G3, 2021, 11(2), jkaa008
DOI: 10.1093/g3journal/jkaa008
Advance Access Publication Date: 22 December 2020
Fungal Genetics and Genomics

Genome sequence of the oyster mushroom *Pleurotus ostreatus* strain PC9

Yi-Yun Lee,[1,2] Guillermo Vidal-Diez de Ulzurrun,[1] Erich M. Schwarz ,[3] Jason E. Stajich ,[4] and Yen-Ping Hsueh [1,2,5,*]

[1]Institute of Molecular Biology, Academia Sinica, Nangang, Taipei 115, Taiwan
[2]Genome and Systems Biology Degree Program, National Taiwan University and Academic Sinica, Taipei, Taiwan
[3]Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY 14853-2703, USA
[4]Department of Microbiology and Plant Pathology, University of California, Riverside, CA 92521, USA
[5]Department of Biochemical Science and Technology, National Taiwan University, Taipei 106, Taiwan

*Corresponding author: Institute of Molecular Biology, Academia Sinica, 128 Academia Road, Section 2, Nangang, Taipei 115, Taiwan. pinghsueh@gate.sinica.edu.tw

Abstract

The oyster mushroom *Pleurotus ostreatus* is a basidiomycete commonly found in the rotten wood and it is one of the most cultivated edible mushrooms globally. *Pleurotus ostreatus* is also a carnivorous fungus, which can paralyze and kill nematodes within minutes. However, the molecular mechanisms of the predator–prey interactions between *P. ostreatus* and nematodes remain unclear. PC9 and PC15 are two model strains of *P. ostreatus* and the genomes of both strains have been sequenced and deposited at the Joint Genome Institute (JGI). These two monokaryotic strains exhibit dramatic differences in growth, but because PC9 grows more robustly in laboratory conditions, it has become the strain of choice for many studies. Despite the fact that PC9 is the common strain for investigation, its genome is fragmentary and incomplete relative to that of PC15. To overcome this problem, we used PacBio long reads and Illumina sequencing to assemble and polish a more integrated genome for PC9. Our PC9 genome assembly, distributed across 17 scaffolds, is highly contiguous and includes five telomere-to-telomere scaffolds, dramatically improving the genome quality. We believe that our PC9 genome resource will be useful to the fungal research community investigating various aspects of *P. ostreatus* biology.

Keywords: oyster mushroom; *Pleurotus ostreatus*; whole-genome sequencing

Introduction

Pleurotus ostreatus is a common edible basidiomycete that ranks second for global mushroom consumption (Cohen *et al.* 2002; Sánchez 2010). In the wild, this fungus prefers temperate habitats, such as subtropical forest (Hilber 1997), where it extracts nutrients from dead or dying trees (Karim *et al.* 2016). In order to degrade the wood, *P. ostreatus* releases several enzymes such as lignin, cellulose, and hemicellulose (Sánchez 2009). Nematicidal toxins produced by *P. ostreatus* under starvation conditions, such as trans-2-decenedioic, can paralyze and kill their nematode-prey (Barron and Thorn 1987; Kwok *et al.* 1992). Another potential toxin, linoleic acid, reduces nematodes head size (Satou *et al.* 2008). When nematodes contact the mycelium of *P. ostreatus*, fungal toxins enter the prey through their sensory cilia, leading to hypercontraction and calcium influx of pharyngeal and body wall muscles, ultimately causing necrosis of the neuromuscular system (Lee *et al.* 2020). However, the identity of the toxins and how exactly *P. ostreatus* induces rapid cell necrosis remain unclear.

The *P. ostreatus* dikaryotic strain N001 that produces fruiting bodies is a common commercial strain (Peñas *et al.* 1998; Larraya *et al.* 1999a). In Larraya *et al.* (1999b), two monokaryotic strains (PC9 and PC15) were differentiated from N001 in order to study reproduction and cloning in this mushroom. These two monokaryotic strains of *P. ostreatus* exhibit differences in growth pattern, with PC9 exhibiting faster growth relative to PC15. Although the genomes of these two monokaryotic strains have already been sequenced (Riley *et al.* 2014; Alfaro *et al.* 2016; Castanera *et al.* 2016), the quality of the PC9 genome is poor for current standards. Whereas the latest version of the PC15 genome comprises only 12 scaffolds, the currently available PC9 genome (hereafter denoted PC9_JGI) is still distributed across 572 scaffolds, most of which size are smaller than 1 kb (Alfaro *et al.* 2016; Castanera *et al.* 2016). Despite PC15 having a more complete genome assembly than PC9, PC9 is often the preferred strain for research due to its robust growth in the laboratory. For instance, PC9 has been used to study ligninolytic activity (Nakazawa *et al.* 2017a, 2017b, 2019), gene modification (Yoav *et al.* 2018), and transformation (Nakazawa *et al.* 2016). Genetic manipulation has proven to significantly increase the production of primary and secondary metabolites in several filamentous fungi that are important in industry (Banerjee *et al.* 2003; Kück and Hoff 2010). Therefore, a high-quality PC9 genome is needed to enable scientists to conduct advanced genomic studies and to facilitate functional analyses of *P. ostreatus* genes.

Consequently, we decided to sequence and annotate an updated genome for PC9. In this study, we used long PacBio reads and Illumina reads to assemble the genome of *P. ostreatus* strain PC9 and functionally annotated its predicted proteins. The final

Received: September 03, 2020. **Accepted:** October 1, 2020

assembly, obtained by merging draft assemblies built with CANU (Koren *et al.*) and FALCON (Chin *et al.*), was distributed across 17 scaffolds, including five telomere-to-telomere scaffolds. In addition, the N50 of our assembly is ~3.5 Mb, which ranks it among the best for currently available *Pleurotus* genomes. In summary, here we present a new high-quality PC9 genome that will be help the fungal research community to study this important mushroom.

