



Article

# Genomic Insights into the Fungal Lignocellulolytic Machinery of *Flammulina rossica*

Young-Jin Park <sup>1,\*</sup> , Chang-Soo Lee <sup>1</sup> and Won-Sik Kong <sup>2</sup>

<sup>1</sup> Department of Biomedical Chemistry, Research Institute for Biomedical & Health Science, College of Biomedical and Health Science, Konkuk University, 268 Chungwon-daero, Chungju-si 27478, Korea; cslee@kku.ac.kr

<sup>2</sup> Mushroom Research Division, National Institute of Horticultural and Herbal Science, Rural Development Administration, 92, Bisan-ro, Eumseong-gun 27709, Korea; wskong@korea.kr

\* Correspondence: yjpark@kku.ac.kr; Tel.: +82-43-840-3601

Received: 20 August 2019; Accepted: 28 September 2019; Published: 8 October 2019



**Abstract:** Next-generation sequencing (NGS) of the *Flammulina rossica* (wood-rotting basidiomycete) genome was performed to identify its carbohydrate-active enzymes (CAZymes). De novo genome assembly (31 kmer) revealed a total length of 35,646,506 bp (49.79% GC content). In total, 12,588 gene models of *F. rossica* were predicted using an ab initio gene prediction tool (AUGUSTUS). Orthologous analysis with other fungal species revealed that 7433 groups contained at least one *F. rossica* gene. Additionally, 12,033 (95.6%) of 12,588 genes for *F. rossica* proteins had orthologs among the Dikarya, and *F. rossica* contained 12 species-specific genes. CAZyme annotation in the *F. rossica* genome revealed 511 genes predicted to encode CAZymes including 102 auxiliary activities, 236 glycoside hydrolases, 94 glycosyltransferases, 19 polysaccharide lyases, 56 carbohydrate esterases, and 21 carbohydrate binding-modules. Among the 511 genes, several genes were predicted to simultaneously encode two different CAZymes such as glycoside hydrolases (GH) as well as carbohydrate-binding module (CBM). The genome information of *F. rossica* offers opportunities to understand the wood-degrading machinery of this fungus and will be useful for biotechnological and industrial applications.

**Keywords:** basidiomycete; carbohydrate active enzyme; *Flammulina rossica*; fungi; genome

## 1. Introduction

*Flammulina rossica* (edible mushroom, *Physalacriaceae*) was first identified in 1999 by Redhead and Petersen [1]. According to a previous report, *F. rossica* is one of the recently described *Flammulina* species of the Northern Hemisphere, which include *Flammulina elastica*, *Flammulina fennae*, *Flammulina ononidis*, and *Flammulina velutipes* [2]. Redhead and Petersen [1] reported that the basidiocarps of *F. rossica* are similar to those of *F. velutipes*, although they possess a very pale pileus—whitish to yellowish ochraceous. Based on their ribosomal ITS sequences, *F. rossica* was found in a large clade with *Flammulina mexicana*, *Flammulina populicola*, and *F. fennae* but not with *F. elastica* [2]. *F. rossica* is found on the trunks of *Alnus* sp., *Populus* sp., *Salix amygdaloides*, *Salix caprea*, and *Salix* sp., from July to January [3].

Enzymes involved in the synthesis and breakdown of glycoconjugates, oligosaccharides, and polysaccharides are known as carbohydrate-active enzymes (CAZymes). CAZymes are divided into 6 classes including glycosyltransferases (GTs), carbohydrate esterases (CEs), glycoside hydrolases (GHs), polysaccharide lyases (PLs), auxiliary activities (AA), and carbohydrate-binding modules (CBMs). These CAZymes are further classified into families, based on their structural and amino acid sequence similarities [4]. CAZymes are widely recognized as one of the keys to biofuel production because they play a significant role in plant cell wall degradation. Thus, CAZymes are attracting attention as an area

of biotechnological and industrial application [5,6]. Basidiomycetes are capable of efficiently degrading lignocellulosic biomass, especially plant-derived lignocellulosic biomass, due to various CAZymes [5,7]. This ability allows the fungus to live in a variety of natural conditions such as wood wastes and forest residues. Wood rot fungi are usually divided into brown rot fungus and white rot fungus. In particular, white rot fungi, which account for more than 90% of wood-rotting basidiomycetes, decompose both polysaccharides and lignin, leaving the residue white or yellowish. [5,6].

We previously performed genome-sequencing of *F. velutipes*, *F. elastica*, *F. fennae*, and *F. ononidis* and reported well-developed wood-degrading machineries in their genomes based on CAZyme identification [8–11]. Similarly, genome-sequencing studies of various organisms have been performed in order to massively excavate genes that encode biomass-degrading enzymes [5,12,13]. Moreover, research on biomass-degrading enzymes in the post-genome era is a major research field for understanding wood-degrading machinery and to describe the CAZyme repertoire of fungal species.

This study aimed to determine the genome sequence of *F. rossica* and to identify biotechnologically and industrially useful CAZyme genes. The genomic information of *F. rossica*, including the genes encoding CAZymes, will help understand this fungus and will be useful for future biotechnological and industrial applications.

## 2. Materials and Methods

### 2.1. Fungal Strain, Culture, and Genomic DNA Isolation

*F. rossica* ASI4194 was obtained from the Mushroom Research Division, National Institute of Horticultural and Herbal Science (Rural Development Administration, Eumseong-gun, Korea) and was grown at 26 °C on potato dextrose agar (PDA) for 15 days. For genomic DNA isolation, extraction buffer (100 mM NaCl, 50 mM EDTA, 0.25 M Tris-HCl, 5% SDS), 2 × CTAB buffer (2% Cetyltrimethylammonium bromide, 1.4 M NaCl, 20 mM EDTA pH 8.0, 100 mM Tris-HCl pH 8.0, 1% polyvinyl pyrrolidone), and phenol:chloroform:isoamyl alcohol (25:24:1) were added to the mycelia and mixed vigorously. Sample was centrifuged at 12,000 rpm at 4 °C for 5 min. Supernatant was mixed with 0.7 volume of isopropanol and then centrifuged for 15 min. The pellet was washed in cold 70% then dried pellet was dissolved in TE buffer containing RNase A (Qiagen, Seoul, Korea).

### 2.2. Genome Sequencing, De Novo Assembly, Gene Prediction, and Annotation

Next-generation sequencing (NGS) of the *F. rossica* genome was performed using the HiSeq 2000 platform (Illumina, Inc., San Diego, CA, USA) according to the manufacturer's protocol. The quality of sequencing data was evaluated using FastQC [14] and was processed using Trimmomatic (version 0.32) [15] to detect sequencing adapters and bad quality reads. Quality checked reads were used for assembly using Velvet software [16]. Gene structure modeling was processed using the AUGUSTUS software [17], trained with *Laccaria bicolor*. The predicted genes of *F. rossica* were then compared with a non-redundant database (National Center for Biotechnology Information, NCBI, Bethesda, MD, USA) using DIAMOND software [18] for functional annotation.

### 2.3. Ortholog Analysis

*F. rossica* genes were analyzed and clustered into orthologous groups using OrthoFinder (version 2.3.3) software [19] with other fungal species. *A. nidulans* FGSC-A4 [20], *Botrytis cinerea* B05.10 [21], *Agaricus bisporus* var. *bisporus* H97 [22], *C. cinerea* okayama7#130 [23], *L. bicolor* S238N-H82 [24], *Lentinula edodes* [25], *C. militaris* CM01 [26], *C. neoformans* var. *grubii* H99 [27], *F. elastica* KACC46182 [9], *F. fennae* KACC46185 [10], *F. ononidis* KACC46186 [11], *F. velutipes* KACC42780 [8], *Phanerochaete chrysosporium* RP78 [12], *S. cerevisiae* S288C [28], *Neurospora crassa* OR74A [29], *S. commune* H4-8 [30], and *T. reesei* QM6a [31].

#### 2.4. CAZyme and Signal Peptide Identification

CAZymes, including those encoded by *GH*, *GT*, *PL*, *CE*, *AA*, and *CBM* genes in *F. rossica* were identified and annotated using the dbCAN meta server including the HMMER (dbCAN CAZyme domain HMM database), DIAMOND (CAZy database), and Hotpep (short conserved motifs in the PRR library database) [32]. Signal peptide prediction in CAZyme genes was carried out using the SignalP 5.0 software [33].

#### 2.5. Data Access

The raw reads were deposited in the Sequence Read Archive (SRA) database at NCBI (SRR9964086).

### 3. Results and Discussion

#### 3.1. De Novo Genome Assembly, Gene Prediction, and Genome Comparisons

The short reads (total of 38,390,380; 100 bp paired-end reads) were analyzed using the Trimmomatic tool [14] for quality control including adapter trimming. The resulting short reads (35,908,618 reads, >Q30) were processed for genome assembly using the Velvet assembly software (kmer-size of 17–31) [16]. The optimized assembly (31 kmer) comprised 15,546 sequence contigs with a total length of 35,645,506 bp (49.77% GC contents) and an N50 length of 48,718 bp. In ab initio gene prediction, 12,588 genes were predicted. The average gene, exon, and intron lengths were 1911, 234.67, and 68.03 bp, respectively. The optimized assembly and gene prediction of the *F. rossica* ASI4194 genome are presented in Table 1.

