

# Draft Genome Sequence of the Pathogenic Oomycete *Pythium insidiosum* Strain Pi-S, Isolated from a Patient with Pythiosis

Thidarat Rujirawat,<sup>a,c,d</sup> Preecha Patumcharoenpol,<sup>b</sup> Tassanee Lohnoo,<sup>c</sup> Wanta Yingyong,<sup>c</sup> Tassanee Lerksuthirat,<sup>a,c,d</sup> Sithichoke Tangphatsornruang,<sup>e</sup> Prapat Suriyaphol,<sup>f</sup> Laura J. Grenville-Briggs,<sup>g</sup> Gagan Garg,<sup>h</sup> Weerayuth Kittichotirat,<sup>b</sup> Theerapong Krajaejun<sup>a</sup>

Department of Pathology, Faculty of Medicine, Ramathibodi Hospital, Mahidol University, Bangkok, Thailand<sup>a</sup>; Systems Biology and Bioinformatics Research Group, Pilot Plant Development and Training Institute, King Mongkut's University of Technology Thonburi, Bangkok, Thailand<sup>b</sup>; Research Center, Faculty of Medicine, Ramathibodi Hospital, Mahidol University, Bangkok, Thailand<sup>c</sup>; Molecular Medicine Program, Multidisciplinary Unit, Faculty of Science, Mahidol University, Bangkok, Thailand<sup>d</sup>; Genomic Research Laboratory, National Center for Genetic Engineering and Biotechnology, National Science and Technology Development Agency, Pathumthani, Thailand<sup>e</sup>; Bioinformatics and Data Management for Research, Office for Research and Development, Faculty of Medicine, Siriraj Hospital, Mahidol University, Bangkok, Thailand<sup>f</sup>; Department of Plant Protection Biology, Swedish University of Agricultural Sciences, Alnarp, Sweden<sup>g</sup>; CSIRO Agriculture Flagship, Centre for Environment and Life Sciences, Floreat, Western Australia, Australia<sup>h</sup>

T.R. and P.P. contributed equally to this work.

***Pythium insidiosum* is an oomycete that causes a life-threatening infectious disease called pythiosis in humans and animals living in tropical and subtropical countries. Here, we report the first draft genome sequence of *P. insidiosum*. The genome of *P. insidiosum* is 53.2 Mb and contains 14,962 open reading frames.**

Received 28 April 2015 Accepted 19 May 2015 Published 18 June 2015

**Citation** Rujirawat T, Patumcharoenpol P, Lohnoo T, Yingyong W, Lerksuthirat T, Tangphatsornruang S, Suriyaphol P, Grenville-Briggs LJ, Garg G, Kittichotirat W, Krajaejun T. 2015. Draft genome sequence of the pathogenic oomycete *Pythium insidiosum* strain Pi-S, isolated from a patient with pythiosis. *Genome Announc* 3(3):e00574-15. doi: 10.1128/genomeA.00574-15.

**Copyright** © 2015 Rujirawat et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](https://creativecommons.org/licenses/by/3.0/).

Address correspondence to Weerayuth Kittichotirat, weerayuth.kit@kmutt.ac.th, or Theerapong Krajaejun, mr\_en@hotmail.com.

Morphologically, oomycetes look like fungi, but they separately and independently evolved from the metazoan ancestor into a group of eukaryotic microorganisms that have unique genetic, biochemical, and physiological characteristics (1, 2). While almost all pathogenic oomycetes infect plants, the oomycete *Pythium insidiosum* is capable of infecting humans and other animals living in tropical and subtropical areas of the world and causes a life-threatening infection called pythiosis (1, 3, 4). Surgical removal of an infected organ (eye or leg) or death is usually the final clinical outcome for patients with pythiosis (4). How evolution contributed to virulence and pathogenicity of *P. insidiosum* needs to be explored at the genome level.

Availability of next-generation sequencing technologies provides opportunities to sequence whole genomes of nonmodel organisms, including *P. insidiosum*. Here, we report the first draft genome sequence of *P. insidiosum*. The conventional extraction protocol (5) was used to extract genomic DNA (gDNA) of the *P. insidiosum* strain Pi-S, which was isolated from a Thai patient with vascular pythiosis (the most common form of *P. insidiosum* infection). Identity of the organism was confirmed by culture analysis and rDNA sequence. The extracted gDNA underwent whole-genome sequencing using a combination of the Illumina HiSeq2000 platform by Yourgene Bioscience, Taiwan (<http://www.yourgene.com.tw>), and the 454 Genome Sequencer FLX Titanium platform by the National Center for Genetic Engineering and Biotechnology, Thailand (<http://www.biotech.or.th>). For one library, the 454 platform provided 342,637 raw reads and 200,682,697 raw bases, which were assembled by Newbler version 2.8 (Roche) to the total contig length of 24,383,569 bases (number

of contigs, 14,455; average contig length, 1,686). The Illumina platform was used to sequence one paired-end (180-bp insert) and three mate pair (5-kb, 8-kb, and 15-kb insert) libraries. After sequences were quality trimmed by CLC Genomics Workbench (<http://www.clcbio.com>), all Illumina libraries yielded a total of 240,448,339 reads and 22,926,078,291 bases (average read length, 95), which were then assembled by ALLPATHS-LG version 44588 (6) to the total contig length of 52,806,015 bases (number of contigs, 1,261; average contig length, 41,876). Gaps were filled using GapCloser (7). Merging the 454- and Illumina-derived assembled sequences, using Minimus (8), provided the draft genome sequence of 53,239,050 bases, comprising 1,192 contigs (average contig length, 44,664 [range, 966 to 655,053];  $N_{50}$ , 146,252), with a G+C content of 52% and an N composition of 10%.

