

Article

# Genome Analysis of *Hypomyces perniciosus*, the Causal Agent of Wet Bubble Disease of Button Mushroom (*Agaricus bisporus*)

Dan Li <sup>1,†</sup>, Frederick Leo Sossah <sup>1,†</sup> , Lei Sun <sup>1</sup>, Yongping Fu <sup>1,\*</sup> and Yu Li <sup>1,2,\*</sup>

<sup>1</sup> Engineering Research Center of Chinese Ministry of Education for Edible and Medicinal Fungi, Jilin Agricultural University, Changchun 130118, China; junwuzhongxin@126.com (D.L.); flsossah@gmail.com (F.L.S.); sunlei@jlau.edu.cn (L.S.)

<sup>2</sup> International Cooperation Research Center of China for New Germplasm and Breeding of Edible Mushrooms, Jilin Agricultural University, Changchun 130118, China

\* Correspondence: yongping.fu@jlau.edu.cn (Y.F.); yuli966@126.com (Y.L.); Tel.: +86-431-8453-2989 (Y.L.)

† These authors contributed equally to this work.

Received: 4 April 2019; Accepted: 27 May 2019; Published: 29 May 2019



**Abstract:** The mycoparasitic fungus *Hypomyces perniciosus* causes wet bubble disease of mushrooms, particularly *Agaricus bisporus*. The genome of a highly virulent strain of *H. perniciosus* HP10 was sequenced and compared to three other fungi from the order Hypocreales that cause disease on *A. bisporus*. *H. perniciosus* genome is ~44 Mb, encodes 10,077 genes and enriched with transposable elements up to 25.3%. Phylogenetic analysis revealed that *H. perniciosus* is closely related to *Cladobotryum protrusum* and diverged from their common ancestor ~156.7 million years ago. *H. perniciosus* has few secreted proteins compared to *C. protrusum* and *Trichoderma virens*, but significantly expanded protein families of transporters, protein kinases, CAZymes (GH 18), peptidases, cytochrome P450, and SMs that are essential for mycoparasitism and adaptation to harsh environments. This study provides insights into *H. perniciosus* evolution and pathogenesis and will contribute to the development of effective disease management strategies to control wet bubble disease.

**Keywords:** *Hypomyces perniciosus*; genome sequence; *Agaricus bisporus*; wet bubble disease; CAZymes; cytochrome P450 enzymes; secondary metabolites

## 1. Introduction

*Hypomyces perniciosus* Magnus (syn: *Mycogone perniciosa* (Magn.) Delacr.) is an ascomycetes fungus in the family Hypocreaceae. It causes wet bubble disease (WBD) on a broad range of basidiomycetes hosts. It was first identified as a pathogen of *Agaricus bisporus* more than 100 years ago [1,2]. WBD is one of the major constraints in the commercial production of *A. bisporus*. Typical disease symptoms of WBD on *A. bisporus* include malformation of basidiome, white, fluffy mycelial growth, copious amber droplets (diseased carpophores exuding a brown malodorous liquid), and flocculent mycelia on most substrates. The infected mushrooms are unmarketable, resulting in significant yield and economic losses to the mushroom industry globally [1,3]. WBD is widespread and occurs in most countries where there is commercial production of *A. bisporus* [1,4]. The genus *Mycogone* also contains species, such as *Mycogone* sp., *M. calospora*, and *M. rosea*, which cause less severe forms of WBD compared to *H. perniciosus* [1]. Previous studies have reported variation in colony morphology and physiology [1] as well as high genetic diversity among *H. perniciosus* isolates [1,5,6].

Despite the economic importance and the considerable losses caused by the pathogen *H. perniciosus*, little is known about the genetic and pathogenic mechanisms underlying the interaction between *H. perniciosus* and the fungal host. Previous research has focused on the disease occurrence, genetic

diversity, pathogenesis, pathogenicity, identification of disease resistance, and integrated control and management of the disease [3,4,7–10]. Genomic analysis of pathogens is also one of the most effective ways to obtain a full understanding of fungal pathogenesis at the molecular level [10]. Currently, many plant pathogen genomes are available, and the putative genes involved in pathogenicity have been investigated. However, the availability of mushroom pathogen genomes is still very scarce. Recently, a few genomes of mycoparasites were released, such as *Cladobotryum protrusum* [11] and *Lecanicillium fungicola* [12,13]. The genome analysis is imperative and will facilitate more rapid identification of genes associated with pathogenicity and pathogen-mushroom interactions.

In this study, we present the de novo assembled genome of *H. perniciosus*, representing the first sequenced genome in the genus *Hypomyces* using the single-molecule real-time sequencing platform (SMRT) of Pacific Biosciences (PacBio). Our specific objectives were the following: (1) to conduct high-quality genome sequencing of a highly virulent *H. perniciosus* strain HP10 and estimate its evolution time relative to species in the order Hypocreales; (2) to perform comparative analyses of genome repertoires among species in the same family. The genomic data of *H. perniciosus* will provide information regarding the genes involved in pathogenicity and can be used to assess the factors involved in host-pathogen interactions.

## 2. Materials and Methods

### 2.1. Fungal Strain

The *H. perniciosus*, strain HP10 is a brown (colony morphology), high virulent strain isolated from wet bubble-diseased *A. bisporus* at a mushroom farm in Wuhan, Hubei Province, China [3]. The identity of the pathogen was confirmed through colony morphology and microscopic analysis, Internal Transcribed Spacers (ITS) rDNA and translation elongation factor 1-alpha sequencing and by pathogenicity testing to observe characteristic symptoms of wet bubble disease on a susceptible strain of *A. bisporus* [3]. For the de novo genome sequencing of HP10, mycelial plugs of a pure culture of the strain HP10 were cultured on potato dextrose agar (PDA) plates overlaid with cellophane sheets for seven days at 25 °C under a light/dark photoperiod (12/12 h) [4]. The fungus was maintained on PDA medium at 4 °C and stored at the Engineering Research Center of Chinese Ministry of Education for Edible and Medicinal Fungi, Jilin Agricultural University (Changchun, China).

### 2.2. Genome Sequencing, Assembly, and Annotation

Genomic DNA of *H. perniciosus* HP10 was extracted from the mycelia using a CWBIOTECH Plant Extraction DNA kit following the manufacturer's instructions (CWBiotech Corporation, Beijing, China). The genomic DNA was further assessed using agarose gel electrophoresis and was quantified using a Qubit 4.0 fluorometer (Invitrogen, Carlsbad, CA, USA). The DNA was fragmented into 20 kb using a BluePippin instrument (Sage Science, Inc., Beverly, MA, USA). The construction of a 20 kb library for HP10 was carried out according to Sossah et al. [11], followed by sequencing with a PacBio Sequel sequencer (Pacific Biosciences, Menlo Park, CA, USA) at the Engineering Research Center of the Chinese Ministry of Education for Edible and Medicinal Fungi. The *H. perniciosus* genome sequence was assembled using SMARTdenovo (<https://github.com/ruanjue/smartdenovo>). The assembly completeness was assessed using the Core Eukaryotic Genes Mapping Approach (CEGMA) [14] and Benchmarking Universal Single-Copy Orthologs (BUSCO) (BUSCO v1.22) [15,16]. The repeat sequences and transposable elements in *H. perniciosus* genome were identified using RepeatMasker v4.0.5 and RepeatProteinMasker [17] based on the Repbase database (<http://www.girinst.org/repbase/>) [18]. Tandem repeats were identified by the Tandem Repeats Finder (TRF) v4.04 (<http://tandem.bu.edu/trf/trf.html>) [19]. The identified repeat sequence and the transposable elements (TEs) within the genome assembly were masked with RepeatMasker [17]. Putative protein-coding genes in *H. perniciosus* genome were predicted using a combination of de novo and homology-based strategy. First, homology-based search was performed with proteomes of *Fusarium redolens*, *Fusarium oxysporum* FOX64, *Neurospora crassa*, and *Trichoderma atroviride* using Genewise [20]. Augustus v2.7 [21], GlimmerHMM

v3.02 [22], Genscan v1.0 [23], and SNAP v 2006-07-28 [24] were used for the de novo prediction. GLEAN (<http://sourceforge.net/projects/glean-gene>) [25] was used to integrate the results from both homology-based and de novo prediction to produce a final consensus gene set. Functional annotation of *H. perniciosus* predicted genes was performed by using BLASTP (e-value of  $1 \times 10^{-5}$ ) to search against several protein databases such as the National Center for Biotechnology Information (NCBI) non-redundant (nr), Cluster of Orthologous Groups (COG) [26], the Gene Ontology (GO) database [27], the Kyoto Encyclopedia of Genes and Genomes (KEGG) database [28], the SwissProt database [29–31] the TrEMBL databases [31], and the InterPro Protein Families Database [32]. In addition, transfer RNA (tRNA), ribosomal RNA (rRNA) and non-coding RNAs (small nuclear RNA (snRNAs) and microRNAs (miRNAs) were annotated using tRNAscan-SE 1.3.1 [33], RNAmmer [34] and the Rfam database [35] respectively. The presence of mating-type idiomorph in *H. perniciosus* genome was assessed by sequence similarity searches using mating-type genes and flanking gene sequences from the order Hypocreales retrieved from NCBI database as query. The mating type sequence was searched against the NCBI's conserved domain database [36] to find the conserved domains. The presence or absence of any gene was rechecked through mapping back the reads on these genes [37].

