
Scaffold N50	3.77	3.20	1.79
G+C contents (%)	58.5	58.4	58.6
Repeat sequence (%)	6.07	6.00	1.68
Protein-coding genes	11,773	11,763	13,266
tRNA genes	90	90	91
Secreted proteins	1410	1448	1276
Glycoside hydrolases	239	253	249
Carbohydrate esterases	32	32	106
Protease	430	443	480
PHI genes	2844	2892	1953
SMs	41	41	46

Strains PLBJ-1 and PLFJ-1 of *P. lilacinum* are obtained in this study, strain TERIBC 1 was sequenced in [24] recently.

doi:10.1371/journal.ppat.1005685.t001



**Fig 2. Genomic synteny of PLBJ-1 and PLFJ-1.** (A) The syntenic genome sequences of PLBJ-1 and PLFJ-1 were analyzed by BLASTN, with an E-value cutoff of  $1e-5$ . The red semicircle represents the scaffolds of PLBJ-1, while the blue semicircle represents the scaffolds of PLFJ-1. Scaffold lengths of  $\geq 100$  Kb were used for this analysis, and the threshold of matched blocks was  $\geq 1000$  bp, which are connected by lines of the same color. (B) An example of a super-scaffold inferred by syntenic analysis.

doi:10.1371/journal.ppat.1005685.g002

(retrotransposons) and Class II (DNA transposons) TEs occupied ~1.80% and ~0.76% of the genome, respectively. The PLFJ-1 isolate harbored a similar number of repeat sequences (6.00%). The distribution of the TE families was similar in the two isolates, with the exception of certain families, e.g., I, Gypsy, Penelope, Tc1-Mariner and hAT (S3 Table). In total, the two isolates of *P. lilacinum* contained a larger number of retrotransposons than DNA transposons. *P. lilacinum* exhibited expansion of repeat content comparable to other ascomycetes fungi, with the exception of *H. minnesotensis*, *Ophiocordyceps sinensis* and *Fusarium oxysporum* (fol), in which the repeat sequences accounted for more than one quarter of the genome (S1 Table). In TERIBC 1, approximately 1.68% of the genome sequence was identified as repeat content.

Among the predicted genes of PLBJ-1, 90.4% were supported by RNASeq data from mycelia cultured in PDB. Both strains exhibited a consistent KOG pattern. Except for the category “General function prediction only”, which was ambiguously sorted to a certain group, the most abundant KOG categories were “Signal transduction mechanisms”, “Posttranslational modification, protein turnover, chaperones”, “Lipid transport and metabolism”, and “Intracellular trafficking, secretion, and vesicular transport” (S2 Fig). A signal peptide analysis showed that 1,410 genes of PLBJ-1 and 1,448 genes of PLFJ-1 encoded putatively secreted proteins.

CAZymes that cleave and build polysaccharides could be required when *P. lilacinum* degraded the structural polysaccharide armor of nematode eggshells, such as chitin, during the course of its parasitism. The protease could stop the development of nematode eggs and drastically alter the eggshell structures when applied individually or in combination with chitinases [28, 29]. A detailed examination of the CAZymes and proteases of *P. lilacinum* was performed and compared with other fungi, including nematode parasitic fungi (*P. chlamydosporia* and *H. minnesotensis*), nematode-trapping fungi (*Arthrobotrys oligospora* and *Monacrosporium haptotylum*), entomopathogenic fungi (*T. inflatum*, *Beauveria bassiana*, *Cordyceps militaris*, *Metarhizium robertsii*, and *O. sinensis*), a mycoparasitic fungus (*T. ophioglossoides*), a saprotrophic fungus (*T. reesei*) and a plant pathogenic fungus (*F. oxysporum*). We identified 53 families containing 239 genes in PLBJ-1 and 55 families containing 253 genes in PLFJ-1 that encoded glycoside hydrolases (GH), which was more than the other fungi (an average of 213) (S4 Table). The most abundant family in PLBJ-1 and PLFJ-1 was GH18, which was represented by 32 and 41 chitinases, respectively, that degrade the chitin present in the chitin protein complex of the nematode eggshell [30]. Consistent with GHs, PLBJ-1 and PLFJ-1 contained relatively more carbohydrate-binding modules (CBMs) (59 and 64, respectively) (S5 Table), which were frequently appended to the enzymes involved in polysaccharide depolymerization. A series of carbohydrate esterase (CE)-encoding genes were also detected in the *P. lilacinum* genomes (33 and 32, respectively), including the most abundant sterol esterases (CE10) and cutinases (CE5), which are virulence factors of some plant pathogens [31] (S6 Table). Another major class of CAZymes, the glycosyltransferases (GT), establish natural glycosidic linkages across a broad range of small and macromolecules, and they were represented in the PLBJ-1 genome with 115 members in 32 families and in the PLFJ-1 genome with 124 members in 32 families (S7 Table). These enzymes’ classification demonstrated that they exhibited less variability in ascomycetes than did GHs, a trend that was maintained in a previous analysis [32]. The *P. lilacinum* genome contained more proteases (430 and 443, respectively) than other fungi (an average of 396). The largest category of proteases encoded in PLBJ-1 and PLFJ-1 were serine proteases (194 and 198, respectively) (S8 Table), 76 and 81 of which were secreted proteins, respectively. Among the serine proteases, we identified 34 subtilisins (S8) and ten serine carboxyproteases (S10) in the PLBJ-1 genome (36 and 11, respectively, in PLFJ-1), which were reported to be involved in infection and the lethal activity of nematodes [28, 33]. The metalloprotease (108 in PLBJ-1 and 109 in PLFJ-1) and cysteine protease (66 in PLBJ-1 and 68 in PLFJ-1) families also accounted for a significant proportion of the proteases.

A whole genome analysis was conducted against the pathogen-host interaction (PHI) gene database to identify potential virulence-associated genes, under the assumption that the homologue of an experimentally validated pathogenic gene suggested that it played a pathogenic role [34]. We demonstrated that 2,844 (24.1%) and 2,892 (24.6%) proteins of PLBJ-1 and PLFJ-1, respectively, showed sequence similarity to those in the PHI database. Among these proteins, 299 and 317 proteins of PLBJ-1 and PLFJ-1, respectively, were classified as putatively secreted proteins. The KOG functional class distribution of genes related to PHI showed a similar pattern to the whole genome KOG analysis (S2 Fig). The PHI database search yielded 195 CAZymes in PLBJ-1 and 217 in PLFJ-1, 28 and 36 of which were chitinases (GH18), respectively. Of the proteases, 125 in PLBJ-1 and 132 in PLFJ-1 were pathogenic genes according to the PHI database, of which 64 and 72 were identified as secreted proteins, respectively, and these proteins were more likely to function during the infection process [35].

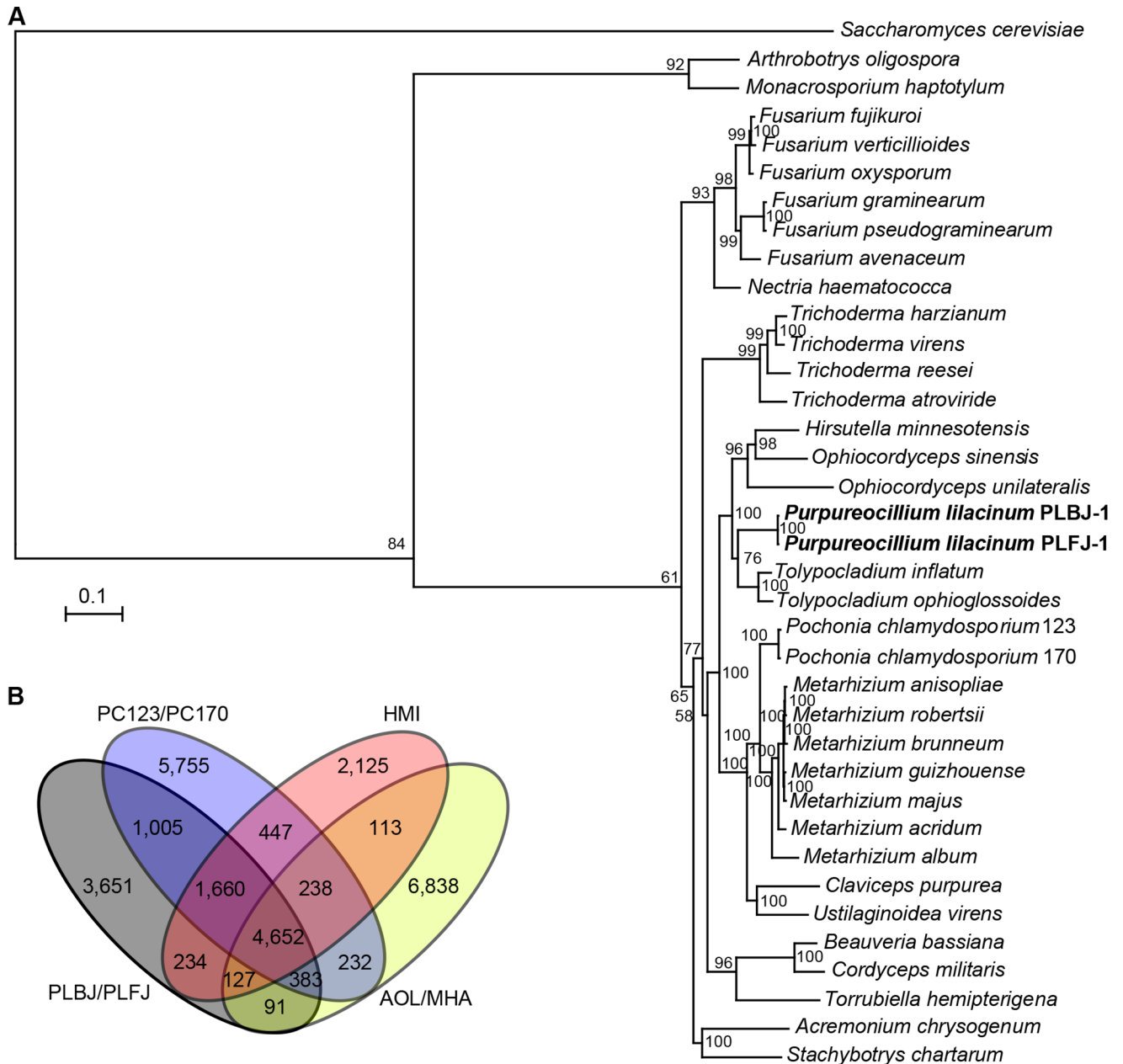
### Phylogenomic relationship and orthologous analysis

A phylogenomic tree was constructed based on 855 single-copy orthologues of *P. lilacinum* and 34 other filamentous fungi, with *Saccharomyces cerevisiae* as the outgroup. The results verified that *P. lilacinum* belongs to Ophiocordycipitaceae, as described by Jennifer Luangsa-ard [8], and it formed a clade with *T. inflatum*, *T. ophioglossoides*, *O. sinensis* [36], *O. unilateralis* [37] and *H. minnesotensis* (Fig 3A). The inferred phylogeny illustrated that *T. inflatum* and *T. ophioglossoides* were most closely related to *P. lilacinum*, and they diverged after their split with *O. sinensis*, *H. minnesotensis* and *O. unilateralis*. This phylogeny also reinforced the previous analysis that found that the split between Cordycipitaceae (including *B. bassiana* and *C. militaris*) and Clavicipitaceae (including *P. chlamydosporium* and *M. anisopliae*) occurred before Ophiocordycipitaceae diverged from Clavicipitaceae (Fig 3A). The three nematode parasitic fungi *P. chlamydosporium*, *H. minnesotensis* and *P. lilacinum* clustered with insect pathogens, indicating that nematode and insect pathogens might share a common ancestor.

A comparative genomic analysis was performed between *P. lilacinum* and other nematode-related fungi (the nematode parasites *P. chlamydosporia* and *H. minnesotensis* and the nematode-trapping fungi *A. oligospora* and *M. haptotylum*). A total of 17,995 orthologous clusters consisting of 76,151 proteins were identified, of which 4,652 clusters containing 35,972 proteins were mapped to all four of the fungi types (Fig 3B). On the whole, the nematode-trapping fungi, which capture nematodes through an entirely different mechanism compared to *P. lilacinum* [26], possessed the largest number of unique gene clusters, although they had a more distant phylogenetic relationship with the other fungi in Hypocreales (Fig 3B). *P. lilacinum* contained a large number (3651) of species-specific clusters, while *P. lilacinum* shared 7,700, 6,673 and 5,253 clusters with *P. chlamydosporia*, *H. minnesotensis* and the nematode-trapping fungi, respectively.

### Analysis of paralogous gene families

Lineage-specific expansions could provide material for the evolution of a specific functional system or adaptation in eukaryotes [38]. To study gene family expansions in *P. lilacinum*, a comparative genomic analysis of 15 fungal species (PLBJ-1, PLFJ-1, *P. chlamydosporia* strain 123, *P. chlamydosporia* strain 170, *H. minnesotensis*, *A. oligospora*, *M. haptotylum*, *T. inflatum*, *B. bassiana*, *C. militaris*, *M. robertsii*, *O. sinensis*, *T. ophioglossoides*, *T. reesei*, and *F. oxysporum*) was performed. In total, 1,963 gene families with more than one gene expansion were identified in both PLBJ-1 and PLFJ-1, of which 1,761 gene families were only present in *P. lilacinum*, and some gene families with significant expansion are listed in S9 Table. However, most families were annotated as reverse transcriptases and transposases, and the others were related to



**Fig 3. Phylogenomic relationships and orthologous gene clusters.** (A) Maximum likelihood phylogeny was computed from a concatenated alignment of 855 groups of single-copy orthologues. Bootstrap values are shown beside the nodes. (B) The number of gene clusters shared by *P. lilacinum* with other major associated ecologies. Gray = *P. lilacinum* isolates PLBJ-1 and PLFJ-1; blue = nematode egg parasite *P. chlamydosporia* isolates 123 and 170; pink = nematode parasite *H. minnesotensis*; and yellow = nematode-trapping fungi *A. oligospora* and *M. haptotylum*.

doi:10.1371/journal.ppat.1005685.g003

transporters or lyases. When the nematode parasitic fungi *P. chlamydosporium* and *H. minnesotensis* were considered, 2,936 orthologous clusters showed expansion in the five isolates. The largest paralogous expansion contained protein families associated with SMs, such as cytochrome P450s, oxidoreductases, and transporters. In addition, these families also contained transcription factors, glycosyl hydrolases, the hAT family, the majority of which are listed in [S10 Table](#).



