Draft Genome Sequencing of the Pathogenic Fungus *Cladosporium phlei* ATCC 36193 Identifies Candidates of Novel Polyketide Synthase Genes Involved in Perylenequinone-Group Pigment Production

Evolutionary Bioinformatics
Volume 15: 1–3

DOI: 10.1177/1176934319831306
SAGE

Yu-Na Kang[1,2,3], Kum-Kang So[2], Do-Wan Kim[1], Dae-Hyuk Kim[2] and Tae-Ho Lee[1]

[1]Genomics Division, National Institute of Agricultural Science, Jeonju, Korea. [2]Department of Molecular Biology, Chonbuk National University, Jeonju, Korea. [3]Department of Agriculture, Chungnam National University, Daejeon, Korea.

ABSTRACT: *Cladosporium phlei*, which causes purple eyespot disease, has been focused on as a source of phleichrome from the perylenequinone group of pigments. Although this agent is important in photodynamic therapy, there are no genome sequences for the species. Here, we sequenced the genome of *C. phlei* and reported the draft sequence. The total length of the draft genome was approximately 31.8 Mb, and 9571 genes were predicted. Phylogenetic analysis showed that *Cladosporium sphaerospermum, Rachicladosporium* sp., and *Rachicladosporium antarcticum* were closely related, and this result corresponded to the taxonomic data. In addition to the draft genome sequence, we report four candidates of new polyketide synthase (PKS) genes, involved in the production of perylenequinone-group pigments.

KEYWORDS: *Cladosporium phlei*, whole-genome sequence, phleichrome, polyketide gene

RECEIVED: January 9, 2019. **ACCEPTED:** January 14, 2019.

TYPE: Short Report

FUNDING: The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was carried out with the support of "the Research Program for Agricultural Science & Technology Development (Project No. PJ01247901 & PJ01355801)" of the National Institute of Agricultural Sciences, Rural Development Administration, Republic of Korea.

DECLARATION OF CONFLICTING INTERESTS: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

CORRESPONDING AUTHOR: Tae-Ho Lee, Genomics Division, National Institute of Agricultural Science, Jeonju-si, Jeollabuk-do 54874, Korea.
Email: thlee0@korea.kr

Cladosporium phlei (George) de Vries belongs to a group of hypomycetous fungi and causes purple eyespot disease, which is one of the most common foliar diseases in timothy (*Phleum pratense*). This disease is easily distinguishable by eye-shaped spots, with light grayish-fawn centers and purple margins on plant leaves.[1] Phytopathogenic fungi, in general, produce toxins to infect the host plant,[2] and the toxins play many roles, such as killing cells, infecting dead cells, or preventing the induction of a defense reaction.[3] Perylenequinones, which are a group of toxins secreted in an infection process, have chromophores that absorb light energy and produce active oxygen. The produced hydroxyl radicals (OH^-), superoxide anions (O_2^-), hydrogen peroxide (H_2O_2), and singlet oxygen ($^1O_2$) impair host cells.[4] Perylenequinones commonly have a 3,10-dihydroxy-4,9-perylenequinone chromophore and different side chains.[5] Since the compounds in the perylenequinone group have light-sensitive properties, they act as photosensitizers or photosensitizing agents. Because of the sensitivity to oxygen and laser light, the agents are used for photodynamic therapy (PDT). Until now, the United States Food and Drug Administration (USFDA) only approved Photofrin as a photosensitizer for the treatment of esophageal cancer and early lung cancer. In addition, the clinical efficacy has been confirmed for superficial early gastric cancer, early cervical cancer, and dendritic cells. Hypocrellin, a major component of Photofrin, is produced from *Hypocrella bambusae*. Hypocrellin is the most promising photoreceptor in the perylenequinone group.[6] However, *H. bambusae* only grows in limited areas, such as the northwest part of Yunnan Province in China and the southeast region of Tibet, and it is difficult to obtain mass production because of the growth conditions. Hypocrellin belongs to the 3,10-dihydroxy-4,9-perylenequinone group, together with phleichrome, cladochrome, cercosporin, and isocercosporin.[5] To overcome the limitation of hypocrellin production, we have focused on phleichrome, which can be used as a precursor of hypocrellin. Phleichrome is a deep red mycelial pigment, produced by *C. phlei* as a secondary metabolite. Perylenequinones, including phleichrome, are produced by the polyketide synthase (PKS) gene family of pigment-related genes.[7] However, the presence of the genes in *C. phlei* is not clear because of the lack of genome information. Thus, we assembled the draft genome of *C. phlei* and identified 9571 genes, including four new PKS genes.