**Table 1.** *Flammulina rossica* Genome Sequencing Statistics.

Hiseq 2000 NGS Analysis	Total reads (100 bp)	38,390,380
	Reads after trimming (%), >Q30	35,908,618 (93.53)
Velvet De Novo Assembly	Optimized Velvet hash value (kmer)	31
	Total number of contigs	15,546
	Number of contigs (>1 kb)	1843
	Contig N50 (bp)	48,718
	Length of longest contig (bp)	418,162
	Total bases in contigs (bp)	35,645,506
	Total bases in contigs (>1 kb)	33,499,878
	GC content (%)	49.79
Gene Prediction	Predicted gene	12,588
	Average gene length (bp)	1911
	Average protein length (aa)	510.29
	Average exon per gene	6.52
	Average exon size (bp)	234.67
	Average intron size (bp)	68.03

Of the 12,588 predicted genes, 83.3% (10,490) had sequence similarity ( $0.001 > e\text{-value}$ ) with the proteins in NCBI-NR (Table S1). The total number of genes in *F. rossica* was comparable to that of its nearest sequenced species, *F. elastica* [9], as well as to that of other basidiomycetes with a similar genome size (Table 2).

Through cluster analysis with other fungal species, 7485 (58.7%) out of 12,756 groups containing at least one *F. rossica* gene were identified (Table 3). In addition, 95.5% of *F. rossica* genes were conserved in the Dikarya including ascomycetes and basidiomycetes (Table 3). Among the set of homologous genes, there were 5 species-specific orthogroups containing 12 species-specific genes in *F. rossica* (Table 3). As shown in Figure 1, *F. rossica* was classified into one group with *F. elastica* and was clustered into one group together with *F. fennae*, *F. onnidis*, and *F. velutipes* by an ortholog-based clustering analysis.

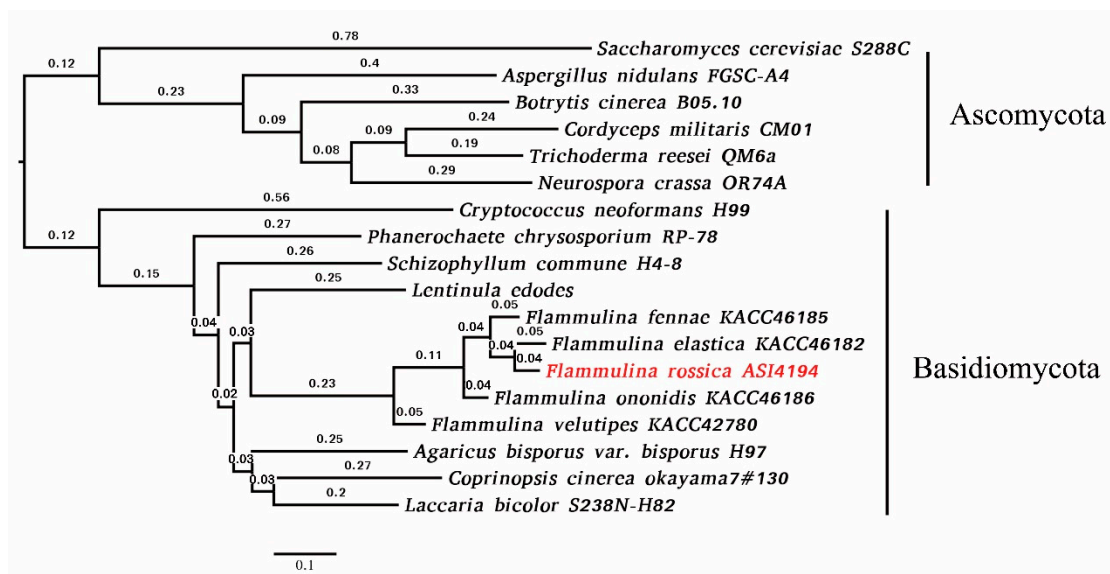
**Table 2.** Comparison of the Genome Characteristics of *Flammulina rossica* and other Basidiomycetes.

<b>Fungal Species</b>	<i>Flammulina rossica</i>	<i>F. ononidis</i>	<i>F. fennae</i>	<i>F. elastica</i>	<i>F. velutipes</i>	<i>Laccaria bicolar</i>	<i>Cryptococcus cinerea</i>	<i>Phanerochaete chrysosporium</i>	<i>Schizophyllum commune</i>
<b>Strain</b>	ASI4194	KACC46186	KACC46185	KACC46182	KACC42780	S238N-H82	Okayama7#130	RP78	H4-8
<b>Genome (Mb)</b>	35.6	34.5	32.4	35	35.6	64.9	37.5	35.1	38.5
<b>Genes</b>	12,588	12,269	11,591	12,536	12,218	20,614	13,544	10,048	13,181
<b>GC (%)</b>	49.79	49.76	39	49.7	48.99	46.6	51.6	53.2	56.6
<b>Average Gene Length (bp)</b>	1911	2009	1980	1973	2294	1533.0	1679.0	1667.0	1794.9
<b>Average Exon Size (bp)</b>	234.67	234.09	230.53	233.91	231.4	210.1	251.0	232.0	249.3
<b>Average Intron Size (bp)</b>	68.03	68.14	68.93	69.29	190.3	92.7	75.0	117.0	79.0

**Table 3.** Ortholog analysis of *Flammulina rossica* and other fungal species.

Fungal Species	Genes	Genes in Orthogroups (%)	Unassigned Genes (%)	Orthogroups Containing Species (%)	Number of Species-Specific Orthogroups	Genes in Species-Specific Orthogroups
<i>Flammulina rossica</i> ASI4194 *	12,588	12,024 (95.5)	564 (4.5)	7485 (58.7)	5	12
<i>Flammulina fennae</i> KACC46185 *	11,591	11,235 (96.9)	356 (3.1)	7307 (57.3)	2	4
<i>Flammulina ononidis</i> KACC46186 *	12,269	11,881 (96.8)	388 (3.2)	7572 (59.4)	1	2
<i>Flammulina elastica</i> KACC46182 *	12,536	11,987 (95.6)	549 (4.4)	7592 (59.5)	1	2
<i>Flammulina velutipes</i> KACC42780 *	12,218	10,774 (88.2)	1444 (11.8)	6679 (52.4)	4	10
<i>Agaricus bisporus</i> var. <i>bisporus</i> H97 *	10,448	9102 (87.1)	1346 (12.9)	5712 (44.8)	29	265
<i>Coprinopsis cinerea</i> okayama 7#130 *	13,356	10,837 (81.1)	2519 (18.9)	6393 (50.1)	59	329
<i>Cryptococcus neoformans</i> var. <i>grubii</i> H99 *	7826	6477 (82.8)	1349 (17.2)	4719 (37.0)	39	129
<i>Laccaria bicolor</i> S238N-H82 *	18,215	12,662 (69.5)	5553 (30.5)	6217 (48.7)	73	524
<i>Lentinula edodes</i> *	14,079	11,678 (82.9)	2401 (17.1)	6674 (52.3)	61	340
<i>Phanerochaete chrysosporium</i> RP78 *	13,602	10,443 (76.8)	3159 (23.2)	6223 (48.8)	36	265
<i>Schizophyllum commune</i> H4-8 *	13,194	10,976 (83.2)	2218 (16.8)	6214 (48.7)	59	437
<i>Aspergillus nidulans</i> FGSC-A4 **	9561	8433 (88.2)	1128 (11.8)	5709 (44.8)	4	24
<i>Botrytis cinerea</i> B05.10 **	16,389	9336 (57.0)	7053 (43.0)	6263 (49.1)	17	64
<i>Cordyceps militaris</i> CM01 **	9651	8369 (86.7)	1282 (13.3)	6131 (48.1)	5	13
<i>Neurospora crassa</i> OR74A **	10,812	8659 (80.1)	2153 (19.9)	6217 (48.7)	10	26
<i>Saccharomyces cerevisiae</i> S288C **	6002	4624 (77.0)	1378 (23.0)	3485 (27.3)	12	33
<i>Trichoderma reesei</i> QM6a **	9115	8332 (91.4)	783 (8.6)	6283 (49.3)	4	9

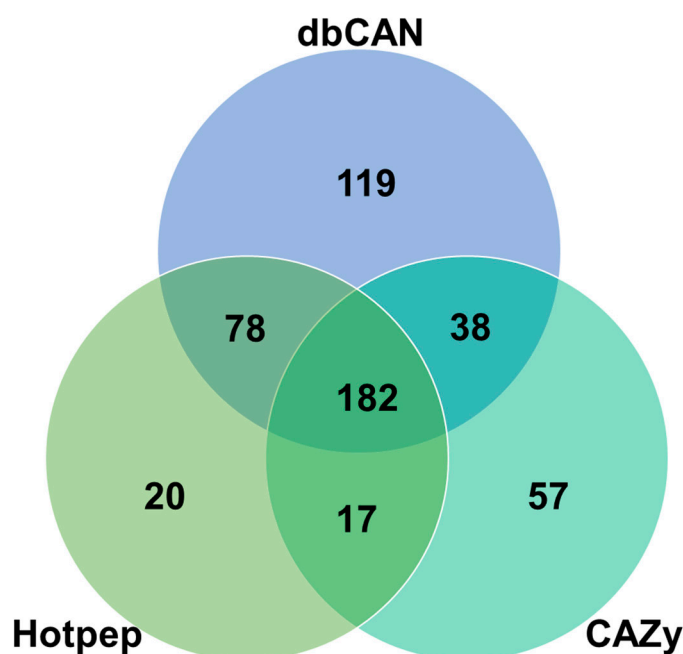
\* Basidiomycota, \*\* Ascomycota.