## SMs analysis based on the sequence of the *P. lilacinum* genome

To evaluate the capability of *P. lilacinum* to produce SMs, we searched the genome of PLBJ-1 and PLFJ-1 for biosynthetic genes encoding the four classes of the main SM-associated synthetases, including polyketide synthase (PKS), non-ribosomal peptide synthetase (NRPS), terpene synthase (TS) and dimethylallyl tryptophan synthase (DMATS) [26]. A uniform SM profile with parallel categories and numbers was presented in the two genomes (S11 Table). In total, 13 PKSs, 10 NRPSs, two PKS-like enzymes, 10 NRPS-like enzymes, one DMATS, 4 TSs and one PKS-NRPS hybrid were identified in the PLBJ-1 genome, as described in S11 Table. Compared to sequenced species in Ophiocordycipitaceae, the number of SMs in *P. lilacinum* (41) was similar to the 45 SMs in *T. ophioglossoides*, 39 SMs in *O. unilateralis*, more than 30 SMs in *Ophiocordyceps sinensis*, fewer than 55 SMs in *T. inflatum* [39], and 101 SMs in the nematode endoparasitic fungus *H. minnesotensis* [26]. These core backbone genes were dispersed among 39 clusters with other enzymes, such as transcriptional regulators, P450s and transporters, as predicted by antiSMASH (antibiotics and Secondary Metabolite Analysis SHell) [40] (S11 Table). According to the BLAST results from the NCBI NR database, no homologues of functionally characterized SMs were detected. Among them, we detected the expression of 29 core genes with FPKM (fragments per kilobase of transcript per million mapped fragments) values > 0.5, using an RNA-seq analysis of PLBJ-1 cultured in PDB medium for 8 days.

A phylogenetic tree was constructed based on the KS domain amino acid sequence of the PKSs in *P. lilacinum* and the products of known PKSs, which were divided into three main clades: non-reducing (NR) PKSs, partially reducing (PR) PKSs and highly reducing (HR) PKSs (S3 Fig). VFPBJ\_05021, VFPBJ\_09342, VFPBJ\_09755, and VFPBJ\_10843 were predicted as NR PKS-encoding genes, and they shared the highest homology with the non-reducing biosynthetic genes, such as citrinin [41] and griseofulvin [42]. VFPBJ\_00212, VFPBJ\_02527, VFPBJ\_02532, VFPBJ\_03442, VFPBJ\_05962, VFPBJ\_06473, VFPBJ\_07567 and VFPBJ\_09314 were distributed in the HR PKS clade in close relationship with HR polyketides, such as fumonisin synthase Fum1p [43]. The phylogenetic analysis was consistent with the domain structure analysis of degree of reduction, in which the HR PKS contained the reductive domains KR (keto-reductase), ER (enoyl reductase) and DH (dehydratase), while the NR PKS did not contain these domains (S3 Fig, S11 Table). VFPBJ\_05021 and VFPBJ\_09342 were grouped with the antibiotics griseofulvin and citrinin with a bootstrap value of 100%, and they shared a common domain structure. This finding suggested that griseofulvin/citrinin or structurally related compounds could be produced by *P. lilacinum*. However, we did not detect these compounds when *P. lilacinum* was cultured in PDB for 8 days.

Among the 10 NRPSs, six contained one module or an incomplete module, which could encode products with one amino acid. Four NRPSs were multi-module enzymes, which could encode products composed of more than one amino acid. To examine the potential NRPS orthologues of *P. lilacinum* and to detect the feasible NRPS evolutionary mechanism in the family Ophiocordycipitaceae, a genealogy was created based on the A-domains from the NRPS of fungi in Ophiocordycipitaceae and several functionally characterized products (S4 Fig). The tree depicted an intricate evolutionary relationship for the NRPS genes. A general trend throughout the tree was that, in Ophiocordycipitaceae, many A-domains clustered with orthologues in other species than with in the same protein. Notably, the 11 A-domains of the cyclosporine synthetases from *T. inflatum* clustered separately (S4 Fig, node 3), indicating that other species were incapable of encoding cyclosporine and that its evolution occurred after *T. inflatum* diverged from these fungi in Ophiocordycipitaceae.

This phylogenetic analysis of the A-domains for *P. lilacinum* detected a series of homologous A-domains: four of the mono-module NRPSs had functionally uncharacterized homologues.

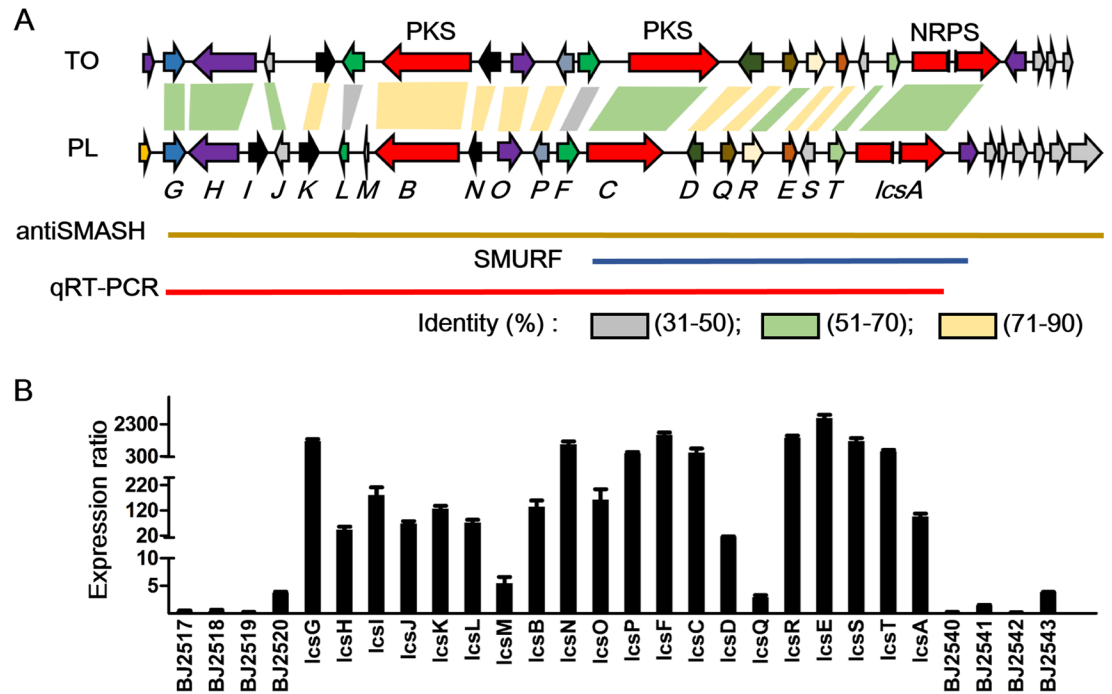
VFPBJ\_05068 was identified as siderophore synthetase, of which three of the A-domains were grouped with homologues to form a sub-clade (S4 Fig, node 2). The three A-domains of VFPBJ\_06596 were grouped with TINF2556, annotated as an ergot alkaloid in *T. inflatum*, while TINF2556 contained four modules.

The peptaibiotics, a class of linear NRPSs that are abundant of AIB [44], were clustered into one sub-clade (S4 Fig, node 1), mainly including the peptaibiotics from *T. ophioglossoides* [45], *T. inflatum* and *P. lilacinum*. The ten A-domains from VFPBJ\_02539 (identified as the leucinostatin biosynthetic gene *lcsA* in this study), clustered with the ten A-domains from the peptaibiotic TOPH\_08469 in *T. ophioglossoides*, with bootstrap values of 100%, and a global BLAST analysis revealed that the sequence identity of the two homologues was 65%. Neither orthologue was identified in other species of Ophiocordycipitaceae. The single A-domain of *lcsA* was scattered in the peptaibiotic sub-clade of the tree, while A2, A5 and A6, which activated Leu or related amino acids, were identified in the subsequent study and were grouped together with a bootstrap value of 60%, suggesting that both lineage-specific changes and module duplication contributed to the evolution of the leucinostatin metabolites. In the previous study, A4, A7 and A8 of TOPH\_08469 were distributed in a sub-clade enriched in A-domains encoding AIB [45], and our study demonstrated that A4, A7 and A8 of *lcsA* were encoded for AIB.

In *T. ophioglossoides*, the TOPH\_08469 gene cluster was predicted to contain 28 genes from TOPH\_08452 to TOPH\_08478 that were located in an ~124 kb region [45]. A comparative analysis of genes surrounding *lcsA* and TOPH\_08469 cluster revealed a high synteny (Fig 4A). VFPBJ\_02521 (designed as *lcsG*) shared 68% sequence identity with TOPH\_08452, and *lcsA* shared 66% sequence identity with TOPH\_08469. Interestingly, no homologues of the genes next to the cluster, VFPBJ\_02510 to VFPBJ\_02520 and VFPBJ\_02540 to VFPBJ\_02550, were identified in the *T. ophioglossoides* genome. Within the *lcs* cluster, two genes, cytochrome P450 *lcsI* and a protein with unknown function, *lcsM*, did not possess homologues in the TOPH\_08469 cluster, while all of the leucinostatin biosynthetic genes in *T. ophioglossoides* (TOPH\_08452 to TOPH\_08469) had homologues within the *lcs* cluster. These results suggested that this nearly 100 kb region might have been horizontally transferred from other fungal or bacterial species. However, leucinostatins have not been reported to be produced by *T. ophioglossoides* to date.

## Identification of the NRPS gene *lcsA*, which is involved in the leucinostatin biosynthesis pathway

The lipopeptide leucinostatin A contains ten amide bonds that divide the molecule into 11 moieties, including 4-methylhex-2-enoic acid, 9 amino acid residues and DPD. The property of the mixture of the polyketide and peptide moieties in the leucinostatins indicated a PKS, NRPS or hybrid PKS-NRPS origin. It is logical to consider that a single reducing PKS encodes the 4-methylhex-2-enoic acid, and a NRPS enzyme encodes the remaining portion, as in the models for emericellamide synthesis in *Aspergillus nidulans*[46] and pneumocandin in *Glarea lozoyensis*[47]. Among the multi-module NRPSs in *P. lilacinum*, VFPBJ\_05068 contains 13 domains grouping into 3 modules, VFPBJ\_06596 contains seven domains grouping into three modules, and VFPBJ\_11400 contains six domains grouping into two modules. These enzymes were insufficient for the assembly of nine amino acids of leucinostatins. Thus, VFPBJ\_02539 was left as the only plausible candidate. VFPBJ\_02539 (*LcsA*) consists of 11,872 amino acids and was encoded by a gene with five introns. The domain structure of *LcsA* was comprised of 10 C-A-PCP modules and carried the correct number of amino acids for the assembly of leucinostatins. The NRSPredictor2 [48] offered little insight into the substrates except that the substrates of A1 and A3 were proline and leucine, respectively (S12 Table). Two PKSs, *lcsB* and

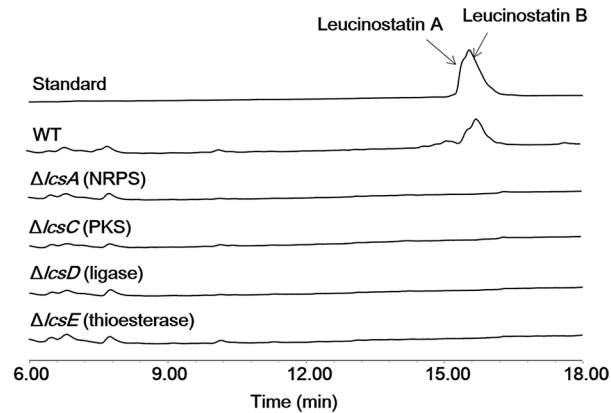


**Fig 4. The boundary of the *lcs* cluster in *P. lilacinum* with its homologues in *T. ophioglossoides*.** (A) Horizontal arrows of the same color represent the orthologous genes. The sequence identity between the homologous genes from two fungi is shown by shaded areas with different colors. TO, *T. ophioglossoides*; PL, *P. lilacinum*. The bars indicate boundaries of the *lcs* cluster predicted by antiSMASH, SMURF, and qRT-PCR. (B) The expression ratio of the genes around *lcsA* when expression in PLBJ-1 cultured in leucinostatin-inducing medium was compared to expression in non-inducing medium. The ratios for different genes demonstrated an extensive range, so the breakpoint was inserted into the Y axis.

doi:10.1371/journal.ppat.1005685.g004

*lcsC*, located not far upstream of *lcsA*, which could encode 4-methylhex-2-enoic acid, indicated that this cluster is responsible for leucinostatin production. Furthermore, *lcsD* (VFPBJ\_02533), located between *lcsA* and the PKSs, was annotated as an acyl-CoA ligase, offering a conceivable route for connecting the fatty acid and peptide.

To verify the associations between the putative *lcsA* and leucinostatins, a gene deletion method was developed for *P. lilacinum* based on the previous method for *Fusarium oxysporum* [49], with the G418 sulfate-resistance gene *neo* as the selection marker. A portion of *lcsA* (2,613 bp, including 236 bp upstream of the ORF) was knocked out by double homologous deletion cassettes with the *neo* marker via PEG-mediated transformation, and the resulting G418 sulfate-resistant isolates (S5A and S5B Fig) were verified by diagnostic PCR, using the primers in *neo* and outside the knockout cassette (S5C Fig) (S13 Table). Finally, one mutant ( $\Delta lcsA$ ) of PLBJ-1 was isolated with correct PCR amplification products from 320 G418 sulfate-resistance mutants (S5C Fig), and the remaining isolates resulted from ectopic integration of the *neo* gene cassette into the genome. The wild type of *P. lilacinum* and the  $\Delta lcsA$  mutant of PLBJ-1 were cultured in PDB medium for 8 days, and the ethyl acetate extracts were analyzed by HPLC-MS. The MS spectrum of the wild type displayed two overlapping peaks at 15.6 and 16.0 min, with  $m/z$   $[M+H]^+$  of 1218.9 and 1204.9, respectively, which were assigned to leucinostatins A and B and were absent in the  $\Delta lcsA$  mutant (Fig 5, S6 Fig). A comparison with the authentic standard confirmed that the missing compounds of the  $\Delta lcsA$  mutant were indeed leucinostatins A and B (Fig 5, S6 Fig). As expected, these results demonstrated the essential roles of *lcsA* in the biosynthesis of the leucinostatins.



**Fig 5. HPLC profiles (UV 210 nm) of culture extracts from the wild type *P. lilacinum* strain and mutants when grown in PDB medium.** Leucinostatins A and B were detected in the wild type isolate, while they were abolished in  $\Delta lcsA$ ,  $\Delta lcsC$ ,  $\Delta lcsD$  and  $\Delta lcsE$ .

doi:10.1371/journal.ppat.1005685.g005

## The boundary determination of the leucinostatin biosynthesis gene cluster

Different boundaries of the *lcs* cluster were defined by the SMURF and antiSMASH programs (Fig 4A). Nine genes flanking *lcsA* from *VFPPL\_02532* to *VFPPL\_02540* spanning 62 Kb were predicted to be in the cluster by SMURF (Secondary Metabolite Unique Regions Finder) [50], while a larger cluster comprising 26 genes from *VFPBJ\_02521* to *VFPBJ\_02546*, spanning 120 Kb, was predicted by antiSMASH. Therefore, it was necessary to explore the genes that were involved in the pathway using a biological approach.

