

Article

Genomic Comparative Analysis of *Cordyceps pseudotenuipes* with Other Species from *Cordyceps*

Yingling Lu ^{1,2,3}, Yi Wang ^{3,*}, Xiaolong Yuan ³, Ou Huang ¹, Quanying Dong ^{1,2}, Dandan Li ¹, Shujin Ding ^{3,4}, Fuxian Ma ^{3,4} and Hong Yu ^{1,*}

¹ Yunnan Herbal Laboratory, College of Ecology and Environmental Sciences, Yunnan University, Kunming 650504, China

² School of Life Science, Yunnan University, Kunming 650504, China

³ Laboratory of Forest Plant Cultivation and Utilization, The Key Laboratory of Rare and Endangered Forest Plants of State Forestry Administration, Yunnan Academy of Forestry and Grassland, Kunming 650201, China

⁴ College of Forestry, Southwest Forestry University, Kunming 650224, China

* Correspondence: wangyi@ynlky.org.cn (Y.W.); hongyu@ynu.edu.cn (H.Y.); Tel.: +86-186-8716-3524 (Y.W.); +86-137-0067-6633 (H.Y.)

Abstract: The whole genome of *Cordyceps pseudotenuipes* was sequenced, annotated, and compared with three related species to characterize the genome. The antibiotics and Secondary Metabolites Analysis Shell (antiSMASH) and local BLAST analysis were used to explore the secondary metabolites (SMs) and biosynthesis gene clusters (BGCs) of the genus *Cordyceps*. The genome-wide basic characteristics of *C. pseudotenuipes*, *C. tenuipes*, *C. cicadae*, and *C. militaris* revealed unequal genome size, with *C. cicadae* as the largest (34.11 Mb), followed by *C. militaris* (32.27 Mb). However, the total gene lengths of *C. pseudotenuipes* and *C. tenuipes* were similar (30.1 Mb and 30.06 Mb). The GC contents of *C. pseudotenuipes*, *C. tenuipes*, *C. cicadae*, and *C. militaris* genomes differed slightly (51.40% to 54.11%). AntiSMASH and local BLAST analysis showed that *C. pseudotenuipes*, *C. tenuipes*, *C. cicadae*, and *C. militaris* had 31, 28, 31, and 29 putative SM BGCs, respectively. The SM BGCs contained different quantities of polyketide synthetase (PKS), nonribosomal peptide synthetase (NRPS), terpene, hybrid PKS + NRPS, and hybrid NRPS + Other. Moreover, *C. pseudotenuipes*, *C. tenuipes*, *C. cicadae*, and *C. militaris* had BGCs for the synthesis of dimethylcoprogen. *C. pseudotenuipes*, *C. tenuipes*, and *C. cicadae* had BGCs for the synthesis of leucinostatin A/B, neosartorin, dimethylcoprogen, wortmanamide A/B, and beauvericin. In addition, the SM BGCs unique to *C. pseudotenuipes* were clavarinic acid, communesin, and deoxynivalenol. Synteny analysis indicated that the scaffolds where the SM BGC was located were divided into more than 70 collinear blocks, and there might be rearrangements. Altogether, these findings improved our understanding of the molecular biology of the genus *Cordyceps* and will facilitate the discovery of new biologically active SMs from the genus *Cordyceps* using heterologous expression and gene knockdown methods.

Keywords: *Cordyceps*; whole-genome sequence; secondary metabolite; biosynthesis gene cluster



Citation: Lu, Y.; Wang, Y.; Yuan, X.; Huang, O.; Dong, Q.; Li, D.; Ding, S.; Ma, F.; Yu, H. Genomic Comparative Analysis of *Cordyceps pseudotenuipes* with Other Species from *Cordyceps*. *Metabolites* **2022**, *12*, 844. <https://doi.org/10.3390/metabo12090844>

Academic Editor: Hunter N. B. Moseley

Received: 17 August 2022

Accepted: 7 September 2022

Published: 8 September 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Cordyceps pseudotenuipes H. Yu, Q. Y. Dong, and Y. Wang is a new fungal species that was published in April 2022 [1]. The species was named after its morphological similarity to *C. tenuipes* (Peck) Kepler, B. Shrestha, and Spatafora. Taxonomically, this species was related to *C. tenuipes* and similar groups, including *C. cicadae* (Miq) Masee and *C. militaris* Fr. [1–3]. Genus *Cordyceps* had many bioactive components, including cordycepin [4], cordyceps polysaccharides [5], alkaloids [6], cordycepic acids [7], pentostatin [8], ophicordin [9], beauvericin [10], myriocin [11], beauveriolide [12], and oosporein [13] (Figure 1). These bioactive components function in various ways. For example, cordycepin was involved in synthesizing nucleic acid, platelet aggregation, and cell apoptosis and exhibits anticancer

activity by incorporating it into RNA synthesis to break it off [14–16]. Moreover, cordycepin also treated mice infected with *Trypanosoma evansi* [17]. Cordycepin and pentostatin were used as chemotherapeutic drugs for leukemia [18], and combining both components increased their antitumor effect by over three times [19]. Oosporein, another bioactive component, was a polyketone compound with antibiotic, antiviral, antifungal, and insecticidal activities [20].

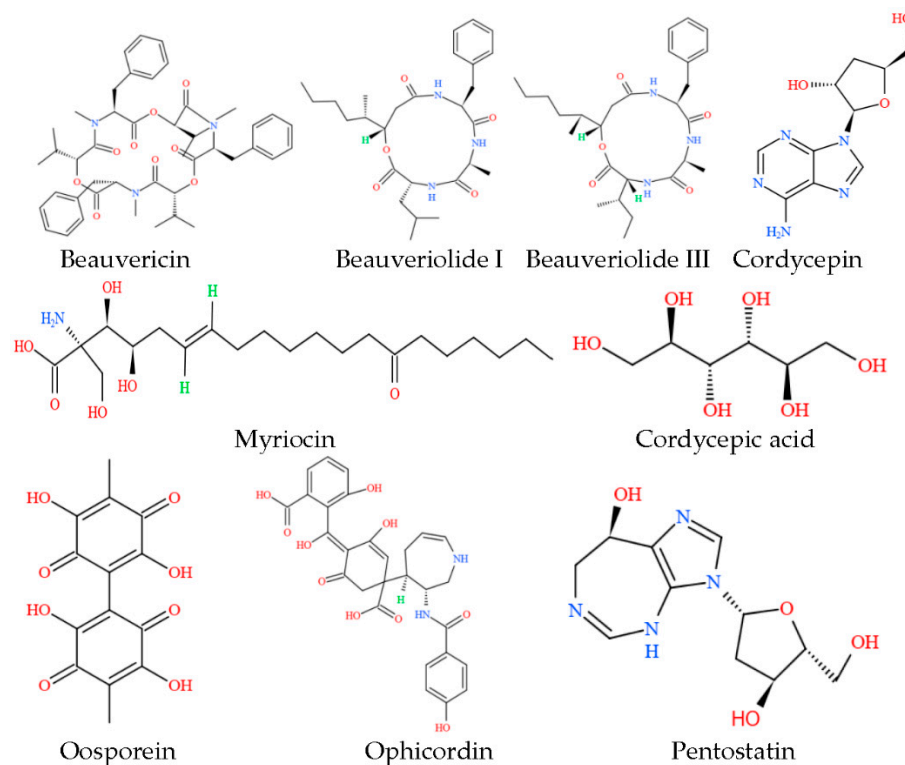


Figure 1. Structures of secondary metabolites from *Cordyceps*.

Genome mining is a computational method that automatically detects and annotates biosynthetic gene clusters (BGCs) from genomic data [21]. Today, the development of genome sequences, such as the identification and characterization of new compounds and metabolic engineering, have increased the range of applications of genome mining [22]. For instance, the antibiotics and Secondary Metabolites Analysis Shell (antiSMASH) web server was a stand-alone tool that was widely used for identifying and analyzing secondary metabolites (SMs) and BGCs in bacterial and fungal genomes [23]. The antiSMASH web server contributed significantly to microbial genome mining for new SM discoveries [24]. Wang et al. showed that the *C. militaris* *cns1-cns3* gene cluster produces cordycepin and pentostatin, where *cns1* and *cns2* were required for synthesizing cordycepin, and *cns3* was essential for pentostatin synthesis, elucidated by their biosynthetic pathways [25]. Moreover, nonribosomal peptide synthetase (NRPS; T-C-A-T-C-C-A-T-C-A-T-C-C), polyketide synthetase (PKS; KS-AT-DH-MT-ER-KR-ACP), acyltransferase, and ATP-dependent long chain fatty acyl-CoA synthetase in the genome of *C. militaris* catalyzes the synthesis of beauveriolide. The compound was generated by heterologous expression in *Aspergillus nidulans* (Eidam) G. Winter [26].