Materials and methods

Fungal strain and DNA extraction

Pleurotus ostreatus strain PC9 was cultured on yeast extract, malt extract, and glucose medium (YMG) at 25° for 7 days. Mycelia were collected from original plates (5 cm) and added to a 100-mL YMG liquid medium shaken at 200 r.p.m. at 25° for 3 days. After 3 days, the culture was transferred to yeast nitrogen base without amino acids liquid medium (YNB), and it was shaken at 200 r.p.m. for 2 additional days. DNA was extracted by using cetyltrimethylammonium bromide and purified with chloroform, isopropanol, and phenol-chloroform.

Genome sequencing and assembly

PacBio long reads were sequenced from the *P. ostreatus* PC9 genome using the Pacbio RSII platform, and a PacBio SMRTBell Template Prep Kit 1.0 SPv3 was used for library preparation. This approach resulted in a total of ~0.3 M reads with a mean length of 14,742 base pairs (bp), representing 161X genome coverage for PC9. The Pacbio reads were used to construct two draft assemblies. The first assembly was built with Canu (v1.7) (Koren *et al.* 2017) using the parameters genomeSize = 36 m, useGrid = false, and maxThreads = 8. The second assembly was built with Falcon (pbalign v0.02) (Chin *et al.* 2016). For this latter assembly, the configuration file was downloaded from https://pb-falcon.readthedocs.io/en/latest/parameters.html, and the parameter Genome_size was set to 36 Mb since published *P. ostreatus* published genome sizes range from 34.3 to 35.6 Mb (Riley *et al.* 2014; Alfaro *et al.* 2016; Castanera *et al.* 2016). Next, we used Quiver (genomicconsensus v2.3.2) (Chin *et al.* 2013) to polish both these PacBio assemblies. Information on the raw polished draft assemblies is presented in Supplementary Table S1.

The Canu assembly contained more telomeric regions, whereas the Falcon assembly was more contiguous. We merged both assemblies using Quickmerge (v0.3) (Chakraborty *et al.* 2016), with Canu as reference and Falcon as donor (Canu–Falcon) or *vice versa* (Falcon–Canu). The Falcon–Canu assembly showed better contiguity, having only 29 scaffolds (Supplementary Table S2), so we selected it as the basis for our final assembly. Redundant contigs were detected using nucmer (Mummer4) (Marçais *et al.* 2018), which were subsequently filtered out of the assembly. Nucmer was also used to detect large overlapping regions within the scaffolds. When a unique overlapping region (scaffolds overlapping at multiple sites were not merged together) exceeded 10 kb, presented high identity (>99%), and lay at the end or beginning of the scaffolds, we manually merged the two scaffolds at the overlapping region (Supplementary Table S3). In total, 10 scaffolds were merged in this way, yielding five larger and more complete scaffolds (scaffolds 2, 4, 6, 7, and 13 of the final assembly) (Supplementary Table S3). Finally, the cleaned assembly consisting of 17 scaffolds was further polished using Illumina reads and Pilon (Walker *et al.* 2014). A total of 100 M Illumina sequence pair-end reads of 151 bp were used for Pilon polishing.

Genome annotation

To annotate the assembled PC9 genome, we used *funannotate* (v1.5.2) (Jon Love *et al.* 2019) with the pipeline described in https://funannotate.readthedocs.io/en/latest/tutorials.html. We used the following commands: *funannotate* mask, to softmask the genome, *funannotate* training, and *funannotate predict* to generate preliminary gene models and consensus gene models [using: AUGUSTUS (Stanke and Waack 2003), GeneMark (Borodovsky and McIninch 1993), and EVidenceModeler (Haas *et al.* 2008)], and *funannotate annotate* to add functional annotation. The functional annotation obtained with funannotate includes Interpro terms, Pfam domains, CAZYmes, secreted proteins, proteases (MEROPS), BUSCO groups, Eggnog annotations, Clusters of Orthologous Groups (COGs), GO ontology, secretion of signal peptides, and transmembrane domains (the full annotation is available in Supplementary Table S4).

Genome analysis and comparison

General assembly statistics for example length and N50 of the scaffolds/contigs were calculated from the assembly fasta file using Perl scripts count_fasta_residues.pl (https://github.com/SchwarzEM/ems_perl/blob/master/fasta/count_fasta_residues.pl). BUSCO completeness was computed using BUSCO 3.0.1 (Simão *et al.* 2015; Waterhouse *et al.* 2018) against the Basidiomycota dataset basidiomycata_odb9 (Simão *et al.* 2015; Waterhouse *et al.* 2018). Repetitive elements were identified using a custom-made repeat library created according to a pipeline described previously (Coghlan *et al.* 2018). The repeat library was used as input for RepeatMasker (Smit *et al.* 2013–2015), and the results were further analyzed using the one_code_to_find_them_all script (Bailly-Bechet *et al.* 2014). The presence of telomeres in the scaffolds was established by searching for the telomeric repeats (TTAGGG)n (Pérez *et al.* 2009). Finally, we used Circos (v0.69.0) (Krzywinski *et al.* 2009) and D-genies (minimap v2) (Cabanettes and Klopp 2018) to illustrate the different genomic features of our assembly and to compare it to the previously published PC9_JGI genome (Alfaro *et al.* 2016; Castanera *et al.* 2016).

Data availability

The final assembled and newly annotated genomes of *P. ostreatus* PC9 (denoted PC9_AS) has been uploaded to NCBI and is available with accession code: JACETU000000000.

The raw PacBio reads of *P. ostreatus* PC9 have been uploaded to Sequence Read Archive (SRA) and the submission code is SRX9217013.