Changes in the culture medium could impact the general metabolic profile of an organism, based on the “OSMAC” (one strain-many compounds) hypothesis[51]. Indeed, we found that *P. lilacinum* produced leucinostatins A and B when cultured with our lab recipe of PDB but did not produce leucinostatins when cultured in PDB-BD (see the [Materials and Methods](#) section). This result provided clues to identify the boundary of the *lcs* cluster using producing versus non-producing media. qRT-PCR analysis was conducted to compare the expression patterns of genes flanking *lcsA* when PLBJ-1 was grown in the two types of media for 8 days. Furthermore, RNA-Seq of PLBJ-1 under leucinostatin-inducing conditions (PDB medium) was performed.

As expected, the expression level of NRPS *lcsA* when *P. lilacinum* was grown in leucinostatin-inducing medium was upregulated 95-fold, compared to those grown in non-inducing medium (Fig 4B). The genes downstream of *lcsA*, including the putative transporter ABC gene *VFPBJ\_02540*, did not display a higher expression level in the leucinostatin-inducing medium, indicating that they were not involved in the leucinostatin biosynthesis pathway. Correspondingly, the RNA-Seq expression profile during leucinostatin production showed a low FPKM value of *VFPBJ\_02540* (2.01) (S7A Fig), while the FPKM value of *lcsA* was 65.4. These results indicated that the 3' edge of the cluster was *lcsA*. The genes upstream of *lcsA* from *VFPBJ\_02520* to *VFPBJ\_02538* (*lcsT*) were upregulated at different levels in the leucinostatin-inducing medium. A 16- to 2692-fold increase in expression was observed (Fig 4B), except for three genes, *VFPBJ\_02520*, *LcsM*, and *lcsQ*, which showed less than 10-fold increase and low FPKM values in the transcriptional data (S7A Fig). *VFPBJ\_02520* was annotated as a phosphohydrolase that appeared to be involved in nucleic acid metabolism and signal transduction, instead of secondary metabolism[52]. Thus, we speculated that the 5' boundary of the cluster was *VFPBJ\_02521* (*lcsG*). To support this hypothesis, the expression patterns of the genes

flanking the cluster were analyzed using qRT-PCR analysis in wild type PLBJ-1 and  $\Delta lcsA$  grown in leucinostatin-inducing medium. We observed an increase in the expression of wild type *P. lilacinum* ranging from four- to 79-fold (S7B Fig). Thus, a series of genes from VFPBJ\_02521 to VFPBJ\_02539, designated as *lcsA* to *lcsT*, included the core enzymes, modifying enzymes and transporter enzymes coding for the biosynthesis of leucinostatins (Fig 4A, Table 2).

### Identification of the leucinostatin biosynthetic pathway

Considering the structural similarities of leucinostatin A with emericellamide A [53] and pneumocandin [47], we reasoned that a similar biosynthetic mechanism might be required to form the skeletons of lipopeptides and peptides. As reported, a single module polyketide synthase iteratively catalyzes the formation of the linear polyketide chain; in daptomycin [54] and echinocandin B [55], acyl-CoA ligase converts the fatty acid to fatty acyl CoASH; in compound W493 B [56], a thioesterase was proposed to hydrolyze the thiol bond and shuttle the product to the first module of NRPS. To determine whether the same enzymes play critical roles in the leucinostatin biosynthesis pathway, we disrupted the PKS (*lcsC*), ligase (*lcsD*) and thioesterase (*lcsE*)-encoding genes in the cluster by homologous recombination (S5A Fig) and verified the mutants by PCR amplification (S8A Fig). After culturing the fungi in PDB medium and comparing the extracts with the PLBJ-1 wild type and  $\Delta lcsA$  by HPLC-MS, we showed that leucinostatins A and B disappeared in  $\Delta lcsC$ ,  $\Delta lcsD$  and  $\Delta lcsE$ , similar to  $\Delta lcsA$  (Fig 5, S6 Fig).

**Table 2. Description of the genes in the leucinostatin biosynthetic cluster.**

Gene ID	Name	Length	Conserved domain	Deduced function
VFPBJ_02519		270	PhyH	Epoxidase subunit
VFPBJ_02520		214	HD	Phosphohydrolase
VFPBJ_02521	<i>lcsG</i>	468	Methyltransf_2	O-methyltransferase
VFPBJ_02522	<i>lcsH</i>	1552	ABC_tran	<a href="#">ABC transporter</a>
VFPBJ_02523	<i>lcsI</i>	516	p450	Cytochrome P450
VFPBJ_02524	<i>lcsJ</i>	309	4HBT_2	Thioesterase-like
VFPBJ_02525	<i>lcsK</i>	548	p450	Cytochrome P450
VFPBJ_11786	<i>lcsL</i>	228	bZIP	Transcriptional regulator
VFPBJ_02526	<i>lcsM</i>	82		<a href="#">Hypothetical protein</a>
VFPBJ_02527	<i>lcsB</i>	2507	KS-AT-DH-MET-ER-KR	PKS
VFPBJ_02528	<i>lcsN</i>	336	p450	Cytochrome P450
VFPBJ_02529	<i>lcsO</i>	593	ABC1	<a href="#">ABC transporter</a>
VFPBJ_02530	<i>lcsP</i>	387	Aminotran_4	<a href="#">Aminotransferase</a>
VFPBJ_02531	<i>lcsF</i>	552		Transcriptional regulator
VFPBJ_02532	<i>lcsC</i>	2332	KS-AT-DH-MET-ER-KR-ACP	PKS
VFPBJ_02533	<i>lcsD</i>	386	AMP-binding domain	Acyl-CoA ligase
VFPBJ_11787	<i>lcsQ</i>	389	tRNA-synt_2c	tRNA synthetases
VFPBJ_02534	<i>lcsR</i>	400	Lactamase_B	Zn-dependent hydrolases
VFPBJ_02535	<i>lcsE</i>	316	Thioesterase	<a href="#">Thioesterase</a>
VFPBJ_02536	<i>lcsS</i>	243		<a href="#">Hypothetical protein</a>
VFPBJ_02538	<i>lcsT</i>	269	CC3_like_SDR_a	<a href="#">Epimerase</a>
VFPBJ_02539	<i>lcsA</i>	11872	PCP-(C-A-PCP)*10-NAD	NRPS
VFPBJ_02540		1444	ABC2_membrane	ABC transporter
VFPBJ_02541		594	Sec14p-like	<a href="#">Phosphatidylinositol transfer protein</a>

VFPBJ\_02519, VFPBJ\_02520, VFPBJ\_02540 and VFPBJ\_02541 are genes flanking *lcs* cluster.

doi:10.1371/journal.ppat.1005685.t002

## Overexpression of the transcription factor *lcsF*

A powerful approach to enhancing the production of leucinostatins was to express transcription factors constitutively that were used for other SMs [57]. *lcsF* encodes a putative transcription factor with a bZIP domain structure, and it is associated with secondary metabolism [58]. To assess the function of *lcsF*, we cloned it into the KSTNP vector under the control of the TrpC promoter. The resulting plasmid, KSTNP-OE*lcsF*, was randomly integrated into the genome of wild type *P. lilacinum* (S8B Fig). The positive transformants were screened by G418 sulfate and were diagnosed by PCR amplification of the expression cassette (S8C Fig). Transformants with an intact overexpression cassette were cultured in leucinostatin-inducing PDB medium for 8 days. The expression level of *lcsF* in the mycelia was analyzed by qRT-PCR, and six of ten transformants demonstrated more than 20-fold upregulation. Finally, three transformants without changes in their physiological indices were selected for the downstream test. As expected, all 20 genes in the cluster were upregulated to some extent by *lcsF*, with the exception of *lcsL* and *lcsP*, which were downregulated three- and five-fold (S9A Fig). In addition to the 30-fold increase in *lcsF* expression, the expression of the three PKS/NRPS synthase encoding-genes (*lcsB*, *lcsC* and *lcsA*) were increased by ~3- to 4-fold. For O-methyltransferase (*lcsG*), ABC transporter (*lcsH* and *lcsO*), thioesterase (*lcsE*), epimerase (*lcsT*) and the unknown function genes *lcsM* and *lcsS*, we observed ten-fold or higher upregulation. The other genes in the cluster displayed a two- to ten-fold increase in expression. Genes adjacent to the *lcs* cluster, VFPBJ\_02520 and VFPPL\_02540, were downregulated three-fold. After the wild type and OE::*lcsF* *P. lilacinum* were grown in PDB medium with shaking for 8 days, the resulting HPLC profile showed that the titers of leucinostatins A and B were elevated by at least 50% (S9B Fig). These results provided evidence that the pathway-specific transcription factor *lcsF* was capable of regulating the entire gene cluster and leucinostatin biosynthesis, further verifying the boundary of this cluster.

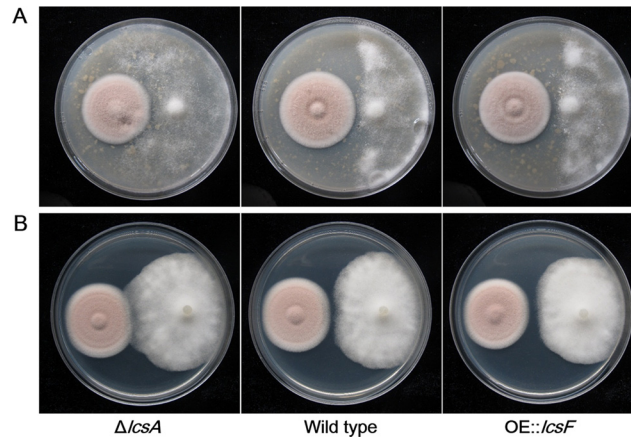
## Antagonism against the oomycetes of *P. lilacinum* depending on leucinostatins

The deletion and overexpression of the genes in the *lcs* cluster had no apparent effects on the fungal hyphae or spore phenotypes of *P. lilacinum* and did not cause any growth defects. It is well known that leucinostatins are antibiotics used to combat fungi and bacteria. Here, we found that leucinostatins contributed to the inhibition of oomycetes, which had not previously been reported. The growth of *P. infestans* and *P. capsici* was inhibited in a confronting incubation with wild type *P. lilacinum* and OE::*lcsF*, while the inhibition disappeared when they were grown in a confronting incubation with  $\Delta lcsA$  (Fig 6). Similar to  $\Delta lcsA$ , *P. infestans* could grow normally in a confronting incubation with  $\Delta lcsC$ ,  $\Delta lcsD$  and  $\Delta lcsE$  (S10 Fig). The results indicated that leucinostatins A and B inhibited the growth of some oomycetes. A gradient inhibitory zone was explored with leucinostatins A and B in different concentrations to find quantitative evidence of inhibition against *P. infestans* (S11 and S12 Figs).

## Discussion

*P. lilacinum* is one of the most important endo-parasites of plant nematodes. We obtained the genome sequence of two *P. lilacinum* strains and compared them with other nematode parasites, nematode-trapping fungi, insect parasites, a mycoparasitic fungus, a saprotrophic fungus and a plant pathogen. This method provided insights into the life strategy and evolution of nematode endoparasites.

Major gene families (GH, protease, SMs) could corroborate each other for the three *P. lilacinum* strains (PLBJ-1, PLFJ-1 and the published TERIBC 1). However, TERIBC 1 was predicted



**Fig 6. The role of leucinostatins in antagonism between *P. lilacinum* and *Phytophthora*.** (A) Cocultivation of *P. infestans* and wild type,  $\Delta lcsA$  and OE::*lcsF* *P. lilacinum* on rye agar medium. (B) Cocultivation of *P. capsici* and wild type,  $\Delta lcsA$  and OE::*lcsF* *P. lilacinum* on PDA medium.

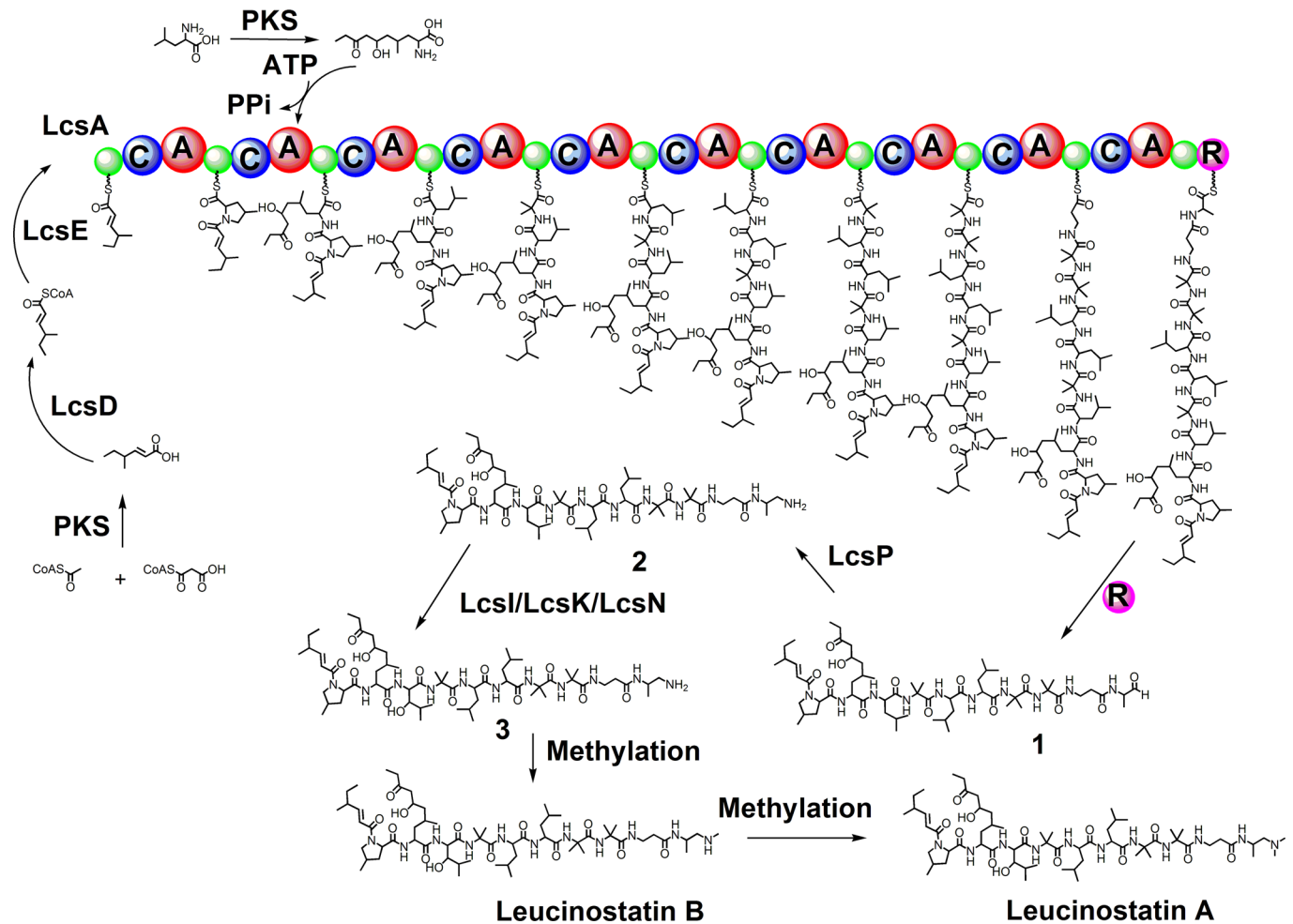
doi:10.1371/journal.ppat.1005685.g006

to encode more CEs and fewer PHI genes in contrast with PLBJ-1 and PLFJ-1 (Table 1). The *lcs* cluster was also detected in the TERIBC 1 genome. The genomic sequence identity of the *lcs* cluster (*lcsG* to *lcsA*) between PLBJ-1 and TERIBC 1 was 98.0%, and the sequence identity was 99.0% between PLBJ-1 and PLFJ-1, with the syntenic relationships shown in S13 Fig.