### 2.3. Functional Annotation of Pathogenicity-Related Genes and Secondary Metabolites

Carbohydrate-active enzymes (CAZymes), protease, transporters, and cytochrome P450s were predicted using the CAZymes database, MEROPS, Transporter Classification database, and the P450 database [38–42], respectively. Secretory proteins, membrane proteins, and Glycosylphosphatidylinositol (GPI)-anchored proteins were identified by SignalP 3.0 [43], TMHMM [44], and PredGPI [45], respectively. Protein localization signals were identified using TargetP [46]. The virulence-associated genes were identified using the pathogen-host interaction database (PHI) [47] and database of fungal virulence factors (DFVF) [48]. Secondary metabolites were predicted by combining antiSMASH 4.0 (<https://fungismash.secondarymetabolites.org#!/about>) [49] and the NaPDoS (<http://napdos.ucsd.edu>) database [50].

### 2.4. Orthology and Phylogenetic Analysis

The protein sequences of *H. perniciosus* (HP) and eight other reported fungal species were first used to identify orthologous gene families using OrthoMCL (v.2.0.9) [51], including *Cladobotryum protrusum* (CP), *Clonostachys rosea* (CR), *Lecanicillium fungicola* (LF) *Metarhizium acridum* (MA), *Tolypocladium inflatum* (TI), *Trichoderma longibrachiatum* (TL), *T. reesei* (TR), and *T. virens* (TV). A total of 4043 single-copy genes were identified among the nine species. The proteins of single-copy gene families were aligned using MUSCLE [52]. The poorly aligned positions and divergent regions of each codon alignment were trimmed with Gblocks [53] and concatenated into one super gene for each species. RAxML-8.0.26 [54] was used to construct a maximum-likelihood (ML) phylogenetic tree using the LG+I+G+F amino acid substitution matrix model selected by ProtTest (v. 3.4) [55] with 1000 bootstrap replicates. The species divergence times among the species were estimated with A Markov chain Monte Carlo (MCMC)Tree module in Phylogenetic Analysis by Maximum Likelihood (PAML) v4.7a [56]. The calibration times obtained from TimeTree (<http://www.timetree.org>) [57]. Computational Analysis of Gene Family Evolution (CAFE) [58] was then used to identify the expansion and contraction of the orthologous gene families across these species with a stochastic birth and death model using a lambda value of 0.314, a *p*-value of 0.01, and 1000 random samples [59]. The evidence of positive selection of genes were tested using the branch-site model of CodeML in PAML [56] software by calculating the dN/dS ratio between the species in the phylogenetic tree ( $p \leq 0.01$ ). The False Discovery Rate (FDR) estimates were computed using the Benjamini–Hochberg procedure and  $FDR \leq 0.05$  were considered significant. The positively selected genes identified were annotated in PFAM, GO and KEGG databases.

### 3. Results

#### 3.1. De Novo Genome Sequencing of *H. pernicius*

The de novo whole-genome sequencing of *H. pernicius* HP10 yielded 8.38 Gb of sequenced data using two SMRT cells (190 X sequence read coverage), which were assembled into 23 scaffolds (Table 1). The de novo assembled genome of HP10 was estimated to be ~44 Mb in size and contained a scaffold N50 length of 5.10 Mbp and a scaffold N90 length of 2.01 Mbp, as well as a 44.39% GC content (Table 1). Its assembly completeness was then evaluated with CEGMA and BUSCO using the fungal database. The analysis showed 98% and 99% for CEGMA and BUSCO, respectively, confirming that the HP10 genome was of high quality. This genome has been deposited into the NCBI database under accession number: SPDT00000000.

**Table 1.** Summary of *H. pernicius* HP10 genome assembly and annotation statistics.

HP10	
Total scaffold number	23
Total length of sequences	44,006,492 bp
N50	5,099,276 bp
N90	2,013,553 bp
GC content	44.39%
Annotated protein-coding genes	10,077
Repeat sequences proportion	25.27%
ncRNA proportion	0.27%
CEGMA	97.98%
BUSCO	99.3%

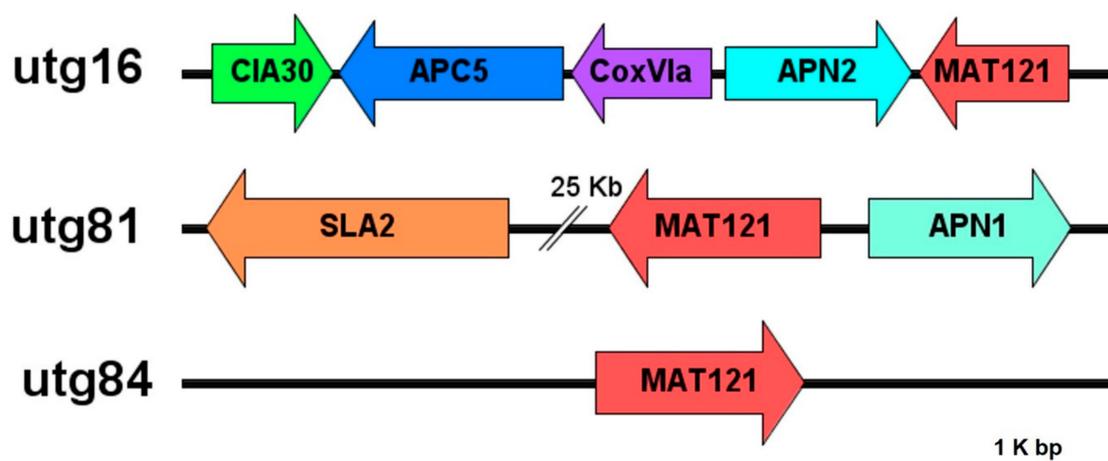
A total of 10,077 genes with an average sequence length of 1683.14 bp in *H. pernicius* HP10 were predicted by the combination of homology-based and de novo-based gene prediction methods. Of these genes, the coding sequences with an average length of 1508.99 bp contained 2.59 exons with an average length of 583.55 bp and introns with an average length of 109.82 bp. We then functionally annotated these protein-coding genes using seven databases and found that a total of 9768 (96.93%) genes could be classified in at least one database. Among them, 10,607 (96.40%) genes were homologous in the NCBI Non-Redundant Protein Sequence Database (NR) database, followed by the TrEMBL database (9736, 96.62%), InterPro database (7112, 70.58%), Swiss-Prot database (6331, 62.83%), KEGG database (6211, 61.64%), GO database (5306, 52.65%), and Clusters of orthologous groups for eukaryotic complete genomes (KOG) database (4343, 43.10%) (Supplementary Figure S1). In addition, combining homology-based and de novo prediction of non-coding RNA (ncRNA) and repeat elements revealed HP10 contained 0.17% of ncRNA and 25.27% of repeat elements. The annotated ncRNAs included 275 transfer RNA (tRNA) genes, 223 ribosomal RNA (rRNA) genes, 31 small nuclear RNA (snRNA) genes, and 53 microRNA (miRNA) genes (Supplementary Table S1). For the annotation of the repeat elements (Supplementary Table S2), the most abundantly characterized elements were long terminal repeats (LTRs), which accounted for 14.56% of the genome, whereas DNA transposons comprised only 1.23%. The other class of retrotransposons, the non-LTR elements (short interspersed elements (SINEs) and long interspersed nuclear elements (LINEs) represented 4.06% of the assembled sequence in which nearly no SINE elements (0.005%) were detected.

#### 3.2. Identification of Mating-Type Idiomorphs in *H. pernicius*

We identified *MAT 1-2* locus (three copies) in *H. pernicius* HP10 genome based on the presence of conserved domains and sequence similarities with mating-type genes of the order Hypocreales, but the *MAT 1* idiomorph was not found. The three copies of *MAT1-2-1* genes in HP10 genome were located on scaffold utg16, utg81 and utg84. No flanking genes were located on scaffold utg84, however, the *MAT1-2-1* gene was flanked by *SLA* (cytoskeleton assembly control) upstream and

*APN1* (DNA lyase) downstream on scaffold 81. The other flanking genes (*APN2* (DNA lyase), *APC5* (Anaphase-promoting complex subunit 5), *CIA30* (Complex I intermediate-associated protein 30), and *CoxVIa* (Cytochrome c oxidase subunit VIa)) were located upstream of the *MAT1-2-1* gene on scaffold *utg16*.