Supplementary material is available at figshare DOI: https://doi.org/10.25387/g3.13023269.

Results and discussion

A new genome assembly of the *P. ostreatus* PC9 strain

Long PacBio and Illumina reads were used to sequence *P. ostreatus* strain PC9 to improve genome quality. The size of our *P. ostreatus* PC9_AS genome assembly is ~35.0 Mb (Table 1), which concurs with the sizes of the currently available *P. ostreatus* genomes: PC9_JGI (35.6 Mb) and PC15 (34.3 Mb) (Riley *et al.* 2014; Alfaro *et al.* 2016; Castanera *et al.* 2016). A comparison with other *Pleurotus* species revealed that our *P. ostreatus* genome is smaller than that of Pleurotus *eryngii* strain 183 (43.8 Mb) (Yang *et al.* 2016), Pleurotus *tuoliensis* strain JKBL130LB (48.2 Mb) (Zhang *et al.* 2018),

Table 1 Genomic features of the three *P. ostreatus* genome assemblies.

General features	PC9_AS	PC9_JGI	PC15
Total, nt	35,032,978	35,630,309	34,342,730
Number of scaffolds	17	572	12
N50 scaffold size, nt	3,500,734	2,086,289	3,270,165
N90 scaffold size, nt	2,134,864	159,002	1,880,400
Scaffold max. nt	4,859,873	4,430,591	4,830,258
Scaffold min. nt	9,086	2,001	280,724
Number of contigs	17	3,272	13
N50 contig size, nt	3,500,734	99,058	3,270,165
N90 contig size, nt	2,134,864	13,737	1.571,664
GC content, %	50.79	50.94	50.95
Genes	11,875	12,206	12,330
BUSCO completeness, %	97.2	97.2	97.6

nt, nucleotides.

and Pleurotus *platypus* strain MG11 (40.0 Mb) (Li *et al.* 2018). PC9_AS is distributed across 17 scaffolds, with the maximum and minimum scaffold sizes being 4.86 Mb and 9.1 kb, respectively. The N50 value of our assembly data is 3.5 Mb, which ranks it highest among the available *Pleurotus* genomes, including for PC9_JGI (N50 = 2.09 Mb) (Alfaro *et al.* 2016; Castanera *et al.* 2016), PC15 (N50 = 3.27 Mb) (Riley *et al.* 2014; Alfaro *et al.* 2016; Castanera *et al.* 2016), CCMSSC03989 (N50 = 2.85 Mb) (Wang *et al.* 2018), *P. tuoliensis* strain JKBL130LB (N50 = 1.17 Mb) (Zhang *et al.* 2018), and *P. eryngii* strain 183 (N50 = 509 kb) (Yang *et al.* 2016). The total annotated gene number in PC9_AS is 11,875, which is slightly fewer than for PC9_JGI (12,206 genes). This phenomenon might cause by the no match between PC9_AS and PC9_JGI genomes, observing that scaffolds 16 and 25 of PC9_JGI are not highly aligned to the PC9_AS genome, showing the gap between PC9_AS and PC9_JGI genomes (Figure 2A). The completeness of our PC9 genome assembly was assessed with BUSCO (Simão *et al.* 2015; Waterhouse *et al.* 2018) using a Basidiomycota dataset. We obtained a 97.2% BUSCO completeness, with 1289 complete BUSCOs, 1284 complete and single-copy BUSCOs (99.6%), 14 complete and duplicate BUSCOs (1.1%), 9 fragmented BUSCOs (0.7%), and 28 missing BUSCOs genes (2.2%). Overall, our statistical analysis suggests that our PC9_AS assembly is more completed and integrated than that of PC9_JGI (Alfaro *et al.* 2016; Castanera *et al.* 2016).

The genome architecture of *P. ostreatus* strain PC9 is shown in Figure 1A. We used the highly conserved sequence (TTAGGG)n to determine telomere locations (Moyzis *et al.* 1988), which represent the ends of the chromosomes in fungi (Schechtman 1990; Farman and Leong 1995). Out of 17 scaffolds, five possess telomeric repeats at both ends, indicating that these scaffolds represent complete chromosomes (Figure 1A). In contrast, four scaffolds have a telomere at only one end, and the remaining scaffolds lack any apparent telomeric repeats. Interestingly, the small scaffold 14 (62,249 bp) have telomeric repeats so scaffold 14 may constitute the ends of other scaffolds lacking telomeres. We also compared the PC9 and PC15 telomere region (Supplementary Figure S1), and we found the telomeric regions of PC15 showed a similar repeat number of about 24 telomeric repeats distributed in tandem at the beginning and the end of its seven telomere-to-telomere scaffolds and the five other half scaffolds. Regarding scaffolds without telomere regions, we found simple repeats at the beginning and the end of scaffolds 2 and 6, as well as the end of scaffold 8 hindering their consolidation into more complete scaffolds. In addition, we observed that the beginning of scaffold 1 contains class II transposable element (TE), while no genes, TE, or simple repeats were found at the beginning of scaffold 3 and the end of scaffold 7. Larraya *et al.* (1999b) used pulsed-field gel electrophoresis to determine how many chromosomes of PC9 and PC15 have and reported a total of 11 chromosomes in *P. ostreatus*, thus suggesting that our assembly is close to the chromosome level.