Although fungi have been screened for activity as bio-control agents against *P. infestans*, the biological control of late blight is dominated by bacterial antagonists. Microbial compounds known as biosurfactants [59] are believed to participate in the process. For example, the cyclic lipopeptide massetolide A produced by *Pseudomonas fluorescens* exhibited destructive effects on the zoospores of oomycetes [60]. The inhibition of *P. infestans* by leucinostatins provided the basis for their chemical application in agriculture and for further biological studies of the antagonist *P. lilacinum* on oomycetes, which had not been researched previously. Leucinostatins also demonstrated inhibition against *P. capsici* [61], another oomyceteous plant pathogen, while the antagonism of *P. lilacinum* against *P. capsici* seemed inferior to that against *P. infestans*, as shown in Fig 6.

The phylogenomic analysis revealed that *P. lilacinum* was a member of Ophiocordycipitaceae, Hypocreales, which includes fungi engaged in various lifestyles, and it was not related to the previously considered *Paecilomyces* in Sordariales. This species' closest relatives, *T. inflatum* and *T. ophioglossoides*, are insect and fungal parasites (Fig 3A), supporting the viewpoint that parasitism might occur due to the formation of novel genes that could be acquired through horizontal transfer or gene duplication and could play specific roles during host infection [62]. Moreover, these results indicated that the nematode pathogens had a strong link with insect pathogens and were distantly related to nematode-trapping fungi, as previously described in [25] and [26]. The large number of hydrolytic enzymes, particularly GHs and proteases, putatively secreted proteins and pathogenesis-related proteins in *P. lilacinum* support its various lifestyles as it encounters diverse nutrient resources [63]. Chitin and proteins comprise a significant proportion of the nematode and insect surface, the degradation of which requires serine proteases and chitinases.

The development of natural compounds from bio-control fungi have recently attracted considerable interest because the production of nematode-toxic SMs could also be a strategy for fungi to infect nematodes [64]. Next-generation sequencing technologies are becoming an essential tool for identifying novel genes for metabolite biosynthesis in fungi. The genome sequence of *P. lilacinum* revealed the potential to produce a rich repertoire of SMs, including



**Fig 7. A putative biosynthetic pathway for leucinostatin A.**

doi:10.1371/journal.ppat.1005685.g007

41 core enzymes. Most of the PKS enzymes could be clustered with the PKS enzymes encoding bioactive polyketides when the PKS tree was constructed based on the KS domains (S3 Fig). The NRPS enzyme tree based on the A-domains demonstrated that some of NRPSs of *P. lilacinum* possessed homologues in closely related species (S4 Fig). However, none of the enzymes' biosynthetic function have been validated previously. The polyketide compounds acremoxanthone C and acremoxidin A belong to the xanthone–anthraquinone heterodimer that was recently isolated from *P. lilacinum*. Structurally related heterodimeric compounds, such as acremoxanthones A and B [65] and xanthoquinodin B3 [66], have been isolated from the genera *Acremonium* and *Humicola*. This category of compounds and their derivatives have remarkable biological and medicinal activities, and their total synthesis has garnered attention worldwide. Unfortunately, there still exist limitations in the current synthesis methods [67], and metabolite regulation based on molecular biosynthesis is not available because the biosynthetic genes have not been identified. One or more non-reducing PKSs (S3 Fig) might be involved in the biosynthesis of acremoxanthone C and acremoxidin A, according to their structures.

By analyzing the genes located in the leucinostatin biosynthetic cluster, combined with the HPLC-MS analysis of gene deletion mutants (Fig 5), we were able to propose a putative



biosynthetic pathway for leucinostatins (Fig 7). This hypothetical biosynthesis initiated with the assembly of 4-methylhex-2-enoic acid by a reducing PKS. However, two reducing PKS encoding genes with 38% sequence identity are present in the cluster, and both contained KS, AT, DH, cMT, ER, KR and ACP domains. We excluded the possibility of a partnership between the two PKSs in a sequential manner or a convergent manner, as has been reported for asperfuranone [68] and azaphilone A [69], based on their structures. Moreover, in the biosynthesis of chaetoviridins and chaetomugilins from *Chaetomium globosum*, a PKS *cazF* (KS-AT-DH-cMT-ER-KR-ACP) encoded an intermediate 4-methylhex-2-enoic acid [70]. The protein sequence identities between *cazF* and *lcsB/lcsC* were both 29%; thus, we could not estimate which PKS was responsible for 4-methylhex-2-enoic acid. The results of RNA-seq and qRT-PCR indicated that both genes contributed to leucinostatin synthesis. The deletion of *lcsC* interrupted leucinostatin biosynthesis, which confirmed that *lcsC* is essential for the synthesis of leucinostatins. Due to the difficult genetic manipulation, we failed to obtain *lcsB* deletion mutant.

The lipopeptide pathways and organizations of their clusters have some striking commonalities. We got clues from the lipopeptides echinocandin B [55], pneumocandin [47] and emericellamide [53]. The polyketide residue might be transferred to the NRPS *LcsA*, mediated by two additional putative enzymes, acyl-CoA ligase (*LcsD*) and thioesterase (*LcsE*). The linear polyketide carboxylic acid, which was released from PKS, was converted to a CoA thioester by *LcsD*, and then *LcsE* hydrolyzed the thiol bond and shuttled the polyketide intermediate to *LcsA*. 4-Methylhex-2-enoic acid was not detected in the culture of the  $\Delta$ *lcsD* isolate, indicating that the triketide might be stuck in the PKS enzyme to prevent its release until the ligase is added for the reaction.

The phylogenetic analysis of the ten A-domains of *LcsA* revealed that *LcsA\_A2*, *LcsA\_A5*, *LcsA\_A6* and *LcsA\_A3* were grouped into one clade, and *LcsA\_A4*, *LcsA\_A7* and *LcsA\_A8* were grouped into another clade (S14 Fig). The conserved domains are believed to have evolved through module duplication, and they activate similar amino acid structures [71]. In the plausible model for leucinostatin synthesis, A5 and A6 incorporated leucine, A2 incorporated AHy-MeOA, the structure of which is equal to a hydroxyl-3-pentone extending at the leucine, and A3 incorporated 3-hydroxyl leucine. A4, A7 and A8 incorporated AIB. Thus, the structural similarity of amino acids activated by conserved A-domains verified that the 4-methylhex-2-enoic acid moiety in the leucinostatins was assembled by a discrete enzyme, instead of *LcsA*. The C domain of the first module catalyzed the condensation of 4-methylhex-2-enoic acid and MePro carried by domain A1, followed by successive condensations of nine amino acids to trigger the elongation of the linear peptide. Next, the peptide scaffold would be released by the NAD(P)H-dependent R domain (thioester reductase) at the C-terminal region of *LcsA*.

In the leucinostatin biosynthetic pathway, it is intriguing that the DPD residue at the C-terminus of leucinostatin A was neither an amino acid nor a carboxy acid, which are incapable of being activated by the A-domain and converting to amino acyl adenylate [72]. The DPD seems to be a modified form of amino acids, whereas the primary form of this moiety cannot be determined based on the domain sequence. However, we could deduce a possible pathway for the modification of the last amino acid according to the structure and the function of the genes in the cluster (Fig 7). Originally, an Ala was likely incorporated into the decapeptide skeleton by the A10 domain, which was attached to *LcsA* via a thioester bond; subsequently, the R domain released this intermediate product. The NAD(P)H-dependent R domains reductively catalyzed to produce linear aldehyde **1** by off-loading peptide thioesters, following completion of the peptide skeleton, as presented in previous studies [73, 74]. Linear aldehydes frequently occurred as intermediates and underwent subsequent reactions, such as macrocyclization, to yield the imine product koranimine [75]; further reduction yielded myxochelin A or

transamination to form an amine myxochelin B by aminotransferase [76]. Regarding the leucinostatins, we speculated that aldehyde **1** would go through a transamination reaction to form compound **2**, which was accomplished by the putative aminotransferase *lcsP*.

In this pathway, the unhydroxylated leucine of intermediate **2** undergoes hydroxylation to form compound **3**. Three putative cytochrome P450-encoding genes (*lcsI*, *lcsK* and *lcsN*) within the cluster alternatively might catalyze this modification. Another scenario equivalent to this pathway for leucinostatin A synthesis was that a leucine was hydroxylated prior to its incorporation into the peptide. In all likelihood, the varying extents of methylation of compound **3** catalyzed to form leucinostatins A and B. It is worth mentioning that, had the methylation reaction not occurred, compound **3** might be the ultimate precursor of leucinostatin C, which is compound in leucinostatin family isolated from *P. marquandii*.

The AHyMeOA in leucinostatin A activated by the A2 of *lcsA* was regarded as a ramification of leucine because leucine is located at this position in leucinostatins C, T, F, D and H, although these compounds were not detected in the PLBJ-1 culture. Based on its structure and the presence of redundant PKSs within the cluster, alternative PKS could be involved in synthesizing the carbon chain. In addition, the leucinostatins contained the nonproteinogenic MePro, incorporated in the synthesis of nostopeptolides in *Cyanobacteria* [77]. A zinc-dependent dehydrogenase, *nosE*, and a P5C reductase, *nosF*, were involved in the oxidation and subsequent cyclization of leucine to form MePro, and the presence of *nosE* and *nosF* recently led to screening for novel MePro-containing peptides [78]. In the pneumocandins from *Zalerion arboricola*, feeding experiments established that leucine was cyclized to produce 3-hydroxy-4-methylproline, whereas MePro might be an intermediate [79]. It was reasonable to assume that the MePro in the leucinostatins originated from leucine cyclization. Although homologues of *nosE* nor *nosF* were not present in the *lcs* cluster, it was plausible that MePro biosynthesis, engaged in a separate pathway, was independent of leucinostatin synthesis. Another nonproteinogenic amino acid,  $\beta$ -Ala, was present in leucinostatins and activated by the A9 of *lcsA*. A previous study of the destruxins in *Metarhizium* proposed that the aspartic acid decarboxylase *dtxS4* triggered the decarboxylation of aspartic acid into  $\beta$ -Ala, as a substrate for the assembly line [80]. A genome-wide blast search for genes encoding aspartic acid decarboxylases in PLBJ-1 revealed the presence of two candidate genes, VFPBJ\_01400 and VFPBJ\_10476, with 68% and 61% sequence identity to *dtxS4*, respectively, which could have catalyzed the biosynthesis of  $\beta$ -Ala in leucinostatins.

## Conclusions

The genomes of *P. lilacinum* strains PLBJ-1 and PLFJ-1 were sequenced, completely assembled, annotated, and comparatively analyzed with related fungi. Phylogenomic analysis showed that *P. lilacinum* was most closely related to *T. inflatum* and *T. ophioglossoides*, and the cluster of nematode parasitic fungi and insect pathogens indicated their common origin. PKS and NRPS-encoding genes were thoroughly characterized and analyzed by phylogenetic analysis, from which we found that *lcsA* was specific to *P. lilacinum* and *T. ophioglossoides*. Furthermore, *lcsA* was proved to be responsible for leucinostatin biosynthesis by homologous deletion. The boundary of the *lcs* cluster was identified by comparison of gene expression levels when *P. lilacinum* was cultured in leucinostatin-inducing and non-inducing medium as well as RNA-Seq analysis. Disruption of *lcsC*, *lcsD* and *lcsE* demonstrated the critical roles of PKS, acyl-AMP ligase and thioesterase in the biosynthetic pathway of leucinostatins. Overexpression of the transcription factor *lcsF* increased the production of leucinostatins A and B through regulated expression levels of genes in the *lcs* cluster. We also demonstrated that leucinostatins could enable the fungus with antagonistic activity against the oomycetes.

## Materials and Methods

### Fungal strains, vectors and reagents

The leucinostatin-producing *P. lilacinum* strain PLBJ-1 (CGMCC3.17492) was isolated from tomato roots in Beijing, China, and PLFJ-1 (CGMCC3.17493) was isolated from tomato roots in Fujian, China. Both strains were sequenced to obtain the common features of *P. lilacinum* and to ensure the information accuracy of the *lcs* cluster. PLBJ-1 was used as the wild type recipient for the subsequently genetic manipulations because PLFJ-1 was insensitive to the antibiotics used as selection markers. *P. infestans* and *P. capsici* were maintained at the Chinese Academy of Agricultural Sciences. The pKOV21 vector used for homologous deletion and the KSTNP vector used for overexpression came from Prof. Youliang Peng, China Agricultural University. The leucinostatin A standard came from Bioaustralis, Inc. (NSW, AUS). G418 sulfate was purchased from Amresco, Inc. (OH, USA).

### Medium and culture conditions

The non-inducing PDB-BD medium (Potato Dextrose Broth) came from Becton, Dickinson and Company (NJ, USA). Leucinostatin-inducing PDB medium was prepared in the lab. Briefly, 200 g of potatoes were boiled for 30 min, and then 20 g of glucose were dissolved into the filtrate and diluted to 1 L. The rye agar medium contained 50 g of crushed rye, 20 g of sucrose and 15 g of agar per liter. PDB cultures with  $1 \times 10^5$  conidia per mL of PLBJ-1 were grown at 28°C on a shaker at 150 rpm for 8 days before DNA/RNA isolation.

### DNA and RNA isolation

The mycelium tissues of the PLBJ-1 and PLFJ-1 isolates were harvested via filtration. Genomic DNA was isolated using a Qiagen DNeasy kit, according to the manufacturer's protocol. The PLBJ-1 tissue for RNA isolation was grown in the leucinostatin-inducing medium. RNA was extracted using TRIZOL reagent (Invitrogen, USA) following the manufacturer's protocol.