The fungus also encodes many other sex-related genes such as pheromone receptor genes, meiotically up-regulated gene, *Clr5* domain, and mitogen-activated protein kinase genes. Figure 1 shows the proposed organization of the mating-type gene structure of HP10 and *Mat1-2-1* gene structure as observed from the genome annotation.



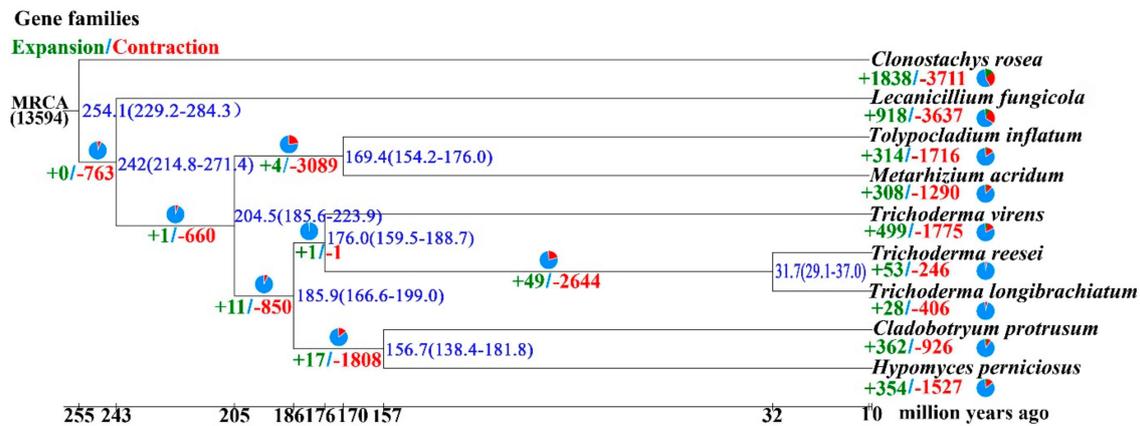
**Figure 1.** Organization of the mating-type loci (MAT) and flanking regions for in *H. perniciosus* HP10 genome. The arrows represent the orientation of the MAT1-2 genes *APC5* (Anaphase-promoting complex subunit 5), *APN1* (DNA lyase), *APN2* (DNA lyase), *CIA30* (Complex I intermediate-associated protein 30), *CoxVIa* (Cytochrome c oxidase subunit VIa) and *SLA* (cytoskeleton assembly control).

### 3.3. Genome Evolution and Phylogenomic Analysis of *H. perniciosus*

The proteomes of HP10 and eight other fungal species (Supplementary Table S3) in the order Hypocreales were clustered into 13,755 gene families using OrthoMCL, including entomopathogenic *Metarhizium acridum* (MA), and *Tolypocladium inflatum* (TI), mycoparasitic *C. protrusum* (CP), *Clonostachys rosea* (CR), *Lecanicillium fungicola* (LF), *T. longibrachiatum* (TL), and *T. virens* (TV), and saprotrophic, *Trichoderma reesei* (TR). A total of 4802 (34.91%) gene families were shared among all nine species. HP10 contained 963 (7.00%) unclustered genes, 63 (0.46%) unique gene families and an average of 1.1 genes per family. Overall, 4043 (38.11%) single-copy genes were identified in all nine species and were used to construct a phylogenetic tree (Figure 2). The tree placed HP and CP as sister taxa, that is closely related to *Trichoderma* spp., consistent with its classification in the family Hypocreaceae. HP is distantly related to MA, TI, LF, and CR. MA, TI, LF, and CR belong to the families Clavicipitaceae, Ophiocordycipitaceae, Cordycipitaceae and Bionectriaceae, respectively. Divergence time analysis showed that all the species diverged from a common ancestor of Hypocrealean fungi around 254.1 million years ago (MYA). HP is most closely related to CP with an estimated divergence time of ~156.7 MYA. The divergence time between *Trichoderma* species and CP and HP is estimated to be 185.9 MYA.

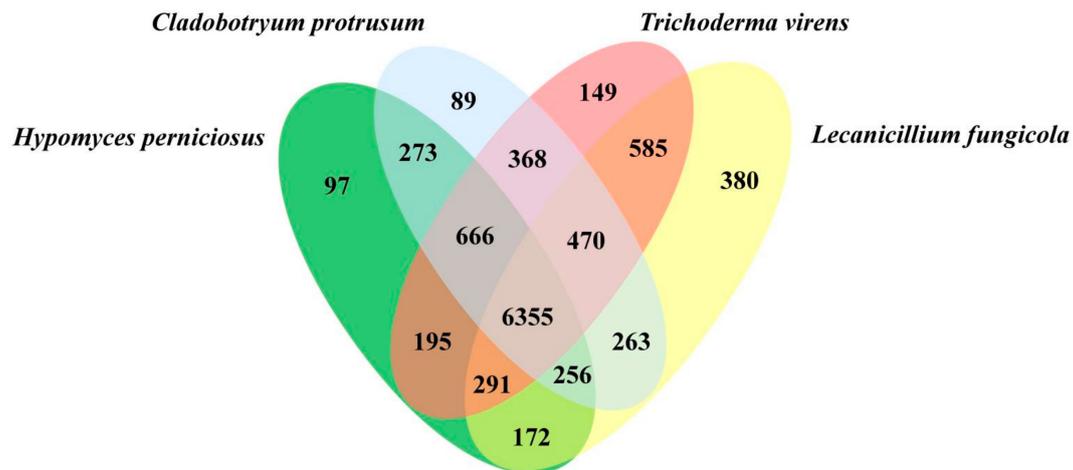
We further identified a total of 10,609 orthologous gene clusters among the closely related Hypocreales (mycoparasites) species that cause disease on *A. bisporus*, including HP, CP, TV, and LF (Figure 3). Among them, 6355 (59.90%) gene clusters were shared among all four species, representing the ancestral gene families. In addition, 97 (0.91%), 89 (0.84%), 149 (1.41%), and 380 (3.58%) gene clusters were unique to HP, CP, TV, and LF, respectively. The numbers of orthologous gene clusters shared among HP and CP, HP and TV, HP and LF were 273 (2.57%), 195 (1.84%), and 172 (1.62%), respectively. The highest number of gene clusters 666 (6.28%) shared uniquely among three species

was HP, CP, and TV. The *H. pernicius* enriched genes were related to cellular processes, CAZymes, peptidase, and Secondary Metabolites (SMs) and showed potential for pathogenicity and adaptation to the environment.



**Figure 2.** The *H. pernicius* genome phylogeny and analysis of contraction and expansion gene families analysis among nine fungal species. The divergence time range (blue), the number of expanded (green) and contracted (red) gene families is shown at each branch. The estimated divergence time (MYA: million years ago) is shown at the bottom. MRCA: most recent common ancestor.

Analysis of 5092 shared gene families based on the constructed phylogenetic tree revealed 364 (6.95%) expanded ( $p < 0.05$ ) and 1527 (29.99%) contracted gene families in HP10 (Figure 2). Gene contraction was widespread during the diversification of Hypocreales lineages, and the largest gene family loss was observed for CR 72.88% (3711). Similarly, the Hypocreaceae lineage was characterized by the loss of gene families 16.69% (850) and the divergence of HP10 and CP 35.51% (1808). With the exception of TV 34.86% (1775) and HP 29.99% (1527), the gene family contraction was smaller in the family Hypocreaceae than in the other species (246–926). The 1527 gene family contractions calculated in HP contained 1535 (15.23%) genes, out of which only eight (8) genes were significant at  $p < 0.5$ . PFAM annotation revealed all the eight genes were ATP-binding cassette (ABC) transporter proteins. Similarly, the expanded gene families in HP 4.31% (354) contained 527 (5.23%) genes with 270 genes significant at ( $p < 0.5$ ). 10.34% (28) of the significant 270 genes were genes with no known annotation in the seven different databases used for the annotation of the predicted-coding genes. PFAM annotation of the expanded genes revealed genes related to peptidases, CAZymes, SMs, and adaptation to the environment. The five most abundant expanded genes were ankyrin repeat (17), phosphorylase superfamily (14), protein kinase domain (12), BCS1 N terminal (9), and protein of unknown function (DUF3638) (9). To identify genes under positive selection, the 5092 orthologous genes from the phylogenetic analysis was used in PAML analysis [56]. The results revealed 8.3% (423 genes at  $p < 0.01$ ) were under positive selection in *H. pernicius* (Supplementary Table S5). Among them, 342 genes were annotated in the PFAM database. The most abundant GO terms for PSGs for biological process, cellular component and molecular function were cellular process (3 genes), membrane (12 genes) and binding (15 genes), respectively (Supplementary Figure S2). The top positively selected genes with enriched GO terms include nucleotide binding, nucleoside phosphate binding, oxidoreductase activity, oxidizing metal ions, and ferroxidase activity.



