Contents lists available at ScienceDirect

Data in Brief

Data Article

Draft genome sequences of the oomycete *Pythium insidiosum* strain CBS 573.85 from a horse with pythiosis and strain CR02 from the environment

Preecha Patumcharoenpol [a,b], Thidarat Rujirawat [c], Tassanee Lohnoo [c], Wanta Yingyong [c], Nongnuch Vanittanakom [d], Weerayuth Kittichotirat [a,*], Theerapong Krajaejun [e,*]

[a] *Systems Biology and Bioinformatics Research Group, Pilot Plant Development and Training Institute, King Mongkut's University of Technology Thonburi, Bangkok, Thailand*
[b] *Department of Biomedical Informatics, University of Arkansas for Medical Sciences, Little Rock, AR, USA*
[c] *Research Center, Faculty of Medicine, Ramathibodi Hospital, Mahidol University, Bangkok, Thailand*
[d] *Department of Microbiology, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand*
[e] *Department of Pathology, Faculty of Medicine, Ramathibodi Hospital, Mahidol University, Bangkok, Thailand*

ARTICLE INFO

ABSTRACT

Pythium insidiosum is an aquatic oomycete microorganism that causes the fatal infectious disease, pythiosis, in humans and animals. The organism has been successfully isolated from the environment worldwide. Diagnosis and treatment of pythiosis is difficult and challenging. Genome sequences of *P. insidiosum*, isolated from humans, are available and accessible in public databases. To further facilitate biology-, pathogenicity-, and evolution-related genomic and genetic studies of *P. insidiosum*, we report two additional draft genome sequences of the *P. insidiosum* strain CBS 573.85 (35.6 Mb in size; accession number, BCFO00000000.1) isolated from a horse with pythiosis, and strain CR02 (37.7 Mb in size; accession number, BCFR00000000.1) isolated from the environment.

---

* Corresponding authors.
  E-mail addresses: weerayuth.kit@kmutt.ac.th (W. Kittichotirat), mr_en@hotmail.com (T. Krajaejun).

Specifications Table

Subject area	Biology
More specific subject area	Microbiology, Genomics
Type of data	Genome sequence data
How data was acquired	IlluminaHiSeq 2000 and IlluminaHiSeq 2500 Next Generation Sequencing Platforms
Data format	Assembled genome sequences
Experimental factors	Genomic DNA was extracted from the *Pythium insidiosum* strains CBS 573.85 (an animal isolate) and CR02 (an environmental isolate).
Experimental features	Genome of the *P. insidiosum* strains CBS 573.85 and CR02 were sequenced and assembled.
Data source location	*Pythium insidiosum* strain CBS 573.85 was isolated from a horse in Costa Rica, and strain CR02 was isolated from the environment in Thailand.
Data accessibility	The draft genome sequences of *P. insidiosum* have been deposited in the DNA Data Bank of Japan (DDBJ) under the accession numbers: BCFO00000000.1 (strain CBS573.85; https://www.ncbi.nlm.nih.gov/nuccore/BCFO00000000.1) and BCFR00000000.1 (strain CR02; https://www.ncbi.nlm.nih.gov/nuccore/BCFR00000000.1).

Value of the data

- Previously, only genome sequence data of *P. insidiosum* isolated from humans is available in the public databases.
- The first draft genome sequences of *P. insidiosum* isolated from a non-human animal with pythiosis and from the environment are now made available.
- The additional genome data will facilitate biology-, pathogenicity-, and evolution-related studies of *P. insidiosum*, through comparative genomic studies of *Pythium* species and related species.

1. Data

Pythium insidiosum is an aquatic oomycete microorganism that causes the lethal infectious condition, pythiosis, in humans and other animals [1,2]. The organism has been isolated from the environment in Australia, Thailand, Brazil and the United States [3–6]. Genome sequences of *P. insidiosum*, isolated from humans, are available and accessible in public databases [7,8]. We report two additional draft genome sequences of the organism isolated from a horse with pythiosis, as well as from the environment.

2. Experimental design, materials and methods

2.1. Genome sequencing and assembly

Genomic DNA samples were extracted from *P. insidiosum* strain CBS573.85 (from an infected horse in Costa Rica) and strain CR02 (from an agricultural area in Thailand), using the conventional extraction method described by Lohnoo and co-workers [9]. rDNA sequence analysis was performed to confirm the identity of the organism [10–12]. The extracted genomic DNA of each of these two strains was subjected to preparation of a paired-end library for genome sequencing, using the IlluminaHiSeq 2500 (strain CBS573.85) or IlluminaHiSeq 2000 (strain CR02) platform (Yourgene Bioscience, Taiwan). Quality trims of the raw reads were executed by CLC Genomics Workbench

(www.clcbio.com) to yield read lengths with at least 35 bases. The adaptor sequences were removed by Cutadapt 1.8.1 [13] to obtain a total of 34,651,034 reads with an average read length of 122 bases (4,238,414,330 total bases) for the strain CBS573.85, and a total of 27,436,541 reads with an average read length of 105 bases (2,888,290,738 total bases) for the strain CR02. All reads were assembled by Velvet 1.2.10 [14] to the total sequence length of 35,561,321 bases (number of contigs, 11,223; average contig length, 3169; N_{50}, 12,261; 'N' composition, 1.2%) for the strain CBS573.85, and 37,673,126 bases (number of contigs, 22,560; average contig length, 1670; N_{50}, 3553; 'N' composition, 2.7%) for the strain CR02. CEGMA analysis with 248 highly-conserved eukaryotic genes [15,16] reported 87% and 77% completeness of draft genomes of the strains CBS573.85 and CR02, respectively. MAKER2 [17] assigned 14,487 (strain CBS573.85) and 15,231 (strain CR02) open reading frames.