### Genome assembly, gene prediction and RNA-seq analysis

The raw sequencing data (Illumina HiSeq 2000) from the PLBJ-1 and PLFJ-1 strains were generated by BGI-Shenzhen (China) and Berry Genomics Co., Ltd. (China), respectively. A total of 13.27 Gb bases for the PLBJ-1 strain from three libraries, with average insert sizes of 165 bp, 758 bp and 5,490 bp, were obtained, and 5.88 Gb bases for the PLFJ-1 strain from two libraries with the average insert sizes of 175 bp and 4,760 bp were obtained. Both of the genomes were assembled using ALLPATHS-LG revision 42305 [81]. The repeat sequences were identified as previously described [82], based on *de novo* and homology methods. For the *de novo* method, Piler [83] and RepeatScout, version 1.0.5 [84], were used to construct the repeat sequence families; then, RepeatMasker, version 4.0.5, was used for repeat analysis. For the homology method, the sequence families from Repbase, version 19.06 [85], were used for annotation by performing RepeatMasker analysis.

For gene prediction, the Augustus algorithm, version 2.7 [86], identified 11,404 and 11,554 complete genes for the PLBJ-1 and PLFJ-1 strains, respectively, and the GeneMark-ES algorithm, version 2.3f [87], discovered 11,001 and 11,070 complete genes for the PLBJ-1 and PLFJ-1 strains, respectively. The comparison showed that 9,509 and 9,562 genes of the PLBJ-1 and PLFJ-1 strains were predicted by both Augustus and GeneMark-ES. These consensus genes were considered to be high quality predicted genes and were used in this study. The additional 2,264 and 2,201 genes of the PLBJ-1 and PLFJ-1 strains were obtained according to the method in [82]. EuGene, version 4.1 [88], was used to integrate multiple sources, including

transcription start sites identified by Netstart [89], homologous proteins identified from the Swiss-Prot database, version 2015-07-22, by BLAST, version 2.2.26, the assembled transcripts generated by IDBA-tran, version 1.1.1 [90], and the exon junctions identified from RNA-seq by Tophat, version 2.0.13 [91]. The gene expression values were presented by the expected FPKMs using Cufflinks, version 2.2.1 [92], based on the Tophat [91] analysis.

### Protein family classification

Proteins were annotated by aligning their sequences to the NCBI fungi refseq, version 2015-07-10, and SwissProt, version 2015-07-22, with an E-value cutoff of  $1e-5$  using BLASTP. In addition, the Pfam database, version 27.0, was used for domain annotation by HMMER, version 3.1b1 (<http://hmmer.janelia.org/>). The putative proteins were further classified by Gene Ontology (GO) [93], using Blast2Go [94], and the euKaryotic Clusters of Orthologous Groups (KOG) [95], using BLAST (E-value of  $1e-5$ ). The Web server of the CAZymes Analysis Toolkit (CAT) [96] was used to identify CAZymes in *P. lilacinum* with E-values  $\leq 1e-50$ .

The proteases were discovered by the MEROPS batch BLAST online server [97]. Proteins with sequences that matched the cytochrome P450 genes [98] with E-values  $\leq 1e-50$  were annotated as P450 enzymes. Candidate pathogenic factors were predicted by sequence alignment against the Pathogen Host Interactions (PHI) database, version 3.5 [99], with E-values  $\leq 1e-50$ . In addition, the secretomes were identified based on recognizing the signal peptide and transmembrane sequences. Proteins were considered to be secreted proteins if the signal peptides were identified by at least two methods among SignalP, version 4.0 [100], TargetP, version 1.1 [101], Phobius, version 101 [102], and Predisi [103], and transmembrane sequences were not identified by at least one of the methods among SignalP, Phobius and TMHMM, version 2.0c [104].

### Orthologous and phylogenomic analysis

Orthologous groups of genes from *P. lilacinum* and the other fungi listed in S14 Table were detected by OrthoMCL, version 2.0.9 [105], and then were filtered to identify the single copy orthologues. The single copy orthologues were aligned with MUSCLE [106]. The poor alignment regions of the concatenated sequences were removed using Gblock, version 0.91b [107], and then the high quality sequences were used for the maximum likelihood phylogeny analysis with the Dayhoff model implemented in the TREE-PUZZLE program [108]. Bootstrap support value was calculated by analyzing 1,000 replicates.

### Phylogenetic analysis of the PKS and NRPS genes

The secondary metabolite genes were discovered by performing SMURF [50] and antiSMASH [40] analyses. PKS and NRPS domain structures were characterized by antiSMASH and Pfam, or were visually identified by multiple alignments. The KS domains extracted from PKS and the A-domain from NRPS were aligned by MUSCLE [106], and then a maximum likelihood phylogeny was constructed by treeBeST (<http://treesoft.sourceforge.net/treebest.shtml>) using 1,000 bootstrap replicates.

### qRT-PCR analysis

Three biological replicates were performed for each analysis of the relative expression levels. The cDNAs were synthesized with a *TIANScript II RT Kit* (TIANGEN, China). The cDNA was analyzed by qRT-PCR using SYBR Premix Ex Taq (TAKARA, Japan) on a BIO-RAD CFX96 (BIO-RAD). The housekeeping actin gene designed from VFPBJ\_07912, which was similar to

the reported GU299860.1, was used for normalization. The relative expression values were calculated using the  $2^{-\Delta\Delta Ct}$  method. The primers are listed in [S13 Table](#).

## Molecular genetic procedures

Polyethylene glycol-mediated protoplast transformation of PLBJ-1 was performed as previously reported [49, 109], with the following modifications: the protoplast was produced by  $20 \text{ gL}^{-1}$  Driselase (Sigma) digestion for 4 h at  $31^\circ\text{C}$ . The regeneration medium was PDA medium containing G418 sulfate ( $400 \mu\text{g/L}$ ), supplemented with molasses ( $10 \text{ g/L}$ ), saccharose ( $0.6 \text{ M}$ ), yeast extract ( $0.3 \text{ g/L}$ ), tryptone ( $0.3 \text{ g/L}$ ), and casein peptone ( $0.3 \text{ g/L}$ ) [10]. The construction of knockout and overexpression plasmids originated from pKOV21 and KSTNP, and the primers are listed in [S13 Table](#). A quick method for isolating the fungal genomic DNA was developed to screen for a large number of transformants. Briefly, a nip of mycelia was transferred to  $50 \mu\text{L}$  of NaOH ( $50 \text{ mM}$ ) and was incubated at  $95^\circ\text{C}$  for 20 min. The solution was directly used for PCR amplification after  $5 \mu\text{L}$  of Tris-HCl ( $1 \text{ M}$ ) were added to neutralize the base.

## Culture extraction and HPLC-MS profiling

Cultures of  $1 \times 10^5$  conidia per mL of *P. lilacinum* and its mutants were grown in leucinostatin-inducing PDB medium at  $28^\circ\text{C}$  on a shaker at 150 rpm for 8 days. Culture medium (7.5 L) was extracted with the same volume of EtOAc three times (each 1 h) ultrasonically. The combined EtOAc extracts were concentrated to afford a crude extract (0.4 g), which was subjected to reversed-phase ODS column chromatography eluting with MeOH- $\text{H}_2\text{O}$  (from 40% to 100%) to afford 6 fractions (Fr.A–Fr.F). Fr.E (40 mg) was passed through a Sephadex LH-20 column (MeOH) and yielded mixtures of 5.0 mg of leucinostatins A and B. The structure of the mixtures was further identified by standard substance using LC-MS analysis. Approximately 200 mL of culture medium were used for comparative LC-MS analysis between PLBJ-1 and its mutants. LC-MS was performed on an Agilent Accurate-Mass-QTOF LC/MS 6520 instrument. HPLC analysis was performed on a Waters HPLC system (Waters e2695, Waters 2998, Photodiode Array Detector) using an ODS column (C18,  $250 \times 4.6 \text{ mm}$ , YMC Pak,  $5 \mu\text{m}$ ). The ODS ( $5 \mu\text{m}$ ) column was produced by YMC Co. Ltd. (Kyoto, Japan). The Sephadex LH-20 was purchased from GE Healthcare. Analytical HPLC was conducted with a Waters HPLC system (Waters e2695, Waters 2998, Photodiode Array Detector) using an ODS column (C18,  $250 \times 4.6 \text{ mm}$ , YMC Pak,  $5 \mu\text{m}$ ) with a flow rate of  $1 \text{ mL/min}$ . The fresh extracts were dissolved in methanol before being separated on a linear gradient of MeOH: $\text{H}_2\text{O}$  (0.1% formic acid) at a flow rate of  $1 \text{ mL/min}$ . Fresh extracts of mutant strains were detected for 30 min using a linear gradient of 20% to 100% (0–20 min), 100% MeOH (20–25 min), and 20% MeOH (25–30 min). The LC-MS analysis method was consistent with analytical HPLC.

## Fungus bioassay

Confronting incubation of *P. lilacinum* (wild type,  $\Delta lcsA$  and OE::*lcsF*) with *P. infestans* was performed on rye agar medium in 9 cm Petri plates, incubated simultaneously and cultured at  $28^\circ\text{C}$  for 24 h and then at  $18^\circ\text{C}$ , the optimum temperature for *P. infestans*, for 9 days, while confrontation with *P. capsici* was performed on lab-made PDA medium, cultured at  $28^\circ\text{C}$  for 7 days. For the inhibitory zone experiment, freshly produced sporangia of *P. infestans* was suspended in sterile water at a concentration of  $2 \times 10^5$  sporangia/mL. One milliliter of the suspension was smeared on 15 cm Petri plates, followed by Oxford cups with a diameter of 1 cm being placed. From 5 to  $60 \mu\text{g}$  (increment  $5 \mu\text{g}$ ) of leucinostatins A and B dissolved in 20% methanol were added to the Oxford cup, and  $20 \mu\text{L}$  of 20% methanol were used a control. Then, the Oxford cups were

removed after the solution was absorbed by media. Five days later, the area was calculated by drawing circles of the inhibitory zone on metric graph paper and counting the number of square millimeters within the circle [110, 111]. Three biological replicates were performed. At the same time, the effects of 50  $\mu\text{g}$ , 33  $\mu\text{g}$ , 17  $\mu\text{g}$  and 8.5  $\mu\text{g}$  of leucinostatins are demonstrated in S11 Fig.

## Accession numbers

The genome sequences of PLBJ-1, PLFJ-1, and *P. chlamydosporium* strain 170 used for comparative analysis have been deposited at GenBank under the accession numbers LSBH000000000, LSBI000000000 and LSBJ000000000, respectively.

## Supporting Information

**S1 Fig. Genomic synteny of PLBJ-1 and TERIBC 1.** Syntenic relationships were analyzed by BLASTN with an E-value cutoff of  $1e^{-5}$ . The red semicircle represents the scaffolds of PLBJ-1, while the blue semicircle represents the scaffolds of TERIBC 1. Scaffold lengths of  $\geq 100$  Kb were used for this analysis, and the threshold of the matched block was  $\geq 1000$  bp, which was connected by lines of the same color.

(TIF)

**S2 Fig. KOG classification of PLBJ-1 and PLFJ-1 genomes and proteins related to the PHI database.**

(TIF)

**S3 Fig. Maximum likelihood phylogeny of the PKSs in PLBJ-1 and other identified fungal PKSs.** The domain structure of PKSs was predicted by Pfam and antiSMASH, and KS domains were used for phylogenetic analysis. Bootstraps values  $>50\%$  are presented on the nodes.

(PDF)

**S4 Fig. Maximum likelihood phylogeny of the NRPS A domains from fungi in Ophiocordycipitaceae.** The tree included NRPSs from Ophiocordycipitaceae (*P. lilacinum*, *T. inflatum*, *T. ophioglossoides*, *H. minnesotensis*, *O. sinensis* and *O. unilateralis*) and some functionally characterized products. The sub-clades of peptaibiotics, siderophore synthetase and cyclosporine synthetase are highlighted by shading.

(PDF)

**S5 Fig. Strategy and identification of the *lcsA* deletion.** (A) Double homologous recombination strategy for deleting 2,613 bp of *lcsA* by introducing the homologous arm into the plasmid pKOV21. (B) Candidate transformants grown on medium resulted from either homologous recombination or ectopic integration of the *neo* gene cassette into the genome. (C) PCR amplification verified the validity of  $\Delta lcsA$ . Lanes 1 and 2, amplified with primer pair targetF and targetR, identified the target gene that was deleted from *P. lilacinum*. Lanes 3 and 4 were amplified with primer pair upcheckF and neoupR, and lanes 5 and 6 were amplified with primer pair neo-downF and upcheckR. The bands in lanes 4 and 6 indicated the *neo* replaced the target gene.

(TIF)

**S6 Fig. MS profile of the extracts of the wild type and mutant *P. lilacinum* cultures.** Peaks of  $[M+H]^+ = 1219$  (leucinostatin A) and  $[M+H]^+ = 1205$  (leucinostatin B) were detected in wild type, but not in the  $\Delta lcsA$ ,  $\Delta lcsC$ ,  $\Delta lcsD$  and  $\Delta lcsE$ .

(TIF)

**S7 Fig. The boundary of the *lcs* cluster identified by RNA-seq and qRT-PCR.** (A) Expression levels in FPKM for genes in and flanking the *lcs* cluster in leucinostatin-inducing medium. (B)

Expression ratio of genes in wild type PLBJ-1 to those in  $\Delta lcsA$  when cultured in leucinostatin-inducing medium.

(TIF)

**S8 Fig. Identification of the  $\Delta lcsC$ ,  $\Delta lcsD$  and  $\Delta lcsE$  transformants and overexpression of transcription factor *lcsF*.** (A) PCR amplification verified the validity of  $\Delta lcsC$ ,  $\Delta lcsD$  and  $\Delta lcsE$ . Lanes 1 and 2 were amplified for target gene products, lanes 3 and 4 were amplified with primer pairs in *neo* and upstream of the knockout cassette, and lanes 5 and 6 were amplified with primer pairs in *neo* and downstream of the knockout cassette. (B) Overexpression cassette of KSTNP-OE*lcsF* including PlcsF and TrpC terminator, using *neo* as the marker. (C) The approximate 3,000 bp fragments in lanes 2, 3 and 4 were amplified from three overexpressing transformants with the primer pair OE*lcsFF* and OE*lcsFR*, and lane 1 was amplified from wild type *P. lilacinum*.

(TIF)

**S9 Fig. *lcsF* overexpression increased the expression of related genes and leucinostatin production.** (A) Overexpression of transcription factor *lcsF* enhanced expression levels of genes in the *lcs* cluster. The expression ratio referred to the genes in OE::*lcsF* to those in the wild type when cultured in leucinostatin-inducing medium. (B) HPLC profiles (UV 210 nm) of culture extracts from the wild type *P. lilacinum* strain and three overexpressing transformants cultured in PDB medium.