**Figure 3.** The Venn diagram shows shared and unique orthologous gene clusters among different mycoparasitic fungi (CP, HP, TV, and LF) that cause disease on *A. bisporus*.

### 3.4. Carbohydrate-Active Enzyme (CAZymes) in *H. perniciosus*

The CAZymes analysis revealed the genome of HP10 contained a total of six classes: glycoside hydrolases (GH = 186), glycosyl transferases (GTs = 83), carbohydrate esterases (CEs = 6), auxiliary activities (AAs = 30), carbohydrate-binding molecule (CBMs = 28), and polysaccharide lyases (PL = 3). These diverse gene families are associated with fungal cell wall synthesis, modification, and degradation. We then used the hidden markov model (HMMs) from dbCAN2 meta server [60] to identify 119 CAZymes encoding genes in the secretome, including 12 CE, 84 GH, 3 GT, 4 PL, and 16 AA (Supplementary Table S6). Among them, the GH18, GH55, and GH75 families associated with fungal cell wall degradation are highly represented in HP10 secretome. We next compared the CAZymes number of HP with other pathogenic fungi with a different lifestyle and host mushrooms (Figure 4). The CAZymes numbers of HP10 were higher than those of the host *A. bisporus*, the entomophagous fungi (MA and TI) and saprophytic TR, but lower compared with the mycoparasitic fungi (TL, TV, and CR) (Supplementary Table S7).

	HP	TL	TR	TV	CR	MA	TI	AB
AA1	6	6	6	7	11	7	7	12
AA11	4	3	3	4	5	5	3	0
AA12	2	1	1	1	3	1	1	1
CBM1	1	16	16	25	40	2	2	15
CBM12	0	0	0	0	2	0	0	0
CBM13	2	0	0	2	12	2	2	9
CBM14	0	1	0	0	0	1	1	0
CBM18	17	7	6	15	8	5	4	1
CE1	0	1	1	4	6	1	0	1
CE12	0	0	0	0	5	0	0	0
CE15	0	1	1	1	1	0	0	1
CE16	1	2	2	3	8	1	1	3
GH18	41	8	21	34	16	6	4	14
GH55	6	1	5	9	3	0	0	0
GH74	0	3	1	1	0	2	2	1
GH75	4	7	3	5	4	12	6	0
GH76	7	1	8	8	11	0	1	0
GH16	12	4	15	15	27	4	5	15
GT1	11	6	7	7	22	12	9	3
GT20	5	4	4	3	3	4	3	3
GT31	8	9	9	10	10	10	7	2
PL20	1	2	2	2	1	1	0	0
PL4	0	0	0	0	4	0	0	1
PL7	2	2	2	3	0	1	1	0

**Figure 4.** CAZyme richness of different fungi with different lifestyles (HP, TL, TV, and CR, mycoparasitic; AB (*A. bisporus*) and TR, saprotrophic; MA and TI, entomopathogenic).

Among genes encoding GH families, GH18 was the most abundant CAZymes module in HP10 compared to all of the other species analyzed. Other GH families ( $\beta$ -1,3-Glucanases) in HP10 involved in mycoparasitism include GH5, 16, 17, 30, 55, 64, and 81. In addition, four genes encoding GH75 (chitosanases) were found in HP10, which are known to degrade chitin [61], but were absent in the host *A. bisporus*. In contrast to the host *A. bisporus*, HP10 also lost gene-encoding cellobiohydrolases (GH6 and GH7) that are key enzymes regulating cellulose degradation and contracted nearly half the number of AA genes involved in lignin degradation. In addition, all of the fungi examined have cellulose degrading enzymes, such as GH74 with the exception of HP10 and CR. Within the family Hypocreaceae, the HP10 genome contained the highest number of CAZymes genes encoding GTs (83). The most abundant GT families in HP10 were GT1, GT31, GT90, GT32, GT8, and GT20. The number of GT2 (chitinase) was similar for all of the species analyzed. These GTs are mainly associated with chitin synthesis, cell wall biosynthesis, and glycosylation. All four Hypocreaceae fungal species possessed members of two families of putative polysaccharide lyases (PL7 and PL20) that may have activity specific to pectin [13].

### 3.5. Secondary Metabolites in *H. pernicius*

The HP10 genome contained a total of 91 SMs gene clusters and 64 SM genes from the antiSMASH and NaPDoS analysis, respectively (Supplementary Tables S8 and S9). We found 14, 13, 9, and 7 clusters belonged to type 1 polyketide synthases (T1PKS), non-ribosomal peptide synthetase hybrids, terpene synthases, and non-ribosomal peptide synthetases (NRPS). The antiSMASH analysis revealed one T1PKS bikaverin biosynthetic gene cluster and four T1PKS-NRPS biosynthetic gene clusters, which are LL-Z1272beta, chaetoglobosins, betaenone (C/A) and leucinostatins, and one putative cluster, nivalenol/deoxynivalenol/3-acetyldeoxynivalenol. The NaPDoS analysis revealed 40 genes for PKS and 24 for NRPS. The top five highest pathway products for NaPDoS were: lovastatin (14), HC-toxin (10), fumonisin (9) compactin (6) and complestatin (10), and naphtopyrone (5). The majority of SMs in HP10 show genes involved in the biosynthesis of toxins, and fungus-host interaction, such as the production of antibiotics, anthelmintics, and antifungals. In addition, the high number of unknown SMs in HP10 shows the potential for the production of bioactive compounds.

### 3.6. Pathogenicity-Related Genes in *H. pernicius*

We performed BLAST analysis of HP10 proteomes against the PHI gene database, Transport Classification Database (TCDB), and DFVF to find potential genes related to virulence (Supplementary Tables S10–S12). The analysis revealed that 945, 422, and 402 protein-coding genes were involved in PHI, TCDB, and DFVF, respectively, which accounted for 9.38%, 4.19%, and 3.99% of the total proteins in HP10. We also found 205 genes of DFVF in the PHI database. The genes in the PHI database were classified as a chemistry target: resistance to chemicals (3) effectors (plant avirulence determinant) (1), enhanced antagonism (2), increased virulence (hypervirulence) (26), lethal (55), loss of pathogenicity (86), reduced virulence (343), and unaffected pathogenicity (429). The HP10 genome encodes 630 secretory proteins (6.25% of the total predicted proteins) with a signal peptide. Annotation of the secretory proteins (Supplementary Table S13) showed that *H. pernicius* secretome includes 180 PHI genes, 186 putative effector proteins (Supplementary Table S14), 119 CAZymes (Supplementary Table S6), and 191 proteases.

A number of HP10 putative PHI genes were found in the cereal pathogenic fungi *Fusarium* spp. (526 genes), *Magnaporthe oryzae* (161), *Cochliobolus* spp. (9), and *Ustilago maydis* (7). Similarly, 26 and 9 genes were similar to entomophagous fungi (*Metarhizium robertsii* (13) and *Beauveria bassiana* (13)) and necrotrophic fungi (*Botrytis cinerea* (7) and *T. vires* (2)), respectively. Interestingly, numerous HP10 genes were found in animal pathogenic fungi such as *Cryptococcus neoformans* (26) and *Candida albicans* (26). This may be due to *H. pernicius* possessing LysM, which masks chitin on the external surface of the hyphae from being detected by the hosts. This phenomenon is present with animal fungal pathogens because they lack infection structures such as appressoria to attach to the host. In addition, we identified some genes that might be involved in breaking host physical barriers and

in pathogenicity, nutrient acquisition, and adaptation to environmental stress, including proteases (399) (Supplementary Table S15), cytochrome P450 (125) (Supplementary Table S16), G protein-coupled receptors (GPCRs), hydrophobins, PacC T2 family RNases, the necrosis-inducing protein NPP1, GLEYA adhesin domain proteins, and MD-2-related lipid-recognition.