Whole-genome analysis of *C. militaris* CM01 revealed a cryptic gene cluster of encoding highly reducing polyketide synthase (HR-PKS), enol reductase (ER), and cytochrome P450. In *A. nidulans*, heterologous expression of the cryptic gene cluster produced two novel polyketide compounds, cordypyrone A and B [27]. From the bioinformatic analysis of the whole-genome sequence of *C. militaris* CM01's two strains, 1630 and DSM 1153, Wang et al. identified two NRPS genes, one PKS, and a hybrid cluster. They predicted the structural characteristics of their potential products [28]. In order to further explore the biosynthetic

potential of strain *C. pseudotenuipes* SM, its whole genome was sequenced and analyzed, while its whole genome data were compared with those of *C. tenuipes*, *C. militaris*, and *C. cicadae*, as well as for gene mining studies. The results showed that *C. pseudotenuipes*, *C. tenuipes*, *C. militaris*, and *C. cicadae* all contained different amounts of BGCs, including PKS, NRPS, terpene synthase (TPS), hybrid PKS + NRPS, and hybrid NRPS + Other.

2. Materials and Methods

2.1. Test Materials

C. tenuipes, *C. cicadae*, *C. militaris*, and *C. pseudotenuipes* strains were collected from Sapa, Vietnam; Dimalo village, Bangdang Township, Gongshan County, Yunnan Province; Zixi mountain, Zixi Town, Chuxiong City, Yunnan Province; and Yeyahu Forest Park, Kunming City, Yunnan Province. The voucher specimens were stored in Yunnan Herbal Herbarium (YHH) of Yunnan University, and the isolated strains were stored in Yunnan Fungal Culture Collection (YFCC) of Yunnan University. The whole-genome data of *C. tenuipes* were obtained from Yunnan Herbal Laboratory, College of Ecology and Environmental Sciences, Yunnan University. The genome-wide data of *C. cicadae* and *C. militaris* were obtained from NCBI (<https://www.ncbi.nlm.nih.gov/> (accessed on 3 March 2022)), accession numbers ASM296887v1 and GCA_000225605.1, respectively.

2.2. Strain Culture

C. pseudotenuipes was cultured on an MY solid medium (21 g/L Yeast Malt Broth, 15 g/L Agar Powder, 1 L H₂O) at 28 °C for 10 days. The mycelium was transferred to fresh MY liquid medium (21 g/L Yeast Malt Broth, 1 L H₂O) at 28 °C at 150 r/min. After 8 days of culture, mycelia were collected, and three biological replicates of each sample were stored in a refrigerator at −80 °C until they were transported on dry ice to Personalbio (Shanghai, China) for high-throughput sequencing.

2.3. Genome Sequencing and Assembly

We used the Illumina NovaSeq 2000 platform (Illumina, CA, USA) to construct the *C. pseudotenuipes* gene library with 400 bp insert fragments. Raw data were processed using FastQC, 3'-terminal DNA junction was decontaminated using AdapterRemoval (version 2) [29], and Soapec (V2.0) software was used to perform quality correction on all reads based on the KMER frequency that the KMER used for correction was set to 17 to obtain high-quality adaptor-free genome sequences. The data were assembled de novo to construct the contigs and scaffolds. The obtained contigs and scaffolds were corrected using pilon v1.18 [30] software.

2.4. Gene Prediction and Annotation

Gene prediction was performed using homology, SNAP, and Augustus software. The predicted genes were annotated through BLAST searches against existing databases, including NCBI non-redundant protein sequences (Nr), Swiss prot, Kyoto Encyclopedia of Genes and Genomes (KEGG), Evolutionary Genealogy of Genes: Non-supervised Orthologous Groups (EggNOg), Pathogen–Host Interactions Database (PHI), and Carbohydrate-Active Enzymes (CAZy).

2.5. Secondary Metabolite Biosynthesis Gene Cluster Analysis

Gene cluster predictions for *C. tenuipes*, *C. cicadae*, *C. militaris*, and *C. pseudotenuipes* scaffolds were performed using the antiSMASH (<https://antismash.secondarymetabolites.org/> (accessed on 27 May 2022)) online program. AntiSMASH detected scaffolds with gene clusters by using *C. militaris* as a parameter. The FGENESH (www.softberry.com/ (accessed on 28 May 2022)) online program predicted gene structures, and the PKS/NRPS online program (nrps.igs.umaryland.edu/ (accessed on 1 June 2022)) determined the gene clusters in contigs containing genes with NRPS/PKS structural domains. Meanwhile, the

BLAST analysis (<https://blast.ncbi.nlm.nih.gov/> (accessed on 8 June 2022)) was used for protein structure comparison to identify contigs harboring the NRPS/PKS gene.

2.6. Cluster Analysis

Known NRPS, PKS, and hybrid PKS-NRPS protein sequences were downloaded from NCBI and compared with protein sequences from this study using the Clustal W program of the MEGA5.0 software of IQ-TREE. The IQ-TREE web server is fast and accurately generates phylogenetic trees using the maximum likelihood method (<http://iqtree.cibiv.univie.ac.at/> (accessed on 25 June 2022)). The analysis involved 1000 bootstraps using default parameters for constructing the cluster tree.

2.7. Synteny Analysis

Scaffolds of *C. cicadae*, *C. militaris*, *C. pseudotenuipes*, and *C. tenuipes* genome containing BGC of SM were analyzed, combining MAUVE v2.4.0 according to the assembly order for collinearity analysis.

3. Results

3.1. Basic Features of the *C. pseudotenuipes* Genome

3.1.1. Genome Sequencing and Assembly

A total of 34,711,656 raw reads were obtained by Illumina sequencing (Illumina Ina, San Diego, CA, USA), yielding 34,001,494 HQ reads. The whole genome of *C. pseudotenuipes* was 30.1 Mb, containing 645 contigs and 527 scaffolds with an N50 of 131,856 bp and 54.11% GC content. The 8705 protein-coding genes with 13.97 Mb gene length were predicted, the average sequence length 1605.2 bp and the longest contig measured 0.41 Mb (Table 1). For non-coding RNA, we used tRNAscan, RNAmmer, and rfam_scan to predict 124 tRNA secondary structures, 31 rRNA, and 37 snRNA, respectively.

Table 1. Genomic assembly and functional annotation of *C. pseudotenuipes* genome.

Item	Value	Item	Count	Percentage (%)
Total length (Mb)	30.1	NR	8596	98.75
Max length (bp)	413,323	SwissProt	6206	71.29
GC content (%)	54.11	KEGG	3614	41.52
Gene number	8705	GO	5861	62.33
Total gene number (bp)	13,973,781	EggNOg	7877	90.49
Average gene number (bp)	1605.2	P450	8480	97.42
Gene/Genome (%)	47.3169	TCDB	1358	15.60
Contigs	645			
Scaffolds	527			
Contigs N50	101,518			
Scaffolds N50	131,856			
Contigs N90	27,054			
Scaffolds N90	32,698			

3.1.2. *C. pseudotenuipes* Genome Annotation

Similarly, analysis of the 8705 non-redundant *C. pseudotenuipes* genes on publicly available protein sequence databases yielded varying results. The NCBI Nr (8596 genes) had 98.75%, Swiss-Prot (6206) had 71.29%, KEGG 3614 had 41.52%, GO (3614 genes) had 41.52%, EggNOg (7877) had 90.49%, Cytochrome P450 (CYP) (8480) had 97.42%, and Transporter Classification Database (TCDB) (1358) had 15.60% similarity to the 8705 predicted proteins (Table 1). The EggNOg database showed that most of the predicted genes were functionally associated with “Function unknown” (3079), “Posttranslational modification, protein turnover, chaperones”, and “Carbohydrate transport and metabolism” in that order (Figure 2a). Nonetheless, the wide variety of posttranslational events and carbohydrate metabolism suggested that they improved regulatory protein bioactivity and energy

conversion efficiency. The KEGG functional classification revealed protein families for genetic information processing (2295), signaling and cellular processing (619), and signal transduction (562) (Figure 2b). Moreover, the rich diversity of genetic information and signaling proteins might facilitate more efficient information exchange and secondary metabolism. GO annotation indicated biological process (5316), cellular nitrogen compound metabolic process (1566), and biosynthetic process (1425) as the enriched terms from biological processes. Furthermore, the annotation revealed molecular function (4861), ion binding (2016), and oxidoreductase activity (767) from molecular function, and cell (2453), intracellular (2352), and cellular component (1972) from the cellular component (Figure 2c). *C. pseudotenuipes* was a wild strain; thus, many metabolic genes might be involved in signal transduction.