The strain used in this study was ATCC 36193. The fungus was grown in V8 juice (20°C), and after 12 days, the mycelium was harvested by filtration and cryogenically pulverized. Genomic DNA was isolated from 1 g of pulverized tissue, according to a previously described protocol.[8] We constructed a sequencing library with 350 and 550 bp inserts with the Illumina Truseq nano DNA library preparation kit. The library was quantified using the Qubit dsDNA high-sensitivity (HS) assay and quantitative polymerase chain reaction (qPCR). The library was sequenced on the Illumina NextSeq500 system. The total amount of raw data was 10.1 Gbase.

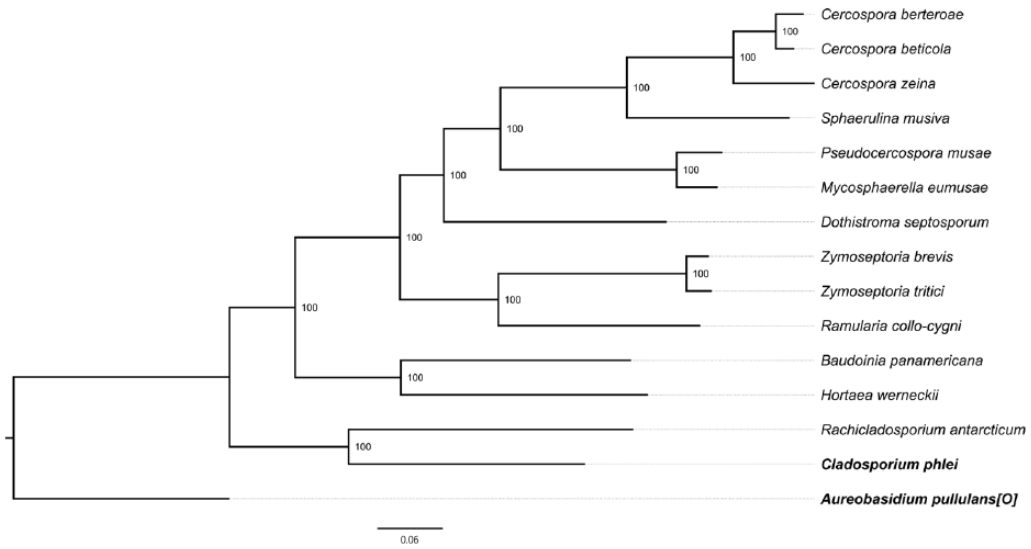

Figure 1. Phylogenetic position of *Cladosporium phlei* based on 71 single-copy genes. *Aureobasidium pullulans* in Dothideales is used as the outgroup.

To generate the best genome sequence, we assembled the genome by three different tools, Platanus (v.1.2.4),[9] SOAPdenovo2,[10] and CLCAssemblyCell (v4.2.0), with the default parameters. Scaffolding and gap closing for the contigs assembled by CLCAssemblyCell were performed using an in-house pipeline employing SSPACE[11] and Gapfiller[12] and considering homology regions in genomes of related species. The number of scaffolds produced by SOAPdenovo (6197) was smaller than that produced by Platanus (23 597), while the NG50 value of the Platanus results (25 429 bp) was, interestingly, larger than that of SOAPdenovo (16 255 bp). CLCAssemblyCell showed the best results, with the least number of scaffolds (5006) and the longest NG50 value (132 918 bp; Supplementary Table 1).