2.2. Data accessibility

The genome sequence data has been deposited in DDBJ under the accession numbers BCFO00000000.1 (strain CBS573.85) and BCFR00000000.1 (strain CR02).

Acknowledgements

Transparency document. Supplementary material

Transparency document associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.dib.2017.11.002.

References

[1] T. Krajaejun, B. Sathapatayavongs, R. Pracharktam, P. Nitiyanant, P. Leelachaikul, W. Wanachiwanawin, A. Chaiprasert, P. Assanasen, M. Saipetch, P. Mootsikapun, P. Chetchotisakd, A. Lekhakula, W. Mitarnun, S. Kalnauwakul, K. Supparatpinyo, R. Chaiwarith, S. Chiewchanvit, N. Tananuvat, S. Srisiri, C. Suankratay, W. Kulwichit, M. Wongsaisuwan, S. Somkaew, Clinical and epidemiological analyses of human pythiosis in Thailand, Clin. Infect. Dis. 43 (2006) 569–576.

[2] W. Gaastra, L.J.A. Lipman, A.W.A.M. De Cock, T.K. Exel, R.B.G. Pegge, J. Scheurwater, R. Vilela, L. Mendoza, *Pythium insidiosum*: an overview, Vet. Microbiol. 146 (2010) 1–16.

[3] R.I. Miller, Investigations into the biology of three "phycomycotic" agents pathogenic for horses in Australia, Mycopathologia. 81 (1983) 23–28.

[4] J. Supabandhu, M.C. Fisher, L. Mendoza, N. Vanittanakom, Isolation and identification of the human pathogen *Pythium insidiosum* from environmental samples collected in Thai agricultural areas, Med. Mycol. 46 (2008) 41–52.

[5] J.W. Presser, E.M. Goss, Environmental sampling reveals that *Pythium insidiosum* is ubiquitous and genetically diverse in North Central Florida, Med. Mycol. 53 (2015) 674–683.

[6] C.G. Zambrano, A.O.S. Fonseca, J.S.S. Valente, C.Q. Braga, E.S.V. Sallis, M.I. Azevedo, C. Weiblen, J.M. Santurio, S.A. Botton, D.I. B. Pereira, Isolation and characterization of Pythium species from swampy areas in the Rio Grande do Sul, Brazil, and evaluation of pathogenicity in an experimental model, Pesqui. Veter. Bras. 37 (2017) 459–464.

[7] M.S. Ascunce, J.C. Huguet-Tapia, E.L. Braun, A. Ortiz-Urquiza, N.O. Keyhani, E.M. Goss, Whole genome sequence of the emerging oomycete pathogen *Pythium insidiosum* strain CDC-B5653 isolated from an infected human in the USA, Genome Data 7 (2016) 60–61.

[8] T. Rujirawat, P. Patumcharoenpol, T. Lohnoo, W. Yingyong, T. Lerksuthirat, S. Tangphatsornruang, P. Suriyaphol, L. J. Grenville-Briggs, G. Garg, W. Kittichotirat, T. Krajaejun, Draft genome sequence of the pathogenic oomycete *Pythium insidiosum* strain Pi-S, isolated from a patient with pythiosis, Genome Announc. 3 (2015).

[9] T. Lohnoo, N. Jongruja, T. Rujirawat, W. Yingyon, T. Lerksuthirat, U. Nampoon, Y. Kumsang, P. Onpaew, P. Chongtrakool, A. Keeratijarut, T.T. Brandhorst, T. Krajaejun, Efficiency comparison of three methods for extracting genomic DNA of the pathogenic oomycete *Pythium insidiosum*, J. Med. Assoc. Thail. 97 (2014) 342–348.

[10] T. Rujirawat, T. Sridapan, T. Lohnoo, W. Yingyong, Y. Kumsang, P. Sae-Chew, W. Tonpitak, T. Krajaejun, Single nucleotide polymorphism-based multiplex PCR for identification and genotyping of the oomycete *Pythium insidiosum* from humans, animals and the environment, Infect. Genet. Evol. 54 (2017) 429–436.

[11] P.R. Badenoch, D.J. Coster, B.L. Wetherall, H.T. Brettig, M.A. Rozenbilds, A. Drenth, G. Wagels, *Pythium insidiosum* keratitis confirmed by DNA sequence analysis, Br. J. Ophthalmol. 85 (2001) 502–503.

[12] A. Chaiprasert, T. Krajaejun, S. Pannanusorn, C. Prariyachatigul, W. Wanachiwanawin, B. Sathapatayavongs, T. Juthayothin, N. Smittipat, N. Vanittanakom, A. Chindamporn, *Pythium insidiosum* Thai isolates: molecular phylogenetic analysis, Asian Biomed. 3 (2009) 623–633.

[13] M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads, EMBnet J. 17 (2011) 10.

[14] D.R. Zerbino, E. Birney, Velvet: algorithms for de novo short read assembly using de Bruijn graphs, Genome Res. 18 (2008) 821–829.

[15] G. Parra, K. Bradnam, Z. Ning, T. Keane, I. Korf, Assessing the gene space in draft genomes, Nucleic Acids Res. 37 (2008) 289–297.

[16] G. Parra, K. Bradnam, I. Korf, CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes, Bioinformatics 23 (2007) 1061–1067. http://dx.doi.org/10.1093/bioinformatics/btm071.

[17] C. Holt, M. Yandell, MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects, BMC Bioinform. 12 (2011) 491.