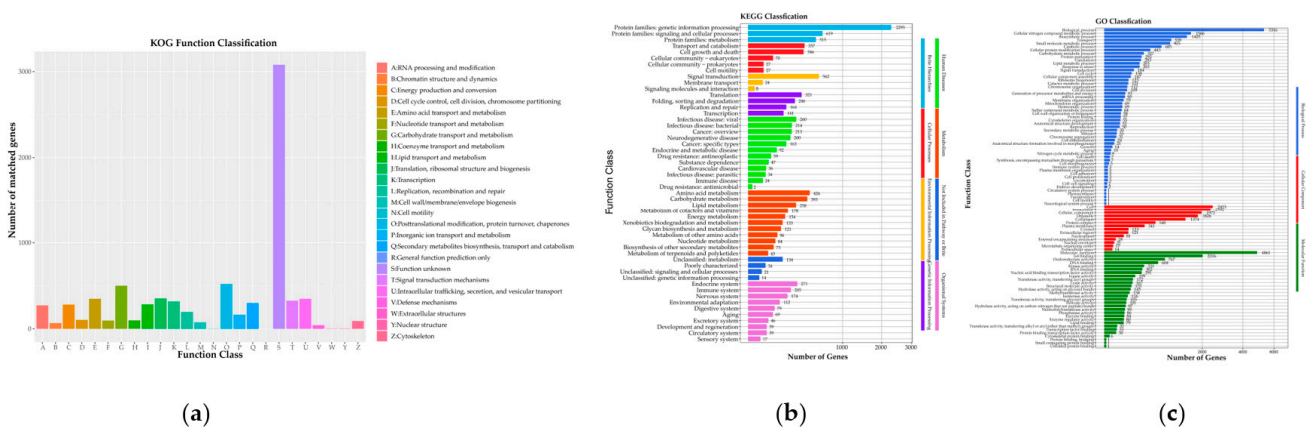


Figure 2. Functional annotation of *C. pseudotenuipes* gene-encoding proteins in the (a) Evolutionary Genealogy of Genes: Non-supervised Orthologous Groups of proteins (EggNOg), (b) Kyoto Encyclopedia of Genes and Genomes (KEGG), and (c) Gene Ontology (GO) databases.

3.1.3. *C. pseudotenuipes* Additional Annotation Pathogen Host Interactions (PHI)

Pathogen and host-interaction database (PHI-base), mainly from fungi-, oomycete-, and bacterial pathogen-infected hosts, included animals, plants, fungi, and insects [31]. Thus, the amino acid sequences of *C. pseudotenuipes* were aligned to obtain annotated results from the PHI-base (Figure 3). *C. pseudotenuipes* contained genes for reduced virulence (1108), unaffected pathogenicity (970), loss of pathogenicity (208), lethal (123), increased virulence (hypervirulence) (89), effector (plant avirulence determinant) (16), sensitivity to chemical (7), resistance to chemical (5), and enhanced antagonism (2). The results showed that the main annotated genes were for reduced virulence and unaffected pathogenicity, indicating that *C. pseudotenuipes* was a mildly pathogenic strain.

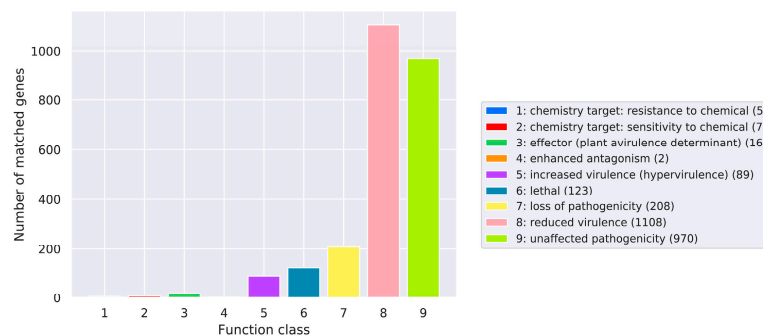


Figure 3. Distribution of the mutation types in the pathogen-PHI phenotype of *C. pseudotenuipes*.

Carbohydrate Genes

Analysis revealed 396 gene-encoding carbohydrate-active enzymes (CAZy) in *C. pseudotenuipes*, including 168 glycoside hydrolases (GHs) and 63 carbohydrate esterases (CEs), 56 auxiliary activities (AAs), 16 carbohydrate-binding modules (CBMs), and 3 polysaccharide lyases (PLs) (Figure 4). CAZy catalyzed the assembly and breakdown of glycans and glycoconjugates [32]. Moreover, CAZy was a database of carbohydrate-active and carbohydrate enzymes [33]. It was suggested that *C. pseudotenuipes* should possibly capture more energy and decompose complex carbohydrates.

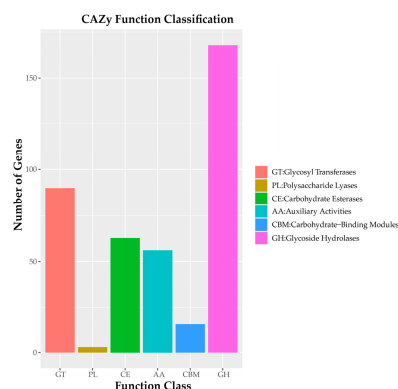


Figure 4. CAZy functional classification chart of *C. pseudotenuipes*.

3.2. Characteristics of *C. pseudotenuipes*, *C. tenuipes*, *C. cicadae*, and *C. militaris* Genomes

The genome of *C. pseudotenuipes* was characterized and compared with its three close taxa (*C. tenuipes*, *C. cicadae*, and *C. militaris*) (Table 2). *C. cicadae* was the largest (34.11 Mb), followed by *C. militaris* (32.27 Mb), *C. pseudotenuipes* (30.1 Mb), and *C. tenuipes* (30.06 Mb). Moreover, *C. cicadae* had the highest number of scaffolds (595), *C. pseudotenuipes* and *C. cicadae* contained an equal number of scaffolds (527), *C. tenuipes* contained 285 scaffolds, and *C. militaris* had the lowest scaffold number (32). The genome containing the highest number of contigs was *C. cicadae* (1799), followed by *C. pseudotenuipes* (645), *C. militaris* (597), and *C. tenuipes* (384). However, the GC contents of *C. pseudotenuipes*, *C. tenuipes*, *C. cicadae*, and *C. militaris* were 54.11%, 53.72%, 52.70%, and 51.40%, respectively.

Table 2. Characteristics of *C. pseudotenuipes*, *C. tenuipes*, *C. cicadae*, and *C. militaris* genomes.

Item	<i>C. pseudotenuipes</i> This Study	<i>C. tenuipes</i> This Study	<i>C. cicadae</i> ASM296887v1	<i>C. militaris</i> GCA_000225605.1
Contigs	645	384	1799	597
Scaffolds	527	285	595	32
Total length (Mb)	30.1	30.06	34.11	32.27
GC content (%)	54.11	53.72	52.70	51.40
Scaffold N50 (bp)	131,856	172,867	212,207	4,551,492
Contig N50 (bp)	101,518	140,681	47,316	10,818

3.3. Analysis of Secondary Metabolite Biosynthesis Gene Cluster

AntiSMASH and local BLAST analyses showed that *C. pseudotenuipes* and *C. cicadae* had the same putative SM BGCs (31), followed by *C. militaris* (29) and *C. tenuipes* (28) (Tables S1 and S2). The SM BGCs presumably contained the highest number of *C. pseudotenuipes* NRPSs (16), followed by *C. cicadae* (15), *C. tenuipes* (13), and *C. militaris* (9). *C. militaris* had six PKS + NRPS hybrids, *C. cicadae* had five, and both *C. pseudotenuipes* and *C. tenuipes* had four. *C. tenuipes* contained the majority of NRPS + Other hybrids (3), *C. cicadae* and *C. militaris* had two, and *C. pseudotenuipes* had one. *C. pseudotenuipes*, *C. cicadae*, *C. militaris*, and *C. tenuipes* contained decreasing amounts of encoding TPS. *C. pseudotenuipes* (1), *C. cicadae* (1), *C. militaris* (2), and *C. tenuipes* (2) had similar numbers of other biosynthetic

genes. *C. pseudotenuipes*, *C. tenuipes*, and *C. cicadae* had the same PKSs (5), while *C. militaris* had the highest number (8). The *C. pseudotenuipes* genome had five PKSs, including three HR-PKSs, two partially reducing (PR) PKSs, and one non-reducing (NR) PKS. The five PKSs obtained from the *C. tenuipes* genome included three NR-PKSs, one HR-PKS, and one PR-NRPS. The genome of *C. cicadae* had five PKSs, including two NR-PKSs, two HR-PKSs and one PR-NRPS. The eight PKSs retrieved from the *C. militaris* genome included five HR-PKSs, two PR-NRPSs, and one NR-PKS. These results showed that the number and type of SM BGCs obtained differed among species of the same genus.