#### 4. Discussion

The button mushroom (*A. bisporus*) is of socio-economic importance for human nutrition globally, despite production constraints by fungal and bacterial diseases [62,63]. *H. perniciosus* is the causal agent of the wet bubble disease that affects *A. bisporus* and other mushrooms around the world, resulting in the loss of millions of tons of mushrooms [1,3,64]. Recent advances in biotechnology have led to the availability of the whole genome sequence of edible mushrooms and pathogenic fungi [65–68]. However, the global genomics and transcriptomic studies of the interaction between *A. bisporus* and *H. perniciosus* have rarely been investigated. In this study, we used the PacBio Sequel platform to sequence the genome of *H. perniciosus*, the first reference genome for the genus *Hypomyces*. In addition, we estimated the genome evolution time, carried out a phylogenomic study on the order Hypocreales. The completeness and quality of *H. perniciosus* genome assembly are comparable to those of the *C. protrusum* genome assembly which was generated with PacBio sequel long-read sequencing technology [11]. The *H. perniciosus* genome size was within the range observed for the family Hypocreaceae (e.g., *C. protrusum* (CP), *T. virens* (TV), and *T. harzianum* (TH)) and other sequenced ascomycetes in the class Sordariomycetes [11,67]. The genome of *H. perniciosus* is composed of 25.27% repeat sequences, the largest in the family Hypocreaceae and other pathogens such as *Pyricularia oryzae* (11.75%) [69], *Valsa mali* (14.05%) [70], and *Ustilago virens* (25%) [71]. The expansion of repeat sequences is a common feature among some fungi with known large genomes, such as *Puccinia graminis* (45%) [72], *Tuber melanosporum* (58%) [73], *Blumeria graminis* (64%) [74], and *Ophiocordyceps sinensis* (74.67%) [75]. The higher number of repeat elements in *H. perniciosus* could be a result of poor or inefficient DNA removal mechanisms compared to other members in the family Hypocreaceae and may explain its relatively large genome size [75]. Previous studies have shown that the repeat sequences have an important role in driving pathogenicity variation in many pathogenic fungi and enable us to better understand the genome architecture of a pathogen and evolution [76,77].

The protein-coding genes in *H. perniciosus* (based on the genome size) were relatively low in number (10,077) compared to those in the genomes of the mycoparasite CP (11,003), TV (12,250), and TH (14,294) [11,67]. This may be due to the high number of repeat sequences in the genome. The genes are mainly located in repeat sequences poor regions, and the gene density is heterogeneous when compared with that of other ascomycetes [73]. *H. perniciosus* shares the highest number of unique orthologous gene families with *Trichoderma* (or members of the family Hypocreaceae especially the genera *Escovopsis*, *Cladobotryum*, or *Trichoderma*). Phylogenetic analysis from the total of 4043 gene families shared among all of the nine species used revealed *H. perniciosus* belonged to the family Hypocreaceae and diverged from its common ancestor ~156.7 million years ago. This divergence time is consistent with those reported by Sung et al. [78].

We identified several expanded gene families, such as different types of ankyrin repeats, glycosyl hydrolases, lipases, peptidases, transporters (ABC, MFS, sugar, ion, and phosphate), and several groups of predicted signaling genes, including kinases and transcription factors; in addition, SMs are lineage-specific genes, suggesting that important protein-coding innovation occurred in these lineages [79]. The gene family expansion in *H. perniciosus* might have contributed to its pathogenicity, host lifestyle, the wide adaptability to varied climates, and worldwide infection of *A. bisporus* and other mushrooms [80]. Moreover, gene family contraction in *H. perniciosus* was more prominent than gene family expansions. We hypothesize that these gene contractions have limited the host range of *H. perniciosus* compared to *Trichoderma* species. Previous studies also have shown the massive gene contraction in pathogens might lead to the adaptation of host changes [81,82]. Fungal effectors, cell wall degrading enzymes and secondary metabolites genes are known to be under

positives selection in pathogenic fungi [83]. The positively selected genes in *H. perniciosus* may have contributed to its mycoparasitic lifestyle and adaptation to a harsh environment [80,84,85]. The comparative analysis between *H. perniciosus* and the other mycoparasites that cause disease in *A. bisporus* revealed the genome structure and overlap among the orthologous clusters, which can elucidate gene function and evolution of the proteins involved in pathogenicity [80]. Identification of those genes is relevant in designing effective disease control structures that can be applied to all of the pathogens examined. Fungi are known to exhibit two different sexual lifestyles, homothallism (self-fertility) and heterothallism (self-sterility) [86]. The organization of *H. perniciosus* HP10 mating-type and flanking genes (*SLA-MAT-APN*) is consistent with other representation in Hypocreales [86]. The analysis of the mating-type structure in *H. perniciosus* HP10 genome suggests the fungus is heterothallic. *Cladobotryum protrusum* [11] and *Trichoderma reesei* [87] in the family Hypocreaceae are heterothallic and also contain the MAT 1-2 idiomorph. The three copies of *MAT1-2-1* may be due to the number of nuclei in the conidia. As of now, there is no report of mating-type idiomorph in *H. perniciosus*. The genome sequence of *H. perniciosus* has provided some insights into the mating structure of the pathogen. However, many fungi can switch between heterothallism and homothallism [85,88], therefore, additional functional characterization of the mating type genes including the identification of the flanking genes and genome resequencing of various populations of *H. perniciosus* is required to ascertain these results.

The total number of secretory proteins, CAZymes, P450, and SMs in *H. perniciosus* is smaller than that of *C. protrusum* and *Trichoderma* species. The secretory proteins and pathogenicity-related genes are crucial effectors responsible for host-pathogen interactions [89,90]. The analysis of *H. perniciosus* secretory proteins and pathogenicity-related genes shows that there is a likelihood that the fungus does not require considerably large numbers of CAZymes and SMs during host infection, but *H. perniciosus* utilizes secretory proteins, key pathogenic GHs (GH18, GH75, GH 76 GH 16), CBM50 (LysM domain), CBM18 and CE4 (chitin deacetylase), peptidases, and SMs to infect its hosts [86,91].

## 5. Conclusions

In summary, the genome analysis of *H. perniciosus* has provided insights into the evolution, pathogenicity, and genes influencing the pathogen interaction with its fungal host. The genomic resource data generated in our study can be further mined and the identified genes characterized to explain the basis of the mycoparasitic lifestyle in *H. perniciosus* and the family Hypocreaceae in general, and it can serve as the basis for developing efficient disease-management strategies to mitigate wet bubble disease.

**Supplementary Materials:** Supplementary materials can be found at <http://www.mdpi.com/2073-4425/10/6/417/s1>. Table S1: Number of RNA genes in the *Hypomyces perniciosus* genome. Table S2: Detailed classification of identified repeat elements in *Hypomyces perniciosus* genome. Table S3: Genome features and background information of *Hypomyces perniciosus* and other representative fungi analyzed in this study. Table S4: Gene family contraction and expansion analysis results. Table S5: Annotation of positive selection genes during the evolution of *Hypomyces perniciosus*. Table S6: List of carbohydrate-active enzyme genes in *Hypomyces perniciosus* secretome. Table S7: (A) Total number of carbohydrate-active enzymes in *Hypomyces perniciosus* genome compared to other eight fungi used in this study. (B) Detailed list of carbohydrate-active enzymes modules in *Hypomyces perniciosus* genome compared to other eight fungi used in this study. Table S8: (A) Total number of antiSMASH secondary metabolites gene clusters in *Hypomyces perniciosus* genome. (B) The number of antiSMASH secondary metabolites gene clusters in *Hypomyces perniciosus* with known annotation. Table S9: Total number of NapDoS secondary metabolites genes in *Hypomyces perniciosus* genome. Table S10: The pathogen-host interaction (PHI) genes and annotation found in *Hypomyces perniciosus* genome. Table S11: List of membrane transport proteins in *Hypomyces perniciosus* genome from Transporter Classification Database (TCDB). Table S12: List of fungal virulence factors genes in *Hypomyces perniciosus* genome from database of fungal virulence factors (DFVF). Table S13: List of secretory protein genes in *Hypomyces perniciosus* genome. Table S14: List of effector protein genes from the secretome of *Hypomyces perniciosus* genome. Table S15: List of protease genes in *Hypomyces perniciosus* genome from MEROPS database. Table S16: List of Cytochromes P450 genes in *Hypomyces perniciosus* genome. Figure S1: Venn diagram representing the annotation of *Hypomyces perniciosus* HP10 in five different databases. Figure S2: Number of genes.

**Author Contributions:** Conceptualization, Y.F. and Y.L.; Funding acquisition, Y.F. and Y.L.; Writing—original draft, D.L. and Y.F.; Writing—review & editing, D.L. and Y.F.; Formal analysis, D.L. and Y.F.; Software, D.L., L.S. and F.L.S.; Methodology, D.L. and F.L.S.; Resources, D.L., L.S. and F.L.S.

**Funding:** This research was funded by National Natural Science Foundation of China (No. 31701977); Science and Technology Research “13th Five-Year Plan” Projects of Education Department of Jilin Province ((2016) 191); Changchun Science and Technology Project (No. 15SS11); National Natural Science Foundation of China (No. 31700012); The Special Fund for Agro-scientific Research in the Public Interest (No. 201503137); The Program of Creation and Utilization of Germplasm of Mushroom Crop of “111” Project (No. D17014); National-level International Joint Research Centre (2017B01011); The International Agro-biological Resources and the Collection of Information about Agro-technic Demands and Agricultural Strategies (2016-X07).