**Figure 1.** Phylogenetic tree of fungal species based on ortholog clustering.

### 3.2. *F. rossica* CAZymes and Genome-Wide Comparisons with Other Fungal Species

The genome sequence of *F. rossica* revealed a series of genes associated with the breakdown (GHs, PLs, CEs) and assembly (GT) of carbohydrate complexes. The *F. rossica* genome was also found to contain several genes encoding lignin degradation enzymes (auxiliary activities; AA) as well as a carbohydrate-binding modules (CBM). CAZyme prediction of *F. rossica* genes through the dbCAN meta server [32] including the HMMER (dbCAN CAZyme domain HMM database), Hotpep (short conserved motifs in the PRR library database), and DIAMOND (CAZy database) revealed 419, 300, and 294 CAZymes, respectively (Figure 2 and Table 4).



**Figure 2.** Identification and annotation of carbohydrate-active enzyme genes in the *F. rossica* genome based on three different databases including the HMMER (dbCAN carbohydrate-active enzymes domain HMM database), DIAMOND (CAZy database), and Hotpep (short conserved motifs in the PRR library database) [32].

**Table 4.** CAZymes of *Flammulina rossica* and other fungal species.

Fungal Species	CAZymes						No. of CAZyme (Annotation DB)	Total	Reference
	AA <sup>1</sup>	GH <sup>2</sup>	GT <sup>3</sup>	CE <sup>4</sup>	CBM <sup>5</sup>	PL <sup>6</sup>			
<i>Flammulina rossica</i> *	93	182	64	56	7	17	419 (Hmmer dbCAN)	528	This study
	61	145	53	21	5	15	300 (Hotpep)		
	39	155	72	11	11	6	294 (CAZy database)		
<i>Flammulina fenae</i> *	86	220	85	57	45	20	-	513	[10]
<i>Flammulina ononidis</i> *	87	228	87	61	40	21	-	524	[11]
<i>Flammulina elastica</i> *	82	218	89	59	42	18	-	508	[9]
<i>Flammulina velutipes</i> *	85	239	84	63	44	25	-	540	[9]
<i>Agaricus bisporus</i> *	81	174	54	33	44	9	-	395	JGI database
<i>Cryptococcus cinerea</i> *	111	195	83	60	105	16	-	570	[9]
<i>Laccaria bicolor</i> *	55	170	96	18	31	7	-	377	JGI database
<i>Lentinula edodes</i> *	82	254	85	44	61	11	-	537	[9]
<i>Phanerochaete chrysosporium</i> *	85	175	65	16	62	4	-	407	JGI database
<i>Schizophyllum commune</i> *	78	241	85	57	37	18	-	516	[9]
<i>Cryptococcus neoformans</i> *	14	97	70	5	12	4	-	202	CAZy database
<i>Cordyceps militaris</i> **	54	165	91	34	39	5	-	388	[9]
<i>Trichoderma reesei</i> **	59	210	90	32	44	5	-	440	[9]
<i>Saccharomyces cerevisiae</i> **	5	57	68	2	12	0	-	144	CAZy database
<i>Aspergillus nidulans</i> **	33	267	91	30	46	23	-	490	CAZy database
<i>Neurospora crassa</i> **	35	177	80	21	42	4	-	359	CAZy database
<i>Botrytis cinerea</i> **	77	287	119	37	89	10	-	619	CAZy database

\* Basidiomycota, \*\* Ascomycota, <sup>1</sup> auxiliary activities (AA), <sup>2</sup> glycoside hydrolase (GH), <sup>3</sup> glycosyltransferase (GT), <sup>4</sup> carbohydrate esterase (CE), <sup>5</sup> carbohydrate-binding module (CBM), <sup>6</sup> polysaccharide lyase (PL).

Among the 511 genes associated with CAZymes, several genes were predicted to simultaneously encode two different CAZymes, such as GH as well as CBM. Therefore, in total, 528 CAZymes including 102 AAs, 236 GHs, 94 GTs, 19 PLs, 56 CEs, and 21 CBMs were identified in the *F. rossica* genome (Table 4 and Table S2). For genome-wide comparison, the annotated CAZymes of eight other fungal species were obtained from the CAZy database [4] and JGI database (<https://genome.jgi.doe.gov/programs/fungi/index.jsf>).

### 3.2.1. Glycosyltransferases (GTs) of *F. rossica* Genome

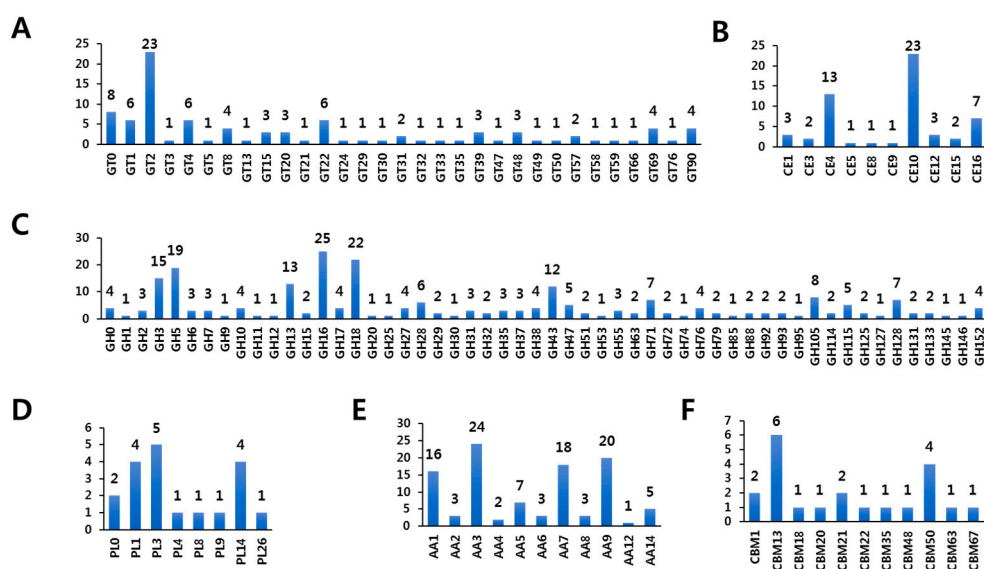
GTs (EC 2.4.x.y) catalyze glycosyl group transfer and glycosidic linkage formation using activated donor sugar phosphates [34–37], which are involved in the biosynthesis of glycoconjugates, oligosaccharides, and polysaccharides [34–37]. CAZyme prediction revealed that *F. rossica* contains a total of 23 GT families in its genome sequence based on the dbCAN meta server search (Figure 3A and Table S3). Among the 94 GTs, 2, 1, and 4 genes predicted to encode GTs were identified uniquely based on the HMMER (dbCAN CAZyme domain HMM database), Hotpep (short conserved motifs in the PRR library database), and DIAMOND (CAZy database) searches, respectively (Table S3). Moreover, the GT2 family was prominent with 23 genes in the *F. rossica* genome (Figure 3A and Table S3). Complete genome sequences of various organisms including bacterial, archaeal, or eukaryotic organisms, reveal that a large number of GTs (about 1–2% of the total number of genes) are present in their genomes (CAZy database) [4]. Genome-wide comparisons also showed a number of genes encoding the GT2 family, suggesting that this family is a major component among GT families in most fungal species (Figure 4A and Table S3). Breton et al. [34] demonstrated that incorporation of newly discovered GT genes would increase the number of families and that not all sequences of GT were present in the public database. About 50% of the total number of GTs in the database is GT2 and GT4 families. At the time of writing (July 2019), the database comprised more than 550,978 classified and 11,654 non-classified GT sequences divided into 107 families (CAZy database). More than 170,000 sequences from various organisms were classified into the GT2 family in databases [4]. Signal peptide prediction revealed six genes comprising the signal peptides in 94 GT genes in *F. rossica* (Table S2). Signal sequence prediction of GT revealed 6 genes possessing the signal peptides in their amino acid sequences (Table S2). Most GTs are resident membrane proteins in Golgi apparatus and the endoplasmic reticulum. All GT proteins have a short N-terminal cytoplasmic tail, large C-terminal catalytic domains, and a signal-anchor domain [38]. Signal-anchor domains act as uncleavable signal peptides [39]. Thus, the predicted signal

peptides in six genes likely act as signal-anchor domains. In this study, 8 genes were annotated as the GT0 family (not yet assigned to a family) in the *F. rossica* genome. GT families were defined based on significant amino acid sequence similarities [35,36]. However, previous studies have described difficulties in classifying GTs based on sequence similarity, because many GTs have divergent activities, even though their sequences are highly similar. Therefore, additional studies based on structural and mutational analyses are needed to elucidate their enzymatic characteristics.