The average coverage depth of PC9_AS is 135.3 reads per 10 kb (Figure 1A). Regions of high coverage depth are apparent at the ends of most scaffolds, perhaps due to the presence of telomeric repeats. Interestingly, regions of low coverage depth are found in some of the smallest scaffolds, such as 16 (45,074 bp) and 17 (9086 bp). However, other small scaffolds show high coverage depth, such as 12 (135,750 bp), 14 (62,249 bp), and 15 (54,081 bp). The regions of higher coverage depth in these smaller scaffolds may correspond to the presence of repeats whereas, because of low repeat numbers in centromeric regions, the regions of lower coverage depth might represent misplaced centromeres. Aneuploidy is characterized by abnormal chromosome number and has been observed in other fungal species (Käfer 1960; Cox and Bevan 1962; Pollard *et al.* 1968). When this phenomenon occurs, differences in the depth coverage of the abnormal chromosomes can be clearly observed. However, in our assembly, we did not observe an evidence for aneuploidy. The gene density of PC9_AS in sliding windows of 100 kb is also illustrated in Figure 1A. Gene density is consistent among the different scaffolds, with an average of ~33 genes per 100 kb. Moreover, we observed regions of low gene density that mostly align with stretches of telomeric repeats and with clusters of TEs. Telomeres consist of several tandemly arrayed (TTAGGG)n repeats, which hinder gene translation at these sites. Similarly, low expression of genes proximal to TEs has been reported previously (Castanera *et al.* 2016), potentially explaining low gene density in their vicinity. Although TEs are located near centromeres in certain fungi (Klein and O'Neill 2018), density of TEs in PC9_AS is highly diverse among most scaffolds (Figure 1A). A similar phenomenon has been reported for the previously published PC9_JGI and PC15 genomes (Castanera *et al.* 2016), wherein regions of low gene density aligned with zones harboring many TEs. We identified a total of 253 TE families in our PC9_AS genome, *i.e.*, more than the 80 TE families reported previously for PC9_JGI and PC15 (Castanera *et al.* 2016). These TE families account for 7.12% of the total PC9_AS genome, greater than the 6.2% and 2.5% cited previously for the PC15 and PC9_JGI genomes, respectively (Castanera *et al.* 2016). This striking difference between PC9_AS and PC9_JGI may be due to the fragmented nature of the latter, a scenario that often hinders TEs identification. In Table 2, we summarize the TEs identified in PC9_AS, and it reveals that Class I transposons account for 89% of the TEs, with LTR-retrotransposons being the most abundant family of Class I elements (Supplementary Table S5). This outcome corroborates findings for the PC15 genome (Castanera *et al.* 2016). More detailed information on the TEs identified in PC9_AS genome—including size, number of fragments and repeats, and total bp—is presented in Supplementary Table S5.

Genome annotation

A total of 11,875 genes were annotated from PC9_AS, which is close to the current gene numbers of *P. ostreatus* PC15 and PC9_JGI genomes (12,330 genes and 12,206 genes, respectively). The full annotation is available in Supplementary Table S4, which contain Interpro terms, Pfam domains, CAZYmes, secreted proteins, proteases (MEROPS), BUSCO groups, Eggnog annotations, COGs, GO ontology, secretion of signal peptides, and