(TIF)

**S10 Fig. Cocultivation of *P. infestans* with *P. lilacinum* wild type,  $\Delta lcsC$ ,  $\Delta lcsD$  and  $\Delta lcsE$  mutants on rye agar medium.**

(TIF)

**S11 Fig. Inhibitory zone of *P. infestans* with leucinostatins A and B of different concentration.** The location of leucinostatins was represented by black circle, of which the dosage is marked. The red dotted line circles the inhibitory zones of *P. infestans*. Twenty microliters of 20% methanol were placed at CK.

(TIF)

**S12 Fig. The area (mm<sup>2</sup>) of the inhibitory zone with different leucinostatin dosages.**

(TIF)

**S13 Fig. Syntenic relationships of *lcs* cluster in PLBJ-1, PLFJ-1 and TERIBC 1.** Syntenic relationships were analyzed by BLASTN, with an E-value cutoff of 1e-5. The *lcs* cluster was located in scaffold00004 (1,048,997–1,162,497 bp) of PLFJ-1, scaffold00002 (1,030,100–1,143,599 bp) of PLBJ-1 and LOFA01000012.1 (1,012,032–1,125,446 bp) of TERIBC 1. The genes in this region are indicated by blue arrows.

(TIF)

**S14 Fig. Maximum likelihood phylogeny of the ten A domains of LcsA.**

(TIF)

**S1 Table. Genome size and repeat sequence in *P. lilacinum* and other fungi.**

(DOCX)

**S2 Table. Super scaffold construction, based on syntenic alignments of PLBJ-1 and PLFJ-1.**

(DOCX)

**S3 Table. Comparison of repeat elements between PLBJ-1 and PLFJ-1.**

(DOCX)

**S4 Table. Glycoside hydrolases in *P. lilacinum* and comparison with other fungi.**  
(XLSX)

**S5 Table. Carbohydrate-binding modules in *P. lilacinum* and comparison with other fungi.**  
(XLSX)

**S6 Table. Carbohydrate esterases in *P. lilacinum* and comparison with other fungi.**  
(XLSX)

**S7 Table. Glycosyltransferases in *P. lilacinum* and comparison with other fungi.**  
(XLSX)

**S8 Table. Proteases in *P. lilacinum* and comparison with other fungi.**  
(XLSX)

**S9 Table. Paralogous expansion of some protein families in PLBJ-1 and PLFJ-1.**  
(DOCX)

**S10 Table. Major paralogous gene expansion in nematode parasitic fungi.**  
(DOCX)

**S11 Table. Biosynthetic gene clusters of secondary metabolites in *P. lilacinum*.**  
(DOCX)

**S12 Table. A-domain selectivity of LcsA predicted by NRSPredictor2.**  
(DOCX)

**S13 Table. Primer sequences used in this study.**  
(DOCX)

**S14 Table. Fungus strains used in the phylogenomic analysis.**  
(DOCX)

## Author Contributions

Conceived and designed the experiments: BX WBY. Performed the experiments: GW ZL EL. Analyzed the data: GW RL JL. Contributed reagents/materials/analysis tools: ZM YY. Wrote the paper: GW.

## References

1. Abad P, Gouzy J, Aury JM, Castagnone-Sereno P, Danchin EG, Deleury E, et al. Genome sequence of the metazoan plant-parasitic nematode *Meloidogyne incognita*. *Nat Biotechnol*. 2008 Aug; 26(8): 909–915. doi: [10.1038/nbt.1482](https://doi.org/10.1038/nbt.1482) PMID: [18660804](https://pubmed.ncbi.nlm.nih.gov/18660804/)
2. Anastasiadis IA, Giannakou IO, Prophetou-Athanasidou DA, Gowen SR. The combined effect of the application of a biocontrol agent *Paecilomyces lilacinus*, with various practices for the control of root-knot nematodes. *Crop Prot*. 2008; 27(3–5): 352–361. doi: [10.1016/j.cropro.2007.06.008](https://doi.org/10.1016/j.cropro.2007.06.008)
3. Noling J, Becker J. The challenge of research and extension to define and implement alternatives to methyl bromide. *J Nematol*. 1994; 26(4S): 573. PMID: [PMC2619561](https://pubmed.ncbi.nlm.nih.gov/PMC2619561/)
4. Singh S, Pandey RK, Goswami BK. Bio-control activity of *Purpureocillium lilacinum* strains in managing root-knot disease of tomato caused by *Meloidogyne incognita*. *Biocontrol Sci Technol*. 2013; 23(12): 1469–1489. doi: [10.1080/09583157.2013.840770](https://doi.org/10.1080/09583157.2013.840770)
5. Castillo Lopez D, Zhu-Salzman K, Ek-Ramos MJ, Sword GA. The entomopathogenic fungal endophytes *Purpureocillium lilacinum* (formerly *Paecilomyces lilacinus*) and *Beauveria bassiana* negatively affect cotton aphid reproduction under both greenhouse and field conditions. *PLoS One*. 2014; 9(8): e103891. doi: [10.1371/journal.pone.0103891](https://doi.org/10.1371/journal.pone.0103891) PMID: [25093505](https://pubmed.ncbi.nlm.nih.gov/25093505/)



6. Fiedler Ž, Sosnowska D. Nematophagous fungus *Paecilomyces lilacinus* (Thom) Samson is also a biological agent for control of greenhouse insects and mite pests. *BioControl*. 2007; 52(4): 547–558. doi: [10.1007/s10526-006-9052-2](https://doi.org/10.1007/s10526-006-9052-2)
7. Goffré D, Folgarait P. *Purpureocillium lilacinum*, potential agent for biological control of the leaf-cutting ant *Acromyrmex lundii*. *J Invertebr Pathol*. 2015; 130: 107–115. doi: [10.1016/j.jip.2015.07.008](https://doi.org/10.1016/j.jip.2015.07.008) PMID: [26205173](https://pubmed.ncbi.nlm.nih.gov/26205173/)
8. Luangsa-Ard J, Houbraken J, van Doorn T, Hong SB, Borman AM, Hywel-Jones NL, et al. *Purpureocillium*, a new genus for the medically important *Paecilomyces lilacinus*. *FEMS Microbiol Lett*. 2011 Aug; 321(2): 141–149. doi: [10.1111/j.1574-6968.2011.02322.x](https://doi.org/10.1111/j.1574-6968.2011.02322.x) PMID: [21631575](https://pubmed.ncbi.nlm.nih.gov/21631575/)
9. Bonants PJ, Fitters PF, Thijs H, den Belder E, Waalwijk C, Hentfling JWD. A basic serine protease from *Paecilomyces lilacinus* with biological activity against *Meloidogyne hapla* eggs. *Microbiology*. 1995; 141(4): 775–784.
10. Yang J, Zhao X, Liang L, Xia Z, Lei L, Niu X, et al. Overexpression of a cuticle-degrading protease Ver112 increases the nematocidal activity of *Paecilomyces lilacinus*. *Appl Microbiol Biotechnol*. 2011 Mar; 89(6): 1895–1903. PMID: [21110018](https://pubmed.ncbi.nlm.nih.gov/21110018/) doi: [10.1007/s00253-010-3012-6](https://doi.org/10.1007/s00253-010-3012-6)
11. Dong L, Yang J, Zhang K. Cloning and phylogenetic analysis of the chitinase gene from the facultative pathogen *Paecilomyces lilacinus*. *J Appl Microbiol*. 2007; 103(6): 2476–2488. doi: [10.1111/j.1365-2672.2007.03514.x](https://doi.org/10.1111/j.1365-2672.2007.03514.x) PMID: [18045433](https://pubmed.ncbi.nlm.nih.gov/18045433/)
12. Park JO, Hargreaves JR, McConville EJ, Stirling GR, Ghisalberti EL, Sivasithamparam K. Production of leucinostatins and nematocidal activity of Australian isolates of *Paecilomyces lilacinus* (Thom) Samson. *Lett Appl Microbiol*. 2004; 38(4): 271–276. doi: [10.1111/j.1472-765X.2004.01488.x](https://doi.org/10.1111/j.1472-765X.2004.01488.x) PMID: [15214724](https://pubmed.ncbi.nlm.nih.gov/15214724/)
13. Teles AP, Takahashi JA. Paecilomide, a new acetylcholinesterase inhibitor from *Paecilomyces lilacinus*. *Microbiol Res*. 2013 May; 168(4): 204–210. PMID: [23219197](https://pubmed.ncbi.nlm.nih.gov/23219197/) doi: [10.1016/j.micres.2012.11.007](https://doi.org/10.1016/j.micres.2012.11.007)
14. Madariaga-Mazon A, Gonzalez-Andrade M, Gonzalez Mdel C, Glenn AE, Cerda-Garcia-Rojas CM, Mata R. Absolute configuration of acremoxanthone C, a potent calmodulin inhibitor from *Purpureocillium lilacinum*. *J Nat Prod*. 2013 Aug; 76(8): 1454–1460. PMID: [23876004](https://pubmed.ncbi.nlm.nih.gov/23876004/) doi: [10.1021/np4002477](https://doi.org/10.1021/np4002477)
15. Arai T, Mikami Y, Fukushima K, Utsumi, Yazawa K. A new antibiotic, leucinostatin, derived from *Penicillium lilacinum*. *J Antibiot*. 1973; 26(3): 157–161. PMID: [4783199](https://pubmed.ncbi.nlm.nih.gov/4783199/)
16. Casinovi C, Rossi C, Tuttobello C, Ricci M. The structure of Leucinostatin C. A minor peptide from *Paecilomyces marquandii*. *Eur J Med Chem*. 1986; 21(6): 527–528.
17. Rossi C, Tuttobello L, Ricci M, Casinovi CG, Radios L. Leucinostatin D, A novel peptide antibiotic from *Paecilomyces marquandii*. *J Antibiot*. 1987; 40(1): 130–133. PMID: [3558114](https://pubmed.ncbi.nlm.nih.gov/3558114/)
18. Radics L, Kajtar-Peredy M, Casinovi C, Rossi C, Ricci M, Tuttobello L. Leucinostatins H and K, two novel peptide antibiotics with tertiary amine-oxide terminal group from *Paecilomyces marquandii* isolation, structure and biological activity. *J Antibiot*. 1987; 40(5): 714–716. PMID: [3610826](https://pubmed.ncbi.nlm.nih.gov/3610826/)
19. Strobel GA, Torczynski R, Bollon A. *Acremonium* sp. -a leucinostatin A producing endophyte of European yew (*Taxus baccata*). *Plant Sci*. 1997; 128(1): 97–108. doi: [10.1016/S0168-9452\(97\)00131-3](https://doi.org/10.1016/S0168-9452(97)00131-3)
20. Martinez AF, Moraes LA. Liquid chromatography-tandem mass spectrometry characterization of five new leucinostatins produced by *Paecilomyces lilacinus* CG-189. *J Antibiot*. 2015 Mar; 68(3): 178–184. doi: [10.1038/ja.2014.120](https://doi.org/10.1038/ja.2014.120) PMID: [25182486](https://pubmed.ncbi.nlm.nih.gov/25182486/)
21. Kawada M, Inoue H, Ohba S, Masuda T, Momose I, Ikeda D. Leucinostatin A inhibits prostate cancer growth through reduction of insulin-like growth factor-I expression in prostate stromal cells. *Int J Cancer*. 2010 Feb; 126(4): 810–818. doi: [10.1002/ijc.24915](https://doi.org/10.1002/ijc.24915) PMID: [19795463](https://pubmed.ncbi.nlm.nih.gov/19795463/)
22. Ishiyama A, Otoguro K, Iwatsuki M, Namatame M, Nishihara A, Nonaka K, et al. *In vitro* and *in vivo* antitrypanosomal activities of three peptide antibiotics: leucinostatin A and B, alamethicin I and tsushimycin. *J Antibiot*. 2009 Jun; 62(6): 303–308. PMID: [19407848](https://pubmed.ncbi.nlm.nih.gov/19407848/). doi: [10.1038/ja.2009.32](https://doi.org/10.1038/ja.2009.32)
23. Fukushima K, Arai T, Mori Y, Tsuboi M, Suzuki M. Studies on peptide antibiotics, leucinostatins. I. Separation, physico-chemical properties and biological activities of leucinostatins A and B. *J Antibiot*. 1983; 36(12): 1606–1612. PMID: [6662804](https://pubmed.ncbi.nlm.nih.gov/6662804/)
24. Prasad P, Varshney D, Adholeya A. Whole genome annotation and comparative genomic analyses of bio-control fungus *Purpureocillium lilacinum*. *BMC Genomics*. 2015; 16(1): 1. doi: [10.1186/s12864-015-2229-2](https://doi.org/10.1186/s12864-015-2229-2)
25. Larriba E, Jaime MDLA, Carbonell-Caballero J, Conesa A, Dopazo J, Nislow C, et al. Sequencing and functional analysis of the genome of a nematode egg-parasitic fungus, *Pochonia chlamydosporia*. *Fungal Genet and Biol*. 2014; 65: 69–80. doi: [10.1016/j.fgb.2014.02.002](https://doi.org/10.1016/j.fgb.2014.02.002)
26. Lai Y, Liu K, Zhang X, Zhang X, Li K, Wang N, et al. Comparative genomics and transcriptomics analyses reveal divergent lifestyle features of nematode endoparasitic fungus *Hirsutiella minnesotensis*. *Genome Biol Evol*. 2014; 6(11): 3077–3093. doi: [10.1093/gbe/evu241](https://doi.org/10.1093/gbe/evu241) PMID: [25359922](https://pubmed.ncbi.nlm.nih.gov/25359922/)