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

## References

1. Fletcher, J.T.; Jaffe, B.; Muthumeenakshi, S.; Brown, A.E.; Wright, D.M. Variations in isolates of *Mycogone perniciosa* and in disease symptoms in *Agaricus bisporus*. *Plant Pathol.* **1995**, *44*, 130–140. [[CrossRef](#)]
2. Smith, F.E.V. Three diseases of cultivated mushrooms. *Trans. Br. Mycol. Soc.* **1924**, *10*, 81–85. [[CrossRef](#)]
3. Fu, Y.P.; Wang, X.X.; Li, D.; Liu, Y.; Song, B.; Zhang, C.L.; Wang, Q.; Chen, M.Y.; Zhang, Z.W.; Li, Y. Identification of resistance to wet bubble disease and genetic diversity in wild and cultivated strains of *Agaricus bisporus*. *Int. J. Mol. Sci.* **2016**, *17*, 1568. [[CrossRef](#)]
4. Zhang, C.L.; Kakishima, M.; Xu, J.; Wang, Q.; Li, Y. The effect of *Hypomyces pernicius* on the mycelia and basidiomes of *Agaricus bisporus*. *Microbiology* **2017**, *163*, 1273–1282. [[CrossRef](#)]
5. Wang, W.; Li, X.; Chen, B.Z.; Wang, S.; Li, C.H.; Wen, Z.Q. Analysis of genetic diversity and development of SCAR Markers in a *Mycogone perniciosa* population. *Curr. Microbiol.* **2016**, *73*, 9–14. [[CrossRef](#)]
6. Zhou, C.; Li, D.; Chen, L.; Li, Y. Genetic diversity analysis of *Mycogone perniciosa* causing wet bubble disease of *Agaricus bisporus* in China using SRAP. *J. Phytopathol.* **2016**, *164*, 271–275. [[CrossRef](#)]
7. Du, P.Q.; Wu, X.H.; He, H.; Zhang, Y.; Xu, J.; Dong, F.S.; Zheng, Y.Q.; Liu, X.G. Evaluation of the safe use and dietary risk of beta-cypermethrin, pyriproxyfen, avermectin, diflubenzuron and chlorothalonil in button mushroom. *Sci. Rep.* **2017**, *7*, 8694. [[CrossRef](#)] [[PubMed](#)]
8. Gandy, D.G.; Spencer, D.M. Fungicides for the control of *Mycogone perniciosa* (Magn.), the cause of wet bubble on the cultivated mushroom. *Sci. Hortic.* **1978**, *8*, 307–313. [[CrossRef](#)]
9. Glamoclija, J.; Soković, M.; Ljaljević-Grbić, M.; Vukojević, J.; Milenković, I.; Van Griensven, L. Morphological characteristics and mycelial compatibility of different *Mycogone perniciosa* isolates. *J. Microsc.* **2008**, *232*, 489–492. [[CrossRef](#)] [[PubMed](#)]
10. Schmidt, S.M.; Lukasiewicz, J.; Farrer, R.; van Dam, P.; Bertoldo, C.; Rep, M. Comparative genomics of *Fusarium oxysporum* f. sp. melonis reveals the secreted protein recognized by the Fom-2 resistance gene in melon. *New Phytol.* **2016**, *209*, 307–318. [[CrossRef](#)]
11. Sossah, F.L.; Liu, Z.H.; Yang, C.T.; Okorley, B.A.; Sun, L.; Fu, Y.P.; Li, Y. Genome sequencing of *Cladobotryum protrusum* provides insights into the evolution and pathogenic mechanisms of the cobweb disease pathogen on cultivated mushroom. *Genes* **2019**, *10*, 124. [[CrossRef](#)] [[PubMed](#)]
12. Berendsen, R.L.; Baars, J.J.P.; Kalkhove, S.I.C.; Lugones, L.G.; Wösten, H.A.; Bakker, P.A.H.M. *Lecanicillium fungicola*: causal agent of dry bubble disease in white-button mushroom. *Mol. Plant Pathol.* **2010**, *11*, 585–595. [[CrossRef](#)]
13. Gruninger, R.J.; Nguyen, T.T.M.; Reid, I.D.; Yanke, J.L.; Wang, P.; Abbott, D.W.; Tsang, A.; McAllister, T. Application of transcriptomics to compare the Carbohydrate Active Enzymes that are expressed by diverse genera of anaerobic Fungi to degrade plant cell wall carbohydrates. *Front. Microbiol.* **2018**, *9*, 1581. [[CrossRef](#)] [[PubMed](#)]
14. Parra, G.; Bradnam, K.; Korf, I. CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. *Bioinformatics* **2007**, *23*, 1061–1067. [[CrossRef](#)] [[PubMed](#)]
15. Kriventseva, E.V.; Simão, F.A.; Klioutchnikov, G.; Seppey, M.; Manni, M.; Ioannidis, P.; Waterhouse, R.M.; Zdobnov, E.M. BUSCO Applications from quality assessments to gene prediction and phylogenomics. *Mol. Biol. Evol.* **2017**, *35*, 543–548.

16. Simao, F.A.; Waterhouse, R.M.; Ioannidis, P.; Kriventseva, E.V.; Zdobnov, E.M. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **2015**, *31*, 3210–3212. [[CrossRef](#)]
17. Tarailo-Graovac, M.; Chen, N. Using repeatMasker to identify repetitive elements in genomic sequences. *Curr. Protoc. Bioinform.* **2009**, *25*, 4–10.
18. Bao, W.; Kojima, K.K.; Kohany, O. Repbase Update, a database of repetitive elements in eukaryotic genomes. *Mob. DNA* **2015**, *6*, 11. [[CrossRef](#)]
19. Benson, G. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res.* **1999**, *27*, 573–580. [[CrossRef](#)]
20. Birney, E.; Clamp, M.; Durbin, R. GeneWise and Genomewise. *Genome Res.* **2004**, *14*, 988–995. [[CrossRef](#)]
21. Stanke, M.; Keller, O.; Gunduz, I.; Hayes, A.; Waack, S.; Morgenstern, B. AUGUSTUS: ab initio prediction of alternative transcripts. *Nucleic Acids Res.* **2006**, *34*, 435–439. [[CrossRef](#)] [[PubMed](#)]
22. Majoros, W.H.; Pertea, M.; Salzberg, S.L. TigrScan and GlimmerHMM: two open source ab initio eukaryotic gene-finders. *Bioinformatics* **2004**, *20*, 2878–2879. [[CrossRef](#)] [[PubMed](#)]
23. Burge, C.; Karlin, S. Prediction of complete gene structures in human genomic DNA. Edited by F.E. Cohen. *J. Mol. Biol.* **1997**, *268*, 78–94. [[CrossRef](#)]
24. Korf, I. Gene finding in novel genomes. *BMC Bioinform.* **2004**, *5*, 59. [[CrossRef](#)] [[PubMed](#)]
25. Elsik, C.G.; Mackey, A.J.; Reese, J.T.; Milshina, N.V.; Roos, D.S.; Weinstock, G.M. Creating a honey bee consensus gene set. *Genome Biol.* **2007**, *8*, 13. [[CrossRef](#)] [[PubMed](#)]
26. Tatusov, R.L.; Fedorova, N.D.; Jackson, J.D.; Jacobs, A.R.; Kiryutin, B.; Koonin, E.V.; Krylov, D.M.; Mazumder, R.; Mekhedov, S.L.; Nikolskaya, A.N.; et al. The COG database: an updated version includes eukaryotes. *BMC Bioinform.* **2003**, *4*, 41. [[CrossRef](#)] [[PubMed](#)]
27. Ashburner, M.; Ball, C.A.; Blake, J.A.; Botstein, D.; Butler, H.; Cherry, J.M.; Davis, A.P.; Dolinski, K.; Dwight, S.S.; Eppig, J.T.; et al. Gene Ontology: tool for the unification of biology. *Nat. Genet.* **2000**, *25*, 25–29. [[CrossRef](#)]
28. Kanehisa, M.; Furumichi, M.; Tanabe, M.; Sato, Y.; Morishima, K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* **2016**, *45*, 353–361. [[CrossRef](#)]
29. Magrane, M.; Consortium, U. UniProt Knowledgebase: a hub of integrated protein data. *Database* **2011**, *2011*, bar009. [[CrossRef](#)] [[PubMed](#)]
30. Bairoch, A.; Apweiler, R. The SWISS-PROT protein sequence data bank and its new supplement TREMBL. *Nucleic Acids Res.* **1996**, *24*, 21–25. [[CrossRef](#)]
31. Bairoch, A.; Apweiler, R. The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. *Nucleic Acids Res.* **2000**, *28*, 45–48. [[CrossRef](#)]
32. Mitchell, A.; Chang, H.-Y.; Daugherty, L.; Fraser, M.; Hunter, S.; Lopez, R.; McAnulla, C.; McMenamin, C.; Nuka, G.; Pesseat, S.; et al. The InterPro protein families database: the classification resource after 15 years. *Nucleic Acids Res.* **2014**, *43*, 213–221. [[CrossRef](#)]
33. Lowe, T.M.; Eddy, S.R. tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* **1997**, *25*, 955–964. [[CrossRef](#)] [[PubMed](#)]
34. Lagesen, K.; Hallin, P.; Rødland, E.A.; Stærfeldt, H.-H.; Rognes, T.; Ussery, D.W. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* **2007**, *35*, 3100–3108. [[CrossRef](#)] [[PubMed](#)]
35. Gardner, P.P.; Eldai, H. Annotating RNA motifs in sequences and alignments. *Nucleic Acids Res.* **2014**, *43*, 691–6987. [[CrossRef](#)] [[PubMed](#)]
36. Marchler-Bauer, A.; Derbyshire, M.K.; Gonzales, N.R.; Lu, S.; Chitsaz, F.; Geer, L.Y.; Geer, R.C.; He, J.; Gwadz, M.; Hurwitz, D.I.; et al. CDD: NCBI’s conserved domain database. *Nucleic Acids Res.* **2014**, *43*, 222–226. [[CrossRef](#)] [[PubMed](#)]
37. Agrawal, Y.; Khatri, I.; Subramanian, S.; Shenoy, B.D. Genome sequence, comparative analysis, and evolutionary insights into chitinases of entomopathogenic Fungus *Hirsutella thompsonii*. *Genome Biol. Evol.* **2015**, *7*, 916–930. [[CrossRef](#)] [[PubMed](#)]
38. Rawlings, N.D.; Barrett, A.J.; Thomas, P.D.; Huang, X.; Bateman, A.; Finn, R.D. The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. *Nucleic Acids Res.* **2018**, *46*, 624–632. [[CrossRef](#)]
39. Saier, M.H.; Reddy, V.S.; Tsu, B.V.; Ahmed, M.S.; Li, C.; Moreno-Hagelsieb, G. The transporter classification database (TCDB): recent advances. *Nucleic Acids Res.* **2016**, *44*, 372–379. [[CrossRef](#)]