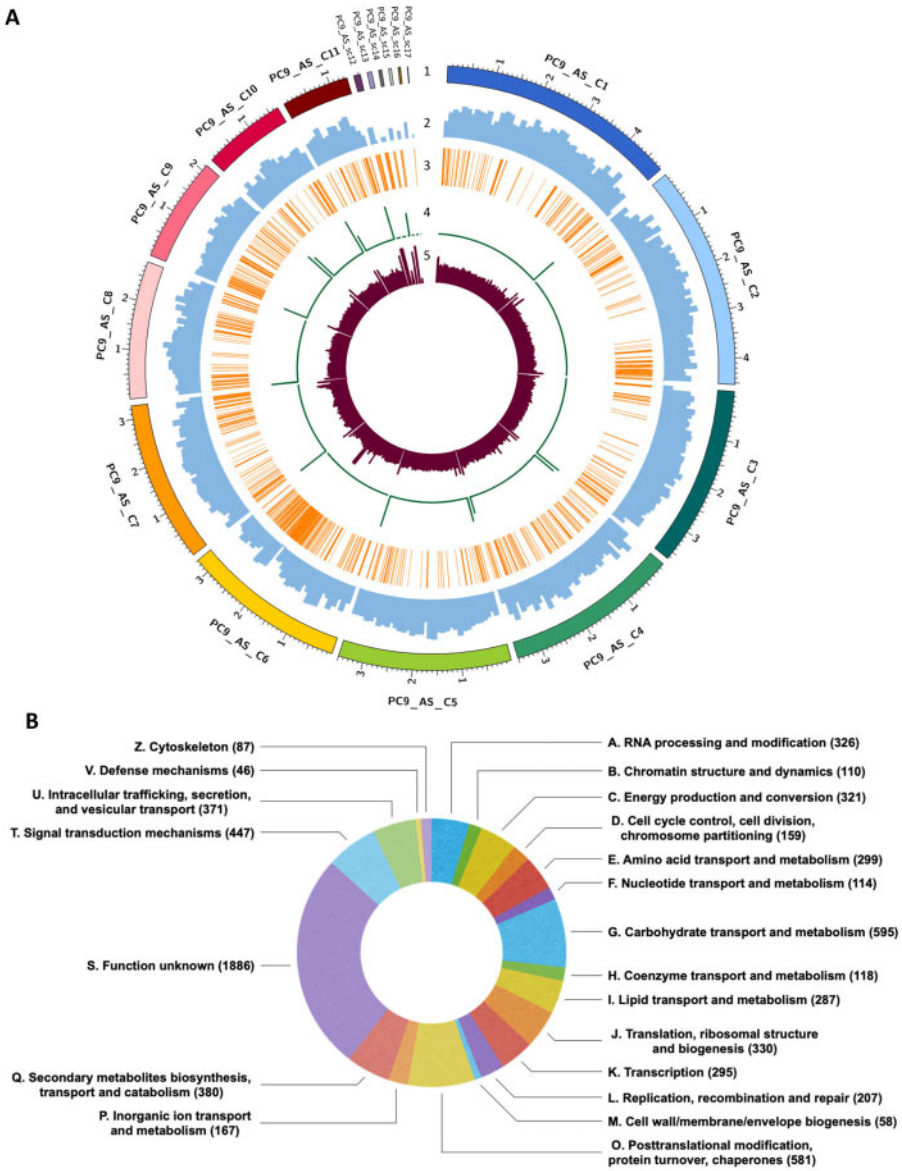

Figure 1 (A) Genome architecture of *P. ostreatus* strain PC9 based on our PC9_AS assembly. Tracks (outer to inner) represent the distribution of genomic features in our PC9_AS assembly: (1) sizes (in Mb) of PC9_AS scaffolds, with numbers prefixed by the letter "C" indicating the order of scaffold size; (2) gene density with 100-kb sliding windows, ranging between 0 and 50 genes; (3) distribution of TEs along the PC9_AS genome; (4) distribution of telomere repeats with 1-kb sliding windows, ranging between 10 and 30 repeats; and (5) depth of sequencing coverage with 10-kb sliding windows, ranging between 0 and 300 depth. (B) Predicted functions of genes identified from the PC9_AS genome, cataloged using COGs database.

transmembrane domains. We used the COGs of proteins database (Tatusov *et al.* 2000) to catalog the functions of those genes (Figure 1B). A considerable number of the gene functions are unknown, but the second biggest cluster in PC9_AS, containing 595 genes, consists of genes coding for putative carbohydrate transport and metabolism, such as transporters, chitinase, and carbohydrate-active enzymes (CAZymes), among others. CAZymes are essential to saprotrophic fungi-like *P. ostreatus* to decay materials that are subsequently used as carbon sources (Mikiashvili *et al.* 2006). These enzymes play important roles in cellulose and hemicellulose degradation (Alfaro *et al.* 2016), with over 130 families described to date (Cantarel *et al.* 2009; Davies and Williams 2016), including glycosyl hydrolases (GHs), carbohydrate esterases (CEs), polysaccharide lyases (PLs), carbohydrate-binding modules (CBMs), glycosyl transferases (GTs), and auxiliary activities (AAs). We further explored the number of CAZymes-related genes in PC9_AS using the CAZymes database (Cantarel *et al.* 2009) and identified 459 such genes, an outcome consistent with our calculations for other *P. ostreatus* strains (Supplementary Table S6; PC15 = 562; PC9_JGI = 408). The EuKaryotic Orthologous Groups (KOG) database, which is a eukaryote-specific version of the COGs, has been used previously to identify the orthologous and paralogous proteins from the PC15 and PC9_JGI genomes (Grigoriev *et al.* 2012). Therefore, we used the KOG classification available from the JGI for those genomes to reveal how gene families are distributed in *P. ostreatus* and to compare with our PC9_AS COG classification. As shown in Supplementary Table S6, the biggest gene class among PC15 and PC9_JGI (both with 662 genes) is signal transduction mechanisms. However, PC9_AS only contains 447 genes coding for signal transduction mechanisms. In the classes of extracellular structures and nuclear structures, PC15 and PC9_JGI both have ~100 genes

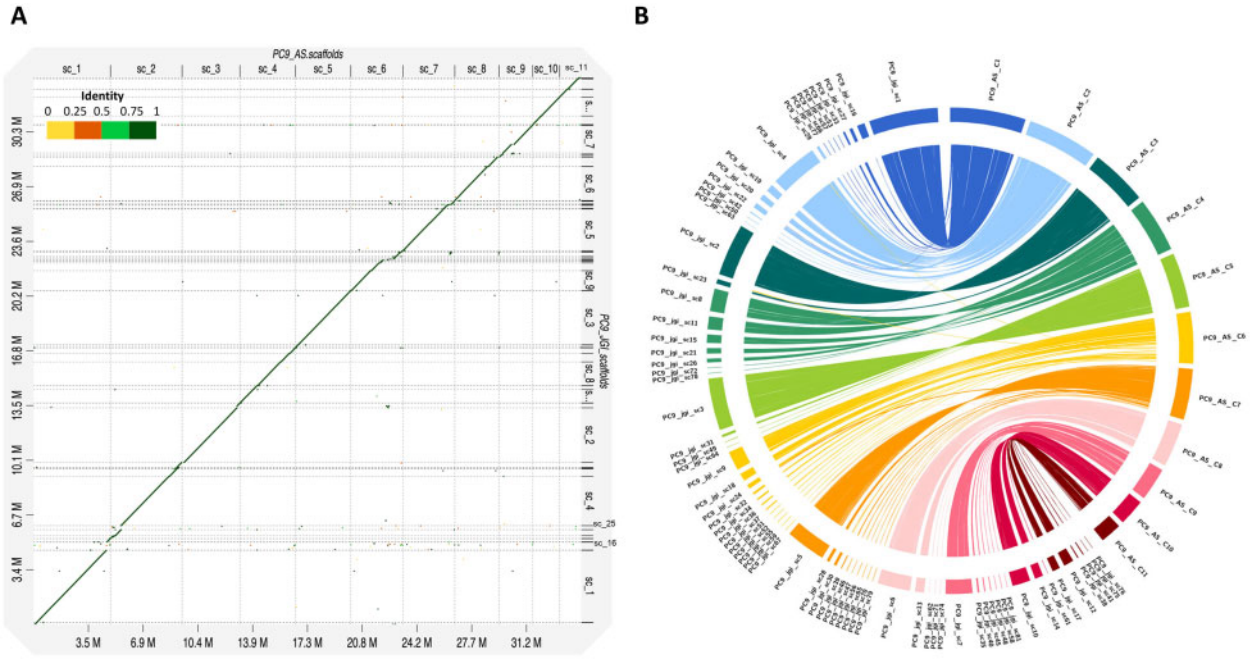

Figure 2 Comparisons between the PC9_AS and PC9_JGI *P. ostreatus* PC9 genomes. (A) Dot-plot alignment of *P. ostreatus* PC9_AS (target) and PC9_JGI (query) generated by D-Genies minimap2. (B) Circos plot showing regions of similarity shared between PC9_AS (scaffolds 1–11) and PC9_JGI (scaffolds 1–81) (identity > 95%, length > 10 kb).