The annotation was performed using the Maker annotation pipeline,[13] resulting in the prediction of 9571 genes. To determine the completeness of the genome sequence and the annotation, BUSCO[14] analysis was performed with the BUSCO fungal data set including 1438 genes. As a result, 1412 complete single-copy BUSCOs, 4 duplicated BUSCOs, and 17 fragmented BUSCOs were found in the genome sequence, while only 5 BUSCOs were missing.

To construct a phylogenetic tree, we downloaded protein sequence files of available 15 species in Capnodiale order from GenBank and determined 71 proteins of single-copy genes that one species only has one ortholog gene by orthoMCL.[15] The tree was constructed using proml in the Phylip package[16] and the phangorn package[17] in the R environment, with the concatenated single-copy protein sequences multiple-aligned by MAFFT. *Aureobasidium pullulans*, belonging to the order Dothideales, was used as the outgroup. In the tree, *C. phlei* was included in a Cladosporiaceae family cluster with *Rachicladosporium antarcticum* (Figure 1), and the results corresponded to the taxonomic information.

The candidates for PKS were selected based on the BLASTP similarity for protein sequences larger than 500 amino acids, with a 1e–10 e-value cutoff against a custom database containing 14 188 known PKS proteins. The PKS genes were determined by a National Center for Biotechnology Information (NCBI) CDD[18] domain search for the candidates. By the NCBI CDD domain search, eight PKSs were determined from 58 candidates. Among them, CpPKS1, CpPKS2, CpPKS3, and CpPKS4 have been known from previous studies.[7] However, four candidates, namely, CpPKS5, CpPKS6, CpPKS7, and CpPKS8, were new in *C. phlei*. Previous studies have shown that CpPKS1 is a non-reducing (NR)-PKS, which contains a duplicated acyl carrier protein (ACP) and thioesterase (TE) domain and is involved in pigment production. CpPKS2, a partially reducing (PR)-PKS, was found to be involved in the synthesis of methylsalicylic acids (MSAs) such as patulin. CpPKS3 and CpPKS4 are highly reducing (HR)-PKSs and contain domains such as dehydratase (DH), ketoreductase (KR), or enoylreductase (ER) domains.[7] Similarly, CpPKS5 had a NR-PKS domain similar to that of CpPKS1; CpPKS7 was similar to CpPKS2; CpPKS6 and CpPKS8 had HR-PKS domains similar to those of CpPKS3 and CpPKS4.

Here, we report the first draft genome of *C. phlei*, isolated from timothy, and four new PKS genes, which are potentially involved in phleichrome metabolism. Despite its importance, only four genomes, *C. cladosporioides* (https://www.ncbi.nlm.nih.gov/assembly/GCA_001931875.2), *C. sphaerospermum* UM843,[19] *Rachicladosporium* sp. CCFEE 5018, and *R. antarcticum* CCFEE5527,[20] in *Cladosporiaceae* family were previously available (Table 1). We believe that the availability of the genome sequence, as well as the PKS gene information, provides a valuable basis for *Cladosporium* studies.

Table 1. Genome assembly results of five *Cladosporium phlei*–related genomes.

SPECIES	PHLEI	CLADO	UM843	5018	ANTAR
Genome status	Draft	Draft	Draft	Draft	Draft
Assembly size (bp)	31.8 M	47.5 M	26.8 M	44.7 M	47.4 M
GC content	52.95%	42.9%	54.3%	54.9%	54.9%
Genome coverage	317×	100×	98.0×	175×	41.0×
No. of scaffolds	5006	843	155	232	267
N50	132 918	186 555	969 659	1 358 697	896 820
N90	25 411	20 301	200 226	194 991	162 724

PHLEI, CLADO, UM843, 5018, and ANTAR in the species row denote *Cladosporium phlei, Cladosporium cladosporioides, Cladosporium sphaerospermum* UM843, *Rachicladosporium* sp. CCFEE 5018, and *Rachicladosporium antarcticum*, respectively.

The draft genome sequence was deposited at NCBI (National Center for Biotechnology Information, https://www.ncbi.nlm.nih.gov/). The accession number is QZFA00000000.