40. Park, J.; Lee, S.; Choi, J.; Ahn, K.; Park, B.; Park, J.; Kang, S.; Lee, Y.H. Fungal cytochrome P450 database. *BMC Genom.* **2008**, *9*, 402. [[CrossRef](#)]
41. Nelson, D.R. The cytochrome p450 homepage. *Hum. Genom.* **2009**, *4*, 59.
42. Lombard, V.; Golaconda Ramulu, H.; Drula, E.; Coutinho, P.M.; Henrissat, B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* **2014**, *42*, 490–495. [[CrossRef](#)] [[PubMed](#)]
43. Bendtsen, J.D.; Nielsen, H.; von Heijne, G.; Brunak, S. Improved prediction of signal peptides: signalP 3.0. *J. Mol. Biol.* **2004**, *340*, 783–795. [[CrossRef](#)] [[PubMed](#)]
44. Krogh, A.; Larsson, B.; von Heijne, G.; Sonnhammer, E.L.L. Predicting transmembrane protein topology with a hidden markov model: application to complete genomes. *J. Mol. Biol.* **2001**, *305*, 567–580. [[CrossRef](#)] [[PubMed](#)]
45. Pierleoni, A.; Martelli, P.L.; Casadio, R. PredGPI: a GPI-anchor predictor. *BMC Bioinform.* **2008**, *9*, 392. [[CrossRef](#)] [[PubMed](#)]
46. 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*, 1005–1016. [[CrossRef](#)] [[PubMed](#)]
47. Urban, M.; Cuzick, A.; Rutherford, K.; Irvine, A.; Pedro, H.; Pant, R.; Sadanadan, V.; Khamari, L.; Billal, S.; Mohanty, S.; Hammond-Kosack, K.E. PHI-base: a new interface and further additions for the multi-species pathogen–host interactions database. *Nucleic Acids Res.* **2016**, *45*, 604–610. [[CrossRef](#)] [[PubMed](#)]
48. Lu, T.; Yao, B.; Zhang, C. DFVF: database of fungal virulence factors. *Database* **2012**, *2012*, 1. [[CrossRef](#)]
49. Blin, K.; Weber, T.; Chevrette, M.G.; Lu, X.; Schwalen, C.J.; Kautsarsa, S.A.; Suarez Duran, H.G.; de Los Santos, E.L.C.; Kim, H.U.; Nave, M.; et al. antiSMASH 4.0-improvements in chemistry prediction and gene cluster boundary identification. *Nucleic Acids Res.* **2017**, *45*, 36–41. [[CrossRef](#)]
50. Ziemert, N.; Podell, S.; Penn, K.; Badger, J.H.; Allen, E.; Jensen, P.R. The natural product domain seeker NaPDoS: a phylogeny based bioinformatic tool to classify secondary metabolite gene diversity. *Plos ONE* **2012**, *7*, e34064. [[CrossRef](#)]
51. Li, L.; Stoeckert, C.J.; Roos, D.S. OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res.* **2003**, *13*, 2178–2189. [[CrossRef](#)] [[PubMed](#)]
52. Edgar, R.C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **2004**, *32*, 1792–1797. [[CrossRef](#)]
53. Talavera, G.; Castresana, J. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst. Biol.* **2007**, *56*, 564–577. [[CrossRef](#)]
54. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **2014**, *30*, 1312–1313. [[CrossRef](#)]
55. Darriba, D.; Taboada, G.L.; Doallo, R.; Posada, D. ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics* **2011**, *27*, 1164–1165. [[CrossRef](#)]
56. Yang, Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **2007**, *24*, 1586–1591. [[CrossRef](#)]
57. Hedges, S.B.; Marin, J.; Suleski, M.; Paymer, M.; Kumar, S. Tree of life reveals clock-like speciation and diversification. *Mol. Biol. Evol.* **2015**, *32*, 835–845. [[CrossRef](#)] [[PubMed](#)]
58. De Bie, T.; Cristianini, N.; Demuth, J.P.; Hahn, M.W. CAFE: a computational tool for the study of gene family evolution. *Bioinformatics* **2006**, *22*, 1269–1271. [[CrossRef](#)]
59. Floudas, D.; Binder, M.; Riley, R.; Barry, K.; Blanchette, R.A.; Henrissat, B.; Martínez, A.T.; Otilar, R.; Spatafora, J.W.; Yadav, J.S.; et al. The paleozoic origin of enzymatic Lignin decomposition reconstructed from 31 Fungal genomes. *Science* **2012**, *336*, 1715–1719. [[CrossRef](#)]
60. Zhang, H.; Yohe, T.; Huang, L.; Entwistle, S.; Wu, P.Z.; Yang, Z.Q.; Busk, P.K.; Xu, Y.; Yin, Y.B. dbCAN2: a meta server for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* **2018**, *46*, 95–101. [[CrossRef](#)]
61. Seidl, V. Chitinases of filamentous fungi: a large group of diverse proteins with multiple physiological functions. *Fungal Biol. Rev.* **2008**, *22*, 36–42. [[CrossRef](#)]
62. Kertesz, M.A.; Thai, M. Compost bacteria and fungi that influence growth and development of *Agaricus bisporus* and other commercial mushrooms. *Appl. Microbiol. Biot.* **2018**, *102*, 1639–1650. [[CrossRef](#)] [[PubMed](#)]
63. McGee, C.F. Microbial ecology of the *Agaricus bisporus* mushroom cropping process. *Appl. Microbiol. Biot.* **2018**, *102*, 1075–1083. [[CrossRef](#)]