The predicted BGCs showed different levels of genetic homology to known clusters in the MIBiG database, with *C. pseudotenuipes* having the highest homology (37.93%), followed by *C. tenuipes* (28.57%) and *C. cicadae* (25.81%). The BGCs predicted from *C. pseudotenuipes*, *C. tenuipes*, *C. cicadae*, and *C. militaris* genomes potentially synthesize dimethylcoprogen. In contrast, the BGCs predicted from *C. pseudotenuipes*, *C. tenuipes*, and *C. cicadae* predicted biosynthesis genes that catalyze the synthesis of leucinostatin A/leucinostatin B, neosartorin, dimethylcoprogen, wortmanamide A/wortmanamide B, and beauvericin. The antiSMASH and local BLAST analyses showed that only *C. tenuipes*, *C. cicadae*, and *C. militaris* potentially produce viriditoxin, while *C. tenuipes* and *C. cicadae* potentially produce trichodiene-11-one. Moreover, *C. pseudotenuipes*, *C. cicadae*, and *C. militaris* possibly catalyzed the synthesis of squalestatin S1; *C. pseudotenuipes* and *C. cicadae* also had BGCs for ilicicolin H synthesis, while *C. pseudotenuipes* and *C. tenuipes* were presumed responsible for epichloenin A synthesis. AntiSMASH analysis showed that ferrichrome was unique to *C. cicadae*, clavatic acid, communesin, and deoxynivalenol were unique to *C. pseudotenuipes*, and phomasetin, fumosorinone, and 1-nonadecene/(14z)-1,14-nonadecadiene were unique to *C. militaris*. Furthermore, similar species of the same genus had different types and quantities of BGCs and catalytically synthesized compounds.

Several BGCs of *C. pseudotenuipes*, *C. tenuipes*, and *C. cicadae* were 100% similar to MIBiG sequences. Dimethylcodogen was siderophores, produced by the *Alternaria* species to obtain extracellular iron [34,35]. The predicted region 358.1 of *C. militaris*, region 71.1 of *C. pseudotenuipes*, region 56.2 of *C. cicadae*, and region 18.2 of *C. tenuipes* might be responsible for dimethylcodogen biosynthesis (Figure 5a). Region 111.1 of *C. pseudotenuipes* and region 34.1 of *C. tenuipes* were responsible for epichloenin A biosynthesis (Figure 5b), and region 2.2 of *C. pseudotenuipes* was responsible for clavatic acid biosynthesis. In contrast, *C. cicadae* and *C. militaris* lacked the potential to synthesize epichloenin A. The enzymes involved in the synthesis of dimethylcoprogen and epichloenin A were NRPS, and the enzyme involved in the synthesis of clavatic acid was terpene. A local BLAST comparison demonstrated that dimethylcoprogen synthesis required enzymes with the A-P-C-P-C structural domain, acetyltransferase, PRK08315 superfamily, and MFS superfamily. Moreover, epichloenin A synthesis required NRPS (A-A-A-P-C), PTZ00265 superfamily, and MFS superfamily.

The five main biosynthetic genes in the putative SM BGCs of *C. pseudotenuipes*, *C. tenuipes*, *C. militaris*, and *C. cicadae* were similar to known gene clusters in the MIBiG database (ranging from 40% to 85%). Region 78.1 of *C. pseudotenuipes*, region 94.1 of *C. cicadae*, and region 426.1 of *C. militaris* are for squalestatin S1 synthesis. Region 38.2 of *C. tenuipes*, region 117.1 of *C. cicadae*, and region 95.1 of *C. militaris* probably synthesized viriditoxin, while region 1.2 of *C. cicadae* and region 346.1 and region 468.1 of *C. militaris* synthesized ferrichrome, phomasetin, and fumosorinone, respectively. The results showed that *C. pseudotenuipes* lacked the potential to synthesize viriditoxin and that catalytic synthesis of viriditoxin required several additional enzymes besides the core enzyme PKS (SAT-KS-AT-PT-ACP-ACP-ACP-TE). The additional enzymes included the Abhydrolase superfamily, MFS_Azrl_MDR_like, SDR, AdoMet_Mtases superfamily, and others of unknown function (Figure 6a). LC361337.1 synthesized phomasetin (Figure 6b). Furthermore, the core gene structural domain (KS-AT-MT-KR-ACP-C-A-P-TE) and modifier genes in the *C. militaris* region 346.1 were similar to LC361337.1. The location and orientation of the modifier genes differed, and the remaining three species lacked the potential to synthesize phomasetin. These results revealed that *C. militaris* region 468.1 was similar to the gene

cluster synthesizing fumosorinone (Figure 6c). Both regions had the KS-AT-MT-KR-ACP-C-A-P-TE core structural domain, the fungal-specific transcription factor domain-containing protein, cytochrome P450, and enoyl-reductase modifier genes, indicating that the positions and directions of functional enzymes differed between species.

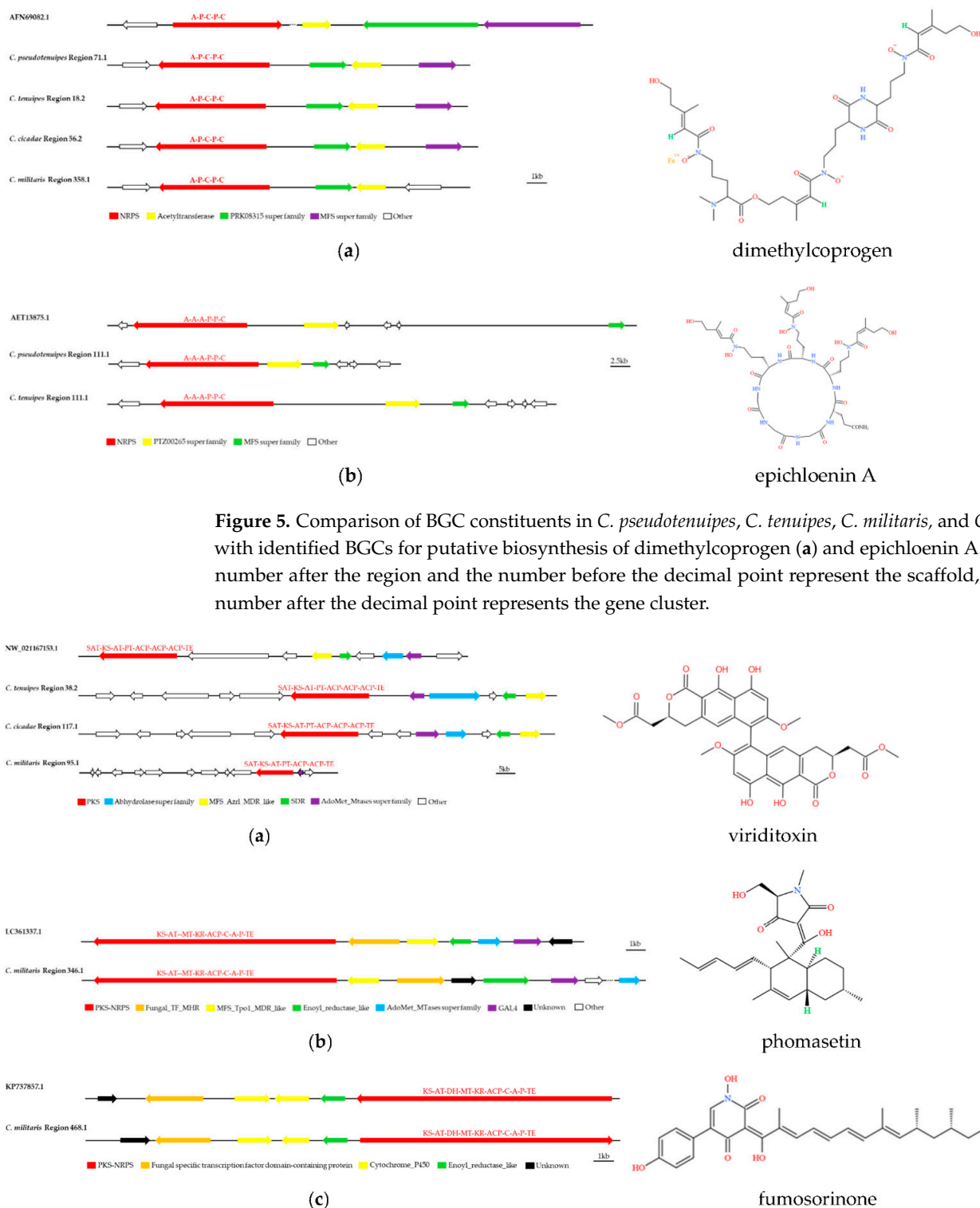


Figure 6. Comparison of biosynthesis of putative viriditoxin (a), phomasetin (b), and fumosorinone (c) biosynthetic gene clusters. The number after the region and the number before the decimal point represent the scaffold, and the number after the decimal point represents the gene cluster.

The antiSMASH and local BLAST analyses showed that the whole genomes of *C. tenuipes*, *C. cicadae*, and *C. pseudotenuipes* contained gene clusters that were highly similar to those responsible for the synthesis of Beauvericin (BEA) BGCs (GenBank: EU886196.1) (Figure 7). The putative NRPS from the *C. tenuipes* region 63.1, *C. cicadae* region 87.1, and *C. pseudotenuipes* region 44.1 had a C-A-P-C-A-MT-P-P-C structural domain, with precursors initiated and terminated by condensed structural domains. The NRPS also required several modifier genes, including the putative PRK06522, Far-17a_AIG1, and PRK08294 superfamilies, WD40, calreticulin, ALDH_F6_MMSDH, 17beta-HSDXI-like_SDR_c, Indigoidine_A, and GINS_A_Sld5, although *C. cicadae* region 87.1 lacked the Far-17a_AIG1 superfamily. The high similarity of the BEA BGC between *Cordyceps* and *Beauveria bassiana* suggested some degree of horizontal gene transfer between these genera and that gene loss/addition might occur between different species, even though the gene sequences varied in orientation and position.