Author Contributions

Conceived and designed the experiments: D.H.K. and T.H.L. Performed the experiments: K.Y.N. and K.K.S. Analyzed the data: K.Y.N. and T.H.L. Wrote the first draft of the manuscript: K.Y.N. and D.W.K. Agree with manuscript results and conclusions: K.Y.N., K.K.S., D.W.K., D.H.K., and T.H.L. Made critical revisions and approved final version: D.W.K., D.H.K., and T.H.L. All authors reviewed and approved of the final manuscript.

Supplemental Material

Supplemental material for this article is available online.

REFERENCES

1. Yoshihara T, Shimanuki T, Araki T, Sakamura S. Phleichrome: a new phytotoxic compound produced by *Cladosporium phlei*. *Agric Biol Chem*. 1975;39:1683–1684.
2. Knogge W. Fungal infection of plants. *Plant Cell*. 1996;8:1711–1722.
3. Wolpert TJ, Dunkle LD, Ciuffetti LM. Host-selective toxins and avirulence determinants: what's in a name? *Ann Rev Phytopathol*. 2002;40:251–285.
4. Daub ME, Ehrenshaft M. The photoactivated Cercospora toxin cercosporin: contributions to plant disease and fundamental biology. *Ann Rev Phytopathol*. 2000;38:461–490.
5. Daub ME, Herrero S, Chung KR. Photoactivated perylenequinone toxins in fungal pathogenesis of plants. *FEMS Microbiol Lett*. 2005;252:197–206.
6. Diwu Z. Novel therapeutic and diagnostic applications of hypocrellins and hypericins. *Photochem Photobiol*. 1995;61:529–539.
7. So KK, Kim JM, Nguyen NL, et al. Rapid screening of an ordered fosmid library to clone multiple polyketide synthase genes of the phytopathogenic fungus *Cladosporium phlei*. *J Microbiol Methods*. 2012;91:412-419.
8. Yi MH, Kim JA, Kim JM, et al. Characterization of a mutant strain of a filamentous fungus *Cladosporium phlei* for the mass production of the secondary metabolite phleichrome. *J Microbiol*. 2011;49:680–683.
9. Kajitani R, Toshimoto K, Noguchi H, et al. Efficient de novo assembly of highly heterozygous genomes from whole-genome shotgun short reads. *Genome Res*. 2014;24:1384–1395.
10. Luo R, Liu B, Xie Y, et al. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *Gigascience*. 2012;1:18.
11. Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W. Scaffolding preassembled contigs using SSPACE. *Bioinformatics*. 2010;27:578–579.
12. Nadalin F, Vezzi F, Policriti A. GapFiller: a de novo assembly approach to fill the gap within paired reads. *BMC Bioinformatics*. 2012;13:S8.
13. Cantarel BL, Korf I, Robb SM, et al. MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Res*. 2008;18:188–196.
14. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*. 2015;31:3210–3312.
15. Li L, Stoeckert CJ Jr, Roos DS. OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res*. 2003;13:2178–2189.
16. Plotree D, Plotgram D. PHYLIP-phylogeny inference package (version 3.2). *Cladistics*. 1989;5:6.
17. Schliep KP. Phangorn: phylogenetic analysis in R. *Bioinformatics*. 2011;27:592–593.
18. Marchler-Bauer A, Derbyshire MK, Gonzales NR, et al. CDD: NCBI's conserved domain database. *Nucleic Acids Res*. 2015;43:D222–D226.
19. Ng KP, Yew SM, Chan CL, et al. Sequencing of *Cladosporium sphaerospermum*, a dematiaceous fungus isolated from blood culture. *Eukaryotic Cell*. 2012;11:705–706.
20. Coleine C, Masonjones S, Selbmann L, et al. Draft genome sequences of the antarctic endolithic fungi rachicladosporium antarcticum CCFEE 5527 and rachicladosporium sp. CCFEE 5018. *Genome Announc*. 2017;5:e00397–e00417.