associated with each class, whereas PC9_AS only harbors two extracellular structure genes and five nuclear structure genes. We acknowledge that these differences may reflect the use of different databases for gene function classification. Interestingly, PC9_AS also contained fewer PFAM domains (6770) compared to the other two *P. ostreatus* genomes (PC9_JGI = 7914; PC15 = 7972).

Genomic/genetic comparison of PC9 genomes

We performed a comprehensive comparison of the PC9_JGI (Alfaro *et al.* 2016; Castanera *et al.* 2016) and PC9_AS genomes. First, we observed that the number of scaffolds and contigs of PC9_AS (17) is smaller than that of PC9_JGI (572) (Table 1). In terms of size, both genomes are similar at ~35 Mb, but PC9_JGI is slightly larger than PC9_AS by ~600 kb. In terms of the N50 scaffold and N50 contig sizes, values for PC9_AS (N50 scaffold size = 3,500,734; N50 contig size = 3,500,734) are larger than those of PC9_JGI (N50 scaffold size = 2,086,289; N50 contig size = 99,058), which reflects greater contiguity in the former. PC9 and PC15 are protoclones of N001 (Larraya *et al.* 1999b), and the genome size of PC9 is slightly larger than that of PC15. The latest updated assembly of PC15 is distributed across 12 scaffolds, with a size of 34.3 Mb that is similar to PC9_AS. We used D-Genies (Cabanettes and Klopp 2018) to construct a dot-plot alignment of the genome between PC9_JGI and PC9_AS assemblies. The result, shown in Figure 2A, demonstrates that the PC9_JGI and PC9_AS genomes are highly similar, with small scaffolds of PC9_JGI corresponding to portions of the larger PC9_AS scaffolds, which indicates that our PC9_AS assembly is more complete. Moreover, when we aligned the two genomes in a circos plot (Figure 2B), we observed long regions of high similarity between the most relevant scaffolds of PC9_AS (scaffolds 1-11) (Supplementary Table S7) and PC9_JGI (scaffolds 1–81) (Supplementary Table S8). Figure 2B also reveals that, in general, more than five PC9_JGI scaffolds can be mapped to single PC9_AS scaffolds, with scaffolds 6 and 7 of PC9_AS incorporating 12 and 10 of the PC9_JGI scaffolds, respectively. However, we can find that scaffolds 16 and 25 of PC9_JGI are not highly aligned to PC9_AS genome (Figure 2A) and among 482,463 bp of scaffold 16 of PC9_JGI only ~10,000 bp can align to scaffold 1 of PC9_AS genome (Figure 2B). Moreover, scaffolds 25, 43, 44, 54, 68, 69, 73, and 80 of PC9_JGI cannot align to PC9_AS genome with criteria of identity over 95% and length over 10 kb. These might be the reason why the gene numbers of our

Table 2 Classification of TEs identified from out *P. ostreatus* strain PC9_AS genome assembly.

Family	Fragments	Copies	Total (bp)
Class I			
LINE/I-Jockey	26	8	18,516
LINE/L1	9	3	5,065
LINE/Tad1	39	27	77,409
LINE/R2-NeSL	4	4	6,740
LINE/LOA	7	2	2,406
LINE/L1-Tx1	15	7	13,873
LTR/Copia	307	182	238,979
LTR/Gypsy	1,468	825	1,681,184
LTR/Ngaro	51	31	38,588
LTR/ERVK	14	5	5,240
Other LINE	11	5	3,390
Total Class I repeat	1,951 (82.4%)	1,099 (80%)	2,091,390 (89%)
Class II			
DNA/hAT-Charlie	11	11	5,823
DNA/hAT-Restless	5	2	3,894
DNA/TcMar-Fot1	4	4	4,431
DNA/TcMar-Tc1	31	26	18,898
DNA/hAT-Ac	1	1	546
DNA/TcMar-Sagan	34	22	10,141
DNA/CMC-EnSpm	69	39	89,604
DNA/TcMar-Pogo	2	2	2,140
DNA/RC	68	49	55,928
DNA/TcMar-Ant1	5	5	2,064
DNA/PIF-Harbinger	144	99	38,476
Other DNA	44	16	25,489
Total Class II repeat	418 (17.6%)	276 (20%)	257,434 (11%)
Total repeat	2,369	1,375	2,348,824

assembly (11,875 genes) are less than that of PC9_JGI genome (12,206 genes).

Conclusions

In this study, we used long PacBio and Illumina reads to assemble a new genome of *P. ostreatus* strain PC9. A combination of high read coverage and the latest bioinformatic tools resulted in a high-quality PC9_AS genome. Compared to the currently available PC9_JGI genome, our new assembly is more complete, comprising only 17 scaffolds that include five telomere-to-telomere scaffolds and the highest N50 values yet achieved for a *Pleurotus* genome. Genomic comparisons between PC9_AS and the currently available assemblies of *P. ostreatus* evidence the high quality of our genome assembly. This new PC9 genome will enable the fungal research community to perform further genomic and genetic analyses of *P. ostreatus* and advance our understanding of this common edible mushroom.