### 3.2.2. Carbohydrate Esterases (CEs) of the *F. rossica* Genome

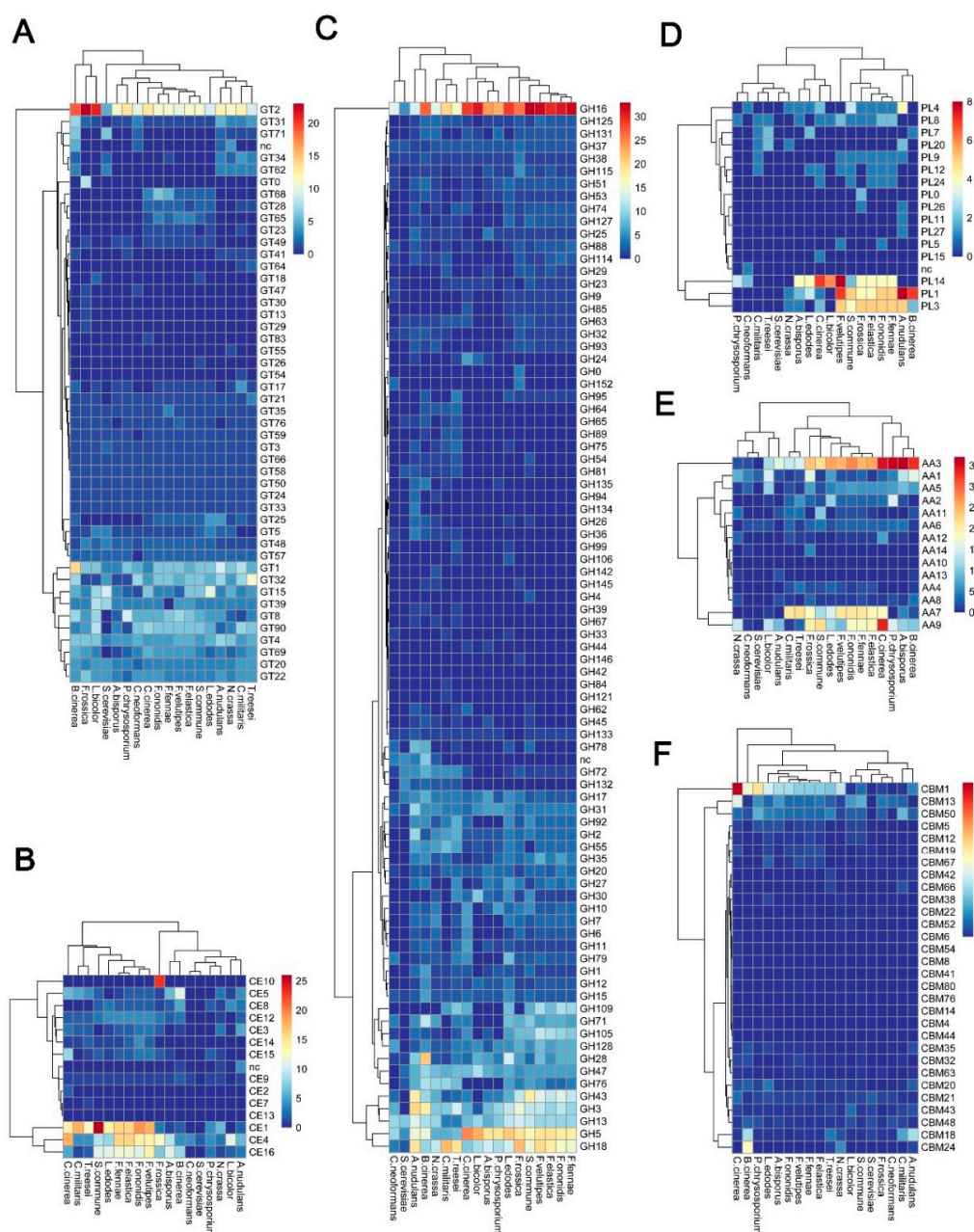
Esterases act on ester bonds and are widely used as biocatalysts in biotechnology and industrial processes. [40,41]. CE represents a family of esters that generally catalyze *N*-deacylation or *O*-deacylation to remove the esters of substituted saccharides [42]. These CEs are classified into 16 families, with more than 67,000 classified (1200 non-classified) CEs in the current CAZy database (CAZy database) [4]. CEs have a variety of substrate specificities, such as specificity for acetic ester, chitin, xylan, feruloyl-polysaccharide, pectin, and peptidoglycan [43].

Our results revealed a total of 56 predicted CEs classified into 10 families in the *F. rossica* genome based on the HMMER (dbCAN CAZyme domain HMM database), Hotpep (short conserved motifs in the PRR library database), and DIAMOND (CAZy database) searches (Figure 3B and Table S3). The CE10 family was prominent with 23 CEs, and the CE4 family was the second largest family with 13 CEs in the *F. rossica* genome (Figure 3B). Genome-wide comparisons showed that the total number of CEs in *F. rossica* was similar to that found in *Flammulina* species, *Coprinopsis cinerea*, and *Schizophyllum commune*. Additionally, CE1, CE4, and CE16 families were also prominent in other basidiomycetes. However, since only 5 CE (4 CE4 and 1 CE9) and 2 CE (CE4) were found in *Cryptococcus neoformans* and *Saccharomyces cerevisiae*, respectively, the CE families were found to vary in fungal species (Figure 4B and Table S3). CAZyme prediction based on three different databases also revealed many CE10 family members in the *F. rossica* genome. However, most members of the CE10 family are reported to act on non-carbohydrate substrates [4,44].



**Figure 3.** Number of CAZymes in *F. rossica*. Number of (A) GT families; (B) CE families; (C) GH families; (D) PL families; (E) AA families; (F) CBM families. AA, Auxiliary Activities; GH, glycoside hydrolase; GT, glycosyltransferase; CBM, carbohydrates-binding module; PL, polysaccharide lyase.





**Figure 4.** Distribution of CAZymes in *F. rossica* and other fungal species. (A) GT families; (B) CE families; (C) GH families; (D) PL families; (E) AA families; (F) CBM families.

Despite the large number of enzymes recently classified as CE, only a small number of members from the CE family have been biochemically and structurally analyzed, and some features of the amino acid sequences have been reported. For instance, CE families including CE1, CE4, CE5, and CE7, have been characterized as possessing the Ser-His-Asp and Gly-Xaa-Ser-Xaa-Gly conserved motifs in their amino acid sequences. In addition, CE2 and CE3 family members have the Gly-Asp-Ser-(Leu) motif rather than the Gly-Xaa-Ser-Xaa-Gly conserved motif. CE16 family members also possess the Gly-Asp-Ser-(Leu) and Ser-Gly-Asn-His motif in their amino acid sequences [45]. In the present study, we identified several CE family members containing the GX SXG conserved motifs in their amino acid sequences (Table S4). Esterase containing a Gly-Xaa-Xaa-Leu (GXXL) motif highly homologous to Class C  $\beta$ -lactamases was also identified [46,47]. Likewise, some CE family members were found to possess the (GXXL) motif (Table S4). Furthermore, in the present study, all genes predicted to encode CE4 family members were found to have conserved motifs such as Phe-Asp-Asp-Gly-Pro (FDDGP),

in their amino acid sequences (Table S4). CE families generally catalyze *N*-deacylation or *O*-deacylation reactions of polysaccharides to promote degradation by GHs and assist biomass saccharification [48]. Therefore, an extensive range of genes encoding CE family members in the *F. rossica* genome suggests the potential for this fungus to be used in industrial applications such as biofuel production.

### 3.2.3. Glycoside Hydrolases (GHs) of *F. rossica* Genome

GHs (glycosidases or glycosyl hydrolases, EC 3.2.1.-) are key enzymes involved in carbohydrate metabolism, which catalyze the hydrolysis of glycosidic bonds in complex carbohydrates. GHs are also common enzymes that degrade the most abundant biomass such as hemicellulose, cellulose, and starch [49,50]. GHs can be assigned to various families based on their sequence similarities.