27. Vleeshouwers VG, Raffaele S, Vossen JH, Champouret N, Oliva R, Segretin ME, et al. Understanding and exploiting late blight resistance in the age of effectors. *Annu Rev Phytopathol.* 2011; 49: 507–531. doi: [10.1146/annurev-phyto-072910-095326](https://doi.org/10.1146/annurev-phyto-072910-095326) PMID: [21663437](https://pubmed.ncbi.nlm.nih.gov/21663437/)
28. Khan A, Williams KL, Nevalainen HK. Effects of *Paecilomyces lilacinus* protease and chitinase on the eggshell structures and hatching of *Meloidogyne javanica* juveniles. *Biol Control.* 2004; 31(3): 346–352. doi: [10.1016/j.biocontrol.2004.07.011](https://doi.org/10.1016/j.biocontrol.2004.07.011)
29. Xiang MC, Yang XH, Wang ZX, Liu XZ, Chen SY, Xiao QM. Variability of morphology, parasitism, and nucleotide sequences among isolates and species of nematophagous *Hirsutella*. *Biol Control.* 2007; 41(1): 110–119. doi: [10.1016/j.biocontrol.2006.12.016](https://doi.org/10.1016/j.biocontrol.2006.12.016)
30. Bird AF, McClure M. The tylenchid (Nematoda) egg shell: structure, composition and permeability. *Parasitology.* 1976; 72(01): 19–28. <http://dx.doi.org/10.1017/S0031182000043158>
31. Xiao G, Ying SH, Zheng P, Wang ZL, Zhang S, Xie XQ, et al. Genomic perspectives on the evolution of fungal entomopathogenicity in *Beauveria bassiana*. *Sci Rep.* 2012; 2. doi: [10.1038/srep00483](https://doi.org/10.1038/srep00483)
32. Martinez D, Berka RM, Henrissat B, Saloheimo M, Arvas M, Baker SE, et al. Genome sequencing and analysis of the biomass-degrading fungus *Trichoderma reesei* (syn. *Hypocrea jecorina*). *Nat Biotechnol.* 2008; 26(5): 553–560. doi: [10.1038/nbt1403](https://doi.org/10.1038/nbt1403) PMID: [18454138](https://pubmed.ncbi.nlm.nih.gov/18454138/)
33. Yang J, Huang X, Tian B, Wang M, Niu Q, Zhang K. Isolation and characterization of a serine protease from the nematophagous fungus, *Lecanicillium psalliotae*, displaying nematocidal activity. *Biotechnol Lett.* 2005; 27(15): 1123–1128. doi: [10.1007/s10529-005-8461-0](https://doi.org/10.1007/s10529-005-8461-0) PMID: [16132863](https://pubmed.ncbi.nlm.nih.gov/16132863/)
34. Baldwin TK, Winnenburg R, Urban M, Rawlings C, Koehler J, Hammond-Kosack KE. The pathogen-host interactions database (PHI-base) provides insights into generic and novel themes of pathogenicity. *Mol Plant Microbe In.* 2006; 19(12): 1451–1462. doi: [10.1094/MPMI-19-1451](https://doi.org/10.1094/MPMI-19-1451)
35. Ohm RA, Feau N, Henrissat B, Schoch CL, Horwitz BA, Barry KW, et al. Diverse lifestyles and strategies of plant pathogenesis encoded in the genomes of eighteen Dothideomycetes fungi. *PLoS Pathog.* 2012; 8(12): e1003037. doi: [10.1371/journal.ppat.1003037](https://doi.org/10.1371/journal.ppat.1003037) PMID: [23236275](https://pubmed.ncbi.nlm.nih.gov/23236275/)
36. Hu X, Zhang Y, Xiao G, Zheng P, Xia Y, Zhang X, et al. Genome survey uncovers the secrets of sex and lifestyle in caterpillar fungus. *Chinese Sci Bul.* 2013; 58(23): 2846–2854. doi: [10.1007/s11434-013-5929-5](https://doi.org/10.1007/s11434-013-5929-5)
37. De Bekker C, Ohm RA, Loreto RG, Sebastian A, Albert I, Mew M, et al. Gene expression during zombie ant biting behavior reflects the complexity underlying fungal parasitic behavioral manipulation. *BMC Genomics.* 2015; 16(1): 620. doi: [10.1186/s12864-015-1812-x](https://doi.org/10.1186/s12864-015-1812-x)
38. Lespinet O, Wolf YI, Koonin EV, Aravind L. The role of lineage-specific gene family expansion in the evolution of eukaryotes. *Genome Res.* 2002; 12(7): 1048–1059. doi: [10.1101/gr.174302](https://doi.org/10.1101/gr.174302) PMID: [12097341](https://pubmed.ncbi.nlm.nih.gov/12097341/)
39. Bushley KE, Raja R, Jaiswal P, Cumbie JS, Nonogaki M, Boyd AE, et al. The genome of *Tolypocladium inflatum*: evolution, organization, and expression of the cyclosporin biosynthetic gene cluster. *PLoS Genet.* 2013 Jun; 9(6): e1003496. doi: [10.1371/journal.pgen.1003496](https://doi.org/10.1371/journal.pgen.1003496) PMID: [23818858](https://pubmed.ncbi.nlm.nih.gov/23818858/)
40. Weber T, Blin K, Duddela S, Duddela S, Kim HU, Bruccoleri R, et al. antiSMASH 3.0—a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Res.* 2015: gkv437. doi: [10.1093/nar/gkv437](https://doi.org/10.1093/nar/gkv437)
41. Shimizu T, Kinoshita H, Ishihara S, Sakai K, Nagai S, Nihira T. Polyketide synthase gene responsible for citrinin biosynthesis in *Monascus purpureus*. *Appl Environ Microb.* 2005; 71(7): 3453–3457. doi: [10.1128/AEM.71.7.3453-3457.2005](https://doi.org/10.1128/AEM.71.7.3453-3457.2005)
42. Chooi YH, Cacho R, Tang Y. Identification of the viridicatumtoxin and griseofulvin gene clusters from *Penicillium aethiopicum*. *Chem Biol.* 2010; 17(5): 483–494. doi: [10.1016/j.chembiol.2010.03.015](https://doi.org/10.1016/j.chembiol.2010.03.015) PMID: [20534346](https://pubmed.ncbi.nlm.nih.gov/20534346/)
43. Proctor RH, Desjardins AE, Plattner RD, Hohn TM. A polyketide synthase gene required for biosynthesis of fumonisin mycotoxins in *Gibberella fujikuroi* mating population A. *Fungal Genet and Biol.* 1999; 27(1): 100–112. doi: [10.1006/fgbi.1999.1141](https://doi.org/10.1006/fgbi.1999.1141)
44. Szekeres A, Leitgeb B, Kredics L, Antal Z, Hatvani L, Manczinger L, et al. Peptaibols and related peptaibiotics of *Trichoderma*. *Acta Microbiol Imm H.* 2005; 52(2): 137–168. <http://dx.doi.org/10.1556/AMicr.52.2005.2.2>
45. Quandt CA, Bushley KE, Spatafora JW. The genome of the truffle-parasite *Tolypocladium ophioglossoides* and the evolution of antifungal peptaibiotics. *BMC Genomics.* 2015; 16(1): 553. doi: [10.1186/s12864-015-1777-9](https://doi.org/10.1186/s12864-015-1777-9)
46. Chiang YM, Szewczyk E, Nayak T, Davidson AD, Sanchez JF, Lo HC, et al. Molecular genetic mining of the *Aspergillus* secondary metabolome: discovery of the emericeamide biosynthetic pathway. *Chem Biol.* 2008; 15(6): 527–532. doi: [10.1016/j.chembiol.2008.05.010](https://doi.org/10.1016/j.chembiol.2008.05.010) PMID: [18559263](https://pubmed.ncbi.nlm.nih.gov/18559263/)

47. Chen L, Yue Q, Zhang X, Xiang M, Wang C, Li S, et al. Genomics-driven discovery of the pneumocandin biosynthetic gene cluster in the fungus *Glarea lozoyensis*. BMC Genomics. 2013; 14:339. doi: [10.1186/1471-2164-14-339](https://doi.org/10.1186/1471-2164-14-339) PMID: [23688303](https://pubmed.ncbi.nlm.nih.gov/23688303/)
48. Röttig M, Medema MH, Blin K, Weber T, Rausch C, Kohlbacher O. NRPSpredictor2- a web server for predicting NRPS adenylation domain specificity. Nucleic Acids Res. 2011 Jul; 39: W362–367. doi: [10.1093/nar/gkr323](https://doi.org/10.1093/nar/gkr323) PMID: [21558170](https://pubmed.ncbi.nlm.nih.gov/21558170/)
49. Liang L, Li J, Cheng L, Ling J, Luo Z, Bai M, et al. A high efficiency gene disruption strategy using a positive–negative split selection marker and electroporation for *Fusarium oxysporum*. Microbiol Res. 2014; 169(11): 835–843. doi: [10.1016/j.micres.2014.03.004](https://doi.org/10.1016/j.micres.2014.03.004) PMID: [24755311](https://pubmed.ncbi.nlm.nih.gov/24755311/)
50. Khaldi N, Seifuddin FT, Turner G, Haft D, Nierman WC, Wolfe KH, et al. SMURF: genomic mapping of fungal secondary metabolite clusters. Fungal Genet and Biol. 2010; 47(9): 736–741. doi: [10.1016/j.fgb.2010.06.003](https://doi.org/10.1016/j.fgb.2010.06.003)
51. Bode HB, Bethe B, Höfs R, Zeeck A. Big effects from small changes: possible ways to explore nature's chemical diversity. ChemBioChem. 2002; 3(7): 619–627. doi: [10.1002/1439-7633\(20020703\)3:7<619::AID-CBIC619>3.0.CO;2-9](https://doi.org/10.1002/1439-7633(20020703)3:7<619::AID-CBIC619>3.0.CO;2-9) PMID: [12324995](https://pubmed.ncbi.nlm.nih.gov/12324995/)
52. Aravind L, Koonin EV. The HD domain defines a new superfamily of metal-dependent phosphohydrolases. Trends Biochem Sci. 1998; 23(12): 469–472. PMID: [9868367](https://pubmed.ncbi.nlm.nih.gov/9868367/)
53. Chiang YM, Szewczyk E, Nayak T, Davidson AD, Sanchez JF, Lo HC, et al. Molecular genetic mining of the Aspergillus secondary metabolome: discovery of the emericellamide biosynthetic pathway. Chem Biol. 2008; 15(6): 527–532. doi: [10.1016/j.chembiol.2008.05.010](https://doi.org/10.1016/j.chembiol.2008.05.010) PMID: [18559263](https://pubmed.ncbi.nlm.nih.gov/18559263/)
54. Wittmann M, Linne U, Pohlmann V, Marahiel MA. Role of DptE and DptF in the lipidation reaction of daptomycin. FEBS J. 2008; 275(21): 5343–5354. doi: [10.1111/j.1742-4658.2008.06664.x](https://doi.org/10.1111/j.1742-4658.2008.06664.x) PMID: [18959760](https://pubmed.ncbi.nlm.nih.gov/18959760/)
55. Cacho RA, Jiang W, Chooi YH, Walsh CT, Tang Y. Identification and characterization of the echinocandin B biosynthetic gene cluster from *Emericella rugulosa* NRRL 11440. J Am Chem Soc. 2012; 134(40): 16781–16790. doi: [10.1021/ja307220z](https://doi.org/10.1021/ja307220z) PMID: [22998630](https://pubmed.ncbi.nlm.nih.gov/22998630/)
56. Sørensen JL, Sondergaard TE, Covarelli L, Fuertes PR, Hansen FT, Frandsen RJN, et al. Identification of the biosynthetic gene clusters for the lipopeptides fusaristatin A and W493 B in *Fusarium graminearum* and *F. pseudograminearum*. J Nat Prod. 2014; 77(12): 2619–2625. doi: [10.1021/np500436r](https://doi.org/10.1021/np500436r) PMID: [25412204](https://pubmed.ncbi.nlm.nih.gov/25412204/)
57. Keller NP, Turner G, Bennett JW. Fungal secondary metabolism—from biochemistry to genomics. Nat Rev Microbiol. 2005 Dec; 3(12): 937–947. doi: [10.1038/nrmicro1286](https://doi.org/10.1038/nrmicro1286) PMID: [16322742](https://pubmed.ncbi.nlm.nih.gov/16322742/)
58. Wang X, Wu F, Liu L, Liu X, Che Y, Keller NP, et al. The bZIP transcription factor PfZipA regulates secondary metabolism and oxidative stress response in the plant endophytic fungus *Pestalotiopsis fici*. Fungal Genet and Biol. 2015; 81: 221–228. doi: [10.1016/j.fgb.2015.03.010](https://doi.org/10.1016/j.fgb.2015.03.010)
59. Ron EZ, Rosenberg E. Natural roles of biosurfactants. Environmental microbiology. 2001; 3(4): 229–236. PMID: [11359508](https://pubmed.ncbi.nlm.nih.gov/11359508/)
60. Tran H, Ficke A, Asimwe T, Höfte M, Raaijmakers JM. Role of the cyclic lipopeptide massetolide A in biological control of *Phytophthora infestans* and in colonization of tomato plants by *Pseudomonas fluorescens*. New Phytol. 2007; 175(4): 731–742. doi: [10.1111/j.1469-8137.2007.02138.x](https://doi.org/10.1111/j.1469-8137.2007.02138.x) PMID: [17688588](https://pubmed.ncbi.nlm.nih.gov/17688588/)
61. Nguyen XH, Naing KW, Lee YS, Kim YH, Moon JH, Kim KY. Antagonism of antifungal metabolites from *Streptomyces griseus* H7602 against *Phytophthora capsici*. J Basic Microb. 2015; 55(1): 45–53. doi: [10.1002/jobm.201300820](https://doi.org/10.1002/jobm.201300820)
62. Tunlid A, Talbot NJ. Genomics of parasitic and symbiotic fungi. Curr Opin Microbiol. 2002; 5(5): 513–519. doi: [10.1016/S1369-5274\(02\)00355-7](https://doi.org/10.1016/S1369-5274(02)00355-7) PMID: [12354560](https://pubmed.ncbi.nlm.nih.gov/12354560/)
63. Gao Q, Jin K, Ying SH, Zhang Y, Xiao G, Shang Y, et al. Genome sequencing and comparative transcriptomics of the model entomopathogenic fungi *Metarhizium anisopliae* and *M. acridum*. PLoS Genet. 2011; 7(1): e1001264. doi: [10.1371/journal.pgen.1001264](https://doi.org/10.1371/journal.pgen.1001264) PMID: [21253567](https://pubmed.ncbi.nlm.nih.gov/21253567/)
64. Niu XM, Wang YL, Chu YS, Xue HX, Li N, Wei LX, et al. Nematodetoxic aurovertin-type metabolites from a root-knot nematode parasitic fungus *Pochonia chlamydosporia*. J Agr Food Chem. 2009; 58(2): 828–834. doi: [10.1021/jf903259n](https://doi.org/10.1021/jf903259n)
65. Isaka M, Palasarn S, Auncharoen P, Komwijit S, Jones EG. Acremoxanthones A and B, novel antibiotic polyketides from the fungus *Acremonium* sp. BCC 31806. Tetrahedron Lett. 2009; 50(3): 284–287. doi: [10.1016/j.tetlet.2008.10.146](https://doi.org/10.1016/j.tetlet.2008.10.146)
66. Tabata N, Tomoda H, Matsuzaki K, Omura S. Structure and biosynthesis of xanthoquinodins, anticoccidial antibiotics. J Am Chem Soc. 1993; 115(19): 8558–8564. doi: [10.1021/ja00072a006](https://doi.org/10.1021/ja00072a006)
67. Yang CH, Li M, Wei ZP, Feng H, Jing G. Advances in isolation and synthesis of xanthone derivatives. Chinese Herbal Medicines. 2012; 4(2): 87–102. doi: [10.3969/j.issn.1674-6384.2012.02.003](https://doi.org/10.3969/j.issn.1674-6384.2012.02.003)