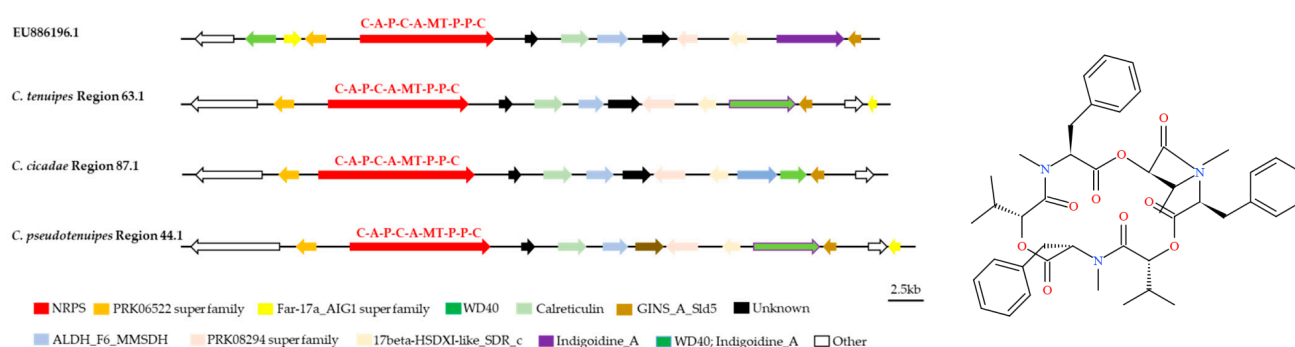


Figure 7. Comparison of putative beauvericin biosynthetic gene clusters and structure. The number after the region and the number before the decimal point represent the scaffold, and the number after the decimal point represents the gene cluster.

Furthermore, the gene clusters of *C. tenuipes* region 36.1, *C. cicadae* region 27.1, and *C. pseudotenuipes* region 112.1 resembled wortmanamide A or wortmanamide B BGC (Figure 8). Wortmanamide A or wortmanamide B was a reduced long-chain polyketide amidated by a specific ω -amino acid 5-aminopentanoic acid (5PA), initially found in *Talaromyces wortmannii*, a long-chain N-acylamide-like signaling lipid. A hybrid PKS-NRPS catalyzed synthesis with thioesterase and cytochrome P450 as the upstream modifier genes and enoyl_reductase and MFS as the downstream modifier gene synthesized wortmanamide A/B [36]. The hybrid PKS-NRPS structural domain of *C. tenuipes* region 36.1, *C. cicadae* region 27.1, and *C. pseudotenuipes* region 112.1 was KS-AT-DH-MT-KR-ACP-C-A-(P)-SDR. KS-AT-DH-MT-KR-ACP-C was the domain for catalytic wortmanamide A/B synthesis. The domains also had slightly different modifier genes. For instance, the modifier genes downstream of *C. tenuipes* region 36.1 was reductase, not enoyl_reductase, and thiolase, instead of thioesterase for *C. tenuipes* region 36.1, *C. cicadae* region 27.1, and *C. pseudotenuipes* region 112.1. Thus, *C. tenuipes* region 36.1, *C. cicadae* region 27.1, and *C. pseudotenuipes* region 112.1 might synthesize wortmanamide or its analogs.

3.4. Cluster Analysis

The clustering results of *Cordyceps* PKS and hybrid PKS-NRPS proteins with other fungal PKS and hybrid PKS-NRPS proteins showed that *C. militaris* region 346.1 clustered with *Pyrenochaetopsis* sp. (BBC43184.1), which catalyzed phomasetin biosynthesis (Figure S1). Moreover, *C. militaris* region 346.1 might catalyze the biosynthesis of phomasetin or its analogs. *C. militaris* region 95.1, *C. cicadae* region 117.1, and *C. tenuipes* region 38.2 clustered with the PKS protein and possibly catalyzes viriditoxin synthesis in *Paecilomyces variotii* (XP_0284818201). Moreover, *C. militaris* region 95.1, *C. cicadae* region 117.1, and *C. tenuipes* region 38.2 possibly catalyze viriditoxin or its analogs. *C. pseudotenuipes* region 112.1, *C. cicadae* region 27.1, *C. tenuipes* region 36.1, and *Talaromyces wortmannii* (QBC19710.1) were

clustered on an independent branch that produced wortmanamide A/B. *C. pseudotenuipes* region 112.1, *C. cicadae* region 27.1, and *C. tenuipes* region 36.1 presumably catalyze the synthesis of wortmanamide A/B or its analogs. *C. militaris* region 468.1 clustered with *C. fumosorosea* (AKC54422.1), which catalyzed fumosorinone biosynthesis. Furthermore, *C. militaris* region 468.1 might catalyze the biosynthesis of fumosorinone or its analogs. *C. pseudotenuipes* region 71.1, *C. cicadae* region 56.2, and *C. tenuipes* region 18.2 clustered with *Alternaria alternata* (AFN69082.1) catalyzed dimethylcoprogen synthesis (Figure S2) and probably produced dimethylcoprogen or its analogs.

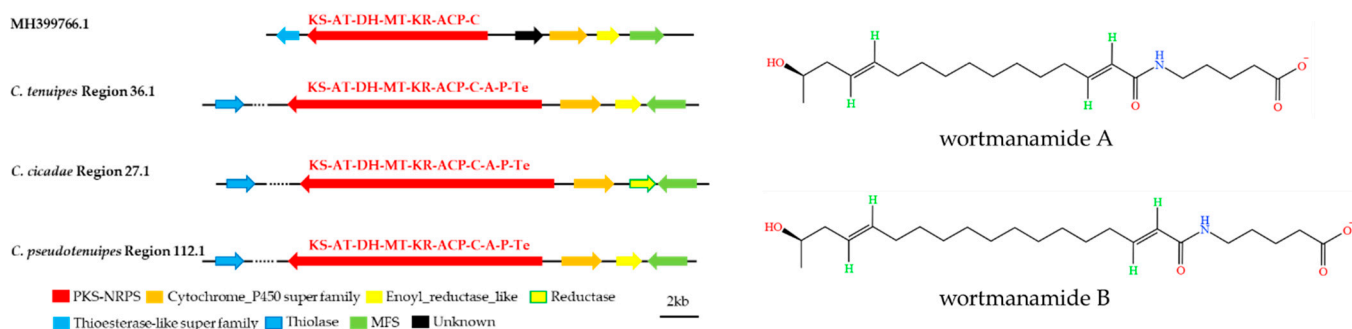


Figure 8. Comparison of putative wortmanamide A or wortmanamide B biosynthetic gene clusters and structure. The number after the region and the number before the decimal point represent the scaffold, and the number after the decimal point represents the gene cluster.

C. pseudotenuipes region 44.1, *C. cicadae* region 87.1, and *C. tenuipes* region 63.1 clustered on a separate branch from *Beauveria bassiana* (ACI30655.1), which catalyzed BEA biosynthesis, and the three regions presumably catalyze the biosynthesis of BEA or its analogs. Likewise, *C. pseudotenuipes* region 111.1 and *C. tenuipes* region 34.1 clustered on an independent branch with *Epichloe festucae* (AET13875.1), an *E. festucae* protein sequence that catalyzed epichloenin A biosynthesis. We hypothesized that *C. pseudotenuipes* region 111.1 and *C. tenuipes* region 34.1 might catalyze the synthesis of epichloenin A or its analogs.

3.5. Synteny Analysis

The scaffolds containing the SM BGC in the genomes of *C. cicadae* (24), *C. militaris* (26), *C. pseudotenuipes* (27), and *C. tenuipes* (28) were subjected to synteny analysis. The scaffolds where the SM BGC were located were divided into more than 70 collinear blocks, and there may be rearrangements (Figure 9). From top to bottom, they were *C. pseudotenuipes*, *C. tenuipes*, *C. cicadae*, and *C. militaris*.