Up to now (July 2019), the CAZy database comprised more than 664,000 classified and 10,000 non-classified GH sequences that were divided into 165 families (CAZy database) [4]. In the present study, a total of 236 GHs classified into 54 families were predicted in the *F. rossica* genome based on based on three different database searches (dbCAN, Hotpep, CAZy databases) (Figure 3C and Table S3). GH family classification also revealed that the GH16 family was prominent with 25 genes in the *F. rossica* genome (Figure 3C and Table S3). In addition, many GH16 family members were also identified in other fungal species except for some ascomycetes, including *Cordyceps militaris*, *Aspergillus nidulans*, *Trichoderma reesei*, and *S. cerevisiae* (Figure 4C and Table S3). Additionally, multiple copies of GH5 and GH18 in *F. rossica* were similar to those in other basidiomycetes.

The GH16 family consists of agarase (EC 3.2.1.81), lichenase (EC 3.2.1.73),  $\kappa$ -carrageenase (EC 3.2.1.83), xyloglucan xyloglucosyltransferase (EC 2.4.1.207), endo- $\beta$ -1,3-glucanase (EC 3.2.1.39), endo- $\beta$ -1,3-1,4-glucanase (EC 3.2.1.6), and endo- $\beta$ -galactosidase (EC 3.2.1.103), and most of these enzymes contain conserved motifs such as Glu-Xaa-Asp-Xaa-(Xaa)-Glu (EXDX[X]E). The first and last glutamic acid (E) residues of the conserved motif are characterized as a nucleophile and Brønsted acid/base, respectively, and play an important role in the catalytic activity of GH16 family enzymes [51–53]. All of the predicted GH16 family members in *F. rossica* showed this conserved motif, and five genes were predicted to encode GH16 contained 2 or more of the conserved motif (EXDX[X]E) (Table S4). Although not all glycosyl hydrolases have signal sequences, many GHs have signal sequences that are secreted or targeted to other cell sites including the periplasmic space or Golgi apparatus [54]. In the present study, *F. rossica* was shown to contain signal peptides in about 50% of the GH genes (98 out of 236 GH), suggesting that these GHs can be secreted (Table S2).

GHs are essential for the processing of cellulose and xylan in plants, as well as chitins in nature (CAZy database) [4]. Cellulases (GH5, -6, -7, -8, -9, -12, -44, -45, and -48), xylanases (GH10, -11, and -30), and chitinases (GH18, -19, and -85) are known to be active against cellulose, xylose, and chitin, respectively [4,55]. *F. rossica* also contains a series of genes related to cellulase (27 genes), xylanase (6 genes), and chitinase (23 genes) in its genome sequence (Figure 3C and Table S3). Cellobiose can be converted to glucose by  $\beta$ -glucosidases (EC 3.2.1.21) involved in GH families such as GH1 and GH3 [4,56]. *F. rossica* also possesses GH family members including 1 GH1 and 15 GH3 in its genome (Figure 3C and Table S3).

Simultaneous actions of several GHs are necessary to effectively degrade plant cell wall complexes composed of cellulose and xylan. Recently, genome sequencing of various bacterial and fungal species has reviewed the various activities of GHs on cellulose, chitin, and xylan degradation and their potential for biotechnological applications and industrial degradation of biopolymers [55,57–61]. In the present study, *F. rossica* with more than 200 genes encoding various GHs showed strong potential for diverse applications, such as biotechnology and industry.

### 3.2.4. Polysaccharide Lyases (PLs) of the *F. rossica* Genome

PLs (EC 4.2.2.-) cleave polymer chains of polysaccharides, essential cellular components of all living organisms, through a  $\beta$ -elimination mechanism to produce unsaturated polysaccharides [62,63]. PLs have been classified into 36 families with more than 19,600 classified and 1000 non-classified PLs in

the database [4]. Our results showed that a total of 19 PLs classified into eight families were predicted in the *F. rossica* genome based on three different databases (dbCAN, Hotpep, CAZy database) searches (Figure 3D and Table S3). Among them, the PL3 family was prominent and four families, including PL4, -8, -9, and -26, consisted of only one PL (Figure 3D and Table S3). In addition, it has been found that other basidiomycetes had many PL14 family members in their genomes (Table S3). PL20 was only found in ascomycetes, whereas PL14 appeared to be specific to the Basidiomycota. Moreover, PL5, -15, and -24 family members are Basidiomycota specific; thus, the distribution of some PL family members was found to be phylum specific (Table S3) [64].

Pectate (partially branched polymers containing 1→4 linked  $\alpha$ -D-galacturonate) or pectin (partially branched polymers containing 1→4 linked  $\alpha$ -D-methylgalacturonate) is an acidic polysaccharide in plant cell walls [65–67]. The enzymes that degrade polygalacturonan (PGA) are known as pectate or pectin lyases [66]. Bacterial species mainly produce pectate lyase, whereas fungal species produce pectate lyases and pectin lyases (EC 4.2.2.10) simultaneously to act on pectin and/or pectate [67]. Pectin and/or pectate lyases have been classified into 6 PL families such as PL1, -2, -3, -9, and -10. Fungal pectate lyases (EC 4.2.2.2 and EC 4.2.2.9) were classified into PL family 1, 3, and 9 in the CAZy database [4]. The major family members of PLs were pectate lyases including PL families 1, 3, and 4 in *F. rossica* and other fungal species. However, according to this study, genes encoding PL families 2 and 10 were not found in *F. rossica* and in other fungal species genomes (Figure 3D and Table S3). Furthermore, PL family member 9 was found only in 5 *Flammulina* species, in *S. commune*, and in two ascomycetes (*A. nidulans* and *C. militaris*) but not in other fungal species (Figure 4D and Table S3).

Biochemical information on enzymes that degrade pectin or pectic bacteria in basidiomycetes is relatively scarce compared to those from other bacterial and fungal species, although genome sequencing reveals many genes encoding PL in several basidiomycetes. It is reported that *S. commune* produces high levels of pectinase and that polygalacturonase is produced at high levels, especially when cultivated in wheat bran [68]. Although no further studies were conducted in this study, there were similar numbers of genes encoding PL family members 1, 3, and 9 in the *F. rossica* genome as those in *S. commune* suggesting that *F. rossica* might be a candidate for biotechnological applications.

### 3.2.5. Auxiliary Activities (AAs) of *F. rossica* Genome

Ligninolytic enzymes such as lytic polysaccharide monooxygenases (LPMOs) are classified into AA families, which are mainly involved in the depolymerization of non-carbohydrate structural components (lignin) of plants [6]. These AAs are classified into 16 families with more than 12,200 classified and 40 non-classified AA sequences in the current CAZy database. In addition, the AAs are presently classified into 6 families of LPMOs and 9 families of ligninolytic enzymes (CAZy database) [4]. In the present study, we identified a total of 11 AA families with 102 AAs in *F. rossica* genome sequence (Figure 3E and Table S3). Among them, AA3 (glucose-methanol-choline oxidoreductase, alcohol oxidase, aryl-alcohol oxidase/glucose oxidase, cellobiose dehydrogenase, pyranose oxidase) and AA9 (lytic polysaccharide monooxygenase) are the major members with 24 AA3 and 20 AA9 in the *F. rossica* genome, respectively (Figure 3E and Table S3). Interestingly, the total number of AAs in the *F. rossica* genome was similar to that in *C. cinerea*, but not other *Flammulina* species (Figure 4E and Table S3). In addition, our results showed that *F. rossica* had more AA1 family members in its genome compared to other basidiomycetes (Figure 4E and Table S3).

Several AA families have been shown to possess conserved motifs necessary for their activities. A laccase (EC 1.10.3.2, AA1 family) has conserved motifs (copper binding motifs) within its amino acid sequence such as His-Xaa-His, His-Xaa-His-Gly, His-Xaa-Xaa-His-Xaa-His, and His-Cys-His-Xaa<sup>3</sup>-His-Xaa<sup>4</sup>-Met/Leu/Phe [69]. Our results also showed that 7 genes predicted to encode the AA1 family contained the copper-binding motifs in their amino acid sequences (Table S4). GMC oxidoreductase protein (AA3 family) require a flavin-adenine dinucleotide (FAD) cofactor and also has a  $\beta$ - $\alpha$ - $\beta$  dinucleotide binding-motif composed of Gly-Xaa-Gly-Xaa-Xaa-Gly-Xaa<sup>18</sup>-Glu that interacts with the FAD cofactor [70–72]. *F. rossica* was also found to contain 24 genes predicted to encode

AA3, and 17 genes contained the  $\beta$ - $\alpha$ - $\beta$  dinucleotide binding motif in their amino acid sequences (Table S4). These results indicate that the gene predicted to encode AA1 and AA3 with motifs may act as a laccase and a GMC oxidoreductase, respectively. Wood degradation by wood rotting fungi generally begins with the depolymerization of lignin, which results in further degradation of the wood polymer by highly reactive lignin radicals [73,74]. Therefore, the extensive range of enzymes belonging to AA families in the *F. rossica* genome observed in this study suggest the great potential for this fungus as a biomaterial and biofuel producer in the future.