Acknowledgments

The authors thank Ursula Kües, Yoichi Honda, and Yitzhak Hadar for their helpful suggestions on growing *Pleurotus ostreatus*. We thank the IMB genomic core for Illumina sequencing.

Funding

This work was supported by Academia Sinica Career Development Award AS-CDA-106-L03 and Taiwan Ministry of Science and Technology grant 106-2311-B-001-039-MY3 to Y.-P.H.

Conflicts of interest: None declared.

Literature cited

Alfaro M, Castanera R, Lavín JL, Grigoriev IV, Oguiza JA, *et al.* 2016. Comparative and transcriptional analysis of the predicted secretome in the lignocellulose-degrading basidiomycete fungus *Pleurotus ostreatus*. Environ Microbiol. 18:4710–4726.

Bailly-Bechet M, Haudry A, Lerat E. 2014. "One code to find them all": a perl tool to conveniently parse RepeatMasker output files. Mob DNA. 5:13.

Banerjee AC, Kundu A, Ghosh SK. 2003. Genetic manipulation of filamentous fungi. In: S Roussos, CR Soccol, A, Pandey C Augur, editors. New Horizons in Biotechnology. Dordrecht: Springer Netherlands. p. 193–198.

Barron GL, Thorn RG. 1987. Destruction of nematodes by species of *Pleurotus*. Can J Bot. 65:774–778.

Borodovsky M, McIninch J. 1993. GENMARK: parallel gene recognition for both DNA strands. Comput Chem. 17:123–133.

Cabanettes F, Klopp C. 2018. D-GENIES: dot plot large genomes in an interactive, efficient and simple way. PeerJ. 6:e4958.

Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, *et al.* 2009. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. Nucleic Acids Res. 37: D233–D238.

Castanera R, López-Varas L, Borgognone A, LaButti K, Lapidus A, *et al.* 2016. Transposable elements versus the fungal genome: impact on whole-genome architecture and transcriptional profiles. PLoS Genet. 12:e1006108.

Chakraborty M, Baldwin-Brown JG, Long AD, Emerson JJ. 2016. Contiguous and accurate *de novo* assembly of metazoan genomes with modest long read coverage. Nucleic Acids Res. 44:e147.

Chin C-S, Alexander DH, Marks P, Klammer AA, Drake J, *et al.* 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat Methods. 10:563–569.

Chin C-S, Peluso P, Sedlazeck FJ, Nattestad M, Concepcion GT, *et al.* 2016. Phased diploid genome assembly with single-molecule real-time sequencing. Nat Methods. 13:1050–1054.

Coghlan A, Coghlan A, Tsai IJ, Berriman M. 2018. Creation of a comprehensive repeat library for a newly sequenced parasitic worm genome. Protoc Exchange. 10.1038/protex.2018.054.

Cohen R, Persky L, Hadar Y. 2002. Biotechnological applications and potential of wood-degrading mushrooms of the genus *Pleurotus*. Appl Microbiol Biotechnol. 58:582–594.

Cox BS, Bevan EA. 1962. Aneuploidy in yeast. New Phytol. 61: 342–355.

Davies GJ, Williams SJ. 2016. Carbohydrate-active enzymes: sequences, shapes, contortions and cells. Biochem Soc Trans. 44:79–87.

Farman ML, Leong SA. 1995. Genetic and physical mapping of telomeres in the rice blast fungus, *Magnaporthe grisea*. Genetics. 140: 479–492.

Grigoriev IV, Nordberg H, Shabalov I, Aerts A, Cantor M, *et al.* 2012. The genome portal of the Department of Energy Joint Genome Institute. Nucleic Acids Res. 40:D26–D32.

Haas BJ, Salzberg SL, Zhu W, Pertea M, Allen JE, *et al.* 2008. Automated eukaryotic gene structure annotation using EVidenceModeler and the program to assemble spliced alignments. Genome Biol. 9:R7.

Hilber O. 1997. The Genus Pleurotus (Fr.) Kummer. Kelheim: Erschienen im Selbstverlag.

Käfer E. 1960. High frequency of spontaneous and induced somatic segregation in *Aspergillus nidulans*. Nature. 186:619–620.

Karim M, Daryaei MG, Torkaman J, Oladi R, Ghanbary MAT, *et al.* 2016. In vivo investigation of chemical alteration in oak wood decayed by *Pleurotus ostreatus*. Int Biodeterior Biodegrad. 108: 127–132.

Klein SJ, O'Neill RJ. 2018. Transposable elements: genome innovation, chromosome diversity, and centromere conflict. Chromosome Res. 26:5–23.

Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, *et al.* 2017. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome Res. 27:722–736.

Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, *et al.* 2009. Circos: an information aesthetic for comparative genomics. Genome Res. 19:1639–1645.

Kück U, Hoff B. 2010. New tools for the genetic manipulation of filamentous fungi. Appl Microbiol Biotechnol. 86:51–62.

Kwok OCH, Plattner R, Weisleder D, Wicklow DT. 1992. A nematicidal toxin from *Pleurotus ostreatus* NRRL 3526. J Chem Ecol. 18: 127–136.

Larraya L, Peñas MM, Pérez G, Santos C, Ritter E, *et al.* 1999a. Identification of incompatibility alleles and characterisation of molecular markers genetically linked to the A incompatibility locus in the white rot fungus *Pleurotus ostreatus*. Curr Genet. 34: 486–493.

Larraya LM, Pérez G, Peñas MM, Baars JJP, Mikosch TSP, *et al.* 1999b. Molecular karyotype of the white rot fungus *Pleurotus ostreatus*. Appl Environ Microbiol. 65:3413–3417.

Lee C-H, Chang H-W, Yang C-T, Wali N, Shie J-J, *et al.* 2020. Sensory cilia as the Achilles heel of nematodes when attacked by carnivorous mushrooms. Proc Natl Acad Sci USA. 117:6014–6022.