The draft genome was assessed to be 92% complete, using CEGMA with a well-defined set of 248 highly conserved genes from eukaryotes (9, 10). Transcript mapping (11) and gene prediction by MAKER2 (12) revealed 14,962 open reading frames (ORFs), of which 3,579, 1,254, and 10,129 ORFs were predicted by both transcript mapping and gene prediction, only transcript mapping, or only gene prediction, respectively. A BLASTp search against the NCBI nonredundant protein database found significant matches ( $E$  value, less than  $-6$ ) in 89.7% of ORF-translated proteins.

**Nucleotide sequence accession numbers.** Related genome sequence data of *P. insidiosum* strain Pi-S has been deposited in the DDBJ under the accession numbers [BBXB01000001](https://www.ncbi.nlm.nih.gov/nuclseq/BBXB01000001) to [BBXB01001192](https://www.ncbi.nlm.nih.gov/nuclseq/BBXB01001192).

## ACKNOWLEDGMENTS

This work is supported by a research fund from Mahidol University (T. Krajaejun), by a Ph.D. scholarship from the Faculty of Medicine, Ramathibodi Hospital, Mahidol University (T. Rujirawat), and by a Royal Golden Jubilee Ph.D. scholarship from the Thailand Research Fund (T. Lerksuthirat).

We thank Thomas D. Sullivan for reviewing the manuscript, and we are grateful to Angkana Chairasert for material support and suggestion.

## REFERENCES

1. Kamoun S. 2003. Molecular genetics of pathogenic oomycetes. *Eukaryot Cell* 2:191–199. <http://dx.doi.org/10.1128/EC.2.2.191-199.2003>.
2. Kwon-Chung KJ. 1994. Phylogenetic spectrum of fungi that are pathogenic to humans. *Clin Infect Dis* 19(Suppl 1):S1–S7. [http://dx.doi.org/10.1093/clinids/19.Supplement\\_1.S1](http://dx.doi.org/10.1093/clinids/19.Supplement_1.S1).
3. Gaastra W, Lipman LJ, De Cock AW, Exel TK, Pegge RB, Scheurwater J, Vilela R, Mendoza L. 2010. *Pythium insidiosum*: an overview. *Vet Microbiol* 146:1–16. <http://dx.doi.org/10.1016/j.vetmic.2010.07.019>.
4. Krajaejun T, Sathapatayavongs B, Prachartam R, Nitiyanant P, Leela-chaikul P, Wanachiwanawin W, Chairasert A, Assanasen P, Saipetch M, Mootsikapun P, Chetchotisakd P, Lekhakula A, Mitarnun W, Kalnauwakul S, Supparatpinyo K, Chaiwarith R, Chiewchanvit S, Tananuvat N, Srisiri S, Suankratay C, Kulwichit W, Wongsaisuan M, Somkaew S. 2006. Clinical and epidemiological analyses of human pythiosis in Thailand. *Clin Infect Dis* 43:569–576. <http://dx.doi.org/10.1086/506353>.
5. Lohnoo T, Jongruja N, Rujirawat T, Yingyon W, Lerksuthirat T, Nampoon U, Kumsang Y, Onpaew P, Chongtrakool P, Keeratijarut A, Brandhorst TT, Krajaejun T. 2014. Efficiency comparison of three methods for extracting genomic DNA of the pathogenic oomycete *Pythium insidiosum*. *J Med Assoc Thai* 97:342–348.
6. Gnerre S, Maccallum I, Przybylski D, Ribeiro FJ, Burton JN, Walker BJ, Sharpe T, Hall G, Shea TP, Sykes S, Berlin AM, Aird D, Costello M, Daza R, Williams L, Nicol R, Gnirke A, Nusbaum C, Lander ES, Jaffe DB. 2011. High-quality draft assemblies of mammalian genomes from massively parallel sequence data. *Proc Natl Acad Sci USA* 108:1513–1518. <http://dx.doi.org/10.1073/pnas.1017351108>.
7. Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J, He G, Chen Y, Pan Q, Liu Y, Tang J, Wu G, Zhang H, Shi Y, Liu Y, Yu C, Wang B, Lu Y, Han C, Cheung DW, Yiu S-M, Peng S, Xiaoqian Z, Liu G, Liao X, Li Y, Yang H, Wang J, Lam T-W, Wang J. 2012. SOAPdenovo2: an empirically improved memory-efficient short-read *de novo* assembler. *GigaScience* 1:18. <http://dx.doi.org/10.1186/2047-217X-1-18>.
8. Sommer DD, Delcher AL, Salzberg SL, Pop M. 2007. Minimus: a fast, lightweight genome assembler. *BMC Bioinformatics* 8:64. <http://dx.doi.org/10.1186/1471-2105-8-64>.
9. Parra G, Bradnam K, Korf I. 2007. CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. *Bioinformatics* 23:1061–1067. <http://dx.doi.org/10.1093/bioinformatics/btm071>.
10. Parra G, Bradnam K, Ning Z, Keane T, Korf I. 2009. Assessing the gene space in draft genomes. *Nucleic Acids Res* 37:289–297. <http://dx.doi.org/10.1093/nar/gkn916>.
11. Krajaejun T, Lerksuthirat T, Garg G, Lohnoo T, Yingyong W, Khositnithikul R, Tangphatsornruang S, Suriyaphol P, Ranganathan S, Sullivan TD. 2014. Transcriptome analysis reveals pathogenicity and evolutionary history of the pathogenic oomycete *Pythium insidiosum*. *Fungal Biol* 118:640–653. <http://dx.doi.org/10.1016/j.funbio.2014.01.009>.
12. Holt C, Yandell M. 2011. MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinformatics* 12:491. <http://dx.doi.org/10.1186/1471-2105-12-491>.