### 3.2.6. Carbohydrate-Binding Modules (CBMs) of *F. rossica* Genome

Amino acid sequence within a carbohydrate-active enzyme with carbohydrate-binding activity is known as CBM [75,76]. CBM generally binds to carbohydrate ligands to facilitate the catalytic activity of CAZyme. [75]. CBMs are commonly found in GHs, PLs, and GTs [77]. Moreover, CBMs are parts of scaffoldin subunit in non-hydrolytic enzymes that organizes the catalytic subunits into a cellulosome (multi-protein complexes) [76]. Enzyme complexes containing CBM in CAZymes are capable of degrading the substrate more efficiently, but it is reported that when CBM is removed from the scaffolding of cellulosomes, catalytic efficiency is reduced [76].

CBMs have been classified into 85 families with more than 180,000 classified and 550 non-classified CBMs in the CAZy database [4]. In this study, we found that a total of 21 CBMs classified into 10 families were predicted in the *F. rossica* genome based on three different database (dbCAN, Hotpep, CAZy database) searches. CBM family 13 was prominent, and six families, including CBM18, -20, -22, -35, -48, -63, and -67, were represented by only one CBM in the *F. rossica* genome (Figure 3F and Table S3). Although multiple copies of CBM13 and -50 were similar to those in other fungal species, several CBM families including CBM5, -12, -19, -32, and -43 were not found, and relatively few members of the CBM1 family were found in the *F. rossica* genome.

Additionally, the abundance of some family members was different between basidiomycetes and ascomycetes. Although CBM family 18 members of ascomycetes were found more than those in other basidiomycetes, CBM 5 and 12 families were not found in ascomycetes as well as in the *F. rossica* genome (Figure 4F and Table S3). These results are consistent with a previous study in which ascomycetes are more abundant in CBM family 18 but have less CBM5 and -12 than basidiomycetes [64]. CBMs have traditionally been regarded as essential modules of cellulases, particularly cellobiohydrolases (family GH6 and -7) [78]. Although the GH6 or GH7 members that contain CBM1 are not found in the *F. rossica* genome, some CAZyme (4 GHs and 1 CE) genes that simultaneously encode CBM suggest that CBM is required for efficient substrate degradation (Table S2).

## 4. Conclusions

This study aimed to improve the understanding of lignocellulolytic machinery in *F. rossica* and to assess its applicability in biotechnology and industry. *F. velutipes* has been previously found to convert disaccharides (cellobiose, maltose, and sucrose), trisaccharide (cellotriose), and oligosaccharide (cellotetraose) to ethanol at similar recovery rates to glucose and to convert glucose to ethanol at a similar level as *S. cerevisiae* [79,80]. This ability of *F. velutipes* is suitable for consolidated bioprocessing (CBP), which is considered an effective process for bioethanol production from lignocellulosic biomass [81–83]. Previously, we demonstrated that *F. velutipes*, the closest white rot fungus to *F. rossica*, is a very attractive model for bioethanol production due to the numerous genes associated with lignocellulolytic enzymes such as CAZymes [8]. In this study, we performed sequencing of the *F. rossica* genome to identify genes involved in lignocellulose degradation. As described above, various CAZyme genes were identified in the *F. rossica* genome including 102 auxiliary activities, 236 glycoside hydrolases, 94 glycosyltransferases, 19 polysaccharide lyases, 56 carbohydrate esterases, and 21 carbohydrate binding-modules. Although further studies on CAZyme genes are needed, this study suggests that *F. rossica* has great potential for future production of biomaterials such as bioenergy.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2076-2607/7/10/421/s1>.

**Author Contributions:** Y.-J.P. conceived and designed the experiments; Y.-J.P., C.-S.L., and W.-S.K. performed the experiments; Y.-J.P. conducted the bioinformatics analysis and analyzed the data; Y.-J.P. wrote the paper.

**Funding:** This research was supported by the Basic Science Research Program of the National Research Foundation of Korea (NRF), funded by the Ministry of Education (2016R1D1A1B03930808).

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

## References

1. Redhead, S.A.; Petersen, R.H. New species, varieties and combinations in the genus *Flammulina*. *Mycotaxon* **1999**, *71*, 285–294.
2. Hughes, K.W.; McGhee, L.L.; Methven, A.S.; Johnson, J.E.; Petersen, R.H. Patterns of geographic speciation in the genus *Flammulina* based on sequences of the ribosomal ITS-5.8S-ITS2 area. *Mycologia* **1999**, *91*, 978–986. [[CrossRef](#)]
3. Adamcik, S.; Ripkova, S. New collections of *Flammulina rossica*. *Czech Mycol.* **2008**, *60*, 113–121. [[CrossRef](#)]
4. 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*, D490–D495. [[CrossRef](#)]
5. Sista Kameshwar, A.K.; Qin, W. Comparative study of genome-wide plant biomass-degrading CAZymes in white rot, brown rot and soft rot fungi. *Mycology* **2018**, *9*, 93–105. [[CrossRef](#)] [[PubMed](#)]
6. Rytioja, J.; Hildén, K.; Yuzon, J.; Hatakka, A.; de Vries, R.P.; Mäkelä, M.R. Plant-polysaccharide-degrading enzymes from basidiomycetes. *Microbiol. Mol. Biol. Rev.* **2014**, *78*, 614–649. [[CrossRef](#)] [[PubMed](#)]
7. Eriksson, K.; Blanchette, R.A.; Ander, P. Morphological aspects of wood degradation by fungi and bacteria. In *Microbial and Enzymatic Degradation of Wood and Wood Components*; Springer: Berlin/Heidelberg, Germany, 1990; pp. 1–87.
8. Park, Y.J.; Baek, J.H.; Lee, S.; Kim, C.; Rhee, H.; Kim, H.; Seo, J.S.; Park, H.R.; Yoon, D.E.; Nam, J.Y.; et al. Whole genome and global gene expression analyses of the model mushroom *Flammulina velutipes* reveal a high capacity for lignocellulose degradation. *PLoS ONE* **2014**, *9*, e93560. [[CrossRef](#)]
9. Park, Y.J.; Jeong, Y.U.; Kong, W.S. Genome sequencing and carbohydrate-active enzyme (CAZyme) repertoire of the white rot fungus *Flammulina elastica*. *Int. J. Mol. Sci.* **2018**, *19*, 2379. [[CrossRef](#)]
10. Lee, C.S.; Kong, W.S.; Park, Y.J. Genome sequencing and genome-wide identification of carbohydrate-active enzymes (CAZymes) in the white rot fungus *Flammulina fennae*. *Microbiol. Biotechnol. Lett.* **2018**, *46*, 300–312. [[CrossRef](#)]
11. Park, Y.J.; Kong, W.S. Genome-wide comparison of carbohydrate-active enzymes (CAZymes) repertoire of *Flammulina ononidis*. *Mycobiology* **2018**, *46*, 349–360. [[CrossRef](#)]
12. Martinez, D.; Larrondo, L.F.; Putnam, N.; Sollewijn Gelpke, M.D.; Huang, K.; Chapman, J.; Helfenbein, K.G.; Ramaiya, P.; Detter, J.C.; Larimer, F.; et al. Genome sequence of the lignocellulose degrading fungus *Phanerochaete chrysosporium* strain RP78. *Nat. Biotechnol.* **2004**, *22*, 695–700. [[CrossRef](#)] [[PubMed](#)]
13. Kämper, J.; Kahmann, R.; Bölker, M.; Ma, L.-J.; Brefort, T.; Saville, B.J.; Banuett, F.; Kronstad, J.W.; Gold, S.E.; Müller, O.; et al. Insights from the genome of the biotrophic fungal plant pathogen *Ustilago maydis*. *Nature* **2006**, *444*, 97–101. [[CrossRef](#)] [[PubMed](#)]
14. Andrews, S. FastQC: A quality control tool for high throughput sequence data. Available online: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/> (accessed on 4 May 2019).
15. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [[CrossRef](#)] [[PubMed](#)]
16. Zerbino, D.R.; Birney, E. Velvet: Algorithms for  $\beta$  short read assembly using de Bruijn graphs. *Genome Res.* **2008**, *18*, 821–829. [[CrossRef](#)] [[PubMed](#)]
17. Stanke, M.; Morgenstern, B. AUGUSTUS: A web server for gene prediction in eukaryotes that allows user-defined constraints. *Nucleic Acids Res.* **2005**, *33*, W465–W467. [[CrossRef](#)] [[PubMed](#)]
18. Buchfink, B.; Xie, C.; Huson, D. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **2015**, *12*, 59–60. [[CrossRef](#)] [[PubMed](#)]
19. Emms, D.; Kelly, S. OrthoFinder: Solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol.* **2015**, *16*, 157. [[CrossRef](#)]