Li H, Wu S, Ma X, Chen W, Zhang J, *et al.* 2018. The genome sequences of 90 mushrooms. Sci Rep. 8:9982.

Love J, Palmer J, Stajich J, Esserand T, Kastman E, *et al.* 2019. nextgenusfs/funannotate: funannotate v1.5.2. 10.5281/zenodo.2576527.

Marçais G, Delcher AL, Phillippy AM, Coston R, Salzberg SL, *et al.* 2018. MUMmer4: a fast and versatile genome alignment system. PLoS Comput Biol. 14:e1005944.

Mikiashvili N, Wasser SP, Nevo E, Elisashvili V. 2006. Effects of carbon and nitrogen sources on *Pleurotus ostreatus* ligninolytic enzyme activity. World J Microbiol Biotechnol. 22:999–1002.

Moyzis RK, Buckingham JM, Cram LS, Dani M, Deaven LL, *et al.* 1988. A highly conserved repetitive DNA sequence, (TTAGGG)n, present at the telomeres of human chromosomes. Proc Natl Acad Sci USA. 85:6622–6626.

Nakazawa T, Izuno A, Horii M, Kodera R, Nishimura H, *et al.* 2017a. Effects of pex1 disruption on wood lignin biodegradation, fruiting development and the utilization of carbon sources in the white-rot Agaricomycete *Pleurotus ostreatus* and non-wood decaying *Coprinopsis cinerea*. Fungal Genet Biol. 109:7–15.

Nakazawa T, Izuno A, Kodera R, Miyazaki Y, Sakamoto M, *et al.* 2017b. Identification of two mutations that cause defects in the ligninolytic system through an efficient forward genetics in the white-rot agaricomycete *Pleurotus ostreatus*. Environ Microbiol. 19: 261–272.

Nakazawa T, Morimoto R, Wu H, Kodera R, Sakamoto M, *et al.* 2019. Dominant effects of gat1 mutations on the ligninolytic activity of the white-rot fungus *Pleurotus ostreatus*. Fungal Biol. 123:209–217.

Nakazawa T, Tsuzuki M, Irie T, Sakamoto M, Honda Y. 2016. Marker recycling via 5-fluoroorotic acid and 5-fluorocytosine counter-selection in the white-rot agaricomycete *Pleurotus ostreatus*. Fungal Biol. 120:1146–1155.

Peñas MM, Ásgeirsdóttir SA, Lasa I, Culiañez-Macia` FA, Pisabarro AG, *et al.* 1998. Identification, characterization, and in situ detection of a fruit-body-specific hydrophobin of *Pleurotus ostreatus*. Appl Environ Microbiol. 64:4028–4034.

Pérez G, Pangilinan J, Pisabarro AG, Ramírez L, 2009. Telomere organization in the ligninolytic basidiomycete *Pleurotus ostreatus*. Appl Environ Microbiol. 75:1427–1436.

Pollard DR, Käfer E, Johnston MT. 1968. Influence of chromosomal aberrations on meiotic and mitotic nondisjunction in *Aspergillus nidulans*. Genetics. 60:743–757.

Riley R, Salamov AA, Brown DW, Nagy LG, Floudas D, *et al.* 2014. Extensive sampling of basidiomycete genomes demonstrates inadequacy of the white-rot/brown-rot paradigm for wood decay fungi. Proc Natl Acad Sci USA. 111:9923–9928.

Sánchez C. 2009. Lignocellulosic residues: biodegradation and bioconversion by fungi. Biotechnol Adv. 27:185–194.

Sánchez C. 2010. Cultivation of *Pleurotus ostreatus* and other edible mushrooms. Appl Microbiol Biotechnol. 85:1321–1337.

Satou T, Kaneko K, Li W, Koike K. 2008. The toxin produced by *Pleurotus ostreatus* reduces the head size of nematodes. Biol Pharm Bull. 31:574–576.

Schechtman MG. 1990. Characterization of telomere DNA from *Neurospora crassa*. Gene. 88:159–165.

Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics. 31: 3210–3212.

Smit A, Hubley R, Green P. 2013–2015. *RepeatMasker Open-4.0.*

Stanke M, Waack S. 2003. Gene prediction with a hidden Markov model and a new intron submodel. Bioinformatics. 19:ii215–ii225.

Tatusov RL, Galperin MY, Natale DA, Koonin EV. 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. Nucleic Acids Res. 28:33–36.

Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, *et al.* 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One. 9:e112963.

Wang L, Gao W, Wu X, Zhao M, Qu J, *et al.* 2018. Genome-wide characterization and expression analyses of *Pleurotus ostreatus* MYB transcription factors during developmental stages and under heat stress based on de novo sequenced genome. Int J Mol Sci. 19: 2052.

Waterhouse RM, Seppey M, Simão FA, Manni M, Ioannidis P, *et al.* 2018. BUSCO applications from quality assessments to gene prediction and phylogenomics. Mol Biol Evol. 35:543–548.

Yang R-H, Li Y, Wáng Y, Wan J-N, Zhou C-L, *et al.* 2016. The genome of *Pleurotus eryngii* provides insights into the mechanisms of wood decay. J Biotechnol. 239:65–67.

Yoav S, Salame TM, Feldman D, Levinson D, Ioelovich M, *et al.* 2018. Effects of cre1 modification in the white-rot fungus *Pleurotus ostreatus* PC9: altering substrate preference during biological pretreatment. Biotechnol Biofuels. 11:212.

Zhang Z, Wen J, Li J, Ma X, Yu Y, *et al.* 2018. The evolution of genomic and epigenomic features in two *Pleurotus* fungi. Sci Rep. 8:8313.

Communicating editor: A. Rokas