68. Chiang YM, Szewczyk E, Davidson AD, Keller N, Oakley BR, Wang CC. A gene cluster containing two fungal polyketide synthases encodes the biosynthetic pathway for a polyketide, asperfuranone, in *Aspergillus nidulans*. *J Am Chem Soc.* 2009; 131(8): 2965–2970. doi: [10.1021/ja8088185](https://doi.org/10.1021/ja8088185) PMID: [19199437](https://pubmed.ncbi.nlm.nih.gov/19199437/)
69. Zabala AO, Xu W, Chooi YH, Tang Y. Characterization of a silent azaphilone gene cluster from *Aspergillus niger* ATCC 1015 reveals a hydroxylation-mediated pyran-ring formation. *Chem Biol.* 2012; 19(8): 1049–1059. doi: [10.1016/j.chembiol.2012.07.004](https://doi.org/10.1016/j.chembiol.2012.07.004) PMID: [22921072](https://pubmed.ncbi.nlm.nih.gov/22921072/)
70. Winter JM, Sato M, Sugimoto S, Chiou G, Garg NK, Tang Y, et al. Identification and characterization of the chaetoviridin and chaetomugilin gene cluster in *Chaetomium globosum* reveal dual functions of an iterative highly-reducing polyketide synthase. *J Am Chem Soc.* 2012; 134(43): 17900–17903. doi: [10.1021/ja3090498](https://doi.org/10.1021/ja3090498) PMID: [23072467](https://pubmed.ncbi.nlm.nih.gov/23072467/)
71. Bushley KE, Ripoll DR, Turgeon BG. Module evolution and substrate specificity of fungal nonribosomal peptide synthetases involved in siderophore biosynthesis. *BMC Evol Biol.* 2008; 8(1): 328. doi: [10.1186/1471-2148-8-328](https://doi.org/10.1186/1471-2148-8-328)
72. Finking R, Marahiel MA. Biosynthesis of nonribosomal peptides. *Annu Rev Microbiol.* 2004; 58: 453–488. doi: [10.1146/annurev.micro.58.030603.123615](https://doi.org/10.1146/annurev.micro.58.030603.123615) PMID: [15487945](https://pubmed.ncbi.nlm.nih.gov/15487945/)
73. Chen Y, McClure RA, Zheng Y, Thomson RJ, Kelleher NL. Proteomics guided discovery of flavopeptins: Anti-proliferative aldehydes synthesized by a reductase domain-containing non-ribosomal peptide synthetase. *J Am Chem Soc.* 2013; 135(28): 10449–10456. doi: [10.1021/ja4031193](https://doi.org/10.1021/ja4031193) PMID: [23763305](https://pubmed.ncbi.nlm.nih.gov/23763305/)
74. Du L, Lou L. PKS and NRPS release mechanisms. *Nat Prod Rep.* 2010; 27(2): 255–278. doi: [10.1039/B912037H](https://doi.org/10.1039/B912037H) PMID: [20111804](https://pubmed.ncbi.nlm.nih.gov/20111804/)
75. Evans BS, Ntai I, Chen Y, Robinson SJ, Kelleher NL. Proteomics-based discovery of koranimine, a cyclic imine natural product. *J Am Chem Soc.* 2011; 133(19): 7316–7319. doi: [10.1021/ja2015795](https://doi.org/10.1021/ja2015795) PMID: [21520944](https://pubmed.ncbi.nlm.nih.gov/21520944/)
76. Li Y, Weissman KJ, Müller R. Myxochelin biosynthesis: direct evidence for two-and four-electron reduction of a carrier protein-bound thioester. *J Am Chem Soc.* 2008; 130(24): 7554–7555. doi: [10.1021/ja8025278](https://doi.org/10.1021/ja8025278) PMID: [18498160](https://pubmed.ncbi.nlm.nih.gov/18498160/)
77. Luesch H, Hoffmann D, Hevel JM, Becker JE, Golakoti T, Moore RE. Biosynthesis of 4-Methylproline in Cyanobacteria: Cloning of nosE and nosF genes and biochemical characterization of the encoded dehydrogenase and reductase activities. *J Org Chem.* 2003; 68(1): 83–91. doi: [10.1021/jo026479q](https://doi.org/10.1021/jo026479q) PMID: [12515465](https://pubmed.ncbi.nlm.nih.gov/12515465/)
78. Liu L, Jokela J, Herfindal L, Wahlsten M, Sinkkonen J, Permi P, et al. 4-methylproline guided natural product discovery: Co-occurrence of 4-hydroxy-and 4-methylprolines in nostoweipeptins and nostopeptolides. *ACS Chem Biol.* 2014; 9(11): 2646–2655. doi: [10.1021/cb500436p](https://doi.org/10.1021/cb500436p) PMID: [25203327](https://pubmed.ncbi.nlm.nih.gov/25203327/)
79. Adefarati AA, Hensens OD, Jones ETT, Tkacz JS. Pneumocandins from *Zalerion arboricola*. V. Glutamic acid-and leucine-derived amino acids in pneumocandin A0 (L-671,329) and distinct origins of the substituted proline residues in pneumocandins A0 and B0. *The J Antibiot.* 1992; 45(12): 1953–1957. PMID: [1362720](https://pubmed.ncbi.nlm.nih.gov/1362720/)
80. Wang B, Kang Q, Lu Y, Bai L, Wang C. Unveiling the biosynthetic puzzle of destruxins in *Metarhizium* species. *P Natl Acad Sci.* 2012 Jan; 109(4): 1287–1292. doi: [10.1073/pnas.1115983109](https://doi.org/10.1073/pnas.1115983109)
81. Gnerre S, MacCallum I, Przybylski D, Ribeiro FJ, Burton JN, Walker BJ, et al. High-quality draft assemblies of mammalian genomes from massively parallel sequence data. *P Natl Acad Sci.* 2011; 108(4): 1513–1518. doi: [10.1073/pnas.1017351108](https://doi.org/10.1073/pnas.1017351108)
82. Zheng A, Lin R, Zhang D, Qin P, Xu L, Ai P, et al. The evolution and pathogenic mechanisms of the rice sheath blight pathogen. *Nat Commun.* 2013; 4:1424. doi: [10.1038/ncomms2427](https://doi.org/10.1038/ncomms2427) PMID: [23361014](https://pubmed.ncbi.nlm.nih.gov/23361014/)
83. Edgar RC, Myers EW. PILER: identification and classification of genomic repeats. *Bioinformatics.* 2005; 21(suppl 1): i152–i158. doi: [10.1093/bioinformatics/bti1003](https://doi.org/10.1093/bioinformatics/bti1003) PMID: [15961452](https://pubmed.ncbi.nlm.nih.gov/15961452/)
84. Price AL, Jones NC, Pevzner PA. De novo identification of repeat families in large genomes. *Bioinformatics.* 2005; 21(suppl 1): i351–i358. doi: [10.1093/bioinformatics/bti1018](https://doi.org/10.1093/bioinformatics/bti1018) PMID: [15961478](https://pubmed.ncbi.nlm.nih.gov/15961478/)
85. Jurka J, Kapitonov VV, Pavlicek A, Klonowski P, Kohany O, Walichiewicz J. Repbase Update, a database of eukaryotic repetitive elements. *Cytogenet Genome Res.* 2005; 110(1–4): 462–467. PMID: [16093699](https://pubmed.ncbi.nlm.nih.gov/16093699/)
86. Stanke M, Keller O, Gunduz I, Hayes A, Waack S, Morgenstern B. AUGUSTUS: ab initio prediction of alternative transcripts. *Nucleic Acids Res.* 2006; 34(suppl 2): W435–W439. doi: [10.1093/nar/gkl200](https://doi.org/10.1093/nar/gkl200)
87. Ter-Hovhannisyan V, Lomsadze A, Chernoff YO, Borodovsky M. Gene prediction in novel fungal genomes using an ab initio algorithm with unsupervised training. *Genome Res.* 2008; 18(12): 1979–1990. doi: [10.1101/gr.081612.108](https://doi.org/10.1101/gr.081612.108) PMID: [18757608](https://pubmed.ncbi.nlm.nih.gov/18757608/)

88. Schiex T, Moisan A, Rouzé P. EuGene: an eukaryotic gene finder that combines several sources of evidence. *Comput Biol.* 2001; 2066: 111–125.
89. Pedersen AG, Nielsen H, editors. Neural network prediction of translation initiation sites in eukaryotes: perspectives for EST and genome analysis. *Ismb.* 1997; Citeseer.
90. Peng Y, Leung HC, Yiu S-M, Lv M-J, Zhu X-G, Chin FY. IDBA-tran: a more robust de novo de Bruijn graph assembler for transcriptomes with uneven expression levels. *Bioinformatics.* 2013; 29(13): i326–i334. doi: [10.1093/bioinformatics/btt219](https://doi.org/10.1093/bioinformatics/btt219) PMID: [23813001](https://pubmed.ncbi.nlm.nih.gov/23813001/)
91. Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics.* 2009; 25(9): 1105–1111. doi: [10.1093/bioinformatics/btp120](https://doi.org/10.1093/bioinformatics/btp120) PMID: [19289445](https://pubmed.ncbi.nlm.nih.gov/19289445/)
92. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, Van Baren MJ, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol.* 2010; 28(5): 511–515. doi: [10.1038/nbt.1621](https://doi.org/10.1038/nbt.1621) PMID: [20436464](https://pubmed.ncbi.nlm.nih.gov/20436464/)
93. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene Ontology: tool for the unification of biology. *Nat Genet.* 2000; 25(1): 25–29. doi: [10.1038/75556](https://doi.org/10.1038/75556) PMID: [10802651](https://pubmed.ncbi.nlm.nih.gov/10802651/)
94. Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics.* 2005; 21(18): 3674–3676. doi: [10.1093/bioinformatics/bti610](https://doi.org/10.1093/bioinformatics/bti610) PMID: [16081474](https://pubmed.ncbi.nlm.nih.gov/16081474/)
95. Koonin EV, Fedorova ND, Jackson JD, Jacobs AR, Krylov DM, Makarova KS, et al. A comprehensive evolutionary classification of proteins encoded in complete eukaryotic genomes. *Genome Biol.* 2004; 5(2): R7. PMID: [14759257](https://pubmed.ncbi.nlm.nih.gov/14759257/)
96. Park BH, Karpinets TV, Syed MH, Leuze MR, Uberbacher EC. CAZymes Analysis Toolkit (CAT): web service for searching and analyzing carbohydrate-active enzymes in a newly sequenced organism using CAZy database. *Glycobiology.* 2010; 20(12): 1574–1584. doi: [10.1093/glycob/cwq106](https://doi.org/10.1093/glycob/cwq106) PMID: [20696711](https://pubmed.ncbi.nlm.nih.gov/20696711/)
97. Rawlings ND, Waller M, Barrett AJ, Bateman A. MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res.* 2013:gkt953. doi: [10.1093/nar/gkr987](https://doi.org/10.1093/nar/gkr987)
98. Nelson DR. The cytochrome p450 homepage. *Human genomics.* 2009; 4(1): 59. doi: [10.1186/1479-7364-4-1-59](https://doi.org/10.1186/1479-7364-4-1-59) PMID: [19951895](https://pubmed.ncbi.nlm.nih.gov/19951895/)
99. Urban M, Pant R, Raghunath A, Irvine AG, Pedro H, Hammond-Kosack KE. The Pathogen-Host Interactions database (PHI-base): additions and future developments. *Nucleic Acids Res.* 2014:gku1165. doi: [10.1093/nar/gku1165](https://doi.org/10.1093/nar/gku1165)
100. Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods.* 2011; 8(10): 785–786. doi: [10.1038/nmeth.1701](https://doi.org/10.1038/nmeth.1701) PMID: [21959131](https://pubmed.ncbi.nlm.nih.gov/21959131/)
101. Emanuelsson O, Nielsen H, Brunak S, von Heijne G. Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J Mol Biol.* 2000; 300(4): 1005–1016. doi: [10.1006/jmbi.2000.3903](https://doi.org/10.1006/jmbi.2000.3903) PMID: [10891285](https://pubmed.ncbi.nlm.nih.gov/10891285/)
102. Käll L, Krogh A, Sonnhammer EL. A combined transmembrane topology and signal peptide prediction method. *J Mol Biol.* 2004; 338(5): 1027–1036. doi: [10.1016/j.jmb.2004.03.016](https://doi.org/10.1016/j.jmb.2004.03.016) PMID: [15111065](https://pubmed.ncbi.nlm.nih.gov/15111065/)
103. Hiller K, Grote A, Scheer M, Münch R, Jahn D. PrediSi: prediction of signal peptides and their cleavage positions. *Nucleic Acids Res.* 2004; 32(suppl 2): W375–W379. doi: [10.1093/nar/gkh378](https://doi.org/10.1093/nar/gkh378)
104. Krogh A, Larsson B, Von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol.* 2001; 305(3): 567–580. doi: [10.1006/jmbi.2000.4315](https://doi.org/10.1006/jmbi.2000.4315) PMID: [11152613](https://pubmed.ncbi.nlm.nih.gov/11152613/)
105. Li L, Stoeckert CJ, Roos DS. OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res.* 2003; 13(9): 2178–2189. doi: [10.1101/gr.1224503](https://doi.org/10.1101/gr.1224503) PMID: [12952885](https://pubmed.ncbi.nlm.nih.gov/12952885/)
106. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 2004; 32(5): 1792–1797. doi: [10.1093/nar/gkh340](https://doi.org/10.1093/nar/gkh340) PMID: [15034147](https://pubmed.ncbi.nlm.nih.gov/15034147/)
107. Castresana J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol.* 2000; 17(4): 540–552. PMID: [10742046](https://pubmed.ncbi.nlm.nih.gov/10742046/)
108. Schmidt HA, Strimmer K, Vingron M, von Haeseler A. TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. *Bioinformatics.* 2002; 18(3): 502–504. doi: [10.1093/bioinformatics/18.3.502](https://doi.org/10.1093/bioinformatics/18.3.502) PMID: [11934758](https://pubmed.ncbi.nlm.nih.gov/11934758/)
109. Li E, Ling J, Wang G, Xiao J, Yang Y, Mao Z, et al. Comparative proteomics analyses of two races of *Fusarium oxysporum* f. sp. *conglutinans* that differ in pathogenicity. *Sci Rep.* 2015; 5. doi: [10.1038/srep13663](https://doi.org/10.1038/srep13663)

110. Singh AK, Singh A, Joshi P. Combined application of chitinolytic bacterium *Paenibacillus* sp. D1 with low doses of chemical pesticides for better control of *Helicoverpa armigera*. *Int J Pest Manage.* 2016; 1–6. doi: [10.1080/09670874.2016.1167267](https://doi.org/10.1080/09670874.2016.1167267)
111. Bohannon RW, Pfaller BA. Documentation of Wound Surface Area from Tracings of Wound Perimeters Clinical Report on Three Techniques. *Phys Ther.* 1983; 63(10): 1622–1624. PMID: [6622538](https://pubmed.ncbi.nlm.nih.gov/6622538/)