20. Galagan, J.E.; Calvo, S.E.; Cuomo, C.; Ma, L.J.; Wortman, J.R.; Batzoglou, S.; Lee, S.I.; Baştürkmen, M.; Spevak, C.C.; Clutterbuck, J.; et al. Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae*. *Nature* **2005**, *438*, 1105–1115. [[CrossRef](#)]
21. Staats, M.; van Kan, J.A. Genome update of *Botrytis cinerea* strains B05. 10 and T4. *Eukaryot. Cell* **2012**, *11*, 1413–1414. [[CrossRef](#)]
22. Morin, E.; Kohler, A.; Baker, A.R.; Foulongne-Oriol, M.; Lombard, V.; Nagye, 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](#)]
23. Stajich, J.E.; Wilke, S.K.; Ahrén, D.; Au, C.H.; Birren, B.W.; Borodovsky, M.; Burns, C.; Canbäck, B.; Casselton, L.A.; Cheng, C.K.; et al. Insights into evolution of multicellular fungi from the assembled chromosomes of the mushroom *Coprinopsis cinerea* (*Coprinus cinereus*). *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 11889–11894. [[CrossRef](#)] [[PubMed](#)]
24. Martin, F.; Aerts, A.; Ahrén, D.; Brun, A.; Danchin, E.G.J.; Duchaussoy, F.; Gibon, J.; Kohler, A.; Lindquist, E.; Pereda, V.; et al. The genome of *Laccaria bicolor* provides insights into mycorrhizal symbiosis. *Nature* **2008**, *452*, 88–92. [[CrossRef](#)] [[PubMed](#)]
25. Chen, L.; Gong, Y.; Cai, Y.; Liu, W.; Zhou, Y.; Xiao, Y.; Xu, Z.; Liu, Y.; Lei, X.; Wang, G.; et al. Genome sequence of the edible cultivated mushroom *Lentinula edodes* (Shiitake) reveals insights into lignocellulose degradation. *PLoS ONE* **2016**, *11*, e0160336. [[CrossRef](#)] [[PubMed](#)]
26. 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](#)]
27. Janbon, G.; Ormerod, K.L.; Paulet, D.; Byrnes, E.J., III; Yadav, V.; Chatterjee, G.; Mullapudi, N.; Hon, C.-C.; Billmyre, R.B.; Brunel, F.; et al. Analysis of the genome and transcriptome of *Cryptococcus neoformans* var. *grubii* reveals complex RNA expression and microevolution leading to virulence attenuation. *PLOS Genet.* **2014**, *10*, e1004261.
28. Fisk, D.G.; Ball, C.A.; Dolinski, K.; Engel, S.R.; Hong, E.L.; Issel-Tarver, L.; Schwartz, K.; Sethuraman, A.; Botstein, D.; Cherry, J.M. *Saccharomyces cerevisiae* S288C genome annotation: A working hypothesis. *Yeast* **2006**, *23*, 857–865. [[CrossRef](#)] [[PubMed](#)]
29. Galagan, J.E.; Calvo, S.E.; Borkovich, K.A.; Selker, E.U.; Read, N.D.; Jaffe, D.; FitzHugh, W.; Ma, L.J.; Smirnov, S.; Purcell, S.; et al. The genome sequence of the filamentous fungus *Neurospora crassa*. *Nature* **2003**, *422*, 859–868. [[CrossRef](#)]
30. Ohm, R.A.; De Jong, J.F.; Lugones, L.G.; Aerts, A.; Kothe, E.; Stajich, J.E.; de Vries, R.P.; Record, E.; Levasseur, A.; Baker, S.E.; et al. Genome sequence of the model mushroom *Schizophyllum commune*. *Nat. Biotechnol.* **2010**, *28*, 957–963. [[CrossRef](#)]
31. Li, W.C.; Huang, C.H.; Chen, C.L.; Chuang, Y.C.; Tung, S.Y.; Wang, T.F. *Trichoderma reesei* complete genome sequence, repeat-induced point mutation, and partitioning of CAZyme gene clusters. *Biotechnol. Biofuels* **2017**, *10*, 170. [[CrossRef](#)]
32. Zhang, H.; Yohe, T.; Huang, L.; Entwistle, S.; Wu, P.; Yang, Z.; Busk, P.K.; Xu, Y.; Yin, Y. dbCAN2: A meta server for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* **2018**, *46*, W95–W101. [[CrossRef](#)]
33. Petersen, T.N.; Brunak, S.; von Heijne, G.; Nielsen, H. SignalP 4.0: Discriminating signal peptides from transmembrane regions. *Nat. Methods* **2011**, *8*, 785–786. [[CrossRef](#)] [[PubMed](#)]
34. Breton, C.; Šnajdrová, L.; Jeanneau, C.; Koča, J.; Imberty, A. Structures and mechanisms of glycosyltransferases. *Glycobiology* **2006**, *16*, 29R–37R. [[CrossRef](#)] [[PubMed](#)]
35. Lairson, L.L.; Henrissat, B.; Davies, G.J.; Withers, S.G. Glycosyltransferases: Structures, functions, and mechanisms. *Annu. Rev. Biochem.* **2008**, *77*, 521–555. [[CrossRef](#)] [[PubMed](#)]
36. Coutinho, P.M.; Deleury, E.; Davies, G.J.; Henrissat, B. An evolving hierarchical family classification for glycosyltransferases. *J. Mol. Biol.* **2003**, *328*, 307–317. [[CrossRef](#)]
37. Campbell, J.A.; Davies, G.J.; Bulone, V.; Henrissat, B. A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities. *Biochem. J.* **1997**, *326*, 929–939. [[CrossRef](#)] [[PubMed](#)]
38. Paulson, J.C.; Weinstein, J.; Ujita, E.L.; Riggs, K.J.; Lai, H. The membrane-binding domain of a rat liver Golgi sialyltransferase. *Biochem. Soc. Trans.* **1987**, *15*, 618–620. [[CrossRef](#)]

39. Wickner, W.T.; Lodish, H.F. Multiple mechanisms of protein insertion into and across membranes. *Science* **1985**, *230*, 400–407. [[CrossRef](#)]
40. Bornscheuer, U.T. Microbial carboxyl esterases: Classification, properties and application in biocatalysis. *FEMS Microbiol. Rev.* **2002**, *26*, 73–81. [[CrossRef](#)]
41. Jaeger, K.E.; Eggert, T. Lipases for biotechnology. *Curr. Opin. Biotechnol.* **2002**, *13*, 390–397. [[CrossRef](#)]
42. Cantarel, B.L.; Coutinho, P.M.; Rancurel, C.; Bernard, T.; Lombard, V.; Henrissat, B. The Carbohydrate-Active EnZymes database (CAZy): An expert resource for glycogenomics. *Nucleic Acids Res.* **2009**, *37*, D233–D238. [[CrossRef](#)]
43. Biely, P. Microbial carbohydrate esterases deacetylating plant polysaccharides. *Biotechnol. Adv.* **2012**, *30*, 1575–1588. [[CrossRef](#)] [[PubMed](#)]
44. Nakamura, A.M.; Nascimento, A.S.; Polikarpov, I. Structural diversity of carbohydrate esterases. *Biotechnol. Res. Innov.* **2017**, *1*, 35–51. [[CrossRef](#)]
45. Adesioye, F.A.; Makhallanyane, T.P.; Biely, P.; Cowan, D.A. Phylogeny, classification and metagenomic bioprospecting of microbial acetyl xylan esterases. *Enzyme. Microb. Technol.* **2016**, *93*, 79–91. [[CrossRef](#)]
46. Wei, Y.; Schottel, J.L.; Derewenda, U.; Swenson, L.; Patkar, S.; Derewenda, Z.S. A novel variant of the catalytic triad in the *Streptomyces scabies* esterase. *Nat. Struct. Biol.* **1995**, *2*, 218–223. [[CrossRef](#)] [[PubMed](#)]
47. Petersen, E.I.; Valinger, G.; Sölkner, B.; Stubenrauch, G.; Schwab, H. A novel esterase from *Burkholderia gladioli* shows high deacetylation activity on cephalosporins is related to  $\beta$ -lactamases and DD-peptidases. *J. Biotechnol.* **2001**, *89*, 11–25. [[CrossRef](#)]
48. Christov, L.P.; Prior, B.A. Esterases of xylan-degrading microorganisms: Production, properties, and significance. *Enzyme Microb. Tech.* **1993**, *15*, 460–475. [[CrossRef](#)]
49. Berlemont, R.; Martiny, A.C. Glycoside hydrolases across environmental microbial communities. *PLoS Comput. Biol.* **2016**, *12*, e1005300. [[CrossRef](#)] [[PubMed](#)]
50. Henrissat, B. A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J.* **1991**, *280*, 309–316. [[CrossRef](#)]
51. Hahn, M.; Olsen, O.; Politz, O.; Borriss, R.; Heinemann, U. Crystal structure and site-directed mutagenesis of *Bacillus macerans* endo-1,3-1,4- $\beta$ -glucanase. *J. Biol. Chem.* **1995**, *270*, 3081–3088. [[CrossRef](#)]
52. Masuda, S.; Endo, K.; Koizumi, N.; Hayami, T.; Fukazawa, T.; Yatsunami, R.; Fukui, T.; Nakamura, S. Molecular identification of a novel  $\beta$ -1,3-glucanase from alkaliphilic *Nocardioopsis* sp. strain F96. *Extremophiles* **2006**, *10*, 251–255. [[CrossRef](#)]
53. Kotake, T.; Hirata, N.; Degi, Y.; Ishiguro, M.; Kitazawa, K.; Takata, R.; Ichinose, H.; Kaneko, S.; Igarashi, K.; Samejima, M.; et al. Endo- $\beta$ -1,3-galactanase from winter mushroom *Flammulina velutipes*. *J. Biol. Chem.* **2011**, *286*, 27848–27854. [[CrossRef](#)] [[PubMed](#)]
54. Horton, P.; Park, K.J.; Obayashi, T.; Fujita, N.; Harada, H.; Adams-Collier, C.J.; Nakai, K. WoLF PSORT: Protein localization predictor. *Nucleic Acids Res.* **2007**, *35*, W585–W587. [[CrossRef](#)] [[PubMed](#)]
55. Berlemont, R. Distribution and diversity of enzymes for polysaccharide degradation in fungi. *Sci. Rep.* **2017**, *7*, 222. [[CrossRef](#)] [[PubMed](#)]
56. Wilson, D.B. Microbial diversity of cellulose hydrolysis. *Curr. Opin. Microbiol.* **2011**, *14*, 259–263. [[CrossRef](#)] [[PubMed](#)]
57. Berlemont, R.; Martiny, A.C. Phylogenetic distribution of potential cellulases in bacteria. *Appl. Environ. Microbiol.* **2013**, *79*, 1545–1554. [[CrossRef](#)] [[PubMed](#)]
58. Berlemont, R.; Martiny, A.C. Genomic potential for polysaccharides deconstruction in bacteria. *Appl. Environ. Microbiol.* **2015**, *81*, 1513–1519. [[CrossRef](#)] [[PubMed](#)]
59. Talamantes, D.; Biabini, N.; Dang, H.; Abdoun, K.; Berlemont, R. Natural diversity of cellulases, xylanases, and chitinases in bacteria. *Biotechnol. Biofuels* **2016**, *9*, 133. [[CrossRef](#)]
60. Eichlerová, I.; Homolka, L.; Žifčáková, L.; Lisá, L.; Dobiášová, P.; Baldrian, P. Enzymatic systems involved in decomposition reflects the ecology and taxonomy of saprotrophic fungi. *Fungal Ecol.* **2015**, *13*, 10–22. [[CrossRef](#)]
61. Treseder, K.K.; Lennon, J.T. Fungal traits that drive ecosystem dynamics on land. *Microbiol. Mol. Biol. Rev.* **2015**, *79*, 243–262. [[CrossRef](#)]
62. Sutherland, I.W. Polysaccharide lyases. *FEMS Microbiol. Rev.* **1995**, *16*, 323–347. [[CrossRef](#)]
63. Yip, V.L.; Withers, S.G. Breakdown of oligosaccharides by the process of elimination. *Curr. Opin. Chem. Biol.* **2006**, *10*, 147–155. [[CrossRef](#)] [[PubMed](#)]