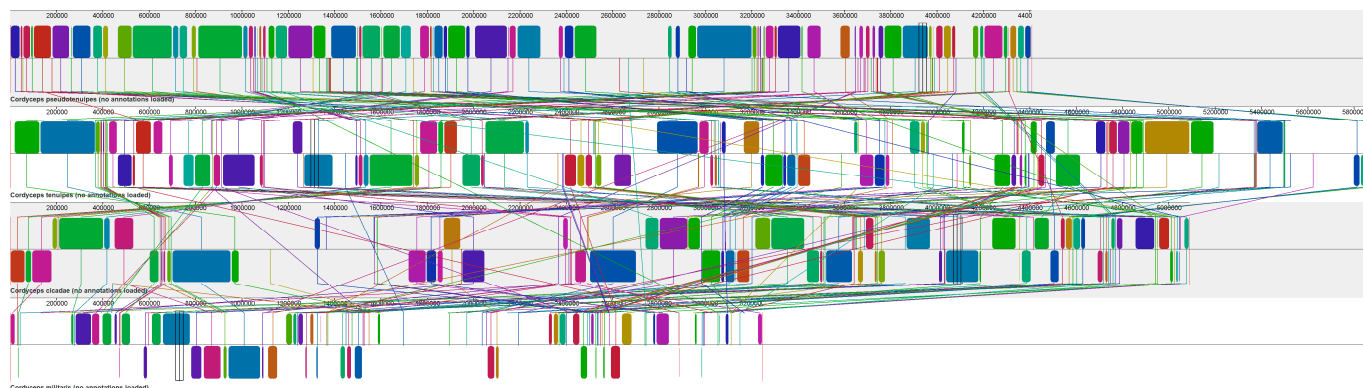


Figure 9. Scaffold containing secondary metabolite biosynthesis gene cluster used for synteny analysis.

4. Discussion

This study showed that *Cordyceps* contains numerous bioactive components, such as cordycepin, cordycepic acids, polysaccharide, and cyclic peptides [9], that have various clinical health effects, including anticancer, antioxidant, anti-inflammatory, immunomodulatory, and antibacterial activities [37]. The recent, significant improvement in next-generation sequencing increased the speed and reduced the cost of sequencing genomes of fungi [38]. Thus, this study presented the basic genomic characteristics of *C. pseudotenuipes* and its approximate taxa, *C. tenuipes*, *C. cicadae*, *C. militaris*, *C. pseudotenuipes*, *C. tenuipes*, *C. cicadae*, and *C. militaris*, revealing differences in genome size, scaffold N50, contig N50, and GC content, although all the species belong to the same genus.

Fungal SMs were generally classified based on their building blocks into polyketides (PKs), such as lovastatin, nonribosomal polypeptides (NRPs), such as penicillin, and terpenes, such as gibberellin [26]. Nonribosomal polypeptides and PKs were the major SMs groups with significantly diverse structures and biological functions [39,40]. New genome sequences of dozens of filamentous fungi suggested that their potential for SM generation far exceeds previous estimates. Moreover, the addition of genome mining has significantly improved bioassay-guided natural product discovery. Yuan et al. found likely existing NRPs and PKs through genome mining of the genome of *C. militaris* [41]. In this study, whole-genome mining of *C. pseudotenuipes*, *C. tenuipes*, *C. cicadae*, and *C. militaris* revealed putative SM BGCs. AntiSMASH and local BLAST analyses revealed 29 putative SM BGCs in the *C. militaris* genome, including 9 NRPSs, 8 PKSs, 6 hybrids PKS + NRPS, 2 hybrids NRPS + Other, 2 terpene, and 2 other gene clusters. Elsewhere, Wang et al. sequenced and predicted the genome of *C. militaris* CM01 (GenBank: GCA_000225605.1) using the antiSMASH online tool and 27 clusters of SM biosynthesis genes, including 7 NRPS, 7 hybrid PKS-NRPS, 5 PKS, 4 terpenoid cyclase (TC), and 4 others [42]. The contrast between this study and Wang et al. was presumably due to different analysis methods. Wang et al. identified a gene cluster in the *C. militaris* genome that catalyzed beauveriolide synthesis using NRPS, PKS, acyltransferase, and ATP-dependent long-chain fatty acyl-coenzymes A synthase [26]. However, the antiSMASH and local BLAST analyses revealed that the modified genes in *C. militaris* regions 67.1 and 67.2 and between 67.1 and 67.2 were consistent with the gene cluster catalyzing the synthesis of the identified beauveriolide. *C. militaris* region 536.1 and its four downstream modifier genes were consistent with the gene clusters that catalyzed the synthesis of two novel polyketide compounds, cordypyrone A and B. [27]. Nonetheless, the wortmanamide BGC differed, suggesting that *C. tenuipes* region 36.1, *C. cicadae* region 27.1, and *C. pseudotenuipes* region 112.1 might catalyze the synthesis of wortmanamide or its analogs [36]. This study suggested that *C. cicadae* potentially catalyzed the synthesis of BEA or its analogs, confirming previous studies where BEA was isolable from *C. cicadae* [13]. Thus, the next step was to verify through heterologous expression or gene knockout whether *C. cicadae* region 87.1 and its modifier genes were BGCs that catalyze BEA biosynthesis.

Moreover, NRPS and PKS also appeared to hybridize PKS-NRPS or NRPS-PKS when catalyzing the synthesis of compounds. Chimeric genes produced chimeric compounds and biologically active hybrid products such as cyclopiazonic acid, pyranonigrin, and cytochalasin [42–45]. Seshime et al. functionally expressed the hybrid PKS-NRPS CpaA involved in cyclopiazonic acid biosynthesis in *A. flavus* [46]. Moreover, Nielsen et al. successfully constructed hybrid PKS-NRPS expressing functional cross-species in *A. nidulans*, indicating that a rational redesign of fungal natural product enzymes was feasible [47]. Theobald et al. also showed the origin of hybrid PKS-NRPS in *Aspergilli* [44], and Zhong et al. identified a hybrid PKS-NRPS gene cluster from a clonal population of the rice blast fungus *Magnaporthe oryzae* [48]. Tang et al. also identified a PKS-NRPS hybrid enzyme and ten pyranterreones products in *A. terreus* through genome mining [49]. Sigrist et al. studied cis/trans hybrid PKS-NRPS of nonlinear biosynthetic assembly Alpiniamided, suggesting that it might take place on two parallel assembly lines, the products of which were then linked together [50]. In that order, *C. pseudotenuipes*, *C. tenuipes*, *C. cicadae*, and

C. militaris genomes analyzed in this study showed that the type and number of putative BGCs differ among species of the same genus. *C. pseudotenuipes*, *C. tenuipes*, *C. cicadae*, and *C. militaris* contained 4, 4, 5, and 6 hybrid PKS-NRPS, respectively. Thus, gene chimerism probably occurred in the PKS and NRPS of genus *Cordyceps* and catalyzed the synthesis of chimeric compounds. The antiSMASH and local BLAST results showed that horizontal gene transfer possibly occurs among these species of the same genus, explaining the variable direction and position of gene sequences. The next step is to validate the SM BGCs of *C. pseudotenuipes*, *C. tenuipes*, *C. cicadae*, and *C. militaris* through heterologous expression or gene knockdown.

The synteny analysis of scaffolds containing SM BGC in *C. pseudotenuipes*, *C. tenuipes*, *C. cicadae*, and *C. militaris* genome showed that the collinearity of *C. militaris* sequences was significantly different from that of *C. pseudotenuipes*, *C. tenuipes*, and *C. cicadae*. It was presumed that the classification of *C. militaris* was far from that of *C. pseudotenuipes*, *C. tenuipes*, and *C. cicadae*.

5. Conclusions

In order to deeply explore the SMs BGCs of the genus *Cordyceps*, the whole genome of *C. pseudotenuipes* was sequenced, annotated, and compared with three related species to characterize the genome. The high-quality whole-genome sequence of *C. pseudotenuipes* was obtained and extensively analyzed by gene prediction and annotation in this work. The results demonstrated that *C. pseudotenuipes* harbored abundant functional genes in regulatory protein bioactivity and conversion, signal transduction and metabolism, carbohydrate transport and metabolism, capturing energy, and decomposes carbohydrates.

The genome-wide basic characteristics of *C. pseudotenuipes*, *C. tenuipes*, *C. cicadae*, and *C. militaris* revealed unequal genome sizes and GC contents. AntiSMASH and local BLAST analyses showed that *C. pseudotenuipes*, *C. tenuipes*, *C. cicadae*, and *C. militaris* had different amounts and types of putative SM BGC. Presumably, only seven of the putative BGCs were highly similar to known gene clusters, indicating a great potential to generate other SMs. These findings opened the possibility of targeted genomic mining, such as gene knockdown, introduction or heterologous expression of microbial genes, promoter regulation, and mutation induction, to awaken the silenced BGC biosynthesis of more novel bioactive SMs for new drug research and development.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/metabo12090844/s1>, Table S1: Putative biosynthetic gene clusters (BGCs) coding for secondary metabolites in three *Cordyceps* species. Table S2: Overview of biosynthetic gene clusters in the genomes of the four studied fungi. Figure S1: Clustering of the PKS/hybrid PKS-NRPS and other related fungal PKS/hybrid PKS-NRPS proteins in the four *Cordyceps* species. Figure S2: Clustering of NRPS and other fungal NRPS proteins in the four species of *Cordyceps*.