64. Kouser, S.; Ahmed, M.; Shah, S. Disease status and yield losses due to wet bubble disease (*Mycogone perniciosa*) associated with the cultivation of white button mushroom at different mushroom units of Kashmir valley. *Plant Pathol. J.* **2013**, *12*, 104–109.
65. Kuo, A.; Salamov, A.; Korzeniewski, F.; Nordberg, H.; Shabalov, I.; Dubchak, I.; Otilar, R.; Riley, R.; Ohm, R.; Nikitin, R.; et al. MycoCosm portal: gearing up for 1000 fungal genomes. *Nucleic Acids Res.* **2013**, *42*, 699–704.
66. Grigoriev, I.V.; Cullen, D.; Goodwin, S.B.; Hibbett, D.; Jeffries, T.; Kubicek, C.P.; Kuske, C.R.; Magnuson, J.K.; Martin, F.M.; Spatafora, J.W.; et al. Fueling the future with fungal genomics. *Mycology* **2011**, *2*, 192–209.
67. Grigoriev, I.V.; Nordberg, H.; Shabalov, I.; Aerts, A.; Cantor, M.; Goodstein, D.; Kuo, A.; Minovitsky, S.; Nikitin, R.; Ohm, R.A.; et al. The genome portal of the Department of Energy Joint Genome Institute. *Nucleic Acids Res.* **2012**, *40*, 26–32. [[CrossRef](#)] [[PubMed](#)]
68. Morin, E.; Kohler, A.; Baker, A.R.; Foulongne-Oriol, M.; Lombard, V.; Nagy, L.G.; Ohm, R.A.; Patyshakuliyeva, A.; Brun, A.; Aerts, A.L.; et al. Genome sequence of the button mushroom *Agaricus bisporus* reveals mechanisms governing adaptation to a humic-rich ecological niche. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 17501–17506. [[CrossRef](#)]
69. Xue, M.; Yang, J.; Li, Z.G.; Hu, S.N.; Yao, N.; Dean, R.A.; Zhao, W.S.; Shen, M.; Zhang, H.W.; Li, C.; et al. Comparative analysis of the genomes of two field isolates of the rice blast fungus *Magnaporthe oryzae*. *Plos Genet.* **2012**, *8*, e1002869. [[CrossRef](#)]
70. Yin, Z.; Liu, H.; Li, Z.; Ke, X.; Dou, D.; Gao, X.; Song, N.; Dai, Q.; Wu, Y.; Xu, J.R.; et al. Genome sequence of Valsa canker pathogens uncovers a potential adaptation of colonization of woody bark. *New Phytol.* **2015**, *208*, 1202–1216. [[CrossRef](#)]
71. Zhang, Y.; Zhang, K.; Fang, A.; Han, Y.; Yang, J.; Xue, M.; Bao, J.; Hu, D.; Zhou, B.; Sun, X.; et al. Specific adaptation of Ustilagoidea vires in occupying host florets revealed by comparative and functional genomics. *Nat. Commun.* **2014**, *5*, 3849. [[CrossRef](#)]
72. Duplessis, S.; Cuomo, C.A.; Lin, Y.C.; Aerts, A.; Tisserant, E.; Veneault-Fourrey, C.; Joly, D.L.; Hacquard, S.; Amselem, J.; Cantarel, B.L.; et al. Obligate biotrophy features unraveled by the genomic analysis of rust fungi. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 9166–9171. [[CrossRef](#)] [[PubMed](#)]
73. Martin, F.; Kohler, A.; Murat, C.; Balestrini, R.; Coutinho, P.M.; Jaillon, O.; Montanini, B.; Morin, E.; Noel, B.; Percudani, R.; et al. Périgord black truffle genome uncovers evolutionary origins and mechanisms of symbiosis. *Nature* **2010**, *464*, 1033. [[CrossRef](#)] [[PubMed](#)]
74. Spanu, P.D.; Abbott, J.C.; Amselem, J.; Burgis, T.A.; Soanes, D.M.; Stuber, K.; Ver Loren van Themaat, E.; Brown, J.K.; Butcher, S.A.; Gurr, S.J.; et al. Genome expansion and gene loss in powdery mildew fungi reveal tradeoffs in extreme parasitism. *Science* **2010**, *330*, 1543–1546. [[CrossRef](#)] [[PubMed](#)]
75. Xia, E.H.; Yang, D.R.; Jiang, J.J.; Zhang, Q.J.; Liu, Y.; Liu, Y.L.; Zhang, Y.; Zhang, H.B.; Shi, C.; Tong, Y.; et al. The caterpillar fungus, *Ophiocordyceps sinensis*, genome provides insights into highland adaptation of fungal pathogenicity. *Sci. Rep.* **2017**, *7*, 1806. [[CrossRef](#)] [[PubMed](#)]
76. Faino, L.; Seidl, M.F.; Shi-Kunne, X.; Pauper, M.; van den Berg, G.C.; Wittenberg, A.H.; Thomma, B.P. Transposons passively and actively contribute to evolution of the two-speed genome of a fungal pathogen. *Genome Res.* **2016**, *26*, 1091–1100. [[CrossRef](#)]
77. Raffaele, S.; Kamoun, S. Genome evolution in filamentous plant pathogens: why bigger can be better. *Nat. Rev. Microbiol.* **2012**, *10*, 417. [[CrossRef](#)]
78. Sung, G.H.; Poinar, G.J.; Spatafora, J.W. The oldest fossil evidence of animal parasitism by fungi supports a Cretaceous diversification of fungal-arthropod symbioses. *Mol. Phylogenet. Evol.* **2008**, *49*, 495–502. [[CrossRef](#)]
79. Druzhinina, I.S.; Chenthamara, K.; Zhang, J.; Atanasova, L.; Yang, D.; Miao, Y.; Rahimi, M.J.; Grujic, M.; Cai, F.; Pourmehdi, S.; et al. Massive lateral transfer of genes encoding plant cell wall-degrading enzymes to the mycoparasitic fungus *Trichoderma* from its plant-associated hosts. *PLoS Genet.* **2018**, *14*, e1007322. [[CrossRef](#)]
80. Karlsson, M.; Atanasova, L.; Jensen, D.F.; Zeilinger, S. Necrotrophic mycoparasites and their genomes. *MicrobiolSpectrum* **2017**, *5*.
81. Xu, Y.; Zhu, Y.; Wang, Y.; Chang, Y.F.; Zhang, Y.; Jiang, X.; Zhuang, X.; Zhu, Y.; Zhang, J.; Zeng, L.; et al. Whole genome sequencing revealed host adaptation-focused genomic plasticity of pathogenic *Leptospira*. *Sci. Rep.* **2016**, *6*, 2002. [[CrossRef](#)] [[PubMed](#)]

82. Baxter, L.; Tripathy, S.; Ishaque, N.; Boot, N.; Cabral, A.; Kemen, E.; Thines, M.; Ah-Fong, A.; Anderson, R.; Badejoko, W.; et al. Signatures of adaptation to obligate biotrophy in the *Hyaloperonospora arabidopsidis* genome. *Science* **2010**, *330*, 1549–1551. [[CrossRef](#)]
83. Aguileta, G.; Refrégier, G.; Yockteng, R.; Fournier, E.; Giraud, T. Rapidly evolving genes in pathogens: Methods for detecting positive selection and examples among fungi, bacteria, viruses and protists. *Infect. Genet. Evol.* **2009**, *9*, 656–670. [[CrossRef](#)]
84. Schweizer, G.; Münch, K.; Mannhaupt, G.; Schirawski, J.; Kahmann, R.; Dutheil, J.Y. Positively selected effector genes and their contribution to virulence in the smut fungus *Sporisorium reilianum*. *Genome Biol. Evol.* **2018**, *10*, 629–645.
85. Seidl, V.; Song, L.; Lindquist, E.; Gruber, S.; Koptchinskiy, A.; Zeilinger, S.; Schmoll, M.; Martínez, P.; Sun, J.; Grigoriev, I.; et al. Transcriptomic response of the mycoparasitic fungus *Trichoderma atroviride* to the presence of a fungal prey. *BMC Genom.* **2009**, *10*, 567. [[CrossRef](#)]
86. Lin, X.; Heitman, J. Mechanisms of homothallism in Fungi and transitions between heterothallism and homothallism. In *Sex in Fungi*; Heitman, J., Ed.; American Society of Microbiology Press: Washington, DC, USA, 2007.
87. Seidl, V.; Seibel, C.; Kubicek, C.P.; Schmoll, M. Sexual development in the industrial workhorse *Trichoderma reesei*. *Proc. Nat. L Acad. Sci. USA* **2009**, *106*, 1390–1391. [[CrossRef](#)]
88. Ni, M.; Feretzaki, M.; Sun, S.; Wang, X.; Heitman, J. Sex in Fungi. *Annu. Rev. Genet.* **2011**, *45*, 405–430. [[CrossRef](#)]
89. Dodds, P.N.; Rafiqi, M.; Gan, P.H.P.; Hardham, A.R.; Jones, D.A.; Ellis, J.G. Effectors of biotrophic fungi and oomycetes: pathogenicity factors and triggers of host resistance. *New Phytol.* **2009**, *183*, 993–1000. [[CrossRef](#)]
90. Levin, E.; Kishore, A.; Ballester, A.R.; Raphael, G.; Feigenberg, O.; Liu, Y.; Norelli, J.; Gonzalez-Candelas, L.; Wisniewski, M.; Droby, S. Identification of pathogenicity-related genes and the role of a subtilisin-related peptidase S8 (PePRT) in autophagy and virulence of *Penicillium expansum* on apples. *Postharvest Biol. Tec.* **2019**, *149*, 209–220. [[CrossRef](#)]
91. Huang, F.; Liu, W.; Xie, B.-B.; Sun, C.-Y.; Li, D.; Rong, J.-C.; Chen, L.-L.; Shi, M.; Qin, Q.-L.; Wang, X.-W.; et al. Comparative genomics provide insights into evolution of *Trichoderma* nutrition style. *Genome Biol. Evol.* **2014**, *6*, 379–390.



© 2019 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 (<http://creativecommons.org/licenses/by/4.0/>).