64. Zhao, Z.; Liu, H.; Wang, C.; Xu, J.R. Correction: Comparative analysis of fungal genomes reveals different plant cell wall degrading capacity in fungi. *BMC Genomics* **2013**, *14*, 274. [[CrossRef](#)] [[PubMed](#)]
65. Van den Brink, J.; de Vries, R.P. Fungal enzyme sets for plant polysaccharide degradation. *Appl. Microbiol. Biotechnol.* **2011**, *91*, 1477–1492. [[CrossRef](#)]
66. Garron, M.L.; Cygler, M. Structural and mechanistic classification of uronic acid-containing polysaccharide lyases. *Glycobiology* **2010**, *20*, 1547–1573. [[CrossRef](#)] [[PubMed](#)]
67. Linhardt, R.J.; Galliher, P.M.; Cooney, C.L. Polysaccharide lyases. *Appl. Biochem. Biotechnol.* **1987**, *12*, 135–176. [[CrossRef](#)] [[PubMed](#)]
68. Xavier-Santos, S.; Carvalho, C.C.; Bonfá, M.; Silva, R.; Capelari, M.; Gomes, E. Screening for pectinolytic activity of wood-rotting basidiomycetes and characterization of the enzymes. *Folia Microbiol. (Praha)* **2004**, *49*, 46–52. [[CrossRef](#)] [[PubMed](#)]
69. Reiss, R.; Ihssen, J.; Richter, M.; Eichhorn, E.; Schilling, B.; Thöny-Meyer, L. Laccase versus laccase-like multi-copper oxidase: A comparative study of similar enzymes with diverse substrate spectra. *PLoS ONE* **2013**, *8*, e65633. [[CrossRef](#)] [[PubMed](#)]
70. Fernandez, I.S.; Ruiz-Duenas, F.J.; Santillana, E.; Ferreira, P.; Martinez, M.J.; Martinez, A.T.; Romero, A. Novel structural features in the GMC family of oxidoreductases revealed by the crystal structure of fungal aryl-alcohol oxidase. *Acta Crystallogr.* **2009**, *D65*, 1196–1205. [[CrossRef](#)]
71. Wierenga, R.K.; Drenth, J.; Schulz, G.E. Comparison of the 3-dimensional protein and nucleotide structure of the FAD-binding domain of parahydroxybenzoate hydroxylase with the FAD-binding as well as NADPH-binding domains of glutathionereductase. *J. Mol. Biol.* **1983**, *167*, 725–739. [[CrossRef](#)]
72. Varela, E.; Martinet, M.J.; Martinez, A.T. Arylalcohol oxidase protein sequence: A comparison with glucose oxidase and other FAD oxidoreductases. *Biochem. Biophys. Acta Protein Struct. Mol. Enzymol.* **2000**, *1481*, 202–208. [[CrossRef](#)]
73. Leonowicz, A.; Matuszewska, A.; Luterek, J.; Ziegenhagen, D.; Wojtas-Wasilewska, M.; Cho, N.S.; Hofrichter, M.; Rogalski, J. Biodegradation of lignin by white rot fungi. *Fungal Genet. Biol.* **1999**, *27*, 175–185. [[CrossRef](#)] [[PubMed](#)]
74. Guillén, F.; Martínez, M.J.; Gutiérrez, A.; Del Rio, J.C. Biodegradation of lignocellulosics: Microbial, chemical, and enzymatic aspects of the fungal attack of lignin. *Int. Microbiol.* **2005**, *8*, 195–204.
75. Boraston, A.B.; Bolam, D.N.; Gilbert, H.J.; Davies, G.J. Carbohydrate-binding modules: Fine-tuning polysaccharide recognition. *Biochem. J.* **2004**, *382*, 769–781. [[CrossRef](#)] [[PubMed](#)]
76. Shoseyov, O.; Shani, Z.; Levy, I. Carbohydrate binding modules: Biochemical properties and novel applications. *Microbiol. Mol. Biol. Rev.* **2006**, *70*, 283–295. [[CrossRef](#)] [[PubMed](#)]
77. The CAZypedia Consortium. Ten years of CAZypedia: A living encyclopedia of carbohydrate-active enzymes. *Glycobiology* **2018**, *28*, 3–8.
78. Várnai, A.; Mäkelä, M.R.; Djajadi, D.T.; Rahikainen, J.; Hatakka, A.; Viikari, L. Carbohydrate-binding modules of fungal cellulases: Occurrence in nature, function, and relevance in industrial biomass conversion. *Adv. Appl. Microbiol.* **2014**, *88*, 103–165. [[PubMed](#)]
79. Mizuno, R.; Ichinose, H.; Honda, M.; Takabatake, K.; Sotome, I.; Takai, T.; Maehara, T.; Okadome, H.; Isobe, S.; Gau, M.; et al. Use of whole crop sorghums as a raw material in consolidated bioprocessing bioethanol production Using *Flammulina velutipes*. *Biosci. Biotechnol. Biochem.* **2009**, *73*, 1671–1673. [[CrossRef](#)]
80. Mizuno, R.; Ichinose, H.; Maehara, T.; Takabatake, K.; Kaneko, S. Properties of ethanol fermentation by *Flammulina velutipes*. *Biosci. Biotechnol. Biochem.* **2009**, *73*, 2240–2245. [[CrossRef](#)]
81. Lynd, L.R.; Van Zyl, W.H.; McBride, J.E.; Laser, M. Consolidated bioprocessing of cellulosic biomass: An update. *Curr. Opin. Biotechnol.* **2005**, *16*, 577–583. [[CrossRef](#)]
82. Van Zyl, W.H.; Lynd, L.R.; Den Haan, R.; McBride, J.E. Consolidated bioprocessing for bioethanol production using *Saccharomyces cerevisiae*. *Adv. Biochem. Eng. Biotechnol.* **2007**, *108*, 205–235.
83. Xu, Q.; Singh, A.; Himmel, M.E. Perspectives and new directions for the production of bioethanol using consolidated bioprocessing of lignocellulose. *Curr. Opin. Biotechnol.* **2009**, *20*, 364–371. [[CrossRef](#)] [[PubMed](#)]