Author Contributions: Conceptualization, project administration, and funding acquisition, H.Y. and Y.W.; methodology, Y.W., Y.L. and D.L.; software, Y.W. and X.Y.; formal analysis, Y.L.; resources O.H.; investigation, Q.D., S.D. and F.M.; writing—original draft preparation, Y.L.; writing—review and editing, Y.L., Y.W. and H.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the National Natural Science Foundation of China under award numbers 31870017 (to H.Y.), the National Natural Science Foundation of China under award numbers 31860177 (Y.W.), General Project of Basic Research Program in Yunnan Province under award numbers 202101AT070218 (Y.W.), and the Reserve Talents for Young and Middle-aged Academic and Technical Leaders of the Yunnan Province under award numbers 202205AC160044 (Y.W.).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available in supplementary material. *C. pseudotenuipes* and *C. tenuipes* genome data are available from NCBI.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Dong, Q.-Y.; Wang, Y.; Wang, Z.-Q.; Tang, D.-X.; Zhao, Z.-Y.; Wu, H.-J.; Yu, H. Morphology and Phylogeny Reveal Five Novel Species in the Genus *Cordyceps* (*Cordycipitaceae*, *Hypocreales*) From Yunnan, China. *Front. Microbiol.* **2022**, *13*, 846909. [\[CrossRef\]](#)
2. Wang, Y.B.; Wang, Y.; Fan, Q.; Duan, D.E.; Zhang, G.D.; Dai, R.Q.; Dai, Y.D.; Zeng, W.B.; Chen, Z.H.; Li, D.D.; et al. Multigene phylogeny of the family *Cordycipitaceae* (*Hypocreales*): New taxa and the new systematic position of the Chinese cordycipitoid fungus *Paecilomyces hepiali*. *Fungal Divers.* **2020**, *103*, 1–46. [\[CrossRef\]](#)
3. Kepler, R.M.; Luangsa-Ard, J.J.; Hywel-Jones, N.L.; Quandt, C.A.; Sung, G.-H.; Rehner, S.A.; Aime, M.C.; Henkel, T.W.; Sanjuan, T.; Zare, R.; et al. A phylogenetically-based nomenclature for *Cordycipitaceae* (*Hypocreales*). *IMA Fungus* **2017**, *8*, 335–353. [\[CrossRef\]](#)
4. Lu, M.Y.; Chen, C.C.; Lee, L.Y.; Lin, T.W.; Kuo, C.F. N (6)-(2-Hydroxyethyl) adenosine in the Medicinal Mushroom *Cordyceps cicadae* Attenuates Lipopolysaccharide-Stimulated Pro-inflammatory Responses by Suppressing TLR4-Mediated NF- κ B Signaling Pathways. *J. Nat. Prod.* **2015**, *78*, 2452–2460. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Ying, M.; Yu, Q.; Zheng, B.; Wang, H.; Wang, J.; Chen, S.; Nie, S.; Xie, M. Cultured *Cordyceps sinensis* polysaccharides modulate intestinal mucosal immunity and gut microbiota in cyclophosphamide-treated mice. *Carbohydr. Polym.* **2020**, *235*, 115957. [\[CrossRef\]](#)
6. Zhang, J.; Jian, T.; Zhang, Y.; Zhang, G.; Ling, J. Dynamic content changes of cordycepin and adenosine and transcriptome in *Cordyceps kyushuensis* Kob at different fermentation stages. *Bioprocess Biosyst. Eng.* **2021**, *44*, 1793–1803. [\[CrossRef\]](#)
7. Yin, Y.; Yu, G.; Chen, Y.; Jiang, S.; Wang, M.; Jin, Y.; Lan, X.; Liang, Y.; Sun, H. Genome-Wide Transcriptome and Proteome Analysis on Different Developmental Stages of *Cordyceps militaris*. *PLoS ONE* **2012**, *7*, e51853. [\[CrossRef\]](#)
8. Zhao, X.; Zhang, G.; Li, C.; Ling, J. Cordycepin and pentostatin biosynthesis gene identified through transcriptome and proteomics analysis of *Cordyceps kyushuensis* Kob. *Microbiol. Res.* **2018**, *218*, 12–21. [\[CrossRef\]](#)
9. Yang, N.-N.; Jiang, N.; Ma, Q.-Y.; Kong, F.-D.; Xie, Q.-Y.; Zhou, L.-M.; Yu, Z.-F.; Zhao, Y.-X. Chemical study of the strain *Cordyceps* spp. from cell fusion between *Cordyceps militaris* and *Cordyceps cicadae*. *J. Asian Nat. Prod. Res.* **2018**, *21*, 449–455. [\[CrossRef\]](#)
10. Zhao, C.; Bu, H.; Zhu, J.; Wang, Y.; Oliver, K.M.; Hu, F.; Huang, B.; Li, Z.; Peng, F. Integration of Untargeted Metabolomics with Transcriptomics Provides Insights into Beauvericin Biosynthesis in *Cordyceps chanhua* under H₂O₂-Induced Oxidative Stress. *J. Fungi* **2022**, *8*, 484. [\[CrossRef\]](#) [\[PubMed\]](#)
11. Li, Y.; Talbot, C.L.; Chandravanshi, B.; Ksiazek, A.; Sood, A.; Chowdhury, K.H.; Maschek, J.A.; Cox, J.; Babu, A.K.S.; Paz, H.A.; et al. *Cordyceps* inhibits ceramide biosynthesis and improves insulin resistance and hepatic steatosis. *Sci. Rep.* **2022**, *12*, 7273. [\[CrossRef\]](#)
12. Yin, Y.; Chen, B.; Song, S.; Li, B.; Yang, X.; Wang, C. Production of Diverse Beauveriolide Analogs in Closely Related Fungi: A Rare Case of Fungal Chemodiversity. *mSphere* **2020**, *5*, e00667-20. [\[CrossRef\]](#)
13. Zhang, X.; Hu, Q.; Weng, Q. Secondary metabolites (SMs) of *Isaria cicadae* and *Isaria tenuipes*. *RSC Adv.* **2018**, *9*, 172–184. [\[CrossRef\]](#)
14. Tuli, H.S.; Sharma, A.K.; Sandhu, S.S.; Kashyap, D. Cordycepin: A bioactive metabolite with therapeutic potential. *Life Sci.* **2013**, *93*, 863–869. [\[CrossRef\]](#)
15. Holbein, S.; Wengi, A.; Decourty, L.; Freimoser, F.M.; Jacquier, A.; Dichtl, B. Cordycepin interferes with 3' end formation in yeast independently of its potential to terminate RNA chain elongation. *RNA* **2009**, *15*, 837–849. [\[CrossRef\]](#)
16. Wen, T.; Jiang, C.; Shen, D.; Li, C.; Lu, Q.; Liu, X.; Kang, Y.; Chen, X.; Liu, B.; Qin, C.; et al. Evaluation of the Antitumor Activity by Cordycepin with Gold Nanoparticle. *J. Nanosci. Nanotechnol.* **2016**, *16*, 7134–7139. [\[CrossRef\]](#)
17. Dalla Rosa, L.; da Silva, A.S.; Gressler, L.T.; Oliveira, C.B.; Dambrós, M.G.; Miletti, L.C.; França, R.T.; Lopes, S.T.; Samara, Y.N.; da Veiga, M.L.; et al. Cordycepin (3'-deoxyadenosine) pentostatin (deoxycoformycin) combination treatment of mice experimentally infected with *Trypanosoma evansi*. *Parasitology* **2013**, *140*, 663–671. [\[CrossRef\]](#)
18. Sarvaria, A.; Topp, Z.; Saven, A. Current Therapy and New Directions in the Treatment of Hairy Cell Leukemia: A Review. *JAMA Oncol.* **2016**, *2*, 123–129. [\[CrossRef\]](#)
19. Yoshikawa, N.; Nakamura, K.; Yamaguchi, Y.; Kagota, S.; Shinozuka, K.; Kunitomo, M. Reinforcement of antitumor effect of *Cordyceps sinensis* by 2'-deoxycoformycin, an adenosine deaminase inhibitor. *In Vivo* **2007**, *21*, 291–295.
20. Feng, P.; Shang, Y.; Cen, K.; Wang, C. Fungal biosynthesis of the bibenzoquinone oosporein to evade insect immunity. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 11365–11370. [\[CrossRef\]](#)
21. Russell, A.H.; Truman, A.W. Genome mining strategies for ribosomally synthesised and post-translationally modified peptides. *Comput. Struct. Biotechnol. J.* **2020**, *18*, 1838–1851. [\[CrossRef\]](#)
22. Chu, L.; Huang, J.; Muhammad, M.; Deng, Z.; Gao, J. Genome mining as a biotechnological tool for the discovery of novel marine natural products. *Crit. Rev. Biotechnol.* **2020**, *40*, 571–589. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Blin, K.; Shaw, S.; Steinke, K.; Villebro, R.; Ziemert, N.; Lee, S.Y.; Medema, M.H.; Weber, T. antiSMASH 5.0: Updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res.* **2019**, *47*, W81–W87. [\[CrossRef\]](#)
24. Blin, K.; Kim, H.U.; Medema, M.H.; Weber, T. Recent development of antiSMASH and other computational approaches to mine secondary metabolite biosynthetic gene clusters. *Brief. Bioinform.* **2017**, *20*, 1103–1113. [\[CrossRef\]](#) [\[PubMed\]](#)

25. Xia, Y.; Luo, F.; Shang, Y.; Chen, P.; Lu, Y.; Wang, C. Fungal Cordycepin Biosynthesis Is Coupled with the Production of the Safeguard Molecule Pentostatin. *Cell Chem. Biol.* **2017**, *24*, 1479–1489.e4. [[CrossRef](#)]
26. Wang, X.; Gao, Y.L.; Zhang, M.L.; Zhang, H.D.; Huang, J.Z.; Li, L. Genome mining and biosynthesis of the Acyl-CoA: Cho-lesterol acyltransferase inhibitor beauveriolide I and III in *Cordyceps militaris*. *J. Biotechnol.* **2020**, *309*, 85–91. [[CrossRef](#)]
27. Gao, Y.-L.; Yu, C.; Li, L. Heterologous expression of a natural product biosynthetic gene cluster from *Cordyceps militaris*. *J. Antibiot.* **2021**, *75*, 16–20. [[CrossRef](#)]
28. Wang, W.-J.; Vogel, H.; Yao, Y.-J.; Ping, L. The nonribosomal peptide and polyketide synthetic gene clusters in two strains of entomopathogenic fungi in *Cordyceps*. *FEMS Microbiol. Lett.* **2012**, *336*, 89–97. [[CrossRef](#)]
29. Schubert, M.; Lindgreen, S.; Orlando, L. AdapterRemoval v2: Rapid adapter trimming, identification, and read merging. *BMC Res. Notes* **2016**, *9*, 88. [[CrossRef](#)]
30. Walker, B.J.; Abeel, T.; Shea, T.; Priest, M.; Abouelliel, A.; Sakthikumar, S.; Cuomo, C.A.; Zeng, Q.; Wortman, J.; Young, S.K.; et al. Pilon: An Integrated Tool for Comprehensive Microbial Variant Detection and Genome Assembly Improvement. *PLoS ONE* **2014**, *9*, e112963. [[CrossRef](#)]
31. Urban, M.; Cuzick, A.; Seager, J.; Wood, V.; Rutherford, K.; Venkatesh, S.Y.; Sahu, J.; Iyer, S.V.; Khamari, L.; De Silva, N.; et al. PHI-base in 2022: A multi-species phenotype database for Pathogen–Host Interactions. *Nucleic Acids Res.* **2021**, *50*, D837–D847. [[CrossRef](#)] [[PubMed](#)]
32. Garron, M.-L.; Henrissat, B. The continuing expansion of CAZymes and their families. *Curr. Opin. Chem. Biol.* **2019**, *53*, 82–87. [[CrossRef](#)] [[PubMed](#)]
33. Drula, E.; Garron, M.L.; Dogan, S.; Lombard, V.; Henrissat, B.; Terrapon, N. The carbohydrate-active enzyme database: Functions and literature. *Nucleic Acids Res.* **2022**, *50*, D571–D577. [[CrossRef](#)] [[PubMed](#)]
34. Chen, L.-H.; Lin, C.-H.; Chung, K.-R. A nonribosomal peptide synthetase mediates siderophore production and virulence in the citrus fungal pathogen *Alternaria alternata*. *Mol. Plant Pathol.* **2013**, *14*, 497–505. [[CrossRef](#)]
35. Jalal, M.A.F.; Love, S.K.; van der Helm, D. N alpha-dimethylcoprogens. Three novel trihydroxamate siderophores from pathogenic fungi. *Biol. Met.* **1988**, *1*, 4–8. [[CrossRef](#)]
36. Hai, Y.; Tang, Y. Biosynthesis of Long-Chain N-Acyl Amide by a Truncated Polyketide Synthase–Nonribosomal Peptide Synthetase Hybrid Megasyntase in Fungi. *J. Am. Chem. Soc.* **2018**, *140*, 1271–1274. [[CrossRef](#)]
37. Ahn, Y.J.; Park, S.J.; Lee, S.G.; Shin, S.C.; Choi, D.H. Cordycepin: Selective growth inhibitor derived from liquid culture of *Cordyceps militaris* against *Clostridium* spp. *J. Agric. Food Chem.* **2000**, *48*, 2744–2748. [[CrossRef](#)]
38. Cacho, R.A.; Tang, Y.; Chooi, Y.-H. Next-generation sequencing approach for connecting secondary metabolites to biosynthetic gene clusters in fungi. *Front. Microbiol.* **2015**, *5*, 774. [[CrossRef](#)]
39. Du, L.; Lou, L. PKS and NRPS release mechanisms. *Nat. Prod. Rep.* **2009**, *27*, 255–278. [[CrossRef](#)]
40. Parsley, L.C.; Linneman, J.; Goode, A.M.; Becklund, K.; George, I.; Goodman, R.M.; Lopanik, N.B.; Liles, M.R. Polyketide synthase pathways identified from a metagenomic library are derived from soil Acidobacteria. *FEMS Microbiol. Ecol.* **2011**, *78*, 176–187. [[CrossRef](#)]
41. Yuan, X.L.; Li, Y.Q.; Yi, W. The Diversity of Polyketide Synthase (PKS) Gene in *Cordyceps militaris*. *J. West China For. Sci.* **2019**, *48*, 97–103+113.
42. Zheng, P.; Xia, Y.; Xiao, G.; Xiong, C.; Hu, X.; Zhang, S.; Zheng, H.; Huang, Y.; Zhou, Y.; Wang, S.; et al. Genome sequence of the insect pathogenic fungus *Cordyceps militaris*, a valued traditional chinese medicine. *Genome Biol.* **2011**, *12*, R116. [[CrossRef](#)] [[PubMed](#)]
43. Awakawa, T.; Yang, X.-L.; Wakimoto, T.; Abe, I. Pyranonigrin E: A PKS-NRPS Hybrid Metabolite from *Aspergillus niger* Identified by Genome Mining. *ChemBioChem* **2013**, *14*, 2095–2099. [[CrossRef](#)]
44. Theobald, S.; Vesth, T.C.; Andersen, M.R. Genus level analysis of PKS-NRPS and NRPS-PKS hybrids reveals their origin in Aspergilli. *BMC Genom.* **2019**, *20*, 847. [[CrossRef](#)] [[PubMed](#)]
45. Wang, Y.; Zhao, S.; Guo, T.; Li, L.; Li, T.; Wang, A.; Zhang, D.; Wang, Y.; Sun, Y. Bioactive PKS–NRPS Alkaloids from the Plant-Derived Endophytic Fungus *Xylaria arbuscula*. *Molecules* **2021**, *27*, 136. [[CrossRef](#)]
46. Seshime, Y.; Juvvadi, P.R.; Tokuoka, M.; Koyama, Y.; Kitamoto, K.; Ebizuka, Y.; Fujii, I. Functional expression of the *Aspergillus flavus* PKS–NRPS hybrid CpaA involved in the biosynthesis of cyclopiazonic acid. *Bioorganic Med. Chem. Lett.* **2009**, *19*, 3288–3292. [[CrossRef](#)]
47. Nielsen, M.L.; Isbrandt, T.; Petersen, L.M.; Mortensen, U.H.; Andersen, M.R.; Hoof, J.B.; Larsen, T.O. Linker Flexibility Facilitates Module Exchange in Fungal Hybrid PKS-NRPS Engineering. *PLoS ONE* **2016**, *11*, e0161199. [[CrossRef](#)]
48. Zhong, Z.; Lin, L.; Zheng, H.; Bao, J.; Chen, M.; Zhang, L.; Tang, W.; Ebbolle, D.J.; Wang, Z. Emergence of a hybrid PKS-NRPS secondary metabolite cluster in a clonal population of the rice blast fungus *Magnaporthe oryzae*. *Environ. Microbiol.* **2020**, *22*, 2709–2723. [[CrossRef](#)]
49. Tang, S.; Zhang, W.; Li, Z.; Li, H.; Geng, C.; Huang, X.; Lu, X. Discovery and Characterization of a PKS-NRPS Hybrid in *Aspergillus terreus* by Genome Mining. *J. Nat. Prod.* **2020**, *83*, 473–480. [[CrossRef](#)]
50. Sigrist, R.; Luhavaya, H.; McKinnie, S.; Ferreira da Silva, A.; Jurberg, I.D.; Moore, B.S.; Gonzaga de Oliveira, L. Nonlinear Biosynthetic Assembly of Alpinamide by a Hybrid cis/trans-AT PKS-NRPS. *ACS Chem. Biol.* **2020**, *15*, 1067–1077. [[CrossRef](#)]